Encoding mode selection in HEVC with the use of noise reduction

Olgierd Stankiewicz [1], Krzysztof Wegner [1], Damian Karwowski [1]
Jakub Stankowski [1], Krzysztof Klimaszewski [1], Tomasz Grajek [1]

[1] Chair of Multimedia Telecommunications and Microelectronics,
Poznan University of Technology,
Polanka 3, 60-965 Poznan, Poland

ostank@multimedia.edu.pl

Abstract – This paper concerns optimization of encoding in HEVC. A novel method is proposed in which encoding modes, e.g. coding block structure, prediction types and motion vectors, are selected basing on noise-reduced version of the input sequence, while the content, e.g. transform coefficients, are coded basing on the unaltered input sequence. Although the proposed scheme involves encoding of two versions of the input sequence, the proposed realization ensures that the complexity is only negligibly larger than complexity of a single encoder. The proposal has been implemented and assessed. The experimental results show that the proposal provides up to 1.5% bitrate reduction while preserving the same video quality.

Keywords – **HEVC encoding, encoder control, mode selection optimization; noise reduction**

I. Introduction

The progress in the development of compression techniques allows for more and more efficient representation of video. The latest video coding technology described in ISO/IEC and ITU standards is HEVC [1]. Relative to its predecessor (AVC [2]) HEVC brings about 50% of bitrate reduction while preserving the same video quality [3]. Such a promising efficiency gain was possible due to significant development of compression tools available to the encoder. For example, it can be calculated that for the largest coding unit (LCU), a HEVC encoder has a choice of about $48 \cdot 10^{18}$ different CU, TU and PU partition combinations, 33 intra prediction directions, and two reference lists with varying pictures in inter prediction [4]. Therefore, the selection of the encoding modes is even more complex than in previous techniques e.g. AVC. Among all the choices available, the encoder is looking for an optimal set of modes that maximize rate-distortion ratio. Brute-force check of all of the opportunities in the whole image is practically impossible, and therefore the most commonly, the mode selection is optimized only locally at the level of one LCU. For an example, such an approach is employed in HEVC software model (HM) of the JCT-VC group [5]. For the new HEVC video encoder the LCU-level optimization of coding modes is currently the subject of numerous research works (which results can be found for example in [11-28]), however in those works different scenario of video encoding is assumed in comparison to that established in this work.

The broadest application of video compression standards is encoding of natural video sequences, registered with real cameras. Such sequences can be characterized by a presence of noise, which inevitably is produced in electrical circuits of the cameras. From the perspective of the encoder, the presence of noise in the images leads to selection of modes (e.g. prediction modes and transform coefficients) optimized for representation of the content with the noise. Therefore, the mode selection optimization is disturbed by presence of unwanted noise components. Moreover, the encoded and reconstructed block of the image with the noise is not the best prediction source for the next block of the picture, because the noise in successive parts of the image is poorly correlated. Such a phenomena is a problem both in inter and intra predictions. At the same time, HEVC encoder has such sophisticated prediction capabilities that usually the prediction error is practically sole noise (noise is random in nature and cannot not be predicted based on the image content). In practice this means that the encoder often unnecessarily loses bits for representation of noise components.

A commonly used approach to solve this problem [6] is to reduce noise in the sequence before the encoding. In such a case the encoder makes decisions about modes based on the sole content without the noise, however the information about the noise is not encoded at all. This is not a good approach, because noise is perceptually important, and its absence causes the recipients to sense unnaturalness in the reconstructed image. In addition, the noise reduction is often imperfect and can lead to removal of important components of the image, which are not noise, e.g. small details like sea waves.

In the literature there are no known techniques that would use the denoised version of the sequence for the mode selection, while preserving the information about the noise.

II. The proposed idea

This paper presents a novel hybrid approach for mode selection in video encoding. Its aim is to avoid the problem of direct encoding of noise-reduced sequences, which leads to loss of information about the noise, while ensuring better selection modes unaffected by the occurrence of noise.

The main idea of this proposal is to perform selection of the modes in the encoder basing on a noise-reduced version of the sequence (denoised) but use the original, unaffected version of the sequence (containing noise) for production of the output bitstream.

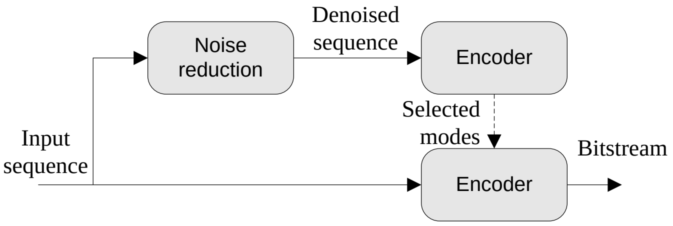

Figure 1. The general scheme of the proposed method of coding modes selection in the encoder.

Implementation of the proposed method (Fig. 1) uses two modified encoders working on two versions of the sequence being encoded, the original input one, and denoised one.

The first encoder works with the denoised version of the sequence and it consists in all encoder functionality with the exception of entropy coding module. During the encoding, it searches and optimizes coding modes, which are then fed to the second encoder. The encoding modes, as mentioned in the introduction, are optimized locally at the level of LCU units, but the decisions are not influenced by the presence of noise which has been removed.

The second encoder works on unaffected input sequence. This encoder is very simple as it does not perform the expensive optimization of the encoding modes. Instead it uses exactly the encoding modes provided by the first encoder. These modes are used to generate prediction signal for each encoded block and to generate transform coefficients. All of the generated syntax components are entropy encoded and outputted in the final bitstream.

Therefore, the bitstream produced as a result of the abovementioned process represents the original input sequence (with noise) but with the encoding modes selected basing on the noise-reduced version of the sequence.

III. IMPLEMENTATION

In order to enable the validation and assessment of the performance, the proposed method was implemented by making modifications to the HM software model [5] (Version 13.0) for HEVC technique developed by JCT-VC group.

Direct realization of the proposed method (Fig. 1) would lead to the use of two encoders, both used to encode a single sequence. Such a redundancy is not required which is discussed below.

As can be easily noticed the first encoder (working on the noise-reduced version of the sequence) does not generate any binary stream, so in that case entropy encoding module is not needed. In HEVC compression, entropy encoding is performed with the use of CABAC technique (Context-Adaptive Binary Arithmetic Coding). CABAC is very complex and it constitutes a substantial portion of runtime of the encoder [7].

On the other hand, the second encoder (working with the unaffected original sequence) lacks the encoding mode optimization, since modes are provided from the first encoder. The encoding mode optimization, which among others include motion vector search, in the terms of computational complexity is practically a second substantial part of the encoding, complementary to the entropy coding.

Therefore, it can be said that the total complexity of both encoders used in the proposal is similar to the complexity of a single, full encoder.

To accomplish the reduction of noise, in the work we have used existing software package for video processing called mv-tools [8]. It allows noise reduction of video sequences in near real time. Although there are many other known noise reduction techniques, we have found out that even such a simple technique is sufficient to achieve good results.

IV. EXPERIMENTAL RESULTS

The experimental verification of the proposed technique was carried out using a set of test sequences developed by MPEG/JCT-VC groups during their works on HEVC technology standard [9]. These sequences are grouped in six classes (A, B, C, D, E, F) with varying resolution and frame-rate (Table I).

TABLE I. TEST SEQUENCES USED IN EXPERIMENTATION

Class	Number of sequences	Resolution	Frame rates
A	3	2560×1600	30; 60
B	5	1920×1080	24; 50
C	4	832×480	30; 50; 60
D	4	416×240	30; 50; 60
E	3	1280×720	60
F	4	1024×768	20; 30

Example of single pictures from the test sequences are presented in Figs. 2a and 3a. In Fig. 2bc and 3bc we show fragments of these exemplary pictures, enlarged and with enhanced contrast, in order to show comparison between the original version of the encoded sequence (b) and denoised one.

The experiments were performed according to the scheme shown in Fig. 4. Each of the sequence was encoded with various values of quantization parameter index QP=[22,27,32,37,42,47]. The encoding was performed in three variants: α, β and γ, as described below.

In the first variant (γ), a single HEVC encoder (unmodified HM software in version 13.0) encodes the input unaltered sequence. This variant is a reference for the other results.

In variant (α), the input sequence is fed to noise-reduction block using technique known from mv-tools [8]. The results are then encoded with unmodified HM 13.0. This version reflects the state of the art known from the literature.

Variant (β) is the proposal, which uses encoding modes optimized basing on the denoised sequence. Two encoders are used. The first one selects encoding modes basing on noise-reduced version of the input sequence. The second one uses encoding modes selected by the first encoder to compress the original, unaltered input sequence. Both encoders implement the approach presented in Section III and are based on HM 13.0 software. It is worth mentioning that thanks to the approach mentioned in Section III, the complexity of the both encoders is practically the same as the complexity of a single HM 13.0 encoder, as it is shown in Fig. 5.

a)

b)

c)

Figure 2. Sample picture from the “BQTerrace” sequence (a), an enlarged and contrast-enhanced fragment (b) and its noise-reduced version (c).

Figure 3.

After the encoding in the three mentioned variants, the produced bitstreams are then decoded and compared to the original, unaltered input sequence (with noise) with the use of objective PSNR metric (Fig. 4).

Some of the obtained results, in the form of a rate-distortion (r-d) curves are illustrated in Fig. 6abc, for an exemplary sequence from class B (BQTerrace), the average of sequences in class B, and for the average for all concerned test sequences, respectively.

Basing on the attained r-d curves, we have measured how much bitrate reduction can be attained with the use of the two latter variants (α) and (β), as compared to the reference (γ). For this purpose we have used Bjøntegaard metric [10]. In particular, Δ(BDRATE) was calculated over quantization parameter index range QP 27...42. The results are presented in Table II and visualized in Fig. 7. The negative Δ(BDRATE) values indicate a reduction in bitrate , related to the reference (γ).

a)

b)

c)

Figure 4. Sample picture from “SteamLocomotive” sequence (a), an enlarged and contrast-enhanced fragment (b) and its noise-reduced version (c).

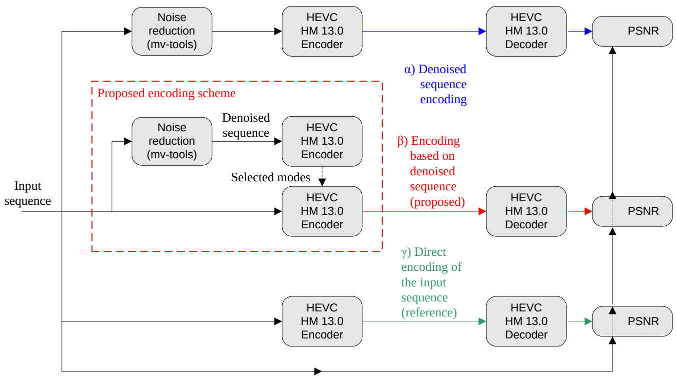

Figure 5. Scheme of the performed experiments

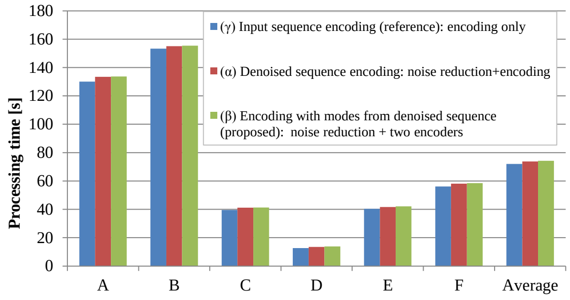

Figure 6. Complexity of the three considered video encoding schemes.

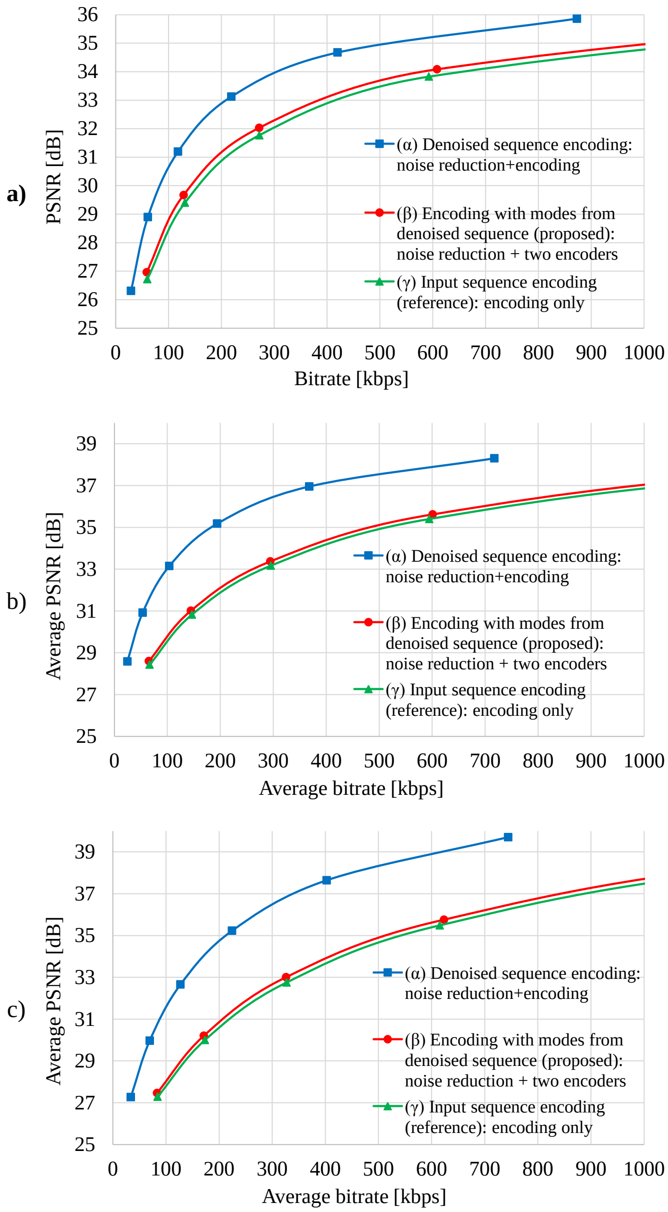

Figure 7. Rate-distortion curves attained for comparison of three tested encoding variants: α , β and γ:
a) for an exemplary BQTerrace sequence from class B,
b) averaged over all sequences from class B,
c) averaged over all the test sequences.

As it can be seen from Table II and Fig. 7, the results are consistent for nearly all of classes. In general, the resulting reduction of the bitrate of the proposed method is about 1.5%.

Of course, the encoding of the noise-reduced version of the test sequence, without any noise information encoding, known from the literature, brings a higher bitrate savings up to an average of about 60%. This indicates that the encoding modes selected based on the noise-reduced version of the sequences are closer to the optimal choice, but still does not allow to bitrate reduction comparable to the case in which the noise is totally omitted.

TABLE II. ATTAINED BITRATE REDUCTIONS COMPARED TO THE REFERENCE, PRESENTED AS BJØNTEGAARD DELTAS Δ(BD RATE)[10]. NEGATIVE Δ(BD RATE) VALUES INDICATE BITRATE REDUCTIONS.

Class	Sequence	Δ(BD-RATE)	
		Proposed (β) vs reference (γ)	**Denoised (α) vs reference (γ)**
A	Traffic	-1.77%	-58.93%
A	PeopleOnStreet	-1.44%	-64.21%
A	SteamLocomotive	-1.31%	-20.62%
	Average – class A	-1.51%	-47.92%
B	Kimono1	-1.01%	-62.80%
B	ParkScene	-1.50%	-67.47%
B	Cactus	-1.80%	-61.02%
B	BQTerrace	-2.40%	-49.08%
B	BasketballDrive	-1.74%	-78.37%
	Average – class B	**-1.69%**	**-63.75%**
C	RaceHorses	-1.44%	-63.09%
C	BQMall	-1.43%	-67.04%
C	PartyScene	-1.61%	-49.09%
C	BasketballDrill	-1.88%	-63.37%
	Average – class C	**-1.59%**	**-60.65%**
D	RaceHorsesLow	-0.30%	-58.30%
D	BQSquare	-1.88%	-56.77%
D	BlowingBubbles	-1.70%	-62.31%
D	BasketballPass	-1.21%	-72.50%
	Average – class D	**-1.27%**	**-62.47%**
E	FourPeople	-1.88%	-56.35%
E	Johnny	-2.50%	-56.08%
E	KristenAndSara	-2.03%	-55.95%
	Average – class E	**-2.14%**	**-56.13%**
F	BasketballDrillText	-1.80%	-64.51%
F	ChinaSpeed	-1.22%	-63.58%
F	SlideEditing	-0.35%	-49.53%
F	SlideShow	-0.98%	-71.00%
	Average – class F	**-1.09%**	**-62.16%**
	Average – all	**-1.53%**	**-59.65%**

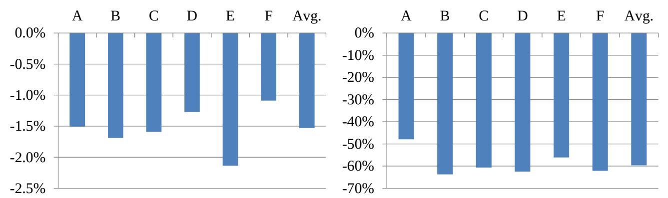

Figure 8. Summary of Bjøntegaard deltas Δ(BD RATE)[10]. Negative Δ(BD RATE) values indicate bitrate reductions.

V. Conclusions

The paper presents a novel method for encoding modes selection, on the example of HEVC coding technology. In the proposed method, the encoding modes are selected in the second encoder, which in parallel encodes noise-reduced version of the input sequence. Therefore, the selected coding modes are optimized for the content without the influence of the noise. The proposal was implemented with the use of HM (version 13.0) software, which is a HEVC model software created by MPEG/JCT group. This allowed experimental evaluation of the proposal. A standard MPEG test sequence set, widely used in the literature, was used for the assessment. The results show that the proposed method enables up to 1.5% reduction in the bit stream. This indicates that the encoding modes selected based on the noise-reduced version of the sequences are closer to the optimal choice, but still does not allow to bitrate reduction comparable to the case in which the noise is totally omitted. This means that the gain coming from directly compressing the noise-reduced version of the sequences, without the noise, comes mainly from omitting transform coefficients related with noise components in the encoded signal and to a lesser extent, with a better choice of encoding modes.

Acknowledgment

Research project was supported by The National Centre for Research and Development, POLAND, Grant no. LIDER/023/541/L-4/12/NCBR/2013.

References

[1] G.J. Sullivan, J.-R. Ohm, W.J. Han, T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, 2012.

[2] T. Wiegand, G.J. Sullivan, G. Bjontegaard, A. Luthra, "Overview of the H.264/AVC video coding standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560-576, 2003.

[3] J.-R. Ohm, G.J. Sullivan, H. Schwarz, TK Tan, T. Wiegand, "Comparison of the coding efficiency of video coding standards - Including High Efficiency Video Coding (HEVC)", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1669-1684, 2012.

[4] F. Bossen, B. Bross, K. Suhring, D. Flynn, "HEVC Complexity Analysis and Implementation", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1685-1696, 2012.

[5] "HM 15.0 Reference Software". 2014. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO / IEC JTC 1 /SC29/WG11, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/ tags / HM-13.0.

[6] B.C. Song, K.W. Chun, “Motion-compensated temporal prefiltering for noise reduction in a video encoder”, 2004 International Conference on Image Processing, ICIP '04, pp. 1221-1224, 2004.

[7] V. Sze, M. Budagavi, " High Throughput CABAC Entropy Coding in HEVC", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1778-1791, 2012.

[8] M. Fizick, A. Balakhnin, T. Schniede "Mv-tools web-page" http://avisynth.org.ru/mvtools/mvtools.html - online 1 Dec 2013.

[9] F. Bossen. "Common test conditions and reference software configurations", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO / IEC JTC1 / SC29 / WG11, Doc. JCTVC-J1100, 2012.

[10] G. Bjøntegaard, "Calculation of average PSNR differences between RD curves", ITU-T SG16 / Q6, Doc. VCEG-M33, 2001.

[11] J. Siast, J. Stankowski, M. Domański, Hierarchical fast selection of intraframe prediction mode in HEVC, International Journal of Electronics and Telecommunications, Vol. 62, No. 2, Electronics and Telecommunications Committee of Polish Academy of Sciences, Warsaw, 2016, pp. 147-151, ISSN: 2300-1933. (www.multimedia.edu.pl)

[12] D. Palomino, E. Cavichioli, A. Susin, L. Agostini, M. Shafique, J. Henkel, “Fast HEVC Intra Mode Decision Algorithm Based on New Evaluation Order in the Coding Tree Block”, Picture Coding Symposium PCS 2013.

[13] T. L. Silva, L. A. Silva Cruz, L. V. Agostini, “HEVC Intra Mode Decision Acceleration Based on Tree Depth Levels Relationship”, Picture Coding Symposium PCS 2013.

[14] W. Jiang, H. Ma, Y. Chen, “Gradient Based Fast Mode Decision Algorithm for Intra Prediction in HEVC”, in Proceedings of 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet), April 2012, pp. 1836-1840.

[15] H. Zhang, Q. Liu, Z. Ma, “Priority Classification Based Fast Intra Mode Decision for High Efficiency Video Coding”, Picture Coding Symposium 2013.

[16] G. Tian, S. Goto, “Content Adaptive Prediction Unit Size Decision Algorithm for HEVC Intra Coding”, in Proceedings of Picture Coding Symposium (PCS), May 2012, pp. 405-408.

[17] Wenjun Zhao, Takao Onoye, Tian Song, Hierarchical Structure based Fast Mode Decision for H.265/HEVC, IEEE Transactions on Circuit and Systems for Video Technology, Volume: PP, Issue: 99, January 2015.

[18] Dongxu Xu, Qiwei Lin, Xiaohui Dong, A novel fast intra-CU size decision algorithm for HEVC, Journal of Electronic Imaging, Volume 23, Issue 1, January 2014.

[19] L. Zhao, L. Zhang, S. Ma, and D. Zhao, “Fast mode decision algorithm for intra prediction in HEVC”, IEEE Visual Communications and Image Processing (VCIP) 2011, pp. 1-4.

[20] Liquan Shen, Zhaoyang Zhang, Zhi Liu, “Effective CU Size Decision for HEVC Intracoding”, IEEE Transactions on Image Processing, vol. 23, no. 10, pp. 4232–4241, October 2014.

[21] Biao Min, Ray C. C. Cheung, “A Fast CU Size Decision Algorithm for the HEVC Intra Encoder”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, no. 5, pp. 892–896, May 2015.

[22] Jung, S.; Park, H., "A Fast Mode Decision Method in HEVC using Adaptive Ordering of Modes", IEEE Transactions on Circuits and Systems for Video Technology, Year: 2015, Volume: PP, Issue: 99 Pages: 1 - 1, DOI: 10.1109/TCSVT.2015.2473303, IEEE Early Access.

[23] K. Lim, J. Lee, S. Kim, S. Lee, Fast PU Skip and Split Termination Algorithm for HEVC Intra Prediction, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 25, No. 8, August 2015, pp. 1335-1346.

[24] T. L. da Silva, L. V. Agostini, L. A. da Silva Cruz, Fast HEVC intra prediction mode decision based on EDGE directional information, in Proceedings of 20th European Signal Processing Conference (EUSIPCO), August 2012, pp. 1214-1218.

[25] H. Sun, D. Zhou, S. Goto, A low-complexity HEVC intra prediction algorithm based on level and mode filtering, in Proceedings of IEEE International Conference on Multimedia and Expo (ICME), July 2012, pp. 1085-1090.

[26] L. Shen, Z. Zhang, P. An, Fast CU size decision and mode decision algorithm for HEVC intra coding, IEEE Transactions on Consumer Electronics, Vol. 59, No. 1, February 2013, pp. 207-213.

[27] S. Cho and M. Kim, Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 23, No. 9, pp. 1555–1564, Sep. 2013.

[28] Hyang-Mi Yoo; Jae-Won Suh, "Fast coding unit decision algorithm based on inter and intra prediction unit termination for HEVC", IEEE International Conference on Consumer Electronics (ICCE) 2013, vol., no., pp.300,301, 11-14 Jan. 2013.