Robust rectification of aerial photographs in an open source environment

Duccio Rocchini, Markus Metz, Alessandro Frigeri, Luca Delucchi, Matteo Marcantonio, Markus Neteler

A R T I C L E   I N   F O

A B S T R A C T

Aerial photographs provide the basis for developing indices of landscape composition and structure as sensitive measures of large-scale environmental change over relatively long periods of time. In view of this, proper image rectification is needed to enable geometrically unbiased application of landscape metrics in order to obtain meaningful results. It is also particularly important to provide researchers with image rectification tools within an open source environment, in order to: (i) guarantee free and robust tools for processing remote sensing data, (ii) facilitate customization, and (iii) provide useful support via forums and email lists. In this paper we provide a complete description of a robust and freely licensed toolchain for orthorectifying images, which is available in the open source software GRASS GIS. We will first sketch the theoretical background behind rectification and then illustrate the workflow of the orthorectification procedure in GRASS GIS.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Remote sensing is a powerful tool for the exploration of the Earth surface since it provides a synoptic view of an area with a high temporal resolution (Gillespie et al., 2008). As an example, multitemporal analysis has been widely used for studying the dynamics of ecological and environmental processes. The literature encompasses several fields of study such as soil loss (e.g., Disperati et al., 2001), natural resources assessment (e.g., Geri et al., 2010), vegetation, and ecological dynamics (e.g., Pelorosso et al., 2009).

Aerial photographs provide the basis for developing indices of landscape composition and structure as sensitive measures of large-scale environmental change (Kepner et al., 2000), and their quantitative analysis is facilitated by geographic information systems (GIS), which offer many new possibilities in this field of research (Baimiters, 1996). In order to be able to overlay remotely sensed data onto each other (image-to-image) or onto topographic maps (image-to-map), geometric correction is needed. Before accurate measurements based on aerial photographs can be made, distortion in the photographs must be removed (Powers et al., 1996). Since all metrics calculated on a landscape (e.g., from land use class coverage to patch shape and size) are sensitive to geometric distortion, geometric correction of aerial photographs is required for a realistic multitemporal study. Thus, orthogonal projection of all points of the image to a reference surface (system) must be carried out in order to correct all types of distortion in the photos (Novak, 1992).

There are several problems with geographical data and their analysis, such as spatial resolution (Jelinski and Wu, 1996) and the definition of thematic classes (Ricotta, 2005). However, few studies have explicitly dealt with the effects of image rectification on the measure of landscape patterns (e.g., Rocchini, 2004).

The spread of software for GIS and remote sensing data analysis could lead to two basic problems. First, an uninformed use of GIS can lead to the application of improper rectification algorithms, which could corrupt subsequent steps of the analytical process. Second, the cost of both software licensing and imagery puts it out of the reach of many researchers (Gillespie et al., 2008), especially those located in developing countries where the need is perhaps greatest (Nagendra and Rocchini, 2008). The development of free tools for robustly rectifying images could, therefore, provide researchers with a valuable resource for landscape analysis. This way, the access to the code allows study of the algorithm implementation, enabling a reproducible environment for processing remote sensing data.

The aim of this paper is to provide researchers with a complete description of a robust and free tool for orthorectifying images. We will firstly sketch the theoretical background behind
rectification and then illustrate the workflow of the orthorectification procedure in GRASS GIS with a working example.

2. Theory behind rectification

Novak (1992) presented an interesting review of the methods used for geometric correction of remotely sensed data. He classified mathematical models used for this kind of correction into three classes: polynomial, projective, and differential rectification. Briefly, the first two classes use polynomial functions, namely mathematical algorithms, which define, for a given point, the transformation from the original coordinates of the point to its rectified coordinates.

The transformation function is defined by ground control points (GCPs) on the incorrect image and their corresponding locations on the map or on an orthorectified image. Once the GCPs have been chosen, polynomial rectification is based on a curve fit, where the order of rectification is defined as the maximum exponent used in the polynomial. For instance, first- (Eq. (1)) and second-order (Eq. (2)) polynomial transformations are usually applied as

\[
\begin{align*}
    x &= a_0 + a_1 x + a_2 y \\
    y &= b_0 + b_1 x + b_2 y
\end{align*}
\]

(Eq. (1)) and

\[
\begin{align*}
    x &= a_0 + a_1 x + a_2 y + a_3 x y + 2 a_4 x^2 + 2 a_5 y^2 \\
    y &= b_0 + b_1 x + b_2 y + b_3 x y + 2 b_4 x^2 + 2 b_5 y^2
\end{align*}
\]

(Eq. (2)) polynomial transformations are usually applied as

where \( x \) and \( y \) are the coordinates of each GCP of the incorrect input image, \( x' \) and \( y' \) are the coordinates of the rectified image, and \( a \) and \( b \) are coordinate transform coefficients. Following common notation, the coordinates of input GCPs \( x, y \) can be translated by a rectification function based on a matrix of coefficients \( A(x) \) for the \( x \) coordinate and \( B(y) \) for the \( y \) coordinate, whose dimensions \( \text{Dim}(A) \) and \( \text{Dim}(B) \) depend on the exponent used in the polynomial function (Konecny, 1979; Novak, 1992).

In general terms, the higher the order of polynomial function, the greater the possibility of correcting images for more complex types of distortion. On the other hand, it is unusual to have terms beyond second power (Russ, 2002). In fact, while higher order polynomials are accurate in the proximity of GCPs, they can lead to significant errors in regions outside the GCP range (Richards and Jia, 1999).

While polynomial rectification can model different kinds of distortion, including translation in \( x \) and \( y \), scale changes in \( x \) and \( y \), skew, and rotation (Jensen, 1996), it cannot correct relief displacements, because no information regarding GCP elevation is provided. Differential rectification, also referred to as orthorectification, takes into account the elevation of the area under study by using a digital elevation model (DEM) or the elevation at each GCP measured by, e.g., a GPS. Following the Konecny (1979) notation, the input elevation \( z \) of each GCP is derived as a function of its spatial location \( z = f(x, y) \) with respect to a DEM grid defined as \( d(x, y) \). In simple terms, orthorectification differentially corrects one part of the aerial photograph at a time as a function of its local elevation, i.e., the elevation of each GCP. This is integrated into the so-called collinearity equations coupling \( z \) with \( x, y \) into the rectification procedure and including additional parameters related to the camera being used, such as its focal length in millimeters or the coordinates of the perspective center, which we will discuss in the following sections (see Section 4, GRASS procedure for orthorectification). Konecny (1979), Novak (1992), Wolf (1983), and Toutin (2004) provide additional mathematical details. In this paper, we will focus on aerial photographs, although the same reasoning can be applied to satellite images once scan-line properties instead of camera parameters are adopted (e.g., Novak, 1992).

Fig. 1. Root mean square error (RMSE) can be viewed as the Euclidean (Pythagorean) hypotenuse of the right-angled triangle, whose catheti are the residual values \( \mu \) and \( \nu \) in the \( x \) and \( y \) axes, respectively (top left). The scatterplot represents residuals in the \( x \) and \( y \) axes achieved by orthorectification (triangles) versus polynomial transformations (circles) using the same ground control points (GCPs). Orthorectification generally outperforms polynomial transformations leading to a lower RMSE. Refer to Rocchini and Di Rita (2005) for an empirical example.

The output bias derives from the error implicit in the rectification procedure being used or in discrepancies between the GCPs recognized in the input photo and the reference map. Output bias can be measured by root mean square error (RMSE). Formally, let \( i \) denote each ground control point (GCP), then RMSE can be expressed as

\[
\text{RMSE}_i = \sqrt{\mu_i^2 + \nu_i^2}
\]

where \( \mu_i \) and \( \nu_i \) are the residuals for the GCP \( i \) in the \( x \) and \( y \) axes, respectively. Geometrically speaking, RMSE represents the Euclidean (Pythagorean) hypotenuse of the right-angled triangle whose catheti are the residual values in the \( x \) and \( y \) axes (Fig. 1).

Once \( n \) GCPs have been collected, the total RMSE is then derived as

\[
\text{Total RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \text{RMSE}_i}
\]

where \( \text{RMSE}_i \) is the error associated with each \( i \)th GCP (Eq. (3)).

As previously stated, orthorectification is the only method which guarantees low spatial error in the geometric properties (e.g., object displacement) of remotely sensed images, taking into account the geomorphological complexity of the landscape under study (Toutin, 2004). Rocchini and Di Rita (2005) empirically demonstrated that, given the same GCPs being used, orthorectification outperforms polynomial transformations leading to a lower RMSE (Fig. 1). Hence, due to its implicit robustness for geometric correction of remotely sensed data, we will rely only on orthorectification in the present paper.

3. Open source GIS: the GRASS GIS software

The idea of open source software has been around for almost as long as software has been developed (Neteler and Mitasova, 2008). The famous “four freedoms” paradigm, developed by
Stallman (1997) in his seminal work which proclaims (i) the freedom to run the program for any purpose, (ii) the freedom to study how the program works and adapt it to one’s own needs, (iii) the freedom to redistribute copies, and (iv) the freedom to improve the program and release such improvements to the public, guarantees that the whole community benefits from software development (also see Fogel, 2009).

It is worth noting that full access to source code is crucial in GIS to ensure robust analysis output, particularly where complex algorithms are concerned (Mitasova and Neteler, 2004). There are well-known examples of Open Source software in other research fields such as statistics (e.g., R Language and Environment for Statistical Computing, R Development Core Team, 2010), while the GIS community has the benefit of a powerful tool named Geographical Resources Analysis Support System (GRASS), which includes more than 350 modules for managing and analyzing geographical data (Neteler and Mitasova, 2008; Neteler et al., in review). GRASS was created in 1982 by the U.S. Army Construction Engineering Research Laboratories, and nowadays it is one of the cutting-edge projects of the Open Source Geospatial Foundation (OSGeo, founded in 2006).

The adoption of the free open source software (FOSS) license in 1999 and the introduction of an online source code repository (Concurrent Versioning System) changed the development process of GRASS, thus allowing worldwide contributions from the scientific community. The modular software design of GRASS facilitates the introduction of new functionalities without affecting the overall performance of the system. Moreover, its scripting capabilities enable automated processing of a large volume of data and a wide-ranging use of the achieved results. In particular, recent developments also allow GRASS users and developers to make use of the Python programming language (Van Rossum, 1995, 1997) to introduce new features.

4. GRASS GIS procedure for orthorectification

GRASS GIS contains a menu-based module named orthophoto for photo orthorectification. The orthorectification procedure in GRASS uses three different coordinate systems: that of the scanned image in pixels (image coordinate system), that of the camera sensor in millimeters (photo coordinate system), and the real-world coordinate system defined by projection parameters (target coordinate system). Orthorectifying an aerial photograph in GRASS involves passing from the image coordinate system to the photo coordinate system (interior orientation) then to the target coordinate system (exterior orientation and rectification), based on the following steps (Fig. 2).

**Step 1: Main parameters input.** First, the GIS operator must indicate the main input parameters used for subsequent orthorectification: (i) the reference system (named “target location” in GRASS) into which the image will be georeferenced, (ii) the DEM used for deriving the elevation of GCPs, and (iii) the camera parameters. As a rule of thumb, a DEM of 10 m is used for photos with an instantaneous field of view (IFOV) of ca. 1 m (e.g., Rocchini and Di Rita, 2005). Camera parameters include: (i) the focal length in millimeters (mm), (ii) the coordinates in millimeters in the photo coordinate system of the principal point of symmetry, which are ideally but not necessarily [0,0], (iii) the number of fiducial marks (usually four or eight) being used for interior orientation (see Step 2: Interior orientation), (iv) the coordinates of the fiducial marks in millimeters in the photo coordinate system. As an example, in the commonly used 230 mm × 230 mm aerial photograph format (Morgan et al., 2010), the four diagonally located fiducial marks should have the following coordinates $F_{SE}$[115, 115], $F_{SW}$[115, −115], $F_{NW}$[−115, −115], and $F_{NE}$[−115, 115]. The operator must graphically individualize them into the scanned image during the interior orientation process. Refer to the main text for additional information.

**Step 2: Interior orientation.** The interior orientation of the photograph is needed to relate the image coordinate system (in pixels) to the photo coordinate system (camera sensor). This is done by graphically marking the fiducial marks (Fig. 3) into the scanned image. Thus, GRASS computes the image-to-photo coordinate transformation parameters.

**Step 3: Exterior orientation.** Exterior orientation of the aerial photograph is based on the selection of $n$ GCPs, taking into account how the photograph is based on the selection of $n$ GCPs, taking into

account their $x$ and $y$ (coordinates of each GCP of the incorrect input image in the image coordinate system), their $x'$ and $y'$ (corresponding GCP coordinates of the rectified image in the target coordinate system defined by projection parameters), and $z$ coordinates (Fig. 4, see Section 2, Theory behind rectification). This is frequently carried out using ancillary information such as topographical maps, GPS field coordinates, and orthorectified photographs (Figs. 2 and 4). Elevation ($z$) of each GCP is derived directly by GRASS from the previously specified DEM.

Those GCPs with a higher bias (RMSE) can be highlighted and rechecked. If they show geometric incongruity with reality they can be removed. Generally, operators tend to eliminate those GCPs that affect total RMSE (Jensen, 1996). GCPs should be distributed over the whole area to ensure an even coverage of the image and to avoid clustering effects. Regarding the number of GCPs to be used, Bernstein (1983) proposed that 16 is a reasonable number if each can be located with an accuracy of one-third of a pixel. Of course, the required number of GCPs is itself a function of the rectification method used. RMSE (Eq. (3)) is calculated by GRASS GIS in map units using the output coordinate system. For instance, if a UTM system has been specified, meters are used as units to report RMSE. As previously stated, according to Eq. (3), the RMSE for each point $i$ represents the Euclidean (Pythagorean) hypotenuse of the right-angled triangle whose catheti are the residual values in the $x(\mu_i)$ and $y(\nu_i)$ axes. As an example, a point with $[\mu_i, \nu_i]=[2.9,4.8]$ meters would bring a contribution of $RMSE_i = \sqrt{\mu_i^2 + \nu_i^2} = \sqrt{8.41 + 23.04} = 5.6m$ to the total RMSE. GRASS GIS reports both the RMSE of each point $i$ and the total RMSE according to Eq. (4).

Step 4: Rectification and resampling. The parameters computed during the interior and exterior orientation steps are used to rectify the image, ultimately passing from the input $x, y$ image coordinates to the $x', y'$ standard map coordinates in the target projection (Fig. 2). The whole image is then resampled with nearest neighbor assignment. Bilinear interpolation and bicubic interpolation by cubic convolution are also possible resampling algorithms. The final result is an orthorectified image in a standard map coordinate system compensated for relief distortions together with the associated RMSE (Figs. 2 and 5). Fig. 5 shows a qualitative example of the output produced by GRASS GIS, using a 1954 aerial photo from Monte Baldo, Province of Trento, in the (Italian) Alps environment, which represents one of the most geomorphologically complex environments worldwide (Rocchini et al., in press).

The supremacy of orthorectification over linear (polynomial) transformation is proven by explicitly comparing the results achieved using the aforementioned aerial photograph (Fig. 6).

Qualitatively speaking, due to the high elevation gradient of the whole photograph which ranges from ca. 130 m to ca. 1500 m, high discrepancies between topographic objects and their corresponding points on the photo were found in both valleys (rectangle A, note the overlap between the street and the river) and upper mountains (rectangle B, note the distance between topographic and photo-based street), when relying on linear polynomial transformation (left panel). Such discrepancies disappeared using orthorectification (right panel).

Quantitatively speaking, this empirical example confirmed the previously described theoretical foundation (Fig. 1). Orthorectification showed a clumped distribution of errors (see the theoretical Fig. 1 and the empirical Fig. 6) compared to first-order polynomial rectification, which, on the contrary, showed high RMSE reaching up to 20 times that achieved by orthorectification.

For the sake of clarity, in this example only a first-order polynomial transformation was compared to orthorectification to explicitly show the benefit of using elevation in the transformation process, but the same concepts hold considering higher order polynomials (data not shown). This example clearly demonstrates that also in morphologically complex environments such as mountain areas GRASS GIS can provide for accurate registration thanks to robust orthorectification.

5. Conclusion

The recent development of user-friendly tools and interfaces to perform simple rectification could lead to misleading input information being fed into subsequent analyses. As stressed by Rocchini and Di Rita (2005), it is essential to explicitly refer to the error (e.g., RMSE) associated with each orthorectified image, which should be accounted for during further analytical steps such as classification, land use map generation, change detection, and landscape metrics calculation.

This paper focused on aerial photography given the lack of long-term satellite image coverages for, e.g., change detection of landscape dynamics over long periods (Morgan et al., 2010). Nonetheless, the same concepts of geometric correction expressed in this paper basically also apply to satellite imagery. Thus, in general terms, without implicit assumptions on remote images.

Fig. 4. Exterior orientation. Ground control points (GCPs) are used to relate image coordinates to real-world coordinates. In this case, an orthorectified aerial photograph (taken in a different year, right panel) is used for choosing the GCPs.

sensing data acquisition and analysis, further discussion of the results obtained is merely trivial (Duggin and Robinove, 1990). Although it may not be necessary to orthorectify images simply to identify the changes between two images (Jensen, 1996), no quantitative information should be extracted from geometrically incorrect photographs until proper tests on geometric accuracy have been carried out. Geometric distortion may alter both (i) the perceived location of features over a landscape (Morgan et al., 2010) and (ii) their size. As an example, Aronoff (2005) showed that landscape objects at higher elevations may seem larger than expected due to relief displacement problems. Furthermore, Rocchini (2004) demonstrated that, during multitemporal analysis, changes in some classes over time could be overestimated by up to double the occupied area if geometric distortions are corrected using an improper rectification method. In this view, GRASS GIS offers an accurate and inexpensive means for obtaining geometrically correct images together with the associated error also in morphologically complex landscapes (Figs. 5 and 6).

Mountainous environments have always presented a challenge for achieving geographically accurate data in a number of cases such as digital elevation model generation (Eckert et al., 2005), weather data interpolation (Roiz et al., 2011), vegetation classification (Zhang et al., 2011), and geometric correction of remote sensing data (Aronoff, 2005). Considering the last point, we provided an explicit proof about the robustness of GRASS GIS for correcting images also in morphologically complex conditions.

A number of studies have demonstrated that GRASS GIS is a powerful tool in many research fields and scientific tasks, such as geomorphology (Grohmann, 2004), landscape ecology (Baker and Cai, 1992; Steinger and Hay, 2009), climate change (Baker et al., 1991), hydrological modeling (Carrera-Hernández and Gaskin, 2008),

Fig. 5. Result of the orthorectification procedure. The parameters computed during the inner and exterior orientations are ultimately used to rectify the image to a standard map coordinate system. The final result is a geometrically correct image with low geometric distortion in both mountainous and flat areas. In this example, using a 1954 aerial photo from a Northern Italy Alpine environment (Monte Baldo, Province of Trento) obtained at a flight altitude of 9000 m (approximate scale 1:33,000).
computer graphics (Sorokine, 2007), geostatistics (Hengl et al., 2008),
distance analysis (Greenberg et al., 2011), and remote sensing
(Rizzoli et al., 2007). The integration of remote sensing tools into
an Open Source Geographic Information System like GRASS GIS is a
major advance over stand-alone solutions (Neteler et al., 2005). This
is true for both image data gathering and statistical analysis,
especially in view of the recent interfacing capabilities of GRASS
GIS with R (R Development Core Team, 2010) based on the
spgrass6 library (Bivand, 2000, 2010; Bivand et al., 2008).

As reported by Stein and van der Meer (2001), spatial statistics
for monitoring landscape changes with remote sensing are being
increasingly used. Developing robust open source software is there-
fore important in order to: (i) guarantee robust algorithms for image
processing, (ii) facilitate customization, and (iii) provide useful
support via forums and email lists (Steiniger and Hay, 2009). With
this paper, we make it known that free and robust tools for image
processing, including image orthorectification, are now available in
an open source environment, bringing remote sensing into com-
pliance with the celebrated “four freedoms” paradigm.

Acknowledgments

We are grateful to the Associate Editor Steve Wise handling
this manuscript, Carlos H. Grohmann, and an anonymous reviewer
for the stimulating comments made on a previous draft of the
paper. The original lortho.photo code was devised by Mike Baba.
D.R. is partially funded by the Autonomous Province of Trento
(Italy) within the ACE-SAP project (University and Scientific
Research Service regulation number 23, June 12, 2008). We thank
Tessa Say for assistance in the English editing of this manuscript.

References

Baker, W.L., Cai, Y., 1992. The r.le programs for multiscale analysis of landscape
of climatic change on the structure of landscapes subject to large disturbances.
Ecological Modelling 56, 109–125.


Bivand, R.S., 2010. spgrass6: Interface between GRASS 6 and R. R package version 0.6-19.


