

# VIDEOGRAMMETRIC RECONSTRUCTION APPLIED TO VOLCANOLOGY: PERSPECTIVES FOR A NEW MEASUREMENT TECHNIQUE IN VOLCANO MONITORING

EMMANUELLE CECCHI<sup>1</sup>, JEAN-MARC LAVEST<sup>2</sup> AND BENJAMIN VAN WYK DE VRIES<sup>1</sup>

<sup>1</sup>Laboratoire Magmas et Volcans, LMV, UMR 6524 du CNRS, Blaise-Pascal University, Clermont-Ferrand,

<sup>2</sup>Laboratoire des Sciences et Matériaux pour l'Electronique, et d'Automatique, LASMEA, UMR 6602 du CNRS, Blaise-Pascal University, Clermont-Ferrand

e-mail: cecchi@opgc.univ-bpclermont.fr

(Accepted February 21, 2002)

## ABSTRACT

This article deals with videogrammetric reconstruction of volcanic structures. As a first step, the method is tested in laboratory. The objective is to reconstruct small sand and plaster cones, analogous to volcanoes, that deform with time. The initial stage consists in modelling the sensor (internal parameters) and calculating its orientation and position in space, using a multi-view calibration method. In practice two sets of views are taken: a first one around a calibration target and a second one around the studied object. Both sets are combined in the calibration software to simultaneously compute the internal parameters modelling the sensor, and the external parameters giving the spatial location of each view around the cone. Following this first stage, a N-view reconstruction process is carried out. The principle is as follows: an initial 3D model of the cone is created and then iteratively deformed to fit the real object. The deformation of the meshed model is based on a texture coherence criterion. At present, this reconstruction method and its precision are being validated at laboratory scale. The objective will be then to follow analogue model deformation with time using successive reconstructions. In the future, the method will be applied to real volcanic structures. Modifications of the initial code will certainly be required, however excellent reconstruction accuracy, valuable simplicity and flexibility of the technique are expected, compared to classic stereophotogrammetric techniques used in volcanology.

Keywords: calibration, digital elevation model (DEM), videogrammetry, volcanology.

## INTRODUCTION

Quantification of morphological changes on volcanoes constitutes an important aspect of monitoring and hazard research. At present, vertical stereophotogrammetric techniques, sometimes using highly advanced equipment Gwinner *et al.* (2000), are used for 3D reconstruction. With such techniques it is difficult to measure a surface of complex geometry, especially steep features on cones and craters, or failure scars. Moreover, these methods are cumbersome owing to the intensive stereopreparation and picture acquisition constraints Villeneuve (2000); Zlotnicki *et al.* (1990). Recent studies in the field of vision have showed the possibility of using new reconstruction tools in volcanology. The reconstruction approach presented here is based on a multi-view sensor calibration stage followed by a 3-D reconstruction process using  $n$  independent views. This reconstruction process consists in deforming a generic meshed model of the studied object. At the moment, the method is being tested in laboratory on analogue experiments dealing with volcano deformation.

## SENSOR MODELLING AND ESTIMATION OF ITS SPATIAL ATTITUDE

### *Multi-view Calibration*

The initial stage of the method consists of modelling the sensor and positioning external geometry. A mathematical relation is established between 3D point coordinates of a scene and the 2D coordinates of the same points projected in the image.

A multi-view calibration method using a 2D target and based on a photogrammetric approach has been chosen Lavest *et al.* (1998); Beyer (1992). This method allow to estimate accurately and rapidly the internal and external parameters of the sensor, which will be used in the reconstruction process. In laboratory, the high reconstruction accuracy expected for the studied objects (sand cones) implies to have an accurate estimation of these parameters.

This multi-view calibration approach offers the following main advantages:

- An accurate knowledge of the calibration target is not necessary: its geometry is reestimated during the calibration process.
- The use of several views allows reliable estimation of distortion parameters.
- Thanks to its simplicity of practical implementation, the method allows sensor modelling (internal parameters) in the field.

The projection model used for image formation process is referred to Pin-hole optical model (e.g Fig. 1).

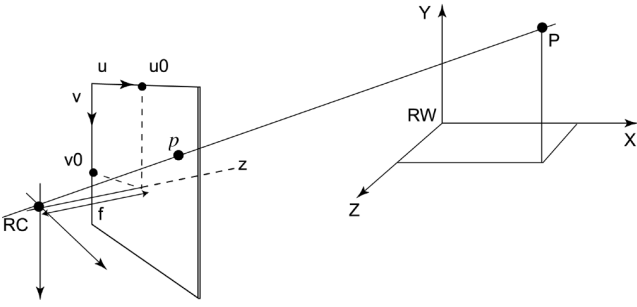


Fig. 1. *Pin-hole model, image geometry and coordinate system.*  $P$  is a 3D point in the world referential  $RW - XYZ$ .  $p_{(u,v)}$  is the projection of  $P$  in the image.  $RC$  camera referential, with  $z_i = f$ , i.e., camera focal length.  $(u_0, v_0)$ : principal point coordinates in the pixel referential.

Following notation will be used:

- $(T_x, T_y, T_z)_T$ : translation vector,
- $R$ : rotation matrix (composed of 9 elements  $(r_{11}, r_{12}, \dots, r_{33})$ , defined by the three Euler angles:  $\alpha$  rotation around the  $X$  axis,  $\beta$  around the  $Y$  axis, and  $\gamma$  around the  $Z$  axis),
- $dx, dy$ : scale factors of the elementary pixel size,
- $(do_{xr}, do_{yr})$  and  $(do_{xt}, do_{yt})$ : elements of the radial and tangential optical distortion according to  $x$  and  $y$ ,
- $(a_1, a_2, a_3)$  and  $(p_1, p_2)$ : polynomial coefficients modelling the radial and tangential distortion respectively.

During calibration, the Euclidean geometry of the observed target, and some extrinsic and intrinsic parameters are estimated. This approach takes into account the image deformation induced by optical distortion phenomena on the lens surface. Let us define the colinearity equations:

$$\left. \begin{aligned} u + e_x &= u_0 + (do_{xr} + do_{xt})/dx + \left. \begin{aligned} &\left(\frac{f}{dx}\right) \frac{r_{11}X + r_{12}Y + r_{13}Z + T_x}{r_{31}X + r_{32}Y + r_{33}Z + T_z} = P(\Phi) \\ v + e_y &= v_0 + (do_{yr} + do_{yt})/dy + \left. \begin{aligned} &\left(\frac{f}{dy}\right) \frac{r_{21}X + r_{22}Y + r_{23}Z + T_y}{r_{31}X + r_{32}Y + r_{33}Z + T_z} = Q(\Phi) \end{aligned} \right\} \end{aligned} \right\} (1)$$

where  $e_x$  and  $e_y$  can be seen as measurement errors related to  $x$  and  $y$  coordinates respectively. Calibrating the sensor is done by estimating the following parameter vector:

$$\Phi_{9+6m+3*n} = \left[ u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, X^1, Y^1, Z^1, \dots, X^n, Y^n, Z^n, T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \dots, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m \right]^T. \quad (2)$$

where:

- $u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y$  are the intrinsic parameters (altogether 9).
- $T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \dots, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m$  are the  $6m$  extrinsic parameters,  $m$  being the picture number,
- $X^1, Y^1, Z^1, \dots, X^n, Y^n, Z^n$  are the  $3n$  target point coordinates,  $n$  being the target point number.

The problem, in a least square sense, is to find  $\Phi$  that minimizes the error:

$$S = \sum_{i=1}^m \sum_{j=1}^n (e_{x_{ij}}^2 + e_{y_{ij}}^2). \quad (3)$$

As  $P(\Phi)$  and  $Q(\Phi)$  are non-linear functions in relation to the parameters of the vector  $\Phi$ , the Levenberg-Marquardt routine is used for the criterion  $S$  optimization.

This approach is known under the name "Bundle adjustments". In order to have an over-determined system, the number of measurements ( $2.m.n$ ) have to be greater or equal to the number of unknown parameters ( $9.6m.3n$ ) (internal, external parameters and target point coordinates).

In practice, pictures are taken around a plate composed of retro-reflective circles. The sensor turns around the target so as to create a convergent view bundle to constrain the 3D position of the target points. Circles should cover the entire sensor image to have the best estimation of the distortion coefficients. Fig. 2 shows four pictures among the  $m$  views used for the sensor calibration. On each view, 25 circles are visible. A subpixelar detection is done Lavest *et al.* (1999) and the parameter vector  $\Phi$  is calculated for each view of the calibration set.

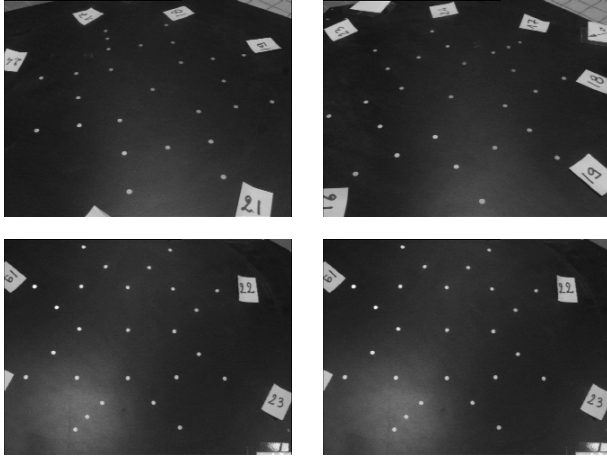


Fig. 2. Some views of the calibration set.

#### Sensor Position and Orientation

Fig. 3 shows a set of views around the cone to model, lying on the calibration target. Only a small number of white circles are visible on the pictures. These circles allow positioning the sensor in the model referential.

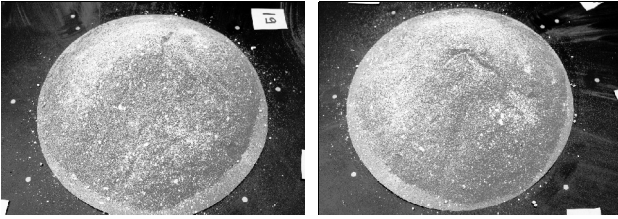


Fig. 3. Two views among height ( $2000 \times 1030$  pixels) of the cone picture set and the surrounding target points.

As the internal parameters of the sensor and the  $(X_i, Y_i, Z_i)$  coordinates of the target points are accurately known, the parameter vector  $\Phi$  to determine becomes:

$$\Phi_{6q} = \left[ T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \dots, T_x^q, T_y^q, T_z^q, \alpha^q, \beta^q, \gamma^q \right]^T. \quad (4)$$

However, it is preferable to combine information from the two view sets (calibration + reconstruction view sets) and optimize the whole system.  $m$  images containing 25 circles and  $q$  images with a few circles ensure the calculation of the sensor position and orientation. So the last  $6q$  extrinsic values of the vector  $\Phi$  at the solution give the different sensor attitudes around the cone.

## RECONSTRUCTION

The reconstruction stage is tested on scaled models, analogous to volcanic structures. Such analogue models make it possible to study the gravitational deformation of a cone induced by a ductile core. Cones, made up of a sand and plaster mixture and containing a silicone inclusion represent an altered core volcanic edifice. As the silicone has a viscous behaviour, cones deform with time.

The position and orientation of the sensor have been given by the calibration stage (*RiTi*) between the camera and the model for each view). So  $q$  views around the object are available, all being referenced in relation to the origin of the model referential. Several ways to reconstruct can be envisaged:

- To erode an initial volume, and only keep the voxels for which the projected luminance is coherent in all views Kutulakos (2000).
- To make a dense correlation of the points in the different images and triangulate to obtain a dense three-dimensional point cloud Koch *et al.* (1998).
- To construct a initial generic model and make an iterative deformation so as the projected texture of each facet of the meshed model be coherent in all views.

The approach using a initial surface model seemed the best adapted for our application. It appeared to be the simplest computational and most direct method, as it deals with a surface.

#### Initialization of the model

As the studied objects in laboratory are cones, the initial model is a conic surface of variable resolution, and is initialized from the information contained in pictures. The diameter, the height and the position of the cone are estimated. As it is shown in Fig. 4, three points selected by the user at the base of the cone in the image give, by ray tracing, their intersection with the support plane of the model (i.e.  $P_1, P_2, P_3$  points). The circle fitting the best these three points is then estimated. So we can define the vector normal to the circle passing by its centre  $\vec{PS}$ . A fourth point selected at the cone summit in the image gives the cone height by calculating the point on the cone normal that is the nearest to the vector  $\vec{c}_i^s$  (e.g. Fig. 4 and Fig. 5).

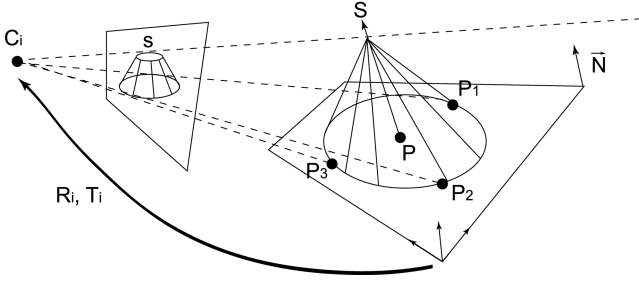


Fig. 4. Initialization of the model with a conic surface.

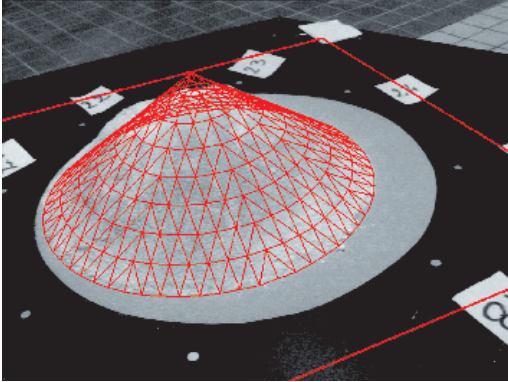


Fig. 5. One of the real cone view and its projected generic model.

### Reconstruction

The meshed model is composed of  $n$  facets and  $p$  3-D points. The number of facets can be defined, so a multi-scale approach is possible. We postulates that the model surface to reconstruct can be represented by a function of  $z = f(x,y)$  type, which is generally the case for the envisaged applications. This approach allows to limit the number of unknown parameters to estimate. After its initialization, the 3D generic model is deformed to fit the real object. This deformation consists, for a given model resolution, in calculating the  $z$  coordinate of each  $p$  points of the model in order to have textured areas, resulting from the projection of the facets in the pictures, coherent in all the views (minimization of a correlation criterion).

First, for each facet of the meshed model, a selection of pictures in which the correlation will be done is carried out. This pictures selection is based on (1) the visibility of the facet in the pictures, (2) the angle between view points. At this stage, each facet can be correlated in a minimum of two views

As the objective is to reconstruct as accurate as possible the studied object, a criterion measuring the similarity between this object (the cone) and the generic model is defined. This criterion, based on texture coherence, has to be minimized in order to have the greater adequacy between the object

and the generic model. As we have opted for a surface parametrization of  $z = f(x,y)$  type, the  $(xi, yi)$  coordinates of each  $p$  point will not be modified during the optimization process. Unknown parameters of the vector  $\Phi$  are the  $zi$  coordinates of each  $p$  point of the meshed model (summit facet).

$$\phi = [z_1, z_2, \dots, z_p]. \quad (5)$$

Let us define the optimization criterion:

$$C(\phi) = \sum_{i=1}^n \left[ \sum_{j=1}^{N_{i-1}} \sum_{k=j+1}^{N_i} (V_{ij} - V_{ik})^2 \right]. \quad (6)$$

Where:

- $N_i$  is the number of views among the picture set where the facet  $i$  is visible ( $N_i \geq 2$ ).
- $V_{ij}$  and  $V_{ik}$  are the centred and normalized luminance vectors of a facet  $i$  in the two images where the measurement is done. When a luminance vector is centred, its mean equal 0 and normalized means that the standart deviation to 1. Centering and normalizing vectors allows to take into account affine variations of light illumination during picture acquisition, that is variations of the mean luminance value and the grey-level range between pictures. The  $V_{ij}$  and  $V_{ik}$  vectors are composed of sampling point luminances of a facet. Their size depends on the model resolution and automatically adapts so that the 3D sampling of a triangular facet corresponds to a step of one pixel in the image. Besides, to reduce noise effect, a filtering is done for each measurement of the vectors, the mask filtering size being linked to the model resolution.

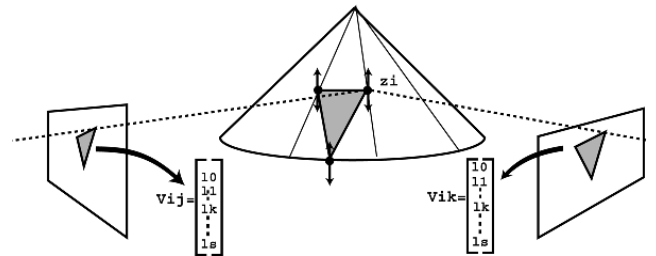


Fig. 6. Surface Adaptation.  $l_0, l_1 \dots l_s$  are the projected sampling point luminances of a facet  $i$ .

To summarize, for a given model resolution, the  $z_i$  coordinates of the  $p$  points are calculated in order that  $C(\Phi)$  be minimum (e.g. Fig 6). As the criterion is non linear in relation to its parameters, its minimization is iteratively done, using the Levenberg-Marquardt method.

## RESULTS AND PERSPECTIVES

In Fig. 7, results from the optimization for two different resolutions of the meshed model (respectively 2048 and 32768 facets) are shown. From an experimental point of view, the reconstructed surface is accurate enough to highlight the main morphological characteristics of the cone, that is to say two nested collapses on one of the flank. This feature can be seen even at a quite low resolution (res.4 = 2048 facets), while at a higher resolution (res.6 > 30000 facets) fine-scale textural irregularities due to grain flow are visible. Increasing the resolution model leads to get progressively closer to the real cone surface. However, the convergence time becomes very important due to the high number of measurements. As we use a non linear optimisation method (Levenberg-Marquardt), we need to give initial conditions close to the solution. This implies that the 3D model initialization is an important stage for the system convergence. However, the multi-scale approach (control of the model resolution) allow to keep a relative flexibility for the initialization. Finally, at the end of the optimization process, calculated camera positions could also be changed to refine the correlation.

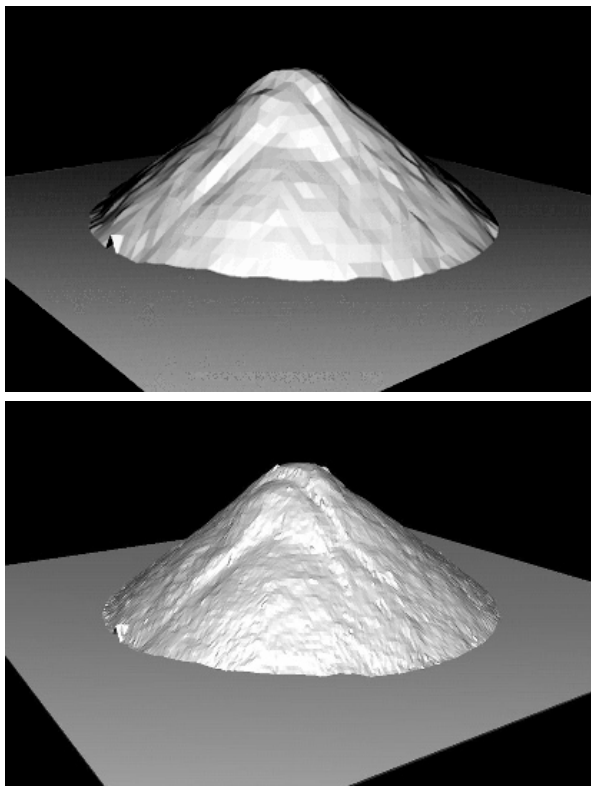


Fig. 7. Reconstruction results for different meshed model resolutions. Resolution (res.)4 = 2048 facets, res.6 = 32768 facets.

## CONCLUSION

The videogrammetric reconstruction technique presented here is new to volcanology. A first step will be to validate the method and test its precision at a laboratory scale. The objective will be then to follow the analogue model deformation with time for accurate quantitative analysis. The technique is also going to be applied to real volcanic structures. The transition between laboratory to natural textures will certainly require important modifications of the initial code, especially concerning the orientation and position sensor estimation. Several situations can be envisaged according to the available input data in the field. This is now in development. However the method promises to provide excellent reconstruction accuracy, a valuable simplicity and flexibility compared to classical stereophotogrammetry techniques currently used. This should make the technique useful and easily applicable for those who need quick and accurate terrain models on rough ground, such as volcanologists engaged in monitoring.

## REFERENCES

- Beyer H (1992). Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System . PhD thesis, Institut für Geodäsie und Photogrammetrie, Nr 51, ETH, Zürich.
- Gwinner K, Hauber E, Jaumann R, Neukum G (2000). High-resolution, digital photogrammetric mapping: A tool for earth science. EOS, 81(44).
- Koch R, Pollefeys M, Gool LV (1998). Multi viewpoint stereo from uncalibrated video sequences. In Proc. of Fifth European Conference on Computer Vision ECCV98, Freiburg, Germany. 55-71.
- Kutulakos K (2000). Approximate n-view stereo. In Proc. of Sixth European Conference on Computer Vision ECCV2000, Dublin, Ireland. 67-83.
- Lavest J, Viala M, Dhome M (1998). Do we really need an accurate calibration pattern to achieve a reliable camera calibration. In Proc. of Fifth European Conference on Computer Vision ECCV98, Freiburg, Germany. 158-74.
- Lavest J, Viala M, Dhome M (1999). Quelle précision pour une mire d'étalonnage? Trait Signal, 16(3):241-54.
- Villeneuve N (2000). Apports multi-sources à une meilleure compréhension de la mise en place des coulées de lave et des risques associés au Piton de la Fournaise. PhD thesis, Institut de Physique du Globe de Paris. Département des Observatoires. Observatoire volcanologique du Piton de la Fournaise, La Réunion.

- Zlotnicki J, Ruegg J, Bachelery P, Blum P (1990).  
Eruptive mechanism on Piton de la Fournaise volcano  
associated with the December 4, 1983, and January 18,  
1984 eruptions from ground deformation monitoring  
and photogrammetric surveys. *J Volcanol Geoth Res*,  
40:197-217.