IEICE TRANS. INF. & SYST., VOL.E97–D, NO.3 MARCH 2014

PAPER

Bitstream-Level Film Noise Cancellation for Damaged Video Playback

Sinwook LEE[†], *Nonmember and* **Euee-seon JANG**[††a)], *Member*

SUMMARY In this paper, we propose a bitstream-level noise cancellation method for playback applications of damaged video. Most analog video data such as movies, news and historical research videos are now stored in a digital format after a series of conversion processes that include analog-to-digital conversion and compression. In many cases, noise such as blotches and line scratching remaining in analog media are not removed during the conversion process. On the other hand, noise is propagated in the compression stage because most media compression technologies use predictive coding. Therefore, it is imperative to efficiently remove or reduce the artifacts caused by noise as much as possible. In some cases, the video data with historical values are to be preserved without correcting the noise in order not to lose any important information resulting from the noise removal process. However, playback applications of such video data still need to undergo a noise reduction process to ensure picture quality for public viewing. The proposed algorithm identifies the candidate noise blocks at the bitstream-level to directly provide a noise reduction process while decoding the bitstream. Throughout the experimental results, we confirm the efficiency of the proposed method by showing *RR* and *PR* values of around 70 percent.

key words: *film noise removal, bistream-level, two pass approach*

1. Introduction

The migration of audio-visual collections from analog to digital formats has been identified as one of the main challenges for long-term preservation. A UNESCO report estimated that there are world audiovisual holdings such as audio, video and film recordings total about 200 million hours. Some expect that these materials will be endangered within the next 20 years. Currently, many audio-visual archives belonging to governments, universities, libraries, museums and private collections are being converted to digital.

One of the major problems in digital conversion of audio-visual collections is handling the noise during the conversion process. Noises in analog audio-visual collections such as film scratches are often resident in the medium. This causes severe degradation of quality during the analog-to-digital conversion and error propagation during compression.

Many researchers have tried noise removal techniques after analog-to-digital conversion before compression. This is because noise removal is better performed at the earliest possible stage, analog-to-digital conversion. Otherwise, the noises are likely to be enlarged or propagated during the compression stage. Eventually, most researches on noise detection and removal are based on the pure film noise before compression [2]–[16].

However, there are several difficulties in the noise removal process right after digitization: the absence of noise-free original data, the variety of noise types (e.g., line scratch noise and blotch noise), and the risk of false detections. Different from conventional digital media processing, noise removal techniques cannot easily be evaluated in an objective manner. This is due to the fact that there is no noise-free original content. Therefore, the evaluation of the noise removal process is mainly done subjectively.

Another difficulty is the variety of noise types, which makes it significant problem to identify what type of noise is present in a given medium even before applying any noise removal method. This is an important problem because there is no generic noise removal technique but several noise removal techniques according to noise type.

Finally, there is a risk of false detection of noise causing removal or alteration of a non-noisy region. This situation may happen when content originally contains many noisy elements such as detailed texture images, foggy scenery, and other complex images. Since there is no noise-free original data, it is very difficult to guarantee the confidence of noise detection and removal.

For most applications, the degradation resulting from the noise removal process and compression can be ignored. The threshold of acceptable quality for such applications is not as high because there is no original noise-free data to compare. Moreover, noise-removed contents are usually more pleasing in terms of subjective quality.

However, there are certain applications that require more caution in noise removal because the noise removal process may alter or remove regions that contain no noise. Such applications include historical archives and government archives in which the noise-embedded contents are considered valuable as is. In these cases, high quality compression without any noise removal is done only to preserve noise-free regions as effectively as possible.

Even for historical and government archives, noise-removed playback services are often necessary for public view. Many national archive organizations such as the National Archives of Korea, the Library and Archives Canada, and the National Archives and Records Administration already provide playback services without further noise treatment. It is often desirable to enhance the visual quality of

Manuscript received September 11, 2013.
Manuscript revised May 24, 2013.

[†]The author is with the Department of Electronics Computer Engineering, Hanyang University, Seoul 133–791, Korea.
[††]The author is with the Department of Computer Science & Engineering, Hanyang University, Seoul 133–791, Korea.
a) E-mail: esjang@hanyang.ac.kr

DOI: 10.1587/transinf.E97.D.562

damaged video contents on the fly. In such a case, the noise removal process should be conducted from the encoded bitstream of the noise-embedded contents. This process after compression should deal with a different set of problems from that before compression: the noises are altered (e.g., enlarged, propagated, or quantized) after compression. Therefore, such characteristics need to be taken into consideration after compression.

In this paper, we propose a bitstream-level noise removal technique that detects and removes noise from embedded noise to compression noise. The novelty of the proposed method lies in the consideration of noise detection and removal after compression by exploiting the noise characteristics at the bitstream level. This paper is organized as follows. In Sect. 2, we review the conventional noise modeling and removal process. The proposed method is described in detail in Sect. 3. In Sect. 4, we provide the experimental results and conclude the paper in Sect. 5.

2. Background

2.1 Noise Analysis

In the previous section, we briefly addressed the two different origins of noise during the digital conversion process: embedded noise (e.g., line scratches, blotches, random white Gaussian noises, and sampling noises) and compression noise (e.g., temporal noise propagation, spatial noise propagation, and quantization noises). More specifically, the digital conversion process with noise can be depicted as shown in Fig. 1.

Analog noises such as line scratches and blotches are caused by improper handling of the original content. A generic noise model can be represented as follows [1]:

$$\hat{I} = \mu * I + \alpha, \tag{1}$$

where $\hat{I}$ denotes the degraded image, I the image before degradation, μ the multiplicative noise, and α the additive noise. At each stage of the analog-to-digital conversion process, both multiplicative and additive noises are added to the content. For the impaired content (A_i') in Fig. 1, (1) can be rewritten as the following function:

$$A_i' = f(A_i, \mu_a, \alpha_a), \tag{2}$$

where i denotes the frame number and μ_a, α_a describes the noises by improper handling. After the analog-to-digital conversion process, the digitized content (D_i) can be represented as the following function:

$$\begin{aligned} D_i &= g(A_i', \mu_d, \alpha_d) \\ &= g(f(A_i, \mu_a, \alpha_a), \mu_d, \alpha_d), \end{aligned} \tag{3}$$

where μ_d and α_d describe the noise using the digital conversion process. From the equation, it is apparent that both analog and digital noises are combined together. Therefore, the noise removal process at this stage has to take this combined noises into consideration. After the compression stage, the degraded image (C_i) can be represented as the following function:

$$C_i = \begin{cases} h(D_i, \mu_c, \alpha_c) \\ = h(g(f(i, \mu_a, \alpha_a), \mu_d, \alpha_d), \mu_c, \alpha_c) \\ \qquad\qquad \text{, if } C_i \text{ is intra frame} \\ h(D_i, \mu_c, \alpha_c, \tau_{i-1}C_{i-1}, \tau_{i-2}C_{i-2}, \ldots, \tau_{i-d+1}C_{i-d+1}) \\ = h(g(f(A_i, \mu_a, \alpha_a), \mu_d, \alpha_d), \mu_c, \alpha_c, \\ \quad \tau_{i-1}C_{i-1}, \ldots, \tau_{i-d+1}C_{i-d+1}) \\ \qquad\qquad \text{, if } C_i \text{ is inter frame} \end{cases} \tag{4}$$

where μ_c, α_c describes the noises by compression process, τ_i a the weighting coefficient of the i-th frame, and d the distance between the current frame and the previous intra frame. Only the spatial compression noises of the current frame are added in intra frames, whereas the spatial compression noises from the previous frames are propagated as temporal compression noises in inter frames. Therefore, the noise removal process should consider the compression noises on top of the embedded noises before compression. In particular, Greater consideration should be given to compression noises because they not only further degrade the content by compression (i.e., spatial compression noise) but also propagate the noise temporally to the following frames (i.e., temporal compression noise).

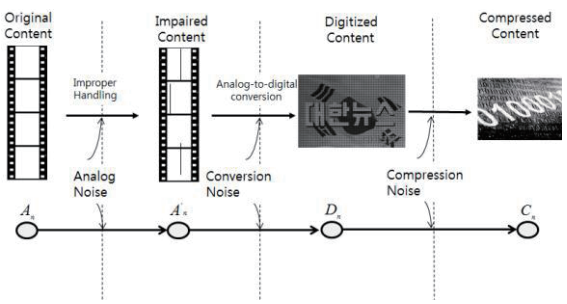


Fig. 1 Noise occurrence in the digital conversion process.

2.2 Conventional Noise Removal Techniques

Conventional noise removal techniques have been focused on the digitized content (D_i) because it is the earliest possible stage to remove the noise. As shown in Eq. (3), the digitized content contains both analog and conversion noises, and most noise removal techniques concentrate on removing these noises.

Noise removal techniques consist of two processes: noise detection and restoration. The noise detection process is usually conducted on a predetermined noise type (e.g., line scratch, blotch, or random white noise). In the restoration process, the region with the detected noises is restored with new pixel values computed using the spatial and temporal neighboring noise-free pixel information.

There are two methods for the detection of the analog noise: spatial noise detection and temporal noise detection. For the spatial noise detection of line scratches, various algorithms are used including the Hough transform, the neural

network-based classifier, morphological filtering, overcomplete wavelet expansion and other techniques [2]–[6]. For the temporal noise detection of line scratches, the Kalman filter and a line scratch tracking algorithm by tracking multiple frames have been proposed [7], [8].

For the spatial detection of blotches, median and morphological filtering is often used [9]–[12]. For the temporal detection of blotches, motion compensation-based methods such as the spike detection index and ranked order difference have been proposed [13]–[16].

The detection of digital noise such as random white noise is often conducted in the spatial domain. The detection of this type of noise can be done using one of three types of detection methods: block-based, filtering-based, and transform-based. In block-based methods, the standard deviation of intensity is computed to find the blocks smoothed by noise [17]–[19]. In filtering-based methods, the white noise is detected by computing the standard deviation between the noisy image and its filtered image [19]–[21]. In transform-based methods, many researchers have analyzed the relation among the coefficients in various transform domains such as DCT, DFT or wavelet [22], [23].

The restoration methods of the detected noise regions are mainly based on the interpolation algorithm using the neighboring pixel information. In interpolation methods, there have been many proposed methods using polynomial interpolation, autoregressive model, Markov random field, directional matching method, transform-based interpolation, and other methods.

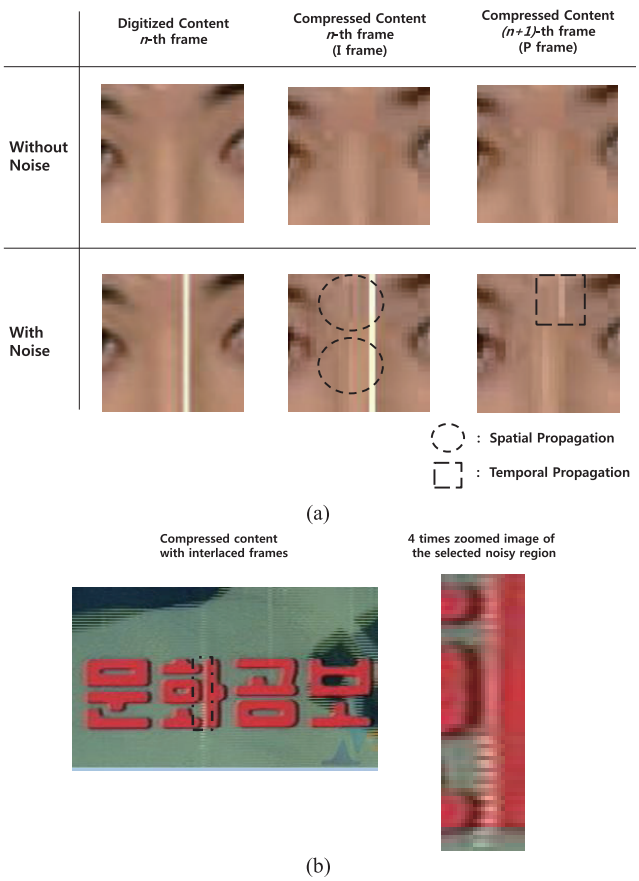


Fig. 2 Examples of spatial and temporal propagations. (a) spatial and temporal propagations, (b) noise contamination in the interlaced frame.

2.3 Characteristics of Compressed Noise

Different from noise in digitized content, noise in the compressed domain has several compression-dependent features: spatial propagation and temporal propagation. In compressed content, coding processes such as motion estimation, intra prediction, and the quantization process may result in spatial and temporal propagation of the noise. Spatial propagation usually appears when the noisy pixels are encoded with the non-noisy pixels as a block for transformation and quantization.

An example of spatial and temporal propagation can be found in Fig. 2-(a), which shows how noise can influence the content through spatial and temporal propagations during the compression process. By comparison in the figure, it is apparent that the existence of noise in the content influences the picture quality in compression due to its edge-intensive nature. We also observed that much digital content is compressed in interlaced format where the noise of a frame may appear in a field level as shown in Fig. 2-(b).

3. Proposed Method

The proposed method is composed of three additional processes on top of the regular decoding process: candidate noise block (CNB) selection, noise detection in the noise block, and a noise removal process as shown in Fig. 3. We exploit the statistics of the decoded discrete cosine transform (DCT) coefficients to select CNB. When a macroblock is selected as CNB, the noise region is identified from the decoded pixels in the noise detection process. The detected noise region is then replaced by the interpolation from the noise-free neighboring pixels in the noise removal process. A detail algorithm of each process is described in the following sections.

3.1 Candidate Noise Block Selection

At the bitstream level, an obvious way to select a CNB is

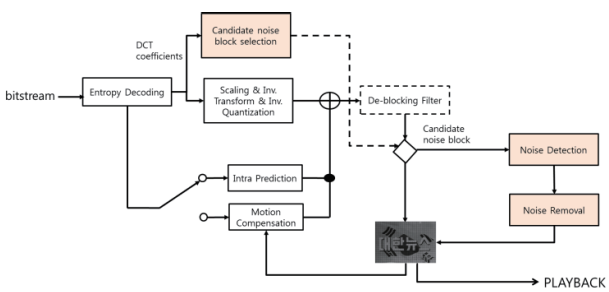


Fig. 3 Block Diagram of the video coding standard with the proposed method.

0	0	0	1	1	1	1	1	F_R
0	0	0	1	1	1	1	1	
0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	

Fig. 4 Definition of F_R in the proposed method.

to exploit the DCT coefficients. Traditionally, the DCT is used in video compression because the most energy of the video signal is concentrated at a low frequency. However, the proposed method exploits the mid-to-high frequency coefficients to select CNB since a detailed pattern of the block texture is well-represented at mid-to-high frequency coefficients. A noise-embedded block is very similar to an edge block in that it contains more non-zero DCT coefficients than a smooth block. Therefore, the number of non-zero coefficients in the mid-to-high frequency range can be used as an indicator to select a noise-embedded block. It is still possible that edge blocks may be selected as CNB, as these will go through a precise detection algorithm in the next stage.

More specifically, CNB can be selected by evaluating the number of non-zero DCT coefficients in the mid-to-high frequency region (N) as follows:

$$f(c) = \begin{cases} 1, & c \neq 0, \quad N = \sum f(c_{\underline{R}}), \quad c_R \in F_R, \\ 0, & otherwise \end{cases} \tag{5}$$

where c_R denotes the DCT coefficients in the mid-to-high frequency region, F_R. The mid-to-high frequency region (F_R) should be defined such that the vertical edge characteristics of the analog film noise can be emphasized in the selected region. Since the characteristics of vertical edge are well represented with the DCT coefficients in the right-top lines, we define FR in the proposed method as Fig. 4. A compact definition of FR is crucial because it affects the time complexity of the proposed method.

The number of non-zero coefficients in the mid-to-high frequency region (N) can be used to determine a CNB by comparing N with a threshold. The threshold value (T) can be determined as follows:

$$T = \arg\max_i \left(\frac{P_{noise}|i}{(1 - P_{noise})|i} \right), \quad 0 \leq i \leq |F_R| \tag{6}$$

where $P_{noise}|i$ denotes the probability that a noise will be detected when the number of DCT coefficients within F_R is larger than i. This equation is drawn from the observation that a noise-embedded block is likely to have a higher value of N than a noise-free block. Once the threshold value (T) is obtained, CNB can be selected when the number of non-zero coefficients at mid-to-high frequency (N) is greater than or equal to the threshold value (T).

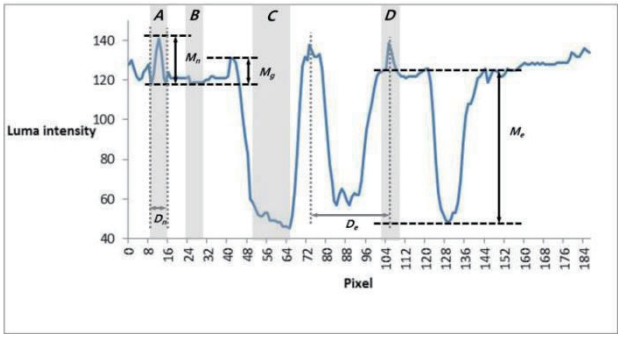


Fig. 5 Magnitude and duration of film noise (Korean flag).

3.2 Noise Detection

The main objective of the CNB selection process is to detect as many noise embedded blocks as CNBs as possible. As a side effect, it is inevitable that some noise-free blocks are also included as CNBs. The purpose of the noise detection process is to filter out the noise-free blocks among the CNBs in the pixel domain.

Specifically, this process distinguishes the noisy pixels from the edge pixels, which is not simple because noisy pixels can be incorrectly identified as edge pixels, as shown in Fig. 5. Another important issue in the noise detection process is distinguishing between film noise and Gaussian noise. Film noise can easily be confused as Gaussian noise because of the fluctuation of the pixel values. However, the magnitude of fluctuation (MOF) in a film noise (M_n) is generally greater than that in a Gaussian noise (M_g). Film noise can be distinguished from the edge region by the duration of film noise (D_n) because the fluctuation created by the film noise is short in duration and returns to the original pixel values, whereas the fluctuation of the edge pixels usually change the pixel values drastically after fluctuation.

In the figure, the pixels can be categorized into three distinct regions (i.e., Gaussian noise region, film noise region, and edge region). Region B indicates a Gaussian noise region, Regions A and D film noise regions, and Region C, the edge region. In the figure, Region B is classified as a Gaussian noise region because the MOF in Region B (M_B) is lower than those in the other regions. Likewise, Region C is classified as an edge region because the duration of Region C (D_c) is much greater than those of Regions A and D (D_A and D_D). As a summary, a film noise map can be determined by the following function ($g(x, y)$).

$$g(x, y) = \begin{cases} 1, & if\ M(x, y) > T_M\ and\ D(x, y) < T_D \\ 0, & otherwise \end{cases} \tag{7}$$

where $M(x, y)$ and $D(x, y)$ denote the functions that return the magnitude and duration values for the corresponding pixel location, respectively. The functions $M(x, y)$ and $D(x, y)$ can be obtained as shown in Fig. 6. T_M and T_D the threshold values empirically obtained from the Gaussian noise region and the edge region of the given image. The Gaussian noise in an 8-bit depth video is assumed to be

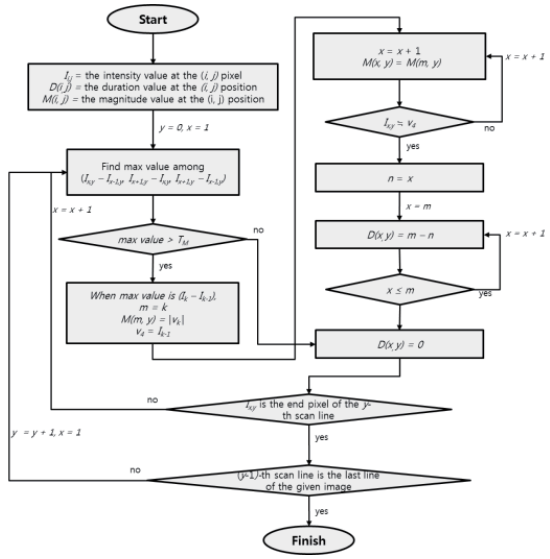


Fig. 6 Flow chart of obtaining the functions $M(x, y)$ and $D(x, y)$.

distributed close to the original pixel value, which is often represented within +/- 8. Therefore, we set T_M to nine in the given sequences. T_D value can be appropriately chosen depending on the durations of the edge regions and is set to six in our simulation.

Generally speaking, it is often difficult to distinguish film noise from Gaussian noise. Even when we choose a good threshold value as in Eq. (7), the likelihood of correct detection is not high for the areas where the distinction between film noise and Gaussian noise is very difficult both objectively and subjectively. One solution to increase the likelihood of correct detection is to exploit a characteristic of film noise: noise connectivity over multiple scan lines. In other words, a pixel region with a relatively low MOF value can be classified as a film noise region if the same area at the previous scan line is a film noise region. Incorporating this concept into Eq. (7), we get the following equation:

$$g(x,y) = \begin{cases} 1, & \text{if } g(x-1,y) = 0 \text{ and } M(x,y) > T_M \\ & \quad \text{and } D(x,y) < T_D \\ 1, & \text{if } g(x-1,y) = 1 \text{ and } M(x,y) > T_{M'} \\ & \quad \text{and } D(x,y) < T_D, T_{M'} < T_M \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

where $T_{M'}$ denotes the additional threshold used when the noise region is detected in the previous scan line.

In Fig. 5, Region D is classified as a film noise region. However, depending on the content characteristics it could still be identified as a part of the edge region. In such a case, an additional process is necessary to determine if it is a film noise region. This is done simply by comparing the chroma values of the film noise region with those of the neighboring pixels.

It should be noted that the noise detection process is conducted in the pixel domain, whereas the CNB selection process is done in the frequency domain. This two-step approach helps finding an active block in the frequency domain and filtering out noise-free blocks (or Gaussian noise blocks) in the pixel domain. The CNB selection process is also helpful to reduce the computational complexity of the noise detection process by selecting fewer blocks to go through the noise detection process. Although all blocks are tested in the CNB selection process, the number of processing samples (e.g., selected DCT coefficients) of the CNB selection process is far less than the number of processing samples (e.g., a block of pixels) of the noise detection process. Empirically, the two-step approach is faster than just the noise detection process by more than two times. This result implies that most blocks are Gaussian noise regions rather than film noise or edge regions.

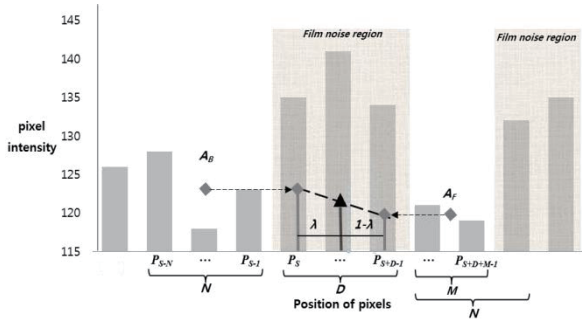


Fig. 7 An example of film noise removal.

3.3 Noise Removal

Once determined as a film noise region, the noise pixels need to be replaced with noise-free pixels. However, the original noise-free pixels are not available because of film degradation over time. In such a case, a useful alternative is to interpolate the noisy region with the non-noisy surrounding pixels. The main objective of interpolation is to approximate the original pixels as much as possible with reasonable computational complexity. Depending on the number of surrounding pixels, there may be many different types of interpolation methods (e.g., nearest-neighbor interpolation, linear interpolation, cubic interpolation).

It is generally true that the performance gain of an interpolation algorithm is proportional to the computational complexity of the algorithm. However, simple interpolation algorithms such as nearest-neighbor interpolation or linear interpolation are preferred to the more sophisticated algorithms in many video coding applications. This is due to the fact that the performance gain difference between algorithms is not so great; instead, the computational complexity often becomes a great concern in order to satisfy real-time performance. In this paper, we chose linear interpolation for a similar reason. In fact, the duration of film noise is often less than the width of a block (i.e., eight pixels). Therefore, we assume that the performance gain difference between linear interpolation and more sophisticated interpolations is very marginal.

Figure 7 shows an example of how noisy pixels are re-

placed by linear interpolation. For linear interpolation, we need two pixel values: one from the backward neighboring pixels and the other from the forward neighboring pixels. In order to minimize the impact of Gaussian noise during linear interpolation, the average value of the neighboring pixels is used for linear interpolation. The backward and forward average values of neighboring pixels (i.e., A_B and A_F) can be computed as follows:

$$\begin{aligned} A_B &= \frac{1}{N}\left(\sum_{k=1}^{N} P_{S-k}\right), \qquad N > 0, \\ A_F &= \frac{1}{M}\left(\sum_{k=0}^{m-1} P_{S+D+k}\right), \quad N \geq M > 0, \end{aligned} \tag{9}$$

where N and M denote the number of backward and forward neighboring noise-free pixels, respectively, and PS is the pixel value at the first pixel of the film noise region. The number of neighboring pixels that are used to compute the backward average (N) does not change, while that of the forward average (M) may be determined between one and N. This is because there may be detected noisy pixels in the forward neighboring pixels. In the case of the backward neighboring pixels, the value of N remains unchanged because all the neighboring pixels are either noise-free or noise-removed.

Additionally, only the available neighboring pixels are used for computing the average when the noise region is close to the left or right boundary of a video frame. If there is no available neighboring pixel to be computed, the other average value can be used instead. After A_B and A_F are obtained, the film noise pixel value at the i-th location (P_i) can be replaced as follows:

$$\begin{aligned} &\lambda = \left(\frac{1}{D-1}\right)(S-i)+1, \quad S \leq i \leq S+D-1, \\ &P_i = \lambda A_B + (1-\lambda)A_F. \end{aligned} \tag{10}$$

3.4 Considerations for Interlaced Material

Much few film content is encoded as interlaced frames even when the analog medium is not interlaced. For this reason, many encoded film contents contain interlace coding, which poses an interesting challenge to noise detection and removal. In interlace video coding, two consecutive video frames are compressed into a single frame where the scan lines of the two frames are interlaced in the same frame. One notable characteristic of interlaced video content is that film noise seldom appears in both frames at the same location, which can be exploited further in noise detection and removal.

The CNB selection process in the proposed method does not require any modification for interlaced video contents because most video coding standards (e.g., MPEG-2, MPEG-4 part 2, and MPEG-4 part 10 AVC) provide two DCT coding types to indicate whether the current macroblock separates even and odd fields or not. When the even and odd fields are separately encoded (e.g., DCT coding type = field DCT), each DCT block can be treated the same as one in a non-interlaced frame.

Otherwise (e.g., DCT coding type = frame DCT), the encoder may choose to combine even and odd fields together because the two fields are homogeneous and, hence, better encoded together. In such a case, we can apply the CNB selection process without any further consideration because such a block seldom contains a noise region—it is not likely that a film noise can be repeated over two frames at the same location. When it comes to the noise detection process, the previous scan line must be referred to in the same field because the reconstructed macroblock is the interlaced image of the two fields. Thus, in the case of the interlaced video content, Eq. (8) should be modified as follows:

$$g(x,y) = \begin{cases} 1, & \text{if } g(x-2,y) = 0 \text{ and } M(x,y) > T_M \\ & \quad \text{and } D(x,y) < T_D \\ 1, & \text{if } g(x-2,y) = 1 \text{ and } M(x,y) > T_{M'} \\ & \quad \text{and } D(x,y) < T_D \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

The noise removal process does not have to be changed because the removal process is conducted only in the same scan line.

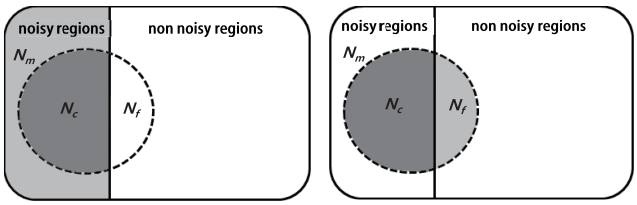
4. Experimental Results

We evaluated the performance of the proposed method with three old film archive data compressed with film noise as described in Table 1. The number of noise regions in each video sequence was subjectively counted as 140, 16 and 87, respectively. These archive films were compressed by the MPEG-2 main profile using interlaced mode. Two of them are grayscale contents without any chroma frames. They were produced in 1982 and 1961, and preserved in the National Archives of Korea (NAK).

However, it is not easy to evaluate objective performance of archival film because there is no noise-free film for comparison. As an efficient evaluation method, we attempted to estimate the recall rate (*RR*) and precision rate

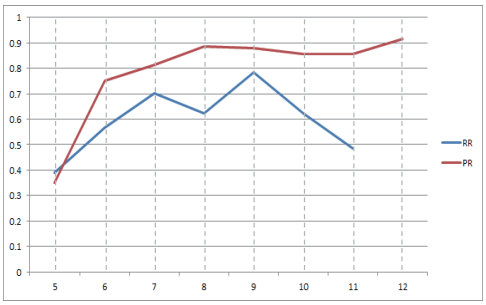
Table 1 Compression features of the test sequences.

Sequences	*Korean flag*	*Panel and projector*	*Speech*
Compression standard	MPEG-2 Main Profile		
Resolution	704 x 480		
Interlaced	Yes		
GOP structure	I : P frames (1 : 11)		
Y:Cb:Cr	4:2:0	4:0:0	4:0:0
Frame rate	29.97		
Production year	1982	1961	1961
Source	National Archives of Korea		
# of Noise	140	16	87

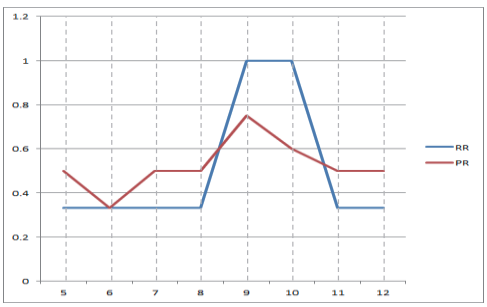


(a) (b)

Fig. 8 Van diagram about recall and precision. (a) *RR*, (b) *PR*.



(a)



(b)

Fig. 9 *RR* and *PR* results depending on the value of T_M. (a) 'Korean flag', (b) 'Panel and projector'.

(*PR*), a method already used in many noise detection papers [4], [14]. This can be described as follows:

$$RR = \frac{N_c}{N_c + N_m}, \tag{12}$$

$$PR = \frac{N_c}{N_c + N_f}, \tag{13}$$

where N_c, N_m, and N_f denote the number of correctly detected, missed and false alarmed regions, respectively. As shown in Fig. 8, *RR* represents the ratio of the correctly detected noise (N_c) over the noisy regions. *PR* represents the ratio of the correctly detected noise over all the detected regions.

Figure 9 shows the *RR* and *PR* results of the proposed method while varying the value of T_M with "Waving Korean flag" and "Panel and projector" sequences. The goal of a good noise detection and removal method is to maximize the *PR* and *RR* values. Therefore, the selection of the proper threshold values of the magnitude and duration (i.e., T_M, $T_{M'}$, and T_D) directly influences the performance of the proposed method. As can be seen from the figure, the *RR* and *PR* performances vary from one sequence to another. However, we observe that the maximum value of *RR* and *PR* combined is found when T_M is set to nine. Similarly, the values of $T_{M'}$ and T_D are set to five and six, respectively.

Table 2 Detection results of the proposed method.

Sequence	Noise type	N_c	N_m	N_f	*RR* (%)	*PR* (%)	∇t (%)
Korean flag	PS	0	0	0	0	0	28.98
	SS	95	42	26	69.3	78.5	
	Blotch	0	3	1	0	0	
	Total	95	45	27	67.9	77.9	
Panel and projector	PS	0	0	0	0	0	27.95
	SS	16	0	26	100	31.6	
	Blotch	0	0	3	0	0	
	Total	16	0	29	100	29.3	
Speech	PS	54	31	0	63.5	100	61.66
	SS	2	0	3	100	40	
	Blotch	0	0	0	0	0	
	Total	56	31	3	64.4	94.9	
Total		163	76	59	68.2	73.4	-
Average		-	-	-	-	-	39.53

PS=Principal Scratch, SS=Secondary Scratch
TM=9, TM'=5, TD=6, 12 frames
∇t = the time with CNB process / the time without CNB process

We tested the proposed method with three compressed film video contents. The performance results in detail are given in Table 2. In this table, it is shown that *RR* and *PR* for all the noises are around 70 percent. The *RR* and *PR* performance obtained in Table 2 cannot be directly compared with other works because the test materials of the proposed method are compressed, while the test materials of other works are not. Considering that the *RR* and *PR* results of other works have been reported ranging from 55 to 80 percent, it can be said that the performance of the proposed method is comparable to other works even when the noise detection and removal is done in the compressed contents. Furthermore, it is confirmed that the CNB process is remarkably efficient in the time complexity reduction from the test results in which the CNB process reduces time complexity of the proposed method by about 60%.

The "Waving Korean flag" sequence contains 12 noises: 11 scratch and one blotch noise, which can be found in the original image as shown in the Fig. 10 (a). The detected noise regions are shown in Fig. 10 (b), whereas Fig. 10 (c) shows the missed and false alarmed noise regions. Figure 10 (d) shows the reconstructed image by the proposed method. The subjective quality of Fig. 10 (d) represents the performance of the proposed method reasonably well. Even if some noises are missed and other non-noisy regions are falsely detected, the proposed method successfully removed the most apparent noise regions, as expected, as shown in Fig. 10 (e) and Fig. 10 (f). One example of the missed noise regions is shown in Fig. 10 (g). In such a case,

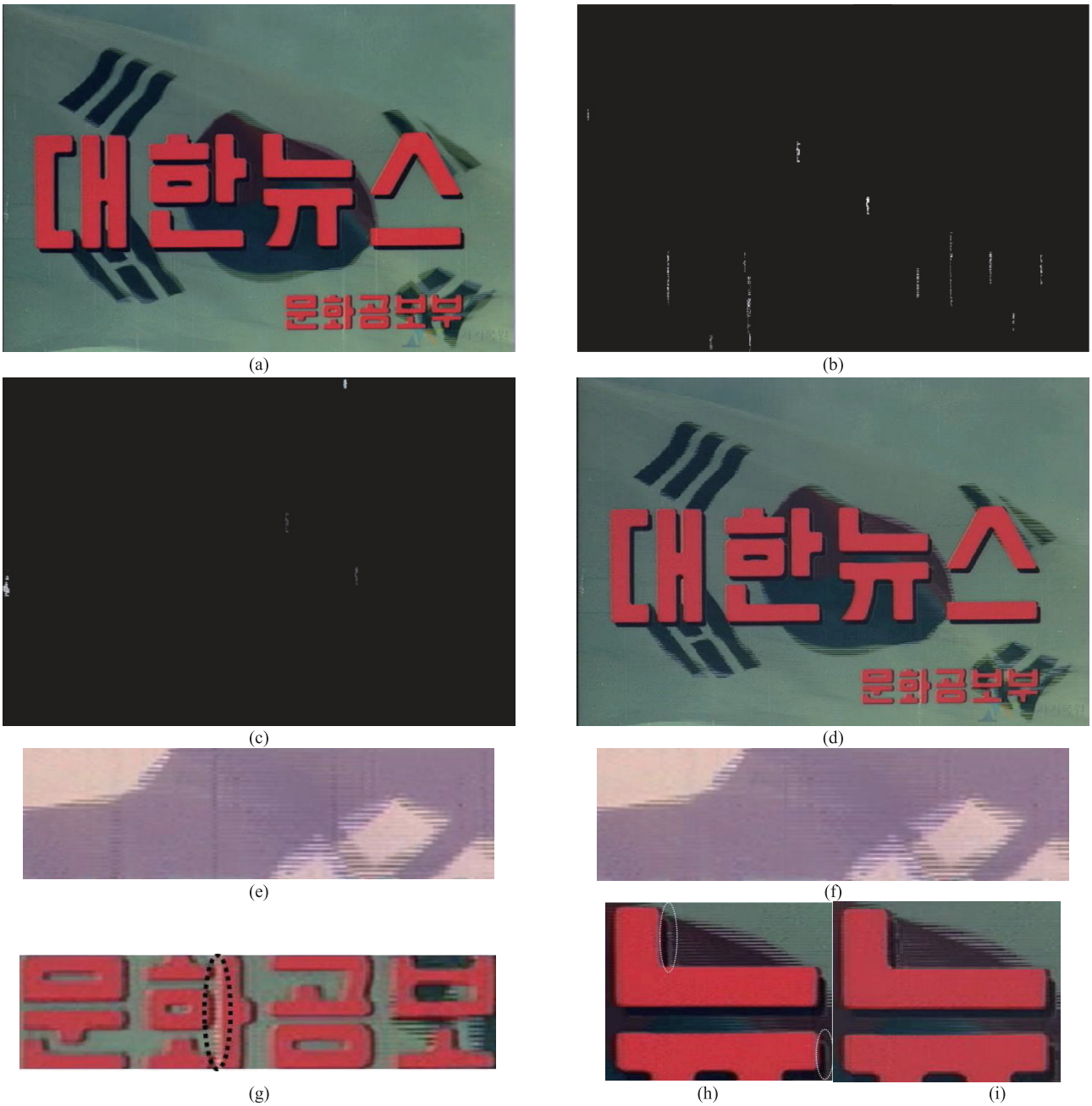


Fig. 10 Noise detection and removal results of the "Korean flag" sequence (1st frame). (a) Original video image, (b) detected noise regions, (c) missed and false alarmed noise regions (white color: missed, gray color: false alarmed), (d) reconstructed video image by the proposed method, (e) detected noise regions in the inversed original image, (f) detected noise regions in the inversed reconstructed image, (g) missed noise region, (h) false alarmed regions in the original image, (i) false alarmed region in the reconstructed image.

the problem is difficult to solve because the film noise is collocated where the edge region is. In the case of a false alarm region as shown in Fig. 10 (h) and Fig. 10 (i), noise detection and removal is conducted as a noise region. This can happen if the pixels show the same characteristics the noise region, which is an inherent problem in noise detection and removal.

Some film contents containing few film noises can be used as a good indicator on how a noise detection and removal method performs. A good example is the "Panel and projector" sequence shown in Fig. 11. Because the number of film noises is very few, the *RR* value of the proposed method is 100 percent. However, we found that some false alarm noises are also detected where the image component is very similar to the film noise. Even when the false alarm regions are corrected, the overall quality of the reconstructed

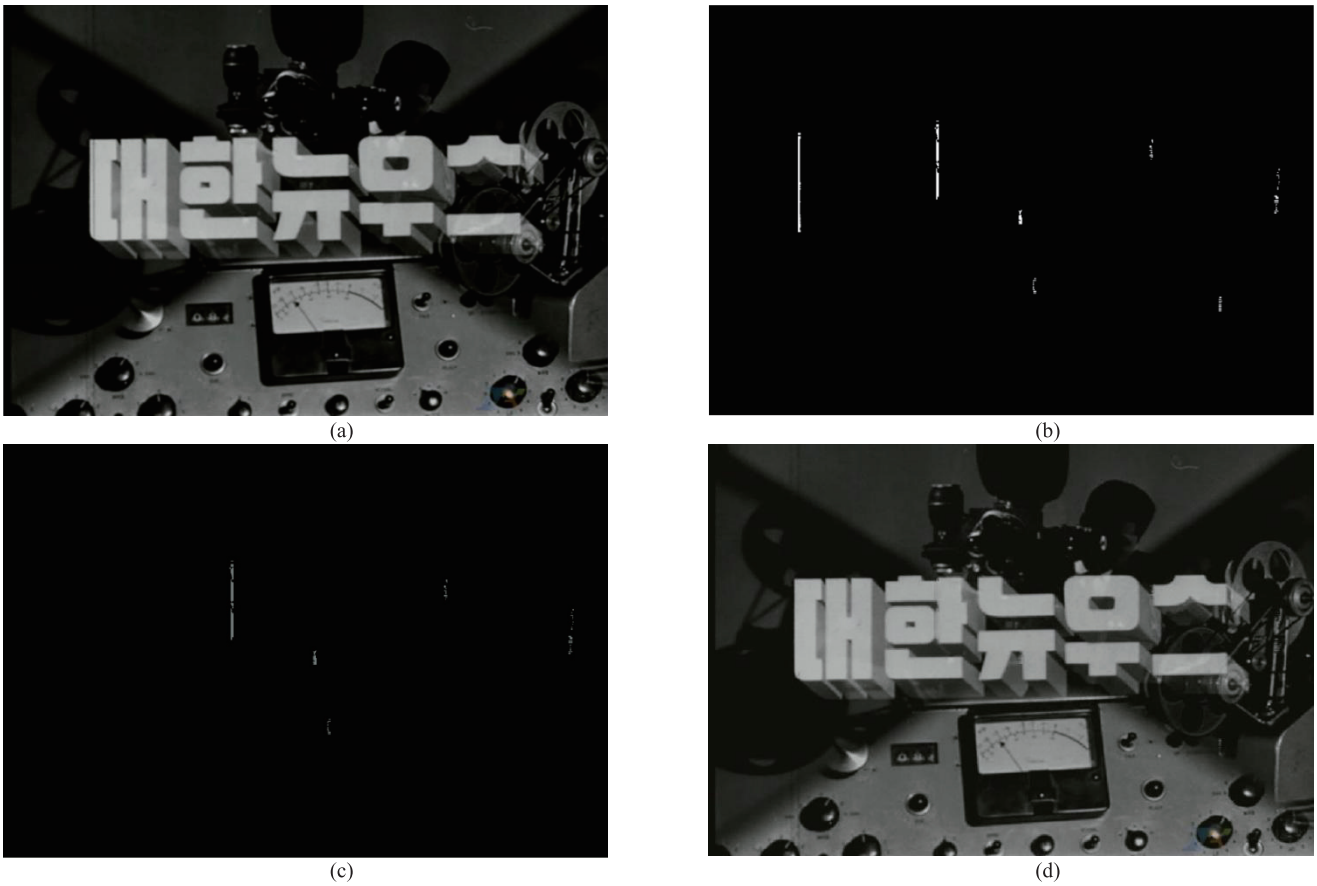


Fig. 11 Noise detection and removal results of the "Panel and Projector" sequence (12th frame). (a) Original video image, (b) detected noise regions, (c) missed and false alarmed noise regions (white color: missed, gray color: false alarmed), (d) reconstructed video image by the proposed method.

image is better than the original compressed content. We provide some snapshots with the "Speech" sequence in Fig. 12. From the reconstructed image (i.e., Fig. 12 (d)), it can be seen that the subjective performance of the proposed method is similar to earlier examples. However, one notable difference is that there are some regions that seem to contain film noises even though they are not clearly visible. This is due to the fact that some film noises that are not detected in the previous frames are propagated through frames by inter frame coding. These temporal noises are apt to be emphasized through inter frame coding. Even though visual quality improvement from the proposed method is apparent, there is room for further improvement as a future study to more effectively cope with temporal noise propagation.

5. Conclusion

We have presented a bitstream-level noise cancellation method as a means to improve visual quality of compressed video contents with film noise. Our proposed method filters out film noise using a two-pass approach: one is conducted in the frequency domain and the other in the pixel domain. In the noise removal process, the searched film noise region is interpolated with the neighboring noise-free pixels. Throughout the experiments, we confirmed the feasibility of the proposed method in removing the film noise embedded in the bitstream. Because many analog video contents are now encoded with film noise, bitstream-level noise detection and removal should be further investigated.

Acknowledgments

The authors thank the editor and the reviewers for their informative and constructive comments. The authors would also like to than National Archives of Korea (NAK) for providing valuable film content. This work is supported by the Seoul R&BD Program (PA100094), Korea.

References

[1] A. McAndrew, "Introduction to digital image processing with MATLAB," Thompson Course Technology, Melbourne.
[2] A.C. Kokaram, "Detection and removal of line scratches in degraded motion picture sequences," Signal Processing VIII, vol.1, pp.5–8, Sept. 1996.
[3] S.K. Kang, E.Y. Kim, K.C. Jung, and H.J. Kim, "Film line scratch detection using neural network," PCM 2004, pp.810–817, 2004.
[4] K.T. Kim and E.Y. Kim, "Film line scratch detection using texture and shape information," Pattern Recognit. Lett., vol.28, no.13, pp.1735–1746, 2007.

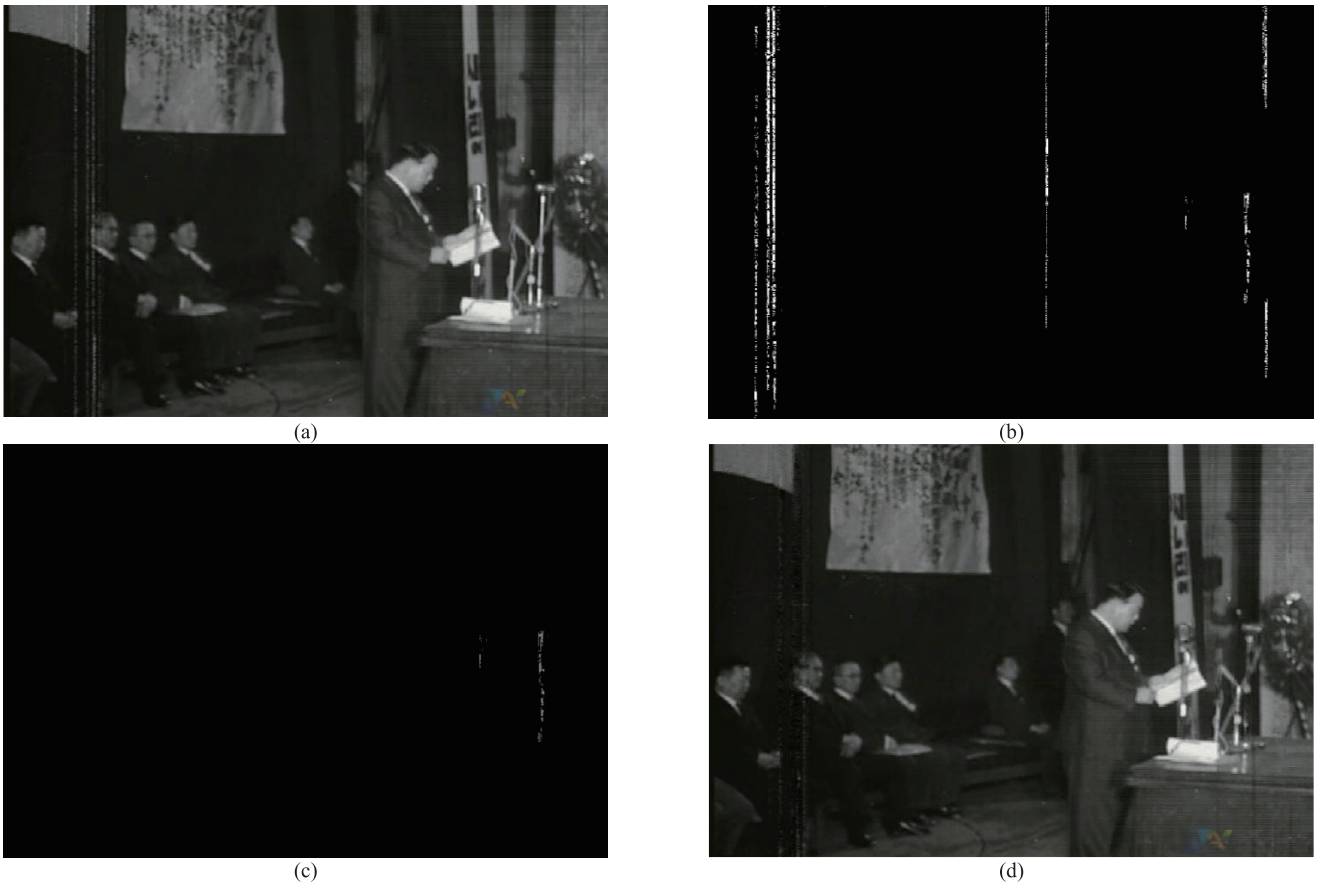


Fig. 12 Noise detection and removal results of the "Speech" sequence (11th frame). (a) Original video image, (b) detected noise regions, (c) missed and false alarmed noise regions (white color: missed, gray color: false alarmed), (d) reconstructed video image by the proposed method.

[5] S.W. Kim and K.H. Ko, "Efficient optimization of inpainting scheme and line scratch detection for old film restoration," PRICAI 2006, pp.623–631, 2006.
[6] Jin Xu, Jun Sun, X. Wang, and Li Chen, "Synthesis algorithm based on overcomplete wavelet expansion for line scratch restoration in old films," Opt. Eng., vol.47, July 2008.
[7] L. Joyeux, O. Buisson, B. Besserer, and S. Boukir, "Detection and removal of line scratches in motion picture films," IEEE Int. Conf. Computer Vision and Pattern Recognition, pp.548–553, June 1999.
[8] B. Besserer and C. Thire, "Detection and tracking scheme for line scratch removal in an image sequence," ECCV 2004, pp.264–275, 2004.
[9] M.S. Hamid, N.R. Harvey, and S. Marshall, "Genetic algorithm optimization of multidimensional grayscale soft morphological filters with applications in film archive restoration," IEEE Trans. Circuits Syst. Video Technol., vol.13, no.5, pp.406–416, 2003.
[10] R.C. Hardie and C.G. Boncelet, "LUM filters: A class of rank-order-based filters for smoothing and sharpening," IEEE Trans. Signal Process., vol.41, no.3, pp.1834–1838, March 1993.
[11] A. Nieminen, P. Heinonen, and Y. Neuvo, "A new class of detail-preserving filters for image processing," IEEE Trans. Pattern Anal. Mach. Intell., vol.PAMI-9, pp.74–90, Jan. 1987.
[12] H.G. Senel, R.A. Peters II, and B. Dawant, "Topological median filters," IEEE Trans. Image Process., vol.11, pp.89–104, Feb. 2002.
[13] A.C. Kokaram, Motion picture restoration, Springer Verlag, 1998.
[14] M.J. Nadenau and S.K. Mitra, "Blotch and scratch detection in image sequences based on rank ordered differences," Proc. 5th Int. Workshop Time-Varying Image Process. Moving Object Recognit., Sept. 1996.
[15] J. Biemond, P.M.B. van Roosmalen, and R.L. Lagendijk, "Improved blotch detection by postprocessing," Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'99) (6), Washington, 1999.
[16] A. Gangal, T. Kayikcioglu, and B. Dizdaroglu, "An improved motion-compensated restoration method for damage color motion picture film," Signal Process.: Image Commun., vol.19, pp.353–368, 2004.
[17] J.S. Lee and K. Hoppel, "Noise modeling and estimation of remotely-sensed images," Proc. IGARSS '89 and Canadian Symposium on Remote Sensing, pp.1005–1008, July 1989.
[18] A. Amer and E. Dubois, "Fast and reliable structure-oriented video noise estimation," IEEE Trans. Circuits Syst. Video Technol., vol.15, no.1, pp.113–118, Jan. 2005.
[19] D.H. Shin, R.H. Park, S. Yang, and J.H. Jung, "Block-based noise estimation using adaptive Gaussian filtering," IEEE Trans. Consum. Electron., vol.51, no.1, pp.218–226, Feb. 2005.
[20] S.I. Olsen, "Estimation of noise in images: An evaluation," CVGIP: Graph. Models and Image Process., vol.55, no.4, pp.319–323, July 1993.
[21] K. Rank, M. Lendl, and R. Unbehauen, "Estimation of image noise variance," IEE Proc. Vis. Image Signal Process., vol.146, no.2, pp.80–84, April 1999.
[22] N. Gupta, M.N. S. Swamy, and E.I. Plotkin, "Wavelet domain-based video noise reduction using temporal discrete cosine transform and hierarchically adapted thresholding," IET Image Process., vol.1, no.1, pp.2–12, March 2007.
[23] S. Yu, M.O. Ahmad, and M.N.S. Swamy, "Video denoising using

motion compensated 3-D wavelet transform with integrated recursive temporal filtering," IEEE Trans. Circuits Syst. Video Technol., vol.20, no.6, pp.780–791, June 2010.

Sinwook Lee received a B.S. in Computer Engineering from Dongguk University, Korea in 2005 and an M.S. in Electrics and Computer Engineering at Hanyang University, Korea in 2007. He is currently working toward a Ph.D. at Hanyang University. He has received three ISO/IEC Certificates of Appreciation for contributions to Rdevelopment. His current research interests include MPEG-4 AVC/H.264, HEVC, reconfigurable graphics coding, and noise processing.

Euee-seon Jang received a B.S. from Jeonbuk National University, Korea and a Ph. D. from SUNY at Buffalo, NY. He is an Associate Professor in the Dept. of CSE, Hanyang University, Seoul, Korea. His research interests include image/video coding, reconfigurable video coding, and computer graphics objects. He has authored more than 150 MPEG contribution papers, more than 30 journal or conference papers, has 35 pending or patented patents, and two book chapters. Dr. Jang has received three ISO/IEC Certificates of Appreciation for contributions to MPEG-4 development. He received a Presidential Award from the Korean Government for his contribution to MPEG standardization.