Corrections to "Transform Coding Techniques in HEVC"

Tung Nguyen, Philipp Helle, Martin Winken, Benjamin Bross, Detlev Marpe, Heiko Schwarz, and Thomas Wiegand

***Abstract*—This is a correction to the "Transform Coding Techniques in HEVC" article published in the IEEE Selected Topics in Signal Processing, vol. 7, no. 6, Dec. 2013**

***Index Terms*—Transform coding, correction, entropy coding, residual quadtree, HEVC, H.265, MPEG-H, video compression.**

I. Corrections

In the text below, the corrections are suggested for the article "Transform Coding Techniques in HEVC" by T. Nguyen *et al.*, which was published in IEEE Selected Topics in Signal Processing in vol. 7, no. 6, December 2013 [1].

Correction 1

On p. 982, the *Subsection C* in *Section IV* should be substituted by the following text.

For binary arithmetic coding, a given non-binary absolute value z has to be binarized, i.e., decomposed into a sequence of binary decisions (bin string). Two information related to transform coefficient level coding are non-binary: the last significant scan position and the absolute levels.

1) Last Significant Scan Position: The binarization of its x and y coordinates is a composition of a truncated unary prefix and a fixed-length suffix. Since the same binarization process is used for both coordinates, the following description uses z as a substitution for x and y. Each prefix bin with an index i denotes the decision for $z > z_i$, with $z_i \in Z = \{0, 1, 2, 3, 4, 6, 8, 12, 16, 24\}$. The maximum prefix length i_{max} is constrained by $2 \cdot \log_2 \mathrm{N} - 1$ and N denotes the TB size. A suffix is present when the prefix consists of more than four bins, i.e., $z > 4$. This is exactly the case when the difference between two neighbouring entries in Z is greater than one. As a consequence for 4×4 TBs, the binarization always results in a bin string without suffix. When a suffix is present, the remaining value $z - z_{i_{max}}$ is decomposed by the fixed-length scheme with a number of bins equal to $\lfloor (i_{max} - 3)/2 \rfloor$.

2) Absolute Level: The binarization of absolute levels, denoted as z in the following, is a decomposition into four different syntax elements. They are referred to as *significance flag* (ω_{sig}), *absolute level greater than 1* (ω_{gr1}), *absolute level greater than 2* (ω_{gr2}), and *remaining absolute level* (ω_{rem}). At decoder side, z can be reconstructed simply by sum up the values of the syntax elements as denoted in the following.

$$z = \omega_{\mathrm{sig}} + \omega_{\mathrm{gr1}} + \omega_{\mathrm{gr2}} + \omega_{\mathrm{rem}} \tag{0}$$

The binary decomposition of an *absolute level* is controlled by the adaptive parameters b_0, b_1, and k. As illustrated in Fig. 5, the binarization process is a concatenation of three different binarization schemes: truncated unary code [24], truncated Golomb-Rice codes [30], [31], and Exp-Golomb codes [24]. Both parameters b_0 and b_1 specify the thresholds between the different binarization schemes. The parameter k specifies the order of the truncated Golomb-Rice codes with the order of the Exp-Golomb codes is equal to $k + 1$. The selection of the three parameters is a backward-adaptive process for each scan position such that the resulting bin strings are already close to a minimum-redundancy code for the given statistics of absolute transform coefficient levels. How these parameters are derived is given in more detail as follows.

Initially, for the first SB in a TB specified by the *last significant scan position*, the first variable threshold b_0 is set to be equal to two. It is kept equal to two up to and including the scan position with the first occurrence of $z > 1$. After passing that scan position, b_0 is set equal to one and is further reduced to be equal to zero after $z > 0$ occurs eight times within an SB. Note that a direct adaptation from two to zero can be performed when $z = 1$ occurs eight times. Similar to b_0, the Rice parameter k is set to be equal to zero before processing an SB. After each scan position, the value z is evaluated and k is incremented by 1 when $z > 3 \cdot 2^k$. However, this incrementation process is stopped whenever k reaches a value equal to four. The second bound b_1 depends on both b_0 and k with the relationship given by $b_1 = 4 \cdot 2^k + b_0$.

All bins resulting from the TU binarization scheme are transmitted in the regular operation mode of BAC. However, due to the adaptive parameter b_0, the ω_{gr2} syntax element can occur once and ω_{gr1} syntax element can occur eight times in an SB only. The syntax element ω_{rem} is transmitted by using the low-complexity bypass mode of BAC. For the transmission of the absolute levels within each SB, the syntax elements are grouped and each group is transmitted in a separate coding step. In the first coding step, ω_{sig} is transmitted, which is also referred to as *significance map*. In the second coding step, up to eight ω_{gr1} syntax elements are transmitted within an SB and the third coding step transmits up to one ω_{gr2} syntax element. After that, the signs ω_{sign} are transmitted for all non-zero entries of the significance map using the bypass mode of BAC. Finally, the ω_{rem} syntax elements for an SB are signaled in the fourth coding step, with all coding steps using the same reverse scan pattern.

Correction 2

On p. 984, (2), the symbol max should be replaced by min. After the correction, the equation should be as follows:

$$\chi_{\mathrm{csf}} = \min\left(1, \omega_{\mathrm{csf}}^{\mathrm{r}} + \omega_{\mathrm{csf}}^{\mathrm{b}}\right), \tag{2}$$

Correction 3

On p. 984, (4), the corrected calculation for the condition $\sum \omega_{\mathrm{gr1}} = C(\omega_{\mathrm{gr1}})$ should be $1+\min(C(\omega_{\mathrm{gr1}}), 2)$. After the correction, the equation should be as follows:

$$\chi_{\mathrm{gr1}} = \begin{cases} 1 + \min(C(\omega_{\mathrm{gr1}}), 2) : & \sum \omega_{\mathrm{gr1}} = C(\omega_{\mathrm{gr1}}) \\ 0 : & \text{otherwise} \end{cases}, \tag{4}$$

Manuscript received September 23, 2014; date of current version nulldate.

T. Nguyen, P. Helle, M. Winken, B. Bross, D. Marpe, and H. Schwarz are with the Image and Video Coding Group, Fraunhofer Institute for Telecommunications—Heinrich Hertz Institute, Berlin 10587, Germany (e-mail: (tung.nguyen@hhi.fraunhofer.de; philipp.helle@hhi.fraunhofer.de; martin.winken@hhi.fraunhofer.de; benjamin.bross@hhi.fraunhofer.de; detlev.marpe@hhi.fraunhofer.de; heiko.schwarz@hhi.fraunhofer.de).

T. Wiegand is with the Department of Image Processing, Fraunhofer Institute for Telecommunications—Heinrich Hertz Institute, Berlin 10587, Germany, and also with the Technical University of Berlin, Berlin 10587, Germany (e-mail: thomas.wiegand@hhi.fraunhofer.de).

Digital Object Identifier 10.1109/JSTSP.2014.2360428

References

[1] T. Nguyen, P. Helle, M. Winken, B. Bross, D. Marpe, H. Schwarz, and T. Wiegand, "Transform coding techniques in hevc," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 978–989, Dec. 2013.