Research paper

An algorithm for automatic detection and orientation estimation of planar structures in LiDAR-scanned outcrops

Robson K. Gomes [a], Luiz P.L. de Oliveira [a,*], Luiz Gonzaga Jr[a,c], Francisco M.W. Tognoli [b,c], Mauricio R. Veronez [b,c], Marcelo K. de Souza [b,c]

[a] Programa Interdisciplinar de Pós-Graduação em Computação Aplicada, Brazil
[b] Programa de Pós-Graduação em Geologia, Brazil
[c] VizLab - Advanced Visualization Laboratory, Universidade do Vale do Rio dos Sinos - UNISINOS, Av. Unisinos, 950 - São Leopoldo - RS - Brazil

ARTICLE INFO

ABSTRACT

The spatial orientation of linear and planar structures in geological fieldwork is still obtained using simple hand-held instruments such as a compass and clinometer. Despite their ease of use, the amount of data obtained in this way is normally smaller than would be considered as representative of the area available for sampling. LiDAR-based remote sensors are capable of sampling large areas and providing huge sets of digitized spatial points. However, the visual identification of planes in sets of points on geological outcrops is a difficult and time-consuming task. An automatic method for detecting and estimating the orientation of planar structures has been developed to reduce analysis and processing times, and to fit the best plane for each surface represented by a set of points and thus to increase the sampled area. The algorithm detects clusters of points that are part of the same plane based on the principal component analysis (PCA) technique. When applied to real cases, it has shown high precision in both the detection and orientation of fractures planes.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The measurement of orientation of linear and planar structures is one of the most common and time-consuming activities of field geologists. Identifying and determining the orientation of lineations and paleocurrents, as well as fractures and others planar discontinuities, provide a dataset that may be used in many applications (e.g., structural and geomechanical analysis, paleocurrents, flow modeling, landslide and rock fall prevention). This information is still collected using standard fieldwork instruments such as a hand-held compass and clinometer, although a number of digital electronic devices are available for collecting, correcting and recording field data (e.g., compass apps, rangefinders and LiDAR-based scanners). A range of remote sensors for rock characterization, digital imaging and surface mapping have been developed since the 1990s. Their main advantage is that new topographic and photogrammetric equipment and techniques enable automatic, rapid acquisition of accurate spatial data. For example, LiDAR-based scanners are among the most widely used instruments for acquiring high-resolution data for geological purposes (Bellian et al., 2005; Abellán et al., 2006; Buckley et al., 2008; Ferrari et al., 2012; Hodgetts, 2013). Unmanned airborne vehicles (Bemis et al., 2014) and ground-based photogrammetry (James and Robson, 2012; Westoby et al., 2012) are also becoming important tools for the same purpose.

Data acquisition based on rigorous positioning criteria using global navigation satellite systems (GNSS) allows accurate data to be processed and used for different purposes, such as structural analysis and orientation estimation of planar and quasi-planar surfaces (Fernández, 2005; Slob et al., 2005; Poppinga et al., 2008; Olariu et al., 2008; Pathak et al., 2009; Sturzenegger and Stead , 2009; GarcaSellés et al., 2011; Pearce et al., 2011; Fisher et al., 2014). However, it is fundamental that the collection of more data involves more data management if the database is to be organized so that the information can be used for analysis and interpretation. This has been the main challenge associated with the advanced capability of technological devices to collect large-to-huge datasets: the acquisition of large, accurate datasets increases the representativeness of the rock exposure, but it also increases processing demands and directly affects the amount of time taken to perform analysis and interpretation. This requires the development of algorithms to automate every type of analysis—such as detecting and estimating the orientation of planar surfaces, which is the focus of this work—or any other feature of interest. Several

* Corresponding author.
    E-mail addresses: robson.kg@gmail.com (R.K. Gomes),
lpluna@unisinos.br (L.P.L. de Oliveira), lgonzagajr@gmail.com (L. Gonzaga Jr),
FTOGNOLI@unisinos.br (F.M.W. Tognoli), VERONEZ@unisinos.br (M.R. Veronez),
marcelo.k.souza@gmail.com (M.K. de Souza).

semiautomatic and automatic methods have been reported in the geoscientific literature in recent years, including a number of (more or less) supervised techniques for variously complex workflow and data (pre)processing, triangulated irregular networks (TINs) or raw point clouds and different segmentation and clustering techniques based on planar regression, or fuzzy or PCA algorithms (Lato et al., 2009, 2010; GarcaSellés et al., 2011; Gigli and Casagli, 2011; Lato and Voge, 2012; Maerz et al., 2013; Voge et al., 2013; Zeeb et al., 2013; Riquelme et al., 2014). The most recently developed algorithms fit 3D data to the best plane following a coplanarity test on the point set, for both truly planar and near-planar surfaces.

In this paper we present a PCA-based algorithm that automatically (a) detects planar and near-planar surfaces on outcrops and (b) estimates their orientation, from 3D point clouds obtained from LiDAR scanning. Once the algorithm parameters are given by the user, the planes are detected and their orientation are estimated with no user interferences. The main novelty in this algorithm is the recursive subdivision process of the input set into subsets for detecting the planar regions. Once the appropriated point cloud is given, the algorithm is simple and straightforward in the sense that it does not require any preprocessing related to plane detection and orientation and is based on only one basic procedure, unlike the algorithms found in the reports listed above. In this paper, we consider that point clouds are constituted only by the relevant points, that is, with no outlier points, or ones from vegetation, from rock waste, or from any other irrelevant source.

2. Study area and dataset acquisition

The study site is on the main wall of the Incopel Quarry, located along BR-116 Highway in Estância Velha municipality (UTME= 485685 m; UTMN=6722017 m; Altitude (h); Zone 22S, SIRGAS 2000, WGS-84), (Fig. 1).

An Optech ILRIS 3D laser scanner integrated with a GPS receptor Topcon Hiper Lite RTK was used to perform a high-resolution LiDAR scan of the wall of interest from two locations approximately 30 m from the rock face and giving 30% overlap. These were the best available sites for minimizing occlusion and positional bias related to the fracture orientations. The point cloud obtained had an average of 36 points/cm^2 and about 6.1 million points. The data was filtered to remove points not part of the rock mass (vegetation, waste rock piles, etc.). Two high-accuracy points were used for georeferencing each point cloud in the IMALign module of the PolyWorks software after aligning and merging. To test the algorithm, the large fracture plane visible at the center of the yellow square in Fig. 1 was identified and separated from the whole point cloud.

During the LiDAR scan, one of the planes observed in the rock face was labeled and measured with a Brunton Pocket Transit compass. Three people systematically measured both strike line and dip angle of the plane; each person measured the plane in turn. After 10 rounds, the dataset of 30 strikes and 30 dips was statistically processed to determine the mean and standard deviation of the strike orientation and dip angle.

3. Algorithm for automatic plane detection and specifications

Let $X = \{P_i(x_i, y_i, z_i) \in \mathbb{R}^3; \ i = 1, 2, \ldots N\}$ be a point cloud obtained from an outcrop scan, with average $\overline{P}(\overline{x}, \overline{y}, \overline{z})$. Compute the principal components (PCs) associated with X, centered with respect to $\overline{P}$, arranged in a $N \times 3$ matrix $[X]$. Therefore, the three columns of $[X]$ comprise samples of variables x, y and z, respectively. The PCs thus found are defined by the three (taken as)

unitary eigenvectors $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$ of covariance matrix $\mathbf{S}$ of X, respectively associated with the eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$.

The set $B = \{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3\}$ is an orthonormal basis for $\mathbb{R}^3$. The important property of B is that its vectors point to mutual null-covariance directions with respect to data X. More precisely, it diagonalizes $\mathbf{S}$. Moreover, the associated eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ measure the variances of X with respect to its average $\overline{P}$ in the direction of the corresponding PCs. Therefore, together with $\overline{P}$, the direction associated with λ_1 gives the straight line that best fits the points of X, while the eigendirections associated with the two largest eigenvalues λ_1 and λ_2 define the plane that best fits the data. In both cases, the fitting is in the orthogonal sense, that is, minimizing the sum of squares of the distances from the points to that line or plane. The direction associated with the smallest eigenvalue gives the deviations from that best-fitting plane. Thus λ_3 gives a measure for dispersion of X with respect to that plane, and is a measure of how well the plane defined by the other two PCs fits the data. The plane is given by the dot product

$$(\mathbf{x} - \overline{P}) \cdot \mathbf{t}_3 = 0, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^3$ is a generic point in the plane.

The proposed algorithm explores the properties of the PC to successively divide the input set X into subsets that can be approximated by planes to the level of precision that must be specified by the user. In the process, set X is successively subdivided into $d = 0, 1, 2, \ldots N_d$. Two reasonable examples of user-specified precision follow. Suppose that $A \subset X$ is a subset resulting from a subdivision d. Specification S_1 uses a threshold Q for the average distance $\overline{D}_A$ from the points in the approximated subset $A \subset X$ to the respective fitting plane Γ_A; that is, S_1 is given by

$$S_1: \quad \overline{D}_A \leq Q, \quad \text{with } \overline{D}_A = \frac{1}{N_A} \sum_{k=1}^{N_A} |(P_k - \overline{P}) \cdot \mathbf{t}_3|, \tag{2}$$

where N_A is number of elements of A; $\overline{P}$ its mean point; and $\mathbf{t}_3$ is the third PC for that approximation.

The second specification S_2 is based on parameters m and k as defined in Fernández (2005), which give a measure for classifying a set of points as being geologically coplanar Both are functions of the eigenvalues λ_1, λ_2 and λ_3 obtained from the PCA. These parameters are defined as

$$m = \ln\left(\frac{\lambda_1}{\lambda_3}\right) \tag{3}$$

and

$$k = \frac{\ln\left(\frac{\lambda_1}{\lambda_2}\right)}{\ln\left(\frac{\lambda_2}{\lambda_3}\right)}. \tag{4}$$

Higher values of m correspond to a smaller average distance between the best-fit plane and the data compared to the dispersion of the points normal to the plane. However, a good fit for a plane Γ_A for a region $A \subset X$ does not necessarily qualify that region as a plane since, instead, it could be a straight line. The parameter k is a measure of the classification quality in this sense.[1] Based on these parameters we have the second specification:

$$S_2: \quad m \geq m_{min}, \quad \text{and} \quad k \leq k_{max}, \tag{5}$$

with the values of m_{min} and k_{max} chosen by the user. In our experience, few or no useful results are obtained out of the ranges

---

[1] Recently, other parameters for determination of planar features from geospatial data have been proposed, which could also be used for a specification definition (Jones et al., 2015).

Fig. 1. Global view of the basalt quarry where the dataset was acquired using a laser scanner (visible at the right-hand side of the photograph). The whole outcrop is about 10 m high and 40 m long. The inset shows the three main fracture families and in situ manual orientation using a compass and clinometer. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)
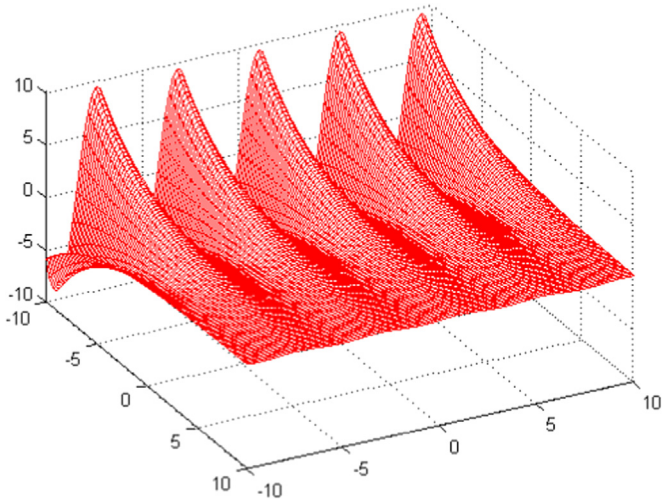

Fig. 2. Plot of function in Eq. (15). A uniform sampling of the graphic was used as a point cloud for testing the algorithm.

$m_{min} \geq 4$ and $0.4 \leq k_{max} \leq 1$. If the chosen specification S (which may be S_1, S_2, or any other specification adopted) is satisfied for X, then the planar approximation Γ^0 is considered acceptable and the process stops. In this case, X is already considered to be a plane with dip and strike specifications given by a third PC ($\mathbf{t}_3$) of that approximation. If X does not satisfy S, however, then the process goes to step $d=1$ where X is written as a union of four subsets (quadrants),

$$X = \bigcup_{i=1}^{4} X_i^1,$$
(6)

defined using the first and second PCs, $\mathbf{t}_1$ and $\mathbf{t}_2$, respectively, as

$$X_1^1 = \{P_k \in X; \quad (P_k - \overline{P})\cdot\mathbf{t}_1 \geq 0 \ \& \ (P_k - \overline{P})\cdot\mathbf{t}_2 \geq 0\}$$
(7)

$$X_2^1 = \{P_k \in X; \quad (P_k - \overline{P})\cdot\mathbf{t}_1 \geq 0 \ \& \ (P_k - \overline{P})\cdot\mathbf{t}_2 < 0\}$$
(8)

$$X_3^1 = \{P_k \in X; \quad (P_k - \overline{P})\cdot\mathbf{t}_1 < 0 \ \& \ (P_k - \overline{P})\cdot\mathbf{t}_2 < 0\}$$
(9)

$$X_4^1 = \{P_k \in X; \quad (P_k - \overline{P})\cdot\mathbf{t}_1 < 0 \ \& \ (P_k - \overline{P})\cdot\mathbf{t}_2 \geq 0\}.$$
(10)

Each subset X_i^1 is approximated by the plane Γ_i^1 given by Eq. (1) obtained from its points. Subsets X_i^1 that do not satisfy the precision specification S go to the next step of approximation $d=2$, leading to its partition as $X_i^1 = \bigcup_{i=1}^{4} X_i^2$ with each X_i^2 defined as above—that is, making $X = X_i^1$ in Eqs. (7)–(10), with $\overline{P}$ replaced by $\overline{P}_i^1$, the mean point of X_i^1. In this case, its approximation Γ_i^1 is ignored. At $d=2$ we then have $X = \bigcup_{i=1}^{N_d} X_i$ with $1 \leq N_d \leq 16$, where X_i are the well-succeeded approximations at step $d=1$ and (all) the approximations of step $d=2$. Generalizing, at step d we have $X = \bigcup_{i=1}^{N_d} X_i$, where X_i are the well-succeeded approximations at steps $1, 2, \ldots d-1$ and (all) the approximations of step d. The process stops when, for D steps, $X = \bigcup_{i=1}^{N_D} X_i$ with $1 \leq N_D \leq 4^D$ where X_i are all well-succeeded approximations for steps $1, 2, \ldots D$. At this stage, the original set X is approximated within criteria S by a piecewise linear model given by

$$X \approx \bigcup_{i=1}^{N_D} \Gamma_i,$$
(11)

where Γ_i is the plane approximation of X_i.

The process of subdivision continues until an acceptable approximation is attained (e.g., as for specifications S_1 or S_2), or until a minimum number of points N_{min} remains, which may also be specified by the user. The number should not be so low that it compromises the statistics of the method, and should also not be so high that it prevents the detection of significant planar regions. Again, in our experience, $N_{min} \sim 50$ should be a good choice for outcrops.

It is interesting to note that the orientation obtained by the proposed method is, in some degree, immune to non biased additive noise. More precisely, let $A \subset X$ be a subset resulted from a subdivision obtained at step d, and $\lambda_1 > \lambda_2 > \lambda_3 > 0$ the three resultant eigenvalues. Let $A' = A + Z$ be a noisy version of A, where the noise Z is a random variable uncorrelated to A with zero mean and standard deviation σ. In this case, the eigenvalues of A' are $\lambda_i' = \sqrt{\lambda_i^2 + \sigma^2}$, $i = 1, 2, 3$. Then, orientations for A and A' given by the method will be the same since $\lambda_1 \geq \lambda_2 \geq_3 > 0$ implies $\lambda_1' \geq \lambda_2' \geq \lambda_3' > 0$, which guarantees that the noise will not cause any change to the directions or to the order of PCs. As a consequence, the method can be applied always when there is no anisotropic
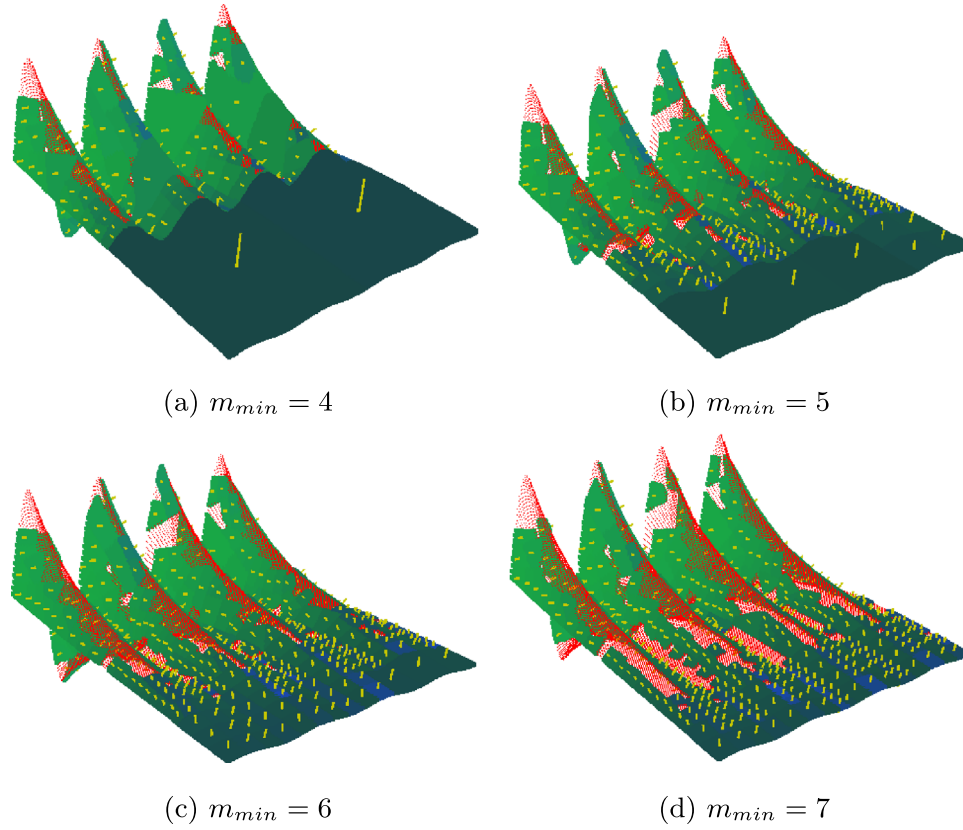
Fig. 3. Portion of plot of Eq. (15) for $k_{max} = 0.8$ and values of m_{min} increasing from (a) to (d). The number of detected planes tends to increase, while their size diminishes. This is seen in the more planar regions, with two planes in (a) $m_{min} = 4$, seven planes in (b) $m_{min} = 5$, and a large number of planes in (c) $m_{min} = 6$ and (d) $m_{min} = 7$. The number of regions that are not well approximated (shown in red) increases from (a) to (d). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

noise that can modify the order of the PCs.

In general, due to the precision of the above procedure and the usual complexity of geological outcrops, it is not easy to distinguish the several planes Γ_i that constitute the obtained approximation in Eq. (11). For that reason, a color code is provided to help the user visually recognize the different planes. Separate RGB colors are used for strike and dip; the unused RGB color then represents points not included in any of the classified planar regions. For example, let us use green for dip, blue for strike, and red for the unclassified points: then, if g_{min} and g_{max} are the minimum and maximum intensity of the green color, and if θ represents dip, we define

$$g(\theta) = \left[1 + \exp\left(\frac{-\theta + \bar{\theta}}{\beta} \right) \right]^{-1}, \qquad (12)$$

where $\bar{\theta}$ is the mean of the lower (θ_{min}) and the upper (θ_{max}) angles that the user wishes to visualize, given by

$$\bar{\theta} = \frac{\theta_{min} + \theta_{max}}{2}, \qquad (13)$$

and β is a measure of the difference within the interval $[\theta_{min}, \theta_{max}]$, giving the inclination of f at $\theta = 0$: $f'(0) = 1/4\beta$. In practice, the differentiable colors (green) tones g range from $g_{min} = g(\theta_{min})$ to $g_{max} = g(\theta_{max})$, in which $g(\theta)$ is quasi-affine, so that we can write the approximation for β as

$$\beta \approx \frac{1}{4} \frac{\theta_{max} - \theta_{min}}{g_{max} - g_{min}}. \qquad (14)$$

Note that the color range entered is already normalized by the quantization level C adopted (say, $C=256$ for 8-bits

representation). So, $0 < g_{min} \leq g(\theta) \leq g_{max} < 1$.

An analogous scheme is adopted for strike, represented by the intensity of the blue color. On the outcrop point cloud, the final result shows the user an image of the set of planes that have been distinguished, each in a combination of green (dip) and blue (strike). Also, a yellow line segment is drawn normal to the plane from the average point to further assist visualization.

4. Results

The algorithm might use the specifications S_1, S_2 as above, or any other specification that the user regards as more appropriate in a given situation. For the point cloud of a highly irregular rock mass, the continuous process of subdivision will result in a number of planes being identified; large numbers of such planes make the visual analysis more problematic. In this section we demonstrate the use of the algorithm both for a synthetic point cloud and for a real point cloud obtained from a LiDAR scan as described in Section 2.

4.1. Synthetic data experiment

The plot of a function $z = z(x, y)$ is taken to be a synthetic landscape where the elevation z is given as a function of the latitude and longitude of positions x and y. The adopted function is

$$z = \rho^x \sin(2\pi f\, y), \qquad (15)$$

which defines an oscillatory pattern of spatial frequency f Hz in the y-direction and of amplitude ρ^x varying exponentially in the x-direction. Fig. 2 shows this surface grading from steep to almost

(a) Point cloud

(b) $m_{min} = 4$
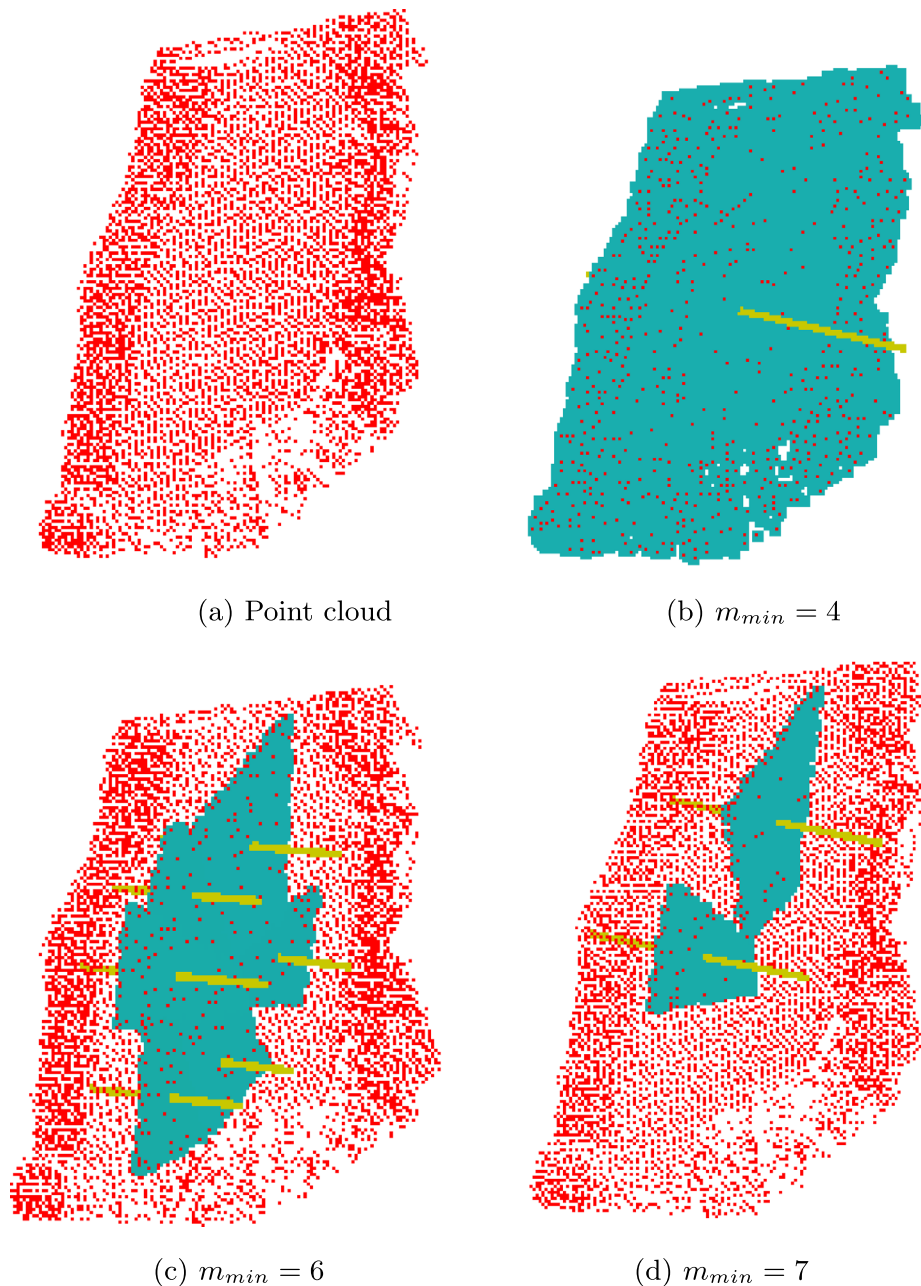
(c) $m_{min} = 6$

(d) $m_{min} = 7$

Fig. 4. Automatic plane detection and measurement by the proposed algorithm at different precisions: (a) point cloud of region C; (b) one plane detected when $m \geq 4$; (c) six planes detected when $m \geq 6$; (d) two planes detected when $m \geq 7$.

horizontal regions, with greater or lesser variation of orientation. The largest oscillations of orientation are seen in the highest-amplitude portions (for $x \approx -10$), gently grading to a more planar surface in the lowest region (for $x \approx 10$).

A uniform sample of surface given by Eq. (15) was used as a synthetic point cloud of a landscape to illustrate plane detection by the algorithm with the parameter $m = 4, 5, 6$ and 7 (Fig. 3) using specification S_2 (Eq. (5)). Only the results for the one value of $k = 0.8$ are shown, since other values of k were found to give similar results. Increasing m makes the method more rigorous when considering a set of points as belonging to the same plane. Therefore, increasing m tends to increase the number of planes detected but reduce their size. Based on this synthetic surface, it was possible to observe how well were the regions approximated by the planar fittings as calculated by the algorithm when using different specifications. The surface also indicates that points are

not aligned sufficiently in the curved regions to fit a plane, and no planes were detected; more subdivisions would be required.

Fig. 3 shows the results for four approximation thresholds: $m \geq m_{min}$ for $m_{min} = 4, 5, 6, 7$ and $k \leq k_{max}$ with $k_{max} = 0.8$. The green areas in the figure represent those points successfully approximated. In each planar approximation, the small yellow straight line segment indicates the direction of the normal to those planes drawn from the mean of the approximated points, as discussed above. In this case, the number of planes increases with higher values of m_{min}, and their size decreases. The red regions indicate points not included in any of the approximation planes; the number of these regions is also seen to increase with higher values of m_{min}.

The synthetic data example has illustrated the extent to which the chosen degree of fit (m_{min}) of the planes affects the results from the proposed algorithm. We observe that, in areas with a
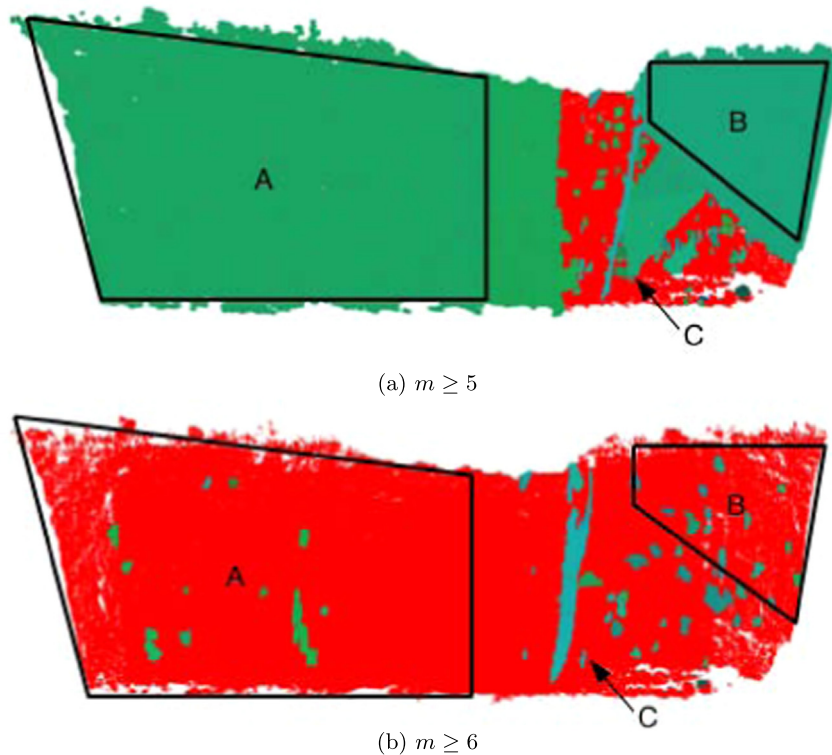
(a) $m \geq 5$

(b) $m \geq 6$

Fig. 5. Automatic plane detection by the proposed algorithm from the entire point cloud: (a) planes detected when $m \geq 5$. A and B are large planes associated with the rock face, and C is part of the large plane measured by compass; (b) planes detected when $m \geq 6$. Planes A, B found in (a) are split into smaller planes, and plane C is still detected.

curved surface with higher chosen degree of fit, the points were not aligned sufficiently and no plane was detected within the chosen minimum of 50 points. Therefore, to find smaller planes more divisions would be necessary which would require the choice of a smaller minimum number of points or, which would be better, a more precise point cloud (obtained with a smaller sampling period) should be considered.

Also notice that some inconsistencies can be identified in the examples of Fig. 3. For example, at the first peak from left to right, there is a region that is not classified as planar while the corresponding regions at the other geometrically identical peaks are classified as planar. This also occurs at other places such as in the third peak. The reason for these inconsistencies is the statistical nature of the method. More specifically, since the method is not geometrically based, it does not take into consideration the geometrical symmetries of the point cloud. Then, the successive divisions does not divides X symmetrically in the geometric sense, but in the statistical sense. So, in general, two geometrically similar regions are not subdivided in the same way, which can result in different classifications since some subdivisions reach the minimum number of points earlier than others. The appropriated way to minimize this is to use a better quality cloud, obtained with a smaller sampling period.

4.2. Real data experiment

The outcrop considered for the next experiments, and the methods applied for data acquisition and in situ measurements of the strike line (SL) and dip angle (DA) of the outcrop were described in Section 2. We tested our method of orientation estimation on the set of points of a plane, indicated as region C in Figs. 5 and 6, measured by traditional in situ methods. Subsequently, we investigated the whole outcrop for global plane detection using the proposed method. In this case, the method classified a great number of planar regions with a variety of

dimensions, including plane C.

4.2.1. Measuring with compass and clinometer

Plane C is shown separately in Fig. 4(d) and (a). The sets of 30 in situ compass measurements and 30 in situ clinometer measurements (listed in Table 1) both show a small standard deviation in SL and DA. These values were taken to represent the true orientation of the plane when testing the calculations performed by the proposed algorithm from the point cloud dataset.

4.2.2. Testing the algorithm with a single plane

The plane whose strike and dip were measured by a hand-held compass and clinometer was identified and segmented, and the resulting point cloud was used as input data for the algorithm. After four runs, each using a different value of m_{min}, one or more planes were automatically detected. The results are shown in Fig. 4.

Fig. 4 shows that the measurements of SL and DA of a set of points assumed to be coplanar depends on the users definition of a plane for the particular case being considered. Table 2 shows different results for SL and DA for different values of m_{min} chosen by the user. This does not imply inconsistency of the proposed method, but rather it illustrates the flexibility of the method from the users point of view.

The differences between the orientations of the selected plane obtained by compass and clinometer and those obtained by the algorithm using the point cloud are defined as $\Delta SL = SL_{compass} - SL_{algorithm}$ and $\Delta DA = DA_{compass} - DA_{algorithm}$). The minimum and maximum variations of ΔSL and ΔDA were found for $m \geq 4$ (Table 3) and the maximum variation of ΔSL was $\Delta SL_{max} = 4.22°$ (for $m \geq 6$) and the minimum variation of ΔDA was less than $1°$ (for $m \geq 5$, $m \geq 6$ and $m \geq 7$), which is an acceptable degree of variability.
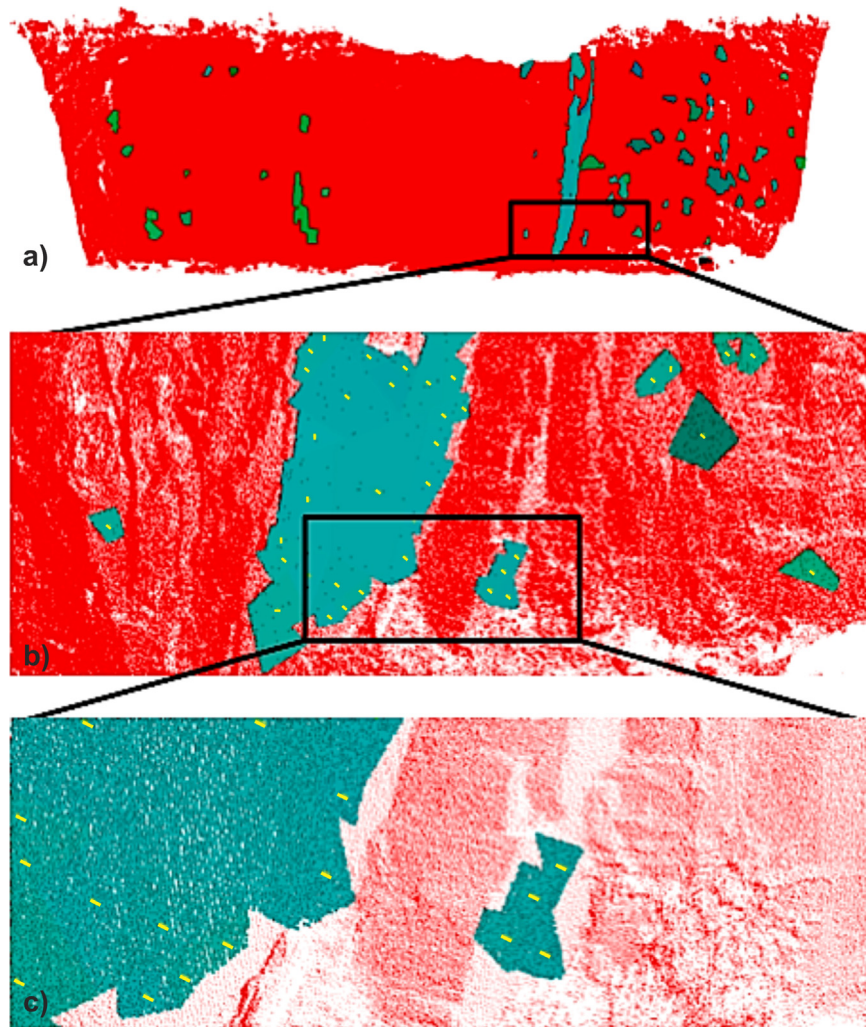
Fig. 6. Detected planes using the point cloud obtained from laser scan of the outcrop (algorithm parameters: $m \geq 6$, $k \leq 0.8$). Notice the sequence of magnifications from (a) to (c), highlighting region C. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Table 1
Mean and standard deviation of in situ measurements of SL and DA for plane C in Fig. 4d.

	SL $(n = 30)$	DA $(n = 30)$
Mean (M)	85.05°	86.23°
St. deviation (σ)	1.24°	1.33°

Table 2
SL and DA obtained by the algorithm for region C of the experimental plane. Note that similar values were obtained for $m_{min} = 5$ and $m_{min} = 6$.

$m \geq m_{min}$	SL: $M(\sigma)$	DA: $M(\sigma)$	No. of planes
$m \geq 4$	87.99°	87.84°	1
$m \geq 5$ or 6	83.83°(1.39°)	85.49°(0.42°)	6
$m \geq 7$	84.39°(1.36°)	85.60°(0.62°)	2

4.2.3. Automatic plane detection and orientation estimation

After testing the automatic detection of planes by the algorithm and how accurately it estimated their orientation, the entire point cloud was used as input data. The algorithm was run with different values of m_{min} as before. The first round ($m_{min} = 4$) detected only two planes in the whole point cloud. These planes were not considered, because they were larger than the fracture planes of interest and represent a simple subdivision of the entire rock face by

Table 3
Comparison between the orientation(s) of the plane(s) calculated by the proposed algorithm and those measured with compass and clinometer ($SL_{compass} = 85.05°$ and $DA_{compass} = 86.23°$).

$m \geq m_{min}$	ΔSL	ΔDA
$m \geq 4$	0.06°	$-1.61°$
$m \geq 6$	4.21°	0.73°
$m \geq 7$	3.65°	0.63°

Table 4
SL and DA obtained by the algorithm when plane C detected from the entire outcrop (cf. Table 2).

$m \geq m_{min}$	SL: $M(\sigma)$	DA: $M(\sigma)$	No. of planes
$m \geq 5$	84.22°(0.49°)	85.44°(0.63°)	3
$m \geq 6$	84.17°(0.42°)	85.43°(0.51°)	4
$m \geq 7$	84.39°(0.53°)	(85.43°(0.51°))	4

the algorithm due to the low value of m_{min}, which did not allow the algorithm to reach the minimum degree of fitting needed to detect smaller planes. Further subdivision was needed to meet the specified limits and detect more planes.

For $m \geq 5$, however, two large planes were detected in smoother areas (planes A and B) and many smaller ones, including

Table 5
Comparison between the orientation(s) of the plane(s) calculated by the proposed algorithm and those measured with compass and clinometer, when plane C was detected from the entire outcrop (cf. Table 3). The small discrepancies indicate the effectiveness of the proposed method.

$m \geq m_{min}$	ΔSL	ΔDA
$m \geq 5$	3.82°	0.79°
$m \geq 6$	3.87°	0.79°
$m \geq 7$	3.62°	0.79°

region C measured in situ in Fig. 5 (corresponding to region C in Section 4.2.2) (Fig. 5(a)). The detection of planes in the point cloud was improved by further processing with $m \geq 6$ (Fig. 5(b)); however, imposing these more restrictive conditions reduced the number of regions classified as planes, including most of the planar regions previously detected.

For $m \geq 6$, plane A was split into 29 additional planes and plane B into 17 additional planes (Fig. 5(b)). Using $m \geq 6$, the three planes C were further subdivided into four planes. The insets in Fig. 6 progressively enlarge region C. The detailed lower image shows that region C was actually interpreted to be four very similar planes indicated by the normals (yellow lines) originating at the average point of each plane. Region C might then be judged by the user to be a single plane because of these similarities, depending on the context of their investigation.

4.2.4. Second experiment: full point cloud

In this case, the full point cloud was used as input data and processed as before for different values of m_{min}. Fig. 6 shows the result for $m_{min} = 6$ for the whole outcrop and zoomed into the studied plane (Section 2).

The values listed in Table 5 indicate that the proposed algorithm is precise both in detecting planar regions and estimating their orientation. Indeed, the similarity of Tables 2 and 4 indicates that the algorithm can be used as a two-stage geological tool, where the user may firstly detect possible planar regions, then determine the orientation of those of interest.

Finally, we mention that the considered examples were performed on a Apple Macbook Pro with an Intel Core i7 2.7 GHz processor, using 4Gb DDR3 1333 MHz, with an Intel HD Graphics 3000 384Mb, using Mac OS X Lion 10.7.5 (11G63). The time required for processing a point cloud depends on the number of analyzed subregions which depends on the parameters m and k. For example, for $m_{min} = 5$ and $k_{max} = 0.8$, the processing lasted about 3 min. For $m_{min} = 6$ and $k_{max} = 0.8$, about 11 min.

4.3. Discussion

Two different situations were noticed when using the algorithm. Firstly, points on highly irregular surfaces did not meet the specifications of the parameters m and k, and no planes were detected. Secondly, the algorithm detected several different planes on a rough surface, the number of planes depending on the value chosen for m and k (see Figs. 3 and 5). As previously shown in Table 3, the maximum difference between measured and calculated planes was 4.22°, which is about half of the maximum 8° mentioned in the some literature for similar studies (Sturzenegger and Stead , 2009; Lato et al., 2009). For shallow-dipping fractures, however (dip less than 25°), the discrepancy in the strike has been reported to exceed 20° (Herda, 1999; Feng et al., 2001). Some conchoidal fractures may be interpreted as false planes, thus adding orientations that do not represent true planar structures.

Statistically, all measurements fell within a small standard deviation: for strike direction, between 1.39° and 0.42°; for dip angle, between 1.33° and 0.42°. An important point is that factors

like roughness of the rock face and curvature, combined with a low-resolution point cloud may produce, in some cases, an apparent inversion of subvertical or subhorizontal planes by 180°. This can happen with the use of any method, including the one proposed here. And this can lead to a problematic inversion of the dip direction. However, our method is less susceptible to that since it is more accurate and relatively robust to the presence of data noise, as discussed in the previous section.

5. Conclusion

In this work, we have proposed a novel algorithm for detecting planes in a point cloud X obtained from outcrop scans. The method uses PCA to detect planar regions in the point cloud and giving their strike line (SL) and dip angle (DA), which were compared to field observations by a hand-held compass and inclinometer.

Unlike other methods proposed in the literature, the algorithm is based on only one progressive scheme of division for planar approximation, defined by the variable structure of the points as given by the three PCs at each approximation step (Section 3), rather than the three, four or even five separate processing steps required by other methods.

Also, in contrast to other methods, the proposed method is entirely automatic. No user intervention is required during the process. The user is required only to initially define the classification parameters for the chosen specification (e.g., specifications S_1 or S_2, Eqs. (2) and (5)) and the appropriated point cloud for analysis. Tests using S_2 indicate that a good starting point is to define $k_{max} = 0.8$ and $m_{min} = 6$. Indeed, this parameter configuration is a useful compromise between seeking good plane approximations and not restricting the classification.

Since the proposed algorithm deals with very large sets of points, geometrical classification must be statistical. PCA deconstructs the space into those directions where the 3D data shows maximum statistical independence, and interprets the optimal piecewise geometrical linear approximations as corresponding to planar surfaces. Successive division of the input set X is determined by the error statistics at each approximation, which also results from PCA. This blend of properties in the algorithm follows a logical pathway to the solution of both problems: detecting planar regions from the LiDAR scan point cloud, and estimating their orientation.

References

Abellán, A., Vilaplana, J.M., Martínez, J., 2006. Application of along-range terrestrial laser scanner to a detailed rock fall study at Vall de Nria (Eastern Pyrenees, Spain). Eng. Geol. 88, 136–148. http://dx.doi.org/10.1016/j.enggeo.2006.09.012.
Bellian, J.A., Kerans, C., Jennette, D.C., 2005. Digital outcrop models: applications of

terrestrial scanning LiDAR technology in stratigraphic modeling. J. Sediment. Res. 75 (2), 166–176. http://dx.doi.org/10.2110/jsr.2005.013.

Bemis, S.P., Micklethwaiteb, S., Turnerc, D., James, M.R., Akcize, S., Thieleb, S.T., Bangashb, H.A., 2014. Ground-based and UAV-based photogrammetry: a multi-scale, high-resolution mapping tool for structural geology and paleoseismology. J. Struct. Geol. (69A), 163–178. http://dx.doi.org/10.1016/j.jsg.2014.10.007.

Buckley, S.J., Howell, J.A., Enge, H.D., Kurz, T.H., 2008. Terrestrial laser scanning in geology: data acquisition, processing and accuracy considerations. J. Geol. Soc. 165 (3), 625–638. http://dx.doi.org/10.1144/0016-76492007-100.

Feng, Q., Sjgren, P., Stephansson, O., Jing, L., 2001. Measuring fracture orientation at exposed rock faces by using a non-reflector total station. Eng. Geol. 59 (1–2), 133–146. http://dx.doi.org/10.1016/S0013-7952(00)00070-3.

Fernández, O., 2005. Obtaining a best fitting plane through 3d georeferenced data. J. Struct. Geol. 27 (5), 855–858. http://dx.doi.org/10.1130/GES00099.1.

Ferrari, F., Veronez, M.R., Tognoli, F.M.W., Inocencio, L.C., Paim, P.S.G., Silva, R.M., 2012. Visualizao e interpretao de modelos digitais de afloramentos utilizando laser scanner terrestre. Geocincias 31 (1), 79–91.

Fisher, J.E., Shakoor, A., Watts, C.F., 2014. Comparing discontinuity orientation data collected by terrestrial LiDAR and transit compass methods. Eng. Geol. 181, 78–92. http://dx.doi.org/10.1016/j.enggeo.2014.08.014.

Garca-Sellés, D., Falivene, O., Arbués, P., Gratacos, O., Tavani, S., Muñoz, J.A., 2011. Supervised identification and reconstruction of near-planar geological surfaces from terrestrial laser scanning. Comput. Geosci. 37, 1584–1594. http://dx.doi.org/10.1016/j.cageo.2011.03.007.

Gigli, G., Casagli, N., 2011. Semi-automatic extraction of rock mass structural data from high resolution LiDAR point clouds. Int. J. Rock Mech. Min. Sci. 48, 187–198. http://dx.doi.org/10.1016/j.ijrmms.2010.11.009.

Herda, H.H.W., 1999. Strike standard deviation for shallow-dipping rock fracture sets. Rock Mech. Rock Eng. 32 (4), 241–255.

Hodgetts, D., 2013. Laser scanning and digital outcrop geology in the petroleum industry: a review. Mar. Pet. Geol. 46, 335–354. http://dx.doi.org/10.1016/j.marpetgeo.2013.02.014.

James, M.R., Robson, S., 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application. J. Geophys. Res.: Earth Surf. 117 (F3), 1–17. http://dx.doi.org/10.1029/2011JF002289.

Jones, R.R., Pearce, M.A., Jacquemyn, C., Watson, F.E., 2015. Robust best-fit planes from geospatial data. Geosphere 12 (1), 196–202. http://dx.doi.org/10.1130/GES01247.1.

Lato, M., Diederichs, M.S., Hutchinson, D.J., Harrap, R., 2009. Optimization of LiDAR scanning and processing for automated structural evaluation of discontinuities in rock masses. Int. J. Rock Mech. Min. Sci. 46, 194–199. http://dx.doi.org/10.1016/j.ijrmms.2008.04.007.

Lato, M.J., Diederichs, M.S., Hutchinson, D.J., 2010. Bias correction for view-limited Lidar scanning of rock outcrops for structural characterization. Rock Mech. Rock Eng. 43, 615–628. http://dx.doi.org/10.1007/s00603-010-0086-5.

Lato, M.J., Voge, M., 2012. Automated mapping of rock discontinuities in 3D Lidar and photogrammetry models. Int. J. Rock Mech. Min. Sci. 54, 150–158. http://dx.doi.org/10.1016/j.ijrmms.2012.06.003.

Maerz, N.H., Youssef, A.M., Otoo, J.N., Kassebaum, T.J., Duan, J., 2013. A simple method for measuring discontinuity orientations from terrestrial LiDAR data. Environ. Eng. Geosci. 19 (2), 185–194.

Olariu, M.I., Ferguson, J.F., Aiken, C.L., Xu, X., 2008. Outcrop fracture characterization using terrestrial laser scanners: deep-water Jackfork sandstone at Big Rock Quarry, Arkansas. Geosphere 4, 247–259. http://dx.doi.org/10.1130/GES00139.1.

Pathak, K., Vaskevicius, N., Birk, A., 2009. Revisiting uncertainty analysis for optimum planes extracted from 3D range sensor point-clouds. In: International Conference on Robotics and Automation (ICRA).

Pearce, M.A., Jones, R.R., Smith, S.A.F., McCaffrey, K.J.W., 2011. Quantification of fold curvature and fracturing using terrestrial laser scanning. AAPG Bull. 95, 771–794. http://dx.doi.org/10.1306/11051010026.

Poppinga, J., Vaskevicius, N., Birk, A., Pathak, K., 2008. Fast plane detection and polygonalizaiton in noisy 3D range images. In: International Conference on Intelligent Robots and Systems (IROS).

Riquelme, A.J., Abellán, A., Toms, R., Jaboyedoff, M., 2014. A new approach for semi-automatic rock mass joints recognition from 3D point clouds. Comput. Geosci. 68, 38–52. http://dx.doi.org/10.1016/j.cageo.2014.03.014.

Slob, S., van Knapen, B., Hack, R., Turner, K., Kemeny, J., 2005. Method for automated discontinuity analysis of rock slopes with three-dimensional laser scanning. Transp. Res. Rec. 1913, 187–194. http://dx.doi.org/10.3141/1913-18.

Sturzenegger, M., Stead, D., 2009. Close-range terrestrial digital photogrammetry and terrestrial laser scanning for discontinuity characterization on rock cuts. Eng. Geol. 106 (3), 163–182. http://dx.doi.org/10.1016/j.enggeo.2009.03.004.

Voge, M., Lato, M.J., Diederichs, M.S., 2013. Automated rock mass discontinuity mapping from 3-dimensional surface data. Eng. Geol. 164, 155–162. http://dx.doi.org/10.1016/j.enggeo.2013.07.008.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. Structure-from-motion photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology 179, 300–314. http://dx.doi.org/10.1016/j.geomorph.2012.08.021.

Zeeb, C., Gomez-Rivas, E., Bons, P.D., Virgo, S., Blum, P., 2013. Fracture network evaluation program (FraNEP): a software for analyzing 2D fracture trace-line maps. Comput. Geosci. 60, 11–22. http://dx.doi.org/10.1016/j.cageo.2013.04.027.