Engineering Geology 172 (2014) 26–40

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

Target Detection and Tracking of moving objects for characterizing landslide displacements from time-lapse terrestrial optical images

J. Gance [a,b,*], J.-P. Malet [b], T. Dewez [a], J. Travelletti [b,c]

[a] BRGM, Bureau de Recherches Géologiques et Minières, 3 Avenue Claude Guillemin, 45060 Orléans, France
[b] Institut de Physique du Globe de Strasbourg, CNRS UM7516, EOST/Université de Strasbourg, 5 rue René Descartes, 67084 Strasbourg Cedex, France
[c] BEG, Bureau d'Etudes Géologiques SA, Rue de la Printse 4, 1994 Aproz, Suisse

ARTICLE INFO

Article history:
Received 4 July 2013
Received in revised form 17 January 2014
Accepted 20 January 2014
Available online 29 January 2014

Keywords:
Displacement monitoring
Time series images
Terrestrial optical images
Landslide
Target Detection and Tracking

ABSTRACT

Terrestrial Optical Photogrammetry (TOP) is a low-cost monitoring technique that is commonly used in change detection studies. For landslide monitoring, image correlation techniques are frequently used to generate Digital Surface Models (DSMs) from stereo-pairs or to measure the 2D surface displacement field from single-view time-lapse sequences. Image correlation allows for detecting spatially continuous displacement fields at a sub-pixel precision. The technique, however, has several limitations for assessing displacement because 1) of its sensitiveness to changes in texture, shape and radiometry in the image pile, 2) of the need of spatially regular sampling grids, and 3) of possible high computation time that can impede the processing of large image datasets.
To address these limitations, an alternative and complementary approach based on a Target Detection and Tracking (TDT) algorithm is proposed for a rapid calculation of the displacement of targets in image time series. The TDT code, developed as a MATLAB-based tool, is able to track natural or man-made targets. The precision of the TDT code is assessed using several image time series acquired at the Super-Sauze landslide (Southern French Alps) and compared to ground based measurements. The computed relative accuracy is between 10^{-3} and 10^{-4} (5 cm at a distance of 115 m). Although the TDT approach does not provide spatially continuous information, it provides 1) a quantification of the object displacements at the same order of precision as image correlation (sub-pixel accuracy) and 2) information in regions where image correlation fails because of too large ground displacements. A sensitivity analysis reveals that the major sources of uncertainty are camera movement and/or lens distortion and not the TDT method itself.

1. Introduction

Digital passive sensors (i.e., sensors that operate in the visible light spectrum), such as Single Lens Reflex (SLR) cameras, are increasingly being used for ground-based geohazard monitoring (e.g., ice glaciers, volcano flanks, landslides) partly because of their low cost compared to expensive terrestrial laser scanner (TLS) or radar imaging (GB-InSAR) systems (Corsini et al., 2006; Jaboyedoff et al., 2010). Indeed, due to the large consumer market, sensor resolution is increasing rapidly; for example, inexpensive 9 to 16 megapixel cameras became available in 2013, allowing arrays of cameras to be set up in the field.

Passive optical sensors provide qualitative information (i.e., identification of large changes in the landscape, weather conditions, presence and position of snow cover; Cardenala et al., 2008; Coe et al., 2009; Sturzenegger and Stead, 2009) and quantitative 3D information using stereo-views (i.e., creation of Digital Surface Models (DSMs), displacement monitoring, and tracking changes in surface state).

In the case of surface displacement monitoring, most recent research has focused on the development of image correlation techniques to determine the average spatial shift by maximizing a cross-correlation function between at least a pair of stereo-images (White et al., 2003). The technique has proven its performance for characterizing the displacement fields of ice glaciers (Maas et al., 2008) and slow-moving landslides (Delacourt et al., 2007; Travelletti et al., 2012) at sub-pixel accuracy (0.2 pixels; Casson et al., 2005; Delacourt et al., 2007) and generates a pseudo-continuous map of the deformation.

However, the performance of image correlation techniques decreases in the case of large changes in radiometry, texture or in the object geometry between two different dates. Therefore, it is difficult to assess complex displacement patterns combining rigid and non-rigid deformation (Lewis, 1995), rapid movements characterized by important surface state changes (Travelletti et al., 2012), or the movement of objects that causes significant perspective distortions (Lewis, 1995). These features can be encountered on certain landslides. Moreover, image correlation can be very sensitive to the illumination conditions (e.g., sun elevation) and low signal-to-noise ratios in the image. The choice of the optimal combination of image correlation processing

* Corresponding author at: BRGM, Bureau de Recherches Géologiques et Minières, 3 Avenue Claude Guillemin, 45060 Orléans, France.
E-mail addresses: j.gance@brgm.fr (J. Gance), jeanphilippe.malet@unistra.fr (J.-P. Malet), t.dewez@brgm.fr (T. Dewez), julien.travelletti@beg-geol.ch (J. Travelletti).
URL: http://www.elsevier.com (J. Gance).

0013-7952/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.enggeo.2014.01.003

parameters (number of images to correlate, size of the correlation window, number of hierarchical steps in the processing, correlation thresholds) must be carefully evaluated for each case study and limits the ability to automate the process, especially if a rapid analysis is needed because of accelerated movement. This constraint is amplified by the usual application of a regular spatial sampling grid on the image (Walter, 2011; Travelletti et al., 2012). The image correlation technique is therefore applied on random points, which are not always remarkable and which may disappear with time or loose their radiometric and textural properties. The results may lead to inconsistent displacement estimates in the area of interest.

For the analysis of long time series, Target Detection and Tracking (TDT; Veeraraghavan et al., 2006) methods may be used to complement image correlation techniques. TDT methods are used to estimate the displacements of discrete features over time, either natural (rock blocks, large fractures) or man-made (benchmarks). The selected objects should stand out by their radiometric properties and ensure their effective and accurate detection. The main advantage of TDT methods over image correlation techniques is the ability to measure complex displacement patterns over space and time for large deformation, high signal-to-noise ratios, changes in the target's shape and the lag of images over time for low view angles. Moreover, the TDT method can be used for rapid assessments, such as in early-warning systems. The main limitation of the TDT method is the possibility of quantifying only local (point-based) displacements at the location of the target. Designs for real-time surveying systems that incorporate TOP have been proposed for monitoring volcanic activity (Honda and Nagai, 2002), but their potential use for monitoring slope movement is restricted to cases where daytime observations offer sufficient information.

This work presents a MATLAB-based TDT method that is adapted to landslide monitoring and can be used as a complement to image correlation and other displacement observation methods or for rapid assessments. The proposed method is able to robustly track a moving target on a background that undergoes large displacements and is only slightly sensitive to target texture changes and rotations.

First, the steps of the method are presented, which include (i) image stack registration, (ii) target detection, and (iii) automatic tracking of the target in the 2D image plane. Second, the method is applied and validated with two datasets acquired at the Super-Sauze landslide (South French Alps) using several natural and man-made targets. Third, the sources of errors are quantified. Finally, guidelines to set up a TOP-based monitoring system are proposed to limit the errors arising from the acquisition geometry and the camera movement.

2. Methodological framework

This section presents the criteria to consider when using a TDT method. The criteria are adapted to the characteristics of the study sites in terms of the geometrical constraints for the camera location, the amplitude and the direction of the displacement vector, and the type of targets in view.

2.1. Criteria and possible approaches for setting up a landslide-oriented TDT method

In his review of the different categories of object tracking algorithms, Yilmaz et al. (2006) identify four questions that can be used to select the most efficient TDT approach according to the type of objects to be detected and the type of deformation to be quantified:

- How is the tracked object represented (i.e., by a line, a point, a surface or by patches)? For landslide displacement monitoring, the TDT algorithm can be applied to simple surface objects (circle, ellipse, rectangle). In the proposed approach, the objects are reduced to a point that is defined as the centroid of the elementary surface. This choice is particularly well suited for small objects that occupy very small regions in the image plane (Yilmaz et al., 2006).
- Which features are tracked? For landslide displacement monitoring, the use of edge and texture information for tracking is difficult because of the presence of highly heterogeneous surfaces. Therefore, objects are tracked using their radiometric properties, which always contrast with those of the surrounding ground (e.g., soil).
- What are the criteria for selecting the most suitable object-tracking algorithm? Classical object detection algorithms such as point detectors (Moravec, 1979; Harris and Stephens, 1988; Lowe, 2004), background subtraction (Wren et al., 1997; Elgammal et al., 2000) or image segmentation (Comaniciu and Meer, 2002; Meer, 2003) cannot be used for landslide displacement monitoring because the objects may exhibit very large displacements in the image time series. Therefore, assuming that the image acquisition rate is appropriate for the ground displacement rates (i.e., such that the pixel shift between two consecutive images is relatively small), simple contrast enhancement algorithms and centroid computation algorithms are employed.
- What are the criteria for selecting the most suitable algorithm to track the object? For landslide displacement monitoring, it can reasonably be assumed that the radiometric properties of the tracked object have sufficient contrast from the surrounding material (i.e., soil). Therefore, it is assumed that the object detected in a search window in image n corresponds to the object detected in a search window centered on the object's coordinates in the previous image $n + 1$. A significant assumption in this method is the uniqueness of the object of interest in the search window.

Based on this conceptual framework, a methodological workflow to track natural and man-made objects in image time series is proposed in Fig. 1. The objective is to obtain a time series of coordinates in the image plane for each tracked object. The input data consist of image time series acquired from a unique camera, for which the radial lens distortion has been calibrated (Horaud and Monga, 1995; Clerc, 2001; this step is not presented in the simplified workflow for clarity). The images are then converted from RGB values to luminance grayscale intensity values by eliminating the hue and saturation information while retaining the luminance. In the following sections, each processing step is described in detail.

2.2. Step 1: Image registration

In many natural environments, time-lapse cameras cannot be maintained absolutely stationary for long periods of time (e.g., months) despite careful field setups. This results in a temporal shift throughout the image pile that needs to be removed. Accurate image co-registration is a key element in the application of change detection methods because it determines the reliability of the displacement estimates. Image co-registration generally consists of correcting the camera motion induced by wind or temperature variations (Goshtasby, 1989) and transferring each image of the time series into the geometry of the reference image. This operation is performed in four stages and has been completely automated (Zitova and Flusser, 2003):

1. Detection of time invariant points in the image: Among the many algorithms available to detect invariant points (Harris and Stephens, 1988; Lowe, 2004; Bay et al., 2008), the Harris corner detector algorithm is used because it is computationally more efficient for studies that use fixed cameras (Tuytelaars and Mikolajczyk, 2008). It is based on the first order image derivatives in the x–y directions in gray-scale images. The areas in the images that correspond to the landslide are masked to register the images only on the assumed stable slopes around the landslide. A cross-correlation in the Fourier domain is then performed on the Harris corner point coordinates (Guizar-Sicairos et al., 2008) to attain a matching resolution of 0.1 pixel.
2. Feature matching: This stage consists of establishing the correspondence between the invariant points on the complete image time series. Putative matches between feature points are computed using a

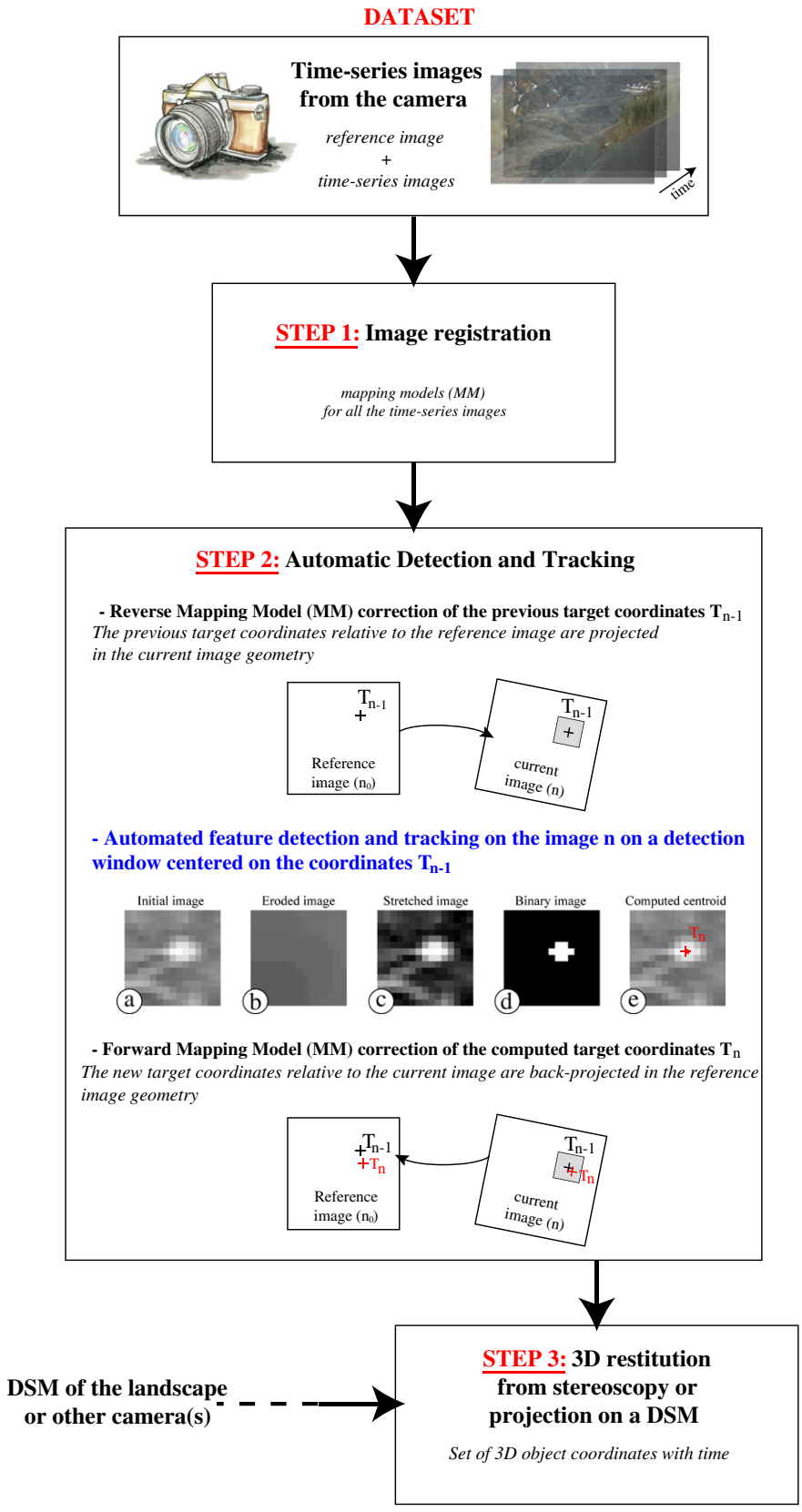

Fig. 1. Flowchart of the methodology.

correlation measure, and points that are strongly correlated in both x–y directions are paired together.

3. Estimation of the mapping model: This stage consists of estimating a global Mapping Model (MM) that best fits the Harris corner point pairs in the image time series. Even if the camera is robustly fixed to a pole cemented to the soil, residual motions of the pole may exist, and temperature cycles can alter internal camera geometry. The most appropriate model is chosen by comparing different

transformations (affine, projective, polynomial non-reflective, reflective similarity) and the associated goodness of fit in terms of the RMS error criteria. In this application, the affine transformation (i.e. translation and rotation) produces the lowest residuals.

4. Projection of the object coordinates: This stage consists of forward projecting the object coordinates (and not the full images) using the best mapping model. This approach avoids errors due to image interpolation and resampling during the projection, preserves the radiometry of the image and reduces the processing time (Zitova and Flusser, 2003).

2.3. Step 2: Automated feature detection and tracking

The objects to be detected and tracked in the image time series are selected manually with an interactive brush tool in the interface. Manual picking is performed on the reference image, and the object coordinates are refined by applying the feature detection and tracking method (Fig. 2). The properties of the target used for the automated detection technique (object size, geometry, luminance) are discussed in the following sections.

Assuming that the position of a point of interest does not change notably from one image to another (Yilmaz et al., 2006), a flexible search window is centered on the coordinates of the object from the previous image. The window size is defined so that the displacement (in pixels) of the object between two consecutive images is less than half the size of the search window. The larger the size, the greater the probability that the object can be detected in the search window but also that it can be mixed with other objects. From the initial search window (Fig. 2a), a background image is then constructed by applying a rank-order filter (Fig. 2b) in which each pixel is replaced with its darkest neighboring pixel. This mathematical morphological operation, called "erosion", only retains the large trend of the initial image (Russ, 2002). A new thumbnail is then created by subtracting the eroded image from the initial image. Radiometric stretching is applied to produce grayscale intensities ranging between 0 and 255 (Fig. 2c). This operation, called "opening", enhances the contrast, allows better object detection by thresholding, and compensates for changes in the illumination conditions (Russ, 2002). A binary image is created by applying the Otsu method (Otsu, 1975; Fig. 2d), which minimizes the grayscale variance inside the two classes assumed to make up the image. The centroid position (e.g., coordinates in the image plane) is computed as the center of mass of the previously created binary image (Fig. 2e). Finally, the coordinates, which are defined relative to the current image, are corrected by applying a radial lens distortion model and the forward mapping model.

2.4. Step 3: 3D photogrammetric restitution

For the first two steps, images were processed independently for each camera. In the third step, 3D object displacements are determined from both pair of images using photogrammetric principles or from independent DSMs (Travelletti et al., 2012). First, the intrinsic parameters (scale factor, image center, radial lens distortion) of the cameras are estimated for the different focal lengths using the PhotoModeler© (self-calibrating bundle adjustment) and Adobe® Lens Profile Creator software. Nine images of two different planar chessboards allow the inversion of the radial lens distortion using the Brown model (Brown, 1971), the focal length and the principal point of the cameras. External orientations of the cameras are then computed (three Euler angles for the orientation, local coordinates for the camera location) from the collinearity equations (Mikhail et al., 2001) using Ground Control Points (GCPs) surveyed in the field with dGNSS measurements:

$$\begin{cases} u = u_0 - \mathrm{f}\dfrac{m_{11}(X-X_0)+m_{12}(Y-Y_0)+m_{13}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)} \\ v = v_0 - \mathrm{f}\dfrac{m_{21}(X-X_0)+m_{22}(Y-Y_0)+m_{23}(Z-Z_0)}{m_{31}(X-X_0)+m_{32}(Y-Y_0)+m_{33}(Z-Z_0)}, \end{cases} \quad (1)$$

where u_0 and v_0 are the coordinates of the central point in the image plane, f is the focal length and m_{ij} is the rotation matrix.

In the case of stereo- (or multiple) views, the 3D restitution is performed by stereoscopy. At the study site, the extrinsic parameters are computed from 21 GCPs. With the camera orientation information, it is then possible to determine the local coordinates (*X*,*Y*,*Z*) of an object from two pairs of coordinates (u_{A1},v_{A1}) and (u_{A2},v_{A2}) by solving the two systems of Eq. (1) for cameras A1 and A2.

For the case of a unique camera, Travelletti et al. (2012) proposed obtaining the 3D coordinates of the objects by projecting a DSM. At the study site, two DSMs (acquired in October 2007 and July 2009) are used to define 45 GCPs and invert the external parameters from the collinearity equations (Eq. (1)). For each point of the DSM, a backward projection is then applied to calculate their associated positions in the image plane so each pixel of the image is associated with 3D coordinates.

3. Application to an image time series of the Super-Sauze landslide

3.1. Characteristics of the Super-Sauze landslide

The Super-Sauze landslide (Southern French Alps) is a flow-like landslide that develops in Callovo-Oxfordian black marls at elevations ranging from 2105 m at the crown and 1740 m at the toe. The landslide is continuously active at average displacement rates of 0.05 to 0.20 m·day^{-1}

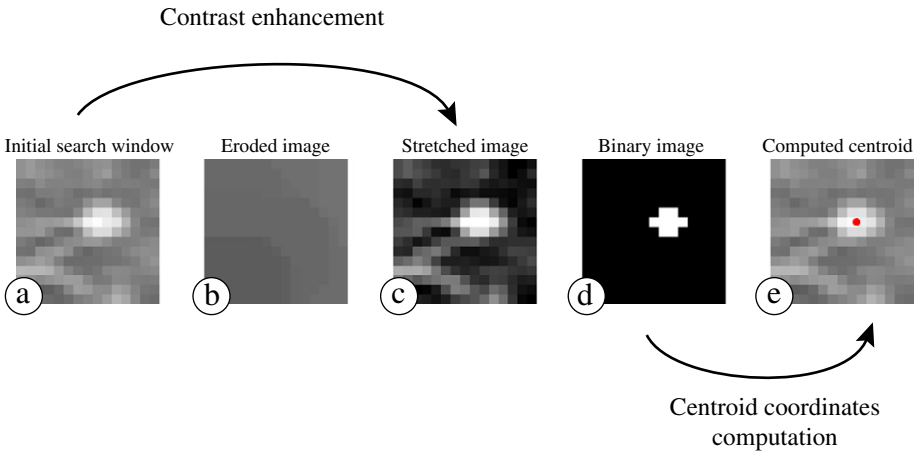

Fig. 2. Automatic tracking flowchart.

(Malet et al., 2005) but maximum rates of several meters per day have been observed (Travelletti et al., 2012). The continuous detachment of large blocks from the main scarp, and their progressive mechanical and chemical weathering to fine particles, maintain the continuous downstream propagation of material on the in-situ black marls. The landslide's activity is mainly controlled by hydrology, with motion acceleration subsequent to rainfall and snowmelt events. The bedrock topography below the slide is complex (Flageollet et al., 2000; Travelletti and Malet, 2012) and partition units with different hydrogeological, rheological and kinematical behaviors.

Surface displacements reflect this buried topography and thus provide useful information for understanding the complex sliding mechanisms. Their measurements may then serve to calibrate hydro-mechanical models.

This complex subsurface topography also induces that sparse point-like observations (measured by GNSS or total stations), beyond their maintenance cost, cannot capture the entire deformation field. Remote landslide observation is both safer for equipment and offers a better spatial observation viewpoint. We therefore installed three cameras in 2007 and 2010 to obtain a global view of the landslide (Travelletti et al., 2012).

3.2. Camera locations and characteristics

Three cameras (A1, A2, B) were installed at different stable locations. Cameras A1 and A2, looking downslope, are paired and provide a close range stereo-view of the upper part of the landslide, which exhibits the highest displacement rates. Camera B is installed on a stable crest in front of the landslide and looks uphill. The characteristics of the cameras are summarized in Table 1.

Cameras A1 and A2 are placed 75 m apart and look out a N–S cross-section composed of man-made targets (Styrofoam spheres) 50 m to 140 m from the cameras (Fig. 3a) is defined across the field of view. The geometry of the setup of Cameras A1 and A2 leads to base to height ratios B/h ranging from 1.6 to 2.1, where B is the distance between the two cameras (base), and h is the distance from the object to the baseline. The cameras are controlled by a Digisnap 2000 intervalometer (Harbortronics); each day, six pictures are acquired at 6:0, 8:00, 10:00, 12:00, 14:00 and 16:00 UTC. Each photograph is stored on a 16 Gb SD Ultra II memory card in RAW format.

Camera B is installed on a concrete pillar located on a stable crest in front of the landslide (Fig. 3a). The camera is located 900 m from the main scarp and 300 m from the lower part of the landslide. Camera B is controlled by a Campbell CR10X datalogger, which is powered by a 40 W solar panel. Each day, six pictures are acquired at 6:00, 8:00, 10:00, 12:00, 14:00 and 16:00 GMT. A complete description of the setup of camera B can be found in Travelletti et al. (2012). Image selection is applied to retain the best photo from each day.

Typical examples of gray-scale images from cameras A1, A2 and B are shown in Fig. 4.

3.3. Characteristics of the targets

The TDT algorithm can track different types of objects with different colors, shapes and sizes. Four types of targets located at distances of 50 to 400 m from the cameras are used to test the performance of the method for two periods.

The first type of targets is a group of thirteen white Styrofoam spheres (10 cm in diameter) that are placed on top of metal rods driven into the soil (Fig. 5a). The metal rods extend 30 cm above the ground and are placed every 0.5, 1.0 and 2.0 m along a N–S profile (Fig. 3d). This type of object was selected because it is particularly easy to track. The spherical shape allows a non-ambiguous center to be defined regardless the view angle. The white color is rarely present in the images, thus ensuring a relatively good contrast to the soil surface. The size of the spheres corresponds to effective pixel diameters of 5 to 25 pixels, which minimize the errors in the computation of the centroid coordinates. The targets are located between 47 m and 88 m from camera A1 and between 77 m and 140 m from camera A2.

The second type of target is an elongated and flattened ellipsoid that serves as a GNSS antenna and is located near the lower part of the profile (Fig. 5b). The antenna is white and is tracked in the same way as the Styrofoam spheres. This target was selected because daily millimetric positioning of the antenna is available (Malet et al., 2013) and can be used to evaluate the performance of the TDT algorithm. The object is located 88.4 m from camera A1 and 139.7 m from camera A2. The effective pixel sizes are 20.9 mm and 28.1 mm, respectively, and the object measures 14.2 pixels and 10.8 pixels in images from cameras A1 and A2, respectively. These properties are comparable to those of the most distant Styrofoam spheres (# 13).

The third type of targets consists of man-made one square meter metal panels that are painted with yellow and red quadrants (Fig. 5c). These targets are not always visible in the gray-scale intensity image. Therefore, an automated tracking algorithm on image I, which is computed from the color bands of the RGB image, is applied. Image I is computed as: $I = R / (R + G + B)$, where R, G and B are the red, green and blue color bands, respectively. It represents the proportion of red per pixel luminance, so the targets contrast sufficiently with the background.

The fourth type of targets consists of natural blocks observed on the slope surface (Fig. 5d). Contrary to the man-made objects, the sizes and shapes of these natural targets can vary and are not known a priori. They are between 313 m and 468 m from the camera, and they range in size between 1 and 2 m.

Table 1
Camera characteristics.

	Camera A1	Camera A2	Camera B
Location	Central stable part	West stable part	Stable crest at the toe
Camera	Pentax K200D	Pentax K200D	Nikon D70
Resolution	3872 × 2592	3872 × 2592	3008 × 2000
Effective pixels	10.2 megapixels	10.2 megapixels	6.1 megapixels
Sensor size	23.5 × 15.7 mm	23.5 × 15.7 mm	23.7 × 15.5 mm
Focal length setup	25.68 m	30.17 mm	52 mm

4. Analysis of the computed displacements

The displacements were monitored during three periods that include different displacement rates of the landslide. The first period, $P1$, was from 16th May to 8th November 2008; it began during a period of melting snow and corresponds to measured displacement rates greater than 3.5 $m \cdot day^{-1}$ that lasted for approximately 20 days. The second period, $P2$, extended from 19th May to 16th September 2009 and also began during a period of melting snow but shows displacement rates of only a few centimeters per day. Image time series from camera B are used for these two periods.

The third period, $P3$, which extended from 28th May to 6th July 2011, also started during a period of melting snow and is characterized by displacement rates of approximately 2.5 cm/day. Image time series from cameras A1 and A2 are used for this period.

For each image, the computation time to track 19 points was less than 40 s on a desktop computer. Most of the computation time is dedicated to the image registration process (approximately 30 s).

4.1. Comparisons of the TDT coordinates to ground-based GNSS observations

The object coordinates computed using the TDT approach are first compared to daily ground-based GNSS observations during period $P3$. The components of the displacements computed from cameras A1 and A2 are shown in Fig. 6.

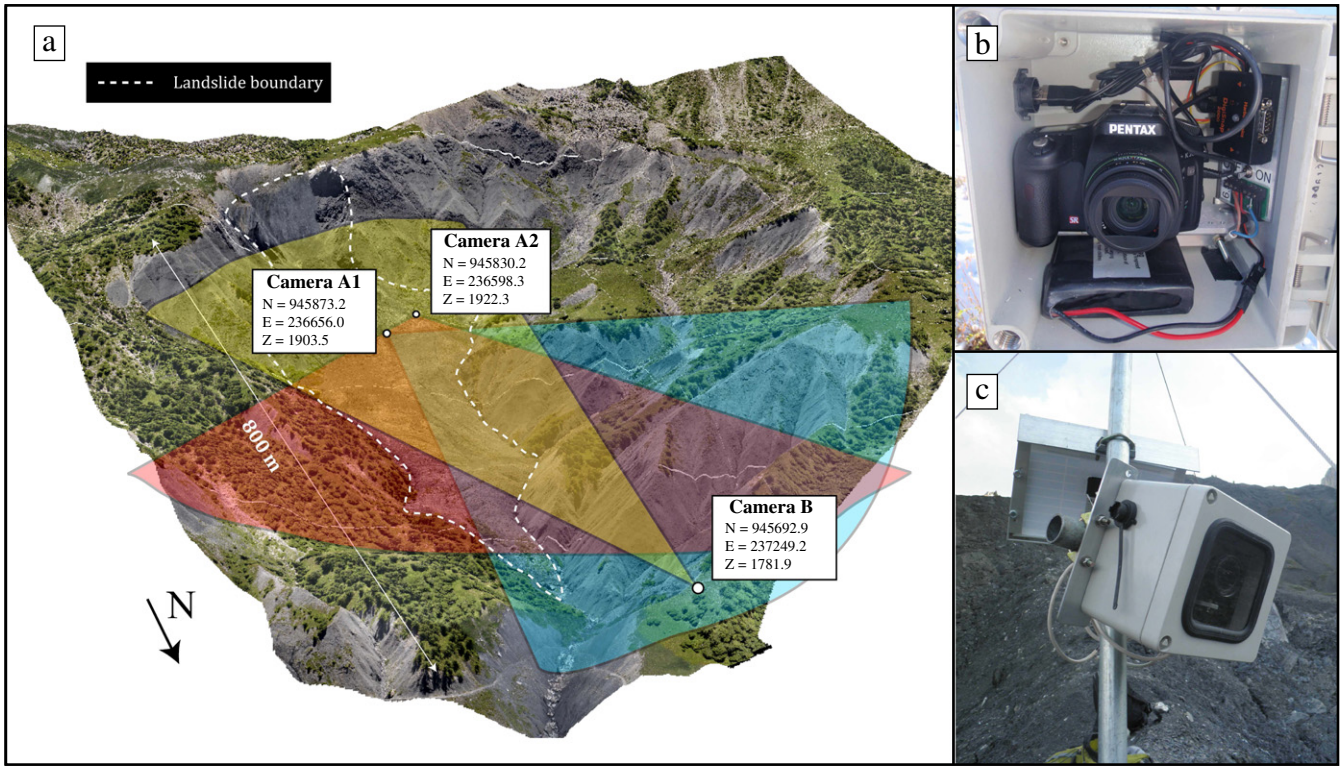

Fig. 3. Super-Sauze landslide site: a) Photo textured 3D rendering of the landslide with the three camera positions and view angles. b) Typical setup of a camera system consisting of a Pentax K200D, a Digisnap 2000 intervalometer (Harbortronics) and a battery installed in a fiberglass housing. c) Setup of Camera A1: the fiberglass housing and the 5 W solar panel are fixed to a metal pole.

The three periods all have correlation coefficients of $R^2 = 0.99$ with a relationship very close to $f(x) = x$. The positions of the GNSS antenna tracked by the TDT method are very similar to the absolute GNSS coordinates; the differences are less than 5 cm in the N and Z directions and less than 20 cm in the E direction, which is apparently less well constrained.

Another means to assess the performance is to test the bounds by tracking two static features on the stable slopes in front of and behind targets of interest. By focusing on the final 3D coordinates, this technique provides an estimate of the minimum and maximum uncertainty associated with the entire processing chain (image registration, TDT, radial lens distortion correction and 3D restitution).

Therefore, the standard deviation of the 3D coordinates is analyzed on the nearest and farthest targets on the stable slopes outside the landslide during period *P*1 (blue points, Fig. 7a). The nearest target is located 205 m from camera B, and the farthest target is located 641 m from the camera, which results in effective pixel sizes of 3.1 cm and 9.7 cm, respectively.

The standard deviation computed from the TDT coordinates, summarized in Table 2, are low compared to the distances between the camera and the objects, with maximum values of 10 cm and 8 cm in the E direction at distances of 205 m and 641 m, respectively, from the camera. Thus, the observed error is not proportional to the distance from the camera to the object and may be due to residual movements of the image that were not corrected by the image registration.

4.2. Comparisons of the TDT displacements with displacements estimated by image correlation

The computed displacements for period *P*1 are also compared to the 3D displacements obtained by the gridded image correlation technique (Travelletti et al., 2012). For the image correlation, the displacements for each time step are computed at the new block position and summed to obtain the cumulative displacements. The differences in the displacements for blocks 4, 5 and 6 are reasonable; they are less than 0.5 m for block 5 and range between 1.4 and 1.8 m for block 4 and from 2.8 to 3.1 m for block 6. This difference can be explained by their locations in different kinematic areas of the landslide. Blocks 4 and 6 are located in an area of secondary rotational slip with respect to the other blocks. Blocks 4 and 6 are rotated and progressively buried, so the image correlation technique is not able to track their movements. Block 5 is located just above this area of rotational slip and is thus less affected by this specific kinematic pattern.

Fig. 4. Image descriptions: images acquired from camera a) A1, b) A2 and c) B.

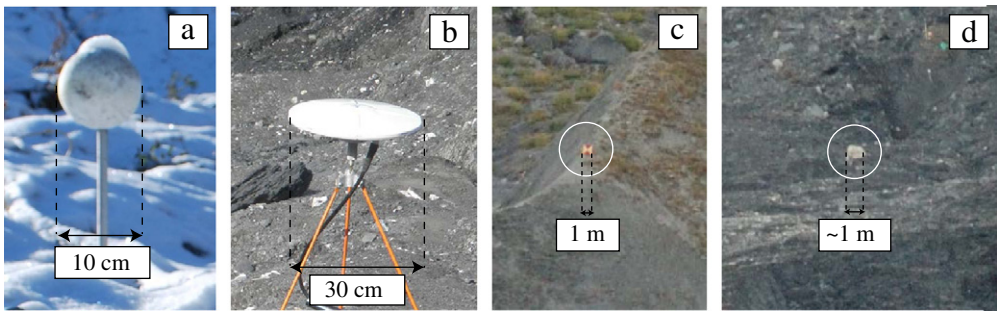

Fig. 5. Example of tracked targets: Styrofoam ball a), dGPS antenna b), metal plate targets c) and natural bright block d).

A comparison of the 2D image plane coordinates of blocks 2, 4 and 6 during period *P*1 obtained from both methods (Fig. 8) indicates large differences in the block coordinates. The position of the block computed by image correlation agrees closely with the position estimated by the TDT approach for the first dates, but blocks are rapidly lost by the IC method, which appears to track another texture pattern moving in the opposite direction. During the same period, the TDT method appears to properly track all three blocks and provides final 2D image plane coordinates that correspond to the block positions on the image. The differences can be explained by the complex motion of the blocks (e.g., rotation, fragmentation, burial) and by changes in their sizes and radiometric characteristics. For these conditions, image correlation is not a suitable technique (Lewis, 1995; Travelletti et al., 2012).

Second, the technique of Travelletti et al. (2012) was not designed for tracking blocks but for correlating regularly distributed points in images. It is possible that only the surrounding ground is characterized and that block displacements occur in the interpolated space (Fig. 8). Moreover, the accumulated displacements obtained from each image position are computed by successive additions of the horizontal and vertical displacements, which can lead to a significant propagation of errors.

By contrast, the TDT approach is able to track the displacements of specific targets with a relative accuracy of between 10^{-3} and 10^{-4}. The maximum differences are 5 cm in the N and Z directions and 20 cm in the E direction at a distance of 115 m. For larger distances, the accuracy of the TDT method can be assessed by computing the standard deviations of the 3D coordinates of targets located on stable ground, which are less than 10 cm at a distance of 587 m from the camera.

Even if no comparison to the GNSS measurements was possible during period *P*1, the 2D image plane coordinates from the TDT method are consistent (the general direction of movement is the same as the sliding direction) and visually agree with the block positions on the image. We note that the aim of Travelletti et al. (2012) was not to track specific blocks but to assess the spatial variability of the amplitude and direction of the displacements. Therefore, the results presented here do not invalidate their work but clearly illustrate the difference between the methods in terms of their accuracy and the ability to track objects; they indicate that the TDT approach is able to track specific targets on landslides that are characterized by complex deformation patterns, where the image correlation technique is not suited.

5. Discussion: the possibility of using Target and Detection Techniques for landslide displacement analysis

5.1. Target choice and size of the search window

Before designing any survey, a suitable target must be selected. As explained previously, the target's radiometric properties must contrast with those of the surrounding soil. Obviously, the more the target contrasts with the surrounding soil, the more accurate the detection will be. We recommend performing several tests before choosing the target.

As shown on the Super-Sauze landslide, the idea is to use a color with an optimal contrast from the color of the ground (white targets for dark ground or black targets for bright ground) or a color has a low occurrence in the image (e.g., red on natural landscape). The TDT method can then be used in different color spaces (grayscale or relative red content as used for the square metal panels visible in Fig. 5c). Natural targets can be used if suitable targets are present (e.g., white blocks for the Super-Sauze landslide); otherwise, artificial targets must be used.

Moreover, in the particular case of fast moving landslides, where the acquisition frequency requires displacement monitoring at night, a slightly luminous artificial target will allow monitoring displacements continuously.

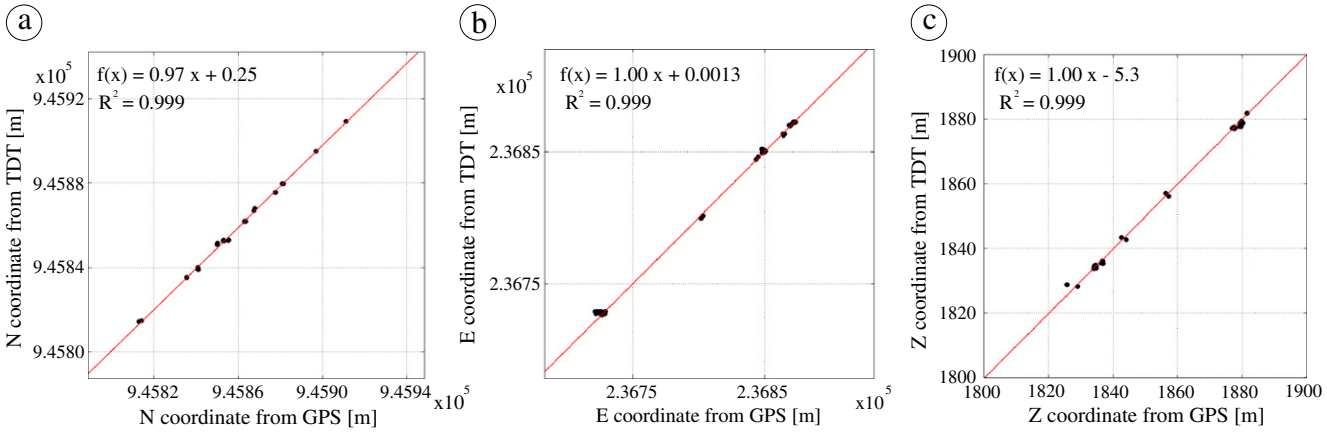

Fig. 6. Comparison of the coordinates estimated by the TDT approach and by the processing of GNSS observations for the three components.

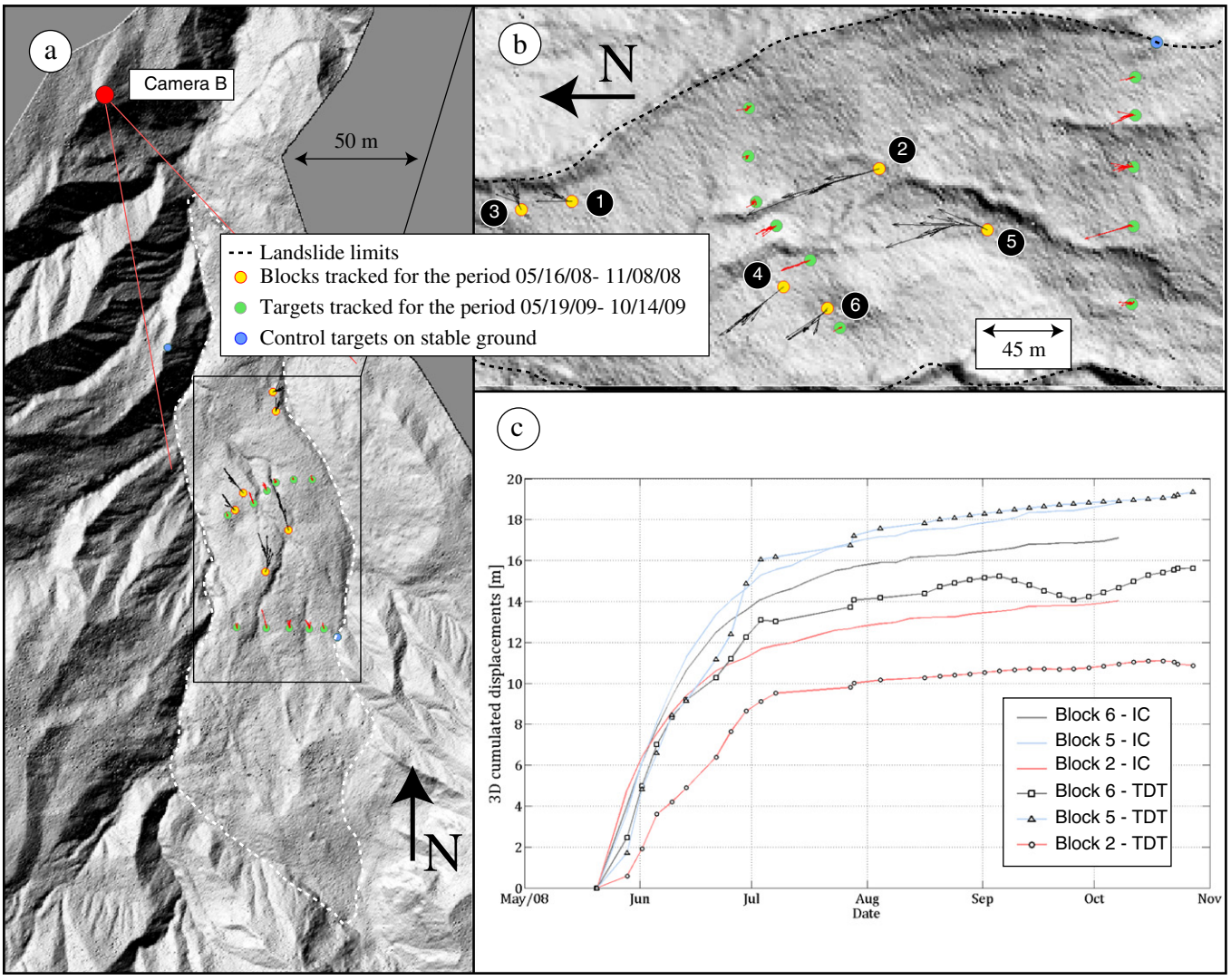

Fig. 7. Displacements obtained from the camera B image time series for the periods *P*1: a) Global map showing the position of the tracked blocks during the first period and the tracked targets during the second period. b) Zoom on the first map showing displacements amplitude and direction for each object. c) Comparison of the cumulated displacement over the two periods.
Panel c is modified from Travelletti et al. (2012).

A second important choice is the size of the search windows that constrain the search area during the automated feature detection and tracking process (Fig. 2). This size is the only calibration parameter of the method; however, it is important to determine. Choosing it requires 1) a priori knowledge of the displacement amplitudes or 2) preliminary tests. Although a priori knowledge may appear to be a major constraint, continuous displacement monitoring is rarely the first study conducted of a landslide, and average displacement rates are relatively easy to obtain.

5.2. Uncertainty assessment of the TDT method

To define the criteria for designing the TDT monitoring system for landslides, the principal sources of errors are assessed. First, the errors in the 2D image plane are assessed, and these are then propagated along the entire processing chain (Fig. 1). The sources are grouped into two main categories: those associated with the automated feature detection and tracking process and those associated with the camera. The uncertainty in the camera orientation is assessed separately.

Table 2
Standard deviation on the 3D coordinates of the nearest and farthest target.

Standard deviation [m]	Camera A1		
	N	E	Z
Nearest target (at 205 m)	0.04	0.10	0.06
Farthest target (at 641 m)	0.07	0.08	0.07

5.2.1. Sources associated with the camera

5.2.1.1. Impact of camera movement. Despite the effort to install the cameras on stable foundations and further co-register the time-lapse sequence, residual movements still persist in the images. The importance of this error is first evaluated by analyzing the residuals on the Harris corner points during the registration process, and points on the slope that are considered to be stable are tracked with the TDT approach and compared to the manual picks.

The registration process involves pair-wise matching of Harris corner points between a reference image and a current image to define an appropriate mapping model. The Harris corner points of the current image are then corrected from this mapping model. Perfect registration of the image time series would lead to a difference between both Harris corner point coordinates of 0. In our case, most of the images are registered with low RMS errors ($\leq$0.15 pixel), but others, mostly from

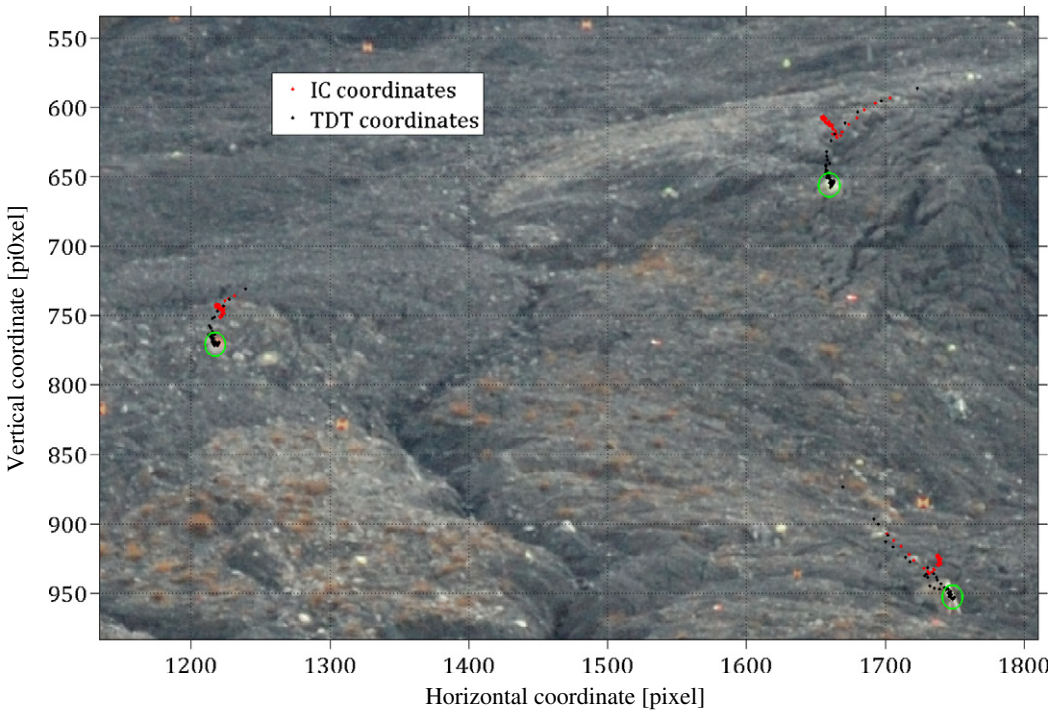

Fig. 8. Difference in the 2D image plane coordinates estimated by the TDT approach and by image correlation for block 2, 4 and 6 during period P1. Both Image Correlation (IC) coordinates (red) and TDT coordinates (black) are represented on an image acquired the 27th of October, 2008. The green circle represents the block positions at this date.

camera A1, show increasing RMS errors with time from the reference image (Fig. 9). This trend can be explained by the difference in views between cameras A1 and A2. Camera A2 looks toward a more distant background landscape (500 m to 10 km) than camera A1 (200 to 500 m), so the changes of vegetation (between May and August) are less perceptible on the images from camera A2. On camera A1, a block detected at the end of May on the reference image may be covered by vegetation in July and thus be detected differently. This assumption is confirmed by the presence of outlier Harris corner points with very large differences of coordinates ($\geq$20 pixels), which are explained by an incorrect matching of the Harris corner points. However, Fig. 9a shows that the error distribution is an approximately 2D Gaussian belt centered on $4.5 \cdot 10^{-14}$. This distribution, which has a mean value very close to zero, indicates that residual minimization by the least-square method is reliable. The standard deviation of 2.05 pixels also indicates a relatively small number of outliers and good reliability of the registration process.

This order of magnitude is confirmed by the comparison of stable points tracked with the TDT method and manual picking on images from the three cameras. For cameras A1 and A2, the standard deviations are greater than 3 pixels (except for the horizontal coordinate of camera A1). The registration process reduces this value by a factor of 3. The final standard deviations can be considered to be low in comparison to the picking uncertainty (Table 3). The coordinates from camera B have lower standard deviations than those from the other cameras. This can be explained by the better stability offered by the concrete pillar on which camera B is fixed and also to the wooden hut that protects the camera from changes in temperature.

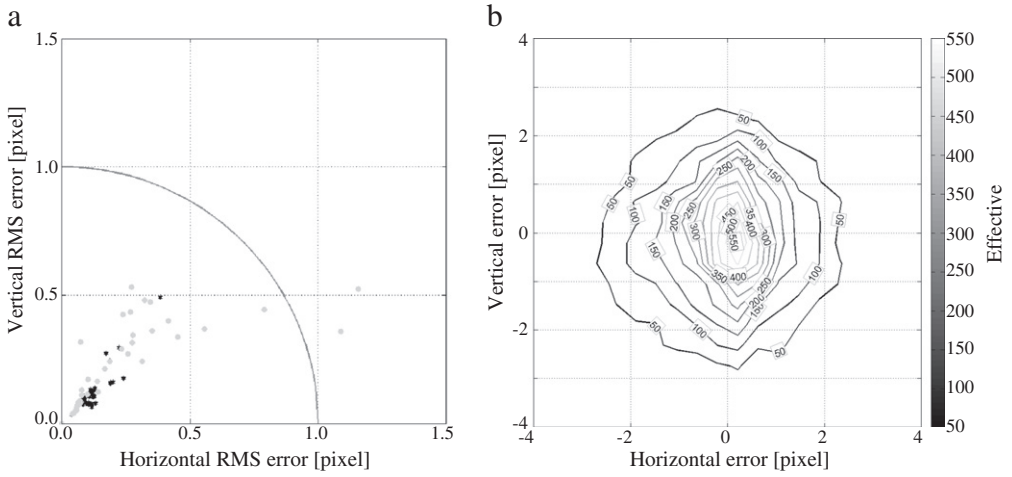

Fig. 9. Impacts of camera movement estimated by the residuals of the computation of the centroid coordinates: a) Vertical and horizontal RMS error for each image of the time series of camera A1 (gray points) and A2 (black stars). The quarter of circle represents the rejection threshold, beyond which the image is not used in the processing. b) Contoured histogram of horizontal and vertical differences between reference and corrected Harris corner point coordinates.

Table 3
Standard deviation of the raw and corrected coordinates of 38 manually picked stable points for cameras A1, A2 and B.

Standard deviation [pixel]	Camera A1		Camera A2		Camera B	
	Horizontal	Vertical	Horizontal	Vertical	Horizontal	Vertical
Raw coordinate	1.7	3.4	3.3	3.8	2.4	2.5
Corrected coordinate	0.4	1.0	0.5	1.2	1.4	1.5

5.2.1.2. Impact of lens distortion. The correction applied to minimize the impact of lens distortion can create important differences in the position estimates (Mikhail et al., 2001). In this study, only radial distortion errors are considered because the tangential distortion errors are comparable to the RMS errors of the model and the source of possible numerical instabilities (Wei and De Ma, 1994).

The accuracy of the radial lens distortion model is assessed by evaluating the maximum difference between two widely used calibration software packages (PhotoModeler© and Adobe® LensProfile Creator). The maximum differences are 0.22 pixel for PhotoModeler© and 0.27 pixel for Adobe® Lens Profile Creator.

However, this accuracy was determined before the installation and may change with the camera setup, the transport and the climatic conditions. Although we were not able to do this, a correction of the lens distortion in each image using GCPs would increase the accuracy.

5.2.2. Sources associated with the automated feature detection and tracking process

5.2.2.1. Impacts of camera–object distance. The camera–object distance affects the estimate of the position in two ways by altering:

- the effective pixel size: The effective pixel size for a camera sensor depends on (i) the camera–object distance and (ii) the angle of incidence, which is defined as the complementary angle between the line of sight of the camera and the normal to the terrain surface (Travelletti et al., 2012). Therefore, the geometry of the camera network must be studied in detail, and the locations of the cameras must be chosen according to the expected displacement rates, the size of the region of interest, and the sizes of the objects being tracked. In addition, stable zones should be integrated in the images to control the stability. For time-lapse monitoring, the effective pixel size may vary significantly with time. For example, for large displacements such as those observed at Super-Sauze, the effective pixel size may be reduced by a factor of four during the period of block tracking because the blocks move away from the camera, and thus the uncertainty is reduced by half. This uncertainty must be taken into account during the 3D restitution process.
- the centroid computation: The computation of the centroid position can reach sub-pixel precision according to the number of pixels that make up the object. Because the centroid is computed as the center of mass of the binary image pixels, pixels of small objects have a larger weight than pixels from large objects. During the creation of the binary image, a missing pixel on the object will therefore create a larger error for a small object than for a large one.

To assess the impact of this error source, synthetic images were created of 18 white Styrofoam spheres overlaid on a background image of the Super-Sauze terrain (Fig. 10a). The diameters of the spheres range from 6 to 28 pixels, and a 5×5 pixel Gaussian kernel smoothing is applied to increase the quality of the image. Fig. 10b shows the L_2 error of the estimate of the centroid coordinates in comparison to the true object diameter. The L_2 error can reach 0.5 pixels for an object with a diameter less than 15 pixels but globally is less than 0.25 pixels for larger objects. This can be explained by the fact that the erosion of large images (Fig. 2b) becomes less and less effective and because several erosion steps should be applied to these images (Russ, 2002).

5.2.2.2. Impact of illumination conditions. The illumination conditions observed in the different images vary significantly according to the time of day, the season, changes in the shadows on the ground, and the weather conditions (sunny or cloudy days). In contrast to the image correlation technique (Berthier et al., 2005; Travelletti et al., 2012), the TDT algorithm is not affected by illumination angles in the tracking step but only in the detection step. The difference of illumination of the object can significantly change the computed binary image so the object is recognized but has a different shape (Fig. 11c and d).

A series of synthetic images were created to assess the impact of illumination on the accuracy of the target positioning. Synthetic images representing white Styrofoam spheres 54 pixels in diameter were

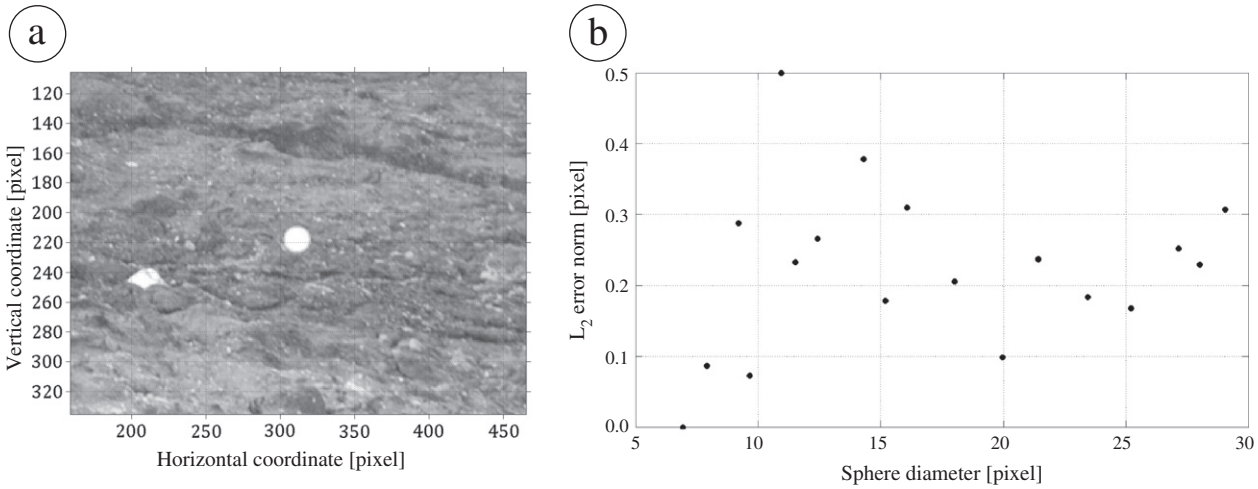

Fig. 10. Impact of the camera–object distance on the object centroid detection: a) Synthetic image of a Styrofoam sphere of 35 pixels in diameter overlaid on a background image of the Super-Sauze terrain, b) Errors on the estimated centroid location in comparison to the true object size.

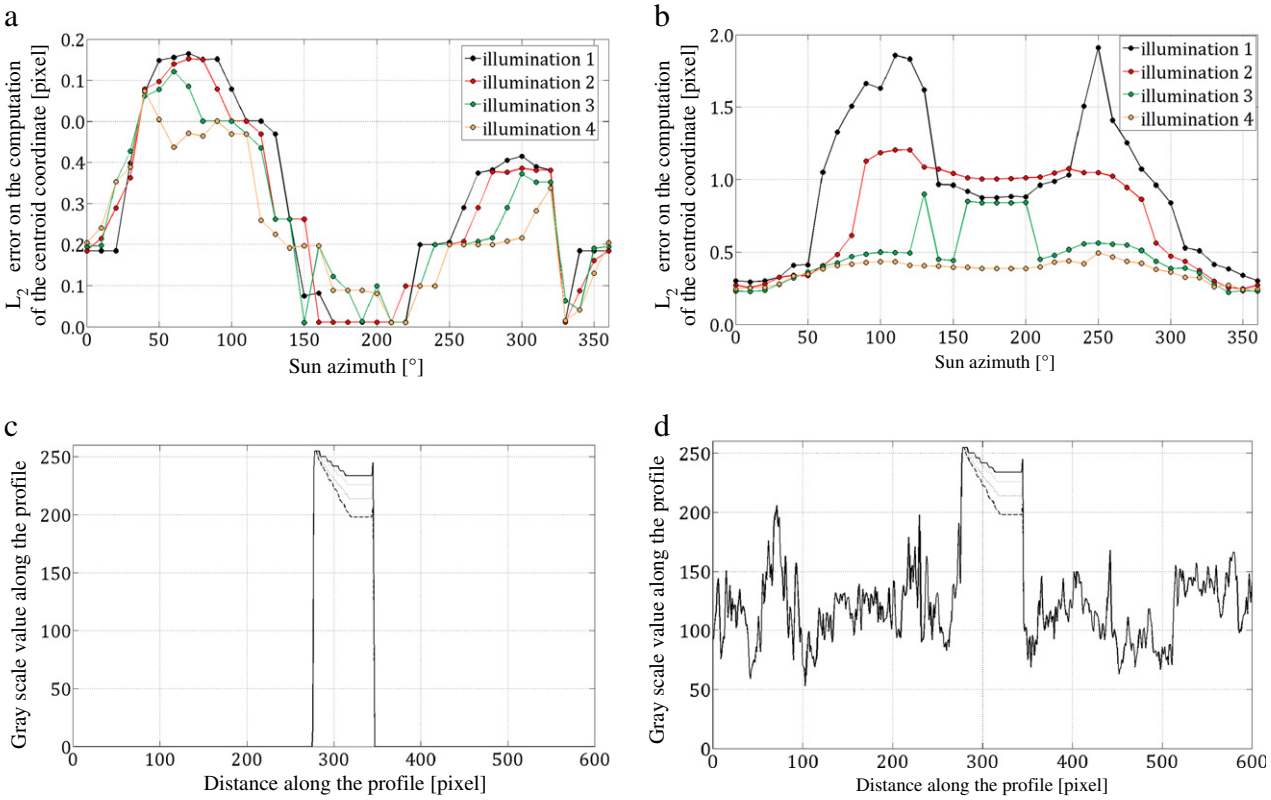

Fig. 11. Influence of illumination condition: Effect of sun azimuth and illumination intensity on a synthetic 54 pixels diameter sphere on a synthetic black and a realistic background. a) Centroid computation error for a sphere of 54 pixels in diameter on a black synthetic background, b) centroid computation error for a sphere of 54 pixels in diameter on a real background, c) gray-scale value along the horizontal profile crossing the sphere center for a black background and for 4 different illuminations and d) gray-scale value along the horizontal profile crossing the sphere center for a realistic background and for 4 different illuminations.

created for different realistic illumination conditions according to the sun azimuth and smoothed using a 5 × 5 pixel Gaussian kernel. The four illuminations are summarized in Table 4. The sun elevation is considered to be 0, and its azimuth varies from 0 to 360. The TDT approach is tested in two experiments:

- overlay of a Styrofoam sphere on a black background (Fig. 11a and c).
- overlay of a Styrofoam sphere on a real background extracted from the image time series (Fig. 11b and d).

The error between the theoretical and computed centroid coordinates varies with the sun azimuth and the illumination conditions (Fig. 11a and b). The L_2 error norm ranges from 0.005 to 0.16 pixels with the synthetic black background and from 0.23 to 1.9 pixels with the real background, corresponding to a maximum relative error of 3.5% with respect to the sphere diameter. The error is due to the gray-scale intensity threshold used to calculate the binary image. As expected, the darker the shadow on the sphere (Fig. 11c), the more pixels belonging to the sphere in the binary image are rejected and, as a consequence, the more the error is important. This phenomenon is amplified with the real background because the sphere has less contrast, and the shadow intensity on the sphere can be very similar to the intensity of the background (Fig. 11d). The maximum error occurs when the illumination is oriented orthogonal (±90) to the camera viewpoint. In the case of a real background, the minimum error is observed for a sun illumination behind the camera, while with the synthetic black background, the minimum error is observed for an illumination in front of the camera.

Finally, the error of the TDT methods is increased by a factor of 10 between the synthetic background and the real background for very bad illumination conditions. However, images characterized by large diffuse reflections are not used in the processing because the color of the Styrofoam sphere is difficult to distinguish from the ground color.

The experiment was also performed for Styrofoam spheres 12 pixels in diameter (corresponding to the average size of the spheres detected from the image time series) and for slope and illumination conditions corresponding to the days and times of acquisition for the image time series from camera A1. The images acquired between 8.00 am and 16.00 UTC, sun elevations between 23.7 and 63.7 and azimuths between 80 and 240 correspond to the worst conditions (Illumination 1, Table 4). Elevations lower than 35 are not computed in the analysis because the topography of the landslide blocks the direct sunlight at lower elevations. The synthetic views of the white spheres are then computed using the orientations of the cameras (azimuth: 145° N and elevation: 30). The sun's azimuth and elevation in Fig. 12 are expressed in the camera's coordinate system.

The errors in this experiment range between 0.06 and 0.33 pixels (Fig. 12). The sun's elevation (relative to the camera view angle) has a significant impact for sun azimuths (relative to the camera view angle) between 0° and 60°. In this range, the minimum error occurs for relative elevation angles between 5°and 20°. The observed error is different from the theoretical error. This might be due to the 5 × 5 Gaussian kernel smoothing that was applied to the small sphere.

Table 4
Combination of illumination criteria used for the experiments.

Relative contribution (%)	Ambient light	Diffuse reflection	Specular reflection
Illumination 1	77.8%	22.2%	0%
Illumination 2	83.3%	16.7%	0%
Illumination 3	88.2%	11.8%	0%
Illumination 4	90.9%	9.1%	0%

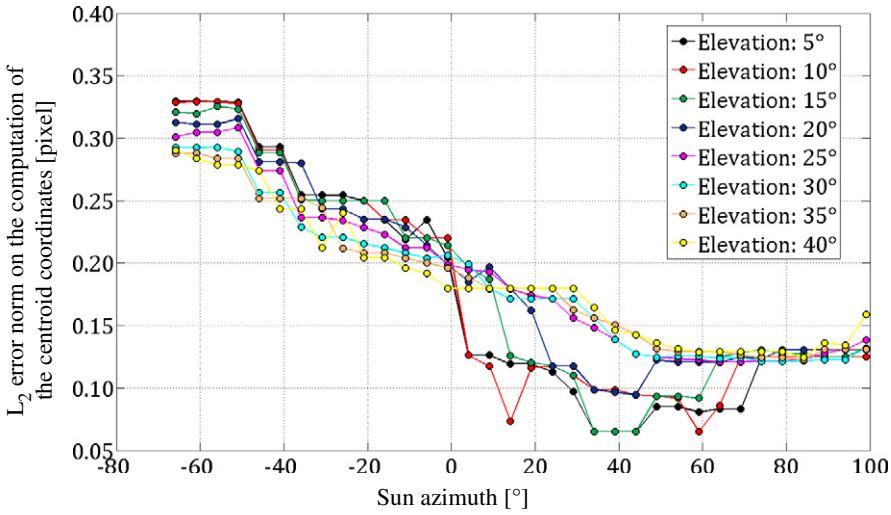

Fig. 12. Error on the centroid coordinate computation with respect to the illumination condition observed at the study site.

5.2.3. Impacts of camera orientation

Another source of uncertainty is the estimate of the camera's external orientation. This error occurs in the final step (step 3; Fig. 1) of the 3D restitution and is directly expressed in metric values for the three components. Therefore, it is assessed separately from the uncertainty from steps 1 and 2 (Fig. 1).

The uncertainty in the camera's orientation largely depends on the quality of the Ground Control Points (GCPs) acquired in the field (number and distribution of GCPs and the intrinsic quality of their coordinate estimates; Fraser, 1997) and on the geometry of the camera network. The error is mainly caused by a large base to height ratio B/h, which complicates the relative orientations of the images. It is generally assessed from the RMS error of the self-calibrating bundle adjustment (Fraser, 1997), which is equal to 4.5 pixels for sphere tracking and 1.6 pixels for block and target tracking. However, this RMS value does not reflect the real error in the 3D coordinates. Therefore, we compute

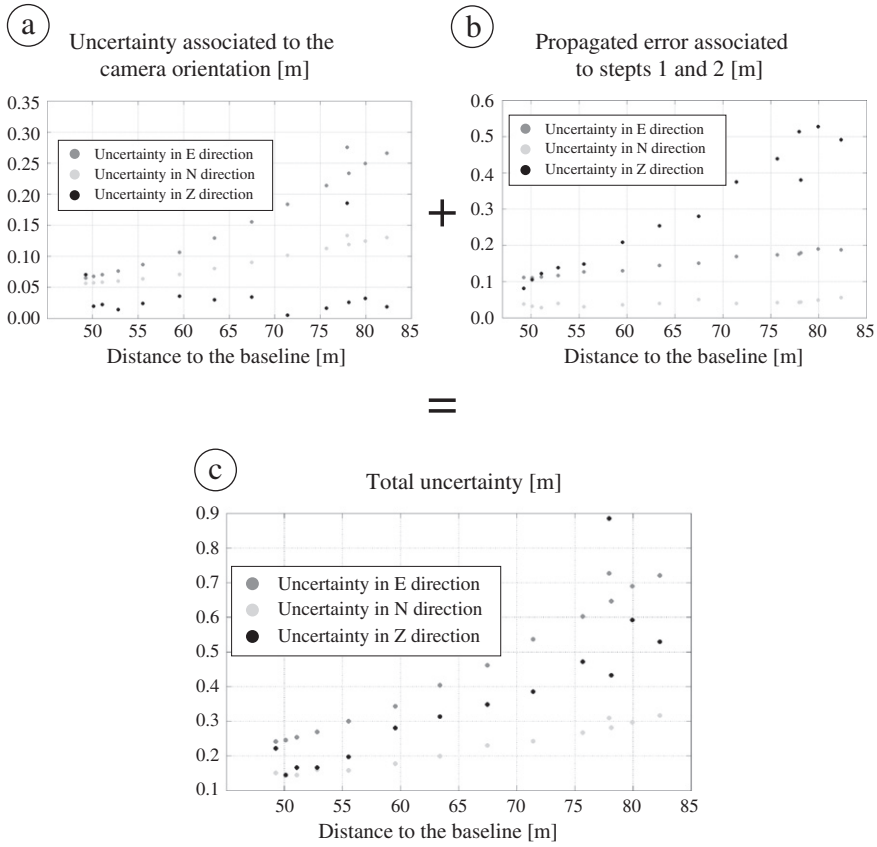

Fig. 13. Estimated uncertainty on the 3D coordinates: a) uncertainty associated to the camera orientation parameters, b) uncertainty associated to steps 1 and 2, after 2D restitution, including lens distortion, camera object distance, camera movement and illumination conditions, c) total maximal uncertainty on the 3D coordinates for the entire methodology, including uncertainty on detection and tracking in the 2D image plane and uncertainty on the camera orientation parameters.

the a posteriori covariance matrix to calculate the standard deviation of the coordinates (Mikhail et al., 2001). The results are presented in Fig. 13a. In this application, the E component is not well constrained compared to the other components, with uncertainties greater than 25 cm.

5.2.4. Evaluation of the propagation of errors and estimate of the uncertainty

In this section, the sources of error are propagated though the entire processing chain to assess the uncertainty of the final 3D coordinates in the most unfavorable case. The maximum errors from the TDT method are summarized in Table 5.

First, the errors determined by the sensitivity analysis (in pixels) from the TDT method are propagated with the stereo-restitution process to obtain an estimate in metric values for the three components. To do this, the errors indicated in Table 5 are summed. Then a 2D image plane error circle with a radius of 3.2 pixels is computed randomly around the object centroid for each image. The 3D coordinates are computed from the extrinsic camera parameters and from the set of Eq. (1) for cameras A1 and A2, which provide 3D error ellipses for each sphere. The errors from the TDT method in the N, E and Z coordinates are computed from the minimum and maximum coordinate in each direction (Fig. 13b).

The previously computed error shows a random behavior that is highly dependent on the image texture and radiometry, lens distortion and camera movement. The final bias introduced in the coordinate is not predictable and not systematic, so we consider this error to be the maximum uncertainty in the methodology (Fig. 13b). Therefore, to obtain the final maximum uncertainty for each object coordinate, including the uncertainties from steps 1 and 2 and the uncertainty from the camera orientation, we simply add both previously computed metric uncertainties. The results are presented in Fig. 13c.

The total uncertainty is reasonable because it increases with the distance from the object to the camera. Globally, the uncertainty ranges between 15 and 90 cm. The N coordinate is well constrained with a total uncertainty ranging from 15 to 32 cm. The uncertainty in the E coordinate is more significant and ranges between 25 and 73 cm. The uncertainty in the vertical coordinate ranges between 15 and 90 cm.

Because the landslide mainly moves to the North (0) in that area, the total 3D displacement can be approximated by the displacements in the N direction. Therefore, the uncertainty in the spheres' displacements is low (15–30 cm), and the larger uncertainty observed in the E direction has little effect on the final result. These results were expected because camera A2, which constrains the displacement in the E direction, is located further upslope than camera A1, which constrains the displacement in the N direction. This study therefore illustrates the importance of a good camera network geometry. Finally, the observed differences between the TDT positions and the GNSS positions are very low (5 cm in the N and Z directions and 20 cm in the E direction), indicating that the error determined with the sensitivity analysis is a maximum value that was computed for the most unfavorable scenario.

6. Setup of a TDT monitoring system

This study indicates that monitoring landslide displacements using TOP is possible using a TDT approach. The computation of the sources of errors shows that it is possible to track objects at a precision comparable to the image correlation techniques. The sensitivity analysis also highlights that the errors related to the camera network configuration and to the choice of the equipment are much greater than those generated by simple feature tracking in the image plane (using TDT or image correlation). We thus propose different criteria to set up a TDT monitoring system based on our experience and on several studies from the literature (Chandler, 1999; Chandler et al., 2005; Wackrow and Chandler, 2010). The different variables used in this section are explained in Table 6. The guidelines are organized according camera positioning, characteristics and configuration.

Table 5
Summary of the errors coming from the TDT method previously computed.

Source of error	Error [pixel]
Camera–object distance	0.5
Camera movement	2
Lens distortion	0.2
Illumination condition	0.5

6.1. Optimal camera positioning

- Camera location: For landslide analysis, appropriate locations for camera are generally limited in practice (unstable slopes around the landslide, complexity of the topography to observe). A first method to select the optimal location is to compare approximately the minimal displacement detectable on the image with the precision expected. Eq. (2) allows determining the approximated minimal displacement that can be tracked on the image (without considering the incidence angle with the terrain).

$$\delta u = \frac{h}{f} \cdot e \cdot p \tag{2}$$

where δu stands for the minimal displacement detectable required, h the maximal distance camera–object, f the focal length, e the sensor pixel size and p the precision of the processing technique used.

- Camera orientation: The camera orientation must be chosen according to four criteria:
  - It should be oriented to maximize the displacement amplitude of the target on the image;
  - It should look at a stable parts, with little vegetation and with a stable geometry and radiometry over time. These stable parts should be, as much as possible, distributed in the image in order to correct for unavoidable image co-registration. Easily detectable GCPs may be added on these stable parts to improve the accuracy;
  - It should be optimized for convergent image configurations in order to minimize the residual systematic errors caused by inaccurately estimated lens distortion parameters (Wackrow and Chandler, 2010);
  - It should be orientated to avoid backlighting.

If the study purpose is object tracking, the camera positions and orientations must be chosen in order that, at the end of the monitoring program, the objects remain in the views of the cameras.

- Number of cameras: The number of camera is determined by the extent of the area of interest and by equipment cost. Several studies show the difficulty of acquiring good quality images, at the same date over long periods in high altitude areas (Major et al., 2009;

Table 6
Meaning of the different variables used in this section.

Variable	Definition	Unit
δu	Minimal displacement detectable	[m]
h	Maximal distance camera–object	[m]
f	Focal length	[m]
e	Sensor pixel size	[m]
p	Precision of the technique	[pixel]
b	Distance between the two cameras	[m]
α	Multiplier ranging between 3 and 5	–
α_{max}	Maximal expected acceleration	$[m \cdot s^{-2}]$
ΔT	Acquisition frequency	[s]

Bernard et al., 2013). For that reason, the selection criteria is to focus on the quality of the equipment, allowing remote access and configuration and data transfer capabilities, and the possibility of efficient power supply (both for data acquisition and data transmission). The B/h ratio is the key parameter for the setup of the array geometry; indeed, a large B/h ratio minimizes the error during the 3D reconstruction but complicates matching process and consequently increases the uncertainty on the extrinsic camera parameters.

6.2. Optimal camera characteristics

- Camera: Most of present-day consumer-grade digital cameras are suitable for digital photogrammetry (Chandler et al., 2005). The camera must be chosen at first according its sensor resolution and size that define the sensor pixel size e in Eq. (2). The sensor pixel size e is calculated by dividing the sensor size by its resolution. Nowadays, it varies between 2 and 7 μm.
- Lens: Concerning lens, a fixed focal length lens, less versatile than variable zoom lens, is preferable. The distortion associated is easier to model and maintain with this type of lens. Avoid extremely wide angle lenses producing large distortions that can reduce the accuracy for the study. Typical used lenses are 35–50 mm (in full frame equivalent.
- Camera fixation: The camera should be fixed on stable grounds. Concrete pillar seems to slightly diminish the camera movement relatively to a metal pole (Fig. 3c). Avoid as far as possible jointed pieces that always end up moving because of the wind or during the maintenance of the camera. The camera should be fixed on its support with quick release pieces, so that it can be detached and repositioned exactly as before.
- Triggering system: Avoid turnkey solutions as the Digisnap 2000 intervalometer (Harbortronics) and its solar controller that are not adapted for long term studies (camera always on standby, and battery charge stopped under 0 °C). Prefer handmade systems if possible or data logger devices, more stable thought more expensive.
- External clock: For long period study, the camera clock slowly drifts and sometimes resets. This problem has also been encountered by (Bernard et al., 2013). All the images acquired for this period are unusable. To remedy this problem, a camera connected to a GPS can be used, provided that the GPS time must be written in an EXIF photo file. Another solution is to use an external GPS and low power computers to save the date a posteriori on the image EXIF.
- Data recovery: The automatic or semi-automatic data recovery is also an important issue when acquiring remotely because it permits a daily check. Teletransmission of a low resolution image is generally sufficient for a simple verification and permits to save communication power. The recovery of the raw data can be done manually and at longer time intervals. However, it might be thought for minimal handling of the camera (e.g. permanent cable or remote transfer).

6.3. Optimal system configuration

- Acquisition frequency: The time between two acquisitions must be defined by the order of magnitude of the maximal acceleration that the user wants to track. Indeed, if the purpose of the study is to track at first shot the landslide acceleration, the data acquisition frequency must be chosen regarding the landslide maximal acceleration ($\alpha_{max} = 3.5\ \mathrm{m \cdot s^{-2}}$) expected using Eq. (3). If the minimal displacement detectable being computed from Eq. (2) is equal to 10 cm, one image may be acquired every 30 s.

Moreover, if it is possible, we advise to multiply the acquisition frequency by the α coefficient, ranging between 3 and 5 to:

— prevent that some images will not be usable because of the bad meteorological conditions

— be able to reduce the high frequency noise due to camera movement in case of slow displacements

$$\Delta T = 2\alpha \cdot \sqrt{\frac{\delta u}{a_{max}}} \tag{3}$$

- Lens calibration: The lens distortion model should be at least calibrated before the installation with the focus used for the study (Fraser, 1997), but a calibration from in place camera images is preferable. The calibration process should be repeated regularly and even automated if possible. We suspect in our dataset that a part of the deformation in the image is due to changes in lens distortion model, certainly with temperature variations. Those variations have also been encountered by Major et al. (2009).
- GCPs: Ground Control Points (GCPs) must be placed on the fields of view of the different cameras in stable parts. Those ground control points must be chosen to be easily recognizable on the images (geometry, color …) and to measure at least 10 pixels on the images. They should be well distributed on the images and the minimum number of control targets necessary is four per stereo pair, but it is prudent to install numerous additional targets. If those GCPs are permanents, they permit to:
  - Calculate the extrinsic camera parameters
  - Verify the quality of image registration
  - Calibrate regularly the lens distortion model.

7. Conclusion

In this paper, we propose a Target Detection and Tracking (TDT) method that is based on simple binary image processing and is designed as a complementary technique to image correlation. The code is developed as an open-source MATLAB-based tool and, like image correlation, allows tracking natural or man-made targets in a time series of images. The difference between the methods is that the TDT method is designed to track radiometrically outstanding targets in the image, while an image correlation technique is often used to track points on a spatially regular grid with random radiometry.

Although it does not permit spatially distributed measurements of displacement, the method allows robust tracking of natural or man-made objects in high displacement areas that are affected by soil reworking and radiometry and texture changes. The object may be affected by rotations and scale changes, and the tracking process remains efficient if the object's radiometric properties remain different from those of the surrounding ground.

The potential of the presented TDT code is assessed using three time-series acquired from the Super-Sauze landslide (South French Alps). The method is validated by comparing the results to those from GNNS measurements and multi-temporal image correlations from Travelletti et al. (2012) for camera-object distances ranging between 100 m and 650 m. TDT allows target tracking with the same order of precision as image correlation techniques (relative precisions ranging between 10^{-4} and 10^{-3}), and we demonstrate its performance in the image plane in areas where image correlation is less efficient.

A sensitivity analysis was used to assess the uncertainties from different error sources (the setup and geometry of the device and the method itself). This study highlights that the minimum object centroid coordinate computation precision is on the order of a pixel. The maximum uncertainty, which is computed as the sum of the camera-object distance and the illumination condition uncertainties, is estimated at 1 pixel (which confirms the method's robustness). The uncertainty due to equipment setup (lens distortion, camera movement) is calculated as 2.2 pixels. This result is then amplified by the uncertainty of the

external orientation parameters due to the geometry of the photogrammetric network and to the precision of the GCPs used for the study. The results of this study highlight that the major source of uncertainty is not inherent to the displacement monitoring method used (Image Correlation or TDT) but is rather due to the acquisition protocol and to the configuration of the photogrammetric network. Based on this finding, we propose guidelines to design an operational system that minimizes those errors.

A perfect comparison between image correlation and the TDT method would certainly demonstrate that the former technique is slightly more precise. However, the proposed TDT method is more robust under real conditions because it is only slightly affected by illumination conditions, easy to parameterize (only one parameter), relatively accurate, does not require data post-processing (contrary to image correlation) and allows near real-time processing. As a result, it is potentially a better option for use in early warning systems than image correlation.

If possible, the complementary use of image correlation and TDT methods on targets could provide information about the reliability of the measured displacements. If both displacements are equivalent, the measurements are reliable, but if the displacements are not consistent, the reliability is lower, and the system may require human intervention.

However, we note the importance of the design for obtaining good efficiency in both the image correlation and TDT methods. Ground-based photogrammetry is generally described as a promising technique because of its light weight, possible total automation and ease of installation. Nevertheless, it is difficult to optimally install cameras in inaccessible areas and is even more difficult to maintain them in good condition in locations where climatic conditions are harsh.

Acknowledgments

This work was supported by the French National Research Agency (ANR) within the Project SISCA (Système Intégré de Surveillance de Crises de glissements de terrains Argileux, 2009–2012), by the BRGM Carnot Institute and by the SNO-INSU OMIV Observatory “Observatoire Multidisciplaire des Instabilités de Versants” at EOST.

References

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-up robust features (SURF). Comp. Vision Image Underst. (CVIU) 110, 346–359.
Bernard, E., Friedt, J., Tolle, F., Griselin, M., Martin, G., Laffly, D., Marlin, C., 2013. Monitoring seasonal snow dynamics using ground based high resolution photography (Austre Lovénbreen, Svalbard, 79N). ISPRS J. Photogramm. Remote Sens. 75, 92–100.
Berthier, E., Vadon, H., Baratoux, D., Arnaud, Y., Vincent, C., Feigl, K., Rémy, F., Legrésy, B., 2005. Surface motion of mountain glaciers derived from satellite optical imagery. Remote Sens. Environ. 95, 14–28.
Brown, D., 1971. Close-range camera calibration. Photogramm. Eng. 37, 855–866.
Cardenala, J., Mataa, E., Perez-Garciaa, J.L., Delgadoa, J., Hernandeza, M.A., Gonzalezb, A., Diaz-de Teranb, J.R., 2008. Close-range digital photogrammetry techniques applied to landslide monitoring. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 37.
Casson, B., Delacourt, C., Allemand, P., et al., 2005. Contribution of multi-temporal remote sensing images to characterize landslide slip surface: application to the La Clapière landslide (France). Nat. Hazards Earth Syst. Sci. 5, 425–437.
Chandler, J., 1999. Effective application of automated digital photogrammetry for geomorphological research. Earth Surf. Process. Landf. 24, 51–63.
Chandler, J.H., Fryer, J.G., Jack, A., 2005. Metric capabilities of low-cost digital cameras for close range surface measurement. Photogramm. Rec. 20, 12–26.
Clerc, P., 2001. Mesure de champs de déplacements et de déformations par stéréovision et corrélation d'images numériques. (Ph.D. thesis) INSA Lyon, France.
Coe, J.A., McKenna, J.P., Godt, J.W., Baum, R.L., 2009. Basal-topographic control of stationary ponds on a continuously moving landslide. Earth Surf. Process. Landf. 34, 264–279.
Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. 24, 603–619.
Corsini, A., Farina, P., Antonello, G., Barbieri, M., Casagli, N., Coren, F., Guerri, L., Ronchetti, F., Sterzai, P., Tarchi, D., 2006. Space-borne and ground-based SAR interferometry as tools for landslide hazard management in civil protection. Int. J. Remote Sens. 27, 2351–2369.
Delacourt, C., Allemand, P., Berthier, E., Raucoules, D., Casson, B., Grandjean, P., Pambrun, C., Varel, E., 2007. Remote-sensing techniques for analysing landslide kinematics: a review. Bull. Soc. Geol. Fr. 178, 89–100.
Elgammal, A., Harwood, D., Davis, L., 2000. Non-parametric model for background subtraction. Computer Vision—ECCV 2000, pp. 751–767.
Flageollet, J.C., Malet, J.P., Maquaire, O., 2000. The 3D structure of the Super-Sauze earthflow: a first stage towards modelling its behaviour. Phys. Chem. Earth B 25, 785–791.
Fraser, C.S., 1997. Digital camera self-calibration. ISPRS J. Photogramm. Remote Sens. 52, 149–159.
Goshtasby, A., 1989. Correction of image deformation from lens distortion using Bezier patches. Comput. Vision Graph. Image Process. 47, 385–394.
Guizar-Sicairos, M., Thurman, S.T., Fienup, J.R., 2008. Efficient subpixel image registration algorithms. Opt. Lett. 33, 156–158.
Harris, C., Stephens, M., 1988. A combined corner and edge detector. Alvey vision conference, p. 50.
Honda, K., Nagai, M., 2002. Real-time volcano activity mapping using ground-based digital imagery. ISPRS J. Photogramm. Remote. Sens. 57, 159–168.
Horaud, R., Monga, O., 1995. Vision par ordinateur: outils fondamentauxHermès.
Jaboyedoff, M., Oppikofer, T., Abellán, A., Derron, M., Loye, A., Metzger, R., Pedrazzini, A., 2010. Use of LIDAR in landslide investigations: a review. Nat. Hazards 61, 5–28.
Lewis, J., 1995. Fast normalized cross-correlation. Vision interface, pp. 120–123.
Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 60, 91–110.
Maas, H.G., Schwalbe, E., Dietrich, R., Bässler, M., Ewert, H., 2008. Determination of spatio-temporal velocity fields on glaciers in West-Greenland by terrestrial image sequence analysis. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 37, 1419–1424.
Major, J., Dzurisin, D., Schilling, S., Poland, M., 2009. Monitoring lava-dome growth during the 2004–2008 Mount St. Helens, Washington, eruption using oblique terrestrial photography. Earth Planet. Sci. Lett. 286, 243–254.
Malet, J.P., Laigle, D., Remaître, A., Maquaire, O., 2005. Triggering conditions and mobility of debris flows associated to complex earthflows. Geomorphology 66, 215–235.
Malet, J.P., Déprez, A., Ulrich, P., Mathieu, A., Jongmans, D., Darras, L., Lebourg, T., Vidal, M., Maquaire, O., Davidson, R., 2013. Continuous monitoring and near-real time processing of GPS observations for landslide analysis: a methodological framework. Landslide Science and Practice 201–209.
Meer, P., 2003. Kernel-based object tracking. IEEE Trans. Pattern Anal. Mach. Intell. 25.
Mikhail, E., Bethel, J., McGlone, J., 2001. Introduction to Modern Photogrammetry, vol. 31. Wiley New York.
Moravec, H., 1979. Visual mapping by a robot rover. vol. 1, pp. 598–600.
Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica 11, 23–27.
Russ, J., 2002. The Image Processing HandbookCRC Press.
Sturzenegger, M., Stead, D., 2009. Close-range terrestrial digital photogrammetry and terrestrial laser scanning for discontinuity characterization on rock cuts. Eng. Geol. 106, 163–182.
Travelletti, J., Malet, J.P., 2012. Characterization of the 3D geometry of flow-like landslides: a methodology based on the integration of heterogeneous multi-source data. Eng. Geol. 128, 30–48.
Travelletti, J., Delacourt, C., Allemand, P., Malet, J., Schmittbuhl, J., Toussaint, R., Bastard, M., 2012. Correlation of multi-temporal ground-based optical images for landslide monitoring: application, potential and limitations. ISPRS J. Photogramm. Remote Sens. 70, 39–55.
Tuytelaars, T., Mikolajczyk, K., 2008. Local invariant feature detectors: a survey. Foundations and Trends® in Computer Graphics and Vision. 3, pp. 177–280.
Veeraraghavan, H., Schrater, P., Papanikolopoulos, N., 2006. Robust target detection and tracking through integration of motion, color, and geometry. Comput. Vis. Image Underst. 103, 121–138.
Wackrow, R., Chandler, J., 2010. Tips for effective use of digital close range photogrammetry and terrestrial laser scanning. ISPRS. website http://isprsv6.lboro.ac.uk/tips.html.
Walter, T.R., 2011. Low cost volcano deformation monitoring: optical strain measurement and application to Mount St. Helens data. Geophys. J. Int. 186, 699–705.
Wei, G., De Ma, S., 1994. Implicit and explicit camera calibration: theory and experiments. IEEE Trans. Pattern Anal. Mach. Intell. 16, 469–480.
White, D.J., Take, W.A., Bolton, M.D., 2003. Soil deformation measurement using particle image velocimetry (PIV) and photogrammetry. Geotechnique 53, 619–632.
Wren, C., Azarbayejani, A., Darrell, T., Pentland, A., 1997. Pfinder: real-time tracking of the human body. IEEE Trans. Pattern Anal. Mach. Intell. 19, 780–785.
Yilmaz, A., Javed, O., Shah, M., 2006. Object tracking: a survey. ACM Comput. Surv. 38, 13.
Zitova, B., Flusser, J., 2003. Image registration methods: a survey. Image Vis. Comput. 21, 977–1000.