Journal of
Electronic Imaging

SPIEDigitalLibrary.org/jei

Improved lossless coding algorithm in H.264/AVC based on hierarchical intraprediction and coding-mode selection

Li-Li Wang
Wan-Chi Siu

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

Journal of Electronic Imaging 20(4), 043001 (Oct–Dec 2011)

Improved lossless coding algorithm in H.264/AVC based on hierarchical intraprediction and coding-mode selection

Li-Li Wang
Wan-Chi Siu
The Hong Kong Polytechnic University
Department of Electronic and Information Engineering
Hung Hom, Kowloon, Hong Kong
E-mail: enwcsiu@polyu.edu.hk

Abstract. *In this paper, an improved lossless intracoding algorithm based on H.264/AVC framework is proposed. In the proposed algorithm, two contributions have been made. One is that samples in a macroblock (MB)/block are hierarchically predicted instead of using block-based prediction as a whole. More specifically, four groups are extracted from the samples in an macroblock/block. Samples in the first group are first predicted based on directional intraprediction method, and then the samples in other groups are predicted using the samples in the first group as the references. As a result, the information left in the residual block can be reduced since the samples can be accurately predicted by using nearer references. The other contribution is that two coding modes are designed to efficiently encode the resulting residual block. A better coding mode can be selected based on the rate optimization. Experimental results show that compared with other methods in the literature the proposed algorithm gives a much better compression performance.* © 2011 SPIE and IS&T. [DOI: 10.1117/1.3644573]

1 Introduction

The lossless coding technique plays an important role in perfectly preserving valuable information using less storage space. Many approaches have been proposed and are mainly developed for image coding. A good example in the standards is the JPEG lossless coding.[1] It has many applications, such as medical images, digital archives, digital documentations, and so on. In the past few years, some further improvements of lossless image coding studies have been done. One of the representatives is an edge-directed prediction method in Ref. 2, which enables the predictor to be adaptive from smooth regions to edge regions. Note that the complexity is usually high for lossless image coding. With the development of video coding standards, high definition videos are gradually entering our daily life. The demand for higher fidelity of encoding videos is ever increasing. It is highly possible that either the whole picture or some regions of a picture in a video need to be represented without any loss of fidelity for some demanding applications. Consequently, it is necessary to integrate the lossless coding technique into video coding standards, such as the H.264/Advanced Video Coding (AVC) coding standard.[3,4] The H.264/AVC standard was originally designed for this purpose of compression. The compression efficiency of video coders can usually be achieved based on the lossy coding technique. However, it also supports lossless video coding for some requirements.

The lossless coding method used in H.264/AVC standard can be considered as the state-of-art representative, which is applied to the video coding standard. Since there is no transform and quantization process in the lossless video coding, the encoding complexity can be substantially reduced but more bits are required to encode the video sequences. Another problem is that, since the H.264/AVC has a block-based coding structure, samples away from the references are not well predicted due to poor correlation. Therefore, further improvement of the lossless coding method in H.264/AVC standard is always desirable for the purpose of compression. The third problem is that the entropy coding method context-based adaptive variable length coding (CAVLC) designed for the quantized discrete cosine transform (QDCT) coefficients does not play a good role in directly coding the residual due to the different distributions between the residual and the quantized DCT coefficients.

Recently, researchers have made various studies on the improvement of the lossless compression algorithm in the H.264/AVC standard. Lee et al.[5] proposed a lossless intraprediction method based on samplewise differential pulse code modulation (DPCM) instead of using a block-based approach. This algorithm achieves better performance since near samples are used as the reference, and it is also adopted into a new enhancement project of the H.264/AVC standard. However, the samplewise DPCM method can only be applied to the modes with one sample predictor. As a result, only four modes (modes 0, 1, 3, and 4) of I4 macroblock (MB)/I8MB type, two modes of I16MB type/Chroma prediction can be performed based on the samplewise DPCM method. Thus, the performance is limited. In Ref. 6, a simple interpolation method is used to make a prediction. As a result, a samplewise DPCM concept can be easily extended to other prediction modes. The experimental results in Ref. 6 and our realization results in Sec. 4 show that further improvement of the lossless compression performance can be achieved compared with the algorithm in Ref. 5. In Ref. 7, a two-layer coding algorithm is proposed to improve the lossless coding performance of H.264/AVC standard. In this method, an MB is first encoded using the lossy coding

---

Paper 11077R received Apr. 21, 2011; revised manuscript received Aug. 16, 2011; accepted for publication Sep. 8, 2011; published online Oct. 14, 2011.

1017-9909/2011/20(4)/043001/10/$25.00 © 2011 SPIE and IS&T

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

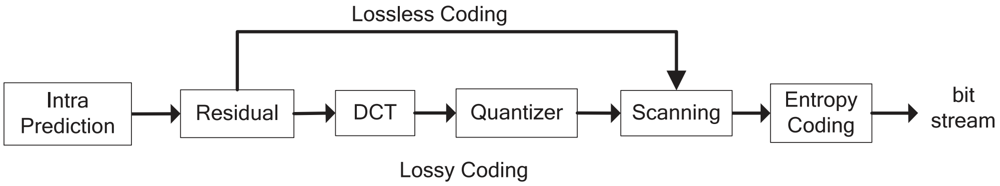

Fig. 1 Lossless and lossy codings adopted in FRExt high profiles of the H.264/AVC standard.

method in H.264/AVC, and then the distortion between the original MB and the reconstructed MB is also encoded to form a so-called D bitstream. The method with a two-layer bitstream can be used in applications with multiple clients for different receiving capabilities. However, our experimental results and those from Ref. 7 show that this method has almost the same compression rate as the method in H.264/AVC,[4] and the coding performance is worse when compared with the method in Ref. 5. Note that this method has a high computational complexity due to two coding processes: lossy and lossless codings. Especially for the lossy coding, the complexity is high due to transformation and the quantization process.

In this paper, a hierarchical intraprediction and coding method is suggested to improve the lossless compression efficiency of an encoder. In this method, samples in an MB/block are classified into different groups. The samples in the first group are first predicted using the directional intraprediction method in the H.264/AVC standard, and then the remaining samples in other groups can be adaptively predicted based on the minimal gradient between neighboring references. After the prediction has been finished, two candidate coding modes: one- and two-layer coding modes, are designed to encode the residual. The final coding mode is, thus, selected based on the rate optimization (RO) method. Experimental results show that this method can solve the inaccuracy problem of the intraprediction, and thereby the coding efficiency can be enhanced when compared with other lossless intracoding methods.[5–7] Furthermore, the compression ratio can be further improved due to a better coding scheme for the residual. Although the proposed method involves two coding processes, it does not still lead to a substantial increase in the complexity of an encoder due to the fact that only part of the samples are predicted using the RO method.

The rest of the paper is organized as follows. First, the lossless intraprediction technique adopted in the H.264/AVC is introduced in Sec. 2. Second, a new lossless intraprediction algorithm is presented in Sec. 3. Subsequently, experimental results are shown in Sec. 4. Finally, a conclusion is provided in Sec. 5.

2 Lossless Intracoding in H.264/AVC

The H.264/AVC standard initially includes three profiles: the baseline, main, and extended profiles, for different applications. In these profiles, lossless coding can be done by an intra-pulse-code-modulation (IPCM) macroblock type, in which the values of the samples are directly encoded without prediction, transformation, or quantization.[3,4] It is clear that this mode is not efficient, since it requires many bits to encode an macroblock without any loss of fidelity.

As a development of the H.264/AVC standard, a more effective lossless coding method was designed in the new profiles called FRExt high profiles.[4] Figure 1 shows the lossless and the lossy coding processes of an intra-residual block/MB in the FRExt high profiles. Note that, in the lossy video coding, the residual block should be transformed and quantized, and then the quantized DCT coefficients are scanned and encoded by entropy coding. Different from that, the residual block produced by intraprediction is directly scanned and encoded by entropy coding without the transformation or quantization process in the lossless coding.

From Fig. 1, we can see that the intraprediction algorithm used to produce the residual is the same for both lossless and lossy codings. In FRExt high profiles, four basic types of intraprediction for luma components are available: 16×16 luma prediction (I16MB type), 8×8 luma prediction (I8MB type), 4×4 luma prediction (I4MB type), and IPCM.[3,4] There are four prediction modes in the I16MB type and nine prediction modes in the I4MB/I8MB types. Chroma components can be predicted using four modes, which are similar to those in the I16MB type. Details about these four types can be found in Refs. 3 and 4.

The new lossless coding method makes the encoder very efficient as compared with the IPCM mode. However, we also note that the intra-block/MB is predicted as a whole by an extrapolation of the neighboring reconstructed pixels in the H.264/AVC standard. As a result, the samples that are far from their references may not be predicted well due to poor correlation. Especially for the blocks predicted using the I8MB or I16MB type, the poor prediction is more obvious since a larger block size is used. Another drawback in the lossless coding of H.264/AVC framework is that the CAVLC entropy coding method is usually not efficient to directly encode the residual block since it is designed based on some characteristics of quantized DCT coefficients of the residual block. In order to resolve these problems, we propose a new lossless coding method based on hierarchical intraprediction and residual coding scheme selection, which will be elaborated in Sec. 3, to improve the coding efficiency of the I8MB and I16MB types in the H.264/AVC standard.

3 Proposed Algorithm based on Hierarchical Intraprediction

To obtain a better prediction for samples away from their references, a hierarchical intraprediction (HIP) method is presented in this section. The details about the HIP method are given from two aspects: the HIP methods for the I8MB type and the I16MB type/Chroma intraprediction. Furthermore,

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

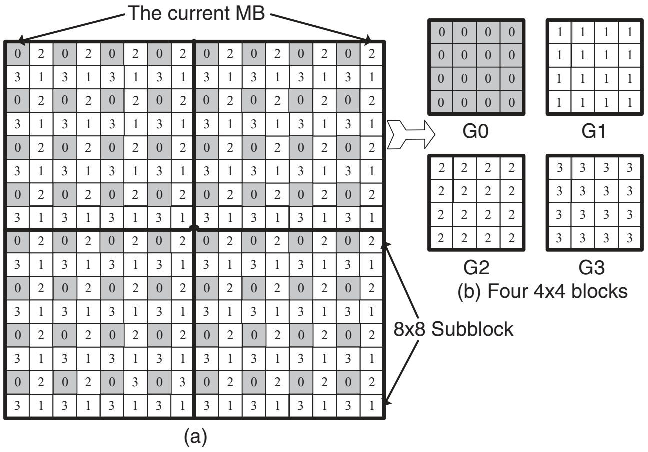

Fig. 2 Structure of HIP for the I8MB type.

a coding scheme selection method is proposed to efficiently encode the resulting residual by the HIP method.

3.1 *Improved Hierarchical Intraprediction Method for I8MB Type*

In FRExt profiles, the I8MB type is introduced to improve the encoding efficiency of the intraprediction in the H.264/AVC standard. An MB is first divided into four 8×8 blocks. For each 8×8 block, nine prediction modes are used to do intraprediction. In order to improve the prediction accuracy for samples away from their references, we suggest that samples in each 8×8 block be hierarchically predicted instead of using the whole-block-based method in the original H.264/AVC standard. As a result, four groups from G0 to G3 in each 8×8 block are classified as shown in Fig. 2. By sampling with the corresponding group numbers as shown in Fig. 2(a), four 4×4 blocks are formed as shown in Fig. 2(b) for the residual entropy coding process.

Figure 3 shows a block diagram of the proposed HIP for an 8×8 block, where *o* denotes an original block G0, *p* denotes the predicted block of G0, and *r* is the difference between *o* and *p*. The detailed prediction strategies of samples in the four groups are presented as follows.

First, select the best prediction mode for samples in G0 from nine prediction modes of the I8MB type. Each 8×8 block in an MB is predicted as a whole according to the nine prediction modes of the I8MB type in the H.264/AVC standard. However, only the predicted values of the 16 samples labeled as 0 are computed based on a samplewise DPCM method instead of using all 64 samples. After the samples in G0 are predicted from their reconstructed neighbors using one prediction mode, the residual block *r* can then be obtained by subtracting the predicted block *p* from the original block *o*. For the residual block *r*, one-layer coding (OLC) and two-layer coding (TLC) schemes are designed to encode it. The rate costs (J_{OLC} and J_{TLC}) corresponding to the two coding schemes can be given by Eqs. (7) and (8), respectively. The coding scheme with the smaller rate cost is selected as the final coding scheme for the residual block *r*, and the corresponding rate cost is chosen as the rate of this prediction mode as shown in Eq. (1). The details about the two coding schemes will be illustrated in Sec. 3.4.

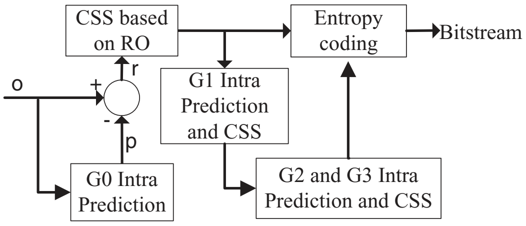

Fig. 3 Block diagram of the proposed HIP.

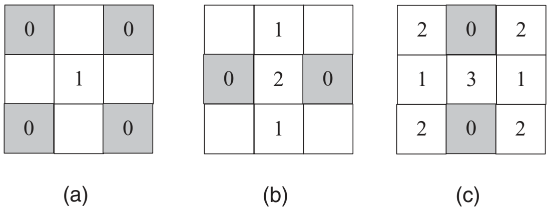

Fig. 4 Predicted structures of samples in G1, G2, and G3.

$$J\,(MODE) = \begin{cases} J_{OLC}, & if\ J_{OLC} < J_{TLC} \\ J_{TLC}, & otherwise \end{cases}, \tag{1}$$

where *MODE* represents one of the nine intraprediction modes of the I8MB type.

By using Eq. (1), each of the nine residual blocks produced by the nine prediction modes for G0 has a smaller coding cost. Finally, the prediction mode with the minimum coding cost is selected as the prediction mode of G0. Since it is a lossless coding process, there is no degradation in the reconstructed videos. The samples can be reconstructed without any loss of fidelity. As a result, the mode with the minimum rate can make the best compression efficiency for an encoder based on the H.264/AVC framework.

Second, make intraprediction for the samples in G1, G2, and G3 based on a measure of the intensity gradient. The remaining 48 pixels in the 8×8 block are classified into three groups (G1, G2, and G3), which correspond to three 4×4 blocks as shown in Fig. 2(b). Samples in G1 are first predicted using samples in G0 as references. The predicted values of samples can be adaptively obtained by Eq. (2) based on a measure of the gradients along different directions (see Fig. 4), and then the residual block can be obtained by subtracting the predicted G1 from the original G1. Similar to the coding method of the residual blocks for G0, the OLC and TLC schemes are adopted to encode the residual block of G1. The only difference of the residual coding schemes between G0 and G1 is that there is only one residual block for G1 since the predicted values of samples in G1 are adaptively obtained based on a measure of the gradients not based on the nine prediction modes. A better coding scheme for the residual block of G1 is thus selected by comparing J_{OLC} with J_{TLC}. Since it is a lossless coding process, G1 can be perfectly reconstructed. Subsequently, samples in G2 and G3 can be predicted using samples in G0 and G1. The predicted values of samples in the G2 and G3 are also adaptively obtained based on the measure of the gradients along different directions. Finally, the coding scheme for the residual block of G2 or G3 can be determined using the method similar to

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

G1. Figure 4 shows the predicted structures of samples in G1, G2, and G3. Referring to Fig. 4, the predicted values of samples in G1, G2, and G3 are calculated by Eqs. (2)–(4), respectively,

$$p_{i,j}^{1}=\begin{cases}\left(f_{i-1,j-1}^{0}+f_{i+1,j+1}^{0}+1\right)\gg 1, & \left|f_{i-1,j-1}^{0}-f_{i+1,j+1}^{0}\right|<\left|f_{i+1,j-1}^{0}-f_{i-1,j+1}^{0}\right|\\ \left(f_{i+1,j-1}^{0}+f_{i-1,j+1}^{0}+1\right)\gg 1, & otherwise\end{cases}, \tag{2}$$

$$p_{i,j}^{2}=\begin{cases}\left(f_{i-1,j}^{0}+f_{i+1,j}^{0}+1\right)\gg 1, & \left|f_{i-1,j}^{0}-f_{i+1,j}^{0}\right|<\left|f_{i,j-1}^{1}-f_{i,j+1}^{1}\right|\\ \left(f_{i,j-1}^{1}+f_{i,j+1}^{1}+1\right)\gg 1, & otherwise\end{cases}, \tag{3}$$

$$p_{i,j}^{3}=\begin{cases}\left(f_{i-1,j}^{1}+f_{i+1,j}^{1}+1\right)\gg 1, & \left|f_{i-1,j}^{1}-f_{i+1,j}^{1}\right|<\left|f_{i,j-1}^{0}-f_{i,j+1}^{0}\right|\\ \left(f_{i,j-1}^{0}+f_{i,j+1}^{0}+1\right)\gg 1, & otherwise\end{cases}, \tag{4}$$

where $p_{i,j}^{m}$ denotes the predicted value of the sample in position (i,j) of group m (Gm) and $f_{i,j}^{m}$ denotes the reconstructed value of the reference sample in the position (i,j) of Gm.

Since the best predicted direction (horizontal or vertical/diagonal down left or diagonal down right) of each sample in the current block can be adaptively determined based on the measure of intensity gradient according to the neighboring reconstructed samples, there is no extra bit required to encode the predicted direction. Moreover, since samples in G1, G2, and G3 are predicted using the neighboring reconstructed samples, which are equal to the original samples in a lossless coding, the correlation between them is generally higher. As a result, the information left in the residual can be significantly reduced compared with the original block-based method.

3.2 *Improved Hierarchical Intraprediction Method for I16MB Type/Chroma Prediction*

Similar to the I8MB type (except the block size), the whole macroblock with a 16×16 size is classified into four groups from G0 to G3 as shown in Fig. 5(a). Samples in G0 are predicted using four modes in the I16MB type based on the samplewise DPCM method. Also two coding schemes are adopted to optimize the residual encoding for each 8×8 block. The coding scheme with the smaller rate cost is selected as the encoding method of the residual block produced by one of four prediction modes. The mode with the minimal rate cost is selected and encoded. Samples in G1, G2, and G3 are then predicted using the neighboring reconstructed samples as Eqs. (2)–(4). The prediction of Chroma cr/cb block can also be performed in a similar way to the HIP method of the I16MB type. The structure of HIP for a Chroma block is shown in Fig. 5(b).

Based on the above analysis, we can see that for each MB, only part of the samples are predicted using the RO process based on the proposed algorithm. As a result, the complexity of the encoder is not dramatically increased in spite of two residual coding schemes. Another observation is that the best prediction directions of the samples in the last three groups can be adaptively selected based on the minimal gradients. Hence, the prediction directions of the samples in a block may be different using the proposed algorithm. This cannot be obtained using the previous intraprediction methods,[5–8] and it is also one of the reasons that the proposed algorithm can show better compression performance. This last point is to be verified in Sec. 6.

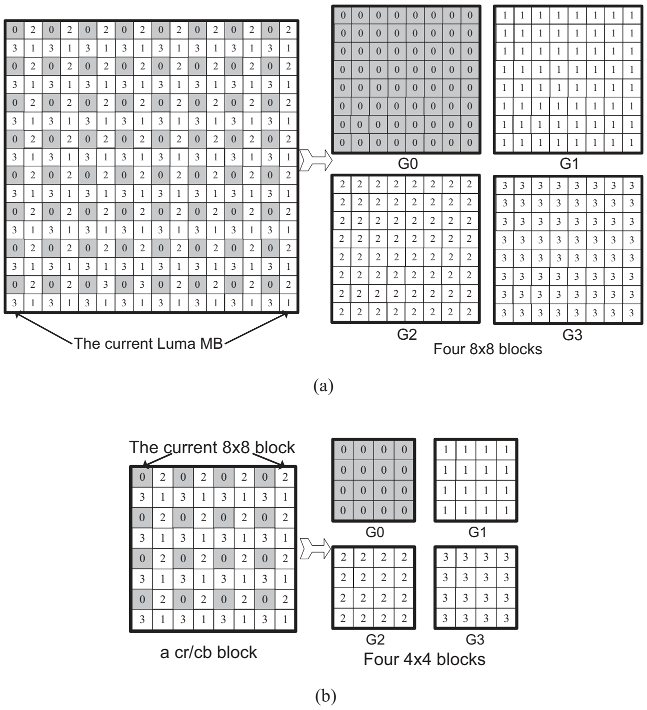

Fig. 5 Structure of HIP: for (a) I16MB type and (b) Chroma prediction.

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

3.3 *Further Improvement of the Hierarchical Intraprediction Method*

Compared with the I8MB and I16MB types, the I4MB type is associated with a 4×4 block, which is small. Experimental results show that the hierarchical intraprediction method does not have a better effect on the I4MB type. Therefore, only the samplewise DPCM method in Ref. 5 is used to improve the prediction of the I4MB type for lossless coding. The difference between them is the prediction of the DC mode. In Ref. 5, the authors said that the samplewise DPCM method is not suitable to the DC mode since more than one reference sample is used to perform interpolation. However, extensive experimental results show that the DC mode cannot show a better performance since the average value is used as the predicted values of the whole MB/block. Note that the DC mode is assigned a relatively small mode index (2) among all prediction modes in the H.264/AVC standard, which means that this mode has a higher probability to be selected as the best prediction mode. Therefore, improving the prediction structure of the DC mode has great significance for an efficient lossless coding. The method as shown in Eq. (5) is proposed to improve the prediction of the DC mode in every intraprediction type (I4MB type, I8MB type, I16MB type, or Chroma prediction) according to the availability of the neighboring reference pixels.

$$p_{i,j}^{0} = \begin{cases} \left(f_{i,j-1}^{0} + f_{i-1,j}^{0} + 1\right) \gg 1, & up\_available \\ \left(f_{i-1,j}^{0} + f_{i-1,j+1}^{0} + 1\right) \gg 1, & left\_available \\ 128, & otherwise \end{cases}, \tag{5}$$

where coordinates (*i*,*j*) correspond to samples in G0. Similarly, in order to use the concept of DPCM for the plane mode in the I16MB type and the Chroma prediction, the predicted value of the sample in G0 for the plane mode is given as follows:

$$p_{i,j}^{0} = \left(f_{i,j-1}^{0} + f_{i+1,j}^{0} + 1\right) \gg 1. \tag{6}$$

In addition, two coding schemes are also applied to the residual blocks produced by the nine prediction modes of the I4MB type in order to efficiently compress the residual. Figure 6 shows an example of an original, reconstructed, predicted, and their corresponding residual frames for several lossless coding algorithms. From the results, it can be observed that the reconstructed frame is the same as the original frame since it is a lossless coding process. Another observation from the predicted frames is that the proposed algorithm makes a best prediction among all of these algorithms, and thereby, the information left in the residual frame (d-2) is the smallest.

3.4 *Two Coding Schemes for the Residual Block and the Selection Criterion of the Better One*

Note that lossless compression techniques usually include two major components: prediction and entropy coding. Having a better prediction method can result in less information in the residual block (see Fig. 6). Also a proper entropy coding is equally important for compressing the residual block. It is well known that the CAVLC method[3,4] is mainly designed to compress the QDCT coefficients, so it is not efficient to directly compress the residual block if the distribution of the residual is different from that of the QDCT coefficients of the residual block. Therefore, the CAVLC method should be correspondingly improved for a lossless coding to obtain a better compression efficiency. Actually, an improved CAVLC method for lossless residual coding has been investigated in Ref. 9. However, this method is proposed based on the statistics of the residual produced by the lossless coding method in the H.264/AVC standard. In other words, the method in Ref. 9 is highly dependent on the distribution of the residual. Different residual distributions require different coding design. As a result, it is hard to directly apply this method to the residual produced by other lossless coding methods. Different from improving CAVLC structure, we propose a general method to improve the compression efficiency from the view of residual block in lossless coding.

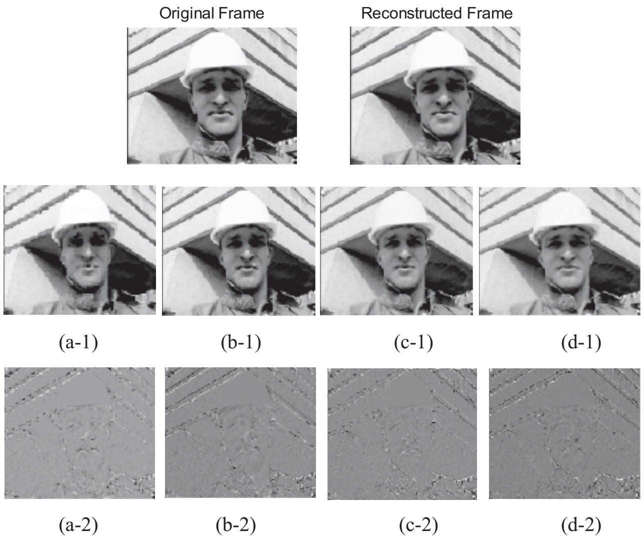

Fig. 6 Experimental results on Luma component for the first frame of Foreman sequence with QCIF (quarter common intermediate format): (a-1/2) predicted/residual frame based on H.264/AVC, (b-1/2) predicted/residual frame based on DPCM, (c-1/2) predicted/residual frame based on SI_DPCM, and (d-1/2) residual/residual frame based on HIP.

After some investigation, we find that some residuals show similar characteristics to that of the QDCT coefficients, and others do not. According to these two cases, we propose two coding schemes to compress the residual block, which

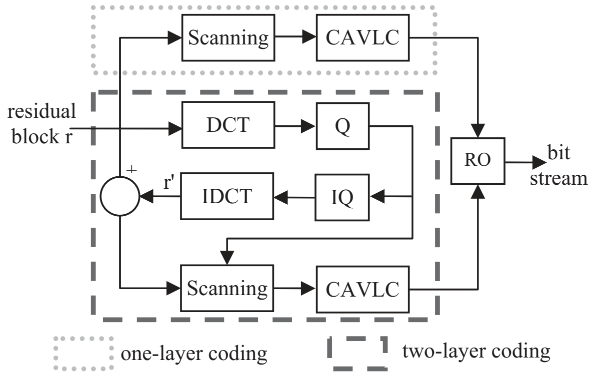

Fig. 7 Coding scheme selection based on the rate optimization.

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

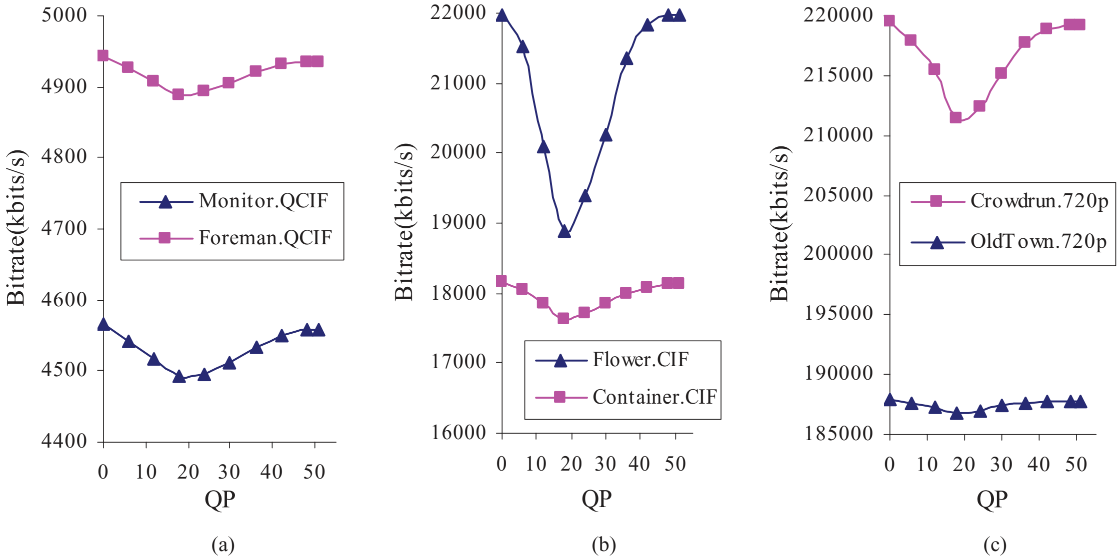

Fig. 8 Relationship between bitrate and QP.

produce efficient results as shown below. Figure 7 shows the block diagram about the coding scheme selection process.

As shown in Fig. 7, there are two coding schemes: one- and two-layer coding schemes, designed for encoding the residual block. In the one-layer coding scheme, the residual block is directly scanned and entropy coded by CAVLC method. While the two-layer coding scheme includes lossy and lossless coding processes. The residual block is first transformed and quantized to obtain the QDCT coefficients in the lossy coding process. The QDCT coefficients are scanned and entropy coded. While for the lossless coding process, inverse quantization and inverse DCT are applied to the QDCT coefficients to obtain the reconstructed residual block r'. The difference between the original residual block r and the reconstructed residual block r' are scanned and entropy coded. Consequently, how to select a better coding scheme from the two candidate schemes is a critical problem during the encoding. In the reference software JM of the H.264/AVC standard, the Lagrangian optimization is introduced to make a prediction mode decision. It aims at minimizing the distortion under a constraint on the rate. However, there is no distortion for a lossless coding. Even though the distortion can be found in the lossy coding process of the two-layer coding scheme, it does not require optimization since the compression ratio is the only element that we care about in a lossless coding. As a result, a rate optimization method is suggested by us to select a better residual coding scheme. The rates corresponding to the two coding schemes can be obtained by Eqs. (7) and (8), respectively. The coding scheme with a smaller rate cost is selected for each residual block. By using the rate optimization method, the largest compression efficiency can be obtained

$$J_{OLC} = R\,(residual|MODE)\,, \tag{7}$$

$$J_{TLC} = R_{lossy}\,(QDCT|MODE,QP) + R_{lossless}\,(distortion|MODE,QP)\,, \tag{8}$$

where QP is a quantization parameter. J_{OLC} ($R(residual|MODE)$) and J_{TLC} represent the number of bits used to encode the residual block by the one- or two-layer coding scheme, respectively. $R_{lossy}(QDCT|MODE,QP)$ represents the number of bits used to encode the QDCT coefficients, and $R_{lossless}(distortion|MODE,QP)$ represents the number of bits used to encode the distortion between r and r'. Note that "$MODE$" is used to represent one of the nine intraprediction modes for G0 block that requires a mode decision process. For other blocks, such as G1, G2, and G3, Eqs. (7) and (8) should be changed into Eqs. (9) and (10), respectively. Because these blocks are adaptively predicted based on a gradient measure method without mode decision process,

$$J_{OLC} = R\,(residual)\,, \tag{9}$$

$$J_{TLC} = R_{lossy}\,(QDCT|QP) + R_{lossless}\,(distortion|QP)\,. \tag{10}$$

Basically, when the residual block has similar characteristics with the QDCT coefficients, the one-layer coding scheme is usually the better selection. By contrast, if the characteristics of the residual are very different from the QDCT coefficients, the two-layer coding scheme shows a better compression performance. Compared with the method in Ref. 7 where only a two-layer coding method is used, our method can provide the better compression ratio by selecting a better one from two coding schemes based on the rate optimization method. Another improvement compared with the method in Ref. 7 is that the rates in both lossy and lossless layers for the two-layer coding scheme are combined together to be optimized for both prediction mode decision and coding scheme selection. This makes the proposed algorithm give the best compression efficiency, which will be experimentally verified in Sec. 6.

As previously mentioned, the two-layer coding scheme associates with a quantization process. Therefore, a quantization parameter must be specified during the encoding process. In order to achieve the best compression efficiency for lossless coding, more than 20 video sequences were encoded with QP from 0 to 51 in our testing. Figure 8 shows an

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

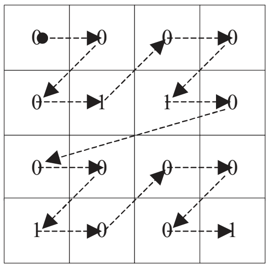

Fig. 9 Map of flags for coding schemes of 16 4×4 blocks.

example about the relationship between bitrate and QP for six video sequences with different frame sizes. According to the experimental results, QP is set to 18 for achieving the best compression efficiency.

3.5 *Coding Method of the Flags to Indicate the Resultant Coding Schemes in an MB*

Due to the introduction of two coding schemes, we have to define a flag (0: one-layer coding scheme; 1: two-layer coding scheme) to indicate which scheme is used during the encoding process, and this flag will also be written into the bit stream. Figure 9 shows an example about the flag map of one macroblock encoded by I4MB or I8MB.

To maximally increase entropy coding efficiency, we propose an improved context-based adaptive variable length coding scheme (ICAVLC) to encode the flags. The flags are first reordered according to the zigzag scanning as shown in Fig. 9 (starting at the top-left) to produce a liner array: 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1. Note that this is a binary map containing "ones" and "zeros" only. Taking this characteristic into consideration, only two syntax elements: *total_zeros* and *run_before* used in CAVLC are used to efficiently encode the flags in ICAVLC and the other three syntax elements are deleted from the original CAVLC method.[3] Similar to CAVLC in the H.264/AVC standard, the "*total_zeros*" denotes the total number of zeros before the last nonzero flag, and "*run_before*" is used to represent the number of consecutive zeros preceding each nonzero flag. In addition, *all_zero_flag* is designed to indicate whether a linear array is only composed of zero values. "$all\_zero\_flag = 1$" denotes that the sixteen flags in a linear array are only composed of zero values, otherwise "$all\_zero\_flag = 0$" means the linear array has at least one element "1." Details of the ICAVLC coding procedure are given as follows:

- Step 1: Encode *all_zero_flag*. If $all\_zero\_flag = 1$, terminate the encoding process. Otherwise go to Step 2.
- Step 2: Encode the number of all zeros before the last nonzero flag (*total_zeros*) using four bits.
- Step 3: Encode the number of consecutive zeros preceding each nonzero flag (*run_before*) in reverse order. Different from CAVLC in the H.264/AVC standard, the last *run_before* is also encoded into the bit stream since there is no syntax element used to represent the number of nonzero flag. In order to incorporate this case, we add a binary codeword "000000000001" to the bottom of Tables 9 and 10 (in Ref. 3) to denote the case when *run_before* is equal to 15.

Table 1 Statistical results for coding schemes of 16 4×4 blocks.

Sequence		Saving (%)	Percentage in total bits (%)
QCIF	Coastguard	14.023	0.124
	Monitor	23.047	0.129
	Foreman	13.243	0.139
	Container	30.960	0.151
CIF	Children	13.497	0.143
	Flower	51.898	0.064
	Hall	21.158	0.141
	Akiyo	37.265	0.138
HDTV (720p)	Sunflower	27.057	0.157
	Crowdrun	9.740	0.113
	Shuttlestart	32.740	0.102
	OldTownCross	29.476	0.108
Average		**25.342**	**0.126**

Different from the HIP methods of the I4MB and I8MB types, where 16 flags are needed for an MB, only four flags corresponding to four 8×8 blocks as shown in Fig. 5(a) are required for an MB based on the HIP method of the I16MB type. In this case, we use four bits to represent the coding schemes of the four 8×8 blocks.

Table 1 lists the percentage of the bits used for coding the flags in total bits of each sequence. Totally, 12 sequences were used with different formats, such as QCIF, CIF (common intermediate format), and HDTV (high definition television). From Table 1, we can see that the percentage of the bits used for coding the flags only accounts for 0.126% of the whole bits on average. This is a very small proportion. Although the bits are increased due to coding the flags, the proposed algorithm can compensate for these bits and provide a higher compression efficiency than other algorithms in the literature as shown in Tables 2–4. Table 1 also shows the saving of bits used in the flags for I4MB and I8MB types, compared with the coding method of "one flag one bit." On average, 25.342% saving can be achieved based on Table 1.

4 Experimental Results

In order to evaluate the performance of the proposed algorithm, the reference software JM 12.2 (Ref. 8) was used to carry out experiments on a number of YUV 4:2:0 format sequences. The test was based on the FRExt high profiles with specifications provided in Ref. 10. All frames were intracoded. Some video sequences referring to Tables 1–3 were used for this work. A total of 150 frames were encoded for each sequence. The frame rate was set to 30. The CAVLC entropy coding method was used for the experiments. Deblocking was not used for lossless coding. In the following

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

Table 2 Results of compression ratios for QCIF sequences.

Sequence	Bitrate (kbits/s)	Δ Bitrate [%]		
(QCIF)	H.264_LS	H.264 _DPCM	SI_DPCM	Proposed method
Coastguard	5855.984	– 10.574	– 11.846	– 14.947
Monitor	5377.597	– 11.853	– 14.863	– 16.509
Foreman	5566.776	– 9.026	– 11.123	– 12.230
Akiyo	4659.207	– 10.843	– 15.904	– 17.732
Mobile	8082.437	– 7.287	– 10.363	– 17.679
Carphone	4945.672	– 11.237	– 14.510	– 15.222
Container	5230.263	– 10.185	– 12.507	– 15.678
TableTennis	5669.637	– 8.190	– 11.269	– 13.551
Salesman	5966.341	– 11.692	– 15.636	– 16.533
Grandmother	5306.482	– 9.132	– 11.572	– 12.687
Average	**5666.040**	– **10.002**	– **12.960**	– **15.277**

tables, positive values mean increase, and negative values mean decrease.

Since there is no degradation in video quality for lossless coding, we only present the compression ratios for different lossless intraprediction methods in terms of encoding bitrates for sequences of different sizes in Tables 2–4. It can be observed from Tables 2–4 that the method based on samplewise DPCM (Ref. 5) gives a higher compression ratio than that of the block-based method in the original H.264/AVC standard. With the improvements of more intraprediction modes, the method based on simple interpolation and DPCM (SI_DPCM) (Ref. 6) achieves a further compression ratio compared with the samplewise DPCM method. Compared with the samplewise DPCM method, the proposed method can achieve a consistent increase in compression ratios for all tested sequences. For most of the sequences, the proposed method also shows a better compression ratio compared to that of the SI_DPCM method. Therefore, we can see that the proposed algorithm can show the best compression efficiency compared with the methods[5–8] based on

Table 3 Results of compression ratios for CIF sequences.

Sequence	Bitrate (kbits/s)	ΔBitrate [%]		
(CIF)	H.264_LS	H.264 _DPCM	SI_DPCM	Proposed method
Children	24264.276	– 15.280	– 23.389	– 22.631
Flower	23235.798	– 3.648	– 5.754	– 18.764
Container	20372.872	– 8.152	– 10.711	– 13.533
Hall	20037.900	– 11.450	– 13.913	– 14.793
Akiyo	16562.680	– 8.763	– 13.217	– 13.258
Foreman	19353.068	– 8.083	– 10.410	– 11.296
Coastguard	23328.964	– 10.615	– 12.444	– 14.840
Silent	21750.124	– 8.450	– 12.036	– 12.274
Tempete	26259.104	– 9.532	– 13.135	– 18.242
Waterfall	27073.632	– 7.946	– 12.964	– 14.871
Average	**22223.840**	– **9.192**	– **12.797**	– **15.450**

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

Table 4 Results of compression ratios for HDTV sequences.

Sequence	Bitrate (kbits/s)	ΔBitrate [%]		
(HDTV)	H.264_LS	H.264 _DPCM	SI_DPCM	Proposed method
Sunflower	168292.240	– 15.361	– 22.372	– 20.656
Crowdrun	241153.440	– 5.462	– 8.035	– 12.402
InToTree	198065.424	– 2.210	– 2.809	– 4.454
OldTownCross	199703.840	– 4.528	– 4.661	– 6.503
Average	**201803.7**	– **6.890**	– **9.469**	– **11.004**

the H.264/AVC framework. Furthermore, the proposed algorithm may be used together with the inter-frame prediction approach used in the H.264/AVC standard, and thereby the lossless compression efficiency of the encoder can also be greatly improved.

Table 5 shows the encoding complexity of the proposed algorithm and the two-layer algorithm in Ref. 7 (labeled by H.264_TL). From Table 5, we can see that the encoding time has been increased compared with JM12.2 for the tested video sequences (such as 64.141% for Coastguard, 52.116% for Children, and 43.864% for Sunflower). However, the rate performance is the only target that we care about when heading for high efficiency video coding (HEVC)/H.265.[11] Recall that one of the objectives of the HEVC (H.265) is to offer a new generation video coding standard, that provides substantially higher compression capability (i.e., around half the bitrate for a similar quality level) even at the expense of significantly higher computational complexity, compared with H.264/AVC.[11] Moreover, some fast algorithms; such as those found in Refs. 12–15, can also be combined with this algorithm, therefore, it gives a possible better selection for real-time applications. Note that only part of samples based on the proposed algorithm are predicted using the multiple prediction modes defined in the H.264/AVC standard. As a result, the complexity of the encoder using the proposed algorithm is still less than the method in Ref. 7 as shown in Table 5 (60.099% increase for the proposed algorithm versus 78.919% increase for Ref. 7 on average).

Table 5 Computational complexity.

Sequence		H.264_TL (%)	Proposed algorithm (%)
QCIF	Coastguard	83.734	64.141
	Monitor	82.956	57.639
	Foreman	87.617	62.348
	Akiyo	76.479	58.570
CIF	Children	67.466	52.116
	Flower	90.614	62.617
	Container	78.167	62.629
	Hall	74.375	58.519
HDTV (720p)	Sunflower	69.703	43.864
	Crowdrun	87.251	64.091
	InToTree	64.793	70.187
	OldTownCross	83.873	64.464
Average		78.919	60.099

5 Conclusion

In this paper, an algorithm based on a hierarchical intraprediction and residual coding scheme selection for lossless coding is presented. Based on the hierarchical prediction, some pixels in an MB/block are first predicted and encoded with the minimal rate cost, and the others are then adaptively predicted based on the minimal gradient approach using the nearer references. The proposed hierarchical prediction method makes the intraprediction more accurate, and thus the information left in the residual block can be significantly reduced. Two coding schemes are then designed to encode the residual block. The optimal one can be selected based on the rate optimization method. Furthermore, a more efficient ICAVLC entropy coding method is proposed to encode the flags of the coding schemes for an MB. Experimental results show that the proposed algorithm can improve the compression efficiency of the lossless coding in the H.264/AVC standard.

In fact, there is still some room for improving the efficiency of the proposed lossless coding method. Much work can be done on the prediction method of samples in G1 to G3. Some new intraprediction and coding schemes, such as those in Refs. 16–19, may be considered to improve the prediction of samples in G1 to G3. This can be considered as an interpolation process. Equations (2)–(4) are simple representatives of an interpolation method, and only two directions are involved in it. If a better interpolation method and multiple directions are used to do prediction, the information left in the residual block should be further reduced. Consequently, only a small number of bits may be required to encode the residual block. The compression ratio can then be further increased.

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms

We would also like to stress that although the algorithm is proposed based on the intracoding method in H.264/AVC, it can also be extended to the intraprediction method[20,21] in HEVC/H.265 by some modification for the prediction of the samples in G0. Moreover, it also has the potential to be used to compress the depth information in 3D video coding. This is a fruitful direction for further research work.

Acknowledgments

This work is supported by the Centre for Multimedia Signal Processing, Hong Kong Polytechnic University (G-U863) and the Research Grant Council of the Hong Kong SAR Government (PolyU 5267/07E).

References

1. W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Data Compression Standard*, Van Nostrand Reinhold, New York (1992).
2. X. Li and M. T. Orchard, "Edge-based prediction for lossless compression of natural images," *IEEE Trans. Image Process.* **10**(6), 813–817 (2001).
3. ITU-T Rec. H.264/ISO/IEC 11496-10, Advanced video coding, Final Committee Draft, Document JVTG050, Mar. 2003.
4. G. J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions," *Proc. SPIE* **5558**, 454 (2004).
5. Y. L. Lee, K. H. Han and G. J. Sullivan, "Improved lossless intra coding for H.264/MPEG-4 AVC," *IEEE Trans. Image Process.* **15**, 2610–2615 (2006).
6. S. T. Wei, S. R. Shen, B. D. Liu, and J. F. Yang, "Lossless image and video coding based on H.264/AVC intra predictions with simplified interpolations," in *Proc., IEEE Int. Conf. on Image Processing, ICIP'2009*, pp. 633–636, Cairo, Egypt, Nov. 2009.
7. J. R. Ding, J. Y. Chan, F C. Yang, and J. F. Yang, "Two-layer and adaptive entropy coding algorithms for H.264-based lossless image coding," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'2008)*, pp. 1369–1372, Las Vegas, Nevada, April, 2008.
8. *H.264/AVC Reference Software, JointModel (JM) Ver. 12.2* [Online]. Available: http://iphome.hhi.de/suehring/tml.
9. J. Heo, S. H. Kim, and Y. S. Ho, "Improved CAVLC for H.264/AVC Lossless Intracoding," *IEEE Trans. Circuits Syst. Video Technol.* **20**(2), 213–222 (2010).
10. G. Sullivan, "Recommended simulation common conditions for coding efficiency experiments revision 1," presented at *31st VCEG-AE010R1 meeting*, Marrakech, Morocco, 13–19 January, 2007.
11. K. Ugur, K. Andersson, A. Fuldseth, G. Bjontegaard, L. P. Endresen, J. Lainema, A. Hallapuro, J. Ridge, D. Rusanovskyy, C. Zhang, A. Norkin, C. Priddle, T. Rusert, J. Samuelsson, R. Sjoberg, and Z. Wu, "High Performance, Low Complexity Video Coding and the Emerging HEVC Standard," *IEEE Trans. Circuits Syst. Video Technol.* **20**(12), 1688–1697 (2010).
12. L. L. Wang and W. C. Siu, "H.264 fast intra mode selection algorithm based on direction difference measure in the pixel domain," in *Proc., Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP'2009*, pp. 1037–1040, Taipei, Taiwan, Apr. 2009.
13. S. Milani, "Fast H.264/AVC FRExt intra coding using belief propagation," *IEEE Trans. Image Process.* **20**(1), 121–131 (2011).
14. K.-C. Hui, W.-C. Siu, and Y.-L. Chan, "New adaptive partial distortion search using clustered pixel matching error characteristic," *IEEE Trans. Image Process.* **14**(5), 597–607 (2005).
15. Y. Zhang, W. C. Siu, and T. Z. Shen, "Comprehensive scheme for subpixel variable block-size motion estimation," *J. Electron. Imaging* **20**(1), 013014 (2011).
16. S. Matsuo, S. Takamura, K. Kamikura, and Y. Yashima, "Extension of Intra Prediction Using Multiple Reference Lines," presented at *32nd VCEG-AF05 meeting*, San Jose, CA, USA, 20–21 April, 2007.
17. D. C. Garcia and R. L. Queiroz, "Least-squares directional intra prediction in H.264/AVC," *IEEE Signal Process. Lett.* **17**(10), 831–834 (2010).
18. A. Masmoudi, W. Puech, and M. S. Bouhlel, "Efficient adaptive arithmetic coding based on updated probability distribution for lossless image compression," *J. Electron. Imaging* **19**(2), 023014 (2010).
19. K. Zhang, S. Lei, and W. Gao, "Enhanced intra prediction and transform video coding," presented at *Proc. IEEE Int. Conf. on Image Processing (ICIP2010)*, pp. 3381–3384, Hong Kong, September 2010.
20. J. Lainema and K. Ugur, "Intra picture coding with planar representations," in *Proc. of Picture Coding Symposium (PCS2010)*, pp. 198–201, Nagoya, Japan, Dec. 2010.
21. K. McCann, B. Bross, S. Sekiguchi, and W. J. Han, "HM3: high efficiency video coding (HEVC) test model 3 encoder description," presented at *JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-E602, 5th Meeting*, Geneva, CH, 16–23 March, 2011.

Lili Wang received her BEng degree from the Shandong Normal University and her MSc degree from the Shandong University in 2004 and 2007, respectively. She has published over 10 research papers in the past six years. Her research interests include digital signal processing and image and video coding. She is now working for her PhD degree in the Centre for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University.

Wan-Chi Siu received his MPhil and PhD degrees from CUHK and Imperial College, UK, in 1977 and 1984, respectively. He joined the Hong Kong Polytechnic University in 1980 and was head of Electronic and Information Engineering Department and subsequently dean of Engineering Faculty between 1994 and 2002. He has been chair professor since 1992 and is now director of the Centre for Signal Processing. He is an expert in DSP, specializing in fast algorithms, video coding, transcoding, 3D-videos, and pattern recognition and has published 380 research papers, over 160 of which appeared in international journals, such as IEEE Transactions on Image Processing. He has been guest/associate editor of many journals, keynote speaker (e.g., PCM'2002), and chief organizer (ISCAS'1997, ICASSP'2003, and ICIP'2010) of many important international conferences. His work on fast computational algorithms (such as DCT) and motion estimation algorithms has been well received by academic peers, with good citation records, a number of which are now being used in hi-tech industrial applications, such as modern video surveillance and video codec design for HDTV systems (http://www.eie.polyu.edu.hk/~wcsiu/mypage.htm).

Downloaded From: http://electronicimaging.spiedigitallibrary.org/ on 08/05/2013 Terms of Use: http://spiedl.org/terms