

MIT Open Access Articles

Adaptive residual DPCM for lossless intra coding

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Cai, Xun, and Jae S. Lim. "Adaptive Residual DPCM for Lossless Intra Coding." Edited by Amir Said, Onur G. Guleryuz, and Robert L. Stevenson. Visual Information Processing and Communication VI (March 4, 2015). © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE)

As Published: <http://dx.doi.org/10.1117/12.2074833>

Publisher: SPIE

Persistent URL: <http://hdl.handle.net/1721.1/100839>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Adaptive Residual DPCM for Lossless Intra Coding

Xun Cai and Jae S. Lim

Research Laboratory of Electronics
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, 02139, USA

ABSTRACT

In the Differential Pulse-code Modulation (DPCM) image coding, the intensity of a pixel is predicted as a linear combination of a set of surrounding pixels and the prediction error is encoded. In this paper, we propose the adaptive residual DPCM (ARDPCM) for intra lossless coding. In the ARDPCM, intra residual samples are predicted using adaptive mode-dependent DPCM weights. The weights are estimated by minimizing the Mean Squared Error (MSE) of coded data and they are synchronized at the encoder and the decoder. The proposed method is implemented on the High Efficiency Video Coding (HEVC) reference software. Experimental results show that the ARDPCM significantly outperforms HEVC lossless coding and HEVC with the DPCM. The proposed method is also computationally efficient.

Keywords: Lossless coding, DPCM, HEVC, adaptive

1. INTRODUCTION

In image and video compression, lossless coding is used for perfect preservation of visual data. The perfect preservation property is desirable in many image and video applications. For example, lossless coding can be used in preserving watermark information embedded in images and videos. In video editing, lossless coding can prevent accumulation of quantization error when the video is repeatedly encoded. In addition, lossless coding can be used to accurately preserve numerical data for further processing, such as in medical imaging applications.

In many image and video compression systems, the Two-dimensional Discrete Cosine Transform (2D-DCT) [1] is used for lossy coding. The 2D-DCT, however, can not be applied to lossless coding in a straightforward manner. DCT coefficients have to be quantized before they are encoded. The quantization error creates a problem for the perfect preservation requirement in lossless coding. As a result, many approaches have been proposed to replace the 2D-DCT in lossless coding.

One approach is based on transforms. In the methods based on this approach, images or prediction residuals are transformed into integer transform coefficients. Transform coefficients are coded without quantization. For example, in JPEG 2000 lossless image coding [2], 5/3 integer lifting wavelet transform is used. In the VP9 video coding standard [3], the Walsh-Hadamard transform [4] is used in lossless coding. In the work reported in [5] and [6], prediction residuals are transformed into two domains and encoded separately.

Another approach is based on the PCM of prediction residuals. This method is used in the HEVC standard. Specifically, prediction residuals are first obtained. The prediction residuals are coded directly by the entropy coder. Transform and quantization steps are bypassed in HEVC lossless coding [7].

The third approach is based on the DPCM. The DPCM is first proposed in [8] for H.264 [9][10][11] intra lossless coding. In this proposal, directional intra prediction modes are replaced with the sample-wise DPCM on image intensities of the same direction. For example, consider the case where the vertical DPCM is used. The image pixel to be coded is subtracted from the one immediately above it. The DPCM prediction error is then encoded. The DPCM can be used for other directions in a similar way. This strategy is adopted in the latest

Further author information:

Xun Cai: E-mail: cx2001@mit.edu

Jae S. Lim: E-mail: jslim@mit.edu

Visual Information Processing and Communication VI, edited by Amir Said, Onur G. Guleryuz,
Robert L. Stevenson, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 9410, 94100A
© 2015 SPIE-IS&T · CCC code: 0277-786X/15/\$18 · doi: 10.1117/12.2074833

Proc. of SPIE-IS&T Vol. 9410 94100A-1

H.264 intra lossless coding standard by using DPCM for the vertical and horizontal directions. The DPCM-based method can be extended to the HEVC system. For example, vertical and horizontal DPCM is proposed for the HEVC system [12]. In the work reported in [13], directional DPCM with different directions is used for 33 directional intra prediction modes, known as the Sample-based Angular Prediction (SAP). In addition to the directional DPCM, two-dimensional DPCM methods are proposed in [14] and [15]. A significant coding gain is observed when these DPCM-based methods are implemented on the HEVC system. We refer to [16] for other DPCM variations proposed for HEVC.

In this paper, the ARDPCM is proposed for intra lossless coding. We apply the DPCM to intra prediction residuals by predicting a residual pixel as a linear combination of a set of surrounding residual pixels. The DPCM prediction is two-dimensional and intra mode dependent. In order to estimate the correct DPCM weights for each mode, we use an adaptive method. The DPCM weights are obtained in the encoding and decoding processes, by minimizing the Mean Squared Error (MSE) of coded data. Simulation results show that the proposed method significantly increases the performance of HEVC lossless coding.

This paper is organized as follows. In Section 2, a lossless coding system based on the ARDPCM is proposed and analyzed. In Section 3, the proposed method is implemented on the HEVC reference software. Simulation results show that the coding performance increases significantly with the ARDPCM. Finally, we conclude in Section 4.

2. ADAPTIVE RESIDUAL DPCM

In the proposed ARDPCM, a pixel is predicted as a linear combination of a set of causal surrounding pixels. To determine the optimal prediction weights associated with the predicting pixels, we estimate the weights in an adaptive manner. The weights are estimated by minimizing the residual MSE of coded data. Because these weights are synchronized at the encoder and the decoder, no additional side information for the weights is transmitted. We will discuss these points in more detail in following subsections.

2.1 High-level view of the ARDPCM system

The ARDPCM system is shown in Figure 1.

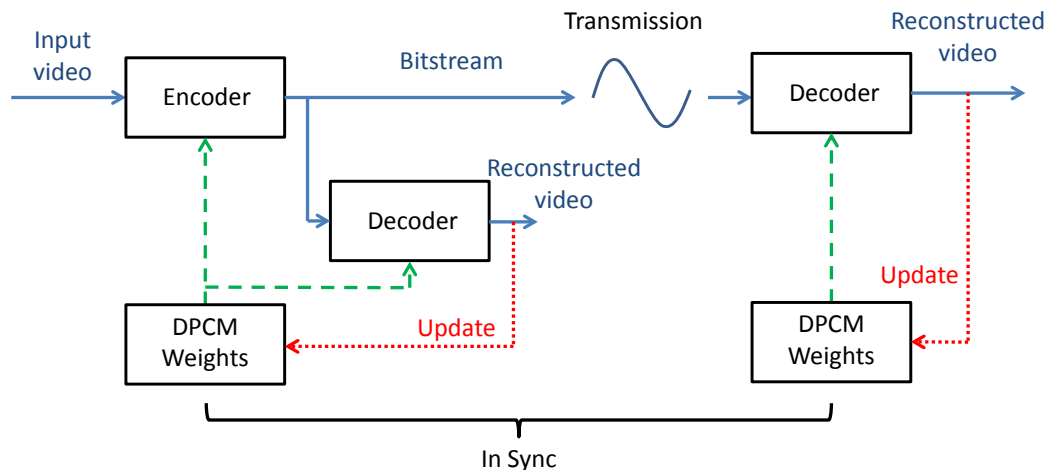


Figure 1. High-level system design

In this system, the video is processed by a video codec with the ARDPCM. The video is encoded and decoded, and the data flow is marked with solid lines. The encoding and decoding processes depend on a set of DPCM weights, marked with dashed lines. These weights are stored and updated from the reconstructed video data at both the encoder and the decoder. The update process is marked with dotted lines. From this system, we can

see that the weights in the encoder and the decoder are always synchronized, if they are initialized to the same initial values and updated with the same rule.

The ARDPCM system in Figure 1 requires two issues for implementation. One is the prediction model used in encoding and decoding video with the given weights. This is discussed in Section 2.2. The other is the method used in updating the weights from the reconstructed video. This is discussed in Section 2.3.

2.2 Prediction model

In this subsection, we discuss how to encode and decode the video based on the DPCM with a given set of DPCM weights. Suppose we obtain the intra prediction residuals and we are interested in encoding the residuals. We denote the residual pixel that is currently being coded as X . We consider predicting X from three of its surrounding pixels [15][17], the one immediately above it, denoted as U , the one immediately left to it, denoted as L , and the one immediately above L , denoted as LU . These three pixels are linearly combined and rounded to form the predictor \hat{X} , with weights w_L, w_U, w_{LU} .

$$\hat{X} = \text{round}(w_L \times L + w_U \times U + w_{LU} \times LU)$$

This is illustrated in Figure 2. The prediction error $X - \hat{X}$ is encoded and transmitted. The rounding process ensures that the prediction error to be transmitted is integer. This is necessary for perfect reconstruction at the decoder side. The rounding process is applied at the last stage of the predictor computation, so that the predictor can be as accurate as possible. At the decoder side, the predictor \hat{X} can be perfectly reconstructed from the same set of coded pixels. The transmitted prediction error is added to \hat{X} to reconstruct X . We note that the weights w_L, w_U, w_{LU} are subject to change as the encoding/decoding process proceeds and the weights are mode dependent. For simplicity, we ignore these dependencies in our notation.



Figure 2. Prediction model used in this paper

We note that this model is a generalization of many sample-wise prediction schemes in the previous research, if mode-dependent weights are allowed. For example, using $w_L = w_U = w_{LU} = 0$ is equivalent to HEVC lossless coding. For directional prediction, using $w_L = 1, w_U = w_{LU} = 0$ for the horizontal mode and $w_U = 1, w_L = w_{LU} = 0$ for the vertical mode is equivalent to H.264 lossless intra coding and the system proposed in [12]. SAP [13] along certain directions can be included in this model, by using up to two non-zero weights. In addition to directional prediction, the proposed model is able to capture two-dimensional correlation in residual signals, by using three non-zero weights. For example, setting $w_L = w_U = \rho, w_{LU} = -\rho^2$ corresponds to the 2-D separable first-order Markov model that is extensively used in image processing. A two-dimensional DPCM coding system with $w_L = w_U = 0.95, w_{LU} = -0.95$ is noted to be effective in [17]. One part of the prediction scheme proposed in [15] is equivalent to using $w_U = w_L = 1, w_{LU} = -1$.

2.3 DPCM weight update

With the prediction model described above, we wish to estimate and update the weights w_L, w_U, w_{LU} from coded data so that the MSE: $\frac{1}{N} \sum (X - \hat{X})^2$ is as small as possible, where N is the number of pixels in the summation. The update rules we use in this paper are designed based on two observations. First, the DPCM weights are mode-dependent. In other words, different weights should be used for residuals of different intra prediction modes. The mode-dependent characteristics of intra prediction residuals have been reported in previous research, for

both lossless and lossy coding [8][13][18][19]. Second, minimizing the prediction error of a set of coded pixels without explicit transmission of weights is simple.

We consider a set of coded pixels, each of them denoted as X_i and their neighbors L_i, U_i, LU_i , similar to the ones shown in Figure 2. We are interested in minimizing the MSE of coded pixels with respect to the DPCM weights.

$$\min_{w_L, w_U, w_{LU}} \sum_i (X_i - (w_L \times L_i + w_U \times U_i + w_{LU} \times LU_i))^2,$$

where X_i, L_i, U_i and LU_i represent previously coded pixels. Ignoring the rounding error for simplicity, the above problem is a linear regression problem and has been well studied. The weights that result in the smallest MSE can be obtained by solving the following linear equation.

$$\begin{bmatrix} \sum LU_i^2 & \sum LU_i \times U_i & \sum LU_i \times L_i \\ \sum LU_i \times U_i & \sum U_i^2 & \sum L_i \times U_i \\ \sum LU_i \times L_i & \sum L_i \times U_i & \sum L_i^2 \end{bmatrix} \begin{bmatrix} w_{LU} \\ w_U \\ w_L \end{bmatrix} = \begin{bmatrix} \sum X_i \times LU_i \\ \sum X_i \times U_i \\ \sum X_i \times L_i \end{bmatrix} \quad (1)$$

The weights that minimize the MSE of coded pixels can be used to encode the current pixels.

In the update rule that we use, we choose a set of coded pixels and group these pixels according to their chosen prediction mode. For each mode, we estimate the best weights that minimize the MSE of these coded pixels of the same mode. The estimated weights are used as the DPCM weights for current pixels.

2.4 Performance analysis

The performance of the proposed system depends on the choice of pixels used for weight estimation and the frequency of weight estimation. To adapt to the local characteristics, we need to use the local region for weight estimation. To obtain a robust set of weights, however, we need a large number of coded pixels. Based on these two considerations, we used a 128x128 pixel moving average for weight estimation in the experiments we discuss in the next section.

Ideally, the weights can be computed and adapted on a pixel-by-pixel basis. This increases the computations substantially both at the encoder and decoder. As a compromise between a higher performance and an increase in computational complexity by a more frequent weight update, we update the weights on a Largest Coding Unit (LCU)-by-LCU basis in the experiments we discuss in the next section.

3. IMPLEMENTATION AND EXPERIMENTAL RESULTS

3.1 Implementation on the HEVC

We implement the ARDPCM on the HEVC reference software 12.0. For the encoder side, the ARDPCM is included in the luminance residual coding with the model proposed in Section 2.2. Pixels outside block boundaries are zero-padded. The decoder side is adjusted accordingly.

The prediction weights are initialized to the same values both at the encoder and decoder. Specifically, we use $w_L = 1, w_U = w_{LU} = 0$ for the horizontal mode, $w_U = 1, w_L = w_{LU} = 0$ for the vertical mode and all zero for other modes. The prediction weights are updated according to the same rule at the encoder and decoder. They are computed and updated from the coded data in a 128x128 moving window for each LCU to be coded.

For weight estimation implementation, we use floating point numbers for the prediction weights. The equation for weight estimation is solved using direct matrix inversion. In the case where the matrix is singular, the prediction weights are set to initial values. This floating point implementation is used for simplicity, and can be modified to integer-based implementation for better numerical stability.

Table 1. List of test sequences

1280x720 (720p) sequences		
Crew	Cyclists	Night
832x480 (WVGA) sequences		
PartyScene	BQMall	RaceHorses
352x288 (CIF) sequences		
Container	MotherDaughter	Foreman
176x144 (QCIF) sequences		
Highway	Mobile	Carphone

Table 2. Average coding gain of the ARDPCM

Resolution	Relative to HEVC	Relative to HEVC with H/V DPCM
1280x720	12.9 %	6.2 %
832x480	3.8 %	1.3 %
352x288	6.8 %	3.0 %
176x144	4.0 %	1.0 %

3.2 Experimental setup and results on HEVC intra lossless coding

We encode a set of intra frames of videos with different resolutions, listed in Table 1. We compare the performance of three systems. The first system is the HEVC system. The second system is the HEVC system with the vertical and horizontal DPCM that is proposed in [12]. The third system is HEVC with the ARDPCM. All experiments are conducted under the HEVC 12.0 intra main profile, with the lossless mode turned on. The average coding gain is shown in Table 2. We observe that HEVC with the ARDPCM outperforms HEVC and the system proposed in [12].

3.3 Effectiveness of the two-dimensional prediction model

To further understand the effectiveness of the proposed prediction model, we investigate the DPCM weights estimated in the system. Figure 3 shows the estimated weights for the vertical mode when the first intra frame of "crew_1280_720" sequence is coded. We plot w_L, w_U, w_{LU} as a function of the LCU index.

From this figure, we observe that all three weights are non-zero. This implies that the prediction error can be reduced by allowing two-dimensional prediction relative to directional prediction. We also notice that w_U in the vertical mode is the most significant weight in most cases. This is consistent with [8], in that the most significant correlation in an intra prediction residual is along the intra prediction direction. We note that w_L is smaller than w_U but still significantly larger than zero. This indicates a weaker correlation along the direction orthogonal to the intra prediction direction. The weaker correlation has been ignored in directional DPCM proposed in [8] and [13]. For other modes, similar weight patterns are observed. Specifically, symmetrical results were obtained for the horizontal mode. For the DC and Planar modes, the weights do not tend to prefer a specific direction and they are much smaller.

3.4 Effectiveness of the adaptive strategy

In order to verify that the proposed adaptive strategy improves the coding performance, we compare the performance of the adaptive DPCM weights with some reasonable sets of fixed DPCM weights. To be specific, we show the coding performance of the following prediction weights:

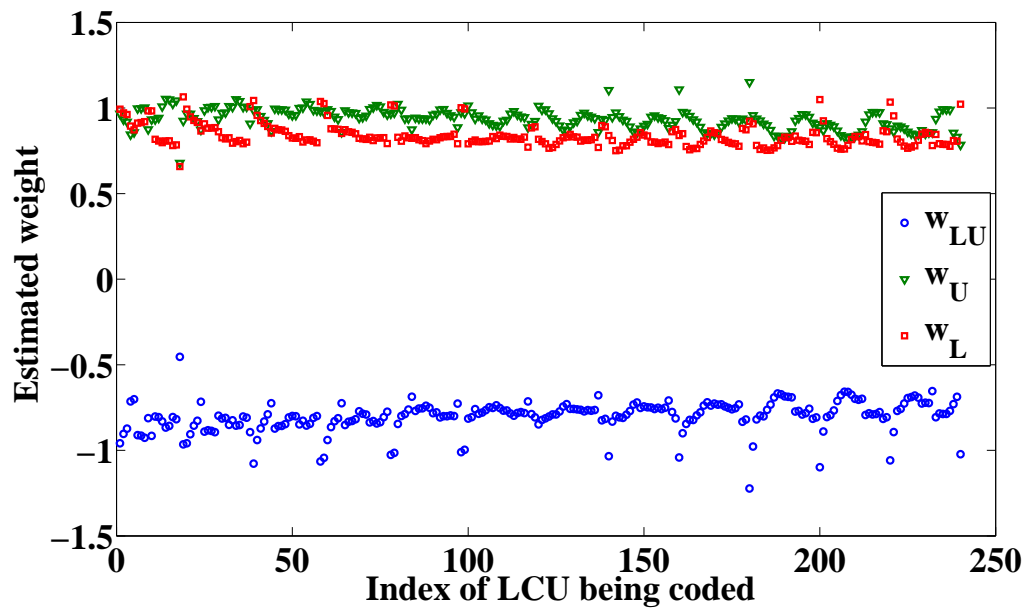


Figure 3. Adaptive weights

- (1) HEVC lossless coding.
- (2) Directional DPCM proposed in [12].
- (3) 2D DPCM with fixed weights $w_L = 1, w_U = 1, w_{LU} = -1$.
- (4) Proposed adaptive weights.

The results are shown in Figure 4. We observe that the proposed adaptive strategy consistently results in significantly better performance compared to the fixed weights. In addition, we note that for the sequences that we tested, the estimated weights change significantly within a frame, and are different for frames particularly with different resolutions. We also note that the 2D fixed weights used in the experiments are not optimized. Better coding performance is possible by choosing a more optimized set of 2D fixed weights. These fixed weights may be obtained, for example, by minimizing the MSE of a training set of sequences. However, we observed that the proposed adaptive strategy always performed better than using the 2D DPCM with any fixed weights in the experiments that we performed.

3.5 Additional comments

The computational complexity of the encoder and the decoder increases due to the use of the ARDPCM and the weight estimation process. Compared to HEVC lossless coding, up to three extra multiplications and three extra additions per pixel are used in the ARDPCM. The weight estimation and update processes are carried out only once for each LCU. These operations are computationally inexpensive in comparison with the current HEVC reference software. As a result, no significant encoder/decoder computational complexity increase is observed in our experiments.

The proposed method can potentially reduce the number of modes in use. In our experiments, we observed that the proposed method is still effective when the ARDPCM is applied only to four intra modes (V,H,DC,Planar) and the residual PCM is applied to other modes. In addition, we observed that the ARDPCM is effective even when other 31 directional intra modes are disabled. In both cases, a very small drop of coding gain is observed. One explanation is that the adaptive weights may help compensate for prediction error increase resulting from using fewer modes.

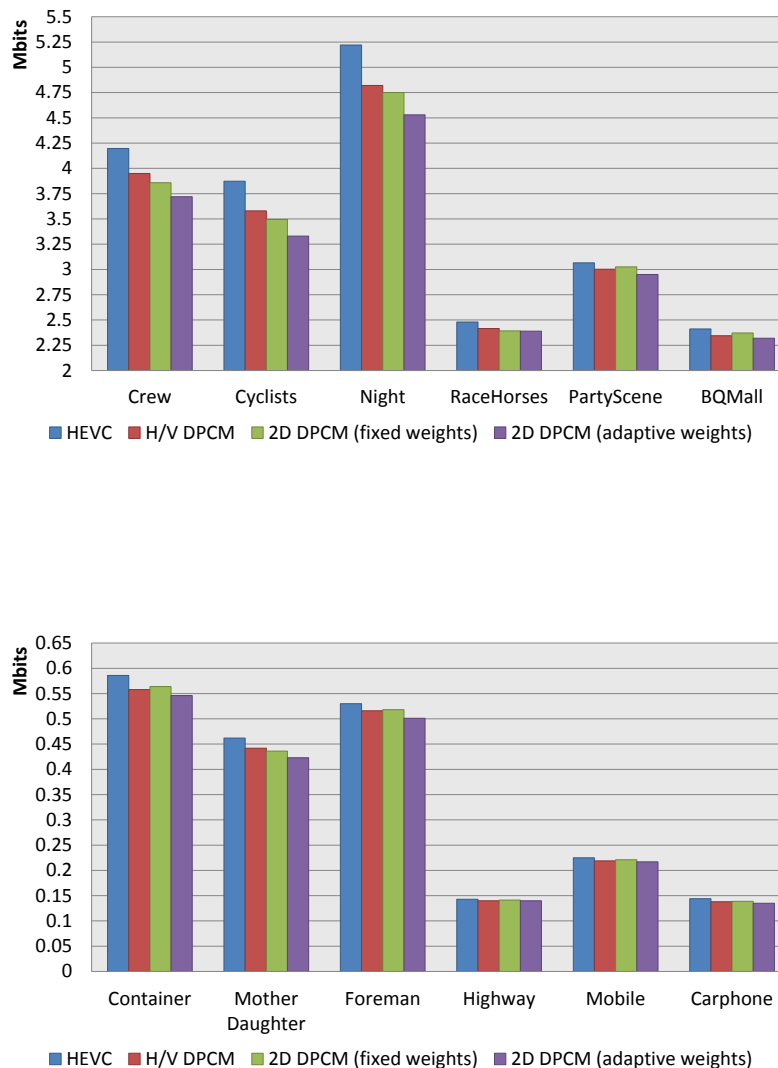


Figure 4. Performance of adaptive and fixed weights

Finally, the ARDPCM does not transmit any extra side information. Therefore, it is HEVC bitstream compliant.

4. CONCLUSIONS

In this paper, we propose the ARDPCM for intra lossless coding. The performance of the proposed system significantly increases due to a two-dimensional prediction model and adaptive prediction weights. The ARDPCM could be used in other applications. For example, the ARDPCM for motion-compensated residuals can be designed as well, by considering the characteristics of inter-prediction residuals.

REFERENCES

- [1] Ahmed, N., Natarajan, T., and Rao, K. R., "Discrete cosine transform," *Computers, IEEE Transactions on* **100**(1), 90–93 (1974).
- [2] Skodras, A., Christopoulos, C., and Ebrahimi, T., "The JPEG 2000 still image compression standard," *Signal Processing Magazine, IEEE* **18**(5), 36–58 (2001).
- [3] Bankoski, J., Bultje, R. S., Grange, A., Gu, Q., Han, J., Koleszar, J., Mukherjee, D., Wilkins, P., and Xu, Y., "Towards a next generation open-source video codec," *IS&T/SPIE Electronic Imaging* (2013).
- [4] Ahmed, N. and Rao, K. R., "Walsh-hadamard transform," in [*Orthogonal Transforms for Digital Signal Processing*], 99–152, Springer (1975).
- [5] Ding, J.-R., Chen, J.-Y., Yang, F.-C., and Yang, J.-F., "Two-layer and adaptive entropy coding algorithms for H.264-based lossless image coding," *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 1369–1372 (2008).
- [6] Cai, X. and Gu, Q., "Improved HEVC lossless compression using Two-Stage coding with Sub-Frame level optimal quantization values," *IEEE International Conference on Image Processing 2014 (ICIP 2014)* (Oct. 2014).
- [7] Sullivan, G., Ohm, J.-R., Han, W.-J., and Wiegand, T., "Overview of the High Efficiency Video Coding (HEVC) standard," *Circuits and Systems for Video Technology, IEEE Transactions on Vol. 22 Iss. 12* (2012).
- [8] Lee, Y.-L., Han, K.-H., and Sullivan, G. J., "Improved lossless intra coding for H.264/MPEG-4 AVC," *Image Processing, IEEE Transactions on* **15**(9), 2610–2615 (2006).
- [9] Wiegand, T., "Draft ITU-T recommendation and final draft international standard of joint video specification," *ITU-T rec. H. 264—ISO/IEC 14496-10 AVC* (2003).
- [10] Wiegand, T., Sullivan, G. J., Bjontegaard, G., and Luthra, A., "Overview of the H.264/AVC video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on* **13**(7), 560–576 (2003).
- [11] Sullivan, G. J., Topiwala, P. N., and Luthra, A., "The H.264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions," *Optical Science and Technology, the SPIE 49th Annual Meeting*, 454–474 (2004).
- [12] Lee, S., Kim, I.-k., and Chanyul, K., "Residual DPCM for HEVC lossless coding," *JCTVC-L0117, Geneva, Switzerland, 14-23 January 2013*.
- [13] Zhou, M., Gao, W., Jiang, M., and Yu, H., "HEVC lossless coding and improvements," *Circuits and Systems for Video Technology, IEEE Transactions on Vol. 22 Iss. 12* (2012).
- [14] Amon, P., Hutter, A., Wige, E., and Kaup, A., "Intra prediction for lossless coding," *JCTVC-L0161, Geneva, Switzerland, 14-23 January 2013*.
- [15] Tan, Y. H., Chuohao, Y., and Zhengguo, L., "Lossless coding with residual sample-based prediction," *JCTVC-K0157, Shanghai, China, 10-19 October 2012*.
- [16] Gao, W., Jiang, M., and Yu, H., "On lossless coding for HEVC," *IS&T/SPIE Electronic Imaging* (2013).
- [17] Lim, J. S., "Two-dimensional signal and image processing," *Englewood Cliffs, NJ, Prentice Hall, 1990*, 710 p. **1** (1990).
- [18] Ye, Y. and Karczewicz, M., "Improved H.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning," *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2116–2119 (2008).
- [19] Han, J., Saxena, A., Melkote, V., and Rose, K., "Jointly optimized spatial prediction and block transform for video and image coding," *Image Processing, IEEE Transactions on* **21**(4), 1874–1884 (2012).