

# Advanced Film Grain Noise Extraction and Synthesis for High-Definition Video Coding

Byung Tae Oh, Shaw-min Lei, *Fellow, IEEE*, and C.-C. Jay Kuo, *Fellow, IEEE*

**Abstract**—A technique for film grain noise extraction, modeling and synthesis is studied and applied to high-definition video coding in this paper. Film grain noise enhances the natural appearance of pictures in high-definition video and should be preserved in coded video. However, the coding of video contents with film grain noise is expensive. In previous art, it was proposed to enhance the coding performance by extracting film grain noise from the input video at the encoder as a preprocessing step, and by resynthesizing and adding it back to the decoded video at the decoder as a postprocessing step. In a novel implementation of this approach, we first remove film grain noise from image/video with a variational denoising approach without distorting its original content. Then, we present a parametric model (consisting of a small set of parameters) to generate film grain noise that is close to the actual one in terms of a couple of observed statistical properties, such as the power spectral density and the crosschannel spectral correlation. Under this framework, the coding gain of denoised video is higher while the visual quality of the final reconstructed video is well preserved. Experimental results are provided to demonstrate the superior performance of the proposed approach.

**Index Terms**—Denoising, film grain noise, image restoration, noise synthesis, parametric model, texture coding, texture synthesis, total variation.

## I. INTRODUCTION

Film grain noise in motion pictures is caused by the developing process of silver-halide crystals dispersed in photographic emulsion [1]. It is unavoidable in the analog film due to the physical process. When we digitize and compress high-resolution movie contents obtained by scanning the analog film, such randomly distributed film grain noise is a major burden to typical video coding methods. Since film grain noise has a relatively large energy level in the high-frequency region, it is more expensive to encode in the DCT domain. Besides, the

Manuscript received April 27, 2008; revised August 6, 2008 and March 17, 2009. First version published July 7, 2009; current version published December 1, 2009. This paper was supported by Sharp Laboratories of America, Camas, WA. This paper was recommended by Associate Editor M. Comer.

B. T. Oh was with Sharp Laboratories of America, Camas, WA 98607-9489 USA. He is now with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: byungoh@usc.edu).

S. Lei was with Sharp Laboratories of America, Camas, WA 98607-9489 USA. He is now with Mediatek, Hsinchu 300, Taiwan (e-mail: shawmin.lei@mediatek.com).

C.-C. Jay Kuo is with the Signal Image Processing Institute, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: cckuo@sipi.usc.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2009.2026974

underlying video suffers from inaccurate motion estimation. A natural idea to overcome this problem is to remove film grain noise as much as possible as a preprocessing step at the encoder so as to achieve a higher coding gain for denoised video [2]–[5].

A lot of efforts have been done for Gaussian noise detection and removal. Under the assumption that film grain noise is one of the Gaussian additive or multiplicative noise, many existing denoising methods could be used. However, this assumption does not hold since film grain noise has the following distinctive properties [1], [3], [4], [6], [7].

- 1) It is temporally independent.
- 2) Its power spectrum density is close to pink noise.
- 3) It is spatially dependent.
- 4) It has strong crosscolor correlation in the RGB domain.
- 5) Its histogram is close to the Gaussian distribution.
- 6) It is dependent on the signal intensity.

Because of these properties, we need to develop a specific algorithm for film grain noise detection, modeling and removal.

To remove the film grain noise from the original video is however not enough. As high-resolution devices such as HDTV are getting popular, film grain noise becomes perceptually important to human eyes since noise-removed video tends to bring an unnatural feeling to people. As a result, we should reproduce and render film grain noise at the decoder. Previous paper following this line of thought will be reviewed in Section II. In this paper, we consider a novel implementation of this approach.

Specifically, we present a method to remove film grain noise from general input video without distorting the underlying video content. In the denoising process at the encoder, we adopt a method based on the principle of *total variation minimization* for film grain noise removal. It suppresses film grain noise effectively so that the video coding gain can be significantly increased. It is important to preserve the quality of the original video content as much as possible. Otherwise, it would lead to false extraction of film grain noise and, consequently, inaccurate noise parameters. To achieve this goal, we detect edges or fine texture regions of input video in advance and extract (or remove) noise in smooth homogenous regions only. In the meantime, we analyze film grain noise using a parametric model and determine its parameters. Based on the model and its parameters, artificial noise (which is close to the extracted one) is generated and added back to

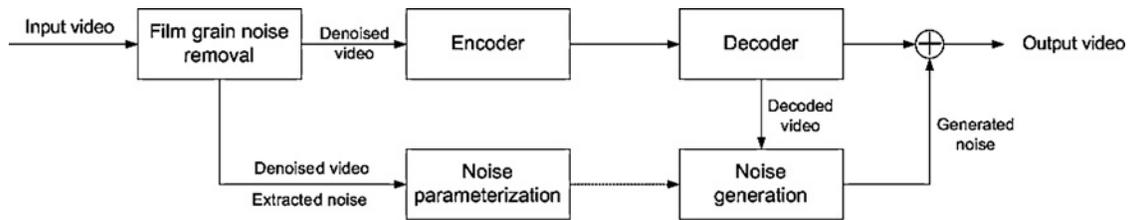


Fig. 1. Overview of a film grain noise processing system.

the decoded video at the decoder. Furthermore, we provide a method to measure the performance of film grain synthesis by comparing the statistical information of distinctive properties of film grain noise between the synthesized and extracted noise.

The rest of this paper is organized as follows. The overall structure of the proposed scheme with previous paper on film grain noise removal and modeling is briefly reviewed in Section II. Then, algorithms for noise removal and synthesis are detailed in Section III and Section IV, respectively. Experimental results are provided in Section V to demonstrate the effectiveness of the proposed scheme with several performance metrics. Finally, concluding remarks are given in Section VI.

## II. REVIEW OF PREVIOUS PAPER

Generally speaking, the film grain noise modeling scheme for video coding consists of two parts: 1) noise removal and extraction at the encoder; and 2) noise synthesis at the decoder as shown in Fig. 1. These can be viewed as the preprocessing and the postprocessing steps in video coding. It is worthwhile to emphasize that it does not modify the encoding and decoding modules in any adopted video coding standard. The only additional information to be transmitted is noise parameters, with which noise can be synthesized. Since a small set of parameters can represent the whole image, or a set of images in a GOP, the overhead of parameter transmission is negligible. One method for parameter transmission was proposed by Gomila *et al.* [3], [4] using the so-called SEI messages, with which no auxiliary transmission channel is needed. The detailed description of these two parts with the proposed algorithms will be given in Section III for noise removal and Section IV for noise synthesis, respectively.

The film grain noise modeling scheme for video coding was first proposed by Gomila *et al.* [3], [4], where noise removal was adopted as a preprocessing step in video encoding and noise synthesis as a postprocessing step in video decoding. This idea has been standardized by AVC [2], and even deployed in commercial products for HD DVD in [8]. However, there is some room for further improvement. First, they did not provide any specific denoising method for film grain noise. As will be described later in this section, most of previous denoising schemes aiming to the conventional Gaussian additive noise is not sufficient to suppress film grain noise efficiently. One approach proposed in [3] uses the reconstructed video as its denoised version, which is attractive since it does not need an additional denoising module. However, it is observed

that residual images usually contain image edges and some structure information besides noise. Poorly extracted film grain noise would lead to false estimation of noise parameters.

In the following, we will review image/video denoising and noise synthesis techniques. Image restoration including noise detection and removal has been one of active research topics in image processing during last several decades. The main objective of these algorithms is suppressing the noise as much as possible without distorting the original image, and various approaches have been proposed so far [9]. When extending the denoising problem from image to video, temporal correlation of noise should be considered. Ozkan *et al.* [10] applied temporal filtering to noise suppression yet preserving image edges at the same time. The integrated spatial and temporal filtering approach was examined in [11], [12]. Temporal filtering methods in these papers were built upon block motion estimation. Boo *et al.* [13] applied the Karhunen–Loeve (KL) transform along the temporal direction to decorrelate dependency between successive frames and then used adaptive Wiener filtering to smooth frames. Most methods using temporal filtering work well for still or slow-motion video since temporal filtering can be done nicely with accurate motion information. However, large motion estimation errors occurring in fast motion video tend to result in erroneous noise estimation. Furthermore, it demands a large memory buffer to implement temporal filtering.

Among hundreds of papers published on this topic, some of them target at film grain noise processing, e.g. [7], [13]–[16]. It was assumed by Yan *et al.* [7], [15] and Al-Shaykh *et al.* [14] that film grain noise is proportional to a certain power,  $p$ , of the background signal intensity, which imposes strong limitation on the extraction and modeling of film grain noise. More recently, Moldovan *et al.* [16] proposed new algorithms using Bayesian model to detect film grain noise, where film grain noise is assumed to be of the beta distribution and spatially independent. Under this assumption, an image can be estimated by an iterative gradient descent method with a predetermined model and proper parameters. This approach is claimed to be robust against different image contents. However, according to our observation, the distribution of film grain noise is close to the Gaussian distribution and it is not spatially independent.

In the denoising process, it is important to find out an edge region of the image, since most of denoising algorithm tends to blur the image, especially around the edge. There has been a large amount of work for edge detection. In particular, Canny [17] proposed an edge detection algorithm for noisy images. More recently, a multilayer approach is used to reduce false

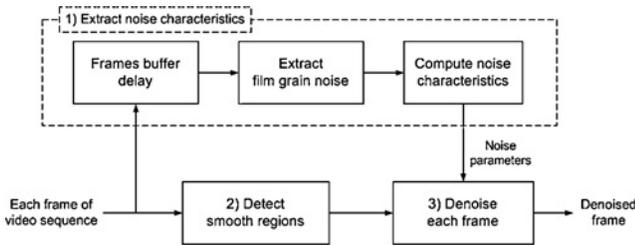


Fig. 2. Block diagram of the preprocessing task.

detection due to noise. For example, Mallet *et al.* [18] used local maxima of over complete wavelet transform coefficients for edge detection. They proved that finding local maxima is equivalent to multiscale Canny’s edge detector. Instead of finding local maxima of wavelet coefficients, Xiong *et al.* [19] used the cross-scale correlation for edge detection by multiplying cross-scale wavelet coefficients under the assumption that the multiscale noise intensity is related in a logarithmic ratio. However, this method would not work well for film grain noise whose energy values deviate from the logarithmic variation assumption. As more specific application for the film grained image, Schallauer *et al.* [20] proposed new algorithms to detect homogeneous image region, in which the block-based 2-D DFT was adopted to detect the directionality of edges and/or fine textures and, then, properties of film grain noise are extracted only from homogenous regions. Since film grain noise is isotropic while edges and fine textures have strong directionality, 2-D DFT provides a good tool for their distinction. However, the decision made based on several points with four-different angles  $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$  and a couple of radial bands is not accurate enough to determine the directionality, since the spectrum of image edges tends to be across a wide range of frequency bands and sampled values at fixed points in the DFT domain may lead to false conclusion.

Two different methods were proposed for film grain synthesis. One is to use the film grain database for the sake of low complexity. The film grain pattern is first identified, and the decoder generates a larger size of film grain from a smaller size of film grain stock. However, the block-based copy-and-paste method might yield artificial boundary artifacts. Besides, the method is workable only when the film stock information is known *a priori*. The other is to use some models for blind film grain synthesis. Several methods have been proposed, e.g., high-order statistics based [7], [15], parametric-modeling-based [21]–[24] or patch-based noise synthesis methods [25]–[28]. Since film grain noise can be viewed as one type of random texture, the conventional texture synthesis method can also be adopted. However, there is one challenge. That is, since film grain noise has special properties as mentioned in Section I, these criteria must be considered and satisfied when synthesizing the film grain. So far, there has been no effort to generate the film grain noise based on the specific film grain properties.

Finally, it is worthwhile to mention that a similar idea was used for speech coding [29], where inactive voice signal is preprocessed before encoding, and noise is added back to the decoded signal for the comfort of human perception.

### III. FILM GRAIN NOISE REMOVAL

For film grain noise removal, we use the total variation minimization method to suppress film grain noise. Since a denoising process might distort areas that have sharp transition between neighboring pixels in the original image, it is important to identify regions of image edges before applying the denoising algorithm. Then, we can perform denoising selectively in smooth regions only. Moreover, the denoising process based on the total variation minimization principle could be more complete with some prior information of noise. Here, we propose to use the independence of film grain in the temporal domain to identify prior noise information. The overall preprocessing task at the encoder can be divided into three steps: 1) extract noise characteristics using the temporal information; 2) identify smooth regions of the image; and 3) denoise each image with prior noise information. The preprocessing module is shown in Fig. 2, and each processing step will be detailed below.

#### A. Extraction of Noise Characteristics

It is important to identify the accurate film grain characteristics, since the proposed denoising algorithm is conducted using the noise prior. Generally, it is nontrivial to identify them in that the film grain noise is not given in this stage. However, what we want to obtain is just statistical properties of film grain noise, and it can be indirectly exploited without computing film grain itself.

To obtain film grain characteristics, we assume an additive model of film grain noise of the following form:

$$U_o^n = U^n + N^n(U^n) \quad (1)$$

where  $U_o$  is the observed image,  $U$  is the original image and  $N$  is the film grain noise, and superscript  $n$  denotes the frame index. Since film grain noise is signal-dependent, noise  $N(\cdot)$  is a function of image  $U$ . If blocks of two consecutive frames are static, we can find the differential noise,  $N_d$ , as

$$\begin{aligned} N_d^n &= N^n - N^{n-1} \\ &= (U_o^n - U^n) - (U_o^{n-1} - U^{n-1}) \\ &\approx U_o^n - U_o^{n-1} \end{aligned} \quad (2)$$

since  $U^n \approx U^{n-1}$ . The static area is identified by analyzing motion vectors of a set of nonoverlapping blocks, where motion vectors are computed in noise suppressed images as done in [11].

Since film grain noise is almost temporally independent and spatially Gaussian distributed as mentioned in Section I, differential noise  $N_d$  is another Gaussian noise with its variance twice as large. Then, all statistical values of film grain noise, such as variance, autocorrelation and crosscolor correlations can be easily obtained from temporally differential noise  $N_d$ . It is also worthwhile to point out that the extracted noise characteristics remain stable for a long sequence of image frames. In practice, they can be use for the whole video sequence or, at least, for a large number of consecutive image frames. In other words, it is unnecessary to estimate the noise parameters frequently. Furthermore, as compared to other previous paper that performs filtering along the temporal

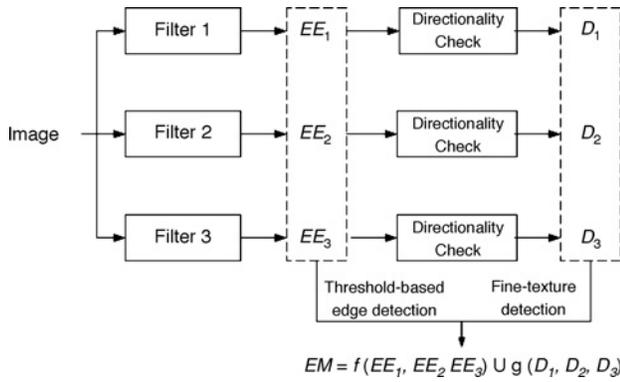


Fig. 3. Detection of nonsmooth regions in an image.

direction directly, the fact that only static areas of consecutive frames are considered in noise extraction makes our algorithm more robust in the presence of object motion.

### B. Enhanced Edge Detection

For input image sequences that contain film grain noise, simple edge detection methods such as the Sobel or the Prewitt filter does not work well since these filters are sensitive to local noise. Note that we need to identify the edge region of noisy image, so that some additional process is necessary to suppress noise as much as possible to facilitate edge detection. In this manner, an enhanced edge detection method using multiresolution filters is described below.

To extract edges from the input image effectively, we consider a set of filters to maximize frequency selectivity. These filters are built upon a pair of low- and high-pass filters as

$$h = \frac{1}{8}[-1 \ 2 \ 6 \ 2 \ -1], \quad g_1 = \frac{1}{2}[1 \ 0 \ -1]. \quad (3)$$

Then, we can construct three filters by  $f_1 = h * g_1$ ,  $f_2 = h * g_2$ ,  $f_3 = h * g_3$  accordingly, where  $*$  is the convolution operation and  $g_2$  and  $g_3$  are the upsampled filters of  $g_1$  as

$$g_2 = \frac{1}{2}[1 \ 0 \ 0 \ 0 \ -1], \quad g_3 = \frac{1}{2}[1 \ 0 \ 0 \ 0 \ 0 \ 0 \ -1].$$

These filters are applied along the horizontal as well as the vertical directions to detect edges of all possible orientations. This process is similar to the overcomplete wavelet decomposition. In [6], we proposed to use a multiresolution overcomplete wavelet decomposition with simple integer-valued low- and high-pass filters. However, we found that its frequency selectivity is not good enough so that it could miss some signal with specific frequency bands. Instead, these filters are chosen to improve frequency selectivity while keeping integer-valued filter coefficients.

By following the terminology used in wavelet theory, the application of filter  $f_i$ ,  $i = 1, 2, 3$ , horizontally and vertically can generate  $LH_i$  and  $HL_i$  output images. Then, the edge energy map is calculated by

$$EE_i = (|LH_i|^p + |HL_i|^p)^{1/p}, \quad i = 1, 2, 3 \quad (4)$$

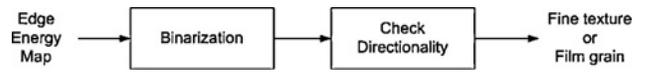


Fig. 4. Process of fine texture detection.

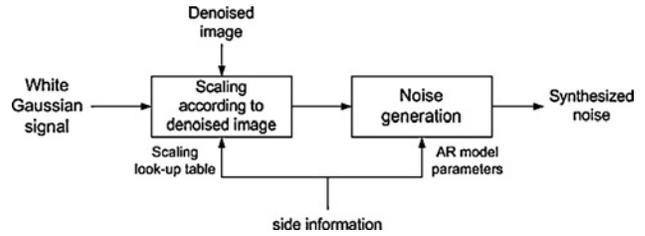
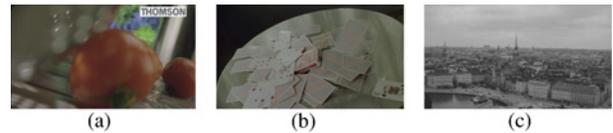


Fig. 5. Film grain noise synthesis with scaled white noise.

Fig. 6. First frames of HD (1920×1080) test sequences. (a) *Rolling tomatoes*. (b) *Playing cards*. (c) *Old town cross*.

and the unified edge energy map is obtained by

$$EE = \max[EE_1, EE_2, EE_3] \quad (5)$$

where the maximum operation is performed pixel-by-pixel. Then, the binary edge map ( $EM$ ) is obtained by thresholding, i.e., if the value of  $EE(i, j)$  at pixel position  $(i, j)$  is larger than a predetermined threshold, it is set to an edge point.

Since the proposed noise extraction and removal algorithm depend on the edge map, it is critical to find a reliable edge threshold value. Recall that the edge threshold depends on the signal intensity due to the unique property of film grain noise. For example, 8-bit image should need 256 different threshold values according to its signal intensity. To the threshold value, we set the initial threshold to 3 for all signal intensities, and adaptively update them according to the noise level in smooth areas (i.e., where the edge map value is equal to 0). The update formula is given by

$$Th_{\text{new}}[L] = (1 - w) \cdot Th_{\text{old}}[L] + w \cdot c \cdot EE \quad (6)$$

where weighting coefficient  $w$  is set to a small number (e.g.,  $10^{-4}$ ) to avoid abrupt change,  $L$  is the background signal luminance, coefficient  $c$  is a scaling factor used to adjust the input value, and  $EE$  is the current edge energy value obtained by the above-mentioned method. Note that the updating process is done pixel-by-pixel manner, i.e., each threshold value corresponding to pixel illumination  $L$  is a scaled mean of  $EE$  by  $c$ . It is observed that threshold values converge after 4–5 frames with the updating process and this algorithm is not sensitive for initial threshold value. Under the assumption that film grain noise is Gaussian distributed, we can show analytically that  $c = 2.5$  would help detect more than 99% of film grain noise. Details of the analysis and its experimental verification are given in the Appendix.

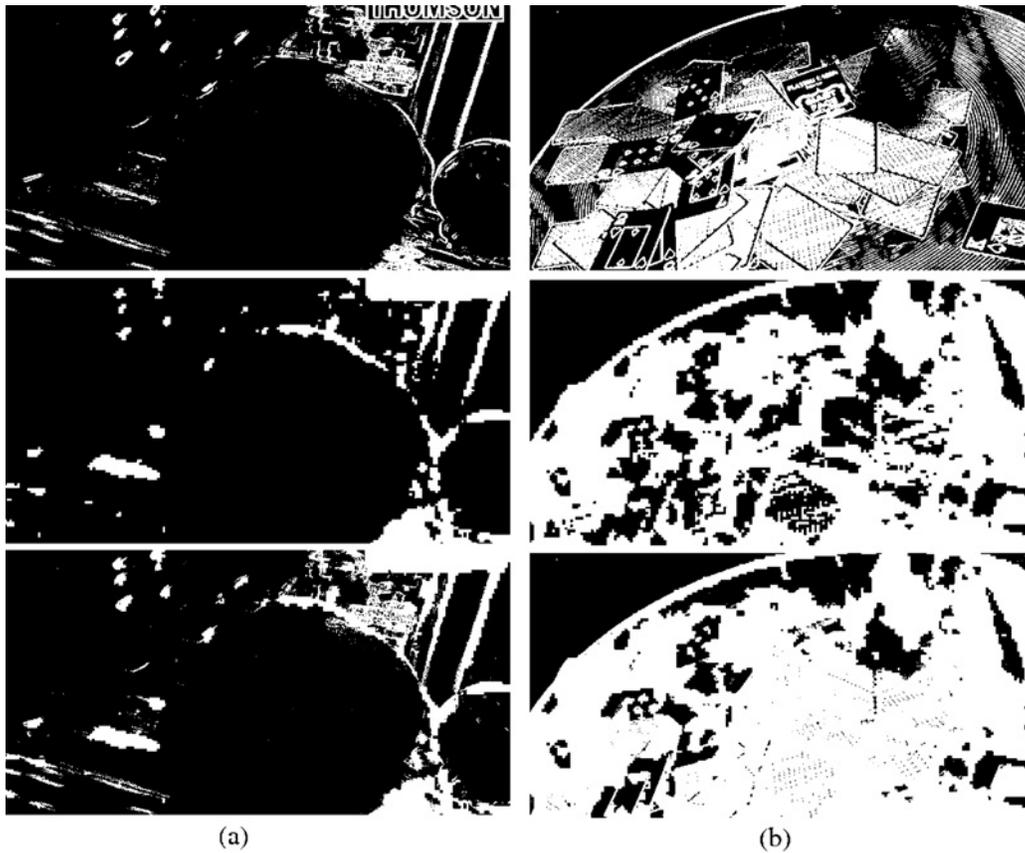


Fig. 7. Edge-regions by threshold method (top); fine texture region (middle) and final edge map (bottom). (a) *Rolling tomatoes*. (b) *Playing cards*.

In the implementation, we use the luminance channel to get the edge map rather than processing each RGB color channel individually, and quantize the signal pixel value  $U(i, j)$  by step size  $Q_s = 8$  via

$$L = \text{floor}(U(i, j)/Q_s). \quad (7)$$

As a result, a total of 32 different threshold values are used to determine the edge map.

In the last stage of edge detection, an additional postprocessing module is added after thresholding to minimize the false detection of edge pixels. That is, all isolated pixels are detected and eliminated under the assumption that edge pixels are connected with each other. In spite of the postprocessing, there might still be misclassified pixels. One possible example is that film grain noise may have a wider spatial dependency than edge detection kernels, by which misclassified pixels are likely to be connected. As a result of false decision, the overall coding gain would be degraded slightly, and the regenerated noise will be added to the original decoded noise as edge regions. However, regions with false decision are often small, and the overall coding performance will not be significantly affected.

### C. Fine Texture Detection

When detecting smooth regions using the edge energy map for denoising, the main difficulty is that we often misclassify fine texture regions into smooth regions. It is desirable to

treat fine texture regions the same as edge regions, since the fine texture pattern is perceptually visible when it has strong periodicity in the spatial domain in spite of its low-edge energy. That is, we should not perform the denoising algorithm on them. However, since the edge energy of these texture pixels tends to be low, they may not be detectable by thresholding. To address this problem, we include the fine texture detection task as shown in Fig. 3, in which periodicity property of fine texture is explored for its detection. Finally, the nonsmooth region is obtained by taking the union of detected edges and fine texture regions.

The fine texture detection process is illustrated in Fig. 4. First, the edge energy map is binarized by its local median to even identify low-intensity texture. Then, we check whether there is strong autocorrelation  $R(m, n)$  along 8 different directions denoted by vector  $(m, n)$ . In the discrete domain, we approximate the 8 directions by

$$\begin{aligned} (m, n) &\in D \\ &= \{(1, 0), (2, 1), (1, 1), (1, 2), (0, 1), (-1, 2), (-1, 1), (-2, 1)\}. \end{aligned}$$

Since film grain noise has no directionality, the maximum correlation value of film grain noise is smaller than that of the fine texture. However, there may be some false alarm if the correlation value of film grain noise becomes larger. To improve the robustness of our algorithm, we check the correlation value twice. That is, if  $|R(p, q)|$  gives the maximum correlation value, we compute the correlation value for

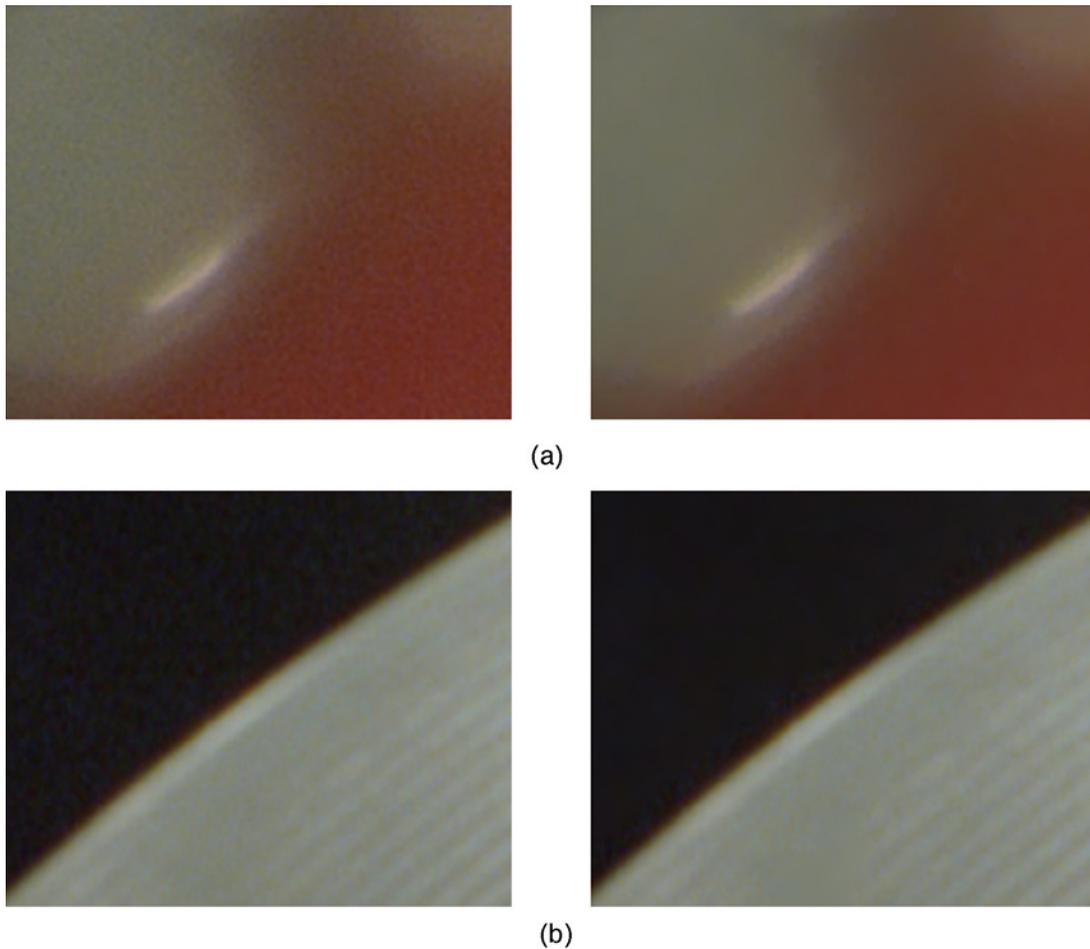


Fig. 8. Close-up view of the original (left) and the denoised (right) images. (a) *Rolling tomatoes*. (b) *Playing cards*.

$|R(2p, 2q)|$ , too. The final decision is made based on the product of  $|R(p, q)|$  and  $|R(2p, 2q)|$ .

This procedure can be explained mathematically as follows. The correlation value of a pixel  $c$  can be written as

$$M_c = |R_c(p, q)| \cdot |R_c(2p, 2q)|, \quad (8)$$

where  $(p, q) = \arg \max_{(m, n) \in D} |R_c(m, n)|$ .

A pixel  $c$  will be determined as fine texture when it satisfies

$$M_c > \beta \cdot M_{N^d} \quad (9)$$

where  $N^d$  is the differential noise obtained in (2). This condition says that the threshold value for fine texture detection is directly dependent on the autocorrelation value of the estimated grain noise. In our implementation, we fix  $\beta = 2$ , which means a pixel is assumed as fine texture if its maximum correlation value is at least twice as large as the maximum correlation of film grain noise.

As discussed above, our approach relies on the periodicity of fine texture under the assumption that strong periodicity boosts the perception of fine texture. On the other hand, low-intensity nonperiodic fine texture would be missed by the proposed algorithm. Spatially nonperiodic but temporally continuous low-intensity texture should be handled by temporal-based filtering, which is out of our current scope.

#### D. Denoising With Total Variation (TV) Minimization

As mentioned in Section I, film grain noise has many specific properties, and it is desirable to exploit these properties in noise removal. Traditionally, linear spatial smoothing filters such as the Gaussian filter have been widely used as the denoising filter. Their performance is however limited since a linear filter is not effective in detecting and extracting low-frequency energy of film grain noise. Here, we proposed to use a nonlinear filtering approach based on the variational principle for film grain denoising, especially total variation (TV) minimization method. Note that the denoising process is only applied in smooth regions, which are detected by methods discussed in Sections III-B and III-C. For more references about the TV method, we refer to [9], [30]–[32].

1) *Denoising With TV Minimization*: The variational approach with TV minimization [32] is an effective nonlinear denoising algorithm. It can preserve edges well while removing background noise. In addition, it can provide a complete solution if the noise prior is given, which is suitable in our current context. The algorithm is detailed below. For the additive noise model in (1), we want to reconstruct the noise-free image  $U^n$  based on observed noisy image  $U_o^n$ . Then, the solution of the ill-posed inverse problem can be obtained by

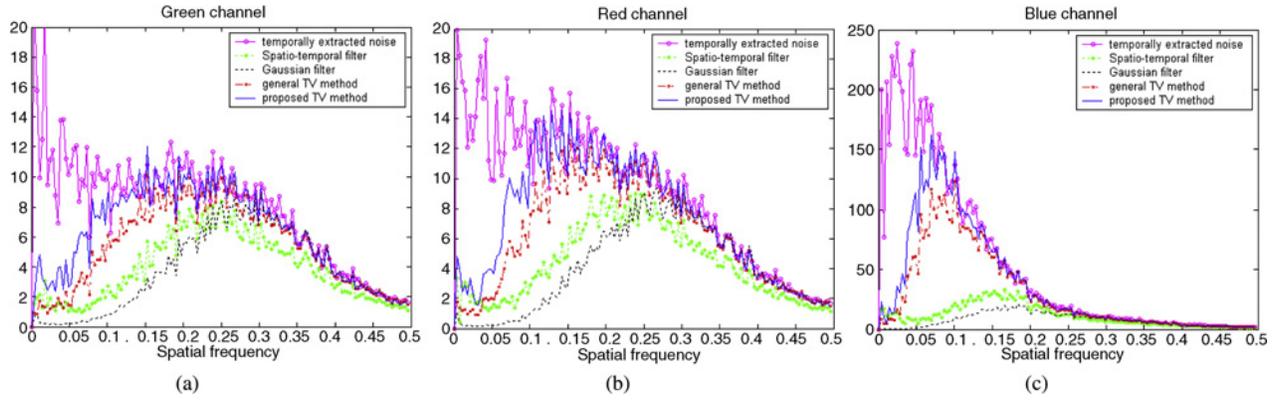


Fig. 9. Comparison of the squared-root of the PSD for extracted noise using several algorithms. (a) Green. (b) Red. (c) Blue.

TABLE I  
COMPARISON OF EXTRACTED NOISE POWER BY DIFFERENT ALGORITHMS

		Temporally Extracted	Spatio-Temporal Filter	Gaussian Filter	General TV Method	Proposed TV Method
<i>Rolling tomatoes</i>	G	5.10	3.95	3.75	4.56	4.89
	R	5.71	4.22	3.97	5.10	5.49
	B	16.72	7.44	5.98	13.50	15.28
<i>Playing cards</i>	G	9.86	4.97	4.40	6.74	6.80
	R	6.95	4.32	3.82	5.07	5.23
	B	38.20	19.56	18.23	28.56	29.64
<i>Old town</i>	Y	13.41	9.65	10.12	12.87	12.88

solving the optimization problem

$$\min_U \int_{\Omega} |\nabla U| dU, \quad \text{s.t.} \quad \|U - U_o\|^2 = \sigma^2 \quad (10)$$

where function  $U : \Omega \rightarrow R$ ,  $\Omega$  is a nonempty bounded open set in  $R^2$ ,  $\nabla$  is the differential operator [30], [32], and superscript  $n$  is dropped for simplicity. In this case, noise is assumed to be white. By Lagrange's theorem, the best estimator can be written as

$$\begin{aligned} \hat{U} &= \arg \min_U \left[ \int_{\Omega} |\nabla U| dU + \frac{\lambda}{2} \|U - U_o\|^2 \right] \\ &= \arg \min_U \left[ \int_{\Omega} \sqrt{U_x^2 + U_y^2} du + \frac{\lambda}{2} \int_{\Omega} (U - U_o)^2 du \right]. \end{aligned} \quad (11)$$

To solve the above problem, the Euler-Lagrange differential equation is used as a necessary condition, and the update process is given by

$$\begin{aligned} U_t &= \text{div} \left( \frac{\nabla U}{|\nabla U|} \right) - \lambda(U - U_o) \\ &= \frac{\partial}{\partial x} \frac{U_x}{\sqrt{U_x^2 + U_y^2}} + \frac{\partial}{\partial y} \frac{U_y}{\sqrt{U_x^2 + U_y^2}} - \lambda(U - U_o) \end{aligned} \quad (12)$$

where  $\lambda$  is the Lagrangian multiplier, which is iteratively updated via

$$\lambda = \frac{1}{\sigma^2} \int_{\Omega} \text{div} \left( \frac{\nabla U}{|\nabla U|} \right) (U - U_o) dU. \quad (13)$$

2) *Film Grain Denoising With TV Minimization*: It is worthwhile to point out that the conventional TV-based denoising algorithm is not enough since it does not take the properties of film grain noise into account. Our main contribution in this paper is to exploit the distinctive properties of film grain and incorporate them as a constraint in the denoising process. Specifically, Properties 2, 3 and 5 can be used in a single channel input image, i.e., the gray (or luminance) channel. With assuming that film grain noise has Gaussian distribution by Property 5, we can find a coloring matrix  $P$  to estimate the spatial correlation of noise among neighborhood pixels by Properties 2 and 3, such that

$$U - U_o = Pw \quad (14)$$

where  $P$  is a linear transform and  $w$  is white Gaussian noise, so that the spectrum of  $Pw$  matches with that of the extracted noise. Then, (11) can be rewritten as

$$\hat{U} = \arg \min_U \left[ F(U) + \frac{\lambda}{2} (U - U_o)^T R^{-1} (U - U_o) \right] \quad (15)$$

where  $R = PP^T$  is the autocorrelation matrix of  $(U - U_o)$ . This whitening process helps estimate the noise behavior in the spatial domain, and it eventually leads to better noise suppression. To reduce the computational complexity, we approximate it by computing only  $9 \times 9$  local blocks in the implementation.

If the input image is a color image of *RGB* three channels, we can use Property 4, i.e., the crosscolor correlation) furthermore to improve the noise estimation. We first find the

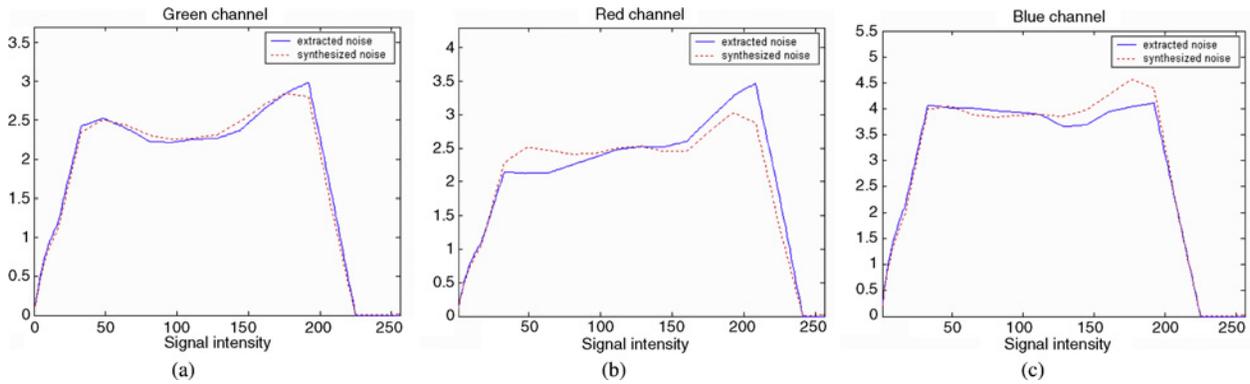


Fig. 10. Comparison of signal dependency between extracted and synthesized noise for the *Rolling tomatoes* sequence. (a) Green. (b) Red. (c) Blue.

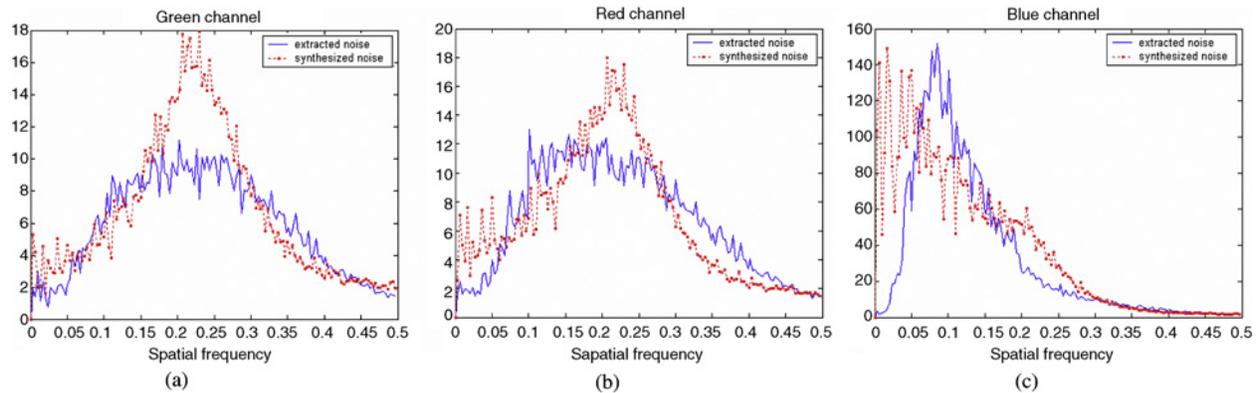


Fig. 11. Comparison of the square root of PSD between extracted and synthesized noise for the *Rolling tomatoes* sequence. (a) Green. (b) Red. (c) Blue.

$G$ -channel noise data as the reference, and obtain  $R$ - and  $B$ -channel noise data based on the given extracted noise information of  $G$ -channel. For this case, we have two constraints at the same time so that the minimization process for  $B$ -channel can be modified as

$$\begin{aligned} \hat{U}_B = \arg \min_{U_B} [ & F(U_B) + \lambda_1 (U_B - U_{Bo})^T R_B^{-1} (U_B - U_{Bo}) \\ & + \lambda_2 (U_G - U_{Go})^T R_{GB}^{-1} (U_B - U_{Bo}) ] \end{aligned} \quad (16)$$

where  $\lambda_1$  and  $\lambda_2$  are updated similarly to that in (13).

#### IV. FILM GRAIN NOISE MODELING AND SYNTHESIS

Film grain noise modeling and synthesis are discussed in this section. For noise modeling, a few parameters are determined to represent the extracted noise and transmitted to the noise synthesizer at the decoder.

##### A. AR Noise Model

There is no commonly agreed objective metric to measure the closeness of the synthesized and the real film grain noise. Thus, this is often done by visual inspection. As a result, film grain noise modeling is a challenging problem. There are several factors to be considered, including the spatial power spectrum density, the noise probability density, and the crosscolor correlation as mentioned in Section I.

TABLE II  
COMPARISON OF THE CROSSCOLOR CORRELATION

	<i>Rolling tomatoes</i>		<i>Playing cards</i>	
	Extracted	Synthesized	Extracted	Synthesized
$G$ - $R$	0.85	0.92	0.21	0.25
$G$ - $B$	0.95	0.97	0.80	1.23

In this paper, for film grain noise modeling, we consider the following AR model:

$$N(i, j, c) = \sum_i \sum_{j'} \sum_{c'} a_{ij'c'} \cdot N(i - i', j - j', c - c') \quad (17)$$

which is a 3-D AR model with the 2-D spatial correlation and the 1-D spectral correlation. Please note that the AR model is an IIR filter, which in general has a better frequency representation than a FIR filter. The power spectrum of synthesized noise can be controlled by the frequency response of the IIR filter with a white input signal. Furthermore, the AR model as given in (17) includes both the spatial and the crosscolor correlation naturally. Generally speaking, the model can capture the desired properties of film grain noise well.

A set of AR parameters is obtained by following the Yule-Walker AR estimation method from the extracted noise at the encoder. Since film grain noise has the same characteristics over a sequence of image frames, a set of AR parameters is sufficient for the noise synthesis purpose. Besides, we

TABLE III  
QUALITY RATING ON A 1–5 SCALE

Rating	Impairment	Quality
5	Imperceptible	Excellent
4	Perceptible, not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

only need a small number of coefficients for the AR model. Empirically, we choose values of  $(i', j', c')$  to be

$$(1, 0, 0), (0, 1, 0), (1, 1, 0), (-1, 1, 0), (2, 0, 0), \\ (0, 2, 0), (0, 0, 1).$$

This choice results in a causal filter in the raster scanning order so that it is convenient for noise synthesis and the overhead of coding these parameters is very low.

### B. Signal-Dependent Noise Synthesis

The synthesized noise by the given AR model has no information about the background signal. Since film grain noise has the signal dependency property, we should modify the synthesized noise according to the decoded signal. Basically, scaling factors are obtained from the extracted noise like AR parameters, and both are transmitted to the decoder as the side information. Then, we can scale synthesized noise based on its background signal intensity according to the transmitted scaling factors. However, it is not easy to preserve the crosscolor correlation by treating the signal-dependent noise directly. That is, if we scale the synthesized noise according to the background signal, it is likely to modify the crosscolor correlation as well. To preserve the crosscolor correlation as much as possible in generating signal-dependent noise, the scaled excitation as shown Fig. 5 is adopted. That is, instead of scaling the synthesized film grain noise, we scale the excitation white signal before noise synthesis.

Film grain synthesis using the AR model was first proposed by Gomila *et al.* [3], [4] and standardized in AVC [2]. However, they did not fully utilize specific film grain properties in the AR model, e.g., the signal-dependent property as we do here. Moreover, they showed the advantage of the AR model only using a subjective approach. In contrast, we will show statistical properties of synthesized film grain noise in this paper (see Section V-B).

### C. Output Image Construction

The noise synthesizer generates film grain noise that has properties similar to the extracted one according to the procedures described above. Then, the final output image is obtained by

$$U_{\text{out}}(i, j) = U_{\text{decod}}(i, j) + N_{\text{gen}}(i, j) \quad (18)$$

where  $U_{\text{decod}}$  is the decoded image and  $N_{\text{gen}}$  is film grain noise generated by the AR model. Since the signal-dependent

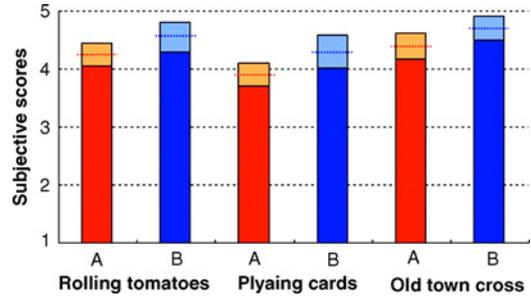


Fig. 12. Comparison of subjective test results, where A denotes the coding result using the conventional H.264/AVC reference codes and B denotes the coding result using the proposed method.

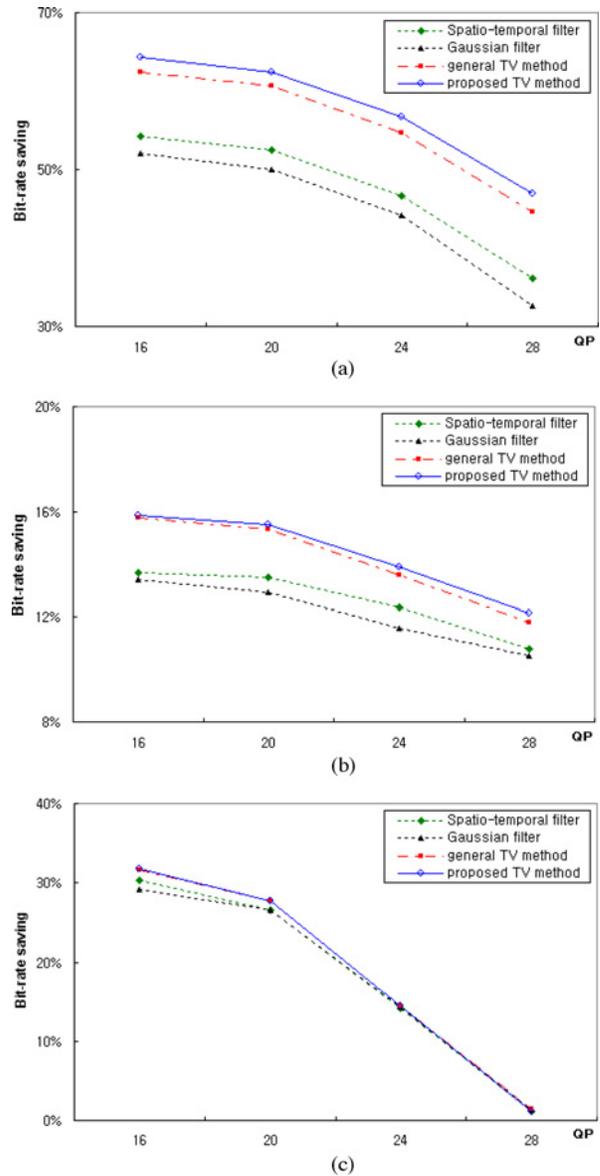


Fig. 13. Coding bit-rate savings comparison of different algorithms as a function of quantization parameters ( $QP$ ). (a) *Rolling tomatoes*, smooth regions: 72.5%. (b) *Playing cards*, smooth regions: 19.5%. (c) *Old town cross*, smooth regions: 37.8%.

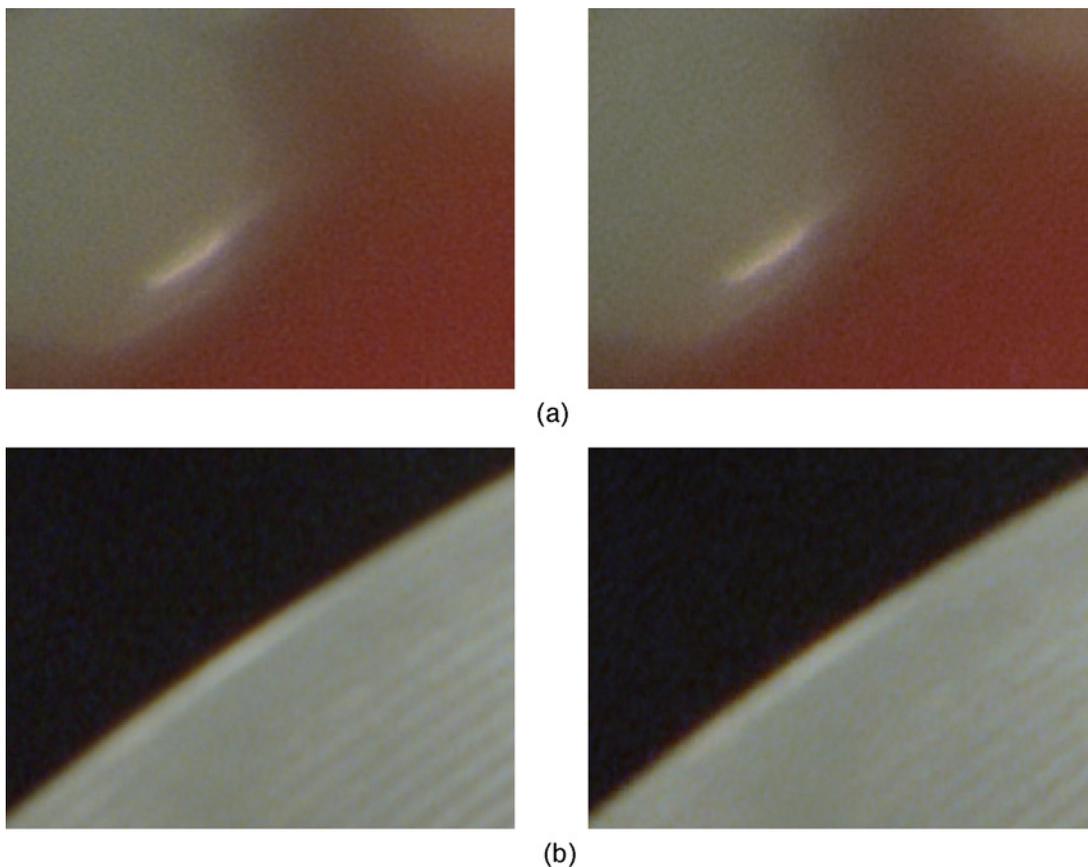


Fig. 14. Close-up of the original images (left) and the resynthesized images (right). (a) *Rolling tomatoes*. (b) *Playing cards*.

property has been considered in noise synthesis, a simple summation of the decoded image and synthesized noise is adopted in (18).

It is worthwhile to mention that, since the film grain noise of edge areas is not removed at the encoder, to add synthesized noise to the edge areas of the decoded image as shown in (18) could cause some problem. However, we observe that noise in edge regions is much less visible than noise in nonedge regions due to the masking effect so that this issue is negligible. Besides, the decoded image is a reconstruction of the smoothed version of the original image. Regardless of the compression method used, film grain noise in edge regions is actually suppressed during the encoding process.

## V. EXPERIMENTAL RESULTS

The film grain noise removal and synthesis processes can be integrated with any typical coding method. We use the JM(ver 11) software [33] of the H.264/AVC reference as the codec module in the experiment. We choose two RGB formatted color and one grey high-definition (HD) video sequences as test sequences. Each sequence has 30 frames, and the first frames of sequences are shown in Fig. 6. Each sequence has different film grain noise characteristics and different type of contents. For example, *Rolling tomatoes* sequence has lots of smooth regions, *Playing cards* sequence mostly consists

of textured region, and *Old town cross* has both smoothed sky and textured buildings so that they provide a set of good test examples to evaluate our proposed scheme. For more experimental results and visual performance evaluation, please visit <http://sites.google.com/site/btoh77/FGN>.

### A. Smooth Region Detection and Denoising

Fig. 7 shows edge maps of the first frame of sequences, where the results of edges extraction and fine texture extraction with the final edge map are given. They demonstrate that our algorithm can detect most of edge regions and fine texture regions successfully. Fig. 8 shows the close-up view of the denoised first frame of each sequence. It is worthwhile to mention that the proposed denoising algorithm is not sensitive to video contents or algorithm parameterization, since the TV minimization method automatically finds and updates the  $\lambda$  value, which is one of the main advantages as compared with other regularization methods.

To demonstrate the superior performance of the proposed denoising algorithm, we consider the power spectrum density (PSD) of extracted film grain noise. That is, we use the temporally extracted noise as the ground-truth and compare it with: 1) Gaussian filtered noise; 2) spatio-temporal filtered noise [11]; 3) noise extracted by the traditional TV algorithm [32]; and 4) noise extracted by the proposed TV algorithm. The squared-roots of PSD of noise extracted by different

algorithms, where the 2-D power spectrum density is projected to 1-D, are shown in Fig. 9. We see that the TV minimization method outperforms the traditional smoothing filter and the spatio-temporal filter.

The superiority of the proposed TV algorithm can be stated below. First, the proposed TV method using spatial correlation and crosscolor correlation detects film grain noise more accurately, especially for low-frequency energy of noise. Since most energy of film grain noise lies in the low-frequency band as mentioned in Section III, the proposed TV method works well. Second, the power of extracted noise using the TV method is closest to the ground truth. To illustrate this point, we compare the power of noise extracted by various algorithms in Table I. To conclude, the proposed TV algorithm can efficiently suppress film grain noise and improve the coding gain, which will be discussed in Section V-C.

### B. Film Grain Noise Synthesis

For synthesized noise evaluation, the conventional metrics such as the MSE or PSNR value are not useful. Here, we consider several criteria to evaluate the performance of different synthesis methods based on unique properties of film grain noise given in Section I. Out of the six properties, temporal independency and Gaussian distribution are automatically satisfied, since we use the independent and identically distributed (i.i.d.) Gaussian noise as the excitation signal. In the frequency domain, the power spectrum density determines the visual appearance of noise. Since the signal with stronger low-frequency components is more visible to human eyes while film grain noise has higher low-frequency components, it is important to resynthesize noise to have a similar power spectrum density. Likewise, the spatial distribution of noise plays an important role for human perception in the spatial domain. In the RGB domain, the correlation between three color channels should be considered in noise generation. Even though pixels in the RGB domain have different values, the same film grain noise is physically created at each pixel and the crosscolor correlation should be preserved. In addition, the background signal with a different intensity has different noise power, which is also perceptually visible. All these criteria will be considered and tested one by one in the following.

Among these criteria, the matching of crosscolor correlation appears to be the most important one since it leads to intensity compensation between color channels. Due to this reason, we use the white signal scaling to preserve the crosscolor correlation as mentioned in Section IV-C. The crosscolor correlation values between the extract and synthesized noise are compared in Table II. The signal dependency property is compared in Fig. 10 while the power spectrum density of the extracted and synthesized noise with the seven-coefficient AR model as described in Section IV-A is compared in Fig. 11. As shown in these two figures, their power spectrum density plots are similar. We can make these curves closer to each other using a higher order AR model or more complicated model such as ARMA model at the cost of higher complexity. Since the performance of a more complicated model is not significantly better, the 7-coefficient AR model is accurate enough as far as the power spectrum density is concerned.

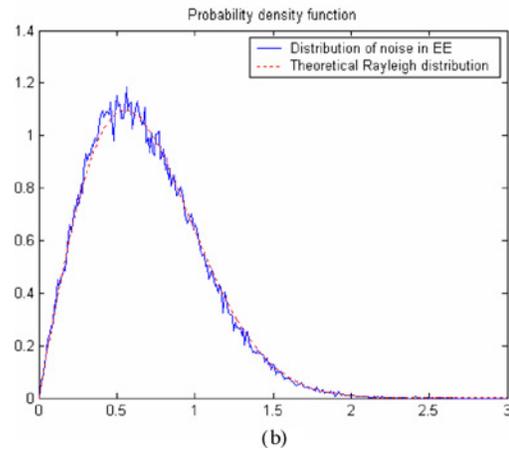
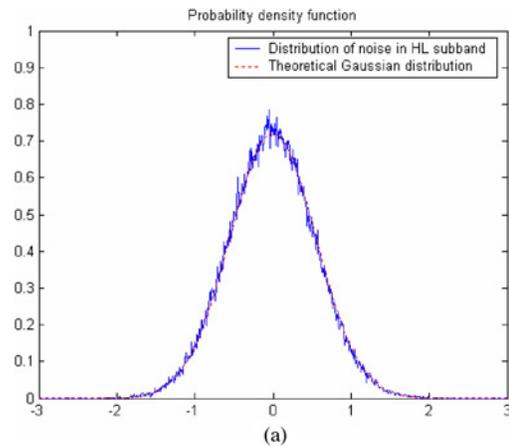


Fig. 15. Distribution of film grain noise and its edge energy values in a typical sub-band, where the *LH* sub-band is used as an example. (a) Distribution of film grain noise in the *LH* subband. (b) Distribution of edge energy values in the *LH* subband.

Finally, we also evaluate the results by subjective quality testing method with Table III as proposed in [34]. Based on the subjective quality rating, we selected 10 experts and 10 nonexperts, and asked them to compare the original video with 1) conventionally decoded video with  $QP = 24$ , and 2) output video by the proposed framework, i.e., denoised, decoded with  $QP = 24$  and noise added sequence. We show the subjective test results in Fig. 12, where each shaded bar and its middle dotted line represent the 95% confidence interval and its mean value, respectively.

### C. Coding Gain Improvement

We show the bit-rate savings as a function of the quantization parameter ( $QP$ ) in Fig. 13 for different denoising algorithms with similar output video quality. Here, we have tested 30 frames (1sec) for each case with one intra-coded (I) frame and 29 predicted coding (P) frames. In Fig. 13, we also provide the portion of smooth regions, since the bit-rate saving is directly affected by this portion. We see that the film grain denoising algorithm significantly reduces the coding bit-rate, especially for smaller  $QP$  values, and denoising with the proposed TV minimization method demands the lowest bit-rate for the same  $QP$ .

Finally, we show parts of two image frames and their corresponding resynthesized counterparts in Fig. 14 for visual performance comparison, where resynthesized images are obtained by adding the synthesized film grain noise to decoded images with  $QP = 24$ . Since the whole image is too big to reveal the advantage of the proposed algorithm, we only show the close-up views that cover the homogenous region (for the sequence of *Rolling tomatoes*) and the edge region (for the sequence of *Playing cards*), respectively.

## VI. CONCLUSION AND DISCUSSION

A novel approach to HD video coding based on film grain noise extraction, modeling, resynthesis was presented. There are several important contributions in our proposed scheme. First, the edge-preserving denoising technique is used for film grain noise extraction. The edge-detection filters and the fine texture detection algorithm allowed us to separate nonsmooth regions from smooth regions in an image. Second, the denoising method based on the modified total variation (TV) minimization method using specific film grain properties was designed to suppress film grain noise efficiently without distorting the original image. Third, noise characteristics were extracted by temporal difference, and they can be used in the denoising process. Last, a novel film grain noise synthesis algorithm was obtained by using the AR model excited by scaled white noise. It was shown that the coding gain is much improved by encoding the denoised image sequence. Furthermore, we show that the synthesized film grain noise is subjectively satisfactory to human eyes and argue that it is objectively similar to observed film grain noise in terms of its statistical properties.

All time-consuming tasks are implemented at the encoder in our proposed algorithm. Only simple noise synthesis and addition is needed at the decoder. Thus, the additional complexity required by consumer electronic devices is negligible. Moreover, it demands only a small number of parameters per frame or per GOP as the overhead. As a result, the proposed film grain noise model can be easily added to the any current video coding standards.

## APPENDIX

To analyze and determine threshold values discussed in Section III-B, we have to determine the distribution of film grain noise first. It is observed that the distribution of film grain noise is close to the Gaussian distribution. Furthermore, the distribution of film grain noise in each sub-band is also close to the Gaussian distribution, which is experimentally verified in Fig. 15(a) for the *LH* sub-band. A similar observation holds for other sub-bands. Note also that sub-bands *LH* and *HL* are independent of each other, since film grain noise is isotropic. For the following analysis, we assume that both *LH* and *HL* sub-bands has the normal distribution with zero-mean and standard deviation  $\sigma$ .

The distribution of the edge energy map of film grain noise can be derived based on the above assumptions. By choosing  $p = 2$  in (4),  $EE$  represents the Euclidian distance of the corresponding pixels in *LH* and *HL* sub-bands. It is known

that  $EE$  has the following Rayleigh distribution

$$P(EE \leq t) = 1 - e^{-\frac{t^2}{2\sigma^2}}. \quad (19)$$

This is experimentally verified in Fig. 15(b). Thanks to the explicit distribution function, threshold value  $th$  can be determined by setting variable  $r$  as

$$\begin{aligned} P(EE \leq th) &= 1 - e^{-\frac{th^2}{2\sigma^2}} = r, \\ th &= \sqrt{-2 \log(1-r)} \cdot \sigma. \end{aligned} \quad (20)$$

In the Rayleigh distribution, its mean is determined by  $\sigma$ , i.e.,  $m = \sigma\sqrt{\pi/2}$ . Thus, the threshold value can be finally determined by (21). For example, if we want to include exactly 99% of grain noise (i.e.,  $r = 0.99$ ), the scaling coefficient  $c$  in (6) should be set to 2.42 due to the following relationship:

$$th = \sqrt{-2 \log(1-r)} \cdot \sqrt{2/\pi} \cdot m. \quad (21)$$

As another option, we may use the median value instead of the mean value since the median is proven to be more robust with respect to outlier signals or noise. In the edge map decision step, there must be false decision, which may lead to undesired perturbation of the threshold value. Under such a scenario, we may use the median to determine the threshold value. Being similar to the previous case, the median is also determined by  $\sigma$ , i.e.,  $\text{med} = \sigma\sqrt{\log 4}$ . Then, the threshold value can be determined as

$$th = \sqrt{-2 \log(1-r)} \cdot \sqrt{1/\log 4} \cdot \text{med}. \quad (22)$$

It is worthwhile to point out that the median-based algorithm needs additional memory to store the signal distribution while the mean-based method can be performed without additional memory due to the dynamic updating process by specified in (6).

## ACKNOWLEDGEMENTS

The authors would like to thank Sharp Laboratories of America for the support of this paper.

## REFERENCES

- [1] L. Stroebel, J. Compton, I. Current, and R. D. Zakia, "Micro-image evaluation," in *Basic Photographic Materials and Processes*, 2nd ed. Boston, MA: Focal Press, Mar. 2000, ch. 11, pp. 257–278.
- [2] "Advanced video coding for generic audiovisual services," document ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC, May 2003.
- [3] C. Gomila and A. Kobilansky, "SEI message for film grain noise," in *Proc. 8th Meeting Joint Video Team*, Document JVT-H022.doc, Geneva, Switzerland, May 2003, pp. 1–14.
- [4] C. Gomila, "SEI message for film grain encoding: Syntax and results," in *Proc. 7th Meeting Joint Video Team*, Doc. JVT-I013.doc, San Diego, CA, 2003, pp. 1–13.
- [5] M. Schlockermann, S. Wittmann, T. Wedi, and S. Kadono, "Film grain coding in H.264/AVC," in *Proc. Joint Video Team 9th Meeting*, Doc. JVT-I034.doc, San Diego, 2003, pp. 1–8.
- [6] B. T. Oh, S. Sun, S. Lei, and C.-C. J. Kuo, "Film grain noise modeling in advanced video coding," in *Proc. SPIE, Visual Commun. Image Process.*, San Jose, CA, vol. 6508, Jan. 2007, p. 650811.
- [7] J. C. K. Yan, P. Campisi, and D. Hatzinakos, "Film grain noise removal and generation for color images," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, 1998, pp. 2957–2960.
- [8] "Film grain technology—Specifications for H.264 | MPEG-4 AVC bit-streams," Society of Motion Picture and Television Engineers (SMPTE) RDD 5-2006, Mar. 2006.

- [9] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simulation (SIAM Interdisc. J.)*, vol. 4, no. 2, pp. 490–530, Jul. 2005.
- [10] M. K. Ozkan, M. I. Sezan, and A. M. Tekalp, "Adaptive motion-compensated filtering of noisy image sequences," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 3, no. 4, pp. 277–290, Aug. 1993.
- [11] H. Y. Cheong, A. M. Tourapis, J. Llach, and J. Boyce, "Adaptive spatio-temporal filtering for video de-noising," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, Oct. 2004, pp. 24–27.
- [12] A. Pizurica, V. Zlokolica, and W. Philips, "Noise reduction in video sequences using wavelet-domain and temporal filtering," in *Proc. SPIE Conf. Wavelet Applcat. Ind. Process.*, 2003, pp. 48–59.
- [13] K. J. Boo and N. K. Bose, "A motion-compensated spatio-temporal filter for image sequences with signal-dependent noise," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 287–298, Jun. 1998.
- [14] O. K. Al-Shaykh and R. M. Mersereau, "Lossy compression of noisy images," *IEEE Transactions on Image Processing*, vol. 7, no. 12, pp. 1641–1652, Dec. 1998.
- [15] J. C. K. Yan and D. Hatzinakos, "Signal-dependent film grain noise removal and generation based on higher order statistics," in *Proc. IEEE Int. Conf. Image Process.*, 1997, pp. 77–81.
- [16] T. M. Moldovan, S. Roth, and M. J. Black, "Denoising archival films using a learned Bayesian model," in *Proc. IEEE Int. Conf. Image Process.*, Atlanta, GA, Oct. 2006, pp. 2641–2644.
- [17] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.
- [18] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 7, pp. 710–732, Jul. 1992.
- [19] Z. Xiong, M. T. Orchard, and Y. Q. Zhang, "A deblocking algorithm for JPEG compressed images using overcomplete wavelet representations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 433–437, Apr. 1997.
- [20] P. Schallauer and R. Morzinger, "Rapid and reliable detection of film grain noise," in *Proc. IEEE Int. Conf. Image Process.*, Atlanta, GA, Oct. 2006, pp. 413–416.
- [21] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. J. Comput. Vision*, vol. 40, no. 1, pp. 49–71, Oct. 2000.
- [22] P. Campisi, D. Hatzinakos, and A. Neri, "A perceptually lossless, model-based, texture compression technique," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1325–1336, Aug. 2000.
- [23] P. Campisi and G. Scarano, "A multiresolution approach for texture synthesis using the circularharmonic functions," *IEEE Trans. Image Process.*, vol. 11, no. 1, pp. 37–51, Jan. 2002.
- [24] P. Schallauer and R. Morzinger, "Film grain synthesis and its application to regraining," in *Proc. Conf. Image Quality Syst. Perform. III*, vol. 6509. 2006, p. 65090Z.
- [25] A. A. Efros and T. K. Leung, "Texture synthesis by nonparametric sampling," in *Proc. Int. Conf. Comput. Vision*, vol. 2, Sep. 1999, pp. 1033–1038.
- [26] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in *Proc. Int. Conf. Comput. Graphics Interactive Tech.*, Aug. 2001, pp. 341–346.
- [27] L. Liang, C. Liu, Y. Xu, B. Guo, and H. Shum, "Real-time texture synthesis by patch-based sampling," *Assoc. Comput. Machinery Trans. Graphics*, vol. 20, no. 3, pp. 127–150, Jul. 2001.
- [28] V. Kwatra, A. Schodl, I. Essa, G. Turk, and S. Bobick, "Graphcut textures: Image and video synthesis using graph cuts," in *Proc. Assoc. Comput. Machinery Special Interest Group Comput. Graphics Interactive Tech. (SIGGRAPH)*, vol. 22, Jul. 2003, pp. 277–286.
- [29] A. Benyassine, E. Shlomot, H. Y. Su, D. Massaloux, C. Lamblin, and J. P. Petit, "ITU recommendation G.729 annex B: A silence compression scheme for use with G.729 optimized for V.70 digital simultaneous voice and data applications," *IEEE Commun. Mag.*, vol. 35, no. 9, pp. 64–73, Sep. 1997.
- [30] A. B. Hamza, H. Krim, and G. B. Unal, "Unifying probabilistic and variational estimation," *IEEE Signal Process. Mag.*, vol. 19, no. 5, pp. 37–47, Sep. 2002.
- [31] D. M. Strong, P. Blomgren, and T. F. Chan, "Spatially adaptive local feature-driven total variation minimizing image restoration," in *Proc. SPIE Annu. Meeting*, vol. 3167, 1997, pp. 222–233.
- [32] L. I. Ludin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenom.*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992.

- [33] Joint Video Team (JVT) reference software version JM [Online]. Available: [http://iphome.hhi.de/suehring/tml/download/old\\_jm](http://iphome.hhi.de/suehring/tml/download/old_jm)
- [34] A. Watson and M. A. Sasse, "Measuring perceived quality of speech and video in multimedia conferencing applications," in *Proc. Assoc. Comput. Machinery Int. Conf. Multimedia*, Sep. 1998, pp. 55–60.



**Byung Tae Oh** received the B.S. degree in electrical engineering from Yonsei University, Seoul, Korea, in 2003 and the M.S. degree in electrical engineering from the University of Southern California, Los Angeles, in 2007. Currently, he is pursuing the Ph.D. degree from the Ming Hsieh Department of Electrical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, and Signal Image Processing Institute.

He was a Summer Intern at Sharp Laboratories of America, Camas, WA, in 2006 and 2007. His research interests include image and video restoration, super-resolution, analysis/synthesis, and pre/postprocessing for coding.



**Shaw-Min Lei** (S'87–M'88–SM'95–F'06) received the B.S. and the M.S. degrees in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1980 and 1982, respectively, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles, in 1988.

From August 1988 to October 1995, he was with Bellcore (Bell Communications Research), Red Bank, NJ, and worked mostly in video compression and communication areas and for a short period of time in wireless communication areas. From October

1995 to March 2007, he was with Sharp Laboratories of America, Camas, WA, as a manager of the Video Coding and Communication Group. Since March 2007, he has been with the MediaTek, Hsinchu, Taiwan, as a Director at the Advanced Technology Division, and working in video/image coding and processing areas. His research interests include video/image compression, processing and communication, picture quality enhancement, multimedia communication, data compression, and error control coding. He has published more than 60 peer-reviewed technical papers and more than 40 contributions to MPEG4, JPEG2000, H.263+, and H.264 international standard meetings. He has been awarded more than 35 patents.

Dr. Lei has received the Best Paper Award (corecipient) of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 1993.



**C.-C. Jay Kuo** (S'83–M'86–SM'92–F'99) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1980, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1985 and 1987, respectively.

He is the Director of the Signal and Image Processing Institute, and a Professor of electrical engineering, computer science, and mathematics at the University of Southern California, Los Angeles.

His research interests include the areas of digital image/video analysis and modeling, multimedia data compression, communication and networking, and biological signal/image processing. He is a coauthor of about 150 journal papers, 780 conference papers and 9 books.

Dr. Kuo is a Fellow of SPIE. He is the Editor-in-Chief of the *Journal of Visual Communication and Image Representation*. He was in the Editorial Board of the IEEE SIGNAL PROCESSING MAGAZINE in 2003, and 2004. He was as an Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING in 1995–1998, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 1995–1997, and IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING in 2001–2003. He received numerous recognition and awards, including the National Science Foundation Young Investigator Award in 1992, the Presidential Faculty Fellow Award in 1993, the USC Junior Faculty Research Award in 1994, the best paper awards from the IEEE Communication Society in 2005, the IEEE Vehicular Technology Fall Conference in 2006, and the IEEE Conference on Intelligent Information Hiding and Multimedia Signal Processing in 2006, and the Okawa Foundation Research Grant in 2007. He was an IEEE Signal Processing Society Distinguished Lecturer in 2006.