



# The Use of Intra Prediction Method in Wavelet-Based Video Coding Systems

Gleb Verba, Kirill Bystrov<sup>(✉)</sup>, Viktor Dvorkovich, and Gennady Gryzov

Moscow Institute of Physics and Technology, Institutsky lane 9, Dolgoprudny, Russia  
{verba,kirill.bystrov}@phystech.edu, v.dvorkovich@mail.ru,  
gryzov@gmail.com

**Abstract.** Modern image and video compression technologies include both lossless compression methods, such as entropy coding, Inter-frame and Intra-frame coding, and lossy compression methods, such as discrete orthogonal transforms with quantization. All these techniques are actively applied in video codecs based on the H.264 and H.265 standards.

Discrete wavelet transform (DWT) is one of the most perspective versions of discrete orthogonal transforms. Both Intra-frame coding and wavelet decomposition of images allow to reduce the volume of transmitted data, but they are not used together in video coding systems. The combined usage of these methods seems to be a promising approach in terms of visual data compression.

Thereby the aim of this research is to develop a new technique based on the combination of Intra-frame coding and the DWT and to test the applicability of the proposed method in the image compression tasks.

The effectiveness of various implementations of the proposed algorithm, including those based on contexts and using several levels of wavelet decomposition of images, was evaluated in the study.

**Keywords:** Intra Prediction · Intra-frame coding · Discrete wavelet transform · Video coding · Data compression · Context-based coding

## 1 Introduction

Intra Prediction is widely used in video coding. It is a transformation that allows to reduce the data entropy at the input of entropy encoder by coding only residual values obtained using the prediction schemes. H.264 [1] and H.265 [2] are the examples of standards which include this technique.

Discrete wavelet transform (DWT) is another efficient approach to image coding. It is based on spatial-frequency analysis of input data and is also applied in video coding standards [3–5].

Combining these two methods opens up new prospects for video compression systems and offers a novelty because, although a number of related studies have been carried out [6, 7], no video coding system performing such concept has been implemented yet.

Thus, the purpose of this study is to attempt to combine Intra Prediction and DWT techniques and to test the applicability of the proposed method in the image compression tasks.

## 2 Intra-frame Prediction in the Space of Wavelet Coefficients

Intra Prediction method matches each coefficient  $s(i; j)$  and its predicted value  $\hat{s}(i; j)$ , which is determined by a prediction scheme using some spatially nearby reference data. In practice, a frame is divided into blocks, for which common reference values are used and which occupy the line above and/or the column to the left of the current block. The array of prediction errors  $e(i; j)$  is fed to the input of entropy encoder. This approach [8–10], which is used in H.264 [1] and H.265 [2] standards, is common for most video codecs. Block-based intra-frame prediction technique was significantly improved during last decade by reducing its computational complexity [11, 12] and by increasing the compression ratio for prediction modes [13, 14].

The forward and inverse prediction processes for wavelet coefficients are described by the following system:

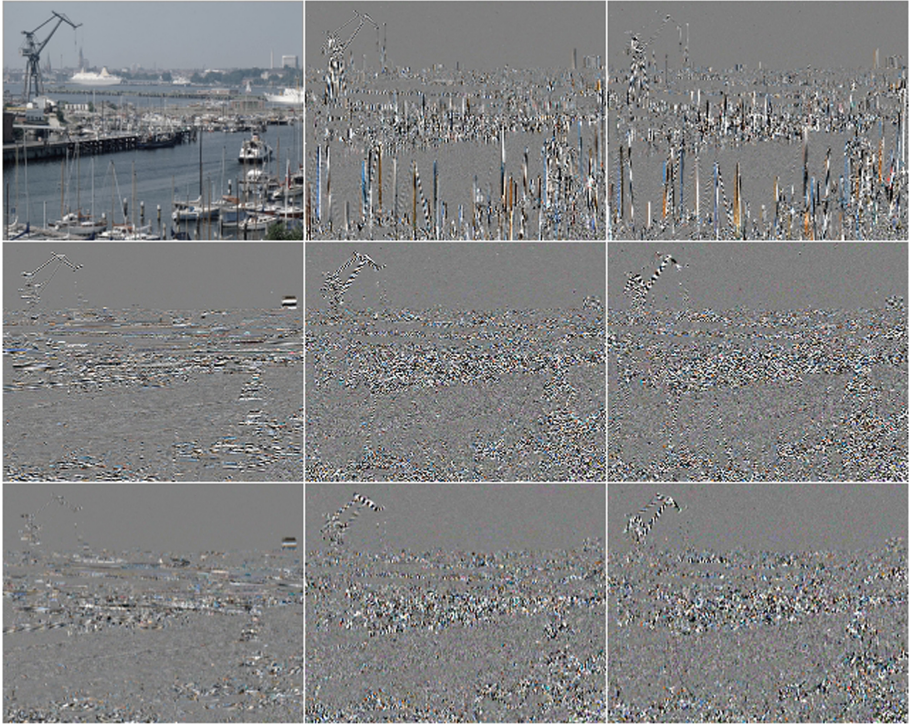
$$\begin{cases} e(i; j) = s(i; j) - \hat{s}(i; j) \\ s(i; j) = e(i; j) + \hat{s}(i; j) \end{cases} \quad (1)$$

Wavelet coefficients have a special correlation pattern: in zones, which are low-frequency in one of dimensions (refer to them as *semi-low-frequency*), the lines with horizontal or vertical orientation (depending on the specific frequency zone) are observed near the peculiarities of the original image—these are lines of the same sign with similar moduli (Fig. 1). This leads to the assumption that prediction with appropriate orientation will also be effective for *semi-low-frequency* subbands. The low-frequency subband is similar to the original image—that allows to suppose that Intra Prediction may be effectively applied to it.

### 2.1 Intra Prediction for Absolute Values of Wavelet Coefficients

As oscillations are characteristic of the high-frequency zones, it is more plausible to expect approximate constancy of the magnitude of these oscillations rather than their values proper, thus this paper proposes a new technique carrying out division the initial data into a field of moduli and a field of signs (the latter of is defined only for non-zero values in order to avoid redundancy introduction) prior to prediction. Note that using this technique requires taking into account the field of sign as well.

The frame of sign of DWT coefficients after quantization is depicted in the Fig. 2. The values oscillate by zero and there are horizontal or vertical lines of constant sign in *semi-low-frequency* subbands.

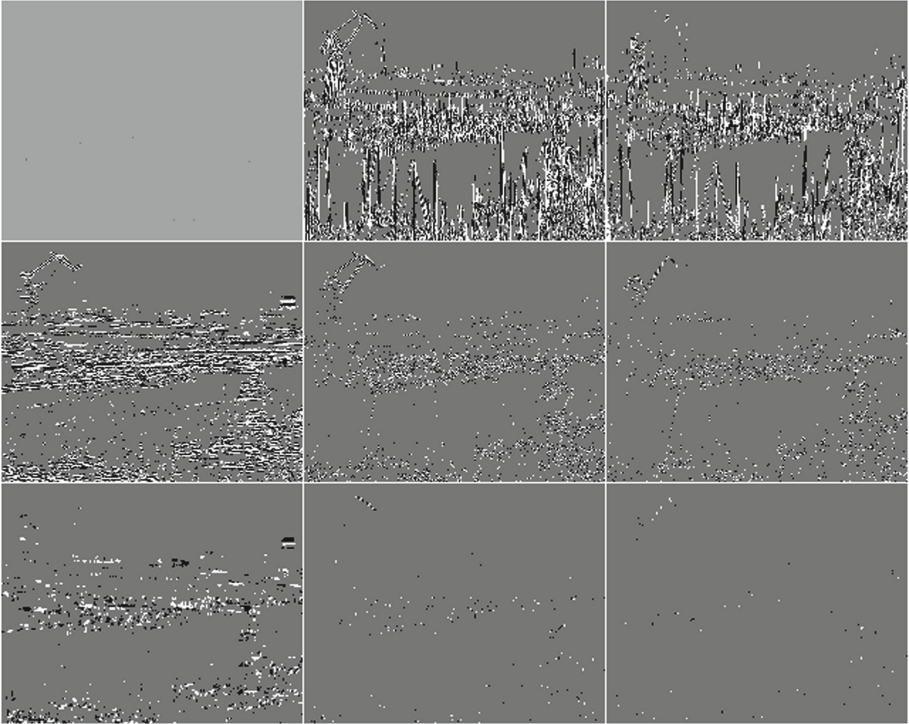


**Fig. 1.** Wavelet decomposition of image *kiel* by three-channel filter bank

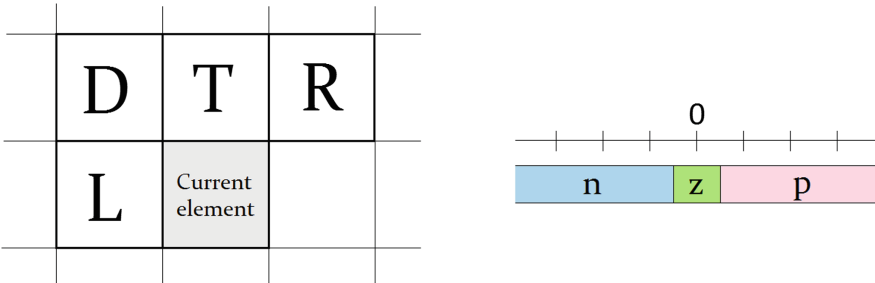
## 2.2 Context-Based Coding for Residual Frame of Intra-predicted Wavelet Coefficients

Context-based coding is an entropy coding technique which is based on computing conditional probabilities of symbols depending on preceding symbol sequences, which allows to improve encoding efficiency. Since the frequency sub-bands of wavelet-decomposed images have specific correlation between their coefficients, their statistical properties can be used for forming contexts of wavelet coefficients. This approach finds an application in a range of papers on compression, including wavelet-based compression [15].

Each wavelet coefficient is assigned a specific context defined by a set of neighboring coefficients. A table of conditional probabilities of symbols appearance is constructed for each context and is applied in subsequent entropy coding. Thus, each coefficient is coded using one of the probability tables depending on the context of the current coefficient. Compression is achieved because the conditional entropy in the general case is not greater than unconditional one and strictly less due to the presence of correlation. The complexity of the context is determined by the number of adjacent elements used, as well as the mapping of the set of their values to the set of contexts.



**Fig. 2.** Field of sign of DWT decomposition of luma component of image *kiel* at quantization step 33. Grey points correspond to zero quantized values, white points—to positive ones, black points—to negative ones

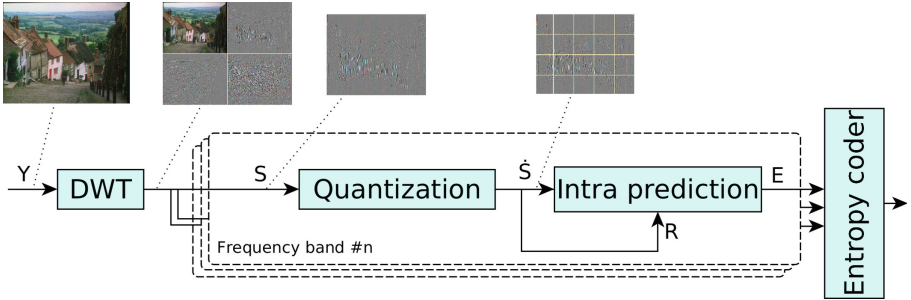


**Fig. 3.** Context model used in study: four neighboring coefficients (left *L*, top-left *D*, top *T* and top-right *R*) and their values range: negative *n*, zero *z* and positive *p* values.

In the framework of this paper it is suggested to apply the following model for context formation (Fig. 3).

### 3 Practical Implementation

Prediction was applied to frames of wavelet coefficients after their quantization. The coder scheme is depicted in Fig. 4.



**Fig. 4.** Coder scheme:  $Y$  - luma component of original image,  $S$  - coefficients of frequency subband of wavelet-decomposed image,  $\hat{S}$  - quantized wavelet coefficients of subband,  $E$  - array of prediction errors,  $R$  - reference data (the row above or/and the column to the left) for Intra Prediction in subband

Prediction process is preceded by picking out reference values taking place in the upper row or/and the left column. The remaining part is divided into blocks of size  $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$  or  $32 \times 32$ .

The following prediction modes, defined in the H.264 recommendation, were considered in the framework of the paper (Table 1): the 8 directed modes and the DC-mode, that had been defined for luma blocks of sizes  $4 \times 4$  and  $8 \times 8$ , and the planar prediction mode, that had been defined for luma and chroma blocks of size  $16 \times 16$ . Also adaptive mode was used for choosing optimal mode for each block. All the modes were extended to all working block sizes.

The efficiency of applying intra prediction to a block was estimated using a certain distortion metric. If applying intra prediction didn't lead to the metric value growth, the block was skipped without prediction. Either a mode was applied to the entire frame, or the adaptive mode was used: each block passed each prediction mode alternately, whereupon the mode minimizing the chosen metric was determined. Using adaptive mode requires transmission of the vector of values denoting the modes chosen for each block. The used metrics are: SAD (sum of absolute differences), SSD (sum of squared differences), SATD (SAD for Hadamard transform of the prediction remainder), \*SAHD (SATD modification: SAD for Haar transform of the prediction remainder), kNN [16].

**Table 1.** Tested intra prediction modes.

Mode index	Prediction mode
0	Vertical
1	Horizontal
2	DC
3	Diagonal down-left
4	Diagonal down-right
5	Vertical-right
6	Horizontal-down
7	Vertical-left
8	Horizontal-up
9	Planar
10	Adaptive

Three-channel filterbank 23/23/23—13/13/13 applied in [17] was chosen for wavelet transform implementation. The investigated quantization techniques were tested on an array of images of different resolutions (from 4CIF to Full HD) [18] using MATLAB modelling environment. Metric “entropy-PSNR”, which characterizes the ratio of the output data volume and the quality of the reconstructed signal, was applied to estimate the effectiveness of the proposed methods. The entropy calculation did not take into account the data necessary to represent the contexts.

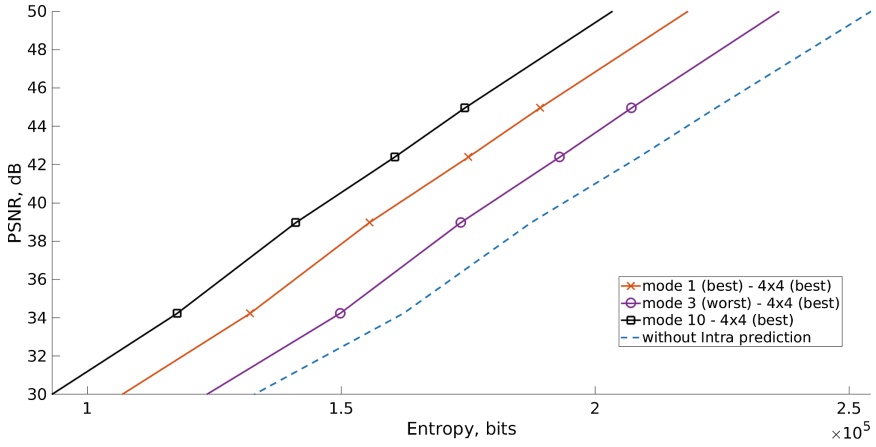
## 4 Results

The use of intra prediction in the domain of wavelet coefficients of LL subband provides guaranteed and noticeable gain in “entropy – PSNR” (Fig. 5). Adaptive mode has a significant advantage over single modes and the smallest block size is optimal. Nonetheless, the execution of an additional wavelet decomposition step of the same subband is more efficient (Fig. 6).

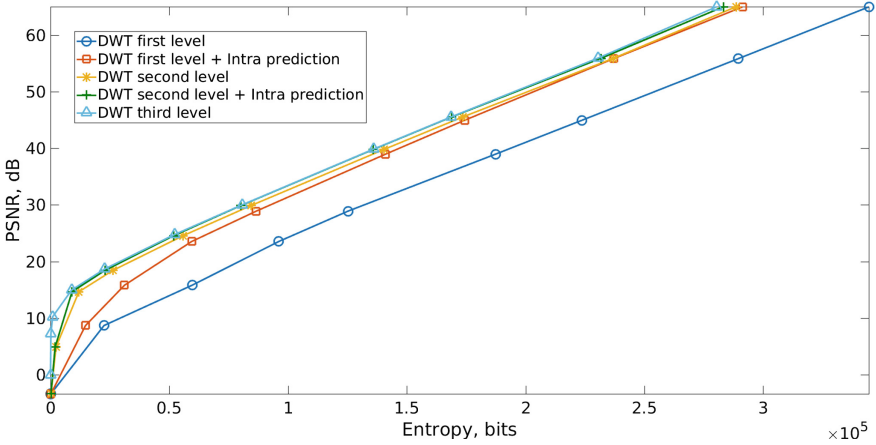
At the same time the benefit of intra prediction usage for *semi-low-frequency* subbands is approximately a few percent in terms of initial entropy of the subband, and only for the corresponding directed mode and adaptive mode (Fig. 7). Use of intra prediction in high-frequency subbands only decreases the compression ratio (Fig. 8).

Division into modulus and sign causes almost no change of the ratio “entropy–PSNR” in any frequency band of the tested images (Fig. 9).



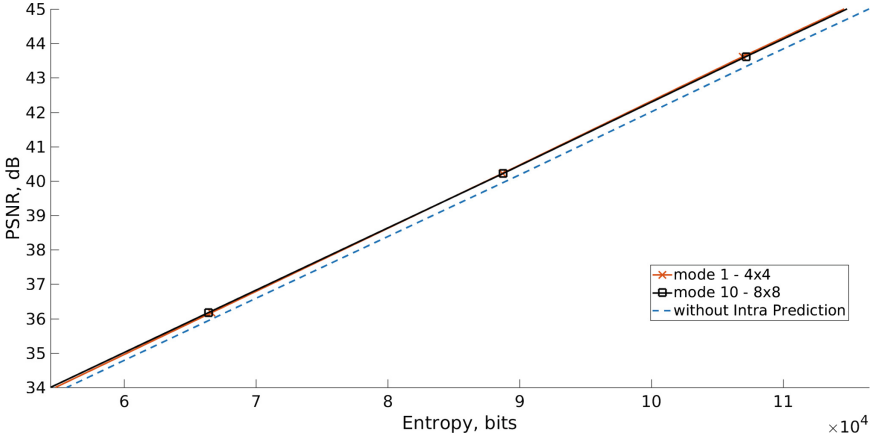


**Fig. 5.** Entropy-PSNR ratio of LL subband of *kiel* image in case of DWT with intra prediction. The size of the block that provides the best compression is chosen as a parameter for each curve and is included in the legend

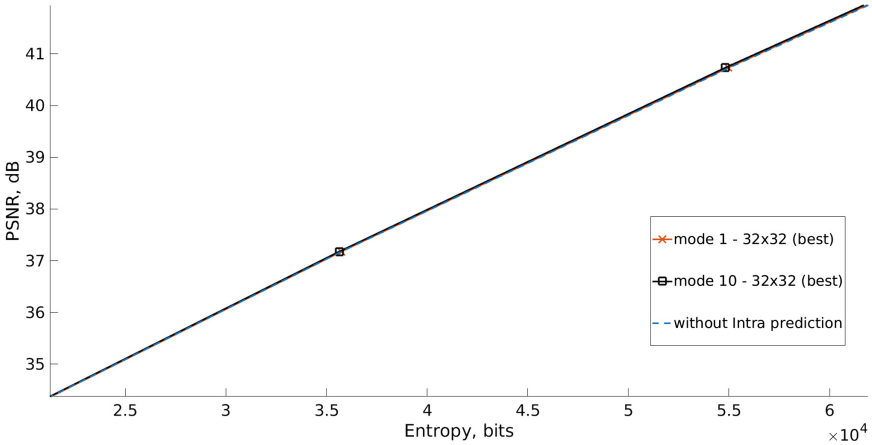


**Fig. 6.** Entropy-PSNR ratio of LL subband of *kiel* image in case of DWT with intra prediction. The block size is  $4 \times 4$

The best results for the proposed method were obtained with a preliminary recalculation of the entropy of the frequency subbands, taking into account the contexts of the wavelet coefficients: the use of intra prediction becomes more efficient for LL zone (Fig. 10), while for other subbands the entropy decreases significantly, both with and without the use of prediction (Fig. 11). Contexts consideration increases the compression ratio by up to 25%, depending on the frequency subband and the image.

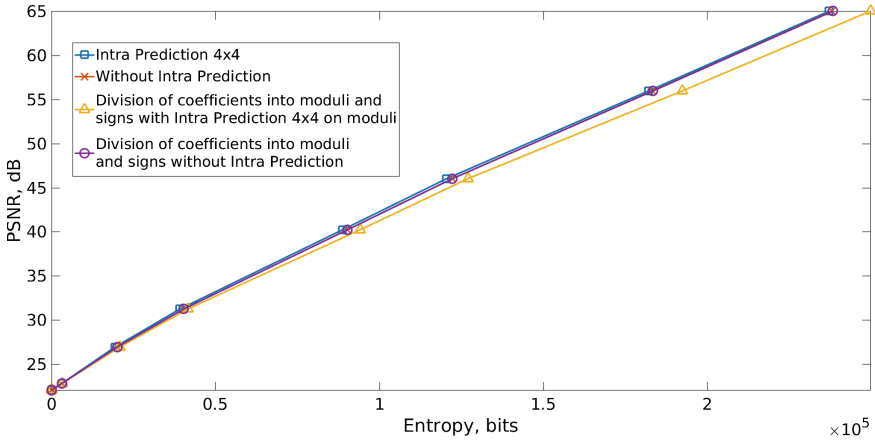


**Fig. 7.** Entropy-PSNR ratio of LM subband of *kiel* image in case of DWT with intra prediction. The size of the block that provides the best compression is chosen as a parameter for each curve and is included in the legend

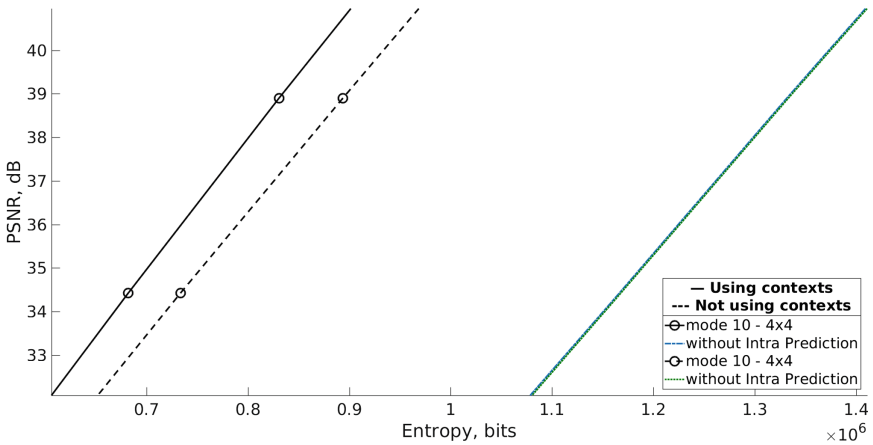


**Fig. 8.** Entropy-PSNR ratio of MM subband of *kiel* image in case of DWT with intra prediction. The size of the block that provides the best compression is chosen as a parameter for each curve and is included in the legend

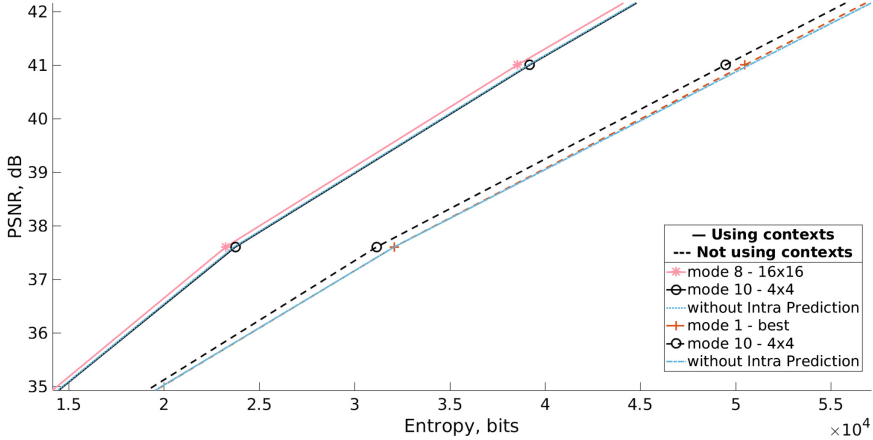




**Fig. 9.** Curves “entropy–PSNR” when using no transforms, after simple prediction, after division into moduli and signs and after prediction in the domain of the derived moduli for LM band of image *kiel*. The curve for the case of moduli extraction coincides with the one for the case of no transforms



**Fig. 10.** Entropy-PSNR ratio of LL subband of *kiel* image in case of DWT intra prediction with context usage and without it. Only best prediction modes are shown in graph. The block size ( $4 \times 4$ ) providing the best compression is chosen



**Fig. 11.** Entropy-PSNR ratio of LH subband of *kiel* image in case of DWT intra prediction with context usage and without it. Only best prediction modes are shown in graph. The block size providing the best compression is chosen

## 5 Conclusion

Application of Intra Prediction modes in the way they are used in the H.264/H.265 standards to the result of the wavelet decomposition of the image is unjustified: in spite of entropy reduction, it is still more efficient to execute additional DWT for low-frequency coefficients (Fig. 6)—both in terms of the ratio of compression and quality of the reconstructed image and in terms of computational complexity; and for other frequency domains it does not guarantee an improvement in compression (Fig. 7 and 8). This is explained by the fact that H.264/H.265 Intra Prediction modes are adapted to be applied to the components of YUV image, meanwhile the correlation between the wavelet coefficients for all frequency subbands (except LL) is completely different than the correlation in YUV. Conversely, the LL zone represents the approximation of the original image, which allows to reduce the entropy with such type of the Intra Prediction.

Modeling (Fig. 9) demonstrates that extracting the moduli of the wavelet coefficients without prediction gives a result, identical to the absence of transformations, while the prediction for absolute values of coefficients gives a negative result. The first observation may be explained simply: if the distribution is initially approximately symmetrical about zero, then the uncertainty of the value is reduced by 1 bit when turning to absolute values, since the probability to meet each value becomes twice as high. This difference of 1 bit is compensated for by introducing the sign that is positive or negative with approximately equal probability. The second observation demonstrates the fallacy of the assumption that the coefficient moduli are correlated.

The results of this study show evidently that the use of contexts of wavelet coefficients not only reduces the initial entropy of a frequency subband of the

wavelet-decomposed image (Fig. 11), but also improves the Intra Prediction results in LL zone (Fig. 10). Thereby it is proved that the use of contexts in wavelet video coding is a promising approach, which, however, requires more thorough research. One of the possible ways to improve the efficiency of the proposed method is to select a context model simultaneously with the choice of prediction mode, as well as taking into account the implementation features of the entropy coder.

## References

1. Wiegand, T., Sullivan, G.J., Bjontegaard, G., et al.: Overview of the H.264/AVC video coding standard. *IEEE Trans. Circ. Syst. Video Technol.* **13**(7), 560–576 (2003)
2. Sullivan, G.J., Ohm, J.-R., Han, W.-J., et al.: Overview of the high efficiency video coding (HEVC) standard. *IEEE Trans. Circ. Syst. Video Technol.* **22**(12), 1649–1668 (2012)
3. Taubman, D.S., Marcellin, M.W.: JPEG2000: standard for interactive imaging. *Proc. IEEE* **90**, 1336–1357 (2002)
4. ISO/IEC 15444–3:2002, Information technology - JPEG 2000 image coding system - Part 3: Motion JPEG 2000 (2002)
5. Onthriar, K., Loo, K.K., Xue, Z.: Performance comparison of emerging Dirac video codec with H.264/AVC. In: *International Conference on Digital Telecommunications (ICDT06)*, p. 22 (2006)
6. Mai, Z., Nasiopoulos, P., Ward, R.: A wavelet-based intra-prediction lossless image compression scheme. In: *2009 Digest of Technical Papers International Conference on Consumer Electronics*, pp. 1–2, Las Vegas (2009)
7. Elarabi, T., Sammoud, A., Abdelgawad, A., Li, X., Bayoumi, M.: Hybrid wavelet - DCT intra prediction for H.264/AVC interactive encoder. In: *2014 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)*, pp. 281–285, Xi’an (2014)
8. Nan, Z., Baocai, Y., Dehui, K., Wenying, Y.: Spatial prediction based intra-coding [video coding]. In: *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*, vol. 1, pp. 97–100, Taipei (2004). <https://doi.org/10.1109/ICME.2004.1394134>
9. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG: New Intra Prediction using Intra-Macroblock Motion Compensation, JVT-C151 (2002)
10. Lainema, J., Bossen, F., Han, W., Min, J., Ugur, K.: Intra coding of the HEVC standard. *IEEE Trans. Circ. Syst. Video Technol.* **22**(12), 1792–1801 (2012). <https://doi.org/10.1109/TCSVT.2012.2221525>
11. Yi, H., Qin, H.: The optimization of HEVC intra prediction mode selection. In: *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, pp. 1743–1748, Changsha (2017). <https://doi.org/10.1109/ICISCE.2017.364>
12. Adireddy, R., Palanisamy, N.K.: Effective approach to reduce complexity for HEVC intra prediction in inter frames. In: *2014 Twentieth National Conference on Communications (NCC)*, pp. 1–5, Kanpur (2014). <https://doi.org/10.1109/NCC.2014.6811337>

13. Sanchez, G., Fernandes, R., Agostini, L., Marcon, C.: DCDM-intra: dynamically configurable 3D-HEVC depth maps intra-frame prediction algorithm. In: 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 1782–1786, Athens (2018). <https://doi.org/10.1109/ICIP.2018.8451620>
14. Agarwal, P., Jiang, M., Ling, N., Zheng, J., Zhang, P.: Enhanced intra prediction mode coding by using reference samples. In: IEEE International Workshop on Signal Processing Systems (SiPS), 296–299, Cape Town (2018). <https://doi.org/10.1109/SiPS.2018.8598421>
15. Jiang, X., Song, B., Zhuang, X.: An enhanced wavelet image codec: SLCCA PLUS. In: International Conference on Audio, Language and Image Processing (ICALIP) (2018)
16. Boltz, S., Wolsztynski, E., Debreuve, E., et al.: A minimum-entropy procedure for robust motion estimation. In: International Conference on Image Processing (2006)
17. Bystrov, K., Dvorkovich, A., Dvorkovich, V., Gryzov, G.: Usage of video codec based on multichannel wavelet decomposition in video streaming telecommunication systems. In: Communications in Computer and Information Science, Distributed Computer and Communication Networks. DCCN 2017, vol. 700, pp 108–119 (2017). [https://doi.org/10.1007/978-3-319-66836-9\\_10](https://doi.org/10.1007/978-3-319-66836-9_10)
18. Tested images. [https://github.com/tovoidcast/dwt\\_test\\_benchmark](https://github.com/tovoidcast/dwt_test_benchmark)