

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



## 3D Modelling from Real Data

Gabriele Guidi<sup>1</sup> and Fabio Remondino<sup>2</sup>

<sup>1</sup>*Politecnico di Milano*

<sup>2</sup>*Fondazione Bruno Kessler, Trento  
Italy*

### 1. Introduction

The genesis of a 3D model has basically two definitely different paths. Firstly we can consider the CAD generated models, where the shape is defined according to a user drawing action, operating with different mathematical “bricks” like B-Splines, NURBS or subdivision surfaces (mathematical CAD modelling), or directly drawing small polygonal planar facets in space, approximating with them complex free form shapes (polygonal CAD modelling). This approach can be used for both ideal elements (a project, a fantasy shape in the mind of a designer, a 3D cartoon, etc.) or for real objects. In the latter case the object has to be first surveyed in order to generate a drawing coherent with the real stuff.

If the surveying process is not only a rough acquisition of simple distances with a substantial amount of manual drawing, a scene can be modelled in 3D by capturing with a digital instrument many points of its geometrical features and connecting them by polygons to produce a 3D result similar to a polygonal CAD model, with the difference that the shape generated is in this case an accurate 3D acquisition of a real object (reality-based polygonal modelling).

Considering only device operating on the ground, 3D capturing techniques for the generation of reality-based 3D models may span from passive sensors and image data (Remondino and El-Hakim, 2006), optical active sensors and range data (Blais, 2004; Shan & Toth, 2008; Vosselman and Maas, 2010), classical surveying (e.g. total stations or Global Navigation Satellite System - GNSS), 2D maps (Yin et al., 2009) or an integration of the aforementioned methods (Stumpf et al., 2003; Guidi et al., 2003; Beraldin, 2004; Stamos et al., 2008; Guidi et al., 2009a; Remondino et al., 2009; Callieri et al., 2011). The choice depends on the required resolution and accuracy, object dimensions, location constraints, instrument’s portability and usability, surface characteristics, working team experience, project’s budget, final goal, etc.

Although aware of the potentialities of the image-based approach and its recent developments in automated and dense image matching for non-expert the easy usability and reliability of optical active sensors in acquiring 3D data is generally a good motivation to decline image-based approaches. Moreover the great advantage of active sensors is the fact that they deliver immediately dense and detailed 3D point clouds, whose coordinate are metrically defined. On the other hand image data require some processing and a mathematical formulation to transform the two-dimensional image measurements into metric three-dimensional coordinates. Image-based modelling techniques (mainly

photogrammetry and computer vision) are generally preferred in cases of monuments or architectures with regular geometric shapes, low budget projects, good experience of the working team, time or location constraints for the data acquisition and processing.

This chapter is intended as an updated review of reality-based 3D modelling in terrestrial applications, with the different categories of 3D sensing devices and the related data processing pipelines.

## **2. Passive and active 3D sensing technologies**

In the following sections the two most used 3D capturing techniques, i.e. photogrammetry (section 2.1) and active range sensing (section 2.2 and 2.3) are reported and discussed.

### **2.1 Passive sensors for image-based 3D modelling techniques**

Passive sensors like digital cameras deliver 2D image data which need to be transformed into 3D information. Normally at least two images are required and 3D data can be derived using perspective or projective geometry formulations (Gruen & Huang, 2001; Sturm et al., 2011). Images can be acquired using terrestrial, aerial or satellite sensors according to the applications and needed scale. Terrestrial digital cameras come in many different forms and format: single CCD/CMOS sensor, frame, linear, multiple heads, SLR-type, industrial, off-the-shelf, high-speed, panoramic head, still-video, etc. (Mass, 2008). Common terrestrial cameras have at least 10–12 Megapixels at very low price while high-end digital back cameras feature more than 40 Megapixel sensors. Mobile phone cameras have up to 5 Megapixels and they could be even used for photogrammetric purposes (Akca & Gruen, 2009). Panoramic linear array cameras are able to deliver very high resolution images with great metric performances (Luhmann & Tecklenburg, 2004; Parian & Gruen, 2004). The high cost of these sensors is limiting their market and thus panoramic images are also generated stitching together a set of partly overlapped images acquired from a unique point of view with a consumer or SLR digital camera rotated around its perspective centre. This easy and low-cost solution allows to acquire almost Gigapixel images with great potential not only for visual needs (e.g., Google Street View, 1001 Wonders, etc.), but also for metric applications and 3D modelling purposes (Fangi, 2007; Barazzetti et al., 2010).

An interesting and emerging platform for image acquisition and terrestrial 3D modelling applications is constituted by Unmanned Aerial Vehicles (UAVs). UAVs can fly in an autonomous mode, using integrated GNSS with Inertial Navigation Systems (INS), stabilizer platform and digital cameras (or even a small range sensor) and can be used to get data from otherwise hardly accessible areas (Eisenbeiss, 2009).

#### **2.1.1 Photogrammetry**

Photogrammetry (Mikhail et al., 2001; Luhmann et al., 2007) is the most well-known and important image-based technique which allows the derivation of accurate, metric and semantic information from photographs (images). Photogrammetry thus turns 2D image data into 3D data (like digital models) rigorously establishing the geometric relationship between the acquired images and the scene as surveyed at the time of the imaging event. Photogrammetry can be done using underwater, terrestrial, aerial or satellite imaging sensors. Generally the term Remote Sensing is more associated to satellite imagery and their use for land classification and analysis or changes detection (i.e. no geometric processing).

The photogrammetric method generally employs minimum two images of the same static scene or object acquired from different points of view. Similar to human vision, if an object is seen in at least two images, the different relative positions of the object in the images (the so-called parallaxes) allow a stereoscopic view and the derivation of 3D information of the scene seen in the overlapping area of the images.

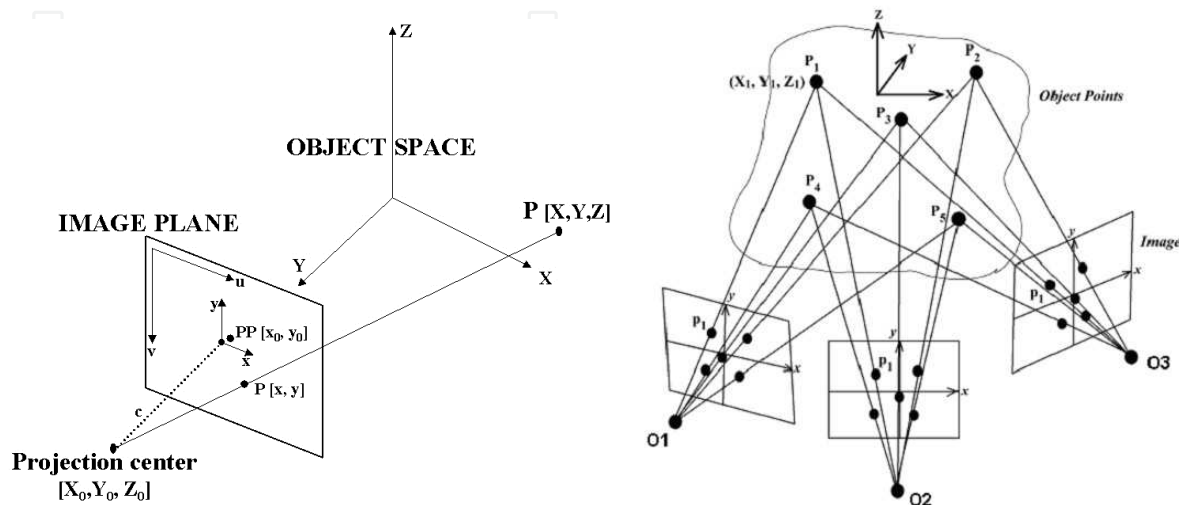


Fig. 1. The collinearity principle established between the camera projection center, a point in the image and the corresponding point in the object space (left). The multi-image concept, where the 3D object can be reconstructed using multiple collinearity rays between corresponding image points (right).

Photogrammetry is used in many fields, from the traditional mapping, monitoring and 3D city modelling to the video games industry, from industrial inspections to the movie production, from heritage documentation to medical field. Photogrammetry was always considered a manual and time consuming procedure but in the last decade many developments lead to a great improvement of the technique and nowadays many semi- or fully-automated procedures are available. When the project's goal is the recovery of a complete, detailed, precise and reliable 3D model, some user interaction in the modelling pipeline is still mandatory, in particular for geo-referencing and quality analysis. Thus photogrammetry does not aim at the full automation of the image processing but it has always as first goal the recovery of metric and accurate results. On the other hand, for applications needing 3D models for simple visualization or Virtual Reality (VR) uses, fully automated 3D modelling procedures can also be adopted (Vergauwen & Van Gool, 2006; Snavely et al., 2008).

The advantages of photogrammetry stay in the fact that (i) images contain all the information required for 3D modelling and accurate documentation (geometry and texture); (ii) photogrammetric instruments (cameras and software) are generally cheap, very portable, easy to use and with very high accuracy potentials; (iii) an object can be reconstructed even if it has disappeared or considerably changed using archived images (Gruen et al., 2004). But a large experience is required to derive accurate and detailed 3D models from images. This has limited a lot the use of photogrammetry in favour of the more powerful active 3D sensors, which allow easily the derivation of dense and detailed 3D point clouds with no user processing.

### 2.1.2 Basic principles of the photogrammetric technique

The basic principle of the photogrammetric processing is the use of multiple images (at least two) and the collinearity principle (Fig. 1). Such principle establishes the relationship between image and object space defining a straight line between the camera perspective center, the image point  $P(x, y)$  and the object point  $P(X, Y, Z)$ . The collinearity model is formulated as:

$$\begin{aligned} x &= -f \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + x_0 \\ y &= -f \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + y_0 \end{aligned} \quad (1)$$

with:

$f$	... camera constant or focal length	} interior orientation parameters
$x_0, y_0$	... principal point	
$X_0, Y_0, Z_0$	... position of the perspective center	} exterior orientation parameters
$r_{11}, r_{12}, \dots, r_{33}$	... elements of the rotation matrix	
$x, y$	... 2D image coordinates	
$X, Y, Z$	... 3D object coordinates	

All measurements performed on digital images refer to a pixel coordinate system while collinearity equations refer to the metric image coordinate system. The conversion from pixel to image coordinates is performed with an affine transformation knowing the sensor dimensions and pixel size.

For each image point measured in at least two images (generally called tie points), a collinearity equation is written. All the equations form a system of equations and the solution is generally obtained with an iterative least squares method (Gauss-Markov model), thus requiring some good initial approximations of the unknown parameters. The method, called *bundle adjustment*, provides a simultaneous determination of all system parameters along with estimates of the precision and reliability of the unknowns. If the interior orientation parameters are also unknowns, the method is named *self-calibrating bundle adjustment*.

The system of equations is iteratively solved with the least squares method and after the linearization and the introduction of an error vector  $e$ , it can be expressed as:

$$-e = A \cdot x \cdot l \quad (2)$$

with:

$e$  = error vector;

$A$  = design matrix  $n \times m$  (numb\_observations  $\times$  numb\_unknowns,  $n > m$ ) with the coefficients of the linearized collinearity equations;

$x$  = unknowns vector (exterior parameters, 3D object coordinates, eventually interior parameters);

$l$  = observation vector;

Generally a weight matrix  $P$  is added in order to weight the observations and unknown parameters during the estimation procedure. The estimation of  $x$  and the variance factor  $s$  is usually (not exclusively) attempted as unbiased, minimum variance estimation, performed by means of least squares and results in:

$$\hat{x} = (A^T P A)^{-1} \cdot A^T P l \quad (3)$$

with the residual  $v$  and the standard deviation a posteriori ( $s_0$ ) as:

$$v = A \cdot \hat{x} - l \quad (4)$$

$$\sigma_0 = \sqrt{\frac{v^T P v}{r}} \quad (5)$$

with  $r$  the redundancy of the system (numb\_observations - numb\_unknonws).

The precision of the parameter vector  $x$  is controlled by its covariance matrix  $C_{xx} = \sigma_0^2 (A^T P A)^{-1}$ .

For  $(A^T P A)$  to be uniquely invertible, as required in (Eq. 3), the network needs to fix an external "datum" i.e. the seven parameters of a spatial similarity transformation between image and object space. This is usually achieved by introducing some ground control points with at least seven fixed coordinate values. Another possibility is to solve the system (Eq. 2) in a free-network mode providing at least a known object's distance to retrieve the correct scale.

Depending on the parameters which are considered either known or treated as unknowns, the collinearity equations may result in different procedures (Table 1).

As previously mentioned, the photogrammetric reconstruction method relies on a minimum of two images of the same object acquired from different viewpoints. Defining  $B$  the baseline between two images and  $D$  the average camera-to-object distance, a reasonable  $B/D$  (base-to-depth) ratio between the images should ensure a strong geometric configuration and reconstruction that is less sensitive to noise and measurement errors.

A typical value of the  $B/D$  ratio in terrestrial photogrammetry should be around than 0.5, even if in practical situations it is often very difficult to fulfil this requirement. Generally, the larger the baseline, the better the accuracy of the computed object coordinates, although large baselines arise problems in finding automatically the same correspondences in the images, due to strong perspective effects. According to Fraser (1996), the accuracy of the computed 3D object coordinates ( $s_{XYZ}$ ) depends on the image measurement precision ( $s_{xy}$ ), image scale and geometry (e.g. the scale number  $S$ ), an empirical factor  $q$  and the number of images  $k$ :

$$\sigma_{XYZ} = \frac{q S \sigma_{xy}}{\sqrt{k}} \quad (6)$$

The collinearity principle and Gauss-Markov model of the least squares are valid and employed for all those images acquired with frame sensors (e.g. a SLR camera). In case of linear array sensors, other mathematical approaches should be employed. The description of such methods is outside the scope of this chapter.



The entire photogrammetric workflow used to derive metric and accurate 3D information of a scene from a set of images consists of (i) camera calibration and image orientation, (ii) 3D measurements, (iii) structuring and modelling, (iv) texture mapping and visualization. Compared to the active range sensors workflow, the main difference stays in the 3D point cloud derivation: while range sensors (e.g. laser scanners) deliver directly the 3D data, photogrammetry requires the mathematical processing of the image data to derive the required sparse or dense 3D point clouds useful to digitally reconstruct the surveyed scene.

Method	Observations	Unknowns
General bundle adj.	tie points, evt. datum	exterior param., 3D coord.
Self-calibrating bundle adj.	tie points, evt. datum	interior and exterior, 3D coord
Resection	tie points, 3D coord.	interior and exterior param.
Intersection	tie points, interior and exterior param.	3D coord.

Table 1. Photogrammetric procedures for calibration, orientation and point positioning.

2.1.3 Other image-based techniques

The most well-known technique similar to photogrammetry is computer vision. Even if accuracy is not the primary goal, computer vision approaches are retrieving interesting results for visualization purposes, object-based navigation, location-based services, robot control, shape recognition, augmented reality, annotation transfer or image browsing purposes. The typical computer vision pipeline for scene’s modelling is named “structure from motion” (Pollefeys et al., 2004; Pollefeys et al., 2008; Agarwal et al., 2009) and it is getting quite common in applications where metrics is not the primary aim.

Other image-based techniques allowing the derivation of 3D information from a single image use object constraints (Van den Heuvel, 1998; Criminisi et al., 1999; El-Hakim, 2000) or estimating surface normals instead of image correspondences with methods like shape from shading (Horn & Brooks, 1989), shape from texture (Kender, 1978), shape from specularity (Healey and Binford, 1987), shape from contour (Meyers et al., 1992), shape from 2D edge gradients (Winkelbach & Wahl, 2001).

2.2 Triangulation-based active range sensing

Active systems, particularly those based on laser light, make the measurement result nearly independent of the texture of the object being photographed, projecting references on its surface through a suitably coded light. Such light is characterized by an intrinsic information content recognizable by an electronic sensor, unlike the environmental diffuse light, which has no particularly identifiable elements. For example, an array of dots or a series of coloured bands are all forms of coded light. Thanks to such coding, active 3D sensors can acquire in digital form the spatial behaviour of an object surface. The output attainable from such a device can be seen as an image having in each pixel the spatial coordinates (x, y, z) expressed in millimetres, optionally enriched with colour information (R, G, B) or by the laser reflectance (Y). This set of 3D data, called “range image”, is generally a 2.5D entity (i.e. at each couple of x,y values, only one z is defined).

At present, 3D active methods are very popular because they are the only ones capable to acquire the geometry of a surface in a totally automatic way. A tool employing active 3D

techniques is normally called range device or, referring in particular to laser-based equipment, 3D laser scanner. Different 3D operating principles may be chosen depending on the object size hence on the sensor-to-object distance. For measuring small volumes, indicatively below a cubic meter, scanners are based on the principle of triangulation. Exceptional use of these devices have been done in Cultural Heritage (CH) applications on large artefacts (Bernardini et al., 2002; Levoy et al. 2000).

### 2.2.1 Basic principles

The kind of light that first allowed to create a 3D scanner is the laser light. Due to its physical properties it allows to generate extremely focused spots at relatively long ranges from the light source, respect to what can be done, for example, with a halogen lamp. The reason of this is related to the intimate structure of light, which is made by photons, short packets of electromagnetic energy characterized by their own wavelength and phase. A laser generates a peculiar light which is monochromatic (i.e. made by photons all at the same wavelength), and coherent (i.e. such that all its photons are generated in different time instants but with the same phase). The practical consequence of the first fact (monochromaticity) is that the lenses used for focusing a laser can be much more effective, being designed for a single wavelength rather than the wide spectrum of wavelengths typical of white light. In other words with a laser it is easier to concentrate energy in space. On the other hand the second fact (coherence) allows all the photons to generate a constructive wave interference whose consequence is a concentration of energy in time. Both these factors contribute to make the laser an effective illumination source for selecting specific points of a scenery with high contrast respect to the background, allowing to measure their spatial position as described below.

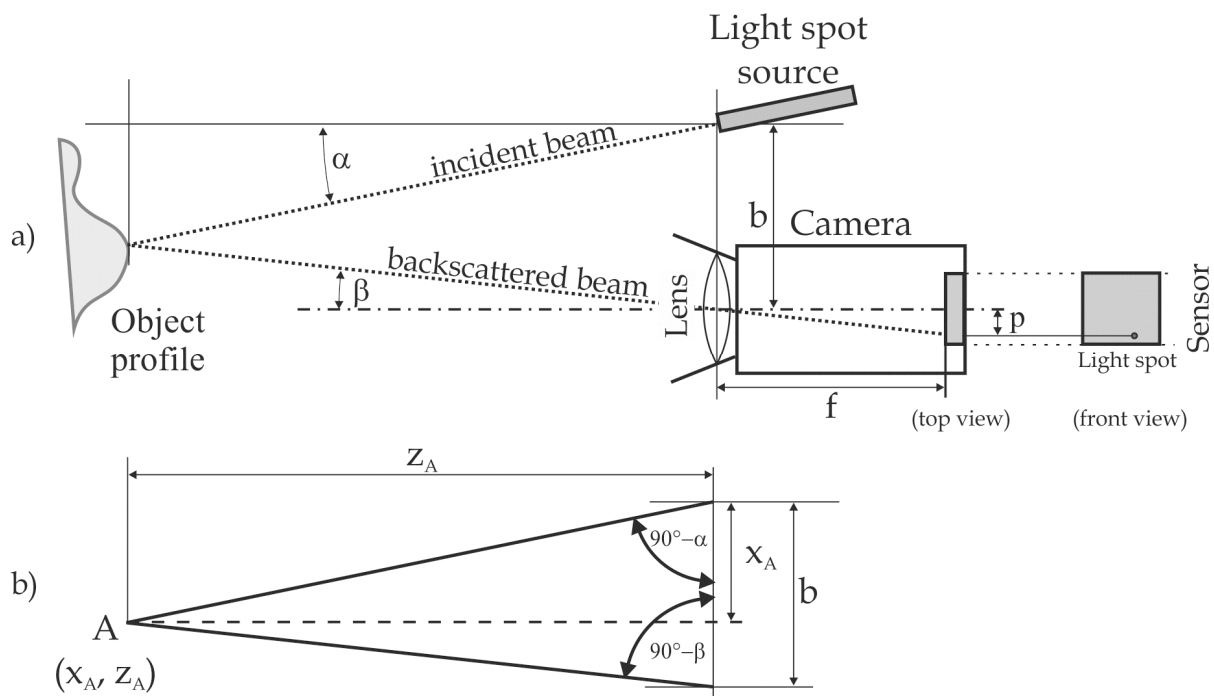


Fig. 2. Triangulation principle: a) xz view of a triangulation based distance measurement through a laser beam inclined with angle  $\alpha$  respect to the reference system, impinging on the surface to be measured. The light source is at a distance  $b$  from the optical centre of an image capturing device equipped with a lens with focal length  $f$ ; b) evaluation of  $x_A$  and  $z_A$ .



Let's imagine to have a range device made by the composition of a light source and a planar sensor, rigidly bounded each other. The laser source generates a thin ray producing a small light dot on the surface to be measured. If we put a 2D capture device (e.g. a digital camera) displaced respect to the light source and the surface is enough diffusive to reflect some light also toward the camera pupil, an image containing the light spot can be picked up. In this opto-geometric set-up the light source emitting aperture, the projection centre and the light spot on the object, form a triangle as the one shown in fig. 2a, where the distance between image capture device and light source is indicated as baseline  $b$ . The lens located in front of the sensor is characterized by its focal length  $f$  (i.e. distance in mm from the optical centre of the lens to the focal plane). On the collected image, a trace of the light spot will be visible in a point displaced with respect to the optical centre of the system.

Depending from the position of the imaged spot respect to the optical axis of the lens, two displacement components will be generated along the horizontal ( $x$ ) and vertical ( $y$ ) directions. Considering that the drawing in fig. 2a represents the horizontal plane ( $xz$ ) we will take into account here only the horizontal component of such displacement, indicated in fig. 2a as  $p$  (parallax). If the system has been previously calibrated we can consider as known both the inclination  $\alpha$  of the laser beam and the baseline  $b$ . From the spot position the distance  $p$  can be estimated, through which we can easily calculate the angle  $\beta$ :

$$\tan \beta = \frac{p}{f} \quad (7)$$

As evidenced in fig. 2b, once the three parameters  $b$ ,  $\alpha$  and  $\beta$  are known, the aforementioned triangle has three known elements: the base  $b$  and two angles ( $90^\circ - \alpha$ ,  $90^\circ - \beta$ ), from which all other parameters can be evaluated. Through simple trigonometry we go back to the distance  $z_A$  between the camera and point A on the object. This range, which is the most critical parameter and therefore gives name to this class of instruments (range devices), is given by:

$$z_A = \frac{b}{\tan \alpha + \tan \beta} \quad (8)$$

Multiplying this value by the tangent of  $\alpha$ , we get the horizontal coordinate  $x_A$ .

In this schematic view  $y_A$  never appears. In fact, with a technique like this, the sensor can be reduced to a single array of photosensitive elements rather than a matrix such as those which are equipped with digital cameras. In this case  $y_A$  can be determined in advance by mounting the optical measurement system on a micrometric mechanical device providing its position with respect to a known  $y$  origin. The Region Of Interest (ROI), namely the volume that can be actually measured by the range device, is defined by the depth of field of the overall system consisting of illumination source and optics. As well known the depth of field of a camera depends on a combination of lens focal length and aperture. To make the most of this area, it is appropriate that also the laser beam is focused at the camera focusing distance, with a relatively long focal range, in order to have the spot size nearly unchanged within the ROI. Once both these conditions are met, the ROI size can be further increased by tilting the sensor optics, as defined by the principle of Scheimpflug (Li et al, 2007).

### 2.2.2 3D laser scanner

The principle described above can be extended by a single point of light to a set of aligned points forming a segment. Systems of this kind use a sheet of light generated by a laser reflected by a rotating mirror or a cylindrical lens. Once projected onto a flat surface such light plane produces a straight line which becomes a curved profile on complex surfaces.

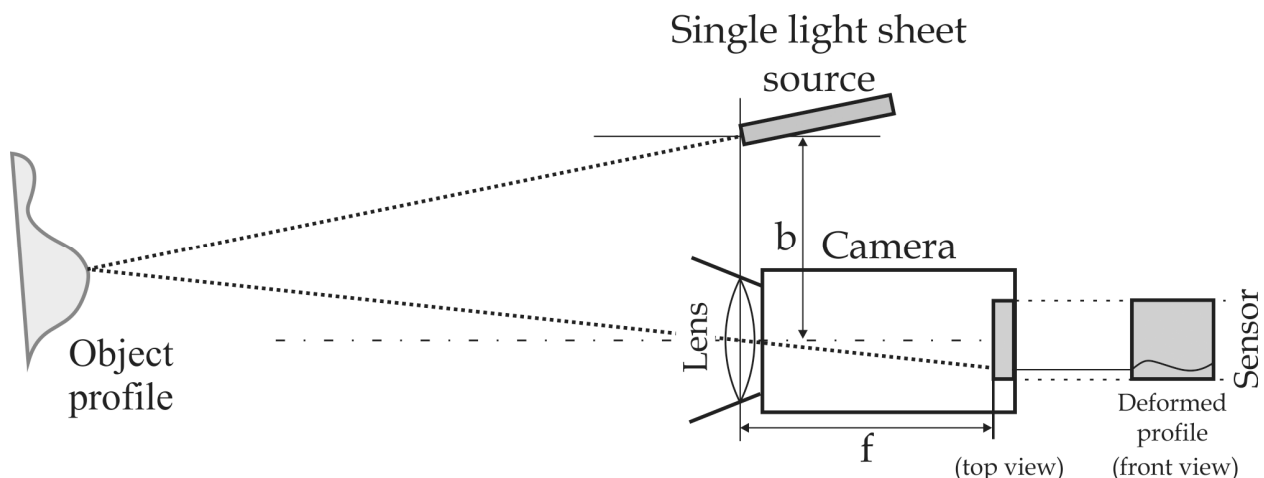


Fig. 3. Acquisition of coordinates along a profile generated by a sheet of laser light. In a 3D laser scanner this profile is mechanically moved in order to probe an entire area.

Each profile point responds to the rule already seen for the single spot system, with the only difference that the sensor has to be 2D, so that both horizontal and vertical parallaxes can be estimated for each profile point. Such parallaxes are used for estimating the corresponding horizontal and vertical angles, from which, together with the knowledge on the baseline  $b$  and the optical focal length  $f$ , the three coordinates of each profile point can be estimated.

This process allows therefore to calculate an array of 3D coordinates corresponding to the illuminated profile for a given light-object relative positioning.

By displacing the light plane along its normal of a small amount  $Dy$ , a different strip of surface can be probed, generating a new array of 3D data referred to an unknown geometrical region close to the first one. The 3D laser scanner is a device implementing the iteration of such process for a number of positions which generates a set of arrays describing the geometry of a whole area, strip by strip. This kind of range image (or range map), is indicated also as structured 3D point cloud.

### 2.2.3 Pattern projection sensors

With pattern projection sensors multiple sheets of light are simultaneously produced thanks to a special projector generating halogen light patterns of horizontal or vertical black and white stripes. An image of the area illuminated by the pattern is captured with a digital camera and each Black-to-White (B-W) transition is used as geometrical profile, similar to those produced by a sheet of laser light impinging on an unknown surface. Even if the triangulating principle used is exactly the same seen for the two devices mentioned above, the main difference is that here no moving parts are required since no actual scan action is performed. The range map is computed in this way just through digital post-processing of the acquired image.

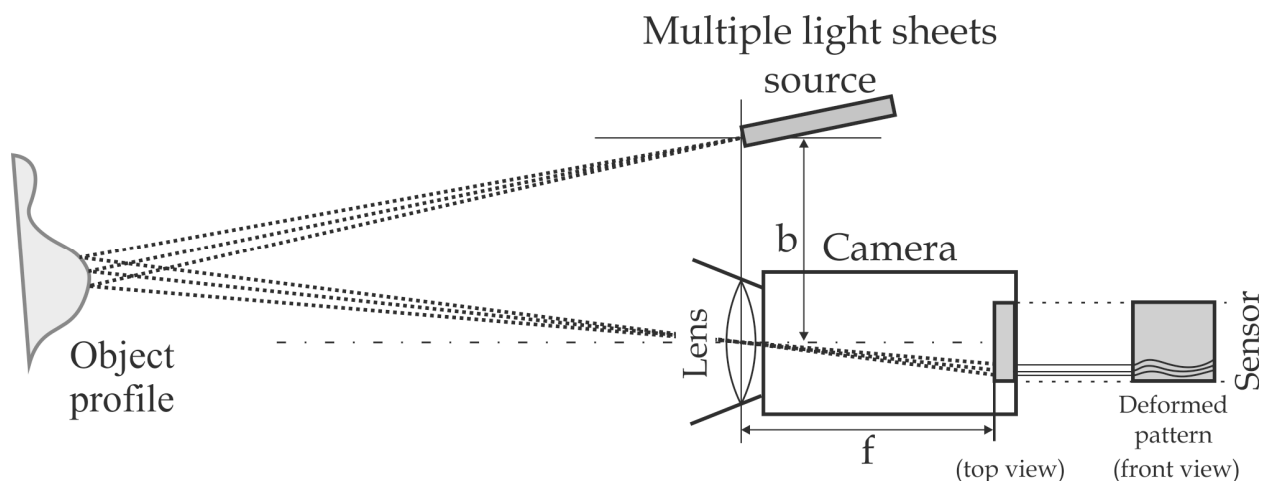


Fig. 4. Acquisition of coordinates along a different profiles generated by multiple sheets of white light.

The more B-W transitions will be projected on the probed surface, the finer will be its spatial sampling, with a consequent increase of the geometrical resolution. Therefore the finest pattern would seem the most suitable solution for gaining the maximum amount of data from a single image, but, in practical terms, this is not completely true. This depends by the impossibility to identify, in an image of an unknown surface with striped patterns projected on it, each single B-W transition, due to the possible framing of an unknown subset of the projected pattern (e.g. for surfaces very close to the camera), or for the presence of holes or occlusions generating ambiguity in the stripes order.

In order to solve such ambiguity this category of devices uses a sequence of patterns rather than a single one. The most used approach is the Gray coded sequence, that employs a set of patterns where the number of stripes is doubled at each step, up to reaching the maximum number allowed by the pattern projector. Other pattern sequences have been developed and implemented, such as phase-shift or Moirè, with different metrological performances.

In general the advantage of structured-light 3D scanners is speed. This makes some of these systems capable of scanning moving objects in real-time.

### 2.3 Time Of Flight (TOF) active range sensing

With active range sensing methods based on triangulation, the size of volumes that can be easily acquired ranges from a shoe box to a full size statue. For a precise sensor response the ratio between camera-target distance and camera-source distance (baseline), has to be maintained between 1 and 5. Therefore framing areas very far from the camera would involve a very large baseline, that above 1 m becomes difficult to be practically implemented. For larger objects like buildings, bridges or dams, a different working principle is used. It is based on optically measuring the sensor-to-target distance, having the *a priori* knowledge of angles through the controlled orientation of the range measurement device.

#### 2.3.1 Base principles

Active TOF range sensing is logically derived from the so-called "total station". This is made by a theodolite, namely an optical targeting device for aiming at a specific point in space, coupled with a goniometer for precisely measuring horizontal and vertical orientations,

integrated with an electronic distance meter. TOF, or time of flight, is referred to the method used for estimating the sensor-to-target distance, that is usually done by measuring the time needed by light for travelling from the light source to the target surface and back to the light detector integrated in the electronic distance meter.

Differently from a total station, a 3D laser scanner does not need that a human operator take aim at a specific point in space, therefore it does not have such sophisticate crosshair. On the other hand it has the capability to automatically re-orient the laser on a predefined range of horizontal and vertical angles, in order to select a specific area in front of the instrument. The precise angular estimations are then returned by a set of digital encoders, while the laser TOF gives the distance. As exemplified in fig. 5, showing a schematic diagram of a system working only on the xz plane analogously to what shown for triangulation based systems, it is clear that if the system return the two parameter distance ( $r$ ) and laser beam orientation ( $\alpha$ ), the Cartesian coordinates of A in the xz reference system are simply given by:

$$x_A = \rho \sin \alpha \quad (9)$$

$$z_A = \rho \cos \alpha \quad (10)$$

In case of a real 3D situation, in addition to the vertical angle an horizontal angle will be given, and the set of coordinate ( $x_A$ ,  $y_A$ ,  $z_A$ ) will be obtained by a simple conversion from polar to Cartesian of the three-dimensional input data.

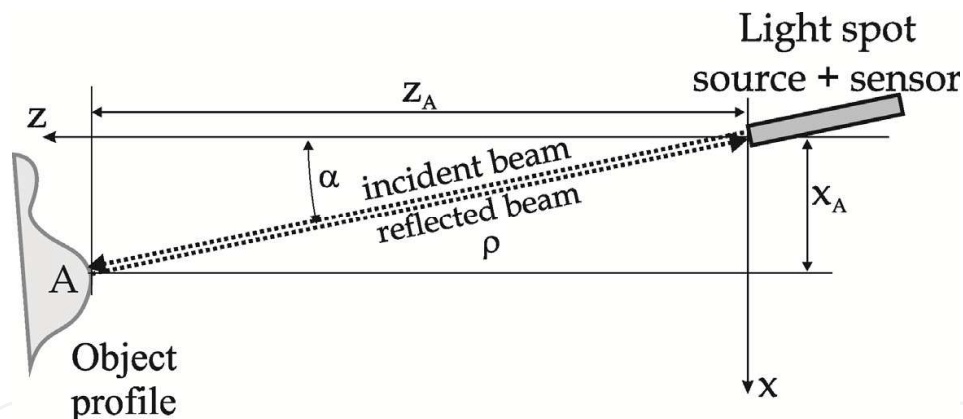


Fig. 5. Acquisition of coordinates of the point A through the *a priori* knowledge of the angle  $\alpha$ , and the measurement of the distance  $r$  through the Time Of Flight of a light pulse from the sensor to the object and back.

Systems based on the measurement of distance are in general indicated as LiDAR (Light Detection And Ranging), even if in the topographic area this acronym is often used for indicating the specific category of airborne laser scanner. The most noticeable aspect of such devices is in fact the capability to work at very long distance from the actual scanning surface, from half meter up to few kilometres, making such devices suitable also for 3D acquisition from flying platforms (helicopters or airplanes) or moving vehicles (boats or cars).

For ground based range sensors the angular movement can be  $360^\circ$  horizontally and close to  $180^\circ$  vertically, allowing a huge spherical volume to be captured from a fixed position. As

for triangulation based range sensors the output of such devices is again a cloud of 3D points originated by a high resolution spatial sampling an object. The difference with triangulation devices is often in the data structure. In TOF devices data are collected sampling an angular sector of a sphere, with a step not always fixed. As a results the data set can be formed by scan lines not necessarily all of the same size. Therefore the device output may be given by a simple list of 3D coordinates not structured in a matrix.

In term of performances, contributions to measurement errors may be given by both angular estimation accuracy and distance measurements. However, due to the very high speed of light, the TOF is very short, and this involves that the major source of uncertainty is due to its estimation that becomes a geometrical uncertainty once time is converted in distance. For this reason angle estimation devices implemented in this kind of laser scanners are similar each other. But different strategies for obtaining distance from light have been proposed for minimizing such uncertainty, all derived by approaches originally developed for radars.

An interesting sensor fusion is given by the Range-Imaging (RIM) cameras which integrate distance measurements (based on the TOF principle) and imaging aspects. RIM sensors are not treated in this chapter as not really suitable for 3D modeling applications.

### 2.3.2 PW laser scanner

Distance estimation is here based on a short Pulsed Wave (PW) of light energy generated from the source toward the target. Part of it is backscattered to the sensor, collected and reconverted in an electric signal by a photodiode. The transmitted light driving pulse and the received one are used as start/stop command for a high frequency digital clock that allows to count a number of time units between the two events. Of course the higher is the temporal resolution of the counting device, the finer will be the distance estimation. However, frequency limitations of electronic counting does not allow to go below a few tens of ps in time resolution, corresponding to some millimetres.

Considering that the speed of light is approximately  $c=3 \cdot 10^8$  m/s, and that the TOF is related to a travel of the light pulse to the surface and back (double of the sensor-to-target distance), the range will be given by:

$$r = \frac{TOF \cdot c}{2} \quad (11)$$

Therefore a small deviation in estimating TOF, for example in the order of 20 ps, will give a corresponding range deviation  $\Delta r = 1/2 \cdot (20 \cdot 10^{-12}) \cdot (3 \cdot 10^8)$  m = 3 mm.

For some recent models of laser scanner based on this principle (Riegel, 2010), the device is capable to detect multiple reflected pulses by a single transmitted pulse, provided by situations where multiple targets are present on the laser trajectory (e.g. a wall behind tree leaves). In this case the cloud of points is not anymore a 2.5D entity.

### 2.3.3 CW laser scanner (phase shift)

In this case distance is estimated with a laser light whose intensity is sinusoidally modulated at a known frequency, generating a Continuous Wave (CW) of light energy directed toward the target. The backscattering on the target surface returns a sinusoidal light wave delayed respect to the transmitted one, and therefore characterized by a phase difference with it. Similarly to the previous approach, the distance estimation is based on a comparison between the signal applied to the laser for generating the transmitted light wave:



$$s_{TX} = \cos(\omega_0 t) \quad (12)$$

and the signal generated by re-converting in electrical form the light backscattered by the surface and received by the range sensor:

$$s_{RX} = \cos(\omega_0 t + \varphi) \quad (13)$$

A CW laser scanner implement an electronic mixing the two signals, that corresponds to a multiplication of these two contributions. It can be reduced as follows:

$$\cos(\omega_0 t) \cdot \cos(\omega_0 t + \varphi) = \frac{1}{2} \cos(2\omega_0 t + \varphi) + \frac{1}{2} \cos(\varphi) \quad (14)$$

The result is a contribution at double the modulating frequency, that can be cut through a low-pass filter, and a continuous contribution, directly proportional to phase difference  $\varphi$ , that can be estimated. Since this angular value is directly proportional to the TOF, from this value the range can be evaluated similarly to the previous case. This indirect estimation of TOF allows a better performance in term of uncertainty for two main reasons: a) since the light sent to the target is continuous, much more energy can be transmitted respect to the PW case, and the consequent signal-to-noise ratio of the received signal is higher; b) the low-passing filtering required for extracting the useful signal component involves a cut also on the high frequency noise, resulting in a further decrease of noise respect to signal.

A peculiar aspect of this range measurement technique is the possibility to have an ambiguous information if the sensor-to-target distance is longer than the equivalent length of a full wave of modulated light, given by the ambiguity range  $r_{amb} = pc/w_0$ , due to the periodical repetition of phase. Such ambiguity involves a maximum operating distance that is in general smaller for CW devices rather than PW.

### 2.3.4 FM-CW laser scanner (laser radar)

In CW systems the need of a wavelength long enough for avoiding ambiguity, influence the range detection performance which is as better as the wavelength is short (i.e. as  $w_0$  grows). This led to CW solutions where two or three different modulation frequencies are employed. A low modulating frequency for a large ambiguity range (in the order of 100m), and shorter modulation frequencies for increasing angular (and therefore range) resolution. By increasing indefinitely the number of steps between a low to a high modulating frequency, a so-called chirp frequency modulation (FM) is generated, with a linear growing of the modulating frequency in the operating range. As light is generated continuously, this kind of instruments are indicated as FM-CW. Since this processing is normally used in radars (Skolnik, 1990), this devices is also known as "laser radar". The peculiar aspect of this approach is the capability to reduce the measurement uncertainty at levels much lower than that of PW laser scanners (typically 2-3 mm), and lower than that of CW laser scanners (less than 1mm on optically cooperative materials at the proper distance), competing with triangulation laser scanners, capable to reach a measurement uncertainty lower than 100 mm. Such devices have therefore the advantage of the spherical acquisition set-up typical of TOF laser scanners, with a metrological performance comparable to that of triangulation based devices, at operating distances from 1 to 20 meters, far larger than the typical triangulation devices operating range (0.5 to 2 m). For this reason such instruments have been experimented in applications where a



wide area and high precision are simultaneously required, like in industrial (Petrov, 2006) and CH (Guidi et al., 2005; Guidi et al., 2009b) applications.

### 3. Digital camera calibration and image orientation

Camera calibration and image orientation are procedures of fundamental importance, in particular for all those geomatics applications which rely on the extraction of accurate 3D geometric information from images. The early theories and formulations of orientation procedures were developed many years ago and today there is a great number of procedures and algorithms available (Gruen and Huang, 2001). Sensor calibration and image orientation, although conceptually equivalent, follow different strategies according to the employed imaging sensors. The camera calibration procedure can be divided in geometric and radiometric calibration but in this chapter only the geometric calibration of terrestrial frame cameras is reported.

#### 3.1 Geometric camera calibration

The geometric calibration of a camera (Remondino & Fraser, 2006) is defined as the determination of deviations of the physical reality from a geometrically ideal imaging system based on the collinearity principle: the pinhole camera. Camera calibration continues to be an area of active research within the Computer Vision community, with a perhaps unfortunate characteristic of much of the work being that it pays too little heed to previous findings from photogrammetry. Part of this might well be explained in terms of a lack of emphasis and interest in accuracy aspects and a basic premise that nothing whatever needs to be known about the camera which has to be calibrated within a linear projective rather than Euclidean scene reconstruction. In photogrammetry, a camera is considered calibrated if its focal length, principal point offset and a set of Additional Parameters (APs) are known. The camera calibration procedure is based on the collinearity model which is extended in order to model the systematic image errors and reduce the physical reality of the sensor geometry to the perspective model. The model which has proved to be the most effective, in particular for close-range sensors, was developed by D. Brown (1971) and expresses the corrections ( $\Delta x$ ,  $\Delta y$ ) to the measured image coordinates ( $x$ ,  $y$ ) as:

$$\Delta x = -\Delta x_0 + \frac{\bar{x}}{f} \Delta f + \bar{x} S_x + \bar{y} a + \bar{x} r^2 k_1 + \bar{x} r^4 k_2 + \bar{x} r^6 k_3 + 2P_1 \bar{x} \bar{y} + (2\bar{x}^2 + r^2) P_1 + 2\bar{x} \bar{y} P_2 \quad (15)$$

$$\Delta y = -\Delta y_0 + \frac{\bar{y}}{f} \Delta f + \bar{x} a + \bar{y} r^2 k_1 + \bar{y} r^4 k_2 + \bar{y} r^6 k_3 + 2P_1 \bar{x} \bar{y} + (2\bar{y}^2 + r^2) P_2 \quad (16)$$

with:

$$\begin{aligned} \bar{x} &= x - x_0; \\ \bar{y} &= y - y_0; \\ r^2 &= \bar{x}^2 + \bar{y}^2; \end{aligned}$$

Brown's model is generally called "physical model" as all its components can be directly attributed to physical error sources. The individual parameters represent:

$Dx_0, Dy_0, Df$  = correction for the interior orientation elements;

$K_i$  = parameters of radial lens distortion;

$P_i$  = parameters of decentering distortion;

$S_x$  = scale factor in x to compensate for possible non-square pixel;

$a$  = shear factor for non-orthogonality and geometric deformation of the pixel.

The three APs used to model radial distortion  $\Delta r$  are generally expressed via the odd-order polynomial  $\Delta r = K_1 r^3 + K_2 r^5 + K_3 r^7$ , where  $r$  is the radial distance. A typical Gaussian radial distortion profile  $\Delta r$  is shown in fig. 6a, which illustrates how radial distortion can vary with focal length. The coefficients  $K_i$  are usually highly correlated, with most of the error signal generally being accounted for by the cubic term  $K_1 r^3$ . The  $K_2$  and  $K_3$  terms are typically included for photogrammetric (low distortion) and wide-angle lenses, and in higher-accuracy vision metrology applications. The commonly encountered third-order barrel distortion seen in consumer-grade lenses is accounted for by  $K_1$ .

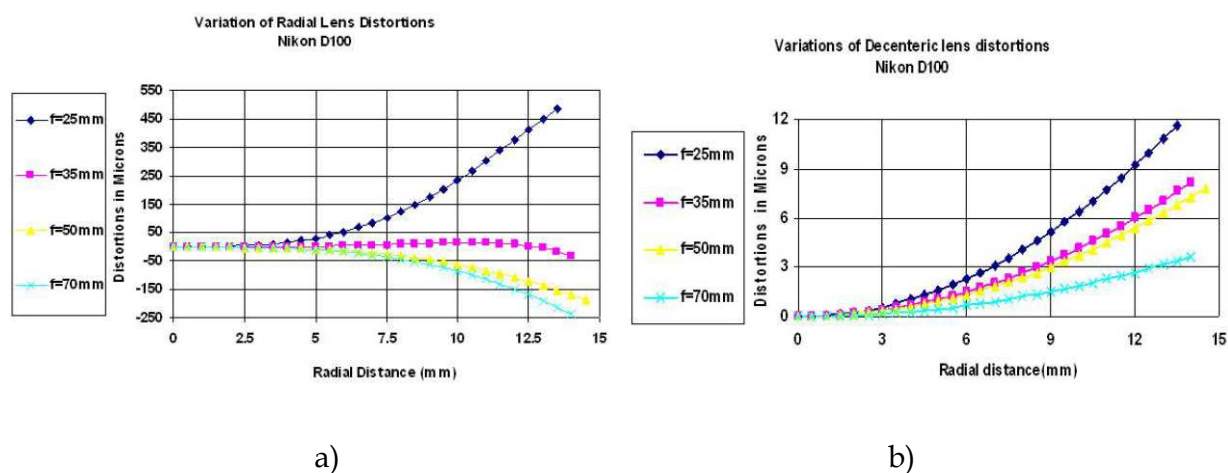


Fig. 6. Radial (a) and decentering (b) distortion profile for a digital camera set at different focal length.

Decentering distortion is due to a lack of centering of lens elements along the optical axis. The decentering distortion parameters  $P_1$  and  $P_2$  are invariably strongly projectively coupled with  $x_0$  and  $y_0$ . Decentering distortion is usually an order of magnitude or more less than radial distortion and it also varies with focus, but to a much less extent, as indicated by the decentering distortion profiles shown in fig. 6b. The projective coupling between  $P_1$  and  $P_2$  and the principal point offsets ( $Dx_0, Dy_0$ ) increases with increasing focal length and can be problematic for long focal length lenses. The extent of coupling can be diminished, during the calibration procedure, through both use of a 3D object point array and the adoption of higher convergence angles for the images.

The solution of a self-calibrating bundle adjustment leads to the estimation of all the interior parameters and APs, starting from a set of manually or automatically measured image correspondences (tie points). Critical to the quality of the self-calibration is the overall network geometry and especially the configuration camera stations. Some good hints and practical rules for camera calibration can be summarized as follows:

- acquire a set of images of a reference object, possibly constituted of coded targets which can be automatically and accurately measured in the images;

- the image network geometry should be favourable, i.e. the camera station configuration must comprise highly convergent images, acquired at different distances from the scene, with orthogonal roll angles and a large number of well distributed 3D object points;
- the accuracy of the image network (and so of the calibration procedure) increases with increasing convergence angles for the imagery, the number of rays to a given object point and the number of measured points per image (although but the incremental improvement is small beyond a few tens of points);
- a planar object point array can be employed for camera calibration if the images are acquired with orthogonal roll angles, a high degree of convergence and, desirably, varying object distances;
- orthogonal roll angles must be present to break the projective coupling between IO and EO parameters. Although it might be possible to achieve this decoupling without 90° image rotations, through provision of a strongly 3D object point array, it is always recommended to have 'rolled' images in the self-calibration network.

Nowadays self-calibration via the bundle adjustment is a fully automatic process requiring nothing more than images recorded in a suitable multi-station geometry, an initial guess of the focal length and image sensor characteristics (and it can be a guess) and some coded targets which form a 3D object point array.

### 3.2 Image orientation

In order to survey an object, a set of images needs to be acquired considering that a detail can be reconstructed in 3D if it is visible in at least 2 images. The orientation procedure is then performed to determine the position and attitude (angles) where the images were acquired. A set of tie points needs to be identified (manually or automatically) in the images, respecting the fact that the points are well distributed on the entire image format, non-coplanar nor collinear. These observations are then used to form a system of collinearity equations (Eq. 1), iteratively solved with the Gauss-Markov model of least squares (Eq. 2).

A typical set of images, acquired for 3D reconstruction purposes, forms a network which is generally not suitable for a calibration procedure. Therefore it is always better to separate the two photogrammetric steps or to adopt a set of images suitable for both procedures.

## 4. Characterization of 3D sensing devices

When a range sensor has to be chosen for geometrically surveying an object shape, independently of its size, the first point to face regards which level of detail has to be recognizable in the final 3D digital model that will be built starting from the raw 3D data, and the acceptable tolerance between the real object and its digital counterpart. These matters are so important that influence all the technological and methodological choices for the whole 3D acquisition project.

The main metrological parameters related to measurement are univocally defined by the International Vocabulary of Metrology (VIM), published by the Joint Committee for Guides in Metrology (JCGM) of ISO (JCGM, 2008). Such parameters are basically Resolution, Trueness (Accuracy) and Uncertainty (Precision).

Although the transposition of these concepts to the world of 3D imaging has been reported in the reference guide VDI/VDE 2634 by the “Association of German Engineers” for pattern projection cameras, a more general international standard on optical 3D measurement is still in preparation by commission E57 of the American Society for Testing Material (ASTM, 2006). Also the International Standard Organization (ISO) has not yet defined a metrological standard for non-contact 3D measurement devices. In its ISO-10360 only the methods for characterizing contact based Coordinate Measuring Machines (CMM) has been defined, while an extension for CMMs coupled with optical measuring machines (ISO 10360-7:2011) is still under development.

#### 4.1 Resolution

According to VIM, resolution is the “smallest change in a quantity being measured that causes a perceptible change in the corresponding indication”. This definition, once referred to non-contact 3D imaging, is intended as the minimum geometrical detail that the range device is capable to capture. This is influenced by the device mechanical, optical and electronic features. Of course such value represents the maximum resolution allowed by the 3D sensor. For its 3D nature it can be divided in two components: the axial resolution, along the optical axis of the device (usually indicated as  $z$ ), and the lateral resolution, on the  $xy$  plane (MacKinnon et al., 2008).

For digitally capturing a shape, the 3D sensor generates a discretization of its continuous surface according to a predefined sampling step adjustable by the end-user even at a level lower than the maximum. The adjustment leads to a proper spacing between geometrical samples on the  $xy$  plane, giving the actual geometrical resolution level chosen by the operator for that specific 3D acquisition action. The corresponding value in  $z$  is a consequence of the opto-geometric set-up, and can't be usually changed by the operator.

In other words it has to be made a clear distinction between the maximum resolution allowed by the sensor, usually indicated as “resolution” in the sensor data sheet, and the actual resolution used for a 3D acquisition work, that the end-user can properly set-up according to the geometrical complexity of the 3D object to be surveyed, operating on the  $xy$  sampling step.

The latter set-up is directly influenced by the lens focal length and the sensor-to-target distance for triangulation devices, using an image sensor whose size and pixel density is known in advance. In that case the sampling step will be attainable for example dividing the framed area horizontal size for the number of horizontal pixels. Since most cameras has square pixels, in general this value is equivalent to (vertical size)/(vertical number of pixels). For TOF devices the sampling can be set-up on the laser scanner control software by defining the angular step between two adjacent point on a scan line, and between two adjacent scan-lines. Of course, in order to convert the angular step in a linear step on the surface, such angle expressed in radians has to be multiplied for the operating distance. Some scanner control packages allow to set directly the former value.

The sampling should be made according to a rule deriving directly by the Nyquist-Shannon sampling theorem (Shannon, 1949), developed first in communication theory. Such theorem states that, if a sinusoidal behaviour has a frequency defined by its period  $T$ , that in the geometrical domain becomes a length (the size of the minimal geometrical detail that we intend to digitally capture), the minimal sampling step suitable for allowing the reconstruction of the same behaviour from the sampled one, is equal to  $T/2$ . Of course it is

not generally true that the fine geometrical details of a complex shape could be considered as made by the extrusions of sinusoidal profiles, but at least this criteria gives a “rule of the thumb” for estimating a minimum geometrical sampling step below which it is sure that the smaller geometrical detail will be lost.

#### 4.2 Trueness (accuracy)

VIM definition indicates accuracy in general as “closeness of agreement between a measured quantity value and a true quantity value of a measurand”. When such theoretical entity has to be evaluated for an actual instrument, including a 3D sensor, such value has to be experimentally estimated from the instrument output. For this reason VIM also define trueness as “closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value”. It is a more practical parameter that can be numerically estimated as the difference between a 3D value assumed as true (because measured with a method far more accurate), and the average of a sufficiently large number of samples acquired through the range device to be characterized. Such parameter refers therefore to the systematic component of the measurement error with respect to the real data (fig.7), and can be minimized through an appropriate sensor calibration. For 3D sensors, accuracy might be evaluated both for the axial direction (z) than for a lateral one (on the xy plane). In general, accuracy on depth is the most important, and varies from few hundredths to few tenths of a millimetre for triangulation based sensors and FM-CW laser scanners, it is in the order of 1mm for CW laser scanners, and in the order of 5 mm for PW laser scanners.

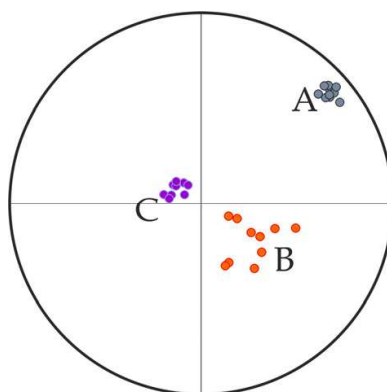


Fig. 7. Exemplification of the accuracy and precision concepts. The target has been used by three different shooters. The shooter A is precise but not accurate, B is more accurate than A but less precise (more spreading), C is both accurate and precise.

#### 4.3 Uncertainty (precision)

Precision is the “closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions” (JCGM, 2008). A practical value for estimating such agreement is to calculate the dispersion of the quantity values being attributed to a measurand through the standard deviation of the measured values respect to their average (or a multiple of it), defined by VIM as uncertainty (fig.7).

As accuracy is influenced by systematic errors, precision is mostly influenced by random errors, leading to a certain level of unpredictability of the measured value, due to thermal



noise in the sensor's detector, and, in case of laser based devices, by the typical laser speckle effect (Baribeau & Rioux 1991).

For a 3D sensor such estimation can be done acquiring several times the same area and analysing the measured value of a specific point in space as a random variable, calculating its standard deviation. This would involve a very large number of 3D acquisitions to be repeated, namely from 10000 to one million, in order to consider the data statistically significant. For this reason a more practical approach (even if not as theoretically coherent with the definition) is to acquire the range map of a target whose shape is known in advance, like for example a plane, and evaluate the standard deviation of each 3D point respect to the ideal shape (Guidi et al., 2010). Since a range map can be easily made by millions of points the statistical significance is implicit.

Precision of active 3D devices ranges from a few tens of micrometres for triangulation based sensors, with an increase of deviation with the square of sensor-to-target distance. It has similar values for FM-CW laser scanners with a much less significant change with distance. For CW laser scanners it has values starting from below 1mm up to a few mm as the sensor is farer from the target, and not less of 2 mm for PW laser scanners (Boehler et al., 2003) with no significant change with distance (Guidi et al., 2011).

For modelling applications the uncertainty level of the range sensor should not exceed a fraction of the resolution step for avoiding topological anomalies in the final mesh (Guidi & Bianchini, 2007). A good rule of the thumb is to avoid a resolution level smaller than the range device measurement uncertainty.

## 5. Photogrammetric 3D point clouds generations

Once the camera parameters are known, the scene measurements can be performed with manual or automated procedures. The measured 2D image correspondences are converted into unique 3D object coordinates (3D point cloud) using the collinearity principle and the known exterior and interior parameters previously recovered. According to the surveyed scene and project requirements, sparse or dense point clouds are derived (fig. 8).

Manual (interactive) measurements, performed in monocular or stereoscopic mode, derive sparse point clouds necessary to determine the main 3D geometries and discontinuities of an object. Sparse reconstructions are adequate for architectural or 3D city modelling applications, where the main corners and edges must be identified to reconstruct the 3D shapes (fig. 8a) (Gruen & X. Wang, 1998; El-Hakim, 2002). A relative accuracy in the range 1:5,000-15,000 is generally expected for such kinds of 3D models.

On the other hand, automated procedures ("image matching") are employed when dense surface measurements and reconstructions are required, e.g. to derive a Digital Surface Model (DSM) to document detailed and complex objects like reliefs, statues, excavations areas, etc. (fig. 8b). The latest development in automated image matching (Pierrot-Deseilligny & Paparoditis, 2006; Hirschmuller, 2008; Remondino et al., 2008; Hiep et al., 2009; Furukawa & Ponce, 2010) are demonstrating the great potentiality of the image-based 3D reconstruction method at different scales of work, comparable to point clouds derived using active range sensors and with a reasonable level of automation. Overviews on stereo and multi-image image matching techniques can be found in (Scharstein & Szeliski, 2002; Seitz et al., 2006). Recently some commercial, open-source and web-based tools were released to derive dense point clouds from a set of images (Photomodeler Scanner, MicMac, PMVS, etc.).



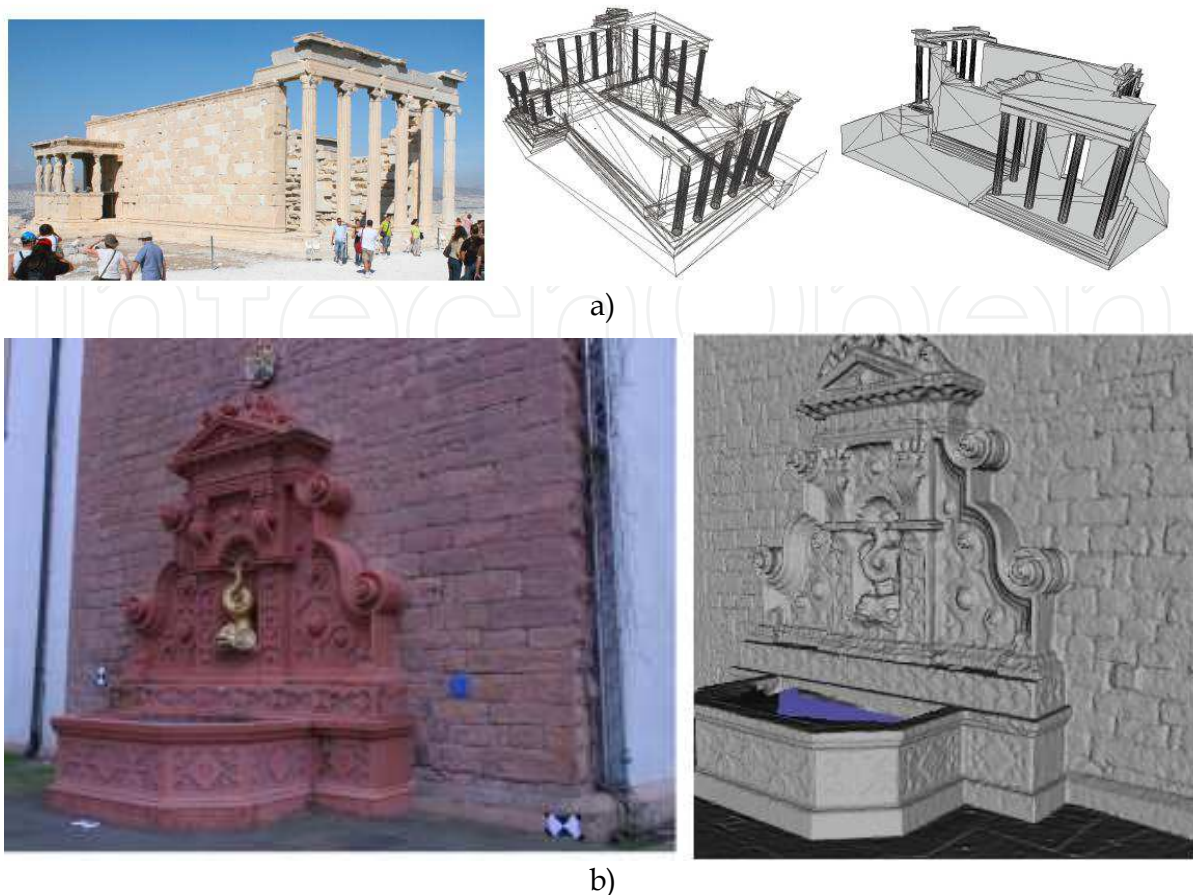


Fig. 8. 3D reconstruction of architectural structures with manual measurements in order to generate geometrical models with the main geometrical features (a). Dense point clouds produced using automated image matching (b).

## 6. Acquisition and processing of 3D point clouds with active sensors

Independently of the active 3D technology used, a range map is a metric representation of an object from a specific point of view through a set of 3D points properly spaced apart, according to the complexity of the imaged surface.

In order to create a model, several views have to be taken for covering the whole object surface. This operation leads to a set of measured points that can be used as nodes of a mesh representing a 3D digital approximation of the real object. Hence, for going from the raw data to the final 3D model, a specific process has to be followed (Bernardini & Rushmeyer, 2002; Vrabel et al., 2009), according to the steps described in the next sections. Many of these steps have been implemented in 3D point cloud processing packages, both open source, like Meshlab (ISTI-CNR, Italy), Scanalize (Stanford University, USA), and commercial, as Polyworks (Innovmetric, Canada), RapidForm (Inus Technology, South Korea), Geomagic Studio (Geomagic, USA), Cyclone (Leica, Switzerland), 3D Reshaper (Technodigit, France).

### 6.1 Project planning

The final scope of the digital model is the first matter to be considered for properly planning a 3D acquisition project. Applications of 3D models may span from a simple

support for multimedia presentations to a sophisticated dimensional monitoring. In the former case a visually convincing virtual representation of the object is enough, while in the latter a strict metric correspondence between the real object and its digital representation is absolutely mandatory. Since parameters as global model accuracy and geometrical resolution have a considerable cost in terms of acquired data and post-processing overhead, a choice coherent with the project budget and final purpose, is a must. Once such aspects have been clearly identified, the object to be acquired has to be analyzed in terms of size, material and shape.

## 6.2 Acquisition of individual point clouds

Once the planning has been properly examined, the final acquisition is rather straightforward. In addition to basic logistics, possible issues may be related with sensor positioning and environmental lighting. Camera positioning for small objects can be solved either by moving the object or the sensor, but when the object is very large and heavy (e.g. a boat), or fixed into the ground (e.g. a building), the only possibility is obviously to move the range sensor. In that case a proper positioning should be arranged through scaffoldings or mobile platforms, and the related logistics should be organized. Another aspect that might influence a 3D acquisition is the need of working in open air rather than in a laboratory where lighting conditions can be controlled. In the former case it has to be considered that TOF laser scanners are designed for working on the field and are therefore not much influenced by direct sunlight. Triangulation based range sensors employ much less light power per surface unit and for this reason give worst or no results with high environmental light. In this case a possible but logistically costly solution is to prepare a set with tents or shields for limiting the external light on the surface to be acquired. However in that conditions a more practical approach for obtaining the same high resolution is dense image matching, that, being a passive technique, works well with strong environmental lighting (Guidi et al., 2009a).

## 6.3 Point clouds alignment

In general each range map acquired from a specific position is given in a coordinate system with the origin located into the range sensor.

Taking range data of a scene or object from different points of view means gathering 3D data representing the same geometry by different reference systems whose mutual orientation is generally unknown. For such reason it is necessary to align all 3D data into the same coordinate system. The process can be achieved in three different ways.

### 6.3.1 Complementary equipment

This approach requires the measurement of the range device position and orientation with a complementary 3D measurement device like a CMM, giving such data in its coordinate system which is assumed as the global reference. These 6 pieces of information (position and orientation) can be used for calculating the roto-translation matrix from the range device coordinate system to the global one. Applying systematically such roto-translation to any 3D point measured by the range device allows to find immediately its representation in the global reference system even for different device-to-target orientations. Although the working volume is limited by the CMM positioning range, such approach is very accurate. This is why it is used in equipment typically employed in high-accuracy industrial

applications with articulated arms (contact CMM) or laser trackers (non-contact CMM) coupled with triangulation based scanning heads (Pierce, 2007; Peggs et al., 2009).

In case of long-range active range sensors (e.g. TOF laser scanners) the complementary device can be represented by a GNSS which is used, for every acquisition, to measure the position of the range sensor in a global reference system.

### 6.3.2 Reference targets

Measuring some reference points on the scene with a surveying system like for example a total station, allows to define a global reference system in which such targets are represented. During the 3D acquisition campaign the operator captures scenes containing at least three targets which are therefore represented in the range device reference system for that particular position. Being their positions known also in a global reference system, their coordinates can be used to compute the roto-translation matrix for re-orienting the point cloud from its original reference system to the global one. The operation is of course repeated up to the alignment of all 3D data of the scene. This approach is used more frequently with TOF laser scanners thanks to their large region of interest.

### 6.3.3 Iterative Closest Point (ICP)

Using as references natural 3D features in the scene is a possible alternative somehow similar to the previous one. The only difference is that no special target has to be fixed on the scene and individually measured by the operator. On the other hand for allowing a proper alignment, a considerable level of overlapping between adjacent range maps has to be arranged, resulting in a large data redundancy and long computational time.

The algorithm for aligning this kind of 3D data sets involves the choice of a range map whose coordinate system is used as global reference. A second data set, partially overlapping with the reference one, is manually or automatically pre-aligned to the main one choosing at least three corresponding points on the common area of both range maps (fig. 9a). This step allows to start an iterative process for minimizing the average distance between the two datasets, initiated by a situation of approximate alignment (fig. 9b) not too far from the optimized one (fig. 9c), that can be reached after a number of iterations as large as the initial approximation is rough. For this reason this class of algorithms is called “Iterative Closest Point” (ICP).

The most critical aspect is that the range maps to be aligned represent different samplings of the same surface, therefore there is not exact correspondence between 3D points in the two coordinate systems. Several solutions have been proposed by considering the minimization of Euclidean distances between points as much corresponding as possible, but it is highly time consuming due to the exhaustive search for the nearest point (Besl & McKay, 1992), or between a point and a planar approximation of the surface at the corresponding point on the other range map (Chen & Medioni, 1992). In both cases the algorithm core is a nonlinear minimization process, being based on a nonlinear feature such as a distance. For this reason the associated cost function has a behaviour characterized by several confusing local minima, and its minimization needs to be started by a pre-alignment close enough to the final solution in order to converge to the absolute minimum.

Once the first two range maps of a set are aligned, ICP can be applied to other adjacent point clouds up to the full coverage of the surface of interest. This progressive pair-wise alignment may lead to a considerable error propagation, clearly noticeable on closed surfaces when the

first range map has to be connected with the last one. For this reason global versions of ICP have been conceived, where the orientation of each range map is optimized respect to all neighbour range maps (Gagnon et al., 1994).

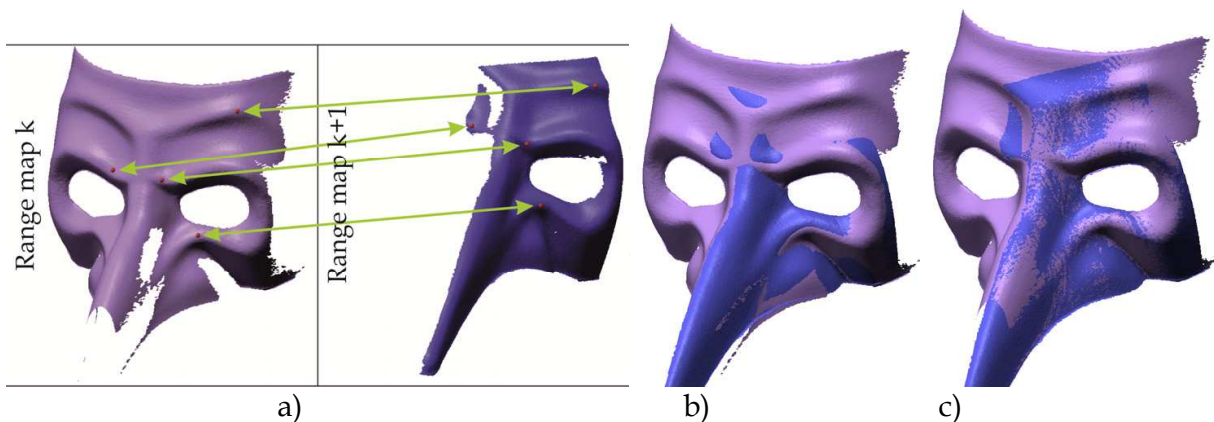


Fig. 9. ICP alignment process: a) selection of corresponding points on two partially superimposed range maps; b) rough pre-alignment; c) accurate alignment after a few iterations.

Several refinements of the ICP approach have been developed in the last two decades for pair-wise alignment (Rusinkiewicz & Levoy, 2001), with the introduction of additional non-geometrical parameters as colour, for solving alignment of object with rich image content but poor 3D structure like flat or regular texturized surfaces (Godin et al., 2001b), and for managing possible shape changes between different shots due to non-rigid objects (Brown & Rusinkiewicz, 2007). A quantitative test of different alignment algorithm has been recently proposed in term of metric performances and processing time (Salvi et al., 2007). For a widespread updated state of the art about alignment algorithms see (Deng, 2011).

## 7. Polygonal model generation

Once a point cloud from image matching or a set of aligned point clouds acquired with an active sensor are obtained, a polygonal model ("mesh") is generally produced. This process is logically subdivided in several sub-steps that can be completed in different orders depending by the 3D data source (Berger et al., 2011).

### 7.1 Mesh generation for structured point clouds

The regular matrix arrangement of a structured point cloud involves an immediate knowledge of the neighbour potential mesh connection for each 3D point, making the mesh generation a rather straightforward procedure. This means that once a set of range maps is aligned, it can be easily meshed before starting the final merge.

This is what is done for example by the Polyworks software package used to create the alignment and meshing shown in fig. 10. For carrying out the following merge, the meshes associated to the various range maps have to be connected with the neighbour meshes. This can be achieved with two different approaches: (i) the so-called *zippering* method (Turk & Levoy, 1994) which selects polygons in the overlapping areas, removes redundant triangles and connects meshes together (zipper) trying to maintain the best possible topology. An



optimized version that uses Venn diagrams for evaluating the level of redundancy on mesh overlaps has been proposed (Soucy & Laurendeau, 1995). Other approaches work by triangulating union of the point sets, like the Ball Pivoting algorithm (Bernardini et al., 1999), which consists of rolling an imaginary ball on the point sets and creating a triangle for each triplet of points supporting the ball. All methods based on a choice of triangles from a certain mesh on the overlapping areas may get critical in case of large number of overlapped range maps; (ii) a *volumetric* algorithm which operates a subdivision in voxels of the model space, calculates an average position of each 3D point on the overlapping areas and re-samples meshes along common lines of sight (Curless & Levoy, 1996). In this case areas with possible large number of overlapped range maps are evaluated more efficiently than with the zippering method, with a reduction of measurement uncertainty by averaging corresponding points.

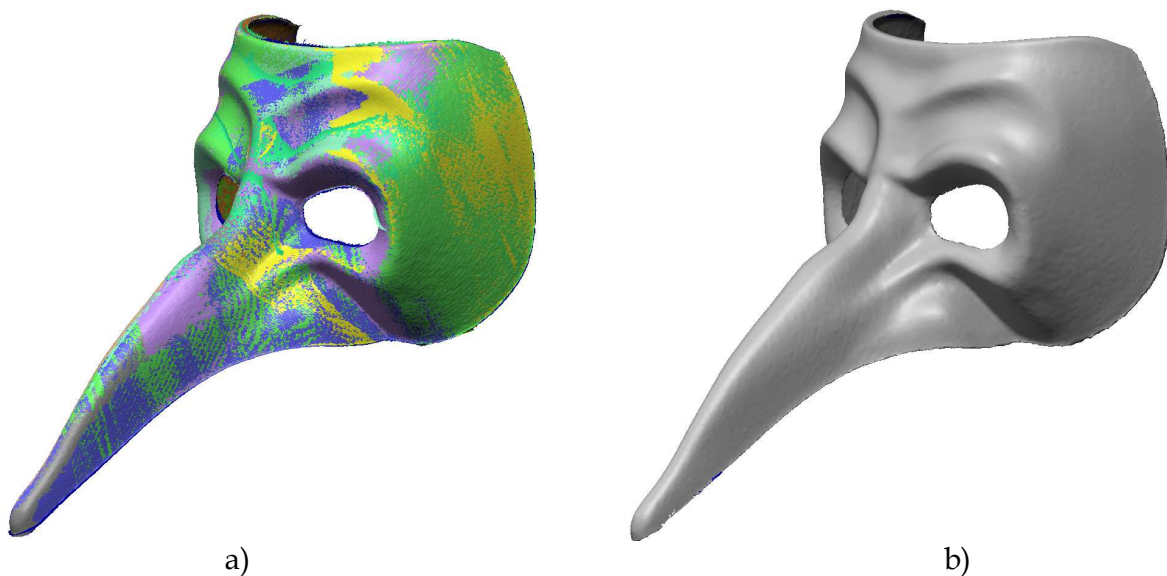


Fig. 10. Mesh generation: a) set of ICP aligned range maps. Different colours indicate the individual range maps; b) merge of all range maps in a single polygonal mesh.

## 7.2 Mesh generation for unstructured point clouds

While meshing is a pretty straightforward step for structured point clouds, for an unstructured point cloud it is not so immediate. It requires a specific process like Delaunay, involving a projection of the 3D points on a plane or another primitive surface, a search of the shorter point-to-point connection with the generation of a set of potential triangles that are then re-projected in the 3D space and topologically verified. For this reason the mesh generation from unstructured clouds may consist in: a) merging the 2.5D point clouds reducing the amount of data in the overlapped areas and generating in this way a uniform resolution full 3D cloud; b) meshing with a more sophisticated procedures of a simple Delaunay. The possible approaches for this latter step are based on: (i) *interpolating surfaces* that build a triangulation with more elements than needed and then prune away triangles not coherent with the surface (Amenta & Bern, 1999); (ii) *approximating surfaces* where the output is often a triangulation of a best-fit function of the raw 3D points (Hoppe et al., 1992; Cazals & Giesen, 2006).

Dense image matching generally consist of unstructured 3D point clouds that can be processed with the same approach used for the above mentioned laser scanner unstructured point clouds. No alignment phase is needed as the photogrammetric process deliver a unique point cloud of the surveyed scene.

### 7.3 Mesh editing and optimization

Mesh editing allows to correct all possible topological incoherence generated after the polygonal surface generation. Generally some manual intervention of the operator is required in order to clean spikes and unwanted features and to reconstruct those parts of the mesh that are lacking due to previous processing stages or to an effective absence of 3D data collected by the sensor.

These actions are needed at least for two purposes: (i) if the final 3D model has to be used for real-time virtual presentations or static renderings, the lacking of even few polygons gives no support to texture or material shading, creating a very bad visual impression and thwarting the huge modelling effort made until this stage; (ii) if the model has to be used for generating physical copies through rapid prototyping, the mesh has to be watertight.

Several approaches have been proposed for creating lacking final mesh as much agreement as possible with the measured object, like radial basis functions (Carr et al., 2001), multi-level partition of unity implicits (Ohtake et al., 2003) or volumetric diffusion (Davis et al., 2002; Sagawa & Ikeuchi, 2008).

In some cases, like for example dimensional monitoring applications, mesh editing is not suggested for the risk of adding not existing data to the measured model, leading to possible inconsistent output.

Optimization is instead a final useful step in any applicative case, where a significant reduction of the mesh size can be obtained. After the mesh generation and editing stages, the polygonal surface has a point density generally defined by the geometrical resolution set by the operator during the 3D data acquisition or image matching procedure. In case of active range sensing as specified in sect. 4.1, the resolution is chosen for capturing the smaller geometrical details and can be therefore redundant for most of the model. A selective simplification of the model can thus reduce the number of polygons without changing significantly its geometry (Hoppe, 1996). As shown in fig. 11a, the point density

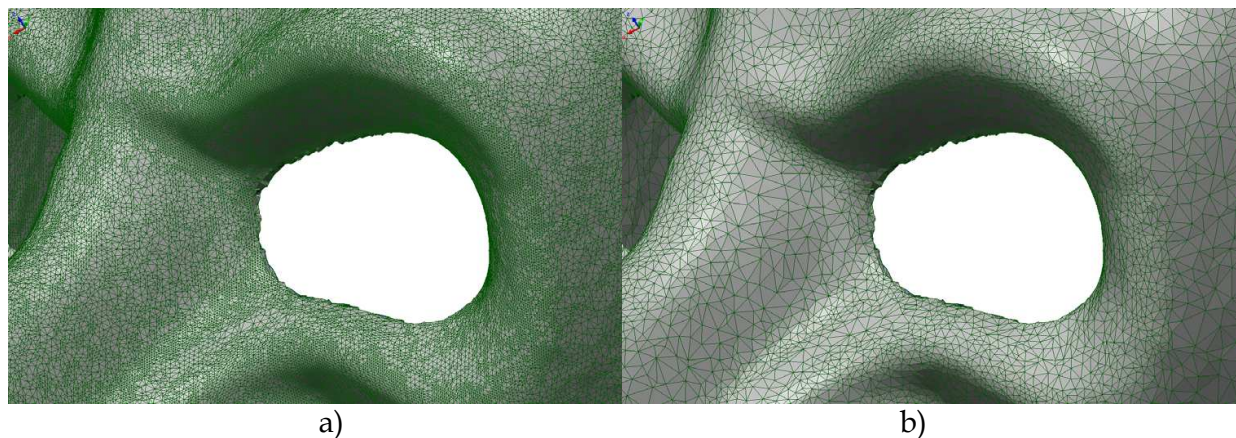


Fig. 11. Mesh optimization: a) mesh with polygon sizes given by the range sensor resolution set-up (520,000 triangles); b) mesh simplified in order to keep the difference with the unsimplified one, below 50mm. The polygon sizes vary dynamically according to the surface curvature and the mesh size drops down to 90,000 triangles.



set for the device appears to be redundant for all those surfaces whose curvature radius is not too small.

A mesh simplification that progressively reduces the number of polygons eliminating some nodes, can be applied up to reaching a pre-defined number of polygons (useful for example in game applications where such limitation holds), or, as an alternative, checking the deviation between simplified and un-simplified mesh and stopping at a pre-assigned threshold. If such threshold is chosen in the order of the 3D sensor measurement uncertainty, this kind of simplification does not practically influence the geometric information attainable by the model (fig. 11b), with a strong data shrinking (nearly six time in the example). Mesh simplification algorithms have been extensively examined and compared by Cignoni et al. (1998).

## 8. Texture mapping and visualization

A polygonal 3D model can be visualized in wireframe, shaded or textured mode. A textured 3D geometric model is probably the most desirable 3D object documentation by most since it gives, at the same time, a full geometric and appearance representation and allows unrestricted interactive visualization and manipulation at a variety of lighting conditions. The photo-realistic representation of a polygonal model (or even a point cloud) is achieved mapping a colour images onto the 3D geometric data. The 3D data can be in form of points or triangles (mesh), according to the applications and requirements. The texturing of 3D point clouds (point-based rendering techniques (Kobbelt & Botsch, 2004) allows a faster visualization, but for detailed and complex 3D models it is not an appropriate method. In case of meshed data the texture is automatically mapped if the camera parameters are known (e.g. if it is a photogrammetric model and the images are oriented) otherwise an interactive procedure is required (e.g. if the model has been generated using range sensors and the texture comes from a separate imaging sensor). Indeed homologue points between the 3D mesh and the 2D image to-be-mapped should be identified in order to find the alignment transformation necessary to map the colour information onto the mesh. Although some automated approaches were proposed in the research community (Lensch et al., 2000; Corsini et al., 2009), no automated commercial solution is available and this is a bottleneck of the entire 3D modelling pipeline. Thus, in practical cases, the 2D-3D alignment is done with the well-known DLT approach (Abdel-Aziz & Karara, 1971), often referred as Tsai method (Tsai, 1986). Corresponding points between the 3D geometry and a 2D image to-be-mapped are sought to retrieve the interior and exterior unknown camera parameters. The colour information is then projected (or assigned) to the surface polygons using a colour-vertex encoding, a mesh parameterization or an external texture.

In Computer Graphics applications, the texturing can also be performed with techniques able to graphically modify the derived 3D geometry (displacement mapping) or simulating the surface irregularities without touching the geometry (bump mapping, normal mapping, parallax mapping).

In the texture mapping phase some problems can arise due to lighting variations of the images, surface specularity and camera settings. Often the images are exposed with the illumination at imaging time but it may need to be replaced by illumination consistent with the rendering point of view and the reflectance properties (BRDF) of the object (Lensch et al., 2003). High dynamic range (HDR) images might also be acquired to recover all scene details

and illumination (Reinhard et al., 2005) while colour discontinuities and aliasing effects must be removed (Debevec et al., 2004; Umeda et al., 2005).

The photo-realistic 3D product needs finally to be visualized e.g. for communication and presentation purposes. In case of large and complex model the point-based rendering technique does not give satisfactory results and does not provide realistic visualization. The visualization of a 3D model is often the only product of interest for the external world, remaining the only possible contact with the 3D data. Therefore a realistic and accurate visualization is often required. Furthermore the ability to easily interact with a huge 3D model is a continuing and increasing problem. Indeed model sizes (both in geometry and texture) are increasing at faster rate than computer hardware advances and this limits the possibilities of interactive and real-time visualization of the 3D results. Due to the generally large amount of data and its complexity, the rendering of large 3D models is done with multi-resolution approaches displaying the large meshes with different Levels of Detail (LOD), simplification and optimization approaches (Dietrich et al., 2007).

## 9. Conclusions

This chapter reported an overview of the actual optical 3D measurements sensors and techniques used for terrestrial 3D modelling. The last 15 years of applications made clear that reality-based 3D models are very useful in many fields but the related processing pipeline is still far from being optimal, with possible improvements and open research issues in many steps.

First of all automation in 3D data processing is one of the most important issues influencing efficiency, time and production costs. At present different research solution and commercial packages have turned towards semi-automated (interactive) approaches, where the human capacity in data interpretation is paired with the speed and precision of computer algorithms. Indeed the success of fully automation in image understanding or 3D point clouds processing depends on many factors and is still a hot topic of research. The progress is promising but the acceptance of fully automated procedures, judged in terms of handled datasets and accuracy of the final 3D results, depends on the quality specifications of the user and final use of the produced 3D model. A good level of automation would make also possible the development of new tools for non-expert users. These would be particularly useful since 3D capturing and modelling has been demonstrated to be an interdisciplinary task where non-technical end-users (archaeologists, architects, designers, art historians, etc.), may need to interact with sophisticated technologies through clear protocols and user-friendly packages.

Sensor fusion has been experimentally demonstrated to be useful for collecting as many features as possible, allowing the exploitation of each range sensing technology capability. Currently available packages allows the creation of different geometric levels of detail (LoD) at model level (i.e. at the end of the modelling pipeline), while this could be performed also at data-level with the development of novel packages capable to deal simultaneously with different sensors and data. Such novel feature should allow also to include new sensors and 3D data in the processing pipeline taking into account their metrological characteristics.

For this reason also the adoption of standards for comparing 3D sensing technologies would help. At present even no common terminology exists for comparing sensors performances.

A smooth connection between a data base and reality-based 3D models is another issue that has to be faced when the model becomes a “portal” for accessing to an informative system associated to the modelled object. Although some experimental systems have been developed, no simple tools suitable for non-expert users are available yet.

The latter open issue is connected with the problem of remotely visualize large 3D models, both for navigation and data access. Despite 3D navigation through the internet has been attempted both with local rendering of downloaded 3D models (possible large initial time lag and poor data security), or with remote rendering and streaming to the client of a sequence of rendered frames (good security but poor real-time navigation), a complete and reliable user oriented solution is still lacking.

## 10. Acknowledgment

The authors would like to thank J. Angelo Beraldin for many useful discussions.

## 11. References

- Abdel-Aziz Y.I. & Karara, H.M. (1971). Direct linear trans-formation from comparator coordinates into object space coordinates in close-range photogrammetry, *Proc. of the Symposium on Close-Range Photogrammetry*, Falls Church (VA) USA, pp. 1-18.
- Agarwal, S.; Snavely, N.; Simon, I.; Seitz, S.; Szelinski, R. (2009). Building Rome in a Day, *Proceedings of the IEEE 12th International. Conference on Computer Vision*, pp. 72-79.
- Akca, D. & Gruen, A. (2009). Comparative geometric and radiometric evaluation of mobile phone and still video cameras. *The Photogrammetric Record*, Vol. 24, pp. 217-245.
- Amenta, N. & Bern, M. (1999). Surface reconstruction by Voronoi filtering, *Discrete and Computational Geometry* Vol. 22, No. 4, pp. 481-504.
- ASTM International (2006). Committee E57 on 3D Imaging Systems; West Conshohocken, PA, USA.
- Barazzetti, L.; Fangi, G.; Remondino, F.; Scaioni, M. (2010). Automation in Multi-Image Spherical Photogrammetry for 3D Architectural Reconstructions, *Proceedings of VAST 2010*, Paris, France, pp. 75-81.
- Baribeau, R. & Rioux, M. (1991), Influence of speckle on laser range finders, *Applied Optics*, Vol. 30, No. 20, pp. 2873-2878.
- Beraldin, J.-A., (2004). Integration of laser scanning and close-range photogrammetry the last decade and beyond, *Proc. IAPRS*, Vol.35, No.5, pp. 972-983.
- Berger, M.; Levine, J.A.; Nonato, L.G.; Taubin, G.; Silva, C.T. (2011). An End-to-End Framework for Evaluating Surface Reconstruction, *SCI Technical Report*, No. UUSCI-2011-001, SCI Institute, University of Utah. Available from: [www.sci.utah.edu/publications/SCITechReports/UUSCI-2011-001.pdf](http://www.sci.utah.edu/publications/SCITechReports/UUSCI-2011-001.pdf)
- Bernardini, F. & Rushmeier, H. (2002). The 3D Model Acquisition Pipeline, *Computer Graphics Forum*, NCC Blackwell, Vol. 21(2), pp. 149-172.
- Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C. & Taubin, G. (1999). The ball-pivoting algorithm for surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 5, pp. 349-359.

- Bernardini, F.; Rushmeier, H.; Martin, I.M.; Mittleman, J.; Taubin, G. (2002). Building a digital model of Michelangelo's Florentine Pieta, *IEEE Computer Graphics Application*, vol. 22, pp. 59-67.
- Besl, P.J. & McKay, N. (1992). A Method for Registration of 3-D Shapes, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 14-2, pp. 239-256.
- Blais F. (2004). Review of 20 Years of Range Sensor Development, *Journal of Electronic Imaging*, Vol 13-1, pp. 231-240.
- Boehler, W.; Bordas V.M. & Marbs, A. (2003). Investigating laser scanner accuracy. *Proceedings of the XIXth CIPA Symposium*, pp. 696-702.
- Brown, B. & Rusinkiewicz, S. (2007). Global non-rigid alignment of 3-D scans. *ACM Transactions on Graphics*, Vol. 26, No. 3, Article 21.
- Brown, D.C. (1971). Close-range camera calibration. *PE&RS*, Vol. 37(8), pp. 855-866
- Callieri, M.; Chica, A.; Dellepiane, M.; Besora, I.; Corsini, M.; Moyé S, J.; Ranzuglia, G.; Scopigno, R.; Brunet, P. (2011). Multiscale acquisition and presentation of very large artifacts: The case of Portalada, *Journal of Computing and Cultural Heritage*, Vol.3, No.4, Article number 14.
- Carr, J. C.; Beatson, R. K.; Cherrie, J. B.; Mitchell, T. J.; Fright, W. R.; McCallum, B. C. & Evans, T. R. (2001). Reconstruction and representation of 3D objects with radial basis functions, *Proc. SIGGRAPH*, pages 67-76.
- Cazals, F. & Giesen, J. (2006). Delaunay triangulation based surface reconstruction. In: *Effective Computational Geometry for Curves and Surfaces*, J.-D. Boissonnat and M. Teillaud, Eds. Springer-Verlag, Mathematics and Visualization, pp. 231-276.
- Chen, Y. & Medioni, G. (1992). Object modeling by registration of multiple range images, *Image and Vision Computing*, Vol 10-3, pp. 145-155.
- Cignoni, P., Montani, C. & Scopigno, R. (1998). A comparison of mesh simplification algorithms. *Computers & Graphics* Vol. 22, No. 1, pp. 37-54.
- Corsini, M.; Dellepiane, M.; Ponchio, F. & Scopigno, R. (2009). Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties. *Computer Graphics Forum*, Vol. 28(7), pp. 1755-1764.
- Criminisi A.; Reid, I. & Zisserman A. (1999). Single view metrology. *Proc. International Conference on Computer Vision*, pp. 434-442.
- Curless, B. & Levoy M. (1996). A Volumetric Method for Building Complex Models from Range Images, *Proc. SIGGRAPH96*, pp. 303-312.
- Davis, J.; Marschner, S. R.; Garr, M. & Levoy, M. (2002). Filling holes in complex surfaces using volumetric diffusion. *Proc. 3DPVT*, pp. 428-438.
- Debevec, P. ; C. Tchou, A. Gardner, T. Hawkins, C. Poullis, J. Stumpfel, A. Jones, N. Yun, P. Einarsson, T. Lundgren, M. Fajardo, and P. Martinez. (2004). Estimating surface reflectance properties of a complex scene under captured natural illumination. *USC ICT Technical Report ICT-TR-06*.
- Deng, F. (2011). Registration between Multiple Laser Scanner Data Sets, *Laser Scanning, Theory and Applications*, Chau-Chang Wang (Ed.), ISBN: 978-953-307-205-0, InTech.
- Dietrich, A.; Gobbetti, A. & Yoon S.-E. (2007). Massive-model rendering techniques: a tutorial. *Computer graphics and applications*, Vol. 27, No. 6, pp. 20-34.



- Eisenbeiss, H. (2009). *UAV Photogrammetry*. Ph.D. Thesis, Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland, p. 235.
- El-Hakim. S. (2000). A practical approach to creating precise and detailed 3D models from single and multiple views. *Int. Arch. Photog. & Rem. Sens.*, Vol. 33(B5), pp. 122-129.
- El-Hakim. S. (2002). Semi-automatic 3D reconstruction of occluded and unmarked surfaces from widely separated views. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 34(B5), pp. 143-148.
- Fangi, G. (2007). The Multi-Image Spherical Panoramas as a Tool for Architectural Survey, *Proceedings of the XXI International CIPA Symposium*, Vol. 36, Part 5/C53.
- Fraser, C.S., 1996: Network design. In: *Close-range Photogrammetry and Machine Vision*, Atkinson (Ed.), Whittles Publishing, UK, pp.256-282.
- Furukawa, Y. & Ponce, J. (2010). Accurate, dense and robust multiview stereopsis, *IEEE Trans. Patt. Anal. Mach. Intell.* Vol. 32, pp. 1362-1376.
- Gagnon, H.; Soucy, M.; Bergevin, R. & Laurendeau, D. (1994). Registration of multiple range views for automatic 3D model building, *Proc. of IEEE Conf. on CVPR*, pp. 581-586.
- Godin, G.; Beraldin, J.-A.; Rioux, M.; Levoy, M. & Cournoyer, L. (2001a). An Assessment of Laser Range Measurement of Marble Surfaces, *Proc. of the 5th Conference on Optical 3-D Measurement Techniques*, Vienna, Austria, pp. 49-56.
- Godin, G.; Laurendeau, D.; Bergevin, R. (2001b). A Method for the Registration of Attributed Range Images. *Proc. of Third Int. Conference on 3D Digital Imaging and Modeling (3DIM2001)*, pp. 179-186.
- Godin, G; Borgeat, L.; Beraldin, J.-A.; Blais, F. (2010). Issues in acquiring, processing and visualizing large and detailed 3D models, *Proceeding of the 44th Annual Conference on Information Sciences and Systems (CISS 2010)*, pp. 1-6.
- Gruen A. & Wang X. (1998). Cc-modeler: a topology generator for 3d city models, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 53, pp. 286-295.
- Gruen, A. & Huang, T.S. (2001). *Calibration and Orientation of Cameras in Computer Vision*, Springer: Berlin/Heidelberg, Germany, p. 239.
- Gruen, A.; Remondino, F.; Zhang, L. (2004). Photogrammetric reconstruction of the Great Buddha of Bamiyan. *The Photogrammetric Record*, Vol. 19, pp. 177-199.
- Guidi, G.; Beraldin, J.-A.; Ciofi, S. & Atzeni, C. (2003). Fusion of range camera and photogrammetry: a systematic procedure for improving 3D models metric accuracy, *IEEE Trans. on Systems Man and Cybernetics Part B*, Vol. 33-4, pp. 667- 676.
- Guidi, G.; Beraldin, J.-A. & Atzeni, C. (2004). High accuracy 3D modeling of Cultural Heritage: the digitizing of Donatello's "Maddalena", *IEEE Transactions on Image Processing*, Vol. 13-3, pp. 370-380.
- Guidi, G.; Frischer, B.; Russo, M.; Spinetti, A.; Carosso, L. & Micoli, L. L. (2005). Three-dimensional acquisition of large and detailed cultural heritage objects, *Machine Vision and Applications*, Special issue on 3D acquisition technology for cultural heritage, Vol. 17-6, pp. 347-426.
- Guidi, G.; Remondino, F.; Morlando, G.; Del Mastio, A.; Uccheddu, F.; Pelagotti, A. (2007). Performance evaluation of a low cost active sensor for cultural heritage documentation, *VIII Conference on Optical 3D Measurement Techniques*, Ed. Gruen/Kahmen, Vol.2, Zurich, Switzerland, pp. 59-69.

- Guidi, G. & Bianchini, C. (2007). TOF laser scanner characterization for low-range applications. *Proceedings of the Videometrics IX – SPIE Electronic Imaging*, San Jose, CA, USA, 29–30 January 2007; SPIE: Bellingham WA, USA; pp. 649109.1–649109.11.
- Guidi, G.; Remondino, F.; Russo, M.; Menna, F.; Rizzi, A. & Ercoli, S. (2009a). A multi-resolution methodology for the 3D modeling of large and complex archaeological areas, *International Journal of Architectural Computing*, Vol. 7-1, pp. 40–55.
- Guidi, G.; Spinetti, A.; Carosso, L. & Atzeni, C. (2009b). Digital three-dimensional modelling of Donatello's David by frequency modulated laser radar, *Studies in Conservation*, Vol. 54, pp. 3–11.
- Guidi, G.; Russo, M.; Magrassi, G. & Bordegoni, M. (2010). Performance Evaluation of Triangulation Based Range Sensors, *Sensors*, Vol. 10-8, pp. 7192–7215.
- Guidi, G.; Russo, M.; Magrassi, G.; Bordegoni, M. (2011). Low cost characterization of TOF range sensors resolution, *Proc. IS&T/SPIE Electronic Imaging*, Vol. 7864, pp. D0–D10.
- Healey G. & Binford, T.O. (1987). Local shape from specularity. *Proc. ICCV*, London, UK.
- Hiep, V.H.; Keriven, R.; Labatut, P.; Pons, J.-P. (2009). Towards high-resolution large-scale multi-view stereo, *Proceedings of IEEE Conference on CVPR*, pp. 1430–1437.
- Hirschmuller, H. (2008). Stereo processing by semi-global matching and mutual information. *IEEE Trans. Patt. Anal. Mach. Intell.*, 30, 328–341
- Hoppe, H. (1996). Progressive meshes, *SIGGRAPH96*, pages 99–108.
- Hoppe, H.; Deroose, T.; Duchamp, T.; McDonald, J. & Stuetzle, W. (1992). Surface reconstruction from unorganized points, *SIGGRAPH92*, Vol. 26, No. 2, pp. 71–78.
- Horn B.K.P. & Brooks. M.J. (1989). *Shape from Shading*. MIT Press, Cambridge.
- JCGM (2008). International vocabulary of metrology – Basic and general concepts and associated terms (VIM), Bureau International des Poids et Mesures (BIPM), France.
- Kender. J.R. (1978). Shape from texture. *Proc. DARPA IU Workshop*.
- Kobbelt L. & M. Botsch. (2004). A survey of point-based techniques in computer graphics. *Computers and Graphics*, Vol. 28, No. 6, pp. 801–814.
- Lensch, H.; W. Heidrich, and H. Seidel. (2000). Automated texture registration and stitching for real world models. *Proc. 8th Pacific Graph. Conf. CG and Application*, pp. 317–327.
- Lensch, H.P.A.; J. Kautz, M. Goesele, W. (2003). Heidrich, and H.-P. Seidel. Image-based reconstruction of spatial appearance and geometric detail, *ACM Transaction on Graphics*, Vol. 22, No. 2, pp. 234–257.
- Levoy, M.; Pulli, K.; Curless, B.; Rusinkiewicz, S.; Koller, D.; Pereira, L.; Ginzton, M.; Anderson, S.; Davis, J.; Ginsberg, J.; Shade, J.; Fulk, D. (2000). The Digital Michelangelo Project: 3D Scanning of Large Statues, *SIGGRAPH00*, pp. 131–144.
- Li, J.; Guo, Y.; Zhu, J.; Lin X.; Xin Y.; Duan K. & Tang Q. (2007). Large depth of view portable three dimensional laser scanner and its segmental calibration for robot vision, *Optics and Laser in Engineering*, Vol. 45, pp. 1077–1087.
- Luhmann, T.; Robson, S.; Kyle, S.; Hartley, I. (2007). *Close Range Photogrammetry: Principles, Techniques and Applications*, Whittles: Dunbeath, UK, p. 528
- Luhmann, T.; Tecklenburg, W. (2004). 3-D Object Reconstruction from Multiple-Station Panorama Imagery. *Proc. of ISPRS Workshop on Panorama Photogrammetry*, Vol. 34, Part 5/W16.
- Maas, H.-G. (2008). Close-range photogrammetry sensors, *ISPRS Congress*, CRC, pp. 63–72.



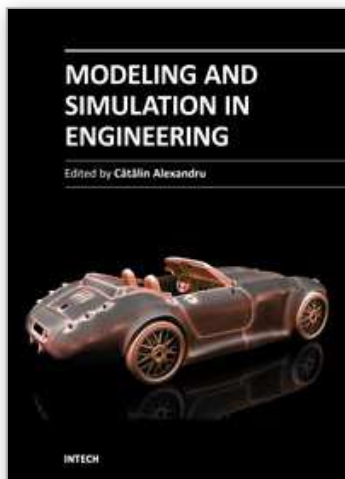
- MacKinnon, D.; Beraldin, J.-A.; Cournoyer, L.; Blais, F. (2008). Evaluating Laser Spot Range Scanner Lateral Resolution in 3D Metrology. Proceedings of the 21st Annual IS&T/SPIE Symposium on Electronic Imaging. San Jose. January 18-22, 2008.
- Meyers, D.; Skinner, S. & Sloan, K. (1992). Surfaces from contours. *ACM Transactions on Graphics*, Vol. 11, No. 3, pp. 228-258.
- Mikhail, E.; Bethel, J.; McGlone, J.C. (2001). *Introduction to Modern Photogrammetry*, Wiley.
- Ohtake, Y.; Belyaev, A.; Alexa, M.; Turk G. & Seidel, H. P. (2003). Multi-level partition of Unity Implicits, *ACM Trans. on Graphics*, Vol. 22, No. 3, pp. 463-470.
- Parian, J.A. & Gruen, A. (2004). An Advanced Sensor Model for Panoramic Cameras, *Proc. of XXth ISPRS Congress*; Vol. 35, Part 5, pp. 24-29
- Peggs, G. N.; Maropoulos, P. G.; Hughes, E. B.; Forbes, A. B.; Robson, S.; Ziebart, M.; Muralikrishnan, B. (2009). Recent developments in large-scale dimensional metrology, *Journal of Engineering Manufacture*, Vol. 223, No. 6, pp. 571-595.
- Petrov, V.; Kazarinov, A.; Cherniavsky, A.; Vorobey, I.; Krikalev, S.; Mikryukov, P. (2006), Checking of large deployable reflector geometry, *Proc. EuCap 2006 - European Space Agency Special Publication*, Vol. ESA SP-626, 4p.
- Pierce, J. (2007). Wider view, *Engineer*, Volume 293, Issue 7735, Pages 36-40.
- Pierrot-Deseilligny, M. & Paparoditis, N. (2006). A Multiresolution and Optimization-Based Image Matching Approach: An Application to Surface Reconstruction from SPOT5-HRS Stereo Imagery. *ISPRS Conf. Topographic Mapping From Space*, Vol. 36 1/W41.
- Pollefeys, M., Nister, D., Frahm, J.-M., Akbarzadeh, A., Mordohai, P., Clipp, B., Engels, C., Gallup, D., Kim, S.-J., Merrell, P., Salmi, C., Sinha, S., Talton, B., Wang, L., Yang, Q., Stewenius, H., Yang, R., Welch, G. and Towles, H. (2008). Detailed Real-Time Urban 3D Reconstruction From Video, *International Journal of Computer Vision*, Vol. 78, No. 2, pp. 143-167.
- Pollefeys, M.; van Gool, L.; Vergauwen, M.; Verbiest, F.; Cornelis, K.; Tops J. & Kock, R. (2004). Visual modelling with a hand-held camera. *International Journal of Computer Vision*, Vol. 59, No. 3, pp. 207-232.
- Reinhard, E.; Ward, G.; Pattanaik, S. & Debevec, P. (2005). *High dynamic range imaging: acquisition, display and image-based lighting*. Morgan Kaufmann Publishers.
- Remondino, F., El-Hakim, S., Girardi, S., Rizzi, A., Benedetti, S., Gonzo, L., (2009). 3D Virtual reconstruction and visualization of complex architectures - The 3D-ARCH project. *ISPRS*, Vol. 38(5/W1).
- Remondino, F. & El-Hakim, S. (2006). Image-based 3D modelling: A review, *The Photogrammetric Record* Vol. 21, pp. 269-291.
- Remondino, F.; El-Hakim, S.; Gruen, A.; Zhang, L. (2008). Turning images into 3D models – Development and performance analysis of image matching for detailed surface reconstruction of heritage objects, *IEEE Sig. Process. Mag.* Vol. 25, pp. 55-65.
- Remondino, F. & Fraser, C. (2006). Digital camera calibration methods: considerations and comparisons, *ISPRS*, Vol. XXXVI, part 5, pp. 266-272.
- Riegl (2010). VZ400 Terrestrial Laser Scanner with on-line waveform processing, Data Sheet, [www.riegl.com/uploads/tx\\_pxpriegldownloads/10\\_DataSheet\\_VZ400\\_20-09-2010.pdf](http://www.riegl.com/uploads/tx_pxpriegldownloads/10_DataSheet_VZ400_20-09-2010.pdf)

- Rusinkiewicz, S. & Levoy, M. (2001). Efficient variants of the ICP algorithm, *Proceedings of 3DIM2001*, IEEE Computer Society, pp. 145-152.
- Sagawa R. & Ikeuchi K. (2008). Hole filling of a 3D model by flipping signs of a signed distance field in adaptive resolution, *IEEE Trans. PAMI*, Vol. 30, No. 4, pp. 686-699.
- Salvi, J.; Matabosch, C.; Fofi, D. & Forest J. (2007). A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, Vol. 25, No. 5, pp. 578-596.
- Scharstein, D. & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Intl. Journal of Computer Vision*, Vol. 47, pp. 7-42.
- Seitz, S.M.; B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. (2006). A comparison and evaluation of multi-view stereo reconstruction algorithm. *Proc. CVPR*, pp. 519-528.
- Shan, J. & Toth, C. (2008). *Topographic Laser Ranging and Scanning: Principles and Processing*, CRC, p. 590.
- Shannon, C. E. (1949). Communication in the presence of noise, *Proc. Institute of Radio Engineers*, vol. 37, no.1, pp. 10-21, Jan. 1949. Reprint as classic paper in: *Proc. IEEE*, Vol. 86, No. 2, (Feb 1998).
- Skolnik M.I. (1990). *Radar Handbook*, chapt. 14, McGraw-Hill.
- Snively, N.; Seitz, S.; Szeliski, R. (2008). Modelling the world from internet photo collections. *Int. J. Comput. Vis.*, Vol. 80, pp. 189-210.
- Soucy, M. & Laurendeau, D. (1995). A general surface approach to the integration of a set of range views, *IEEE Trans. on PAMI*, Vol. 17, No. 4, pp. 344-358.
- Stamos, I.; Liu, L.; Chen, C.; Wolberg, G.; Yu, G.; Zokai, S. (2008). Integrating automated range registration with multiview geometry for the photorealistic modelling of large-scale scenes. *Int. J. Comput. Vis.*, Vol. 78, pp. 237-260.
- Stumpfel, J.; Tchou, C.; Yun, N.; Martinez, P.; Hawkins, T.; Jones, A.; Emerson, B.; Debevec, P. (2003). Digital reunification of the Parthenon and its sculptures. *Proceedings of VAST*, pp. 41-50
- Sturm, P.; Ramalingam, S.; Tardif, J.-P.; Gasparini, S.; Barreto, J. (2011). Camera models and fundamental concepts used in geometric Computer Vision. *Foundations and Trends in Computer Graphics and Vision*, Vol. 6, Nos. 1-2, pp.1-183
- Tsai, R.Y. (1986). An efficient and accurate camera calibration technique for 3d machine vision, *Proc. CVPR 1986*, pp. 364-374.
- Turk, G. & Levoy, M. (1994). Zippered polygon meshes from range images, *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*, pp. 311-318.
- Umeda, K.; M. Shinozaki, G. Godin, and M. Rioux. (2005). Correction of color information of a 3d model using a range intensity image, *Proc. of 3DIM2005*, pp. 229-236.
- Van den Heuvel, F.A. (1998). 3D reconstruction from a single image using geometric constraints, *ISPRS Journal for Photogrammetry and Remote Sensing*, Vol. 53, No. 6, pp. 354-368.
- Vergauwen, M. & Van Gool, L. (2006). Web-based 3D reconstruction service, *J. Mach. Vis. Appl.*, Vol. 17, pp. 411-426.
- Vosselman, G. & Maas, H.-G. (2010). *Airborne and Terrestrial Laser Scanning*, CRC, p. 318.
- Vrubel, A.; Bellon, O.R.P.; Silva, L. (2009). A 3D reconstruction pipeline for digital preservation, *Proceedings of IEEE Conference on CVPR*, pp. 2687-2694.

- Winkelbach S. & F.M. Wahl. (2001). Shape from 2D Edge Gradient, *Lecture Notes in Computer Science*, Vol. 2191, pp. 377-384
- Yin, X.; Wonka, P.; Razdan, A. Generating. (2009). 3D building models from architectural drawings. *IEEE Comput. Graph. Appl.*, Vol. 29, pp. 20-30.

IntechOpen

IntechOpen



## **Modeling and Simulation in Engineering**

Edited by Prof. Catalin Alexandru

ISBN 978-953-51-0012-6

Hard cover, 298 pages

**Publisher** InTech

**Published online** 07, March, 2012

**Published in print edition** March, 2012

This book provides an open platform to establish and share knowledge developed by scholars, scientists, and engineers from all over the world, about various applications of the modeling and simulation in the design process of products, in various engineering fields. The book consists of 12 chapters arranged in two sections (3D Modeling and Virtual Prototyping), reflecting the multidimensionality of applications related to modeling and simulation. Some of the most recent modeling and simulation techniques, as well as some of the most accurate and sophisticated software in treating complex systems, are applied. All the original contributions in this book are joined by the basic principle of a successful modeling and simulation process: as complex as necessary, and as simple as possible. The idea is to manipulate the simplifying assumptions in a way that reduces the complexity of the model (in order to make a real-time simulation), but without altering the precision of the results.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Gabriele Guidi and Fabio Remondino (2012). 3D Modelling from Real Data, Modeling and Simulation in Engineering, Prof. Catalin Alexandru (Ed.), ISBN: 978-953-51-0012-6, InTech, Available from: <http://www.intechopen.com/books/modeling-and-simulation-in-engineering/3d-modeling-from-real-data>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen