

BENDING THE DOMING EFFECT IN STRUCTURE FROM MOTION RECONSTRUCTIONS THROUGH BUNDLE ADJUSTMENT

L. Magri^a, R. Toldo^{b,*}

^a Dip. Informatica - Università di Verona - Strada Le Grazie 15, Verona, Italy - magri.luca.l@gmail.com

^b 3Dflow SRL, c/o Computer Science Park - Strada Le Grazie 15, Verona, Italy

KEY WORDS: Bundle Adjustment, Doming effect, Ground detection

ABSTRACT:

Structure from Motion techniques provides low-cost and flexible methods that can be adopted in aerial surveying to collect topographic data with accurate results. Nevertheless, the so-called “*doming effect*”, due to unfortunate acquisition conditions or unreliable modeling of radial distortion, has been recognized as a critical issue that disrupts the quality of the attained 3D reconstruction. In this paper we propose a novel method, that works effectively in the presence of a nearly flat soil, to tackle *a posteriori* the doming effect: an automatic ground detection method is used to capture the doming deformation flawing the reconstruction, which in turn is wrapped to the correct geometry by iteratively enforcing a planarity constraint through a Bundle Adjustment framework. Experiments on real word datasets demonstrate promising results.

1. THE DOMING EFFECT

In recent years Structure from Motion techniques have enjoyed increasing popularity in the context of aerial surveying – where have been used in several applicative scenarios, ranging from earth science to heritage documentation – thanks to their ability to provide detailed imagery and accurate 3D models, by exploiting a flexible and low cost equipment, such as consumer-grade cameras.

Unfortunately, despite this appealing feature, Structure from Motion methods may be affected with two major problems. First, the incremental nature of Structure from Motion, which iteratively builds up a 3D model by alternating between camera pose and scene points estimates, suffers from *drifting* (Cornelis et al., 2004): i.e. the propagation of errors, due to noisy measurements, accumulated through the reconstruction process. As a remedy, Bundle Adjustment is routinely adopted to jointly refine the 3D structure of the scene together with the viewing parameters of the acquisition setup. Second, inaccuracies in modeling the radial distortion of camera lens (Fryer and Mitchell, 1987) (which may commonly occur in many self-calibration procedure) are reflected in the so called “*doming*” effect: a broad-scale systematic deformation of the reconstructed surface which often appears as a rounded-vault-distortion of flat surface – as recognized experimentally for example in (Rosnell and Honkavaara, 2012, Javernick et al., 2014, Ouédraogo et al., 2014). Interestingly, from a theoretical point of view, this phenomenon is related to critical loci arisen in the context of radial distortion self calibration (Wu, 2014). In practice, (James and Robson, 2014, Smith and Vericat, 2015) observe that the doming effect is particular evident for UAV systems that take images from near-parallel directions, causing systematic errors in the associated DEM. The rundown of these studies is that the doming effect may be alleviated with a careful design of a distributed network of ground points or by taking slightly off-vertical convergent imagery. However, even if these solutions succeed in reducing the amount of deformation, they do not remove its systematic nature and are limited only to laborious controlled setting.

*Corresponding author

Contributions In this work we present a novel solution to deal with the doming effect that mitigates the need of these precautions during the acquisition phase by correcting the deformation during the reconstruction step. In particular, under the assumption of a known geometry of the terrain, *e.g.* an approximate flat soil, we show how it is possible to compensate the doming deformation by enforcing a planarity constraint over a set of automatically selected points on the ground within the Bundle Adjustment framework. The main idea is to extract the relevant information to guide the Bundle Adjustment optimization by leveraging on a simple-to-implement robust model fitting technique that automatically assess and quantify the amount of doming deformation. In this way, we are able to allow accurate reconstructions in a wider array of circumstances. For example, this method will increase the quality of the attained 3D model when the adopted cameras manifest some unavoidable radial distortion or have been modeled in a defective way. In addition, the proposed approach can also deal with challenging acquisition scenarios, *e.g.* when oblique images are not available, or when it is difficult to define a careful designed net of control points.

This manuscript is structured as follows, a first overview of the method in Section 2 is followed by the description of the two main steps of the suggested solutions, namely ground detection in Section 3 and constrained Bundle Adjustment in Section 4. Sample results attained on real datasets are provided in Section 5.

2. OVERVIEW

Our method in a nutshell can be presented as follows. Starting from a tentative 3D reconstruction of the scene, before Bundle Adjustment is carried on, a robust model fitting technique is used to detect the ground, which is geometrically approximate either as a plane or as a second order surface, depending on the amount of the deformation. In this work we employ elliptic paraboloids to describe the domed deformation of the terrain, and we derive an *ad hoc* constrained-type least square fitting method. As far as ground detection is concerned, the points of the scene that do not belong to the terrain must be regarded as outliers,

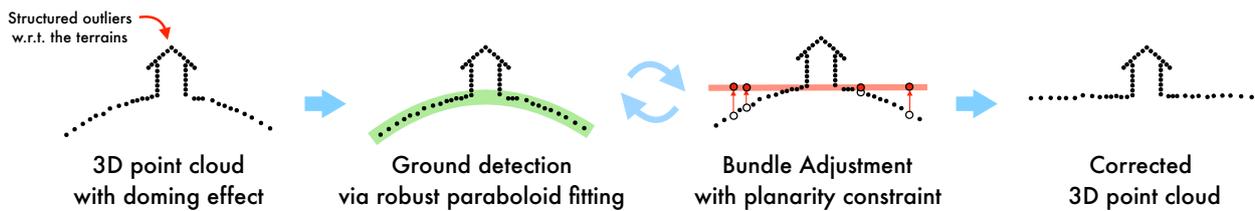


Figure 1. An overview of the main steps of the proposed method. At first the domed terrain is extracted from the point cloud via a robust paraboloid fitting technique. Hence points on the ground are selected and used to enforce a planarity constraint within an iterative Bundle Adjustment framework.

therefore, we exploits a robust framework, based on LMedS, combined with an outlier diagnostic criterion, to extract from the 3D point cloud the ground, either in the form of a plane or a paraboloid. The structure that has the least median of squared residuals is used to approximate the soil geometry. Residuals between points and the attained structures are computed, hence, a Forward Search procedure, based on the scrutiny of this residual distribution, is used to prune outliers and identify points laying on the ground. If the recovered structure happens to be a plane, the doming effect can be considered negligible, otherwise some of the inliers of the estimated paraboloid are used as control points and constrained to be coplanar in the subsequent Bundle Adjustment optimization. The procedure is iteratively repeated until the detected ground happen to be correctly flattened to the soil plane. An high-level illustration of this strategy is pictorially sketched in Figure 1.

3. GROUND DETECTION

We split up the problem of automatically extract the terrain from a 3D reconstruction of a scene in two sub-problems which are respectively addressed in the following two sections.

The first issue is the geometric description of the deformed terrain. While flat soil can be straightforwardly represented by a plane, fitting a second order surface to a domed terrain is a more challenging task, because, without further restraints, the estimation of a generic quadric surface may yield, in presence of noise and occlusions, to different kinds of quadrics and may also get stuck in degenerate cases that do not capture adequately the doming effect. For these reasons, in the next section, we conceive a simple and direct specific-paraboloid least-square fitting method that avoids these difficulties.

The second problem to deal with is robustness, which is discussed in Section 3.2 where LMedS and Forward Search are presented.

3.1 Specific paraboloid fitting

Elliptic paraboloids are manageable parametric models that well represent the doming deformation. They are special instances of quadric surfaces. A general quadric surface can be represented as the vanishing locus of a second order polynomial of the form

$$F(\mathbf{Q}, \mathbf{X}) = \mathbf{X}^\top \mathbf{Q} \mathbf{X} = 0 \quad (1)$$

where \mathbf{Q} is a 4×4 real symmetric matrix and $\mathbf{X} = [x, y, z, 1]^\top$ denotes a point in the euclidean space in homogeneous coordinates. $F(\mathbf{Q}, \mathbf{X})$ represents the algebraic distance of a point \mathbf{X} to the quadric \mathbf{Q} . Given a sparse cloud of n points $X = \{\mathbf{X}_i\}_{i=1}^n$, the general problem of fitting a quadric may be tackled via a least square approach by minimizing the sum of squared algebraic distances

$$\operatorname{argmin}_{\mathbf{Q}} \sum_{x_i \in X} F(\mathbf{Q}, \mathbf{X}_i)^2. \quad (2)$$

In order to avoid the trivial solution $\mathbf{Q} \equiv 0$, the entries of \mathbf{Q} may be constrained to satisfy a relation of some kind. We recall that two matrices \mathbf{Q}, \mathbf{Q}' that differ for a multiplicative factor, represent the same quadric, therefore we have the freedom to arbitrarily scale the entries of \mathbf{Q} .

This feature can be favorably exploited to derive a quadratic constrain to force the quadric to be an elliptic paraboloid. This idea was firstly suggested by (Fitzgibbon et al., 1999) for the problem of fitting an ellipse in the plane. Here we revisit and generalize to our paraboloid fitting scenario this approach. The main advantages are: (i) ensure some desirable features* that well describe the geometry of the soil to the fitted surface – namely, the surface is composed by a single connected component, the nature of the points is *elliptical* (i.e. a point of the surface is elliptical if the surface in its neighborhood is on the same side of the tangent plane) (ii) avoid degenerate cases in the presence of noise and occlusions.

We recall that \mathbf{Q} represents an elliptic paraboloid if and only if the following two conditions on the so called *orthogonal invariants* hold

$$I_4 = \det(\mathbf{Q}) < 0, \text{ and } I_3 = \det \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{12} & q_{22} & q_{33} \end{bmatrix} = 0. \quad (3)$$

We confine our discussion to paraboloids having the principal axis vertically aligned, as we assume that the reference system has the z -axis pointing towards the zenith. In this case Equation (1) can be further simplify as

$$a_1 x^2 + a_2 xy + a_3 y^2 + a_4 x + a_5 y + a_6 z + a_7 = 0, \quad (4)$$

and the corresponding quadric matrix turns out to be

$$\mathbf{Q} = \begin{bmatrix} a_1 & a_2/2 & 0 & a_4/2 \\ a_2/2 & a_3 & 0 & a_5/2 \\ 0 & 0 & 0 & a_6 \\ a_4/2 & a_5/2 & a_6 & a_7 \end{bmatrix}, \quad (5)$$

in this way the condition $I_3 = 0$ is always satisfied. As regards the constraint $I_4 < 0$, with some calculations, it can be reduced to

$$a_1 a_3 - \frac{a_2^2}{4} > 0. \quad (6)$$

Rather than working with this inequality, which in general is difficult to deal with, we take advantage of the freedom to arbitrarily scale the entries of \mathbf{Q} : we simply integrate the scaling factor into

* Also ellipsoids can be used to describe the terrain, and similar method can be derived also for these surfaces. Here we prefer elliptic paraboloid as they depend on fewer parameters.

Equation (6) by enforcing the equality

$$a_1 a_3 - \left(\frac{a_2}{2}\right)^2 = 1. \quad (7)$$

By wrapping up the quadric parameters in vectorial form as $\mathbf{a} = [a_1, a_2, a_3, a_4, a_5, a_6, a_7]^\top$, we can rephrase more compactly this condition as a quadratic constraint $\mathbf{a}^\top \mathbf{C} \mathbf{a} = 1$ of the form

$$\mathbf{a}^\top \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{a} = 0. \quad (8)$$

In summary, starting from the problem in Equation (2), we end up with the following constrained fitting problem

$$\min \|\mathbf{D}\mathbf{a}\|^2 \text{ constrained to } \mathbf{a}^\top \mathbf{C} \mathbf{a} = 1 \quad (9)$$

where \mathbf{D} is the $n \times 7$ design matrix

$$\mathbf{D} = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & z_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n & z_n & 1 \end{bmatrix}. \quad (10)$$

This problem can be efficiently solved, as explained in (Fitzgibbon et al., 1999), by retaining as a solution the eigenvector $\hat{\mathbf{a}}$ that corresponds to the single negative generalized eigenvalue of the associated generalized eigenproblem

$$\begin{cases} \mathbf{D}^\top \mathbf{D} \mathbf{a} = \lambda \mathbf{C} \mathbf{a} \\ \mathbf{a}^\top \mathbf{C} \mathbf{a} = 1. \end{cases} \quad (11)$$

3.2 Robust estimation technique

The reconstructed scene may include many structured points that do not lie on the soil and that act as outliers during the terrain extraction phase, as illustrated at the top of Figure 2. In presence of such outlying measurements, our specific-paraboloid fitting method, being a simple constrained least squares technique, will break down. Therefore, in order to compensate for its fragility, we integrate the ground detection step in a robust estimation framework based on the Least Median of Squares (LMedS) coupled with Forward Search to prune outliers at the end. An example of the results yielded by this robust approach on an instance of synthetic data is depicted at the bottom of Figure 2.

Least Median of Squares LMedS is a robust regression technique that achieves the maximum theoretical breakdown point of 0.5, *i.e.*, it can tolerate up to 50% percent of outliers in the data without being affected. At high level, if we denote by $r_{i,\mathbf{a}}$ the Euclidean distance between the i -th point \mathbf{x}_i and a model determined by \mathbf{a} , the LMedS estimate is given by:

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \text{median}_{\mathbf{x}_i \in X} r_{i,\mathbf{a}}^2, \quad (12)$$

that is, the estimate must yield the smallest value for the median of squared residuals computed for the entire data set.

The parameters \mathbf{a} are obtained via random sampling: a number of subsets $S \subset X$ composed by the minimum number of points

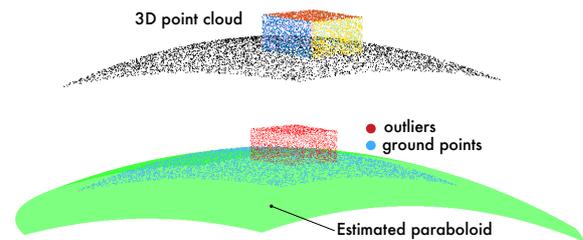


Figure 2. Ground detection with robust paraboloid fitting. A toy example with synthetic data TOP: the input data are given by the points on a paraboloid and points on a box lying on top of it (colored for visualization purpose). Gaussian noise is added on point coordinates BOTTOM: the paraboloid estimated via LMedS is displayed in green. Forward Search is used to dichotomize between inliers in light blue and outliers in red.

necessary to instantiate a model (e.g. 3 points for instantiate a plane or 7 to fit a paraboloid) are randomly selected from the data, and a parameter vector \mathbf{a}_S is fitted to the points in each subset. Hence, each \mathbf{a}_S is tested by calculating the squared residual distance r_{i,\mathbf{a}_S}^2 of each of the $n - k$ points in $X - S$ and the median of the residuals is computed. The \mathbf{a}_S corresponding to the smallest median is chosen as the estimate $\hat{\mathbf{a}}$. If n is the number of points in the initial set of data, there are $\binom{n}{k}$ different subsets of k points that should be considered. In order to speed up the process, the number of subsets to test can be reduced to

$$M = \frac{\log(1 - p)}{\log(1 - (1 - \epsilon)^k)}, \quad (13)$$

where p is the probability of calculating the correct estimate and ϵ is the percentage of outliers.

In our scenario we found beneficial to adopt a simple bucketing technique to encourage the sampling of well separated points, in order to reduce the standard error on the estimated parameter by leveraging on data that are more spread out.

It is known that, when in addition to outliers, Gaussian noise is present, the relative statistical efficiency of LMedS is low.

Therefore, to compensate for this deficiency, it is important to refine the fit on a outliers-free set of data. To this end, a common practice is to adopt scale estimation strategies to dichotomize between noisy inliers and gross outliers. Along this line, common solutions consist in univariate single-step outlier detection procedures that reject those points whose residuals are greater than a certain threshold. This threshold can be fixed in advance or computed adaptively from the data, for example using the median absolute deviation (MAD). Other solutions involve the use of sequential procedures that monitor the distribution of residuals either by *inward testing*, that performs successive elimination of points with high residual, or by *outward testing*, that adds sequentially points to the inlier set until a certain criterion is met.

Forward Search Forward Search (FS) (Hadi and Simonoff, 1993, Atkinson and Riani, 2000) is an outlier detection method that can be ascribed to the latter category. The algorithm starts from an initial subset of observations of size $m < n$ that is known to be outlier free. In our case this is the minimum sample set \hat{S} that corresponds to the $\hat{\mathbf{a}}$ obtained by LMedS, and $m = k$. Then, the iterative process starts and at every i^{th} iteration the s samples ($s = m + i - 1$) with the lowest residual are used to estimate

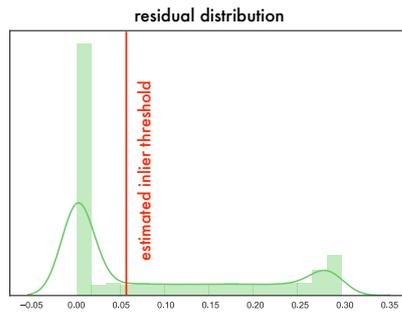


Figure 3. The distribution of the residuals computed w.r.t. the paraboloid estimated from LMedS on the toy example of Figure 2, together with the inlier threshold (red line) obtained via Forward Search.

the parameters of the model \mathbf{a}_s and the $s + 1$ residual r_{s+1, \mathbf{a}_s} is checked to detect the point when the iteration starts including outliers.

Among the several monitoring heuristics proposed in literature, we consider here one of the simplest, that dates back to (Hadi and Simonoff, 1993). The forward search stops when:

$$r_{s+1, \mathbf{a}_s}^2 \geq (t_{(1-\alpha/2)(s+1), s-k} \sigma_s)^2 \quad (14)$$

where $t_{(1-\alpha/2)(s+1), s-k}$ is the $1 - \alpha/2(s + 1)$ quantile of the t-distribution with $s - k$ degrees of freedom, and σ_s is the standard deviation of the least s residuals (the residual divided by σ_s is also called “studentized” residual). The main idea is to recognize the first point whose residual does not follow the gaussian distribution of the inlier to the model.

The i -th iteration of the algorithm can be summarized as follows:

1. Estimate the parameters of the model \mathbf{a}_s using the $s = m + i - 1$ points in the clean subset;
2. Calculate the residuals r_{s+1, \mathbf{a}_s} of all the n points and arrange them in ascending order;
3. Test the $(s + 1)^{th}$ residual:

If $r_{s+1, \mathbf{a}_s}^2 \geq (t_{(1-\alpha/2)(s+1), s-k} \sigma_s)^2$ then declare all the remaining $n - s$ observations as outliers and stop.

Otherwise increment i (include the $(s + 1)^{th}$ point in the inliers).

4. If $m + i - 1 = n$, then stop, otherwise go to step 1.

We applied this procedure to the synthetic data presented in Figure 2, and the results achieved are presented in Figure 3 where it can be appreciated the multi-modality of the residual distribution and how the estimated inlier threshold accurately dichotomized between inliers and structured outliers.

3.3 Model selection: flat or domed?

Our objective is to enforce a linear constraint on the domed ground, warping progressively into a plane the inliers points of the paraboloid retrieved by FS. Due to the iterative nature of this wrapping process, we need a mechanism to automatically assess when points

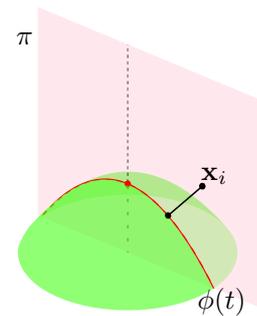


Figure 4. The distance between a point and a paraboloid is reduced to computing the distance between a point and a parabola in the plane.

on the paraboloid happen to be planar, or in other words, we need the ability to robustly detect both elliptic paraboloid and planes*.

In general, this issue turns to be a tricky model selection problem that requires some kind of regularization based on suitable model complexity measures. Fortunately, for our purpose, we find that a simple extension of the LMedS framework, aimed at handling at the same time different kind of parametric models, works profitably. Specifically, we tailor LMedS to simultaneously instantiate tentative paraboloids and planes. In this way, by expressing residuals in the same geometric unit of measure, we are able to neatly compare the goodness of fit of sampled structures and to choose the proper geometric primitive. Please note that the adoption of our specific-paraboloid fitting method is crucial for the success of this approach: a generic quadric fitting method would produce a degenerate estimate in case of flat terrain leading to inconsistent results.

Mixed sampling strategy: The adjustments to LMedS apply exclusively to the sampling step. Paraboloids are instantiated from minimum sample set of 7 points whereas 3D planes are obtained by sampling triplets of points. The geometric residual between a point \mathbf{x}_i and a plane, that is represented as a 4 parameter vector \mathbf{a} , is easily computed as

$$r_{i, \mathbf{a}} = \frac{|\mathbf{a}^\top \mathbf{x}_i|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}. \quad (15)$$

When an elliptic paraboloid is considered, \mathbf{a} becomes a 7 parameter vector, and the geometric residual of \mathbf{x}_i is computed as follows (for reference see Figure 4):

1. the plane π passing through the axis of the paraboloid and the point \mathbf{x}_i is computed;
2. the parabola given by the intersection of π and the paraboloid is determined and parametrized as $\phi(t)$;
3. the distance between \mathbf{x}_i and the parabola is calculated by finding the minimum of the squared norm $f(t) = \|\phi(t) - \mathbf{x}_i\|^2$. This boils down to finding the roots of the third degree polynomial obtained by differentiating $f(t)$ with respect to t , that can be easily computed in closed form.

In order to avoid 7-tuple of bad conditioned points, every time a minimum sample set is drawn, a planarity check is performed

* As a byproduct we will be able to determine if the input cloud of points is not affected by the doming distortion and there is no need of correction.

and, if it is deemed as planar, a plane is instantiated and added to the pool of putative models.

LMedS, aside from the described mixed sampling strategy, is implemented in a straightforward way. The model selection simply boils down to select the structure, either plane or paraboloid, that obtains the least median of squares, which in turn is used to approximate the ground and feed to Forward Search.

4. BUNDLE ADJUSTMENT WITH PLANAR CONSTRAINT

It is customary to adopt Bundle Adjustment (Triggs et al., 1999) to minimize the propagation of errors in Structure from Motion techniques in order to provide a jointly optimal estimate of 3D structure and viewing parameters. The terms “bundle” comes from the bundle of light rays that connects each camera center with the set of visible 3D points of the scene. This bundle of rays is adjusted via a proper optimization routine aimed at minimizing the overall residual error, *i.e.*, the distance between the re-projected 3D points and the corresponding measured image points, in formulas:

$$\min_{\mathbf{X}_i, \mathbf{P}_j} \sum_{i,j} w_{i,j} d(\mathbf{P}_j \mathbf{X}_i, \mathbf{x}_{i,j})^2. \quad (16)$$

\mathbf{P}_j is a 3×4 matrix encapsulating the interior and exterior orientations together with the radial distortion parameters of the j -th camera, \mathbf{X}_i represents the i -th points of the scene and $\mathbf{x}_{i,j}$ its 2D homogeneous coordinates in the j -th image, $d()$ indicates the Euclidean distance in the image plane. The coefficients $w_{i,j}$ are weights that indicates if a points \mathbf{X}_i is seen by a camera \mathbf{P}_j and can also be used to mitigate the effect of gross measurements, down-weighting the influence of outliers (Zach, 2014). This non linear least weighted square problem can be efficiently resolved through several numerical implementations rooted in Gauss-Newton methods, that conveniently exploit the structure of the Jacobian associated to Equation (16), see *e.g.* (Agarwal et al., 2010, Konolige and Garage, 2010) for a review and a more detailed discussion.

We want to modify Equation (16) in order to guide Bundle Adjustment to correct the camera parameters (including radial distortion values) and the point locations towards a reconstruction in which the terrain is plane. Under the flat soil assumption, the positions of the inliers of the attained paraboloid are certainly inaccurate. Therefore, enforcing a planarity constraint on their locations, is a useful cues that can be encompassed in the optimization process. With this idea in mind, we perform the following three steps:

1. At first we automatically pick a subset of ground points $\mathbf{X}_{i \in \mathcal{U}}$ indexed by $\mathcal{U} \subset \{1, \dots, n\}$ among the inliers obtained via Forward Search.
2. Then, we linearly project the selected points onto a designed plane τ , obtaining new 3D positions – say \mathbf{Y}_i with $i \in \mathcal{U}$.
3. At the end, we force the bundle of rays corresponding to the selected points to pass through the 3D positions \mathbf{Y}_i .

This means that Bundle Adjustment minimizes the re-projection error with respect to both the camera parameters and the 3D point positions for all the points, except the selected \mathbf{X}_i $i \in \mathcal{U}$ for

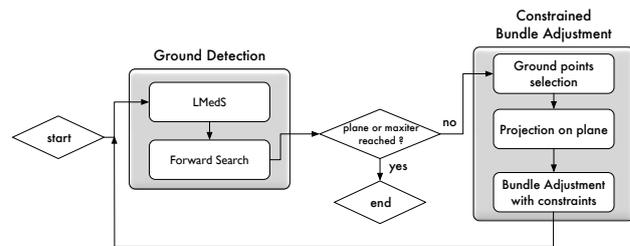


Figure 5. The flow chart of the proposed method

which the coordinates are kept fixed and equal to their planar projections \mathbf{Y}_i . Thus Equation (16) becomes

$$\min_{\mathbf{X}_i, \mathbf{P}_j} \sum_{i \notin \mathcal{U}, j} w_{i,j} d(\mathbf{P}_j \mathbf{X}_i, \mathbf{x}_{i,j})^2 + \sum_{i \in \mathcal{U}, j} w_{i,j} d(\mathbf{P}_j \mathbf{Y}_i, \mathbf{x}_{i,j})^2. \quad (17)$$

Selection of ground points The subset \mathcal{U} corresponding to the selected ground points is obtained through a decimation heuristic based on a bucketing technique. Only points that are inliers to the attained quadric are taken into account. Hence, the space of the scene is divided in regular cells with predetermined width. For each nonempty cell the point that is nearest to the cell-center is selected and the corresponding index is put in \mathcal{U} . In this way the selected points are well spread through the whole scene. An example of selected ground points in a real scene is presented in Figure 7.

Planar projection The plane τ onto which points are projected is chosen to be the tangent plane to the estimated quadric \mathbf{Q} at a certain point \mathbf{M} on \mathbf{Q} . The point \mathbf{M} can be picked in several ways: the vertex of the paraboloid is a good candidate, alternatively it can be specified by the user. Given the coordinates of \mathbf{M} , the coefficients of the plane τ are simply found as $\mathbf{t} = \mathbf{M}^T \mathbf{Q}$. The projection \mathbf{Y}_i on τ are computed via matrix multiplication $\mathbf{Y}_i = \mathbf{H} \mathbf{X}_i$, where

$$\mathbf{H} = (\mathbf{t} \mathbf{d}^T - (\mathbf{t}^T \mathbf{d}) \mathbf{I})^T, \quad (18)$$

\mathbf{d} represents a direction of projection expressed in homogeneous coordinates, *e.g.* $\mathbf{v} = [0, 0, 1, 0]^T$, and \mathbf{I} denotes the 4×4 identity matrix.

Weights setting If the i -th points is visible from the j -th camera, the corresponding addend that appears in Equation (17) is robustly weighted by the coefficients $w_{i,j}$. The weight is defined in terms of the re-projection error so to reduce the influence of contributions higher than a certain tolerance ϵ_i :

$$w_{i,j}(u) = \begin{cases} 1 & \text{if } u < \epsilon_i \\ 1/u & \text{otherwise.} \end{cases} \quad (19)$$

We set an higher tolerance ϵ_i for $i \in \mathcal{U}$ because we want to increase the importance of the error relative to the selected ground points. In this way we are able to guide Bundle Adjustment towards a planar solutions. In practice, if we express the re-projection error in pixel, we fix $\epsilon_i = 5$ if $i \in \mathcal{U}$ and $\epsilon_i = 1$ otherwise.

The whole process is iterative, and can be summarized through the block diagram shown in Figure 5.

5. EXPERIMENTAL VALIDATION

Most Structure from Motion pipelines are either sequential –*i.e.* a two-view reconstruction is sequentially augmented by adding in a new image at each iteration – or hierarchical: input images are at first organized in a dendrogram based on a similarity measure and then, following the order defined by this structure, smaller reconstructions are progressively merged in a complete 3D model. In order to assess the results of our constrained Bundle Adjustment, we used a Structure from Motion technique belonging to the latter category. Specifically, we employed 3DF Zephyr Aerial which implements the hierarchical approach of SAMANTHA – a state-of-the-art Structure from Motion technique described in (Toldo et al., 2015). It is worth noting that, the camera parameters were kept fixed and jointly optimized by Bundle Adjustment for blocks of similar images grouped together by Samantha.

We took in consideration three datasets of sparse reconstructions obtained from aerial survey: *village* composed by 6.323 points captured by 136 cameras, *country lane* composed by 18.059 points registered by 49 cameras, and *buildings* made up of 18.655 points and 103 cameras. The first two acquisitions were taken from an higher distance with respect to the last one. In all the three cases the internal orientations of the cameras were flawed, in particular, the radial distortion parameters were erroneously estimated. These unfavorable conditions resulted in a set of 3D points affected by the doming deformation, as can be seen in the first column of Figure 6 where a top view of the three scenes together with the color-coded elevation of the points is presented. In *village*, where the vertex of the paraboloid of the doming effect is in the middle of the scene, it is particular easy to recognize the parabolic geometry of the terrain.

The point cloud obtained by our constrained Bundle Adjustment are collected in the right column of Figure 6, from which it can be appreciated that the doming effect has been considerably attenuated, confirming the usefulness of this approach. One can also observe that, at the same time, the overall details and the other structures of interest present in the scene have been preserved thanks to the careful choice of ground points to guide Bundle Adjustment. For example, one can compare, the bottom-right corner of the *building* dataset and recognized that the positions of the yellow points – which corresponding to a roof in the scene – have not been moved by Bundle Adjustment as they have been deemed as outliers with respect to the ground paraboloid.

In Figure 7 a comparison between the input and the output point cloud on the *village* datasets, qualitatively illustrate the transformation induced by Bundle Adjustment which wraps the points towards a flat-soil geometry.

6. CONCLUSION AND FUTURE WORK

We have presented an effective method aimed at mitigating the doming effect. Our approach recover a correct reconstruction by properly integrating geometric clues in an iterative Bundle Adjustment framework. Further investigations about the connections between inaccuracies in the estimation of radial distortion parameters and the induced geometry of the deformed scene are in plan for future work. Moreover, a further advancement of this approach would be the introduction of a rematching phase of the keypoints at the end of the whole process. Also a careful study of the direction \mathbf{d} through which the planar projection is realized, would increase the accuracy of the attained reconstruction.

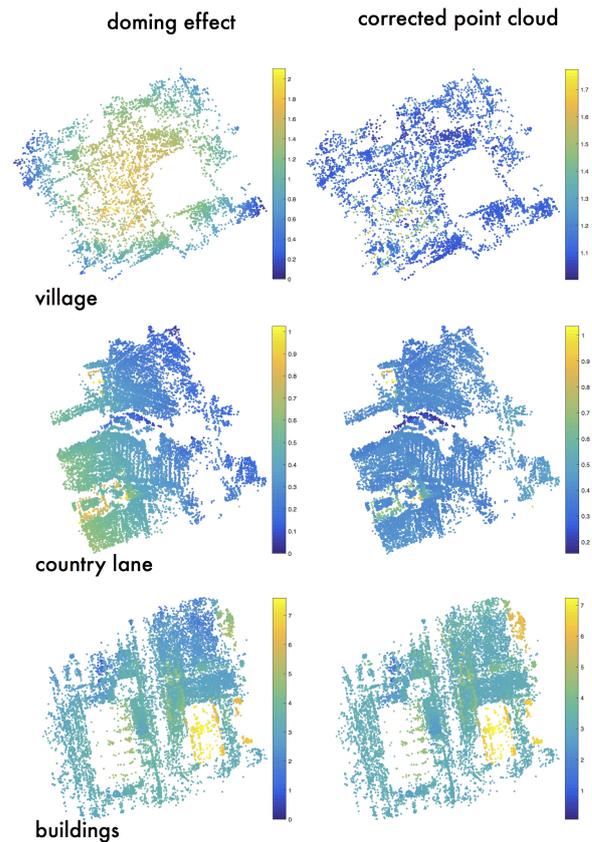


Figure 6. Top views of the doming effect on three datasets. The elevation values of the points is color-coded.

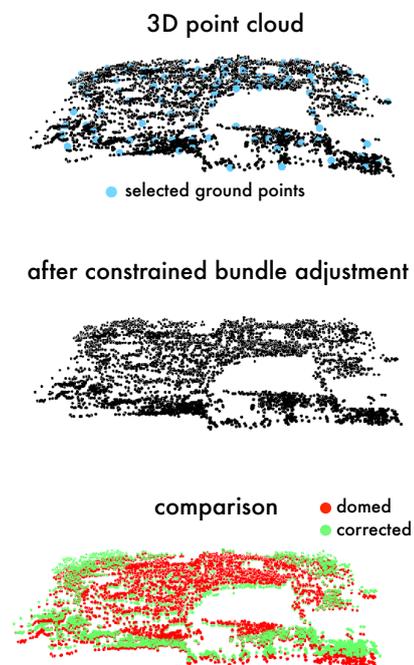


Figure 7. Sample results on the *village* dataset. TOP: the sparse data given in input, selected points on the ground X_i , $i \in \mathcal{U}$ are denoted as blue dots. MIDDLE: constrained Bundle Adjustment succeeds in mitigating the doming effect. BOTTOM: comparison between the input and the output point clouds.

REFERENCES

- Agarwal, S., Snavely, N., Seitz, S. and Szeliski, R., 2010. Bundle adjustment in the large. *Computer Vision–ECCV 2010* pp. 29–42.
- Atkinson, A. and Riani, M., 2000. *Robust Diagnostic Regression Analysis*. Springer Series in Statistics, Springer.
- Cornelis, K., Verbiest, F. and Van Gool, L., 2004. Drift detection and removal for sequential structure from motion algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(10), pp. 1249–1259.
- Fitzgibbon, A., Pilu, M. and Fisher, R. B., 1999. Direct least square fitting of ellipses. *IEEE Transactions on pattern analysis and machine intelligence* 21(5), pp. 476–480.
- Fryer, J. and Mitchell, H., 1987. Radial distortion and close range stereophotogrammetry. *Australian Journal of Geodesy, Photogrammetry and Surveying* 46(47), pp. 123–138.
- Hadi, A. S. and Simonoff, J. S., 1993. Procedures for the identification of multiple outliers in linear models. *Journal of the American Statistical Association* 88(424), pp. pp. 1264–1272.
- James, M. R. and Robson, S., 2014. Mitigating systematic error in topographic models derived from uav and ground-based image networks. *Earth Surface Processes and Landforms* 39(10), pp. 1413–1420.
- Javernick, L., Brasington, J. and Caruso, B., 2014. Modeling the topography of shallow braided rivers using structure-from-motion photogrammetry. *Geomorphology* 213, pp. 166–182.
- Konolige, K. and Garage, W., 2010. Sparse sparse bundle adjustment. In: *BMVC*, Vol. 10, pp. 102–1.
- Ouédraogo, M. M., Degré, A., Debouche, C. and Lisein, J., 2014. The evaluation of unmanned aerial system-based photogrammetry and terrestrial laser scanning to generate dems of agricultural watersheds. *Geomorphology* 214, pp. 339–355.
- Rosnell, T. and Honkavaara, E., 2012. Point cloud generation from aerial image data acquired by a quadcopter type micro unmanned aerial vehicle and a digital still camera. *Sensors* 12(1), pp. 453–480.
- Smith, M. W. and Vericat, D., 2015. From experimental plots to experimental landscapes: topography, erosion and deposition in sub-humid badlands from structure-from-motion photogrammetry. *Earth Surface Processes and Landforms* 40(12), pp. 1656–1671.
- Toldo, R., Gherardi, R., Farenzena, M. and Fusiello, A., 2015. Hierarchical structure-and-motion recovery from uncalibrated images. *Computer Vision and Image Understanding* 140, pp. 127–143.
- Triggs, B., McLauchlan, P. F., Hartley, R. I. and Fitzgibbon, A. W., 1999. Bundle adjustment: a modern synthesis. In: *International workshop on vision algorithms*, Springer, pp. 298–372.
- Wu, C., 2014. Critical configurations for radial distortion self-calibration. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 25–32.
- Zach, C., 2014. Robust bundle adjustment revisited. In: *ECCV* (5), pp. 772–787.