

PREFILTER CONTROL SCHEME FOR LOW BITRATE TV DISTRIBUTION

Ryoichi KAWADA Atsushi KOIKE Yasuyuki NAKAJIMA

KDDI R & D Laboratories
2-1-15 Ohara, Fujimino-shi, Saitama 356-0003, Japan

ABSTRACT

In IP-based TV distribution, coding degradation is sometimes evident in critical scenes because the bit rate for compression is rather low. Prefiltering is an effective countermeasure since it replaces the coding noise with the degradation more difficult to detect visually, though it has the drawback that excessive smoothing might occur. This paper proposes a scene-adaptive method to control a prefilter separate from the encoder. By calculating block-wise motion-compensated predictive error variances and correlation coefficients, it estimates the coding noise as well as the potential improvement by prefiltering each frame, realizing a control scheme which performs prefiltering only when effective.

1. INTRODUCTION

IP-based TV distribution requires high compression coding since the available bitrate is usually low. Annoying coding degradation especially in the form of blocky artifacts unique to MC-DCT schemes such as MPEG-2 and H.264 is sometimes visible, as TV contents include various kinds of pictures.

One countermeasure in such cases is prefiltering which is often used to give the perceptual visual quality priority over PSNR and applies a low-pass type filter to the input video signals to reduce entropy and coding degradation. However, prefiltering has the side effect of blurry pictures so adaptive control is essential.

Lin[1] used a Gaussian filter whose size is determined by recent quantization parameters. Segall and Karunaratne[2][3] further optimized filter control by positioning a filter just prior to quantization (after motion-compensated prediction). These schemes control the prefilter parameters along with quantization control, and so require that the prefilter be integrated with the encoder.

This paper, in order to realize high-quality TV distribution utilizing simple inexpensive encoders, studies a prefilter separate from the encoder. A prefilter separate from the encoder cannot obtain feedback information such as quantization parameters relating to the input pictures. Therefore its adaptive control must be based only on the input video signals.

This study therefore is aimed at finding an appropriate method to control a prefilter that is separate from the encoder

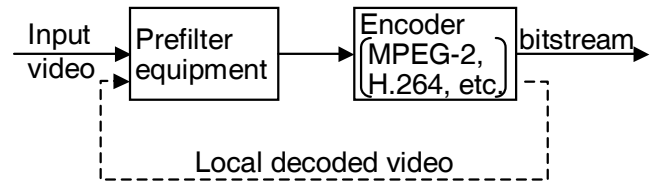


Fig. 1. Concept of prefilter separate from the encoder.

in order to improve the perceptual visual quality of the decoded pictures.

Another application of prefilters is improving the coding efficiency by reducing noise added to input pictures[4]. However that topic is beyond the scope of this study.

2. OBJECTIVES OF STUDY

To achieve high perceptual quality TV distribution utilizing simple inexpensive encoders, the authors consider use of prefilter equipment separate from the encoder equipment, as described in 1 (Fig.1).

Locally decoded pictures can be input as reference information as shown in Fig.1. Codecs such as MPEG-2 and H.264 however usually have a coding and decoding delay of several dozen frames. Therefore unlike prefilters integrated with the encoder where feedback is fast, the feedback control may not work very well.

It is therefore important that control is solely based on analysis of the input pictures.

Principles of filter control in this study are as follows:

- Prefilters are not needed when the available bitrate is sufficiently large. It might in fact be better to avoid using prefilters since they cause a blurring effect. (Using a prefilter usually does not improve the total PSNR (filter degradation plus coding degradation)).
- Prefilters may prove effective when the available bitrate is insufficient and there is large coding noise such as blocky degradation. Even in such cases, because of the blurring it causes, the prefilter should be applied only when a large reduction in coding noise can be obtained.

3. ANALYSIS BY MODELING

In this section, we use theoretical modeling to quantitatively analyze the coding degradation and find how it is improved by prefiltering.

Here we consider a linear FIR filter consisting of m coefficients $a_i (i = 1, \dots, m)$. \hat{X} , the output pixel of the filter, is $\hat{X} = \sum_{i=1}^m a_i X_i$. The motion-compensated predictive error signal, x , is

$$x = \hat{X} - (\hat{X}^{ref} + x_q^{ref}) = \sum a_i (X_i - X_i^{ref}) - x_q^{ref}$$

where x_q is the coding noise and the reference picture (motion-compensated predictive picture) is represented by $(\cdot)^{ref}$. When x_q^{ref} is small enough to be ignored, the predictive error variance σ_x^2 is

$$\begin{aligned} \sigma_x^2 &= Var(x) = \sum_{i=1}^m a_i^2 Var(X_i - X_i^{ref}) \\ &+ 2 \sum_{1 \leq i < j \leq m} a_i a_j Cov(X_i - X_i^{ref}, X_j - X_j^{ref}) \end{aligned} \quad (1)$$

where $Var(\cdot)$ and $Cov(\cdot, \cdot)$ respectively denote the variance and covariance. $Var(X_i - X_i^{ref})$ is the variance of the motion-compensated predictive error between original pictures without noise and is denoted by $\sigma_{x_0}^2$ here. Then,

$$Cov(X_i - X_i^{ref}, X_j - X_j^{ref}) = \rho(i, j) \sigma_{x_0}^2$$

where $\rho(i, j)$ represents the inter-pixel correlation coefficient of the motion-compensated predictive error.

Here we consider a horizontal one-dimensional filter for the sake of simplicity. When the correlation coefficient between horizontally neighboring pixels is denoted by ρ , the general characteristics of the image signals yield $\rho(i, j) = \rho^{|i-j|}$. Since the inter-pixel correlation coefficients of the predictive error signals are generally small, we consider only the first power of ρ . We then obtain the following from Eq.(1):

$$\sigma_x^2 = \left(\sum_{i=1}^m a_i^2 + 2\rho \sum_{i=1}^{m-1} a_i a_{i+1} \right) \sigma_{x_0}^2 \quad (2)$$

On the other hand, given the bitrate and the block-wise predictive error variance, the coding degradation (variance of coding noise) for block-based MC-DCT coding is represented by Eq.(7) in the Appendix.

Let us now apply Eq.(2) to each block. From Eq.(7), the coding degradation without prefiltering, n_{nofil} , is

$$n_{nofil} = \frac{\sqrt{\prod_{k=1}^M \sigma_{x_0 k}^2}}{2R/M} \quad (3)$$

Similarly, from Eq.(2), the coding degradation with prefiltering, n_{fil} , is

$$n_{fil} = \frac{\sqrt{\prod_{k=1}^M \left(\sum_{i=1}^m a_i^2 + 2\rho \sum_{i=1}^{m-1} a_i a_{i+1} \right) \sigma_{x_0 k}^2}}{2R/M} \quad (4)$$

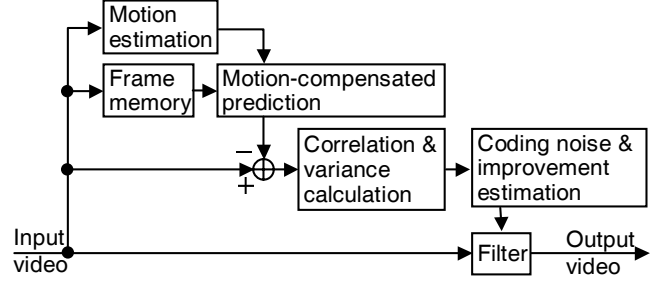


Fig. 2. Proposed prefilter control scheme.

Eq.(3) and Eq.(4) allow deriving the improvement of coding degradation by prefiltering as follows.

$$\frac{n_{fil}}{n_{nofil}} = \sqrt[M]{ \prod_{k=1}^M \left(\sum_{i=1}^m a_i^2 + 2\rho \sum_{i=1}^{m-1} a_i a_{i+1} \right) } \quad (5)$$

4. PROPOSED SCHEME

The authors propose a prefilter control scheme using the following steps(Fig.2).

1. Perform processing frame by frame.
2. Perform simple motion estimation. One example is the iterative gradient method in [5] which can estimate optical flows with simple hardware.
3. Generate a motion-compensated predictive picture using the motion information.
4. Generate a difference picture between the current frame and the motion-compensated predictive frame, and calculate the block-wise variance and correlation coefficient between neighboring pixels. A typical block size is 16x16.
5. Estimate coding degradation without prefiltering from the block variance values and information about the encoder in use, referring to Eq.(3).
6. Estimate the improvement in coding degradation by prefiltering from the block correlation coefficients and Eq.(5).
7. Judge the prefilter as effective when both the estimated coding degradation and its potential improvement are large, and apply prefiltering to the input picture.

In prefilter control it is better to adjust the intensity of the filter continuously rather than by simple on-off control. Continuous adjustment yields better perceptual visual quality since on-off control might cause flickering in the picture.

5. COMPUTER SIMULATION EXPERIMENTS

This section discusses the accuracy when estimating coding degradation and the improvement obtained from prefiltering which is the core of our proposed scheme.

A one-dimensional horizontal 1/3 low-pass prefilter was used (This bandwidth is selected as a ratio of horizontal pixels of SDTV and HDTV, 720:1920). Tbl.1 shows the coefficients.

Table 1. Prefilter coefficients used in experiments (1/3 horizontal low-pass filter)

a_1	a_2	a_3	a_4	a_5	a_6
-0.0046	-0.0163	0.0000	0.0994	0.2546	0.3338
a_7	a_8	a_9	a_{10}	a_{11}	
0.2546	0.0994	0.0000	-0.0163	-0.0046	

The motion estimation method and block size shown in 4 were utilized.

Various kinds of test pictures were selected from an ITU-R library[6]: These included Balls of wool, Birches, Cheerleaders, Flower garden, Horse riding, Kiel Harbour-4, Mobile and calendar, Popple, Summer flowers, and Tempête. There were a total of 10 sequences, each having about 900 frames.

An H.264 encoder, JM9.8 was used with a picture structure of (IBPBPB....) and spatial direct mode. Three bitrates, 0.5Mbps, 1.0Mbps, and 2.0Mbps were examined. The results hereafter use about 500 frames of luminance signals from the 400th frame onward where buffer control is already stable.

Fig.3(a) shows the relationship between the frame PSNR and its estimated value according to Eq.(3). The bitrate is 1Mbps. The approximating line is $y = x - 0.65$ with an average estimation error of 2.6 dB. Similar results were obtained for 0.5Mbps and 2Mbps. At a bitrate of 0.5Mbps, the approximating line is $y = x - 2.78$ with an average estimation error of 2.5 dB. At a bitrate of 2Mbps, the approximating line is $y = x + 1.44$ with an average estimation error of 2.7 dB.

These results show that the frame PSNR by H.264 coding can be estimated with an estimation error of about ± 3 dB, although the approximating line depends on the bitrate. This confirms that our proposed control scheme, which applies a prefilter only when coding degradation is large, can be realized.

Next, Fig.3(b) shows the relationship between the improvement in frame PSNR and its estimated value according to Eq.(5). The bitrate is 1Mbps. The approximating line is $y = 0.98x$ with an average estimation error of 1.0 dB. Similar results were obtained at 0.5Mbps and 2Mbps. At a bitrate of 0.5Mbps, the approximating line is $y = 0.85x$ with an average estimation error of 0.9 dB. At a bitrate of 2Mbps, the approximating line is $y = 1.13x$ with an average estimation error of 1.0 dB.

These results show that an improvement in the frame PSNR can be estimated with an error of about ± 1 dB, although the approximating line depends on the bitrate. This

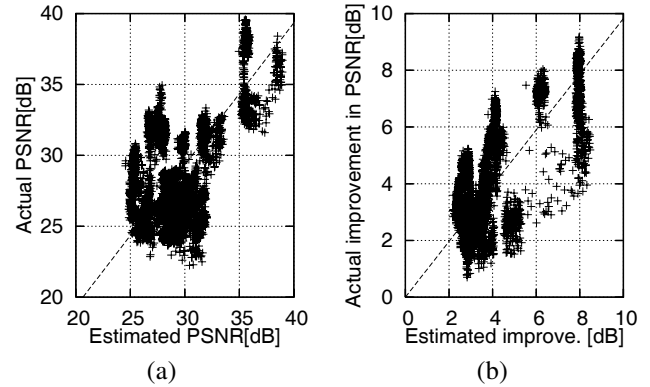


Fig. 3. (a) Actual and Estimated PSNR [dB] (b) Actual and Estimated improvement in PSNR by prefiltering. Bitrate: 1Mbps

Table 2. Coding PSNR without prefilter and its improvement Δ by prefiltering [dB]. Bitrate: 1Mbps

	Horse	Summer	Balls	Popple	Flower
PSNR	35.2	31.9	31.1	28.1	27.6
Δ	2.7	3.5	6.9	3.3	2.9
	Mobile	Tempete	Birches	Kiel4	Cheer
PSNR	27.3	26.7	26.3	26.0	25.3
Δ	5.5	2.5	2.7	3.3	2.0

confirms that our proposed control scheme, which applies a prefilter only when the potential improvement in PSNR is large, can be realized.

The reason why the slope of the approximation line becomes smaller as the bitrate becomes smaller, is that Eq.(7) on which Eq.(5) is based, assumes a rather high bitrate as described in the Appendix. When the bitrate is low, all DCT coefficients in some blocks are discarded, and such blocks do not contribute to the multiplication in Eq.(5). In other words, the improvement in PSNR becomes smaller.

Tbl.2 shows the average coding PSNR without a prefilter and its improvement by prefiltering. As can be seen in the table, significant improvement (over 5dB) can be obtained for Balls and Mobile sequences, for example, whereas about 2-3 dB gain can be obtained in other sequences. Cheerleaders, which involves complicated motion, yields the smallest improvement, since the motion-compensated predictive error picture has rather large inter-pixel correlation. These results confirm that a low PSNR without a prefilter does not necessarily mean that prefiltering will prove effective for the picture, and so that an appropriate control is important.

In addition, we subjectively evaluated the decoded pictures using the prefilter of Tbl.1. As a result, the prefilter proved effective for Mobile, Popple, and Balls especially when these were encoded at 0.5Mbps and 1Mbps. Without a prefilter, blocky degradation was clearly observed around the paintings on the wall in "Mobile", the columns of the cage in

“Popple”, and the woolen yarn in “Balls”. On the other hand, the prefilter showed little significant effect for Cheerleaders.

The approximation lines in Fig.3 and especially that for the estimated PSNR (coding degradation) depend on the encoder being used. So from a practical viewpoint, use of customized estimation tables specifically made for the encoder from test pictures is appropriate. In the case of this experiment, such a control would be possible that the prefilter is applied when the estimated PSNR is below 30dB and the estimated improvement in PSNR is over 3dB.

The above demonstrates the effectiveness of our proposed control scheme.

6. CONCLUSIONS

This paper has proposed a method to control a prefilter separate from the encoder in order to improve perceptual visual quality of low bitrate TV. Our proposed method analyzes the input pictures to estimate the coding degradation as well as potential improvement obtained by prefiltering.

After simple motion compensation, the predictive error variance and correlation coefficient are derived for each block. Based on these, the coding degradation and extent of improvement from prefiltering are estimated for each frame. By considering the bitrate and the encoder being used, the frame PSNR of the compression coding and its improvement from prefiltering can be estimated with the respective average estimation error of ± 3 dB and ± 1 dB. Based on this information a decision can be made to skip prefiltering in order to avoid blurring side effects or to apply prefiltering to reduce visually annoying coding noise.

Further topics for study include improving the estimation accuracy, optimizing the process by subjective evaluation, and refining the filter control according to the picture texture.

The authors hope that this study helps to reduce costs and contributes to the spread of IP-based TV.

7. REFERENCES

- [1] Liang-Jin Lin and Antonio Ortega, “Perceptually based video rate control using pre-filtering and predicted rate-distortion characteristics,” in *Proc. IEEE International Conf. on Image Processing*, Washington, DC, USA, Oct. 1997, vol. 2, pp. 57–60.
- [2] Passant V. Karunaratne, C. Andrew Segall, and Aggelos K. Katsaggelos, “A rate-distortion optimal video pre-processing algorithm,” in *Proc. IEEE International Conf. on Image Processing*, Thessaloniki, Greece, Oct. 2001, vol. 1, pp. 481–484.
- [3] C. Andrew Segall, Passant Karunaratne, and Aggelos K. Katsaggelos, “Pre-processing of compressed digital

video,” in *Proc. SPIE Visual Communications and Image Processing*, San Jose, California, USA, Jan. 2001, pp. 163–174.

- [4] Li Yan, “Noise reduction for MPEG type of codec,” in *Proc. IEEE International Conf. on Acoustics, Speech, and Signal Processing*, Apr. 1994, vol. 5, pp. V/429–V/432.
- [5] Ryoichi Kawada, Takahiro Hamada, and Shuichi Matsumoto, “Improvement in motion-compensated TV standards conversion,” *Journal of ITE (Japanese Edition)*, vol. 51, no. 9, pp. 1577–1586, Sept. 1997.
- [6] ITU-R Recommendation BT.802-1, *Test pictures and sequences for subjective assessments of digital codecs conveying signals produced according to Recommendation ITU-R BT.601*, July 1994.
- [7] R. J. Clarke, *Transform coding of images*, Academic Press, 1985.

A. OPTIMIZED ALLOCATION OF INFORMATION BITS TO EACH BLOCK

Let us assume that the motion-compensated predictive error picture to be encoded consists of M Macroblocks (MB’s). Let σ_k^2 denote the predictive error variance of the k -th MB ($k = 1, \dots, M$).

Let n_k denote the coding noise for each MB.

The objective is to minimize the total coding noise, i.e., $\sum_{k=1}^M n_k \rightarrow \min$. The constraint here is that the amount of total generated bits is constant. As the bits generated from the k -th MB can be assumed as $0.5 \log_2(\sigma_k^2/n_k)$ bpp[7],

$$\sum_{k=1}^M \log_2 \frac{\sigma_k^2}{n_k} = R \quad (\text{constant}) \quad (6)$$

where $\sigma_k^2 \geq n_k$ ($\forall k$). The following is then obtained by solving this by means of Lagrange multipliers.

$$n_k = \frac{\lambda}{\ln 2} \quad \text{or} \quad n_k = \sigma_k^2 \quad (\lambda : \text{constant})$$

The optimized coding noise for each block is therefore the same for all blocks except for blocks where all DCT coefficients are discarded (the coding noise equals the original signal variance).

From Eq.(6),

$$\sum_{k=1}^M \max(0, \log_2 \frac{\sigma_k^2 \ln 2}{\lambda}) = R$$

determines the λ and n_k . When R is large enough to yield $\log_2 \frac{\sigma_k^2 \ln 2}{\lambda} > 0$ for all blocks, the following is obtained.

$$n_k = \frac{\sqrt[M]{\prod_{k=1}^M \sigma_k^2}}{2^{R/M}} \quad (\forall k) \quad (7)$$