Novel Real-Time Coordinate Transformations based on N-Dimensional
Geo-Registration Parameters' Matrices

Sagi DALYOT, Ariel GERSHKOVICH and Yerach DOYTSHER, Israel

Key words: coordinate and datum transformation, geo-registration matrix, interpolation, map projection

SUMMARY

Digital maps and datasets provide the positioning of variety types of geo-spatial entities. These are given in different reference coordinate systems based on different geodetic projections and datums. Simultaneous use of this varied data requires geo-oriented systems to apply real-time data manipulation, and mechanisms enabling the support of immediate and precise processes (for example RTK GPS systems calculating an exact location on a given map). Transforming location-based data between two given coordinate-systems may be time consuming, and hence not always appropriate for systems that require fast decision making. Furthermore, not all datum-to-datum transformation algorithms are known, and some has to be implemented "on the go" - hence involving data uncertainty. This paper suggests a novel method that simplifies these complexities, while presenting a much faster process with no accuracy loss. A short pre-processing procedure is proposed, in which a sparse grid describing the geo-spatial relations - in this case 2D translations - between the two datasets (each with different coordinate system) is created. Based on these translations parameters, N-dimensional geo-registration parameters' matrices are established. By implementing designated interpolation algorithms to handle correctly the data stored in these N-dimensional matrices, it is feasible to calculate immediately the exact transformation for any required location within the maps bounds, and with no calculation complexity. As a result, a location identified in one map is easily transformed precisely to its location in the other map regardless differences in their geodetic projections and/or datums, as well as to transformation complexities exist between them. A case study is presented here that tested this novel concept. It was proved to be accurate - preserving the original transformation accuracy - and 5 times faster than the common approach of implementing complex and time-consuming coordinate and datum transformations. Consequently, this novel approach is suitable for real-time decision making geo-oriented applications, such as web-based ones.
1. INTRODUCTION

The ability and need to connect projects and data throughout a common coordinate system is increasingly necessary with the automation, exchange, update, and merging of different geographic information. However, converting data has become more complicated than in the past due to an increasing number of datums, adjustments, and coordinate systems. New data collection technologies and geo-oriented decision-making applications yield frequent updating of geo-spatial maps and datasets. Some applications - mainly the ones that are particularly designed for the web-domain - demand a fast response time, while being relatively light in data manipulation. One of the most common geo-oriented systems today are the personal Global Position Systems (GPS) that give fast and relatively precise knowledge on the location of its user (Ware, 1995). This system serves as basis for a variety of decision-making applications - dependent on their precision reliability. These systems and devices are becoming more personal and less technically sophisticated to use, so today they can be found in mobile-phones or car navigation systems. These are mostly combined (interwined) with Geographic Information Systems (GIS), which present roads, streets, buildings, etc, and hence serve as the map infrastructure for the GPS positioning data. Still, it is important to remember that while GPS nav-systems give the user a 3D based-positioning (mostly in geographic or geocentric coordinates in respect to WGS84 datum), the GIS systems will usually present a 2D map-projected view of the 3D reality. While there exist only one world ("global") 3D datum used by the GPS nav-systems, several hundred of 2D Cartesian map projected reference datums exist today. These reference datums are usually related to a certain country coordinate system. Still, a certain country can be covered by several known different reference datums. Consequently, a transformation between these different map projections (or coordinate systems) is application-driven, and hence important to know and must be feasible to calculate - fast and accurately. It is important to remember that coordinate systems serve as the framework by which geographic data are referenced to the earth's surface. As such, coordinate systems are the infrastructure, and hence are vital to mapping, surveying, and engineering tasks.

1.1 Geodetic Datum

A geodetic datum is the mathematical modeling of the shape and size of the earth - and its origin and orientation. This defines a geographic coordinate system (i.e., axes) that enables the measurement and description of an object on the earth's surface in latitude and longitude coordinates - or any other coordinate system for that matter (for example: northing and easting). This mathematical modeling process takes into consideration several aspects and parameters - mainly derived from the geographic "shape" needed to be mapped. There are hundreds of locally-developed reference datums that are in use today - each referencing a desired local reference frame while utilizing different modeling techniques and
considerations. The importance of knowing the reference datum is important, for example, when using a GPS where a caution must be taken when checking the maps' datum utilized in order to prevent distortions or wrong location measurement.

1.2 Map Projection

Map projection is the modeling (or "flattening") of the earth surface (or sphere) on a plane - from 3D reality to a 2D look-alike. This modeling, based mainly on a set of mathematical considerations and functions transforming coordinates from curved surface to the projected plane, will always distort the shape of the surface in some fashion (area, shape, direction, scale - to name a few), and in a certain magnitude. As a result, certain map projections exist designated to preserve some basic metric properties of the sphere body that are essential - at the expense of other. It is worth noting that the majority of these projections are inspired by geometric principles, and thus are designated for a (relatively) local and small-scale area mapping. Still, some projections have no geometric or physical interpretation, and are described purely by mathematical formulae - usually specified for large scale areas. Though several hundred of user-defined map projections exist today, still there is no limit to the number of possible map projections. Consequently, each map projection has its advantages and disadvantages, hence the appropriate map projection depends on the scale of the map, and on the purposes for which it will be used. The use of a "projected reality", i.e., map, other than a globe is essential mainly because it is more compact and easier to store; it can be stored as digital map data; it can be viewed easily on computer displays (referring here to computer or GPS/GIS nav-systems screens); it facilitates measuring properties that are more cognitive to human; and more.

Three main steps are involved when modeling the earth's surface and establishing a map projection (Wisconsin Spatial Reference Systems, 2007):

- Selecting the earth's model to be projected: sphere (i.e., geoid) or ellipsoid;
- Reducing the scale;
- Assigning a set of mathematical formulae, geometric principles for transforming the 3D geographic coordinates to 2D plane coordinates.

Several projection shapes can be considered, while the common ones are: cylinder, cone, and plane - all are considered as developable surfaces (a simple geometric form capable of being flattened without stretching) - as depicted in Figure 1. The orientation of the projected shape relative to the mapped surface can be: normal, transverse, and oblique (any given angle). Another consideration is whether the projected shape is tangent to the mapped surface - touching each other, or secant - the mapped surface is sliced by the projected shape - as depicted in Figure 2.

Each map projection utilized to map the earth's surface, i.e., projection shape, orientation and position, has its own set of parameters needed to achieve the desired modeling. This yields that for any given projection a geodetic coordinate system is calculated. These coordinate systems are usually referred to as national grids, mainly because they are designated to map a certain country while minimizing distortions. This means that for any given object on the earth's surface its location in one map projection has coordinates that are completely different when measured while using an alternative map projection. The difference, i.e., 2D translation, between the two different Cartesian coordinates (both in x and y - or northing and easting) is
referred to as datum transformation (or shift). This "value" between any two given particular datums (derived from the utilized projection) can vary from one place to another within one country or region. It can reach the value of several hundreds - and even thousands - of meters, dependent on the two reference datums and map projections used (i.e., parameters used to define the datum and projections).

Figure 1. Several constellations of the earth surface and its shape of projection (taken from geology.isu.edu).

Figure 2. Tangent and secant cylindrical projection of the globe (taken from nationalatlas.gov).

1.3 Coordinate and Datum Transformations

The procedure of "translating" a certain coordinate value, i.e., x, y position, from one defined coordinate system into another while utilizing a set of algorithms and formulae is known as
coordinate transformation (or conversion). When both coordinate systems are based on the same geodetic datum this transformation is precise due to the known exact mathematical relationship. This is sometimes not the case when the transformation deals with different geodetic datums - due to various datum adjustments, errors and distortions exist, and the different mathematical model used. Here, a datum transformation is required, which takes into account ellipsoid (or geoid) translation, inter-mediate coordinate conversion, reverse engineering calculation - just to name a few (in many cases the transformation to a "common" geodetic datum - such as WGS84 - is carried out). A block diagram of this is depicted in Figure 3, where the left side depicts a precise and direct coordinate transformation, while the right side depicts another coordinate system that requires in addition the use of datum transformation, thus introducing uncertainty and ambiguity to the process.

![Figure 3. Coordinate and datum transformation block diagram.](image)

Coordinate and datum transformations can be carried out while utilizing various methods. Some of these methods support highly accurate geodetic and surveying work, and others are approximately adequate for mapping and other purposes. Usually, the coordinate data is exactly converted via a set of mathematical formulae, which are derived from the specific geodetic datums and projections used. These procedures are mostly time consuming and involve mass mathematical formulae and a-priori considerations.

2. PROPOSED CONCEPT

As discussed in section 1.3, the procedure of coordinate and datum transformations is time consuming and usually demands mass computer resources due to its algorithmic complexity. In the case where a mass amount of coordinates are needed to be transformed an alternative
approach that will shorten the procedure time might be required. Instead of transforming all given coordinates from source coordinate system to target coordinate system we suggest implementing an interpolation algorithm on an N-dimensional geo-registration matrices that stores the values of both coordinate systems translation values, as proposed by (Dalyot and Doytsher, 2006). This novel and alternative concept is designated to handle mass coordinates transformation process while not being affected by their mathematical complexities.

2.1 General Processes

The concept is divided into two main stages:

- Pre-processing (it is worth noting that the pre-processing stage is carried out only once for each two given reference datums within the desired bounding area):
  - Dividing the bounded region by all coordinates needed for transformation (source) into a grid (because this process might be suitable for national datums - the entire country bound might be used instead). This grid resolution is derived from the accuracy required for the process (as will be explained in section 3). Each grid node has the coordinates of the source coordinate system (it is worth emphasizing that this set of grid-nodes are not the coordinates needed for transformation);
  - The calculation of the known direct coordinate and datum transformations is carried out - but only for the grid-nodes represented by the grid;
  - Due to the nature of datum transformation, the outcome of the previous stage is an irregular structure of points - each storing a coordinate in the target coordinate system. The connection (inter-relation) between a given source coordinate and its target coordinate is known - expressing the translation;
  - Calculation of the geo-registration values - i.e., in the case of 2D map projection: $dx$ and $dy$ translation - of all the inter-related coordinates (source and target) is carried out, and stored as a matrix structure. These geo-registration values include both coordinate and datum transformation processes.
  - These matrices serve as the basis for calculating the "corrections" needed, and can be depicted as if superimposed on top the source grid;

- Transformation parameters (translation) calculation:
  - A loop is implemented to calculate the target coordinates for all source coordinates to be transformed - without implementing coordinate and datum transformations. First, the grid-cell that bounds the source coordinate is located. The same grid-cell in the matrices that store the translation values are located, simultaneously;
  - A novel interpolation algorithm is implemented to calculate the exact translations (or corrections) needed for the source coordinate. This algorithm is outlined in further detail in section 2.2.
  - The result of adding the calculated translations to the source coordinates is the exact and precise positioning of the target coordinates.

2.2 Interpolation Algorithm

Due to the linear behavior of the translation values stored in the geo-registration matrix - $dx$ and $dy$ - an interpolation on these values can be performed by most of the available
interpolation methods. Nevertheless, due to the data structure - a DEM look like database (matrix) - an algorithm that will ensure continuous representation of the geo-registration parameters has to be chosen. Interpolating DEM heights using bi-directional third degree parabolic equations is described by (Doythser and Hall, 1996). This algorithm ensures a smooth interpolation within a grid-cell, as well as excluding surface representation discontinuities on cell borders while moving to the neighboring cell. This algorithm is an improved version of the simple bi-linear interpolation that produces a jagged and unsmooth surface representation. The algorithm works within a grid-cell, utilizing heights of 4 by 4 nodes, as depicted in Figure 4.

Figure 4. Geometrical scheme showing 16 nodes utilized for interpolation (bold black) (left), with point T located within inner cell (right).

For each location within a grid-cell (T) a normalization between the values of (0,0) and (1,1) of the local coordinates x and y is carried out (ti and tj, respectively). Knowing the translation values of the 16 neighboring nodes, the calculation of the specific point translation (i.e., dx and dy, respectively) is achieved, as depicted in Equation 1, while using third degree parabolic equations:

\[
\begin{align*}
F_1(t) &= -0.5 \cdot t + 1.0 \cdot t^2 - 0.5 \cdot t^3 \\
F_2(t) &= +1.0 - 2.5 \cdot t^2 + 1.5 \cdot t^3 \\
F_3(t) &= +0.5 \cdot t + 2.0 \cdot t^2 - 1.5 \cdot t^3 \\
F_4(t) &= -0.5 \cdot t^2 + 0.5 \cdot t^3 \\
Z_T &= \sum_{i=0}^{4} \sum_{j=0}^{4} F_i(t) \cdot F_j(t) \cdot H(i,j)
\end{align*}
\]

where \(F1(t) - F4(t)\) denotes the third-degree parabolic equations, \(Z_T\) denotes the interpolated translation value at location \(T\), \(t\) denotes the normalized inner cell coordinates \(\{0 \leq t \leq 1\}\), \(H(i,j)\) denotes the translation values of nodes, and \(i, j\) denotes index of 4 by 4 neighboring nodes.

This algorithm describes a weighted interpolation scheme, in which the inner grid-cell heights have the highest weight, while the remaining 12 nodes' heights have inferior weights. The same concept is implemented on the geo-registration matrix; instead of interpolating height values we implement the suggested interpolation in 2 separate processes on the translation values stored in the matrices' nodes: \(dx\) and \(dy\). By implementing this process we were able to
accurately define for each required source coordinates the precise corrections - based on the values stored in the matrices - in order to calculate the target coordinates.

3. CASE STUDY

The two most common projections used as reference surfaces for Cartesian coordinate systems are the Lambert conformal conic and the transverse Mercator, both depicted in Figure 5. Both projections have varying scale but retain the correct shape of the mapped surface. Scale variation is greatest in north-south directions for Lambert projections, and the east-west directions for transverse Mercator projections. These two projections cover western Europe. As a result, France was chosen for evaluating the proposed concept as it is covered by both map projections: Lambert datum - as well as Universal Transverse Mercator (UTM) zones 31-33. The bounds of France in respect to a Geographic global datum is depicted in Figure 6.

Transformation of a given coordinate from France's Lambert datum, defined by Clarke 1880 ellipsoid, to UTM datum, defined by WGS84 ellipsoid, involves the following stages:
- Transforming projected Lambert coordinates \((x, y)\) to Geographical ones given as Latitude and Longitude \((\phi, \lambda)\) (Clarke 1880);
- Datum transformation - from Clarke 1880 to WGS84;
- Transforming the calculated Geographical Latitude and Longitude \((\phi, \lambda)\) (WGS84) to projected UTM coordinates \((x, y)\) - dependent on the required zone.

![Figure 5. Lambert conformal conic projection (left), and transverse Mercator projection (right) (taken from sco.wisc.edu).](image)

The complete set of coordinate and datum transformations from Lambert projection to UTM - and vice versa - was implemented. Tests showed an accuracy of less than 1 cm when calculating the transformation from one projection to the other - and then back. This validated the complete known mathematical formulae utilized for this process, and was referenced to analyze the proposed concept's accuracy.
The implementation of the suggested geo-registration matrices concept was examined by writing a computer program in VB.NET code. A GUI, depicted in Figure 7, enabled the user to define the desired values of the grid resolution, as well as the point needed to be calculated, i.e., transformed from source to target. The output of the GUI is:

- Values difference in \(X\) and \(Y\) directions of the inserted coordinate calculated via the two procedures: interpolation while utilizing the geo-registration matrices; and, the direct coordinate and datum transformations;
- Clock time of these two procedures (it is worth noting that the clock time of the suggested concept did not take into account the pre-processing stage).

It is obvious that the higher the resolution chosen is - the more accurate the calculation via the interpolation will be. Nevertheless, this will cost in a longer pre-processing time and computer resources to store this matrices (probably as a database). A thorough analysis of the accuracy vs. resolution is presented in Table 1, showing the dependency between the two. The error value varies from approximately \(1: 9\times10^{-9}\) to \(1: 250,000\) the resolution value chosen. The higher the resolution - the more accurate the calculation is because a known value is located nearby. As a result, a pre-consideration of the desired accuracy is required prior to carrying out the pre-processing stage. Generally, it can be considered that a resolution value denser than 25,000 m for precise geodetic coordinate calculation purposes, and 100,000 m for graphic ones is unnecessary from a precision viewpoint.
Figure 7. GUI developed to validate and analyse the proposed concept.

<table>
<thead>
<tr>
<th>Resolution value (grid spacing) [m]</th>
<th>Diagonal difference [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>5.65E-08</td>
</tr>
<tr>
<td>1,000</td>
<td>3.37E-07</td>
</tr>
<tr>
<td>5,000</td>
<td>5.58E-05</td>
</tr>
<tr>
<td>10,000</td>
<td>0.0005</td>
</tr>
<tr>
<td>25,000</td>
<td>0.006</td>
</tr>
<tr>
<td>50,000</td>
<td>0.0645</td>
</tr>
<tr>
<td>100,000</td>
<td>0.3846</td>
</tr>
</tbody>
</table>

Table 1. Dependency of grid resolution chosen for the pre-processing stage, and the derived accuracy of the process.

Three main conclusions can be stated:

- For most geodetic purposes the accuracy accepted is adequate enough even when using a resolution of 25,000 m, i.e., less than 1 cm. In the case study of France, the country's entire coverage area is approximately 1,200x1,600 km. This means that a relatively small grid - approximately 50x65 nodes - is suitable. Hence, a short pre-processing process (approximately 70 milli-seconds long) is required. Moreover, the relatively small matrix yields a small database storage, feasible in any hand-held device.

- It is clear that relatively large variations exist in the geo-registration matrices cells, ranging from few dozens to hundreds of meters (shown in Figure 7 on the right GUI side; values are denoted as $dX$ and $dY$ for each grid node in cell). However, the
interpolation concept was accurate enough and reliable to predict the local trends exist in the matrices "topography", thus resulting in a correct translation values calculation needed for each location analyzed and inspected in this case study.

The time required for the interpolation process is approximately 5 times faster. For example, while utilizing the computer configuration used for programming the code - OS: WindowsXP; processor: Intel Core2 Duo 1.8Ghz; RAM: 1 GB - the direct coordinate and datum transformation of 1,000,000 (1E+6) coordinates will take approximately 1 minute, while when utilizing the proposed concept it will take less than 10 seconds (running times are depicted in the lower part of Figure 7). This is significant when a real-time (or web-based) decision-making application is considered.

4. CURRENT RESEARCH

Abd-Elmotaal (1994) presents several reasons that can cause significant distortions while trying to perform datum transformations. Moreover, he states that the mathematical models exist for datum transformation cannot minimize these distortions because they do not contain sufficient transformation parameters to adequately describe the relationship between these datums. While trying to improve the accuracy of the existing models, the author utilized polynomial methods for a limited area, thus still implementing relatively complex formulae as well as restricting the proposed solution to a relatively small area.

Grafarend and Awange (2003) suggest a tool for solving the three-dimensional datum transformation problem by utilizing the weighted Procrustes algorithm that does not require any initial datum parameters for linearization or any iteration procedure. Still, the authors state that some prior information about the variance-covariance matrix of the two sets of Cartesian coordinates (pseudo-observations) is required in order to accomplish this.

Vanícek and Steeves (1996) outline in their research a list of topics that needed to be handled when trying to transform between two horizontal geodetic datum. The authors conclude that these transformations are normally accomplished through the geocentric coordinate system (inter-mediate calculation), while they include the transformation parameters of the two datums. The authors clearly state that several problems might be encountered when trying to put these transformations together.

Several papers, such as (Yanalak et. al. 2005) or (Lippus, 2004) suggest a local transformation-parameters estimation that exist between maps, in order not to coerce a global one. Though these prove to be accurate and reflect correctly local trends exist, these mechanisms are usually time consuming and demand algorithmic complexities, especially in wide areas, such as the one used in the case study presented here.

5. CONCLUSION

Different maps and datasets are given in different reference coordinate systems based on different geodetic projection and datum. In order to be able using simultaneously several datasets, geo-oriented systems have to apply real-time data manipulation. These are essential in order to support immediate and precise processes and decision making.
Direct coordinate and datum transformations include calculation complexity, mainly as these algorithms are based on high order projection and datum transformation formulae that involve several phases, iterations, and a great number of mathematical and geometric considerations: ellipsoids, projections, numerous parameters declarations - just to name a few. As a result, this is usually time consuming and requires large amounts of computer resources and storage capacity, and hence not always appropriate for all systems.

A novel "transformation" method was presented here that simplifies the complexities mentioned before, and suggests a faster process while preserving the required accuracy needed from the process. A short pre-processing procedure is required, which produces matrices of translation values between two coordinate systems, that proved not to be time consuming and does not require special computer resources. By implementing a designated interpolation algorithm to handle correctly the values stored in these matrices, it was feasible to calculate immediately the exact 2D transformation for any required location within the maps bounds, and with no calculation complexity. As a result, a location identified in one map is easily transformed precisely to its location in the other map regardless differences in their geodetic projections and/or datums. It is also worth noting that this concept is free from any transformations complexities that might exist otherwise.

A case study showed that this process preserves the accuracy needed for any geodetic implementation while shortening the time needed for the process - in respect to the common direct transformations. Though the suggested concept was implemented on coordinate systems defined by Lambert Conformal conic projection and Transverse Mercator one, it is clear that it can be utilized to any given type of datum and projections defined by the coordinate systems that are needed for transformation.

In this paper the concept was analyzed based solely on 2D coordinate systems, but with slight modifications it can be evaluated for any N-dimensional transformation, such as 3D one that utilizes 3 translations and 3 rotations (and even scaling). It is worth noting that special considerations must be taken when interpolating rotations. Furthermore, in case where the direct transformation needed for pre-processing grid assembly is not known, the suggested concept can still be implemented. Here, a process that will identify counterpart unique entities that exist in both given maps can be utilized, hence calculate the translation values exist between the two coordinate systems. It is worth noting that this process produces irregular geo-registration structure, hence an adaptation to a grid is required - resulting in a geo-registration matrices - that will enable to utilize the interpolation mechanism presented here. The proposed concept can then be implemented, while preserving the required geodetic accuracy.

REFERENCES


Wisconsin Spatial Reference Systems, 2007, Published by State's Cartographer Office.


BIOGRAPHICAL NOTES

Sagi Dalyot is currently a PhD candidate in Mapping and Geo-Information Engineering at the Technion - Israel Institute of Technology, where he received his MSc (2000). Until 2004 he was involved in geodetic and mapping projects within a private IT company. His main areas of interest are three-dimensional surface modeling, LiDAR and computer vision.

Ariel Gershkovitch is an undergraduate student in Mapping and Geo-Information Engineering at the Technion - Israel Institute of Technology.

Prof. Yerach Doytsher graduated from the Technion - Israel Institute of Technology in Civil Engineering in 1967. He received a M.Sc. (1972) and D.Sc. (1979) in Geodetic Engineering also from the Technion. Until 1995 he was involved in geodetic and mapping projects and consultation within the private and public sectors in Israel. Since 1996 he is a faculty staff member in Civil and Environmental Engineering at the Technion, and is currently the Dean of the Faculty of Architecture and Town Planning. He also heads the Geodesy and Mapping Research Center at the Technion.
CONTACTS

Sagi Dalyot
Mapping and Geo-Information Engineering
Faculty of Civil and Environmental Engineering
Technion – Israel Institute of Technology
Technion City
Haifa 32000
ISRAEL
Tel. +972-4-8295838
Fax +972-4-8295708
Email: dalyot@technion.ac.il

Ariel Gershkovich
Mapping and Geo-Information Engineering
Faculty of Civil and Environmental Engineering
Technion – Israel Institute of Technology
Technion City
Haifa 32000
ISRAEL
Tel. +972-52-6612134
Email: sarielg@t2.technion.ac.il

Prof. Yerach Doytsher
Mapping and Geo-Information Engineering
Faculty of Civil and Environmental Engineering
Technion – Israel Institute of Technology
Technion City
Haifa 32000
ISRAEL
Tel. +972-4-8294001
Fax +972-4-8295641
Email: doytcher@technion.ac.il