

GAIN ADAPTED QUANTIZATION IN HEVC CODING APPLIED TO DRONE REMOTE SENSING

Ulrike Pestel-Schiller, Paul-Robert Meinicke, Jörn Ostermann

Johannes Busch

Institut für Informationsverarbeitung
Leibniz Universität Hannover
Appelstr. 9A
30167 Hannover
Germany

HAIP Solutions GmbH
Escherstr. 23
30159 Hannover
Germany

ABSTRACT

For storing or transmitting hyperspectral images (HSI) in drone remote sensing, an efficient data compression with low computational cost has to be done onboard. Many scenarios do not allow any loss of information except noise which is not interpreted as information. We present an HSI data compression scheme using H.265/HEVC Main10 Profile Hardware, already integrated on the camera system of a drone. Using reference software, we determine, for each test data investigated, the so called best quantization step size which holds the constraint of losing no information at the smallest possible data amount. We map the analog sensor gain to the best quantization step size and find a linear dependency which allows a correct setting of the quantization step size in real-time. Finally, we verify the conformity of the reference software used for the investigations with hardware simulation results. We achieve compression ratios between 11 and 24.

Index Terms— hyperspectral imaging, drone remote sensing, data compression, coding, quantization, HEVC

1. INTRODUCTION

Materials can be distinguished by their spectral characteristics absorption or reflectance. Based on this, hyperspectral remote sensing allows the detection and identification of specific materials, e.g. for monitoring agriculture and forest status, for environmental studies, for search and rescue services, for disaster management or for geological mapping.

Drone remote sensing with hyperspectral sensors collects a large amount of data that has to be either stored onboard or transmitted to ground. Both scenarios benefit from a hyperspectral data compression with the constraint of eliminating nothing but noise.

First approaches of hyperspectral image compression extend established 2D image coding techniques [1] into 3D, e.g. the wavelet-based techniques 3D Set Partitioning in Hierarchical Trees (3D SPIHT) [2] as extension of 2D SPIHT [3],

3D Set Partitioning Embedded block (3D SPECK) [4] as extension of 2D SPECK [5] or 3D JPEG2000 [6] as extension of JPEG2000 [7].

Other approaches work separably and typically apply a Principal Component Analysis (PCA) across the spectral bands followed by still image coding like JPEG2000 to reduce the spatial correlation [8], [9]. Approaches using PCA or KLT either need to buffer rather large groups of bands or have to do parallelizing at the cost of lower data compression.

Yet other approaches [10], [11], [12], [13] use video coding standards H.264/MPEG-4 AVC [14], H.265/HEVC [15] or VVC [16] for compression of hyperspectral and multispectral image data. Here, the hyperspectral data cube is interpreted as an image sequence. The temporal direction of the video codec is assigned to the spectral direction of the hyperspectral image data, such that the spectral correlation is exploited by the temporal prediction of a video coding standard. Using classification criteria to evaluate HEVC compressed HSI data is presented in [17].

Hardware oriented coding approaches for remote sensing focus on speed and low complexity. In [18], an FPGA based solution is presented, in [19] the Hyper-LCA is used on embedded GPUs.

In our approach we use dedicated H.265/HEVC hardware, already integrated in the hyperspectral camera from HAIP Solutions [20], which can be mounted on any DJI Matrice drone [21]. The camera covers a wavelength range from 500 to 1000 nm with 100 spectral bands and a spatial resolution of 540x540 pixels. The amplitude depth of the hyperspectral image data (HSI) is 14 bit. HAIP Solutions BlackBird V2 drone uses an NVIDIA Jetson NANO, while in this work we investigate the successor NVIDIA Jetson XAVIER NX on which the video codec HEVC Main10 Profile is integrated and can be applied for compression of the HSI. This approach benefits from all the powerful techniques of a mature standard implemented in the camera system.

With this approach, realtime compression of HSI data is possible without using CPU or GPU resources which can be

used for other onboard tasks. This makes hyperspectral remote sensing possible when a large amount of data is captured which now can be compressed to save disk usage or even livestreamed over long range wireless link for realtime analysis. Since today most uses of hyperspectral data are still scientific, information must be retained as good as possible. This does not mean lossless but allows the reduction of noise.

The paper is organized as follows: Section 2 presents the test data. In Section 3 the experimental set up is presented and coding results are shown. Section 4 focusses on the relation between the analog gain of the hyperspectral sensor and the quantization step size of the codec. Coding results using the H.265/HEVC hardware encoder (NVENC) from NVIDIA Jetson Xavier NX are presented and discussed in Section 5.

2. TEST DATA

With the BlackBird V2 drone we recorded four hyperspectral data cubes. Each cube is spatially cropped to a size of 512x512 pixels to fit directly into the block based HEVC codec and is called test data further on. The four cubes together comprise the test data set. During the recording we put emphasis on extreme illumination conditions by varying the analog gain G of the sensor in the range of 1 to 15.5. The test data set is publicly available at [22]. The test data set contains landscape, houses and trees as shown in Fig. 1.



Fig. 1. Test data set (band 60)

3. EXPERIMENTAL SET UP AND RESULTS

Fig. 2 shows the block diagram of the experimental set up.

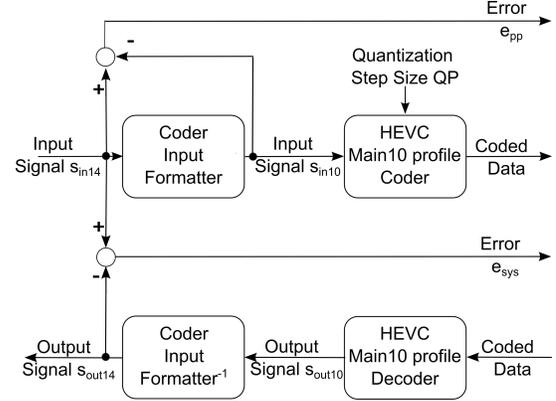


Fig. 2. Experimental set up

The Coder Input Formatter discards the four least significant bits, thus converting the 14 bit sensor signal s_{in14} into a valid HEVC Main10 Profile input signal s_{in10} , input into the HEVC Main10 Profile codec. The decoded output signal s_{out10} is converted to the output signal s_{out14} with 14 bit amplitude depth which allows us to evaluate the overall error e_{sys}

$$e_{sys} = s_{in14} - s_{out14} \quad (1)$$

of the whole system. Furthermore, the preprocessing error e_{pp}

$$e_{pp} = s_{in14} - 16 * s_{in10} \quad (2)$$

caused by the conversion to 10 bit amplitude depth is measured.

We have to make sure that no information is lost due to coding or the other way around we have to verify that only noise is eliminated by the coding.

3.1. Preprocessing of the Input Signal

The conversion of the input signal s_{in14} from 14 bit to 10 bit amplitude depth can be interpreted as a uniform quantization. It has to be proven that the least significant bits 1...4 do not contain any information but just white noise or in other words that the uniform quantization can be assumed as fine quantization.

Assuming uniform fine quantization, the variance $E[q^2]$ of the quantization error q is given by

$$E[q^2] = \frac{1}{12} \delta^2 \quad (3)$$

with δ representing the quantization interval. By shifting 4 bits, the theoretical variance of the quantization error can be averaged to

$$E[q^2] = \frac{1}{12} (2^4)^2 = 21.33. \quad (4)$$

The measured variances per spectral band of the preprocessing error e_{pp} of the test data are shown in Fig. 3. The

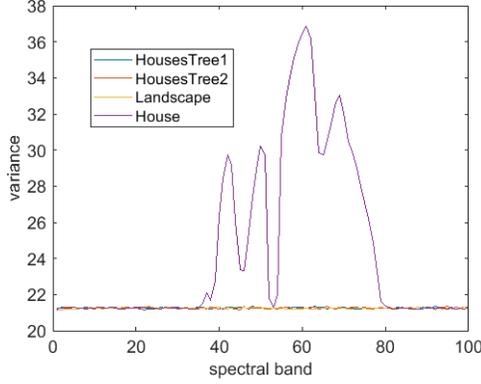


Fig. 3. Variance of error e_{pp}

preprocessing error e_{pp} of test data 'House' contains more than just noise which means that the condition of fine quantization is not given. By investigating the data cube of 'House' we found saturated areas.

By measuring the variance of the quantization error e_{pp} only in areas without saturation we get Fig. 4 where additionally the mean value per band over the test data set is shown. Averaging this mean we get a preprocessing error variance of 21.25 which is very close to the theoretical value 21.33. Furthermore, it was verified that a uniform distribution of the preprocessing error e_{pp} is given to be expected for fine quantization.

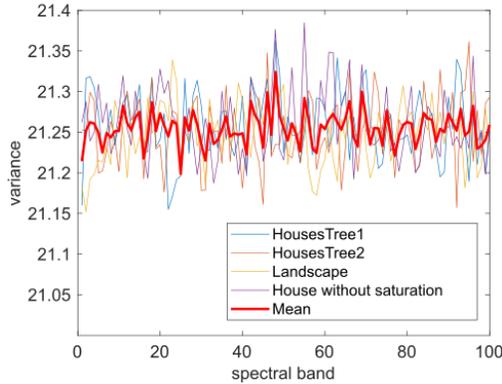


Fig. 4. Variance of error e_{pp} in areas without oversaturation

Because of the high dynamic range of 14 bit of the input signal s_{in14} the spectral signatures of the input signal s_{in14} and the shifted signal s_{in10} will not differ when displayed on a monitor. Therefore we displayed the error signals e_{pp} and e_{sys} on the monitor, both in original range and scaled to maximum display range of 8 bit. By viewing all error signals we verified that only noise structures are visible, except for oversaturated bright areas or very dark areas. In dark areas the signal is mostly zero so that specific quantization artifacts

occur. These quantization artifacts as well as the oversaturation violate the condition of fine quantization, but extremely dark or oversaturated areas never contain any interesting information. Thus, we showed that no information disappears by converting from 14 to 10 bit amplitude depth.

3.2. HEVC Codec

We investigate the standardized coding system High Efficiency Video Coding (HEVC) Main10 Profile [15] using the reference software for Rec. ITU-T H.265 | ISO/IEC 23008-2 High efficiency video coding (HEVC) version HM-16.6. This profile is available in the hardware encoder (NVENC) from NVIDIA Jetson Xavier NX. Due to the hardware constraints the coding structure 'low-delay-P' is chosen with one intra-coded I-frame followed by 24 forward predicted P-frames. The missing color components are substituted by the value 512. Each test data is coded with different quantization step sizes QP in the range of -6...18 resulting in a set of coded test data. Each coded test data is assigned a pair of quantization step size and data amount.

For evaluating the whole system the rate distortion function (RDF) is calculated where the 'Peak Signal to Noise Ratio' (PSNR) in [dB] over the data amount in bit per pixel per band [bpppb] is given as defined in (5) by

$$PSNR = 10 \log \frac{(2^{14})^2}{e_{sys}^2} \quad (5)$$

Fig. 5 shows the RDFs of the test data set. According to Eq. 3, with fine quantization the RDF shows a straight line. With increasing the step size QP , the quantization gets coarser, the data amount smaller and, at some stage, the RDF falls in a nonlinear way. It should be mentioned that in the range of fine quantization the RDF of the test data differs by an offset because of the different statistics like entropy.

For each test data, we are interested in that quantization step size QP which holds the constraint of losing no information at the smallest possible data amount.

The resulting quantization step sizes QP_{best} , corresponding data amounts and compression ratios are shown in Tab. 1.

Table 1. QP_{best} , data rate and compression ratio

test data	QP_{best}	data rate [bpppb]	compression ratio
HousesTree1	-2	0.68	23.4
HousesTree2	0	1.25	12.8
Landscape	6	0.93	17.2
House	10	1.38	11.6

Without any data compression each pixel per band has to be stored in two bytes. Tab. 1 shows that we achieve compression ratios of roundabout 11 to 24. Furthermore, there is

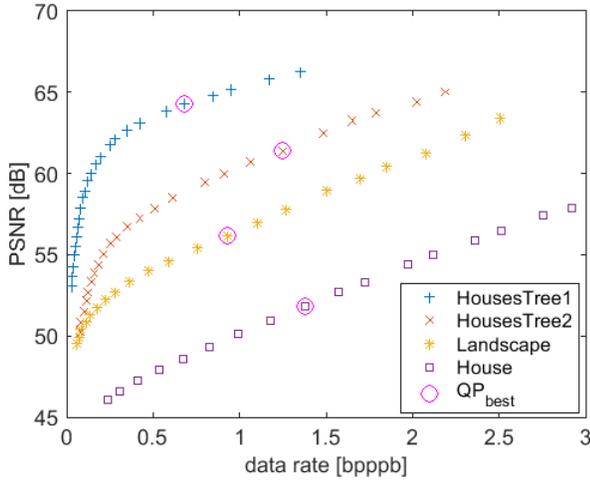


Fig. 5. Rate distortion function of test data using reference software

not one QP_{best} for the whole test data set but one for each test data.

4. RELATION BETWEEN GAIN AND BEST QUANTIZATION STEP SIZE

We observe for each test data a different step size QP_{best} which can be due to having a different amount of noise in each test data in which a corresponding amount of quantization noise can be masked. Therefore it is worth to have a look at the parameter settings of the analog gain G of the recording. We use extremely different illumination scenarios, for which the analog gain G of the sensor has to be adapted. A low analog gain G produces a small amount of noise, a high analog gain G a larger amount of noise. Thus, a large system error q_{sys} can be masked by the noise of a high analog gain G . It should be mentioned that the amount of noise caused by an analog high gain is smaller than that caused by amplification of a low-gain-recorded signal after discretization.

This attempt to explain the different QP_{best} values is checked by inspecting the graph of QP_{best} subject to gain G . Fig. 6 shows that this relation can be assumed as linear.

Due to coder constraints the quantization step size has to be an integer value. To be on the safe side, we map a given analog gain G to a quantization step size $QP_{gainSoft}$ of the reference software by rounding down the corresponding quantization value given by the straight line as

$$QP_{gainSoft} = \lfloor 0.8 * G - 2 \rfloor \quad (6)$$

We achieve compression ratios between 11 and 24.

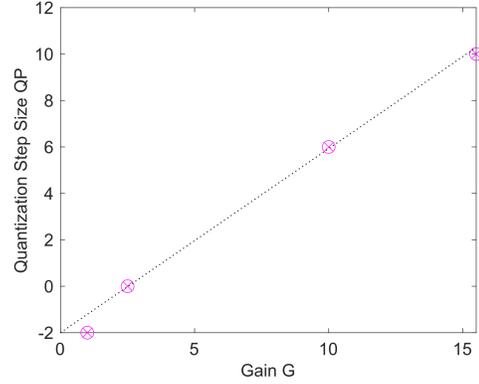


Fig. 6. Quantization step size QP_{best} subject to gain G

5. RESULTS WITH HARDWARE CODER

Using the NVENC HEVC Main 10 Profile with variable bit rate we coded the test data. The decoding was done with the reference software. Due to implementation, only positive QPs are possible with NVENC HEVC Main 10 Profile on Jetson Xavier NX. Therefore only the test data 'House' and 'Landscape' are investigated further on.

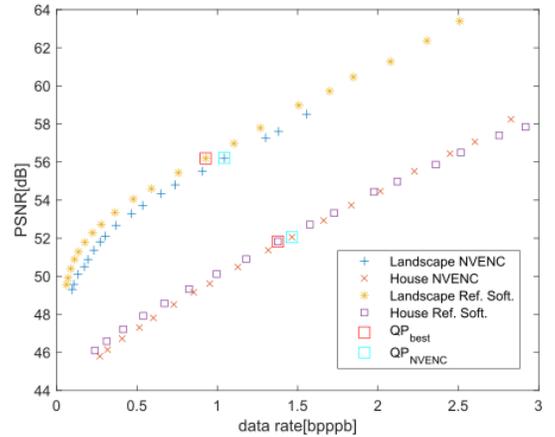


Fig. 7. RDFs of Reference software and NVENC

The rate distortion curves of NVENC are shown and compared to those of the reference software in Fig. 7. The NVENC-RDFs fit well to the RDFs of the reference software. Comparing quantization step size QP_{best} of the reference software (marked with red squares) with the quantization step size QP_{NVENC} of the NVENC HEVC main 10 profile having similar PSNR respective image quality (marked with cyan squares), we find a bias of 2 between QP_{best} and QP_{NVENC} . Thus, we get $QP_{gainNVENC}$ given as

$$QP_{gainNVENC} = QP_{gainSoft} - 2 \quad (7)$$

and finally, using Eq. 6

$$QP_{gainNVENC} = \lfloor 0.8 * G - 4 \rfloor \quad (8)$$

6. CONCLUSION

We presented an HSI data compression scheme using a dedicated H.265/HEVC Main10 Profile Hardware, already integrated on board of a drone. Using reference software, we determined, for each test data investigated, that quantization step size QP_{best} which holds the constraint of losing no information at the smallest possible data amount. We mapped the analog sensor gain to the best quantization step size and found a linear dependency which allows a correct setting of the quantization step size in real-time. Finally, we verified the conformity of the reference software of the standardization process used for the investigations with the NVENC results. We achieve compression ratios of at least 11.

7. REFERENCES

- [1] J. Shapiro, "Embedded image coding using zerotress of wavelet coefficients," *IEEE TSP*, vol. 41, pp. 3445–3462, Dec 1993.
- [2] B.-J. Kim and Z. Xiong et al., "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," *IEEE T-CSVT*, vol. 10, pp. 1374–1387, Dec 2000.
- [3] A. Said and W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE T-CSVT*, vol. 6, pp. 243–350, June 1996.
- [4] X. Tang and W. A. Pearlman et al., "Hyperspectral image compression using three-dimensional wavelet coding," *Proc. SPIE*, vol. 5022, pp. 1037–1047, Jan 2003.
- [5] A. Islam and W.A. Pearlman, "An embedded and efficient low-complexity hierarchical image coder," *Proc. SPIE*, *VCIP '99*, vol. 3653, pp. 294–305, Jan 1999.
- [6] ISO/IEC and ITU-T, *Recommendation ITU-T T.809 and ISO/IEC 15444-10: Information technology - JPEG 2000 image coding system: Extensions for three-dimensional*, Geneva, Switzerland, May 2011.
- [7] D.S. Taubmann and M.W. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, Boston, USA, 2002.
- [8] B. Ramakrishna and J. Wang et al., "Spectral/spatial hyperspectral image compression with virtual dimensionality in conjunction," *Proc. SPIE*, vol. 5806, pp. 772–781, March 2005.
- [9] Q. Du and J.E. Fowler, "Hyperspectral image compression using JPEG2000 and principal component analysis," *IEEE GRSL*, vol. 4.2, pp. 201–205, April 2007.
- [10] L. Santos and S. López et al., "Performance evaluation of the H.264/AVC video coding standard for lossy hyperspectral image compression," *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, vol. 5, no. 2, pp. 451–461, 2012.
- [11] F. Gao and X. Ji et al., "Compression of Multispectral Image using HEVC," *Proc. SPIE*, vol. 9273, Oct 2014.
- [12] A. Meyer and N. Genser et al., "Multispectral image compression based on HEVC using pel-recursive inter-band prediction," *IEEE MMSP*, pp. 1–6, 2020.
- [13] P. Seltsam and P. Das et al., "Adaptive and scalable compression of multispectral images using VVC," *Data Compression Conference*, 2023.
- [14] ISO/IEC and ITU-T, *Recommendation ITU-T H.264 and ISO/IEC 14496-10 (MPEG-4 Part 10): Advanced Video Coding (AVC)-3rd Ed.*, Geneva, Switzerland, Jul 2004.
- [15] ISO/IEC, ITU-T: Recommendation ITU-T H.265, and ISO/IEC JTC1/SC 29 23008-2:2015-05-01 MPEG-H Part 2/, "High efficiency video coding HEVC," 2015.
- [16] ITU-T and ISO/IEC JTC1, "Versatile Video Coding, ITU-T Recommendation H.266 and ISO/IEC 23090-3 VVC," 2020.
- [17] U. Pestel-Schiller and K. Vogt et al., "Impact of hyperspectral image coding on subpixel detection," *Proc. of 32nd PCS*, 2016.
- [18] J. Caba and M. Díaz et al., "FPGA-based on-board hyperspectral imaging compression: Benchmarking performance and energy efficiency against gpu implementations," *Remote Sens.*, vol. 12, no. 22, 2020.
- [19] M. Díaz and R. Guerra et al., "Real-time hyperspectral image compression onto embedded GPUs," *IEEE J. Sel. Top. Appl. Earth Obs. and Remote Sens.*, vol. 12, no. 8, pp. 2792–2809, 2019.
- [20] HAIP Solutions GmbH, <https://www.haip-solutions.com/hyperspectral-camera-drone/>, 2023.
- [21] DJI, <https://enterprise.dji.com/de/matrice-300/>.
- [22] HAIP Solutions GmbH, <https://cloud.haip-solutions.com/s/qEknRiXzXo97NmA>, 2023.