

Compression of GPS trajectories

Version 0.2

Author

Author: Minjie Chen

Goal

We consider the problem of lossy compression for GPS trajectories with latitude, longitude and timestamp information, under a given error tolerance, i.e., *synchronous Euclidean distance* between the original and the compressed trajectories.

About this demo

This demo is written in matlab with both command line format and GUI.

Format

I use txt file as input and output format now. An example of the input file is:

```
40.059846 116.597059 2009-04-19 12:10:44
40.059740 116.597094 2009-04-19 12:10:48
40.059640 116.597137 2009-04-19 12:10:53
40.059559 116.597130 2009-04-19 12:10:58
40.059482 116.597124 2009-04-19 12:11:03
40.059400 116.597108 2009-04-19 12:11:08
40.059364 116.597106 2009-04-19 12:11:13
40.059355 116.597108 2009-04-19 12:11:18
40.059356 116.597108 2009-04-19 12:11:23
40.059358 116.597107 2009-04-19 12:11:28
...
```

Command line:

Encode : Run encode_GPS_traj (inputfile, outputfile, maxtolerance)

Decode: Run decode_GPS_traj (inputfile, outputfile, maxtolerance)

There are GPX->TXT TXT->GPX conversion, which are:

txt2gpx(txtfile,gpxfile)

gpx2txt(gpxfile,txtfile)

Note: Maybe not all GPX files are supported now. A typical GPX format we use is:

```
<trkpt lat="63.181689" lon="22.022990">
```

```
<time>2011-7-21T18:45:34Z</time>
```

```
</trkpt>
```

GUI:

Run test_GUI

Encode: enter input file, set Max SED, press “Encode”.

Decode: enter output file, set Max SED, press “Decode”.

Visualization: enter input file, decoded file and compressed file, “visualize” to plot in matlab figure, “visualize on Google Map” to show the trajectory on Google Map if there are internet connection.

txt2gpx and gpx2txt are also included in GUI. You can convert the decoded txt file to GPX format.

Testing Datasets

The proposed algorithm is tested on 866 trajectories with 4,526,030 points from Microsoft geolife dataset. (We select the trajectories with more than 3,600 points. Every file includes one single trajectory, not multiple ones collected at different time.)

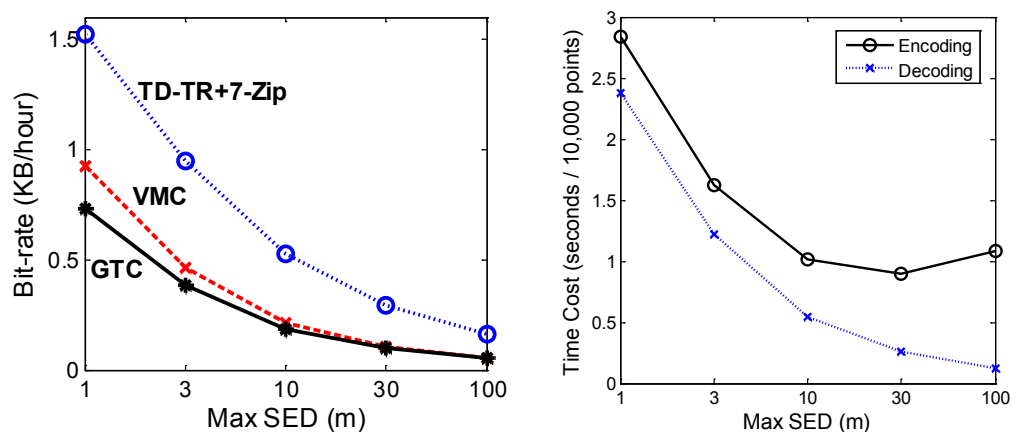
In our algorithm, experiments demonstrate that the storage cost of the proposed methodology is only with a bit-rate 0.39 KB/hour and 0.19 KB/hour for 3m and 10m max SED correspondingly.

We also test it on our MOPSI dataset, which can be downloaded at:

http://cs.joensuu.fi/sipu/MOPSI_GPSTraj_TXTv1.zip

Performance

The average bit-rate (KB/h), time cost of coding process is plotted as follows:



TD-TR: N. Meratnia and R. A. de By. "Spatiotemporal Compression Techniques for Moving Point Objects", *Advances in Database Technology*, vol. 2992, pp. 551–562, 2004.

VMC: M. Chen, M. Xu and P. Fränti, "Fast dynamic quantization algorithm for vector map compression", *IEEE Int. Conf. on Image Processing*, pp. 4289-4292, September 2010.

GTC: Proposed

A comparison with other commercial software can also be seen on:
<http://www.droyd.org/gps-trajectory-compression/>

Trip Length	GPS Records	M4 Binary Accuracy		M4 Binary (zipped) Accuracy	
		[3m]	[10m]	[3m]	[10m]
1 hour	3.600	2.556	1.116	1.404	612
1 day	86.400	61.344	26.784	33.696	14.688
1 week	604.800	429.408	187.488	235.872	102.816
1 month	2.592.000	1.840.320	803.520	1.010.880	440.640
1 year	31.536.000	22.390.560	9.776.160	12.299.040	5.361.120

The compression performance of every file is reported at:

http://cs.joensuu.fi/~mchen/GPSTraj_perf.xlsx

http://cs.joensuu.fi/~mchen/GPSTraj_perf_mopsidata.xlsx

Filtering

If the GPS data is noisy, a pre-filtering can be run beforehand, which is:

filterGPSTraj(inputtxtfile, outputtxtfile)

However, the tolerance will be measured between the compressed and the filtered trajectories later.

Note

This algorithm is for single GPS trajectory compression, not multiple trajectories collected at different time.

Note

Additional infos: see <http://cs.joensuu.fi/~mchen/GPSTrajComp.htm>

Contact

If you have any problems, please send email to mjchen134@gmail.com