Real-time lossless and lossy compression of MBES water column

Jordi Portell, Dept. Quantum Physics and Astrophysics, Institute of Cosmos Sciences - ICCUB, University of Barcelona - IEEC-UB (Spain) and DAPCOM Data Services (Spain), jportell@fqa.ub.edu
David Amblas, Scott Polar Research Institute, University of Cambridge (UK) and CRG Marine Geosciences, Dept. of Earth and Ocean Dynamics, University of Barcelona (Spain), damblas@ub.edu

Multibeam echosounders are excellent for the efficient and accurate mapping of the seafloor, also allowing to collect water column acoustic images [1]. The latter, however, leads to vast amounts of data, for which existing compression algorithms do not provide good enough results. In this work we present a new data compression solution, FAPEC, inherited from the Gaia space astrometry mission.

FAPEC is a highly optimized and versatile two-stage data compression system [2]. The first stage can be adapted to the specific kind of data being handled, aiming at the output of small signed values. These are then compressed by the second stage, an outlier-resilient entropy coder which can outperform an optimum Huffman coder. We have implemented a tailored pre-processing stage for Kongsberg MBES water column data. It arranges backscatter samples in a two-dimensional matrix and codes the differences between neighbour values. It provides lossless compression, allowing to recover the original data file without any loss. The user can also select lossy compression with different quality levels, allowing to achieve a larger reduction in data volume at the cost of some degradation in the image quality. This is achieved by a non-biased quantization of backscatter data samples, reducing the number of grey shades in the reconstructed image but keeping the full image resolution. It is a very interesting option which still allows detecting fish shoals, gas seeps or sunken structures, for example. FAPEC, including this tailored water column stage, is fully operational and provides block-based operation, multi-threading and encryption. Its ANSI C implementation makes it portable, including ARM-based computing platforms such as those present in AUVs.

Test setup and results obtained

We have tested FAPEC on Kongsberg EM302, EM710 and EM2040 water column files. EM2040 files, kindly provided by Kongsberg, were collected in a multibeam survey inside the harbour of Barcelona, presenting several structural elements and shoaling fish. EM302 data, kindly provided by Fugro, were collected in the Gulf of Mexico at depths around 1000m and present gas seeps. For comparison, lossless compression with *gzip*, *bzip2*, *7-zip* and *Zstandard* has also been tested. Lossy FAPEC compression has been tested with option 1 (128 levels or shades of grey) to 5 (just 8 shades of grey, including black and white).

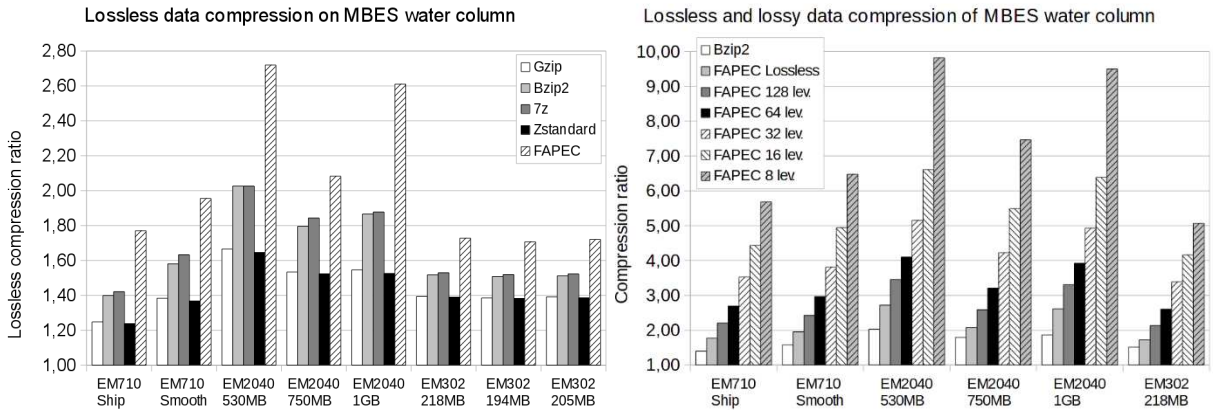

Figure 1: Lossless and lossy compression ratios of FAPEC and other solutions on Kongsberg MBES water column data.

As can be seen in the left panel of Figure 1, FAPEC provides the best lossless compression ratios on all this variety of sonar models and scenes. Furthermore, lossy compression achieves excellent ratios, as seen in the right panel. Besides this, FAPEC compresses at a speed only comparable to that of

Zstandard (which gives much worse ratios), exceeding 60 MB/s even in single-thread mode on a typical desktop computer (Intel® Core™ i5 3.1 GHz). For comparison, *gzip* just reaches 20 MB/s, whereas *bzip2* and specially *7-zip* are even slower. Compressed file sizes with 32 shades of grey or less are very similar (even smaller) than raw bathymetry files (Figure 2). Together with the high compression speed, it means that continuous water column acquisition is finally feasible with FAPEC.

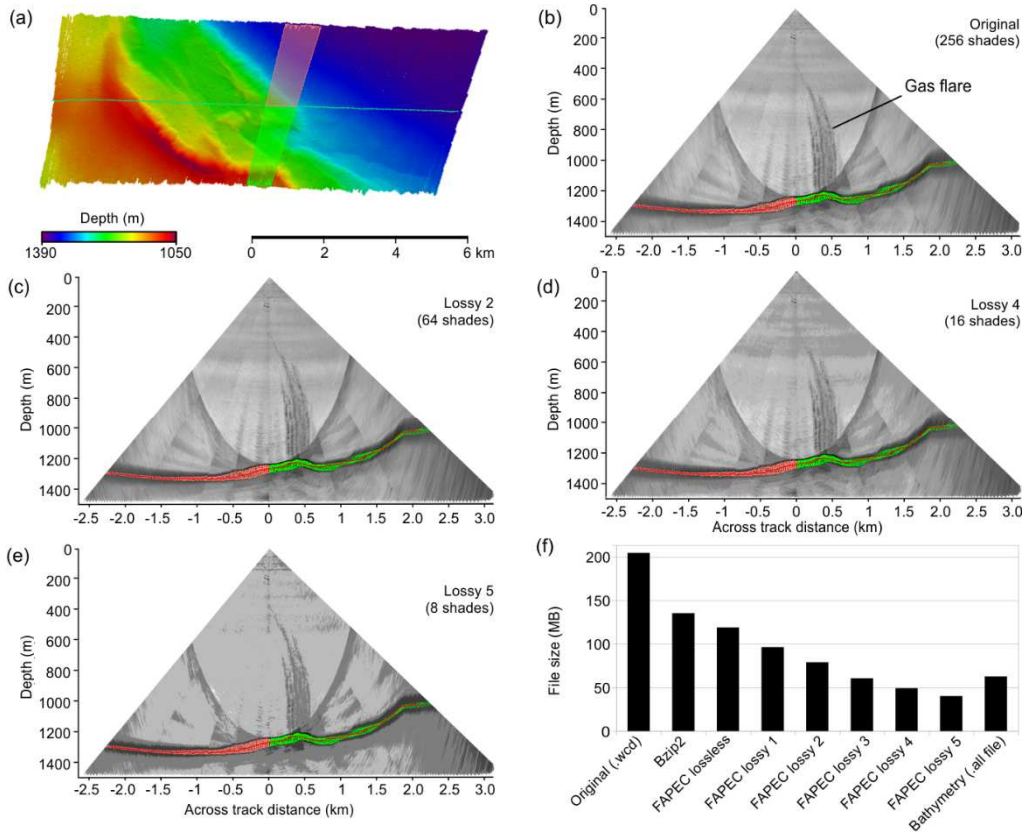

Figure 2: EM302 scene (a) showing the original water column (b) and that obtained with different levels of FAPEC lossy compression (c), (d) and (e). Panel (f) shows the corresponding file sizes, including bzip2 compression for comparison. Features in the water column such as gas seeps can still be detected even with low-quality compression levels. Actually, such levels seem to provide an even clearer evidence of the gas-related features.

Conclusions and forthcoming work

We present a new data compression solution for Kongsberg MBES water column files, which could also be extended to other manufacturers. It provides significantly better lossless ratios and speeds than any other existing solution, which makes it very interesting for massive data archiving or transfer. Lossy compression provides excellent ratios still with very good quality, meaning that it could be used in real-time to allow continuous water column acquisition. The block-based operation of FAPEC provides resilience in front of data corruption, minimizing data loss in such case. It also allows for quick detection of features in the water column from variations in the ratios.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 658358 (D. Amblas), MINECO (Spain) through ESP2016-80079-C2-1-R (MINECO/FEDER, UE) and ESP2014-55996-C2-1-R (MINECO/FEDER, UE), and MDM-2014-0369 of ICCUB (Unidad de Excelencia 'María de Maeztu').

References

[1] K.Colbo, T.Ross, C.Brown, T.Weber, A review of oceanographic applications of water column data from multibeam echosounders, Estuarine, Coastal and Shelf Science 145 (2014) 41-56.

[2] J.Portell, R.Iudica, E.García-Berro, A.G.Villafranca, G.Artigues, FAPEC, a versatile and efficient data compressor for space missions, International Journal of Remote Sensing 39 (7) (2018) 2022-2042.