Impact of GDAL JPEG 2000 Lossy Compression to a Digital Elevation Model

Jan Růžička and Kateřina Růžičková

Abstract Mobile platforms bring possibilities to use distributed services in several users' environments. One of the conditions of the mobile platform is to transfer as less as possible amount of data via network. When dealing with geodata in mobile environment it is always necessary to use compression algorithms to lower their size. Digital elevation models in a GRID form are usually compressed with LZW or DCT technique, but the best compression ratio is now available with wavelet algorithms. The compression can be loss or loss less. The paper describes results of tests of JPEG 2000 loss compression technique. We studied impact of JPEG 2000 loss compression on quality of DEM. We have tested several options of OpenJPEG library to find how these options can change resulting compressed DEM. The main part of the research was focused on pre-filtering of DEM before compression. We have find out a way how to minimise average error of loss compression when keeping the same compression ratio.

Keywords Compression · JPEG 2000 · Filtering · Digital elevation model

1 Introduction

The issue of influence of raster compression on data quality had been solved in many articles. For example Vatolin et al. (2005) deal with comparing several software tools for compression data according to JPEG 2000 format. As an introduction on compressing DEM can be given in thesis of Inanc (2008), witch provides overview of compressions techniques applied for DEM. Mittal et al. (2013)

J. Růžička · K. Růžičková (✉)
Institute of Geoinformatics, HGF, VSB-Technical University of Ostrava,
17. listopadu 15, 708 33 Ostrava, Czech Republic
e-mail: katerina.ruzickova@vsb.cz

J. Růžička
e-mail: jan.ruzicka@vsb.cz

K. Růžičková and T. Inspektor (eds.), *Surface Models for Geosciences*,
Lecture Notes in Geoinformation and Cartography,
DOI 10.1007/978-3-319-18407-4_17

studied, how data for DEM creation can be changed via compression of source satellite images. Gladkova and Grossberg (2006) and Gladkova et al. (2007) focused their research on analysing several compression techniques including JPEG 2000 to MODIS and hyperspectral data.

A lot of papers describe technical specification prepared by software producers. There are also results of several compression ratios and their impact to error in the resulting DEM published by Microimages, Inc. (2009). They analysed impact of loss JPEG 2000 compression to ASTER data, which is quite high.

Ben-Moshe et al. (2007) describes a new technique for simplification of DEM based on analyses of DEM before compression. It is named Image Compression Terrain Simplification (ICTS). The ITCS technique is compared with another techniques including JPEG 2000 compression technique. The paper shows results that ICTS gives similar or even better results for simplification of DEM than JPEG 2000. What is missing in the paper are parameters used for JPEG 2000 compression, mainly compression ratio.

2 Methods and Software

2.1 Methods

The research was focused on two basic ideas:

- How sensitive is the error in the compressed data to parameter BLOCKSIZE and can we use it to reduce the impact of the compression to the DEM quality.
- Can the mean filtering of the data before compression produce less error in the compressed data.

The comparison was based on the simple procedure:

- The original DEM was filtered by mean filter.
- The original DEM and the result of filtering were compressed by loss wavelet compression and stored to JPEG 2000 format (with different parameters described later).
- The compressed files were uncompressed and compared with original data using simple map algebra (difference = original_dem − compressed_dem).
- For the difference layers was calculated value that represents average error realised by data compression.

Used mean filter has the following values:

```
MATRIX 3
1 1 1
1 1 1
1 1 1
DIVISOR 9
```

The compression has been done in four ratios:

- 200:1
- 100:1
- 20:1
- 10:1

The BLOCKSIZE parameter has been specified in three values:

- 512
- 1024
- 2048

3 GDAL

The GDAL allow to use three different open source libraries for wavelet compression according to JPEG 2000 specification. Unfortunately the only one was used for our research purposes. The following table shows problems with each library (Table 1).

The used library was OpenJPEG. There were problems with several tested files, but most of the tested files were compressed correctly. This is probably a bug in the library and should be corrected in the future.

The OpenJPEG library was build from SVN repository to have a latest version of the library.

The basic command for compression to JPEG2000 with OpenJPEG library was:

gdal_translate -of “JP2OpenJPEG” -co “QUALITY = 1” -co “BLOCKSIZEX = 1024” -co “BLOCKSIZEY = 1024” demfilter1.tif demfilter1.tif.j2k

The parameter QUALITY has been set to values: 10, 5, 1, 0.5 to reach the specified compression ratios. The parameters BLOCKSIZE has been set to values: 512, 1024, 2048. Other parameters were used with default values, because they do not have an impact to a quality of a compression.

Table 1 Libraries for GDAL

Library	Problem
Libjasper	The compression ratio can be specified, but it has no effect between 10:1 and 500:1. The library specifies the ratio itself
Kakadu	Compilation process is quite complicated and has not been finished by author of the paper
OpenJPEG	Some of the tested files were compressed in the results with a lot of noise that was not possible to filter out. This is probably a bug in the library

4 Grass GIS

GRASS GIS was used for filtering the original DEM with mean filter, for calculating with map algebra and for counting average error.

The command for filtering the data was:

r.mfilter input = dem output = demfilter1 filter = filter1.txt

The command for map algebra was:

r.mapcalc 'dem_dem.tif.j2k.tif = dem-dem.tif.j2k.tif'

5 Data

The library was tested on five tiles from ASTER DEM version 2. Each tile has a resolution 3601 × 3601 pixels and spatial resolution about 0.01°. In the following table are listed basic characteristics for selected tiles.

From Table 2. is obvious that mainly the flat areas were selected. That selection was made by expectations that for the flat areas the filtering should produce the best results.

The data were available in GeoTIFF format without compression, with encoding of values using range of UInt16 domain. The size of each tile in original format was about 28 MB.

6 Results

The following five tables show the results of calculations. The results are discussed in the chapter discussion and conclusion. The results are based on average error that was counted as sum of errors for each individual pixel divided by number of pixels.

The average error can be used for several purposes, but for other several purposes can be important distribution of the error in the whole DEM and maximal error. The following two tables show distribution of error for compression ratio

Table 2 Libraries for GDAL

Tile	Minimal elevation	Maximal elevation
N23E026	409	1098
N33E081	4342	6375
N49E017	88	1163
N51E021	1	341
S24E125	295	477

100:1 for the tile N23E026 with BLOCKSIZE 1024. The maximum error for not filtered data was 47 m and for filtered data it was 53 m.

The following two tables show distribution of error for compression ratio 20:1 for the tile N49E017 with BLOCKSIZE 1024. The maximum error for not filtered data was 6 m and for filtered data it was 33 m.

7 Discussion and Conclusion

From Tables 3, 4, 5, 6 and 7 is obvious that the BLOCKSIZE 1024 gives the smallest error for all levels of compression and for all tested tiles. So there is not need for ASTER DEM data to use another than default BLOCKSIZE value that is 1024. For other discussion we used only results with BLOCKSIZE = 1024.

Table 3 Impact of compression to tile N23E026

Filtered	Compression ratio	Block size (pixels)	Average error (m)
No	200:1	512	5.8413434788
Yes	200:1	512	5.3532582706
No	200:1	1024	5.4606829955
Yes	200:1	1024	4.8807861465
No	200:1	2048	5.4813481336
Yes	200:1	2048	4.885607773
No	100:1	512	4.0456516406
Yes	100:1	512	3.1500869
No	100:1	1024	3.7438103258
Yes	100:1	1024	2.795019141
No	100:1	2048	3.7461034189
Yes	100:1	2048	2.7379252469
No	20:1	512	0.7462615101
Yes	20:1	512	0.8957982528
No	20:1	1024	0.4930430245
Yes	20:1	1024	0.6542141207
No	20:1	2048	0.609507017
Yes	20:1	2048	0.7624123356
No	10:1	512	0.3767458374
Yes	10:1	512	0.8122391255
No	10:1	1024	0.0963257221
Yes	10:1	1024	0.5737475651
No	10:1	2048	0.2021722344
Yes	10:1	2048	0.6926780112

Table 4 Impact of compression to tile N33E081

Filtered	Compression ratio	Block size (pixels)	Average error (m)
No	200:1	512	8.7217949348
Yes	200:1	512	8.6787087668
No	200:1	1024	8.3779604403
Yes	200:1	1024	8.3703773081
No	200:1	2048	8.3392846305
Yes	200:1	2048	8.3425146259
No	100:1	512	6.7825648727
Yes	100:1	512	6.8121760432
No	100:1	1024	6.6063102592
Yes	100:1	1024	6.6317583108
No	100:1	2048	6.6271325631
Yes	100:1	2048	6.6606450382
No	20:1	512	3.732998432
Yes	20:1	512	2.9196588377
No	20:1	1024	3.5608164784
Yes	20:1	1024	2.7426009669
No	20:1	2048	3.6206660944
Yes	20:1	2048	2.8138112458
No	10:1	512	1.7225701985
Yes	10:1	512	2.6346513022
No	10:1	1024	1.5241541332
Yes	10:1	1024	2.4730518174
No	10:1	2048	1.6048564374
Yes	10:1	2048	2.5526437047

Tables 3, 4, 5, 6 and 7 show that for compression ratio 200:1 the filtered data give smaller error than original data. The difference between errors is between 0.007 and 0.579 m.

For the compression ratio 100:1 are the results similar (except the tile with highest elevation). The difference between errors is between −0.025 and 0.948 m. The result for the tile N33E081 is close to zero, but what is more interesting is that the result for tile N23E026 is close to 1 m. For the tile N23E026 is the average error for not filtered data about 4 m and for filtered data 3 m. That is improve in 25 % and it could be interesting for several applications. When considering the distribution of the error we can see in the Tables 8 and 9 that the distribution is not so different.

For the compression ratio 20:1 are the results (except the tile N23E026, where are the filtered data with higher error) even better when comparing average error. The difference between errors is between −0.161 and 1.065 m. For two tiles

Table 5 Impact of compression to tile N49E017

Filtered	Compression ratio	Block size (pixels)	Average error (m)
No	200:1	512	7.2025538896
Yes	200:1	512	7.1445872552
No	200:1	1024	6.9104690365
Yes	200:1	1024	6.8661799104
No	200:1	2048	6.9288282028
Yes	200:1	2048	6.8633231643
No	100:1	512	6.1518125616
Yes	100:1	512	5.836271914
No	100:1	1024	5.9580645816
Yes	100:1	1024	5.6325006453
No	100:1	2048	6.0004222962
Yes	100:1	2048	5.6750086622
No	20:1	512	3.2997735595
Yes	20:1	512	2.2283458859
No	20:1	1024	3.0743997876
Yes	20:1	1024	2.0134329683
No	20:1	2048	3.1570105993
Yes	20:1	2048	2.1086771154
No	10:1	512	1.1898974189
Yes	10:1	512	2.0523082044
No	10:1	1024	0.9489432608
Yes	10:1	1024	1.8444519369
No	10:1	2048	1.0514626094
Yes	10:1	2048	1.9479006302

(N49E017 and S24E125) is the difference between errors about 1 m and that is improve in 33 % and it could be interesting for several applications. But we have to consider the distribution of the error and from Table 11 is obvious that the number of pixels with error higher than 3 m is enormous in comparison to the results of not filtered data described in the Table 10. Also the maximum error is 33 m for filtered data.

For the compression ratio 10:1 are all the results better for not filtered data.

The test should be done for more tiles of ASTER data and for another DEM data as well. Even from presented results we can conclude, that when are DEM data (from ASTER source) compressed into JPEG 2000 format with GDAL tool and OpenJPEG library that the user should consider filtering the data. When the compression ratio is from 20:1 to 200:1 then the filtering can improve the average error

Table 6 Impact of compression to tile N51E021

Filtered	Compression ratio	Block size (pixels)	Average error (m)
No	200:1	512	6.535689699
Yes	200:1	512	6.337613414
No	200:1	1024	6.3057875019
Yes	200:1	1024	6.0749517957
No	200:1	2048	6.3383421758
Yes	200:1	2048	6.1166968107
No	100:1	512	5.6986560939
Yes	100:1	512	5.0205892544
No	100:1	1024	5.4986851056
Yes	100:1	1024	4.7879202304
No	100:1	2048	5.5468173124
Yes	100:1	2048	4.8213145613
No	20:1	512	2.4868854119
Yes	20:1	512	1.8504729741
No	20:1	1024	2.2364937507
Yes	20:1	1024	1.6295754959
No	20:1	2048	2.3128838675
Yes	20:1	2048	1.7339232268
No	10:1	512	0.9868178954
Yes	10:1	512	1.7198964526
No	10:1	1024	0.7394989867
Yes	10:1	1024	1.5018379834
No	10:1	2048	0.8564899241
Yes	10:1	2048	1.6074981794

or left the average error at the same size. The improvement of the error can reach 33 % of the average error. When the compression is 100:1 then the distributions of the error for not filtered and filtered data are almost equal.

The final conclusion is that the filtering should be used mainly for ratio 100:1 for areas with small variance in elevation.

As described in Ben-Moshe et al. (2007) ICTS technique gives similar or even better results for simplification of DEM than JPEG 2000. What is missing in the paper are parameters used for JPEG 2000 compression, mainly compression ratio. We were not able to test ICTS technique yet, so it would be very interesting to do a research in that area. We would like to recommend to compare our technique with ICTS technique for DEM compression (simplification).

Table 7 Impact of compression to tile S24E125

Filtered	Compression ratio	Block size (pixels)	Average error (m)
No	200:1	512	6.2306892598
Yes	200:1	512	6.0936592253
No	200:1	1024	6.0593004612
Yes	200:1	1024	5.909516942
No	200:1	2048	6.11625323
Yes	200:1	2048	5.9609665185
No	100:1	512	5.6848097751
Yes	100:1	512	5.109245781
No	100:1	1024	5.492256116
Yes	100:1	1024	4.8706002167
No	100:1	2048	5.552220483
Yes	100:1	2048	4.9080530949
No	20:1	512	3.1804059334
Yes	20:1	512	2.0791826239
No	20:1	1024	2.9353852848
Yes	20:1	1024	1.8694861752
No	20:1	2048	3.0062762966
Yes	20:1	2048	1.9709853345
No	10:1	512	1.1675331477
Yes	10:1	512	1.9444637281
No	10:1	1024	0.9293747355
Yes	10:1	1024	1.7352174151
No	10:1	2048	1.0397344809
Yes	10:1	2048	1.8395679222

Table 8 Error distribution for not filtered data for the tile N23E026 with compression ratio 100:1

Error (m)	Number of pixels
0	1,479,984
1–5	8,258,059
6–25	3,228,018
25–40	1127
>40	13

Table 9 Error distribution for filtered data for the tile N23E026 with compression ratio 100:1

Error (m)	Number of pixels
0	1,934,438
1–5	9,192,064
6–25	1,839,507
25–40	1142
>40	50

Table 10 Error distribution for not filtered data for the tile N49E017 with compression ratio 20:1

Error (m)	Number of pixels
0	7,067,717
1	5,423,233
2–3	4,75847
>3	404

Table 11 Error distribution for filtered data for the tile N49E017 with compression ratio 20:1

Error (m)	Number of pixels
0	6,119,138
1	5,601,369
2–3	1,169,941
>3	76,753

Acknowledgments The ASTER GDEM data were obtained through the online Data Pool at the NASA portal http://asterweb.jpl.nasa.gov/gdem-wist.asp.

References

Ben-Moshe B, Serruya L, Shamir A (2007) Image compression terrain simplification. In: Proceedings of the 19th Canadian conference on computational geometry (CCCG2007), pp 125–128. Available at. http://cccg.ca/proceedings/2007/05b1full.pdf

Gladkova I, Gottipati S, Grossberg M (2007) A new lossless compression algorithm for satellite earth science multi-spectral imagers. In: Proceedings of the SPIE, compression, communications, and archiving III, 6683 (668307). Available at. http://glasslab.engr.ccny.cuny.edu/u/grossberg/pubs/imageralgSPIE_07.pdf. Aug 2007

Gladkova I, Grossberg M (2006) A lossless compression algorithm for hyperspectral data. In: Proceedings of the SPIE, satellite data compression, communications, and archiving II, 6300 (630001). Aug 2006

Inanc M (2008) Compressing terrain elevation datasets. Thesis submitted to the graduate faculty of Rensselaer Polytechnic Institute. New York, 62 p. Available at. http://www.ecse.rpiscrews.us/Homepages/wrf/theses/inanc-phd.pdf

MicroImages, Inc. (2009) JPEG2000 Compression for global ASTER DEM. Available at. http://www.microimages.com/documentation/TechGuides/75jp2AsterDEM.pdf

Mittal ML, Singh VK, Krishnan R (2013) Effect of data compression on DEM accuracy derived from satellite images. Docstoc, 6 p. Available at. http://www.docstoc.com/docs/160718730/effect-of-data-compression-on-dem-accuracy-derived-from-satellite

Vatolin D et al (2005) JPEG 2000 image codecs comparison. CS MSU Graphics and Media Lab. Available at. http://compression.ru/video/codec_comparison/jpeg2000_codecs_comparison_en.html