

RATE-COMPLEXITY TRADE-OFF IN MINIMUM RATE PREDICTORS LIGHT FIELD LOSSLESS ENCODING

João M. Santos^{*,†}, Pedro A. A. Assuncao^{*,‡}, Luis A. da Silva Cruz^{*,†},
Luis M. N. Tavora[‡], Rui Fonseca-Pinto^{*,‡} and Sergio M. M. Faria^{*,‡}

^{*}Instituto de Telecomunicações, Portugal

[†]University of Coimbra, Coimbra, Portugal

[‡]ESTG, Politécnico de Leiria, Leiria, Portugal

e-mails: {joao.santos, amado, luis.cruz, sergio.faria}@co.it.pt, {luis.tavora, rui.pinto}@ipleiria.pt

ABSTRACT

The Multi-Reference Minimum Rate Predictors (M-MRP) is currently one of the most efficient algorithms for Light Field lossless encoding, but its high computational complexity is still a concern. This article describes an experimental study on the trade-off between the encoding performance and computational complexity of M-MRP Light Field encoding, taking the HEVC standard as a benchmark reference. The results show that even with an average complexity reduction of approximately 35 times, M-MRP still achieves better compression ratios than other state-of-the-art lossless standard encoding schemes, such as HEVC or JPEG-LS. Furthermore, for a similar computational complexity, M-MRP outperforms the HEVC reference software by 13.5%, on average.

Index Terms — Lossless Compression, Light Field Coding, Minimum Rate Predictors

1. INTRODUCTION

Light Field (LF) imaging has been attracting increased attention from different research domains. One of such domains is light field image compression, currently under standardization in the ISO/IEC JPEG group JPEG-Pleno [1]. The characteristics of plenoptic images (*i.e.*, light fields) enable several post-processing operations, such as image rendering with different focal planes, depth-of-field or viewing perspectives [2, 3, 4], and the extraction of depth maps [5]. These operations have a multitude of possible applications, ranging from medical applications [6, 7] to 3D recording of cultural artefacts [8].

LF coding deals with the problem posed by the large amount of data that is necessary to represent the whole visual information captured from real-world scenes (*e.g.* light intensity, colour and direction). Although, conventional image/video encoders such as HEVC [9] can be used for compression of LF data, they are not very efficient because they do not take into account the intrinsic redundancy of LF images. Furthermore, applications with strict precision requirements may not tolerate the artefacts resulting from irreversible compression, therefore requiring the use of lossless compression. For instance, this might be the case of medical imaging or industrial applications where highly accurate depth measurements are required.

This work was supported in part by the Fundação para a Ciência e a Tecnologia (FCT), Portugal, under PhD Grant SFRH/BD/114894/2016, FCT/MEC through national funds, under project PlenoISLA PTDC/EEI-TEL/28325/2017. and when applicable co-funded by FEDER – PT2020 partnership agreement under the project UID/EEA/50008/2019.

In the context of lossless coding, in [10] the authors performed a systematic study of the influence of various LF data arrangements in the compression efficiency of several state-of-the-art encoders, which has been further expanded in [11] to include several reversible colour transforms. Recently, the authors focused on exploiting the characteristics of a Minimum Rate Predictors (MRP) encoder for lossless compression of LF images in Micro Image (MI) array and Sub-Aperture Image (SAI) array data formats using multiple references for each sub image encoding [12].

Other techniques that can be found in the literature usually fall into one of the following two approaches: either compression of lenslet (or Micro Image (MI) array) such as [13, 14] or compression of Sub-Aperture Images (SAI), in array or Pseudo Video Sequence (PVS) [15, 16].

Despite the good compression efficiency presented by MRP in previous works in comparison with other state-of-the-art encoders, the computational complexity is still a concern. In this work, the performance of the MRP algorithm is further analysed across the whole range of operational points in terms of coding rate and computational complexity. This simulation study uses standard image and video encoders for comparison and analysis.

2. MULTI-REFERENCE MINIMUM RATE PREDICTORS

The Minimum Rate Prediction (MRP) algorithm minimises the coding rate of the prediction error, rather than minimising any rate-distortion cost, as implemented in most lossy encoders. The MRP algorithm estimates the amount of information conveyed by prediction errors and then designs the predictors with the goal of minimising the rate. To minimise the encoding rate, MRP performs the search for linear predictors, seeking to minimize the number of bits required to encode the prediction residues and auxiliary information.

In [12], the MRP algorithm was extended to exploit the inherent redundancy of the LF images. This is achieved by using inter prediction specifically designed to exploit multiple references, such as the neighbouring MI or SAI. Since MI or SAI are specific data structures with intrinsic high cross-correlations due to small disparities of respective image data, efficient prediction is achieved between spatially close MI or SAI images. To accomplish such prediction approach the M-MRP prediction uses the nearest four MI or SAI, as represented in Figure 1.

3. EXPERIMENTAL APPROACH

In this work a Multi-Reference Minimum Rate Predictors (M-MRP) encoder is used for the lossless compression of lenslet and High-

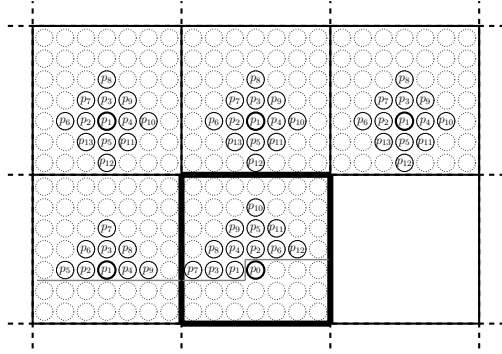


Figure 1. Example of the proposed M-MRP prediction using multiple neighbouring SAI as reference for the prediction.

Density Camera Array light fields, respectively from the EPFL and HCI datasets, both described in the JPEG Pleno Common Test Conditions (CTC) [17]. The purpose is to assess the coding efficiency and computational complexity of M-MRP, using multiple-references at the encoder for efficient lossless compression of LF, both in MI and SAI array formats.

The experimental approach followed in this study was defined to establish a relationship between the input parameters of M-MRP and their impact on the rate-complexity performance. This is achieved by varying three main encoder parameters, namely, the number of optimisation iterations, the number of prediction modes, *i.e.* classes, and the number of reference pixels. The HEVC Reference Software (HM) [18] encoding time is used as a reference for the comparative study of the trade-off between M-MRP coding efficiency and computational complexity. The x265, which is an optimised implementation of the HEVC standard, is also used for comparison. The analysis is based on both compression efficiency, measured in bits-per-pixel (bpp), and complexity, measured by the processing time.

4. RESULTS

The experiments were performed according to the JPEG Pleno CTC [17], with the appropriate adaptations for lossless encoding. Figure 2 shows the pre-processing stage performed to obtain the images to be encoded, *i.e.* the SAI Array. In the case of the HCI dataset, the pre-processing starts at the 4DLF stage.

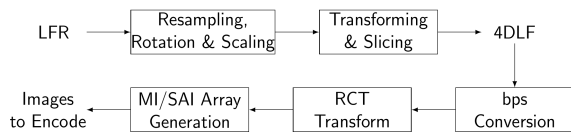


Figure 2. Lytro camera pre-processing diagram.

The CTC defines four operations for the pre-processing of the Lytro dataset. However, as this work deals with lossless coding, the *Color Correction* and *Rectification* steps are skipped as they can be performed after the decoding process. Once the 4DLF is obtained from the LFR, the sample value range is converted to the bits-per-sample (bps) as defined in the CTC, and a Reversible Colour Transform (RCT) [11] is applied to decorrelate the RGB colour components.

The three input parameters were tested with multiple values resulting in a set of operation points defined by coding efficiency (bpp) versus computational complexity (in minutes), as shown in Figures 3 to 5, for the Stone Pillars Outside LF, and in Figures 6 to 8, for the Greek LF. In these figures, the red dashed line represents the limits of HM PVS (pseudo video sequence formed from the

SAI), whose intersection is represented by a black cross, and the x265 reference is represented by a black dot.

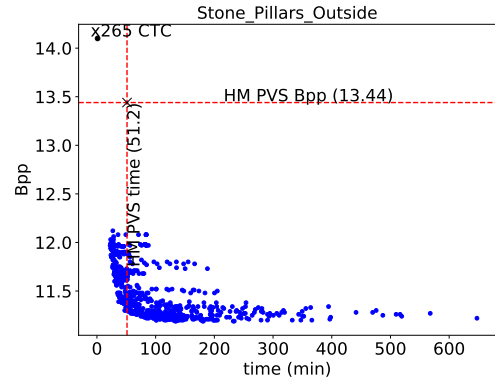


Figure 3. Operational points for Stone Pillars Outside image of EPFL dataset.

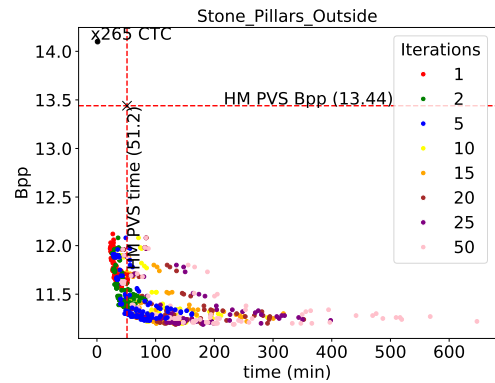


Figure 4. Operational points with the number of iterations sorted by colour.

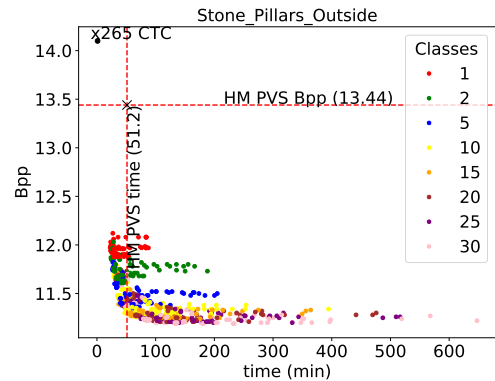


Figure 5. Operational points with the number of classes sorted by colour.

Figures 3 and 6 show all the resulting operational points when varying the previously described parameters. For an easier analysis in Figures 4 and 7 the number of iterations is represented by different colours, and the same is done for the number of classes in Figures 5 and 8. The figures show that with a proper selection of the coding parameters, M-MRP can achieve competitive coding efficiency, while maintaining a similar complexity to HM.

Considering the number of classes for both datasets, it is observed in Figures 5 and 8 that using just one class generally leads to lower compression ratios. This was expected, as using just one class greatly reduces the adaptability of the encoder. For the remaining number of classes the analysis is not straightforward. However, considering the encoding time of HEVC as reference it can be seen that for both cases the highest compression efficiency is obtained for ten to fifteen classes.

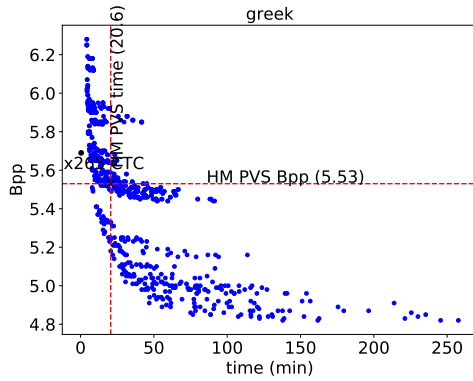


Figure 6. Operational points for Greek image of HCI dataset.

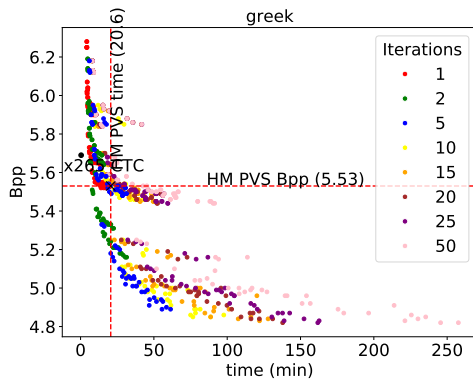


Figure 7. Operational points with the number of iterations sorted by colour.

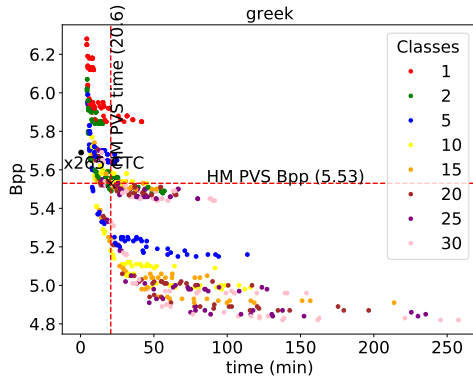


Figure 8. Operational points with the number of classes sorted by colour.

Regarding to the number of iterations, the results show that using a lower number of iterations or classes (1 or 2), generally leads to lower compression efficiency, as can be seen for the Greek image. Once again, considering the encoding time of HEVC as a benchmark, using five iterations provide the highest compression efficiency for both images.

The colour representation of the parameters and visual analysis of the number of reference pixels does not result in sharp conclusions, unlike the case of the number of classes and the number of iterations. Nevertheless, the analysis of the results led to the selection of the following encoding parameters:

- EPFL dataset:
 - Iterations: 5
 - Classes: 10
 - References¹: 6, 5, 5, 5, 5
- HCI dataset:
 - Iterations: 5
 - Classes: 15
 - References: 6, 13, 13, 13, 13

¹The references parameter represents the number of references in the

These M-MRP parameters are selected such that, for a similar complexity, M-MRP achieves higher coding efficiency than HM, originating the results shown in Tables 1 to 3², with bit-rate and complexity results. In these tables, M-MRP stands for the encoding results with the selected parameters and M-MRP *Optimal* the highest compression ratio for M-MRP (non-constrained).

The experimental results show that both M-MRP and M-MRP *Optimal* present better compression ratios than the other encoders. Although there is a slight loss of compression efficiency in M-MRP when compared to M-MRP *Optimal*, as expected, because the parameter values have been decreased to speed up the encoding time. Thus, M-MRP and M-MRP *Optimal* achieve a compression efficiency that is, on average, 13.5% and 15.2% higher than that of HM (HEVC), respectively.

In regard to the encoding and decoding time, Table 3 shows that the encoding time of M-MRP (*i.e.* 105.5%) is quite similar to HEVC, with M-MRP being just 5.5% slower, on average (note that 100% corresponds to the encoding time of HEVC). This means that the encoding time of M-MRP *Optimal* can be reduced by 27 to 48 times, while still achieving an average compression efficiency gain of 13.5% in comparison with the HEVC HM Reference implementation (see M-MRP column in Table 2). Additionally, table 3 shows that the decoding time of M-MRP is just 0.92 times that of HM, on average. In comparison, M-MRP *Optimal* average decoding time is 1.67 times higher than that of HM, which shows that the change of the input parameters have a much lower influence on the decoding process.

5. CONCLUSIONS

This work studies the rate-complexity trade-off of the multi-reference MRP algorithm³. These are novel results which establish a relationship and trade-off between LF lossless coding efficiency and computational complexity between a non-standard, yet more efficient encoder, and the standard HEVC Reference Encoder. Since these results are obtained using LF from JPEG-Pleno CTC, they can be used as benchmark for further studies where both lossless encoding efficiency and computational complexity are jointly considered in the overall performance.

6. REFERENCES

- [1] WG1, “JPEG Pleno Call for Proposals on Light Field Coding,” in *Doc. ISO/IEC JTC 1/SC29/WG1/N74014*, 74th JPEG Meeting, Geneva, Switzerland, Jan. 2017.
- [2] M. Harris, “Focusing on everything,” *IEEE Spectr.*, vol. 49, no. 5, pp. 44–50, May 2012.
- [3] D. Donatsch, S. A. Bigdeli, P. Robert, and M. Zwicker, “Hand-held 3D light field photography and applications,” *The Visual Computer*, vol. 30, no. 6-8, pp. 897, June 2014.
- [4] D. G. Dansereau, O. Pizarro, and S. B. Williams, “Linear volumetric focus for light field cameras,” *ACM Transactions on Graphics (TOG)*, vol. 34, no. 2, pp. 1–20, Mar. 2015.
- [5] H. G. Jeon, J. Park, G. Choe, J. Park, Y. Bok, Y. W. Tai, and I. S. Kweon, “Accurate depth map estimation from a lenslet light field camera,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015, pp. 1547–1555.

SAI and in the neighbour SAIs.

²The computational complexity was measured as the running time of the encoding and decoding operations for the codecs, using a dedicated server running Ubuntu Server 18.04 with an Intel Xeon Silver 4114 @ 2.20GHz CPU and 62.6 GB of RAM.

³Binaries available at: <https://tinyurl.com/MRP-ITLeiria>.

Table 1. Compression results (in bpp) for the JPEG-Pleno CTC using the selected images (YUV444).

Type	Light Fields	HM	x265	JPEG-LS	M-MRP	M-MRP Optimal
LL	Bikes	13.69	14.59	16.42	11.55	11.47
	Danger_de_Mort	13.03	13.75	16.18	10.93	10.80
	Fountain_and_Vincent_2	14.43	15.54	16.66	12.24	12.06
	Stone_Pillars_Outside	13.44	14.10	16.77	11.30	11.20
HDCA	Greek	5.53	5.69	7.25	5.09	4.77
	Sideboard	8.16	8.75	12.19	7.99	7.63
Average		11.38	12.07	14.24	9.85	9.65

Table 2. Bit rate savings for the JPEG-Pleno CTC using the selected images (YUV444, 11 bps and 9 bps) when compared to HM.

Type	Light Fields	x265	JPEG-LS	M-MRP	M-MRP Optimal
LL	Bikes	6.6%	19.9%	-15.6%	-16.2%
	Danger_de_Mort	5.5%	24.2%	-16.1%	-17.1%
	Fountain_and_Vincent_2	7.7%	15.5%	-15.2%	-16.4%
	Stone_Pillars_Outside	4.9%	24.8%	-16.0%	-16.6%
HDCA	Greek	2.9%	31.0%	-7.9%	-13.8%
	Sideboard	7.3%	49.4%	-2.1%	-6.5%
Average		6.1%	25.2%	-13.5%	-15.2%

Table 3. Encoding and decoding processing time relative to HM (in %) for the JPEG-Pleno CTC using the selected images.

Type	Light Fields	x265		JPEG-LS		M-MRP		M-MRP Optimal	
		Enc.	Dec.	Enc.	Dec.	Enc.	Dec.	Enc.	Dec.
LL	Bikes	2.1%	94.8%	0.6%	150.2%	106.5%	85.6%	3202.4%	146.1%
	Danger_de_Mort	2.0%	95.2%	0.6%	155.5%	106.1%	87.5%	3205.7%	148.9%
	Fountain_and_Vincent_2	2.0%	96.8%	0.6%	149.8%	101.2%	83.7%	3185.9%	143.5%
	Stone_Pillars_Outside	2.1%	95.1%	0.6%	161.5%	102.4%	89.3%	4251.1%	152.8%
HCDA	Greek	1.9%	104.2%	0.7%	193.3%	114.6%	143.2%	4795.1%	307.8%
	Sideboard	1.5%	101.8%	0.7%	186.7%	109.8%	117.3%	2706.3%	271.2%
Average		2.0%	96.5%	0.6%	159.0%	105.5%	92.2%	3494.4%	166.7%

- [6] N. Bedard, T. Shope, A. Hoberman, M. A. Haralam, N. Shaikh, J. Kovačević, N. Balram, and I. Tošić, "Light field otoscope design for 3D in vivo imaging of the middle ear," *Biomedical Optics Express*, vol. 8, no. 1, pp. 260–272, 2017.
- [7] D Piccolo, A Ferrari, K. Peris, R Daidone, B Ruggeri, and S Chimenti, "Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study," *British Journal of Dermatology*, vol. 147, no. 3, pp. 481–486, 2002.
- [8] X. Huang, E. Uffelman, O. Cossairt, M. Walton, and A. K. Katsaggelos, "Computational Imaging for Cultural Heritage: Recent developments in spectral imaging, 3-D surface measurement, image relighting, and X-ray mapping," *IEEE Signal Process. Mag.*, vol. 33, no. 5, pp. 130–138, Sept. 2016.
- [9] G. J. Sullivan, J. M. Boyce, Y. Chen, J. R. Ohm, C. A. Segall, and A. Vetro, "Standardized extensions of high efficiency video coding (HEVC)," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1001–1016, Dec. 2013.
- [10] J. M. Santos, P. A. A. Assuncao, L. A. da Silva Cruz, L. Távora, R. Fonseca-Pinto, and S. M. M. Faria, "Lossless coding of light field images based on minimum-rate predictors," *Journal of Visual Communication and Image Representation*, vol. 54, pp. 21–30, 2018.
- [11] J. M. Santos, P. A. A. Assuncao, L. A. da Silva Cruz, L. Távora, R. Fonseca-Pinto, and S. M. M. Faria, "Lossless light-field compression using reversible colour transformations," in *Int. Conf. Image Processing Theory, Tools and Applications (IPTA)*, Nov. 2017, pp. 1–6.
- [12] J. M. Santos, P. A. Assunção, L. A. da Silva Cruz, L. M. Távora, R. Pinto, and S.M.M. Faria, "Lossless compression of light fields using multi-reference minimum rate predictors," in *IEEE Data Compression Conf. - DCC*, March 2019.
- [13] I. Tabus and P. Helin, "Microlens image sparse modelling for lossless compression of plenoptic camera sensor images," in *European Signal Processing Conference (EUSIPCO)*, Aug. 2017, pp. 1907–1911.
- [14] A. Miyazawa, Y. Kameda, T. Ishikawa, I. Matsuda, and S. Itoh, "Lossless coding of light field camera data captured with a micro-lens array and a color filter," in *Proc. Int. Workshop Advanced Image Technology (IWAIT)*, Jan. 2018, pp. 1–4.
- [15] P. Helin, P. Astola, B. Rao, and I. Tabus, "Sparse modelling and predictive coding of subaperture images for lossless plenoptic image compression," in *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, July 2016, pp. 1–4.
- [16] P. Helin, P. Astola, B. Rao, and I. Tabus, "Minimum description length sparse modeling and region merging for lossless plenoptic image compression," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 7, pp. 1146–1161, Oct. 2017.
- [17] WG1, "JPEG Pleno Light Field Coding common Test Conditions," in *Doc. ISO/IEC JTC 1/SC29/WG1/N81022*, 81st JPEG Meeting, Geneva, Switzerland, Mar. 2019.
- [18] "HEVC reference software encoder HM-16.12+RExt," <https://hevc.hhi.fraunhofer.de/>, [Online; accessed 22 April 2019].