A LOSSLESS RE-ENCODING SCHEME FOR MPEG-1 VIDEO

Ichiro Matsuda, Kei Wakabayashi, Yu Ikeda and Susumu Itoh

Department of Electrical Engineering, Faculty of Science and Technology,
Science University of Tokyo

2641 Yamazaki, Noda-shi, Chiba 278-8510, JAPAN
phone: +81 4 7124 1501 (ext.3740), fax: +81 4 7124 9367
email: matsuda@ee.noda.tus.ac.jp

ABSTRACT

This paper proposes a lossless re-encoding scheme for video data which are already encoded by the MPEG-1 video coding standard. In order to avoid any additional loss of quality, quantized DCT coefficients and motion vectors extracted from the original MPEG-1 bitstream are losslessly re-encoded using an adaptive arithmetic coder. Probability models used in the arithmetic coder are iteratively optimized so that the number of coding bits required for the whole group of pictures (GOP) can be a minimum. Furthermore, H.264-like block-adaptive intra prediction is introduced into not only I pictures but also P and B pictures to exploit inter-block correlations of the DCT coefficients. Simulation results indicate that the proposed scheme can reduce the coding rate of monochrome MPEG-1 video by 12–22 %.

1. INTRODUCTION

MPEG-1 [1] was an early video coding standard published in 1993 and formed the basis of the current video compression technology. It was mainly designed for recording video signals in CD media at bit rates of about 1.5 Mbps, but its range of application has been widely expanded with the spread of multimedia environment. Consequently, a huge amount of video contents have been stored using the MPEG-1 standard. Recently, video coding technology makes remarkable progress and the latest standard called H.264/AVC [2] provides an overwhelming advantage in coding efficiency. Unfortunately, existing video contents already encoded by the old standard cannot enjoy such a state-of-the-art video coding technology. Of course, transcoding of video data from MPEG-1 to H.264/AVC is often performed to adapt the video format for the recent video clients. In this case, however, additional loss of image quality is unavoidable due to nature of lossy coding schemes. As a result, the advantage of the H.264/AVC standard is considerably spoiled.

A similar situation is also seen in the area of still image coding where the JPEG standard is most commonly used. To resolve the situation, lossless re-encoding schemes which can further compress existing JPEG images without any loss of quality have been studied in recent years [3–5]. These schemes extract quantized DCT coefficients from the original JPEG image and re-encode them using more efficient entropy coding techniques based on arithmetic coding. Moreover, inter-block correlations of the DCT coefficients which were not used in the JPEG standard with the exception of DC coefficients are exploited to improve the coding efficiency. According to our study [6], this approach can reduce coding rates of monochrome JPEG images by 18–28 % and would be a promising option for new image coding technology.

In this paper, we propose a lossless re-encoding scheme for video data which are already encoded by the MPEG-1 video coding standard. Basic algorithm of the proposed scheme is similar to our lossless re-encoding scheme developed for JPEG images [6] because both MPEG-1 and JPEG standards share many common components such as DCT, quantization and Huffman encoding. A most different point is that the MPEG-1 standard employs motion-compensated prediction for predicted (P) and bi-directional predicted (B) pictures to exploit temporal correlations in video signals. Therefore, not only DCT coefficients but also motion vectors used for P and B pictures must be losslessly re-encoded in the proposed scheme.

2. PDF MODELING OF QUANTIZED DCT COEFFICIENTS

Figure 1 shows an overview of the proposed re-encoding scheme. The original MPEG-1 bitstream compressed by Huffman encoding is decoded halfway to extract quantized DCT coefficients and motion vectors as well as other header information. To enable efficient re-encoding of the quantized DCT coefficients, probability density functions (PDF) of the DCT coefficients are modeled by the following sixteen kinds

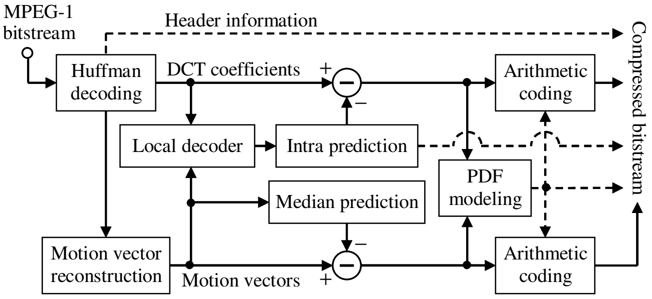


Figure 1: Block diagram of the proposed scheme.

of the generalized Gaussian functions [7]:

$$P_n(t) = \frac{c_n \cdot \eta(c_n, \sigma_n)}{2\Gamma(1/c_n)} \cdot \exp\left\{ -\left| \eta(c_n, \sigma_n) \cdot t \right|^{c_n} \right\},$$

$$\eta(c_n, \sigma_n) = \frac{1}{\sigma_n}\sqrt{\frac{\Gamma(3/c_n)}{\Gamma(1/c_n)}} \qquad (n = 1, 2, \ldots, 16), \quad (1)$$

where $\Gamma(\cdot)$ is the gamma function, σ_n is standard deviation of the DCT coefficients and c_n is a shape parameter which controls sharpness of the function $P_n(t)$. In this paper, values of σ_n are chosen in advance to cover a wide range of activity of the DCT coefficients. On the other hand, values of the shape parameter c_n can be changed according to the actual PDFs. Moreover, all the blocks composed of 8×8 pels are classified into one of twenty classes ($m = 1, 2, \ldots, 20$). Each class has a look-up table $n_m(i, j) \in \{1, 2, \ldots, 16\}$ which describes assignment of the variances $\{\sigma_n^2\}$ to the two dimensional frequencies $(i, j = 0, 1, \ldots, 7)$ of the DCT coefficients. We call the table $n_m(i, j)$ a variance map here after. It should be noted that the above function $P_n(t)$ corresponds to a PDF model of the DCT coefficients before the quantization. In practice, a probability of each quantization level of the DCT coefficient $y = y(i, j) \in \mathbb{Z}$ is given by:

$$\Pr\big(y \,|\, n, q\big) = \int_{T(y-1,q)}^{T(y,q)} P_n(t)\,\mathrm{d}t, \qquad (2)$$

where $n = n_m(i, j)$ is an element of the variance map. q is a quantization step-size determined by a quantization parameter (QP) and two kinds of quantization matrices $\{q_k(i, j)\}$ specified for intra coded ($k = 0$) and non-intra coded ($k = 1$) macroblocks in the MPEG-1 standard. In addition, $T(y, q)$ represents a threshold between adjacent quantization levels y and $y + 1$ as shown in Figure 2. The MPEG-1 standard employs different types of quantization for intra coded and non-intra coded macroblocks. Therefore, definitions of $T(y, q)$ is described as follows,

- For intra coded macroblocks:

$$T(y, q) = (y + 0.5) \cdot q, \qquad (3)$$

- For non-intra coded macroblocks:

$$T(y, q) = \begin{cases} (y + 1) \cdot q & (y \geq 0) \\ y \cdot q & (y < 0) \end{cases}. \qquad (4)$$

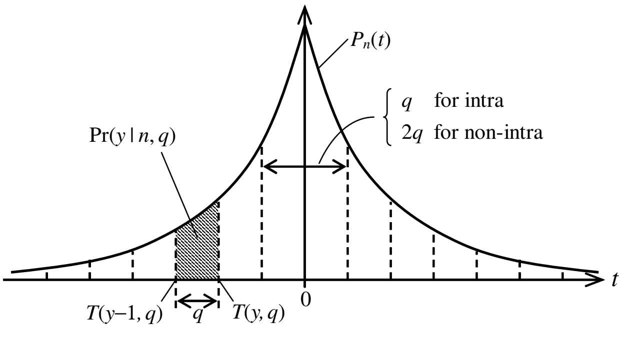


Figure 2: PDF model for quantized DCT coefficients.

In this way, probabilities for all possible quantization levels of the DCT coefficients can be adaptively calculated according to values of the quantization parameter QP, quantization matrix $q_k(i, j)$ and variance map $n_m(i, j)$. These probabilities are used for entropy coding of the actual values of the DCT coefficients $\{y(i, j)\}$. In our implementation, the rangecoder [8] which is known as a fast multisymbol arithmetic coder is employed for the entropy coding process.

3. CODING OF MOTION VECTORS

In addition to the DCT coefficients, motion vectors are also needed for lossless reconstruction of P and B pictures. In the MPEG-1 standard, the motion vectors are detected with half-pel accuracy and differences from the left macroblocks are encoded using a fixed Huffman table. Instead of such a simple DPCM approach, we employ a median prediction technique which has been popular in the recent video coding standard [2]. For this purpose, the differential motion vectors extracted from the original MPEG-1 bitstream are once reconstructed for every macroblock, and again predicted by a median value of three motion vector components taken from the left, above and above-right macroblocks. Finally, the obtained prediction residuals, or differential motion vectors, are encoded using the range coder. In this paper, a PDF model for the differential motion vector $(\Delta v_x, \Delta v_y)$ is defined by the following function:

$$P_v(\Delta v_x, \Delta v_y) = \alpha \cdot \exp\left\{ -\left| \eta(c_v, \sigma_v) \cdot \sqrt{\Delta v_x^2 + \Delta v_y^2} \right|^{c_v} \right\}, \quad (5)$$

where α is a normalization factor making the sum of probabilities equal to one. σ_v and c_v are parameters which control a property of the PDF model. Values of these parameters are optimized for each P or B picture in the way mentioned later. The function $P_v(\Delta v_x, \Delta v_y)$ is rotationally symmetrical and its shape is given by the generalized Gaussian function on the radius coordinate as shown in Figure 3. By using this joint PDF model of two variables together with the rangecoder, the differential motion vector components can be directly encoded in a two-dimensional form [9].

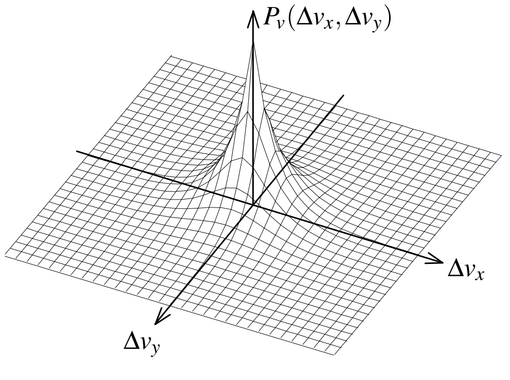


Figure 3: Joint PDF model for differential motion vectors.

4. BLOCK-ADAPTIVE INTRA PREDICTION

In the MPEG-1 standard, coding efficiency is generally poor in intra-coded (I) pictures because these pictures cannot use the motion-compensated prediction. To cope with this problem, the latest video coding standard H.264/AVC [2] employs a block-adaptive intra prediction technique for intra-coded macroblocks. In this paper, we introduce a similar technique to exploit spatial correlations between adjacent blocks. In fact, eight directional prediction modes and a DC prediction mode used in our intra prediction technique are almost same as the nine modes of 4×4 intra prediction specified in the H.264/AVC [2], but block size for the prediction is extended to 8×8 pels in conformity with the size of DCT used in the MPEG-1 standard. Since the intra prediction must be performed in spatial domain, the proposed scheme has a local decoder to reconstruct image values in blocks of causal neighbors. Figure 4 illustrates positions of pels used for the intra prediction as shaded boxes.

In the lossless re-encoding scheme for JPEG images [6], predicted values $\{\hat{x}(i,j)\}$ calculated by applying one of the above nine prediction modes to the current block are transformed into DCT domain. The obtained coefficients $\{\hat{y}(i,j)\}$ are then quantized and subtracted from the original DCT coefficients. By calculating the prediction residuals in DCT domain like this, we can ensure lossless reconstruction of the original quantized DCT coefficients. It makes the reconstructed image also lossless if the same reconstruction process as the standard decoder is applied.

In the proposed scheme, such an approach is taken not only in I pictures but also in P and B pictures. For intra coded macroblocks, the prediction residuals in DCT domain is calculated by:

$$\Delta y(i,j) = y(i,j) - Q_0\left[\hat{y}(i,j)\right], \tag{6}$$

where $Q_0[y]$ means a linear quantizer whose thresholds are given by Eq.(3). On the other hand, since the motion-compensated prediction is already carried out in non-intra coded macroblocks, DCT coefficients $\{\tilde{y}(i,j)\}$ of the motion-compensated predicted values must be considered in calculation of the residuals:

$$\Delta y(i,j) = y(i,j) - Q_1\left[\hat{y}(i,j) - \tilde{y}(i,j)\right], \tag{7}$$

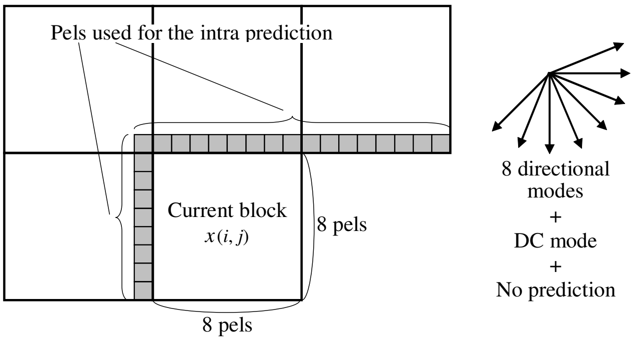


Figure 4: Block-adaptive intra prediction.

where $Q_1[y]$ is another type of the quantizer used for non-intra coded macroblocks and its thresholds are defined by Eq.(4). By encoding values of $\{\Delta y(i,j)\}$ instead of the original DCT coefficients $\{y(i,j)\}$, we can make use of inter-block correlations even in the non-intra coded macroblocks designated in the original MPEG-1 bitstream. If the motion-compensated prediction of the MPEG-1 standard is found to be more efficient than any of the nine intra prediction modes, another mode of 'No prediction' ($d = 0$), that means $\Delta y(i,j) = y(i,j)$, can be selected. Consequently, the prediction mode $d \in \{0,1,2,\dots,9\}$ must be designated at each block of 8×8 pels in the proposed scheme.

5. OPTIMIZATION OF CODING PARAMETERS

In the proposed scheme, parameters listed below must be encoded together with header information of the original MPEG-1 bitstream.

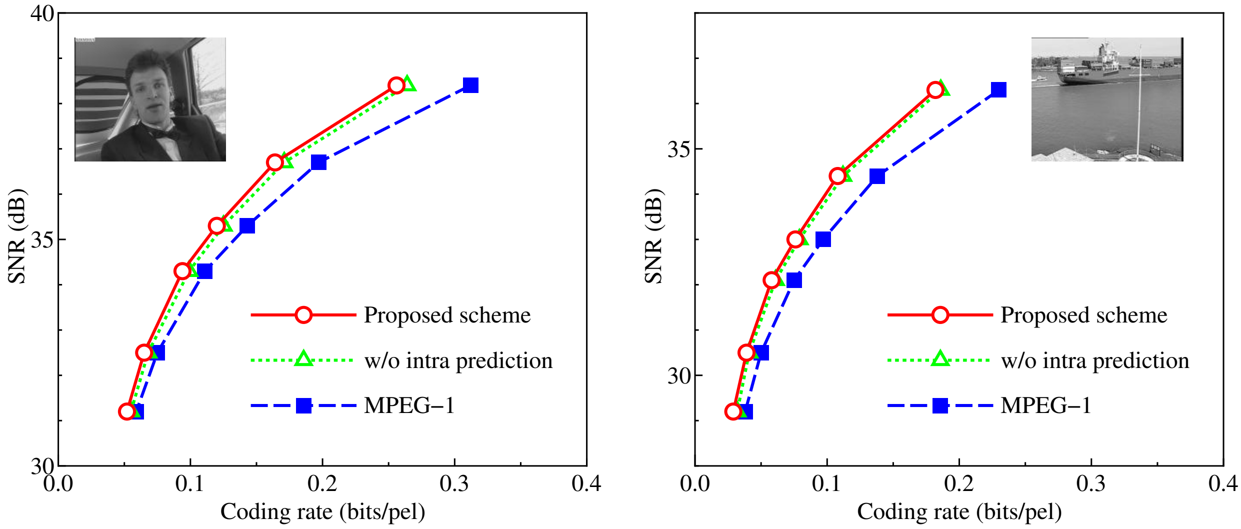
- Prediction mode $d(k,l) \in \{0,1,2,\dots,9\}$ for each block.
- Class label $m(k,l) \in \{1,2,\dots,20\}$ for each block.
- Variance map $n_m(i,j) \in \{1,2,\dots,16\}$ for each class.
- Parameter c_n for each PDF model $P_n(t)$.
- Parameters σ_v and c_v of the 2D PDF model $P_v(\Delta v_x, \Delta v_y)$ for each frame.

These parameters are iteratively optimized prior to the actual re-encoding process so that the following cost function calculated in the whole GOP can be a minimum.

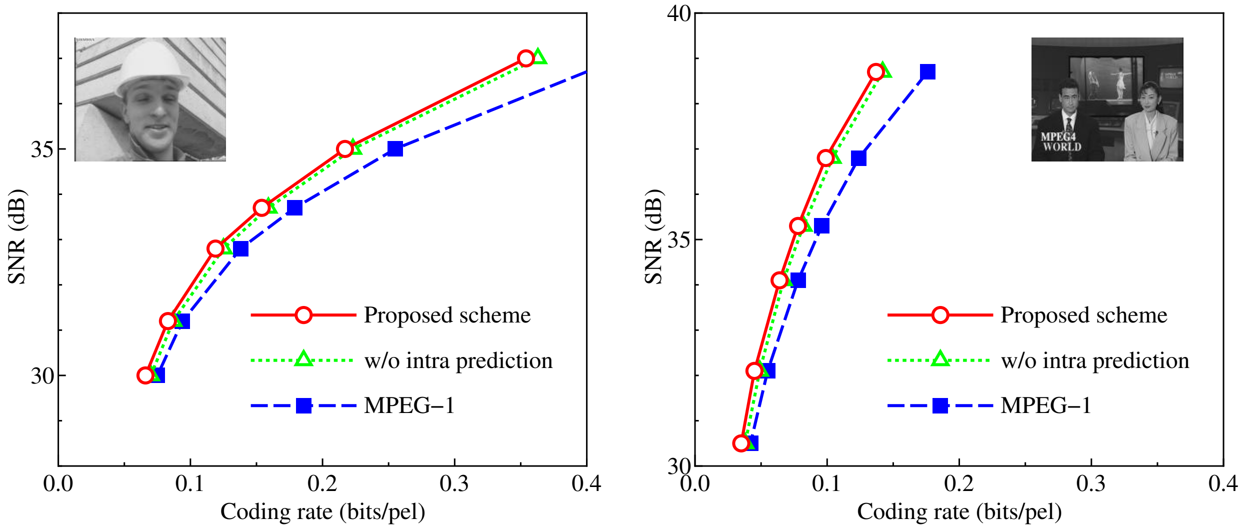
$$\begin{aligned} J = &-\sum \log_2 \Pr(\Delta y(i,j)\,|\,n,q) \\ &-\sum \log_2 \Pr(\Delta v_x, \Delta v_y) \;+\; B_{\text{side}} \end{aligned} \tag{8}$$

The first and second terms in the right hand side of this equation represent the number of coding bits required for entropy coding of the quantized DCT coefficients and the differential motion vectors, respectively. The third term (B_{side}) indicates the amount of side information on the above parameters. Concrete procedures of the optimization for each GOP is as follows:

(1) Initial values of the prediction mode d and the class label m are assigned for each block of 8×8 pels.
(2) Optimum combination of parameters σ_v and c_v is determined for each P or B picture.
(3) For each class (m), the variance map $n_m(i,j)$ is optimized by selecting one of sixteen PDF models $P_n(t)$ ($n = 1,2,\dots,16$) at every frequency (i,j) of DCT coefficients.
(4) Optimum value of the shape parameter c_n is chosen for each PDF model.
(5) Re-classification of the blocks is performed by selecting the optimum variance map $n_m(i,j)$ at every block.
(6) Optimum prediction mode (d) is re-selected for each block.
(7) Procedures (3)–(6) are repeated until all of the parameters converge



(a) Carphone (b) Container



(c) Foreman (d) News

Figure 5: Coding performance.

6. EXPERIMENTAL RESULTS

In the experiments, CIF-sized monochrome video sequences (352×288 pels, 64 frames long) are once encoded into MPEG-1 bitstreams with a GOP structure of IBBPBBP$\cdots$ ($N = 16, M = 3$). We utilize the Berkley MPEG Tools [10] for this process. Then the obtained bitstreams are further compressed by the proposed re-encoding scheme. Figure 5 compares rate-SNR curves of the original MPEG-1 using a fixed quantization parameter of $QP \in \{4, 6, 8, 10, 15, 21\}$ and those of our re-encoding schemes. In the figure, 'w/o intra prediction' means a variant of the proposed scheme where the block-adaptive intra prediction technique is omitted. It is shown that the block-adaptive prediction technique described in section 4 provides 2–6 % coding gains for the proposed scheme. As a result, bit-rate savings of 12–22 % over the original MPEG-1 bitstreams are attained for the tested sequences.

Table 1: Detailed coding results for 'Carphone' (bits/frame).

Item	Proposed scheme	MPEG-1
DCT coefficients	20040.9	28119.7
Motion vectors	1690.8	1843.6
Variance maps	222.5	—
Class labels	1882.4	—
Prediction modes	343.3	—
Others	1766.6	1644.2
Total	*25946.4*	*31607.5*

Table 1 reports breakdowns of the number of coding bits for the 'Carphone' sequence when the quantization parameter is set to $QP = 4$. We can see that the number of coding bits required for the DCT coefficients is remarkably reduced by the proposed scheme.

Table 2: Comparison with H.264/AVC-based tandem coding ('Carphone' at 1.5 Mbps).

	Coding rate (kbps)	SNR (dB)
MPEG-1 (original)	1500.39	39.2
Proposed scheme (lossless)	**1191.51**	**39.2**
H.264 High 4:4:4 (lossless)	7662.27	**39.2**
H.264 Main (lossy)	1501.01	38.8

Additionally, the proposed scheme is tested for MPEG-1 bitstream encoded at a constant bit rate of 1.5 Mbps. Table 2 lists the re-encoding results for the 'Carphone' sequence. In this table, 'H.264 High 4:4:4' and 'H.264 Main' indicate re-encoding results obtained by lossless mode (High 4:4:4 profile) and lossy mode (Main profile) of the H.264/AVC standard, respectively. In both cases, full decoding of MPEG-1 bitstream is needed before the respective re-encoding processes. This kind of approach is often referred to as tandem coding. Although it is possible to realize lossless re-encoding using High 4:4:4 profile of the H.264/AVC standard, such a tandem coding approach considerably increases the coding rate as shown in the table. On the other hand, re-encoding at the same bit rate (1.5 Mbps) by Main profile of the H.264/AVC standard slightly decreases SNR because it is lossy coding.

7. CONCLUSIONS

In this paper, we have proposed a lossless re-encoding scheme for MPEG-1 video. The scheme extracts quantized DCT coefficients and motion vectors from the original MPEG-1 bitstream and re-encodes them using the modern entropy coding technique based on arithmetic coding. Moreover the block-adaptive intra prediction technique inspired by the H.264/AVC standard is introduced to exploit spatial correlations between adjacent blocks. The effectiveness of the proposed scheme is confirmed through the experiments using monochrome video sequences.

The basic concept of the proposed scheme can be applied to other video coding standards such as H.261, MPEG-2, 4 and so on. It could be useful technology because video contents stored in the conventional video formats are still widely available.

REFERENCES

[1] ISO/IEC 11172-2, "Information technology — Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1,5 Mbit/s – Part 2: Video," 1993.

[2] ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC, "Advanced Video Coding for Generic Audiovisual Services," 2003.

[3] I. Bauermann and E. Steinbacj, "Further Lossless Compression of JPEG Images," *Proc. of Picture Coding Symposium (PCS 2004)*, San Francisco, USA, Dec. 15–17, 2004.

[4] N. Ponomarenko, K. Egiazarian, V. Lukin and J. Astola, "Additional Lossless Compression of JPEG Images," *Proc. of the 4th Intl. Symposium on Image and Signal Processing and Analysis (ISPA 2005)*, Zagreb, Croatia, pp.117–120, Sep. 15–17, 2005.

[5] M. Stirner and G. Seelmann, "Improved Redundancy Reduction for JPEG Files," *Proc. of Picture Coding Symposium (PCS 2007)*, Lisbon, Portugal, Nov. 7–9, 2007.

[6] I. Matsuda, Y. Nomoto, K. Wakabayashi and S. Itoh, "Lossless Re-encoding of JPEG Images Using Block-Adaptive Intra Prediction," *Proc. of 16th European Signal Processing Conference (EUSIPCO 2008)*, Lausanne, Switzerland, L3-6, Aug. 25–29, 2008.

[7] I. Matsuda, N. Ozaki, Y. Umezu and S. Itoh, "Lossless Coding Using Variable Block-Size Adaptive Prediction Optimized for Each Image," *Proc. of 13th European Signal Processing Conference (EUSIPCO 2005)*, Antalya, Turkey, WedAmPO3, Sep. 4–8, 2005.

[8] M. Schindler, "A Fast Renormalization for Arithmetic Coding," *Proc. of Data Compression Conference (DCC '98)*, Snowbird, Utah, p.572, Mar. 30–Apr. 1, 1998.

[9] I. Matsuda, T. Siodera and S. Itoh, "Lossless Video Coding Using Variable Block-Size MC and 3D Prediction Optimized for Each Frame," *Proc. of 12th European Signal Processing Conference (EUSIPCO 2004)*, pp.1967–1970, Vienna, Austria, Sep. 6–10, 2004.

[10] http://bmrc.berkeley.edu/frame/research/mpeg/