# GeoSolutions Technical Document

**Optimizing Coverage Reprojection** 



Ing. Andrea Aime Ing. Simone Giannecchini GeoSolutions S.A.S. 17/12/2010



#### Contents

Overview of coverage data reprojection	4
Driving the JAI warp operation	6
Approximating a generic transformation	7
Affine transformation math	9
Grid warp math1	0
The approximate transformation algorithm1	3
The implementation	6
Some benchmarks	0
First test: simple repeated benchmark 2	0
Second test: FOSS4G 2010 WMS shootout 2	4
Comparing the outputs 2	7
Conclusions	5

Page **2** of **35** 

# **GeoSolutions S.A.S**





No table of figures entries found.

Page **3** of **35** 

## **GeoSolutions S.A.S**





#### **Overview of coverage data reprojection**

In GeoTools a grid coverage is the combination of three elements:

- a raster image, or a regular grid of number,
- a transformation that do map each row/col position in the raster space to a real world position (a.k.a a position in the model space)
- a coordinate reference system in which the real world positions are expressed



Transforming a grid coverage from a coordinate system to another means:

• mapping the real world space of the grid to another real world space expressed in a different CRS (bounds mapping)

Page **4** of **35** 

## **GeoSolutions S.A.S**





- building a new grid of pixels, whose size might be different from the original, to accommodate for shape distortions happening during reprojection
- mapping each pixel in the original grid to a pixel in the destination grid taking into account the change of projection and the change of grid size

The latter step is normally carried out by a math transformation, represented in GeoTools by MathTransform2D. In order to transform the raster into the new one the math transformation has to be applied to each pixel, this is normally done by the JAI WarpImageOp operation.



Page **5** of **35** 

# **GeoSolutions S.A.S**





#### Driving the JAI warp operation

A JAI warp transformation takes a **Warp<sup>1</sup>** object, whose sole purpose is to map source grid pixels to a rectangular area of the destination image (as one pixel can grow to occupy more than one destination pixel).

Warp is actually a base class with several subclasses, which can be used alternatively depending on how the programmer wants to express the mapping.

GeoTools at the time of writing does not use any of the built-in JAI Warp subclasses, but provides its own **WarpAdapter** class which wraps a **MathTransform** and performs exact transformations.

While being the most accurate solution we have, it's also the slowest one. This is due to a couple of issues:

- The operation has to perform all the computation for each point. In the worst case that involves two affine transformations, a change of datum and two cartographic projects, the latter very often using trigonometric functions.
- JAI can use native code implementations of most operations, but those engage exclusively if one of the JAI Warp subclasses is used

As a consequence in order to speed up the operation a JAI subclass of Warp must be used. Common subclasses are **WarpGrid**<sup>2</sup> and **WarpAffine**<sup>3</sup>.

WarpGrid uses a regular grid of points to be transformed, and for each, has the transformed position: for

<sup>1</sup> http://tinyurl.com/2ewvoy3

<sup>2</sup> http://tinyurl.com/29qbe3a

<sup>3</sup> http://tinyurl.com/232lxxf

Page **6** of **35** 

# **GeoSolutions S.A.S**





each pixel that is not on the grid a linear interpolation between the four closest points is used.



**WarpAffine** uses a simple affine transformation instead, that is, a composition of a translation, scale, rotation, and skew. It is a simpler transformation compared to the warp, in fact, a warp can be conceived as a set of many small affine transformations operating side by side, one in each cell of the grid. For this reason the warp grid is also sometimes known as piecewise affine transformation, as it assign a different affine transformation to each area of the plane.

#### Approximating a generic transformation

Generally speaking a reprojection is rarely an affine or piecewise affine transformation, normally it makes use of trigonometric and high order polynomial element that makes it impossible to represent it **exactly** with a simpler math.

However in GIS and visualization very often a certain error can be tolerated, this is due to a number of reasons:

- the original data comes with a certain built in error, which depends on the method of survey, but that is normally in the range of the centimeters, and ofter even in the the meters range
- mapping is about perception, the naked eye can hardly distinguish any error that is below the pixel size (on screen) or below 0.2mm (on paper)

Page **7** of **35** 

# **GeoSolutions S.A.S**





This means approximations in the math used to perform a transformation are acceptable provided that they are significantly smaller than the data and display method resolutions.

It also helps to notice that very often transformations over data are performed over small areas, whilst the non linear deformations occurring in most transformations are very visible only on large areas, and only if far away from the "center" of the projection.

Drawing a parallel with one dimensional functions it is possible to approximate a curve with a straight line provided the range in which we want to perform the approximation is small enough:



In case the range is larger a piecewise linear approximation can be used: by making the length of each segment small enough the piecewise approximation can be within any desired maximum distance from the

Page 8 of 35

# **GeoSolutions S.A.S**





Going back to the Warp case, this means it's possible to use a simpler **WarpAffine** or a **WarpGrid** that matches the general projection math within a certain, user specified tolerance.

#### Affine transformation math

An **affine transformation** is a linear transformation mapping between two vector spaces. The transformation is a generalization of common transformations such as scaling, rotation, translation and shears, which can be all represented with specific affine transformation layouts. Affine transform have the following properties:

- 1. the <u>collinearity</u> relation between points; i.e., the points which lie on a line continue to be collinear after the transformation
- Ratios of distances along a line; i.e., for distinct collinear points p1,p2,p3, the ratio | p2 p1 | / | p3 p2 | is preserved

Page **9** of **35** 

# **GeoSolutions S.A.S**





Normally affine transformations are represented in matrix form as:

$\begin{bmatrix} x' \end{bmatrix}$		m00	m01	m02	$\begin{bmatrix} x \end{bmatrix}$		m00x + m01y + m02
y'	=	m10	m11	m12	y	=	m10x + m11y + m12
$\lfloor 1 \rfloor$		0	0	1	1		1

The simplicity of the expression allows for rather efficient software implementations of the above.

#### **Grid warp math**

A grid warp describe a transformation by providing a regular set of points and their transformed relatives. Any generic point falling in a cell of the grid transformation is linearly interpolated based on the transformed positions of the four nearby points.

The following diagram shows the superimposition of a regular source grid (red points) and the transformed

Page **10** of **35** 

## **GeoSolutions S.A.S**





#### grid (blue points):



In a warp transformation the operation has to **map back** from a certain pixel in the transformed grid into the original pixel, in order to gather its value. Under more complex interpolation schemes, such as bicubic, in fact the algorithm has to backtrack the position of a group of nearby cells in order to compute the results.

Page 11 of 35

## **GeoSolutions S.A.S**

![](_page_10_Picture_8.jpeg)

![](_page_11_Figure_0.jpeg)

Given a destination pixel coordinate (x, y) that lies within a cell having corners at (x0, y0), (x1, y0), (x0, y1), and (x1, y1), with source coordinates defined at each respective corner equal to (sx0, sy0), (sx1, sy1), (sx2, sy2), and (sx3, sy3), the source position (sx, sy) that maps onto (x, y) is given by the following equations:

$$\begin{split} xfrac &= \frac{x-x0}{x1-x0}\\ yfrac &= \frac{y-y0}{y1-y0}\\ s &= sx0 + (sx1-sx0)xfrac\\ t &= sy0 + (sy1-sy0)xfrac\\ u &= sx2 + (sx3-sx2)xfrac\\ v &= sy2 + (sy3-sy2)xfrac\\ sx &= s + (u-s)yfrac\\ sy &= t + (v-t)yfrac \end{split}$$

Page 12 of 35

## **GeoSolutions S.A.S**

![](_page_11_Picture_6.jpeg)

![](_page_12_Picture_0.jpeg)

Basically the source x and y values are first interpolated horizontally along the top and bottom edges of the grid cell, resulting in (s,t) and (u,v), and the results are interpolated vertically getting to (sx, sy)

#### The approximate transformation algorithm

In order to replace the generic projection math with a grid warp it's important to be able to estimate what is the maximum error introduced by the simplified grid.

Deciding on the proper grid size is also important speed and memory wise, as a larger grid takes more time to compute, to use, and to store in memory: it's clear that a properly designed algorithm should stop once the target accuracy is achieved, and eventually decide to give up and not use the approximation in case the transformation grid becomes either as dense as the grid to be transformed or just too big to safely fit in memory#.

In order to figure out if an affine transformation can satisfy the desired accuracy a sampling along some major lines in the area to be reprojected is performed:

![](_page_12_Picture_7.jpeg)

For each of these lines the start, end and mid points are reprojected, generating a set of three points, not necessarily collinear. Then the mid point of the transformed start and end points is computed as well, and the distance between this "predicted" mid point and the actual transformed mid point is evaluated: if such distance is lower than the tolerance then the transformation along that line is mostly affine, otherwise we can conclude an affine transformation on the target area.

Page 13 of 35

# **GeoSolutions S.A.S**

![](_page_12_Picture_12.jpeg)

![](_page_13_Picture_0.jpeg)

In case one of the tests fails we have proof the current area is too large to be approximated with a single affine transform, this means the area has to be split into smaller parts (piecewise) and the procedure reiterated until all of the pieces are small enough to allow a single affine transform to represent them.

In practice many map projections tend to deform the area much more in one direction than in another, so it may be that the overall grid required more rows than columns, or viceversa, to properly match the desired accuracy.

For example **Mercator** tends to stretch geometries much more towards the pole and has pretty much constant deformation in the east/west direction: it means the transformation grid will probably have many rows but few columns.

**Transverse Mercator**, used by the UTM system, behaves the opposite way so we'll expect to see a grid with many columns and fewer rows.

This characteristic is matched by the splitting rules implemented in the estimation algorithm:

- if only the horizontal lines fail the accuracy test the area will be split into two columns
- if only the vertical lines fail the accuracy test the area will be split into two rows
- if both fail the tests, or if any of the diagonal lines does, the area will be split into four areas

Page **14** of **35** 

## **GeoSolutions S.A.S**

![](_page_13_Picture_12.jpeg)

![](_page_14_Figure_0.jpeg)

Each recursive split doubles the number of rows or columns (or both) needed to represent the warp grid.

Once the evaluation phase completes a split depth for rows and columns has been computed: if the depth is 0 in both directions a single **WarpAffine** can be used, otherwise a grid made by 2<sup>coudepth</sup> rows by 2<sup>collepth</sup> columns is generated, each grid point is then reprojected and given to **GridWarp** as a reference.

If, during the recursive depth evaluation, the depth reaches a too high level, or the density of the grid mathes the original grid size, the approximation attempt is abandoned and the existing **WarpAdapter** Page **15** of **35** 

# **GeoSolutions S.A.S**

![](_page_14_Picture_6.jpeg)

![](_page_15_Picture_0.jpeg)

transformation is used: this is because building the approximation grid is proving to be so expensive both in terms of computation and memory usage to offset its speed advantages.

#### **The implementation**

The implementation of the above logic has been encapsulated in the GeoTools WarpBuilder class, which takes care of implementing the above logic and come up with either a WarpAffine or a WarpGrid, falling back to the Geotools WarpAdapter only when the warp grid is getting either dense than the original data or too big to be stored in memory.

The following shows how a simple digital elevation model gets reprojected from EPSG:26713 to EPSG:3785, and operation that involves both a change of datum (from NAD27 to WGS84) and a change of projection (from to UTM zone 13N to Mercator):

Source projection	Target projection
<pre>PROJCS["NAD27 / UTM zone 13N", GEOGCS["NAD27", DATUM["North American Datum 1927", SPHEROID["Clarke 1866", 6378206.4, 294.9786982138982, AUTHORITY["EPSG","7008"]], TOWGS84[2.478, 149.752, 197.726, 0.526, - 0.498, 0.501, 0.14129139227926102], AUTHORITY["EPSG","6267"]], PRIMEM["Greenwich", 0.0, AUTHORITY["EPSG","8901"]], UNIT["degree", 0.017453292519943295], AXIS["Geodetic langitude", EAST], AXIS["Geodetic latitude", NORTH], AUTHORITY["EPSG","4267"]], PROJECTION["Transverse Mercator", AUTHORITY["EPSG","9807"]], PARAMETER["central_meridian", -105.0], PARAMETER["scale_factor", 0.9996], PARAMETER["false_easting", 500000.0], PARAMETER["false_northing", 0.0], UNIT["m", 1.0], AXIS["Easting", EAST], AXIS["Northing", NORTH], AUTHORITY["EPSG","26713"]]</pre>	<pre>PROJCS["Popular Visualisation CRS / Mercator", GEOGCS["Popular Visualisation CRS", DATUM["Popular Visualisation Datum", SPHEROID["Popular Visualisation Sphere", 6378137.0, 0.0, AUTHORITY["EPSG", "7059"]], TOWGS84[0.0, 0.0, 0.0, 0.0, 0.0, 0.0], AUTHORITY["EPSG", "6055"]], PRIMEM["Greenwich", 0.0, AUTHORITY["EPSG", "8901"]], UNIT["degree", 0.017453292519943295], AXIS["Geodetic longitude", EAST], AXIS["Geodetic latitude", NORTH], AUTHORITY["EPSG", "4055"]], PROJECTION["Mercator (1SP) (Spherical)", AUTHORITY["EPSG", "9841"]], PARAMETER["latitude_of_origin", 0.0], PARAMETER["central_meridian", 0.0], PARAMETER["false_easting", 0.0], UNIT["m", 1.0], AXIS["Easting", EAST], AXIS["Northing", NORTH], AUTHORITY["EPSG", "3785"]]</pre>

Page 16 of 35

# **GeoSolutions S.A.S**

![](_page_15_Picture_10.jpeg)

![](_page_16_Picture_0.jpeg)

![](_page_16_Picture_2.jpeg)

It is evident from the images that the overall main effect to the image is a small clockwise rotation, yet the overall effect cannot be simply represented as a simple affine transform due to other hard to see to the naked eye alterations.

In fact the optimized code, using a 0.333 tolerance, generates a 3x3 grid warp to perform the transformation. In the following picture the red dots represent the original position of the grid points, and the blue dots the transformed ones:

Page 17 of 35

## **GeoSolutions S.A.S**

![](_page_16_Picture_8.jpeg)

![](_page_17_Picture_0.jpeg)

Transformation involving higher distortions will result in denser grids and easier to tell, to the naked eye, changes, like the following, which is obtained by applying a polar stereographic transformation instead:

Page **18** of **35** 

## **GeoSolutions S.A.S**

![](_page_17_Picture_5.jpeg)

![](_page_18_Picture_0.jpeg)

The tolerance used by the approximating algorithm can be set by provided via a global hint, for example:

Hints.putSystemDefault(Hints.RESAMPLE\_TOLERANCE, 0.333);

Page **19** of **35** 

**GeoSolutions S.A.S** 

![](_page_18_Picture_6.jpeg)

![](_page_19_Picture_0.jpeg)

#### **Some benchmarks**

We run a couple of benchmarks to gauge the performance improvement obtained with the improved warp. To do so we used GeoServer 2.1 beta2 with the modified GeoTools package and run a set of WMS GetMap requests with and without the warp optimizations.

The benchmarks have been run on a Ubuntu 10.04, 64 bit workstation. The hardware includes a core i860 quad core processor, 8GB of memory, a fast 2TB hard drive, Western Digital Caviar Black 2TB (120MB/s peak throughput), and a GTX 260 Nvidia video card.

#### First test: simple repeated benchmark

This test involves issuing repeatedly the same request with 1, 2, 4, 8 and 16 clients on the "spearfish dem" GeoServer sample layer, a single band float GeoTiff rendered using a color map to turn numbers into colors.

Page **20** of **35** 

## **GeoSolutions S.A.S**

![](_page_19_Picture_9.jpeg)

![](_page_20_Picture_0.jpeg)

![](_page_20_Picture_2.jpeg)

The request is http://localhost:8080/geoserver/wms/reflect?layer=sfdem&srs=EPSG:900913, which makes GeoServer fill in the missing WMS GetMap parameters automatically, resulting in a request in the cited srs, showing the full layer, without having to compute the bbox manually.

The following chart shows the results with and without the optimization:

Page **21** of **35** 

# **GeoSolutions S.A.S**

![](_page_20_Picture_8.jpeg)

![](_page_21_Picture_0.jpeg)

![](_page_21_Figure_2.jpeg)

Clients	Base (req/s)	Optimized (req/s)	Speedup
1	2,86	13,01	4,56
2	5,54	22,95	4,14
3	10,87	39,89	3,67
8	9,42	34,67	3,68
16	10,57	40,58	3,84

Page 22 of 35

# **GeoSolutions S.A.S**

![](_page_21_Picture_7.jpeg)

![](_page_22_Picture_0.jpeg)

![](_page_22_Figure_2.jpeg)

Concurrent clients

Clients	Base (ms)	Optimized (ms)	Speedup
1	350	76	4,61
2	351	77	4,56
3	354	84	4,21
8	510	133	3,83
16	822	203	4,05

Page 23 of 35

# **GeoSolutions S.A.S**

![](_page_22_Picture_8.jpeg)

![](_page_23_Picture_0.jpeg)

#### Second test: FOSS4G 2010 WMS shootout

During FOSS4G 2010 a performance comparison between WMS servers was held, one of the tests involved taking a mosaic of high resolution imagery in TIFF format (no compression, with overviews) and in EPSG:25831 (UTM 31N, ETRS89 datum) and rendering it in the Google projection (Mercator projection, WGS84). The overall mosaic size is of 110GB.

The benchmark used growing number of concurrent clients and fully randomized requests to make sure no caching was possible.

![](_page_23_Picture_5.jpeg)

Page 24 of 35

## **GeoSolutions S.A.S**

![](_page_23_Picture_9.jpeg)

![](_page_24_Picture_0.jpeg)

The following table shows the results of the benchmark before and after the warp optimizations:

![](_page_24_Figure_3.jpeg)

Clients	Base (r/s)	Optimized (r/s)	Speedup
1	1,85	9,82	5,30
2	3,15	19,72	6,26
3	6,04	36,60	6,06
8	9,64	47,29	4,90
16	9,52	39,73	4,17

Page 25 of 35

# **GeoSolutions S.A.S**

![](_page_24_Picture_8.jpeg)

![](_page_25_Picture_0.jpeg)

![](_page_25_Figure_2.jpeg)

Concurrent clients

Clients	Base (ms)	Optimized (ms)	Speedup
1	538	101	5,33
2	607	90	6,74
3	605	95	6,37
8	766	136	5,63
16	1591	356	4,47

Page 26 of 35

# **GeoSolutions S.A.S**

![](_page_25_Picture_8.jpeg)

![](_page_26_Picture_0.jpeg)

### Comparing the outputs

A set of tests was run to compare the images generated before and after the change. To do so the same WMS request# was run with the optimization set as in the benchmarks, with a 0.333 tolerance factor, and without the optimization, using the accurate but slow algorithm.

The two images similarity is stunning, it's quite hard to tell any difference by the naked eye:

#### Image obtained without any optimization

![](_page_26_Picture_5.jpeg)

Page 27 of 35

## **GeoSolutions S.A.S**

Via Carignoni 51 55041 Camaiore (LU) Italy Tel: +390584983027 Fax: +390584983027 http://www.geo-.solutions.it info@geo-solutions.it

![](_page_26_Picture_9.jpeg)

**Optimizing Coverage** 

**Reprojection** 

![](_page_27_Picture_0.jpeg)

#### Image obtained with 0.333 tolerance

![](_page_27_Picture_3.jpeg)

The two images have been then compared using the compare ImageMagick utility, which extract image differences a red pixel mask.

The tool was instructed to extract a map of absolute differences (red pixel if even a small difference was Page 28 of 35

## **GeoSolutions S.A.S**

![](_page_27_Picture_8.jpeg)

![](_page_28_Picture_0.jpeg)

found), then a map of pixels with a difference above 5%, then 10, 20 30 and 40. The difference tolerance is set via the fuzz factor, defined as "a 'similarity' match in multi-dimensional spherical distance between colors, using whatever color space the image is using", that is, in our case, a distance between the RGB triplets of the two images.

A second ImageMagick call was then made to add a black border around the images. Example command line invocation used to perform the comparison:

compare -compose src original.png optimized.png -fuzz 10% difference-10.png convert difference-10.png -bordercolor Black -border 1 difference-10b.png

Here are the resulting difference maps:

Page **29** of **35** 

# **GeoSolutions S.A.S**

![](_page_28_Picture_9.jpeg)

![](_page_29_Picture_0.jpeg)

![](_page_29_Picture_2.jpeg)

**Absolute difference** 

Page 30 of 35

## **GeoSolutions S.A.S**

![](_page_29_Picture_7.jpeg)

![](_page_30_Picture_0.jpeg)

![](_page_30_Figure_2.jpeg)

Pixels with a difference above 5%

Page **31** of **35** 

**GeoSolutions S.A.S** 

![](_page_30_Picture_7.jpeg)

![](_page_31_Picture_0.jpeg)

![](_page_31_Figure_2.jpeg)

Pixels with a difference above 10%

Page **32** of **35** 

**GeoSolutions S.A.S** 

![](_page_31_Picture_7.jpeg)

![](_page_32_Picture_0.jpeg)

![](_page_32_Figure_2.jpeg)

#### Pixels with a difference above 20%

Looking at the pictures we can draw a few conclusions:

• It's quite hard to see any difference in the original images by the naked eye

Page **33** of **35** 

# **GeoSolutions S.A.S**

![](_page_32_Picture_9.jpeg)

![](_page_33_Picture_0.jpeg)

• An automated tool can spot differences by computing color distance, yet even in that case the differences are few and there is basically none left to speak of if we consider differences above 20%

Page 34 of 35

## **GeoSolutions S.A.S**

![](_page_33_Picture_6.jpeg)

![](_page_34_Picture_0.jpeg)

#### **Conclusions**

- An approximate method to optimize coverage reprojection has been devised and implemented in GeoTools
- It provides a 4-6 times speedup factor
- Visually indistinguishable results using a tolerance of 0.333
- Possible to control the tolerance in code using Hints, for example:

Hints.putSystemDefault(Hints.RESAMPLE\_TOLERANCE, 0.333);

Page **35** of **35** 

## **GeoSolutions S.A.S**

![](_page_34_Picture_11.jpeg)