HEVC-BASED COMPRESSION OF HIGH BIT-DEPTH 3D SEISMIC DATA

Miloš Radosavljević[1,2], Zixiang Xiong[2], Ligang Lu[3], Detlef Hohl[3], and Dejan Vukobratović[1]

[1]Dept of Power, Electronics and Comm. Eng., University of Novi Sad, 21000 Novi Sad, Serbia
[2]Dept of ECE, Texas A&M University, College Station, TX 77843
[3]Shell International Exploration and Production Inc., Houston, TX 77082

milos.r@uns.ac.rs; zx@ece.tamu.edu; {ligang.lu,Detlef.Hohl}@shell.com; dejanv@uns.ac.rs

ABSTRACT

In a previous work [1] we applied the idea of HEVC intra coding to compression of 32 b/p seismic images. Results are significantly better than those from a licensed commercial wavelet-based codec that is currently used at Shell for seismic image compression, which performs on par with JPEG-XR. Building upon [1], this paper exploits HEVC-based inter predictive coding for 3D 32 b/p seismic data. We propose a new model for the Lagrange multiplier in R-D optimization to accommodate 32 b/p bit-depth and extended quantization parameter range. We also focus on reducing the complexity of motion estimation to meet application needs. Experiments with our new codec show a 95% complexity reduction of encoding time at 10:1 compression ratio with only 1 dB loss on average PSNR. The compression performance and subjective quality of our codec received high evaluation marks from Shell geologists.

Index Terms— HEVC, high bit-depth video compression, 3D volumetric seismic data, Lagrange optimization, motion estimation and complexity reduction.

1. INTRODUCTION

In the oil and gas industry, development and deployment of new technologies to improve productivity and reduce costs have become increasingly important. In seismic data processing, the cost and productivity are highly associated with the data volumes and data transfer speed between the computer system and storage disks. As new technologies emerge, the size of the seismic data set also becomes larger, such as 3D seismic data, deep subsurface seismic survey, and high-resolution seismic imaging. Compression is often required to enable and facilitate the deployment of new technologies to reduce the required storage and processing time. One option for the industry is to license commercial compression software libraries for their compression needs in seismic processing. However, as the seismic processing technology advances, the requirement for the computing hardware and compression from new applications often change. Without the ownership of the source code, the compression library cannot be customized unless paying more in addition to the licensing fee to meet the requirements arising from new seismic processing applications. Thus a more viable and flexible solution is to develop seismic data compression library by adapting the state-of-the-art image or video compression standards' test model codes to meet one's compression requirements stemming from emerging new technologies and their applications. JPEG-XR [2] was adopted for compression of 1D and 2D 32-bit floating-point seismic data in [3] because it performs on par with JPEG 2000 [4] but with computational complexity similar to JPEG [5]. Although JPEG-XR cannot compress all 32 bits of an input image, low complexity makes it a good candidate for commercial use. Another low-complexity scheme is presented in [6], based on an algorithm that has been tailored specifically for high bit-depth seismic image compression. It uses wavelet-based approach, with performance comparable to JPEG-XR. Our previous work in [1] presents a new codec for 32 b/p image compression under the framework of HEVC intra coding [7, 8]. It uses BinDCT transform [9] of flexible block sizes ranging from 4x4 to 32x32 pixels, 12 intra prediction modes, JPEG-XR's uniform quantization scheme (with increased quantization parameter range), and slightly modified CABAC at the end for efficient entropy coding [10]. JP3D, an extension of the JPEG2000, has been proposed to add support for 3D volumetric datasets [11]. However, it has not been often considered as a compression scheme for seismic image applications.

The results reported in [1] significantly outperforms that of the previously presented compression schemes. However the proposed scheme is image based. It does not exploit temporal redundancy in 3D data. In order to improve the compression performance we treat 3D seismic data as a sequence of frames and employ HEVC-based inter prediction. To the best of our knowledge there is no codec on the market for 32 b/p seismic data that exploits temporal redundancy for increased performance. In addition, to the best of our knowledge the resulting codec is the first that compresses all 32 bits of an input image. The experiments are conducted on 2D slices of 3D wavefield seismic data. The proposed work is compared with compression schemes used in the industry (including our previous codec that uses only intra prediction) A significant effort has been dedicated on reducing the codec's computational complexity with a focus on applications of 3D seismic data compression.

2. CODEC DESCRIPTION

The HEVC codec accepts 8-16 b/p data. In this section, we describe a new codec for 32 b/p 3D volumetric seismic data under the framework of HEVC. Proposed method is an extension of our previous work [1] that uses only intra prediction. Motion estimation (ME), as the most important part of the new proposed method, is described with an analysis of the motion estimation complexity reduction in order to meet application needs. Furthermore, we focus on redefining the Lagrange multiplier in order to empower the use of the extended bit-depth and extended quantization parameter (QP) range (as necessitated by 32 bit depth, see more details in [1]).

2.1. Complexity reduction on motion estimation

Motivated by the superior performance of HEVC, we treat 3D seismic data as a sequence of frames, similar to natural video, and exploit ME for predictive coding. ME is based on block matching which consist of finding a displaced block within a reference frame defined with motion vector (MV), while at the same time minimizing a matching distortion measure. A basic prediction unit (PU) represents the region with the same prediction-related data [16].

For a particular PU, ME comprises of the following steps. Advance motion vector prediction (AMVP) scheme is exploited [16, 17]. Based on the standardized competition scheme only two spatial candidates are chosen among the adjacent blocks. In case of unavailability of the spatial candidates, one temporally collocated neighbor can be selected. If the final number of the available candidates is less than two, zero-MV can be added to the competition list so at least two candidates compete for the motion vector predictor [18]. They are further passed to the rate-distortion optimization (RDO) module, and based on the minimal cost, one MV candidate is selected as a predictor for the particular PU. Later on, this candidate is used as a starting point in motion search (MS). Although temporal candidate can be useful to improve performance, it requires significant amount of the memory to store the motion-related data for each PU in reference frames. However, the encoder can control the usage of the temporal predictor by using a flag within the picture parameter set. Similar motion vector prediction scheme has been used in MERGE and SKIP mode (special case of the merge mode when coded block flag is zero). The same candidates as in AMVP have been used, except at most five are selected for the RDO. The number of competitors in MERGE mode can be controlled by the parameter *NumMergeCands* within slice header and can be between 1 and 5. Non-merge ME continues with a predefined set of integer points (search window) within integer-sample precision MS is performed. By default full search is performed where each integer point displacement has been checked in RD sense. Thereafter, fractional ME is performed within the neighborhood of the selected integer displacement. Eight surrounding points are evaluated using half-sample precision, followed by eight quarter-sample search at the end to fine tune previously selected points. It is important to note that fractional samples use seven-tap and eight-tap interpolation, additionally increasing computational complexity. Also multiple partitions sizes (e.g., PU blocks of square, symmetric or asymmetric sizes) are in use, which increase encoding time.

Thus, ME is computationally the most time consuming part of the HEVC [19]. However, compared with natural video, 3D seismic data have low-motion activity. Taking this into account, we first reduce the search window range from the default of 64 all the way to 1.

Next, instead of full search, we use test zone search (TZSearch) [20], which combines diamond search, raster search, and star refinement methods to speed-up encoder with negligible loss in performance by omitting some integer-precision point RDO.

Although fractional sample precision may improve coding performance for natural video, interpolation used there introduce additional calculations. More importantly, in order to make fine MV tuning, RDO has to be performed for every half-sample and quarter-sample precision point. In order to reduce the number of RDO searches, we disable fractional-sample precision estimation.

Additionally, to test the influence of the asymmetric block partitioning on the coding performance, we disable SIZE_2NxnU, SIZE_2NxnD, SIZE_nLx2N, SIZE_nRx2N modes. Hence, four RDO are omitted at each CU level.

To further simplify ME and to relax decoded picture buffer requirements, we avoid the usage of multiple reference pictures. Based on exhaustive research we also avoid the usage of bi-direction ME. In the current setup and with current application requirements it does not justified the price of complexity vs. performance increase, and any other performance improvement for the price of the complexity increase is not needed.

In order to show the importance of using even minimal (truncated) ME, we evaluate performance when no-motion estimation is used. Option to skip all motions is provided in order to use only 1-1 prediction. Prediction in inter coded PU is obtained as a difference between current PU and the corresponding block at the same position in the reference frame. Therefore, residual signal can be computed on the frame level subtracting the whole frame with its reference frame (e.g., in preprocessing stage).

2.2. A new model for the Lagrange multiplier

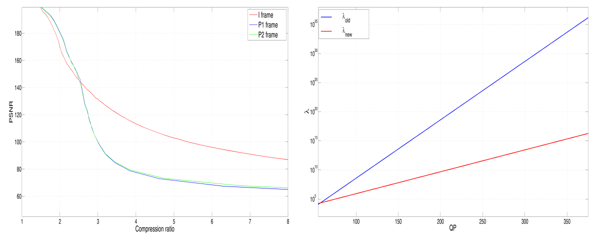
To decide on an optimal quadtree structure, prediction mode, etc., HEVC uses Lagrange optimization [12–15]. The optimal decision is reached by minimizing the cost function given with $J = D + \lambda * R$, where D is the distortion measure between the original and reconstructed samples, R represents the number of bits required to represent various parameters after entropy coding, and λ is the Lagrange multiplier which indicates the slope of the R-D curve. Used distortion measure is the sum of squared differences (SSD), except for ME where in use is the sum of absolute differences (SAD). The Lagrange multiplier plays crucial role in selecting optimal coding parameters. However, all three parameters are subject to the optimization and compromise has to be made. We can give more importance to D or to R with adjusting λ in the cost function J. In HEVC, λ is calculated based on R-D model that is provided for natural images, and is given with the following formula:

$$\lambda(QP) = 0.57 * 2^{\frac{QP-12}{3}} \tag{1}$$

For certain higher values of QP, λ is very high which gives more importance to rate minimization without caring about distortion loss much. As a result, encoder tends to choose parameters that will reduce rate as much as possible. In HEVC, the most rate conservative mode is the SKIP mode. There are two ways how SKIP mode can be selected. First, if the current block completely matches with the reference block so there is no residual data. Second, when there is no perfect match but encoder, based on the cost minimization, choose to force residual block to be all zero. HEVC encoder may do this in order to reduce the number of bits that represent one block for the price of higher distortion, however with smaller overall cost. Since the λ given in (1) is high and does not fit current application, it will enforce rate minimization and hence choose SKIP mode more often. Therefore, as a result of choosing SKIP without making compromise in R-D sense, a huge drop in performance is detected for QP values larger than 51 (see Fig. 1 (a)). The figure shows performance for the first three frames of the same sequence (first I frame, second and third P frames that may be coded using SKIP mode). Naturally P frames should have better performance than I frames which is not case here due to λ non-optimality. From now on we refer to λ given in (1) as λ_{old}.

In order to achieve a good coding performance, proper λ has to be correctly determined for 3D seismic data. In this paper we use a heuristic approach based on our experimental search. Theoretically, λ can go from 0 (highest rate, lowest distortion), to ∞ (lowest rate, highest distortion). We refer to those values as to λ_{min} and λ_{max}, respectively. In our case, λ_{min} corresponds to $QP = 0$ and λ_{max} corresponds to $QP = MAX\_QP$. Using our previously proposed extended QP and quantizer mapping function, we have $MAX\_QP = 400$. However, to maintain the practical meaning of Lagrange optimization, suitable values for λ_{min} and λ_{max} that will balance D and R have been found. We manually tuned them to λ_{min} = 0.0356 and λ_{max} = $1.702 * 10^{18}$. Thus, we were able to reduce the search range from all real positive values $\mathbb{R}^+$ to $\lambda \in (\lambda_{min}, \lambda_{max})$ for any other QP in between. Also, it stands that $\lambda(QP-1) \leq \lambda(QP) \leq \lambda(QP+1)$ so bisection approach can be used to reduce the search range in addition. However, using this approach requires to store experimentally obtained λ for each QP in the look-up table.

To reach a solution that will not require look-up table, and thus relax memory requirements, we compare such obtained values with previous values given with Eq. (1) and propose a



(a) (b)

Fig. 1: (a) Example of performance drop after $QP > 51$ using λ_{old} for the first three frames; (b) λ_{old} compared with redefined λ_{new}.

correction function (CF) to λ_{old}. The purpose of the function is to compensate the large values of λ_{old} when $QP > 51$ and hence provide better compromise between D and R. We refer λ after we apply CF as λ_{new} and assume that it is still in the form of $\lambda_{new}(QP) = k * 2^{\frac{QP-a}{b}}$ (the same form as λ_{old}). We propose a correction function

$$CF(QP) = 3.6842 * 2^{\frac{243-5*QP}{24}}. \tag{2}$$

Hence new recalculated λ is given by

$$\lambda_{new}(QP) = CF(QP) * \lambda_{old}(QP), \tag{3}$$

leading to

$$\lambda_{new}(QP) = 2.1 * 2^{\frac{QP+49}{8}}, \tag{4}$$

where λ_{old} is calculated as in HEVC and given with Eq. (1), $k = 2.1$, $a = -49$, and $b = 8$ are constants.

Fig. 1 (b) shows the comparison between the λ_{old} and the redefined values λ_{new}. It reveals a gap indicating that λ_{old} is very high compared with λ_{new} (the gap is increasing with QP) hence producing low performance. Note that λ_{new} is still suboptimal, and can be further improved.

3. RESULTS

To validate effectiveness of the proposed compression scheme, experiments were performed on the modified HEVC reference software HM15.0 [21] that exploits new λ function in (4). The codec performance has been evaluated using standard R-D curves (PSNR vs. compression ratio), and complexity performance has been evaluated measuring the average time reduction given with $\Delta T = \frac{T_{anchor} - T_{truncated}}{T_{anchor}} * 100\%$.

For the purpose of this research, we use several seismic sequences, that were collected by Shell during real use-case seismic surveys. However, due to nature of the seismic images, sequences are structurally almost identical, and hence performance comparison resulted in the same coding gain. One sequence was randomly selected, and it has been used to evaluate the results presented in this paper. Fig. 2 shows the first two frames of such a 589x236x30 wavefield sample. Notice that they are very similar with low-motion activity. Hence, only the first of the 30 frames is coded as an I-frame. The rest are P-frames, with B-frames not considered. Each P-frame reference to an adjacent previous frame, and only one reference has been used considering the nature of the seismic data.

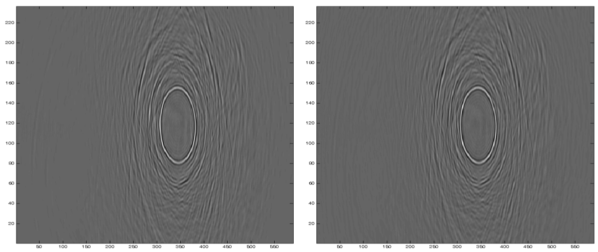


Fig. 2: The first two frames of a 589x236x30 seismic/wavefield sample.

We first compare the performance between an anchor (a version without any complexity reduction modification) and the proposed truncated ME. The proposed simplifications are based on large scale collection of the encoder statistics (e.g. encoder selections of the MV distance, merge candidates, AMVP candidates, etc.), that led us to the following modifications. First, we reduce search range to 1, therefore integer-precision MV search has been reduced significantly. In order to additionally reduce the number of integer-precision points we use modified TZSearch. We use only diamond search without allowing raster search and star refinement. That lead us to only 4 integer-precision points that have to be checked with RDO. In addition, we disable half-precision and quarter-precision MV estimation after integer displacement is selected. Next, we allow the use of median predictor and zero-MV predictor to compete for MV prediction. Temporal predictor and spatial predictors are not used in our setup. Thereafter, within MERGE mode we reduce *NumMergeCands* to 2. Also, the temporal predictor is not used in MV predictor competition scheme. At the end, asymmetric block partitioning is disabled.

The performance difference due to complexity reduction is plotted in Fig. 3. It summarize 3D coding approaches with different levels of complexity. When used all together, simplifications proposed within truncated ME can reduce encoding time by 95% at the targeted compression ratio 10:1 with only 1 dB loss on performance (e.g. from 372s to only 17s to encode one slice). Additionally, for comparison purpose, performance of the 1-1 prediction has been given. Clearly, the 1-1 prediction performance is degraded (compared with truncated ME that uses low-complexity ME). Therefore, the use of truncated ME (minimal low-complexity ME) has been shown to be substantial for performance improvement. At the end, the performance loss of the truncated ME has been analyzed by geologists, and it has been marked as acceptable, with a significant encoding time reduction. Since the time reduction is substantial, we thus only use truncated ME in the following experiments.

Fig. 4 illustrates that our new codec performs around 5 dB better than our previous intra-based codec [1] and 15 dB better than both codecs that have been used in the industry [2, 6]. Also, compared with intra-based coding, truncated ME increases 2x encoding time (e.g. 17s for P-frame vs. 8s for I-frame). However, the overall performance of

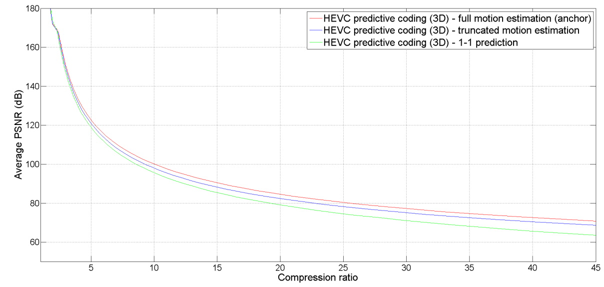


Fig. 3: Performance comparison between an anchor (a version without any complexity reduction modification), the proposed truncated ME and 1-1 prediction.

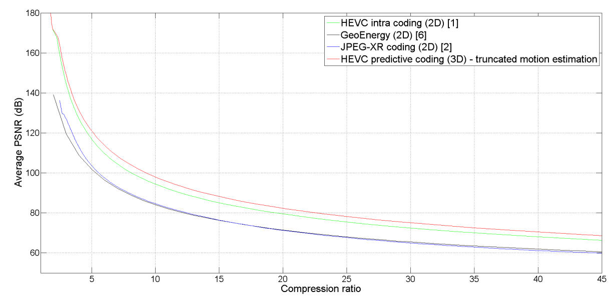


Fig. 4: Performance comparison of our new predictive codec vs. 2D intra coding schemes in terms of PSNR vs. compression ratio.

the new proposed 32 b/p codec shows that truncated ME under the slightly modified HEVC framework is suitable for seismic data compression with additionally improved coding gain. Current emphasis is on GPU acceleration (to make the codec run in real time), after which we plan to provide detailed speed up numbers and complexity comparisons with other methods.

4. CONCLUSIONS AND FUTURE WORK

In this paper we have described a predictive codec for high bit-depth 3D seismic data based on HEVC. We proposed a new model for λ to handle extended bit-depth and QP range. We also focused on complexity reduction in ME for low-motion seismic data. Using truncated ME we were able to significantly reduce the encoding time at acceptable performance loss. We showed that using even minimal ME search our predictive codec outperforms intra coding by a large margin. Future research shall focus on more precise R-D modeling for the seismic data. In addition, proper R-D model should lead to more precise redefinition of λ and its analytical solution. At the end, efficient block size selection using prior knowledge of the quadtree structure from the previous frame and taking into account characteristics of the seismic data will be explored. Additionally, intra/inter mode selection prior to previously coded CUs will be explored, since we believe it could additionally reduce the encoding time due to structural similarity between two consecutive slices.

5. REFERENCES

[1] M. Radosavljević, Z. Xiong, L. Lu, D. Hohl, and D. Vukobratović, "High bit-depth image compression with application to seismic data," Proc. VCIP'16, Chengdu, China, Nov. 2016.

[2] "Recommendation T.832 (01/2012): Information technology - JPEG XR image coding system - Part 2: Image coding specification." ITU-T. January 2012.

[3] Y. Liu, Z. Xiong, L. Lu, D. Hohl, "Fast SNR and Rate Control for JPEG XR," Proc. ICSPCS'16, Golden Coast, Australia, Dec. 2016.

[4] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*, Kluwer , 2001.

[5] Wallace, Gregory K, *The JPEG still picture compression standard*, *IEEE Trans. Consumer Electronic*, vol. 3, pp. 18-34, February 1992.

[6] A. Averbuch, F. Meyer, J.-O. Stromberg, R. Coifman, and A. Vassiliou, "Low bit-rate efficient compression for seismic data", *IEEE Trans. Image Processing*, vol. 10, pp. 1801-1813, December 2001.

[7] G. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, pp. 1649-1668, December 2012.

[8] J. Lainema, F. Bossen, W. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, pp. 1792-1801, December 2012.

[9] J. Liang and T. Tran. "Fast multiplierless approximations of the DCT with the lifting scheme," *IEEE Trans. Signal Processing*, vol. 49, pp. 3032-3044, Dec. 2001.

[10] V. Sze and M. Budagavi, "High throughput CABAC entropy coding in HEVC," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, pp. 1778-1791, Dec. 2012.

[11] Information Technology-JPEG 2000 Image Coding System: "Part 10-Extensions for three-dimensional data," no. 15444-10, ISO/IEC JTC1/SC29/WG1 FDIS, 2007.

[12] Everett III, Hugh, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations research*, vol. 11.3, pp. 399–417, Jun 1963.

[13] Wiegand, Thomas and Girod, Bernd, "Lagrange multiplier selection in hybrid video coder control," Proc. ICIP'01, vol. 3, pp. 542–545, 2001.

[14] Sullivan, Gary J and Wiegand, Thomas, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15.6, pp. 74–90, 1998.

[15] Ohm JR, Sullivan GJ, Schwarz H, Tan TK, Wiegand T., "Comparison of the Coding Efficiency of Video Coding Standards–Including High Efficiency Video Coding (HEVC)," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, pp. 1669–1684, Dec. 2012.

[16] Kim IK, Min J, Lee T, Han WJ, Park J., "Block partitioning structure in the HEVC standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, pp. 1697-1706, Dec. 2012.

[17] K. McCann, W.-J. Han, I.-K. Kim, J.-H. Min, E. Alshina, A. Alshin, T. Lee, J. Chen, V. Seregin, S. Lee, Y.-M. Hong, M.-S. Cheon, and N. Shlyakhov, "Samsungs Response to the Call for Proposals on Video Compression Technology," document JCTVC-A124, Apr. 2010.

[18] T. Sugio and T. Nishi, "Parsing Robustness for Merge/AMVP," document JCTVC-F470, Jul. 2011.

[19] Bossen F, Bross B, Suhring K, Flynn D., "HEVC Complexity and Implementation Analysis," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, pp. 1685-1696. Dec. 2012.

[20] Purnachand N, Alves LN, Navarro A., "Improvements to TZ search motion estimation algorithm for multiview video coding," Proc. IWSSIP'12, pp. 388-391, Apr 2012.

[21] HEVC reference software version HM15.0, https://hevc.hhi.fraunhofer.de/svn/svn.