Analysis of Residual data for High Efficiency Video Coding (HEVC)

Anis BenHajyoussef, Tahar Ezzedine

SYSCOM Laboratory, National School of Engineers of Tunis, Tunisia

Summary
In this paper, we propose an analysis of the residual information of the HEVC. We focus specifically on analyzing the coding of the DC coefficient and its effect on bit consumption. We investigate on the possibilities and the limits of coding the DC coefficient, in a lossless coding way, by changing the process to the encoding a spatial translation value or the residual blocks instead of the transform/quantization process. Experimental results show that 8 bits can be approximated as the limit of DC bit cost of a lossless coding, under which we may obtain a gain in the BD-rate.

Key words:
HEVC, video compression, residual, distortion. DC coefficients

1. Introduction

The demand of multimedia services is growing especially after the emergence of the H.264/AVC, the current video standard developed by the Joint Video Team (JVT). In fact the AVC has achieved a significant improvement in video coding efficiency allowing an expansion on digital applications like video telephony, storage, streaming and broadcasting. Moreover, the growth of applications dealing with the high and ultra-high definitions is increasing the customers need for more video compression and video quality.

Actually, ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) have been exploring on more video efficiency since 2004. And in 2010, they have established the Joint Collaborative Team on Video Coding (JCT-VC) [1] to develop the new video compression standard, the High Efficiency Video Coding (HEVC) [2], aiming to deliver the final specification by January 2013.

So far, many features have been proposed and many improvements and refinements are being achieved in JCT-VC meetings, on HEVC Test Model (HM)[3][8]. HM keeps however the traditional hybrid coding diagram with the main coding tools compared to AVC such prediction, residual coding, transform and quantization.

In addition to the focus on video quality improvement, many efforts are being done to simplify the coding tools and providing hardware friendly implementation of the video features, becoming so complex compared to AVC, even at the expense of video quality in some cases.

Manuscript received August 5, 2012
Manuscript revised August 20, 2012

As an example of such effort, the transform process, dealing in HM with blocks of size from 4x4 to 32x32 have been carefully designed in that sense[4]. Among its properties, the norms of the basis vectors are sufficiently close so that the same scaling can be used for all transform coefficients during quantization/dequantization. The coefficients and basis vectors of the smaller transforms are a subset of coefficients and basis vectors of larger transforms to allow the re-use of logic for multiply/adds in hardware design. Matrices also present the same symmetry properties as the DCT transform.

In this paper, we focus on two properties of the transform process that deals with the residual signal.

• One distinguishing property of the current HM compared to H.264/AVC standard is the remove of the hierarchical transform. Actually, the AVC uses the Hadamard transform for coding the DC coefficients [5] in addition to the first step DCT transform. The idea behind that is that, especially for blocks with mostly flat pixel value, the correlation among the DC coefficient becomes important, so an additional transform level permits to code that correlation efficiently.

• The DC coefficients have in most cases, important values compared to the AC values. The DC coefficients are coded through the processes of transform and quantization, which means that these values are coded with a loss in video quality.

We propose to investigate on better coding of the residual information and we focus specifically on this paper on investigating on how possible to code the DC information lossless.

This paper is organized as follows. The first section, propose an analysis of the residual information in video coding. The second section, investigate on how possible to code the DC coefficient in lossless manner and effects of such approach. The experimental evaluations are shown in the section 3, followed by the conclusion in the section 4.

2. Analysis of residual information

In video compression, the residual represents the texture information of the signal to be coded. It is resulting from the difference of the original and predicted signals. The more the prediction is good, the small residual we have and the less information we have to code. In the case of intra prediction, the residual is so important because the spatial prediction is limited to predict the correlation compared to inter case. We focus on this work on Intra case.

Our analysis starts from the notice that intra residual blocks values have a mean value that is in most case so different to zero, which is the target value of the prediction.

We present in figure 1 the distribution of the mean values of residual taken from an I frame. It shows that the most important values are concentrated in the interval [-100, 100] and the average value of the mean values of residual blocks is 88 for our sample of blocks. We precise here that the values of residual blocks are in a context of Internal bit depth interval (IBDI)of 4 bits in the HM.

In fact, such value depends in fact on prediction mode and essentially on the original signal of the video. Even if intra prediction has been improved, it remains limited facing to the no stationary properties of the real world video sequences.

This mean value of the prediction error will exit spatially in each pixel of the block. But after the transform process of the block, that information we be represented by only one sample which is the DC coefficient.

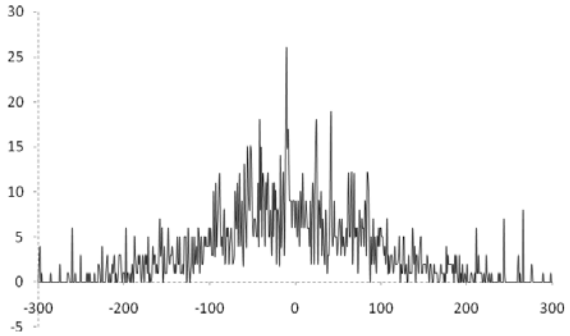
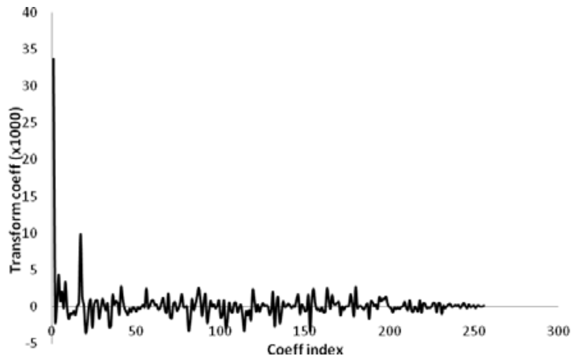


Fig. 1 Distribution of absolute mean of 2000 Intra frame residual blocks from RaceHorses sequence

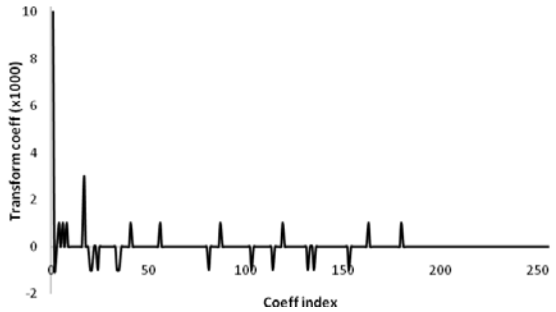
The figure 2 represents the transform and quantization matrices values of a 16x16 TU block. It can be noticed that that DC coefficients has relatively an important value compared to the AC values and that value lead to high quantized coefficient that take more bits to be coded than AC values.

So, since the value of the DC coefficient is proportional to the mean of residual block, that supposes that the more important the mean of residual block is, the more bits we need to code that DC coefficient.

These remarks make worth to investigate on coding the DC information, in other way than the transform and quantization processes. We focus especially in the analysis on effects of coding the DC losslessly on video quality and bits consumption.



(a)



(b)

Fig. 2 Transform (a) and quantization (b) matrices in normal scan order of a 16x16 block, from RaceHorses sequence

2. Coding of residual data

The approach here is to remove the DC component from the residual block by deducing from all the block pixel values the mean value of the block as if we simply translate the bloc spatially from the mean value to zero. We refer, in the rest of the paper, to that value as a *translation value*.

We propose here to code that translation information losslessly TU based and avoiding coding the first coefficient of each TU since it will have automatically a value equal to zero. In this investigation, we simply code the mean value with exponential-Golomb codes, used to code the quantized coefficients [6]. We precise that in this work, we do not concentrate on coding that value efficiently rather than investigate on how possible to code the DC value losslessly and the limits of such approach.

Table 1: Residual coding motivation

	mean Residual	mean Tranf.	mean Quant.
original	100,284	144,593	0,048
translated residual	0,284	133,476	0,029

To help understand the mean residual deep impact on coding, we focus here on TU based information and not the overall impact, not to include the results the interaction between blocks coding such the rate control impact. So we present in the table 1, the average value for respectively residuals, transform and quantization blocks. As we can see the remove of the DC component represent almost the half amount of quantization blocks.

For the implementation of the approach, we consider the High efficiency configuration of HM, and thus the presented syntax element considers the CABAC as entropy coding scheme.

We code the translation value in the bitstream after the coding of the coefficients as shown in figure 3. We use a flag *coded_mean_residual* to indicate if we are coding a translation residual value or not for the current TU. In case, this flag value is equal to 1, we code the mean residual level and the value sign. The level is coded using two coding symbols. The *mean_resid_abs* which used to code the absolute value of the residual mean using Exp-Golomb code and *mean_resid_sign* which one bit symbol with value of 1 for negative coefficients).

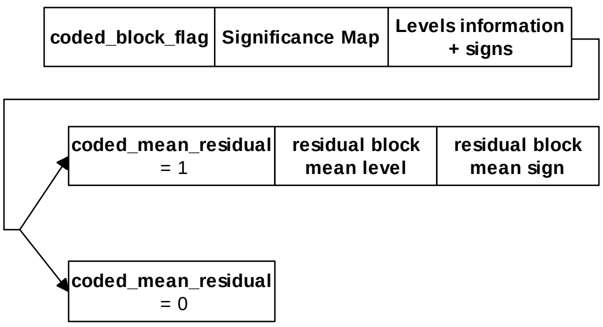


Fig. 3 Syntax element for mean residual value

To evaluate the impact of the approach in term of bits consumption and in term of distortion, we present in the table 2, the amount of bits per pixel used to code transform coefficients, bits need to code the translation residual value and the distortion per pixel, through 3 scenarios; the original configuration, translated residual without coding which is used here as a motivation scenario of coding the translation and translated residual with exponential-Golomb coding.

Table 2: Luma residual coding motivation

Config.	**Bits / pix**		**Distorsion/pix**
	Coeff	**Mean resid.**	
HM1.0	45,60	0	86,53
No coding	37,32	0	83,92
Ex-Glmb coding	38,59	59,84	85,15

From Table 2 we notice a motivation gain given by reducing the DC coefficient, and coding it with Exp-Golomb give an improvement in distortion since the value is coded lossless but that consumes more bits compared to transform/quantization process especially for small blocks like 4x4.

2. Experimental results

We can see so from results that such coding of translated value will be a relatively high consuming approach, we focus so on the next step on exploring the limits of a such lossless coding of the translation value that give an overall gain in BD-rate.

For that we consider 3 simulation scenarios in addition to the two configurations used as anchors. The first consists on translate residual without coding used as a motivation case and the second consists on coding the translation residual value with Exp-Golomb.

In the 3 simulation scenarios, we did not really code the translation value but we force the encoder, instead, to consume an additional 4, 6 and 8 bit per TU. That added bit quantity is supposed to simulate a coding of the translation value that consumes that quantity. We precise here that, 1 bit is used to code the flag *coded_mean_residual* and 1 bit is used to code the sign and the other bit to code the translation value level.

For experimental evaluation, we implemented the proposed approach on the HEVC test model HM2.0. We used the Intra configuration setting with high efficiency (HE) level according to the Common test conditions and software reference configurations [7], incorporating CABAC as entropy coding scheme.

We used 5 sequences representing the 5 classes A, B, C, D and E.

The motivation scenario shows an average -26.2 potential gain of 26.2 in Y BD-rate. And that chroma component also has a potential of gain even if we translated only the luma residual blocks. Such scenario proves, in fact, that is worth to investigate on the possibilities of coding the DC coefficients.

The second configuration shows the limits of Ex-Golomb codes as used in the approach with an average loss of 9.0 in Y BD-rate.

In the simulation scenarios with added 4 and 6 bits gives and respectively an average improvement of 7.7 and 3.5 while adding 8 bits give a negligible loss of 0.2 in Y BD-rate.

We can conclude from these results that 8 bits can be considered as the limit of coding the translation value of the residual blocks, lossless way, under which we obtain an improvement in video coding.

More investigations can deal with possible adaptive translation of residual blocks, more efficient coding ways and translating chroma component.

Table 3: Intra RB-RSNR results

Cfg	Seq/classe	Fr m nbr	Intra BD-rate		
			Y	U	V
no coding	Traffic_2560x1600_30	4	-27,9	-14,5	-12,2
	Kimono1_1920x1080_24	4	-31,6	-18,2	-11,8
	BQMall_832x480_60	50	-19,1	-12,0	-11,3
	RaceHorses_416x240_30	50	-19,0	-12,8	-12,4
	vidyo1_1280x720_60	10	-33,5	-7,8	-4,5
Exp. Glmb	Traffic_2560x1600_30	4	7,3	7,5	8,4
	Kimono1_1920x1080_24	4	4,1	4,7	4,9
	BQMall_832x480_60	50	13,7	12,3	12,3
	RaceHorses_416x240_30	50	12,3	10,8	11,9
	vidyo1_1280x720_60	10	7,4	14,8	13,4
4 bits	Traffic_2560x1600_30	4	-7,5	1,3	3,9
	Kimono1_1920x1080_24	4	-13,8	-5,9	0,4
	BQMall_832x480_60	50	-3,5	1,8	2,4
	RaceHorses_416x240_30	50	-2,9	1,3	1,6
	vidyo1_1280x720_60	10	-10,9	9,3	8,1
6 bits	Traffic_2560x1600_30	4	-3,3	3,4	5,3
	Kimono1_1920x1080_24	4	-10,2	-3,8	1,8
	BQMall_832x480_60	50	0,9	5,0	5,4
	RaceHorses_416x240_30	50	0,9	4,1	4,5
	vidyo1_1280x720_60	10	-5,9	10,0	9,2
8 bits	Traffic_2560x1600_30	4	0,1	5,1	6,6
	Kimono1_1920x1080_24	4	-6,3	-1,5	2,9
	BQMall_832x480_60	50	4,7	7,7	8,0
	RaceHorses_416x240_30	50	4,0	6,2	6,8
	vidyo1_1280x720_60	10	-1,4	11,4	10,1

Conclusions

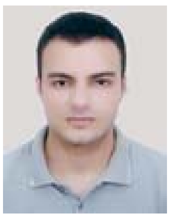
In this paper, we proposed an analysis of the residual information coding on HEVC. We show that the DC coefficient, in most cases, represents almost the half amplitude of the quantization matrices, affecting directly the bit consumption in entropy coding.

As an investigation on the possibilities of coding the DC coefficient, we proposed to analyze the limits of lossless coding way, but transforming problem to the encoding of spatial translation value or the residual blocks instead of the transform/quantization process.

Results show that 8 bits can be approximated as the limit of DC bit cost per TU of a lossless coding, under which we may obtain a gain in the BD-rate.

References

[1] T. Weigand et al, " Special Section on the Joing Call for Proposals on High Efficiency Video Coding (HEVC) Standardization, IEEE Trans. Cir. and Sys. for Video Tech. Nov. 2010.
[2] T. Wiegand, B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, "WD3: Working Draft 3 of High-Efficiency Video Coding", JCTVC-E602, Geneva, CH, March 2011.
[3] K. McCann, S. Sekiguci, B. Bross,W.-J. Han, "HEVC Test Model 3 (HM3) Encoder Description", JCTVC-E602, Geneva, CH, March 2011.
[4] A. Fuldseth, G. Bjøntegaard (Cisco),M. Sadafale, M. Budagavi (TI), "Transform design for HEVC with 16 bit intermediate data representation", JCTVC-E243, Geneva, CH, March 2011.
[5] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560–576, Jul. 2003.
[6] Marpe, D.; Schwarz, H.; Wiegand, T. , "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standar," IEEE Trans. Circuits Syst.Video Technol., vol. 13, no. 7, pp. 620-636, Jul. 2003.
[7] F. Bossen, Common test conditions and software reference configurations, JCTVC-D500, Daegu, Korea, January, 2011.
[8] https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware



Anis Benhajyoussef received his Dip. Ing. degree in telecommunication engineering and the M. degree in communication systems from the National School of Engineers of Tunis in 2005, 2007 respectively. He is currently pursing the Ph.D. degree.
His research interests include digital video coding and rate control.

Tahar Ezzedine received the electrical engineering in 1986, the M.S degree in Telecommunications in 1990, the Ph.D. degree in telecommunications in 2005, from the National Engineer School of Tunis (ENIT). He joined the National School of Engineers of Tunis from 1990 as research assistant. His research interests include adhoc sensor networking, embedded system and image and video coding.