# INTEGRATION OF TOF CAMERA AND MULTI-IMAGE MATCHING APPROACH FOR CULTURAL HERITAGE SURVEY

F. Chiabrando [a], F. Nex [b], D. Piatti [b,] *, F. Rinaudo [b]

[a] Dept. of Human Settlements Science and Technology, Politecnico di Torino, Viale Mattioli 39, 10125 Torino, Italy
- filiberto.chiabrando@polito.it
[b] Dept. of Land, Environment and Geo-Engineering , Politecnico di Torino, C.so Duca degli Abruzzi 24, 10121 Torino, Italy
- (francesco.nex, dario.piatti, fulvio.rinaudo)@polito.it

**Commission V, SS-2**

**KEY WORDS:** Time-of-Flight, Multi-Image Matching, Integration, Breakline Extraction, Automatic Drawing, Calibration

**ABSTRACT:**

In the last few years, several papers dealing with the integration between different surveying techniques (such as photogrammetry, LiDAR, NIR image processing, 2D digital maps, etc.) have been already presented. The main idea of these works is the possibility to overcome the individual weakness of each technique through their combined use, exploiting their complementary nature. In this work, the new integration between data provided by Time-of-Flight (ToF) cameras and a multi-image matching technique is presented. In particular, this approach is considered for the cultural heritage data acquisition and processing. The main advantage is given by the quickness in the data acquisition (only few minutes are requested) and the reduced cost of the instruments, if compared to terrestrial laser scanners. According to the proposed approach, ToF camera and photogrammetric techniques continuously share information in order to extract the geometric breaklines necessary for the cultural heritage survey. One of the first experiences applying this new method on an architectural artefact is presented, in order to show its potentiality for metric survey and architectural drawing purposes.

## 1. INTRODUCTION

The production of architectural and cultural heritage surveys usually requires complete and reliable geometrical information about the object to be described. Traditionally, this information is obtained using time-consuming surveys (total stations, levels, etc.) or performing manual photogrammetric plottings. In the last few years, the introduction of LiDAR instruments has allowed to quickly acquire complete point clouds of the object to be surveyed. Nevertheless, this data cannot be directly used for 2D representations (plans and sections) because of the difficulty in the plotting on a 3D point cloud. Moreover, several manual interventions have to be usually performed on the extracted data since no automatic and reliable procedures have been performed up to now.

In order to represent metrically correct object geometries it is necessary to measure and extract from the metric survey data the breaklines that allow the artifact to be described at the requested representation scale.

All the previous techniques can be useful and each of them allows an object to be surveyed with a certain accuracy; nevertheless, in order to achieve a more accurate and complete survey it is usually necessary to integrate the information coming from more than one technique.

The total station survey is often employed for accurately measure the breaklines and the GCP (Ground Control Points) useful for the photogrammetric process and for the LiDAR survey and, consequently, to define the reference system (usually local in the architectural survey).

On the other hand, a complete survey with topographic instruments could request lot of time; for this reason, in many situations, it could be important to integrate this technique with photogrammetry and/or LiDAR surveys.

The photogrammetry approach is advantageous for the data acquisition time; nevertheless, the data processing needs a lot of time as the majority of the process is manually performed.

The LiDAR technique allows a huge number of points to be quickly acquired with quite good precision; however, the main drawbacks are the high cost of the instruments and the long time needed to process the acquired data.

The main objective of this work is to propose a low cost approach that follows two different aims: quickness in the data acquisition and automatic breakline extraction in order to ease and speed up the drawing production.

The approach is based on the employment of a Time of Flight camera (also called 3D camera) for the generation of an approximate Digital Surface Model (DSM) to be used in a multi-image-matching photogrammetric approach that allows the object breaklines to be automatically extracted. This method has already given good results considering the integration between LiDAR data and multi-image matching techniques (Nex and Rinaudo, 2009). In this paper, the extension to the ToF data is presented. In the following sections the acquisition

* Corresponding author.

phases, the data processing steps and first results obtained from the Geomatics research group of the Politecnico di Torino are reported.

## 2. TOF CAMERAS

### 2.1 State of the art

As it is well known, ToF cameras represent a rather new 3D measurement way to obtain 3D point clouds, which are almost comparable with those of traditional LiDAR acquisitions. Using these cameras, a bundle of distances is determined simultaneously (at video frame rates) for each pixel of a two-dimensional sensor array. Although ToF cameras are usually characterized by no more than a few thousands of tens of pixels, a maximum unambiguous measurement range up to thirty meters can be reached and complete 3D point clouds of the analyzed objects can be quickly acquired. These devices allow to generate point clouds such as in the case of the LiDAR technique and photogrammetry but with the great advantage of real time acquisition, low cost and handiness. Unlike photogrammetric techniques, 3D cameras allow a point cloud to be obtained of the object which has to be surveyed from even only one point of view, without the need of any particular lighting conditions, since these cameras are active sensors that work outside of the visible spectrum.
ToF cameras usually deliver a range image and an amplitude image with infrared modulation intensities at video frame rates: the range image (or depth image) contains for each pixel the radial measured distance between the considered pixel and its projection on the observed object, while the amplitude image contains for each pixel the strength of the reflected signal by the object. In some cases an intensity image is also delivered, which represents the mean of the total light incident on the sensor (reflected modulated signal and background light of the observed scene).
Nevertheless, ToF camera measurements usually suffer from some systematic errors that have to be corrected by using suitable calibration procedures in order to refine the measurement accuracy (Anderson et al., 2005, Kahalmann et al., 2006, Lichti, 2008, Lindner et al., 2006, Rapp et al., 2008, Weyer et al., 2008).

### 2.2 SR-4000 camera

In this work the Swiss-Ranger-4000 camera (SR-4000) was employed. In order to give an idea of its characteristics, the principal specifications of the SR-4000 camera are reported in Table 1, while a picture of this camera is reported in Figure 1. For more details about camera specifications see (www.mesaimaging.com).

| | |
|---|---|
| ***Pixel array pitch [-]*** | 176 (h) × 144 (v) |
| ***Field of view [°]*** | 43.6 (h) × 34.6 (v) |
| ***Pixel pitch [μm]*** | 40 |
| ***Illumination wavelength with standard settings [nm]*** | 850 |
| ***Working range with standard settings [m]*** | 0.3 ÷ 5.0 |
| ***Maximum frame rate [fps]*** | 54 |
| ***Dimensions [mm]*** | 65 × 65 × 68 |
| ***Weight [g]*** | 470 |

Table 1. SR-4000 specifications.

In the presented work, all the acquired data were corrected with the distance error model proposed in (Chiabrando et al., 2009), which refines the camera distance measurement accuracy: in particular, it has been shown that the distance measurements of the SR-4000 camera can reach accuracies approximately up to 1 cm. Moreover, in (Chiabrando et al., 2009) it has been shown that a camera warm-up of about 40 minutes should be performed in order to achieve a good stability of the camera distance measurements.

Figure 1. SR-4000 camera.

In order to test the potentiality of the camera for architectural surveys, some tests on a window of the *Castello del Valentino* (the headquarter of the Architecture Faculty of the Politecnico di Torino), have been performed. In particular, the 3D point clouds acquired with the ToF camera were processed and used as approximate DSM for the multi-matching approach described in section 3.

## 3. MULTI-IMAGE MATCHING APPROACH

The automated extraction of objects from photogrammetric images has been a topic of research for decades. Nowadays, image matching techniques allow a great number of points to be extracted in a quick way. In particular, the multi-image matching techniques allow an improvement in the geometric precision and the reliability with respect to image pairs, by extracting points and edges from images and projecting their match in the space (Zhang, 2005). Actually, this kind of solution requires an approximate DSM in order to “drive” the solution to the correct match; the more accurate is this model, the more correct (without blunders) is the solution.
Image matching techniques have shown good results in aerial applications. In particular, multi-image techniques have improved the results in terms of precision and reliability (Zhang, 2005), and allowed a point cloud density comparable to the LiDAR one to be obtained. These techniques consider the epipolar geometry between images in order to reduce the search area in adjacent images, and thus decreasing the number of blunders to a great extent. The run on the epipolar line is further reduced by the z-value which is provided by an approximate DSM. Nevertheless, the z value varies smoothly in aerial cases, compared to the relative flight height. The height differences between the top and the bottom of the buildings are small, compared to the taking distance: a 40 m high building is only 1/20 of an 800 m flight height. In these conditions, the approximate DSM can be achieved by the feature extraction and the matching of interest operators.
In contrast, the terrestrial case shows greater z-values differences: façades are usually rough with balconies, columns or decorations that protrude of several meters. These variations are more relevant as they can be 1/5 of the taking distance. In this situation, z-values provided by approximate DSM (created

through a photogrammetric approach) are not sufficient to limit the run of the epipolar line: blunders are more frequent during the matching procedure and it is still difficult to filter them. Furthermore, the façade texture is often not good enough to allow effective matching techniques to be performed: blank areas in automatically extracted point cloud are very common in correspondence of painted walls. Until now, fully matching techniques have only achieved good results in bas-relief or in limited area surveys.
In order to overcome these problems the photogrammetric process can be helped by the ToF camera data. The main idea is to use the reliable information provided by this instrument as DSM in the matching algorithms.

### 3.1 Breakline extraction algorithm

The proposed algorithm can be summarized into several steps shown in Figure 2. In particular, the steps where the ToF data gives additional information to the multi-image matching approach are highlighted in orange colour.

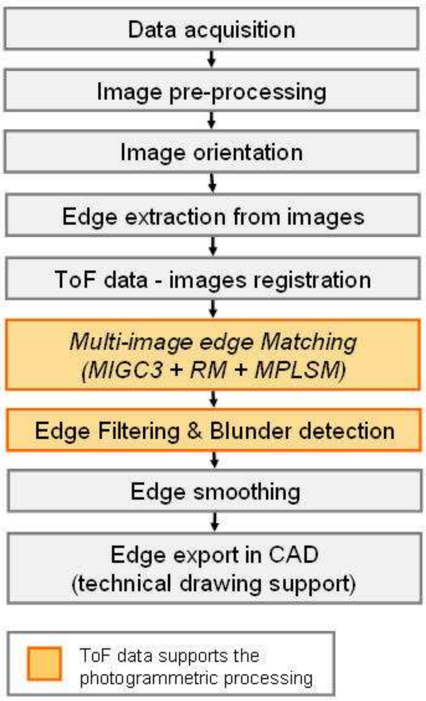


Figure 2. Breakline extraction workflow.

The images are acquired according to an *ad hoc* taking configuration (Figure 3): several images are acquired and the most central one is considered as reference image during the matching process. The ToF point cloud is acquired from a central position with respect to the image acquisition in order to have approximately the same occluded areas in the ToF data and in the reference image.
The acquired images are pre-processed according to an auto-adaptive smoothing. Then, they are enhanced using a Wallis filter (Wallis, 1976); this filter is able to sharpen the radiometric boundaries and to enhance the edges.
The orientation is performed in a proper reference system in order to have the z-coordinate normal to the main plain of the façade. In this step, the $A^2$SIFT (Auto-Adaptive Scale Invariant Feature Transform) operator (Lingua et al., 2009) is adopted in the tie-point extraction and a robust (Least Median Square) relative orientation is then performed in order to eliminate the mismatches (Lingua and Rinaudo, 2000). Finally, a bundle block adjustment is performed. After that, the edge extraction is performed by the Canny operator (Canny, 1986) on the reference image. The extracted edges are then approximated, by identifying the pixels where the edge changes in direction as knots and linking these dominant points by straight edges.

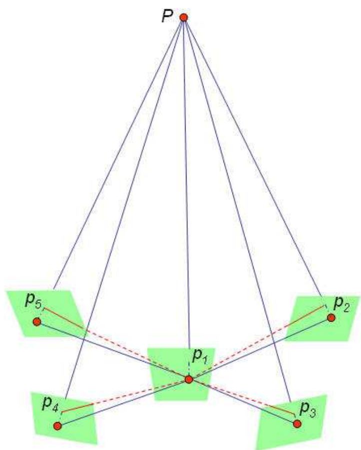


Figure 3. *Ad hoc* image taking configuration.

The point cloud is registered in the photogrammetric reference system by means of a spatial similarity transformation. In this way, it is possible to share the information between the images and the point cloud. Then, a multi-image matching algorithm is set up. The algorithm is similar to the Geometrically Constrained Cross Correlation ($GC^3$) (Zhang, 2005): it uses a multi-image approach, that is, it considers a reference image and projects the image patch (of each dominant point) of the reference image onto the DSM (ToF point cloud), and then, using the approximate z-value achieved by the DSM, back-projects it onto the other images. Through this algorithm, the dominants points of each edge are matched in all the images in order to reconstruct the breakline positions in 3D. The images are preliminarily undistorted (using the camera calibration) in order to ease them into a central perspective. The epipolar constraint limits the search space in the images. The length of this line could be achieved considering the z-value given by the ToF point cloud; then, in order to find the homologous point in all the images, this value is varied into a range ($\Delta z$). This work is enforced and improved through the position of already matched points: the z-value of two adjacent dominant points on the same edge must be similar. In this way it is possible to reduce the run of the epipolar line on the façade to a few centimetres. In order to improve the rate of the successfully matched points, a relational matching has been developed. This algorithm allows several ambiguities during the matching phases to be solved by imposing a smoothness constraint. A Multi-photo Least Square Matching (MLSM) (Baltsavias, 1991) has been performed for each extracted point, in order to improve the accuracy up to a sub-pixel dimension.
During the matching process, some blunders can be generated. These blunders are firstly deleted from the extracted edges using a filter which considers the reciprocal point positions on the same edge: in particular, the position of a point is predicted considering the neighbouring points of the edge and, then, the difference between the predicted and the real position of the point is evaluated. If the difference value is higher than a

predefined threshold the point is deleted. This filter is not robust: it will work well if the blunders are isolated from each other. For this reason, a second filter could be used to clean the edges when several blunders are close together: this algorithm uses the ToF information to verify the correctness of each dominant point: when it is farther than a defined threshold from the point cloud, it is deleted.

Then, the image matching allows radiometric edges to be extracted. Most of these edges are due to shadows or radiometric changes but they have no a geometric correspondence. Only geometric boundaries are of interest in the surveying graphic drawings and for modelling purposes. For this reason, the position of each dominant point on the extracted edges is considered with respect to the ToF point cloud: it is verified whether a geometric discontinuity occurs in the ToF data close to the edge point.

The edges extracted by the matching algorithm are random noise affected and they cannot be directly used in the drawing production. For this reason, the noisy edges are split in basic elements (linear and curved elements) and each element is smoothed and eased, in an automatic way, into lines and second order curves by means of a polynomial fitting. Finally, the basic elements are recollected in a unique smoothed edge (Nex, 2010).

Finally, geometric edges are exported in CAD in order to give preliminary data for the graphic drawing realization of the survey and for a rough evaluation of the achieved results.

## 4. EXPERIMENTAL TEST

A first test of the proposed approach was realized on a window of the *Castello del Valentino* (Figure 4). The façades are painted and the texture is generally not good enough for the traditional image matching approach to be performed. In the following sections, the data acquisition and processing information are reported for both the ToF camera and the multi-image matching approach.

Figure 4. Three views of the surveyed window of the *Castello del Valentino.*

### 4.1 ToF data acquisition

After the camera warm-up, the SR-4000 was positioned on a photographic tripod and moved to different positions in order to achieve a complete coverage of the window to be surveyed (Figure 5). Thirty frames were acquired from each position using the software delivered with the camera (SR_3D_View software), adjusting the integration time in order to have the minimum possible number of saturated pixels but a low noise level of the distance measurements.

Figure 5. SR-4000 data acquisition.

According to the SR-4000 specifications, since the average distance between the camera and the window was about 3.5 m, the acquired area dimensions for each range image were about 3.00 m x 2.50 m.

In order to obtain a complete 3D model of the window (3.0 m large, 5.00 m high), the ToF data was acquired from six different positions, with a good overlap between the acquired range images.

In Figure 6 two screen-shots of the processed point clouds are reported.

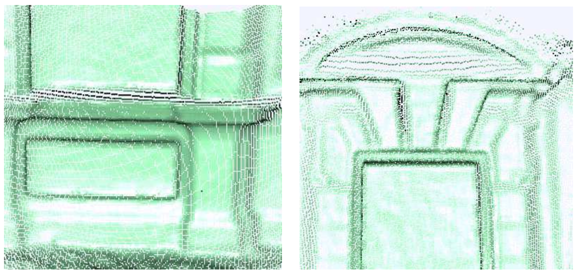

Figure 6. Screen-shots of the processed point clouds acquired from different positions.

### 4.2 ToF data processing

The thirty frames acquired from each of the six positions were averaged in order to reduce the measurement noise.

Then, the distance of each pixel of the averaged frames was corrected with the distance error model proposed in (Chiabrando et al., 2009), using a custom-made *Matlab*® application.

The obtained point clouds were registered using the ICP algorithm implemented in the *Geomagic Studio*® software in order to obtain a unique 3D model of the artefact.

Figure 7 shows some screen-shots of the point cloud obtained after the registration phase.

From this product a dense DSM (168551 points) was generated and then employed for the multi-image -matching approach in order to extract the breaklines needed for the 2D representation of the surveyed object.

Figure 7. Screen-shots of the complete 3D ToF point cloud.

### 4.3 Image acquisition

The image acquisition was performed using the *Canon Eos-5D Mark II* digital camera equipped with a 24 mm lens. The taking distance was about 6 m. Five images were acquired according to an *ad hoc* configuration (Nex and Rinaudo, 2010). In Figure 8 an example of epipolar lines and correlation patches on the five images employed for the multi-image matching approach is reported.

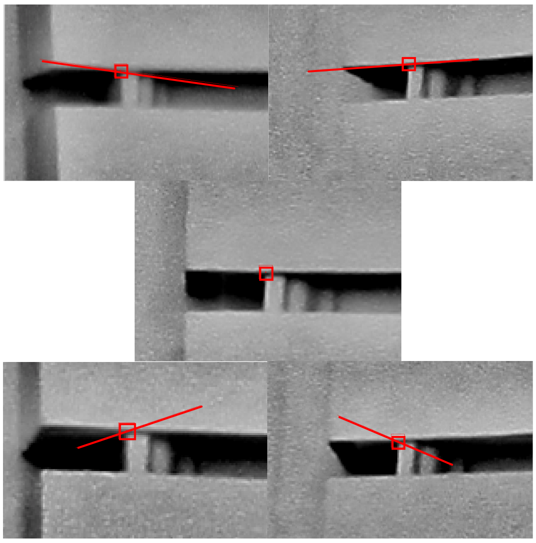

Figure 8. Epipolar geometry of the acquired digital images.

According to this particular configuration the epipolar lines run in tilted direction with respect to the main lines of the façade (horizontal and vertical), and the homologous points can be determined in an unambiguous way. Thanks to the image dimension (5616 x 3744 pixels) and the short taking distance, an excellent image resolution was reached (less than 2 mm of Ground Sample Distance (GSD)).

### 4.4 Data integration and results

The integration between ToF data and digital images has been performed according to the workflow reported in Figure 2.

The edge extraction allowed a complete set of lines to be defined from the reference image: Figure 9 shows the extracted edges, which are described by 45636 dominant points.

Figure 9. Extracted edges on the reference image.

After the matching process, the position in the space of 32566 dominant points was defined. Only a percentage of 3% of these points was deleted after the blunder detection process. The majority of these deleted points was concentrated on the gray wall of the central part of the window. The resulting data was smoothed in order to ease the edges in lines and curves.
The result of this work is reported in Figure 10: it can be noticed that the window geometry is complete and only some parts of the arcs are missing.

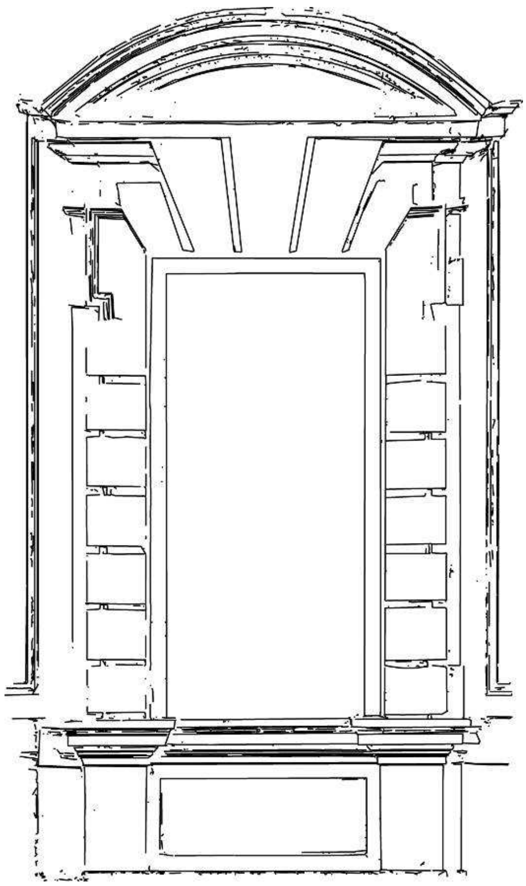

Figure 10. Smoothed edges.

The smoothing correctly eased all the elements of the façade. In Figure 11 a zoom of the final result is shown.

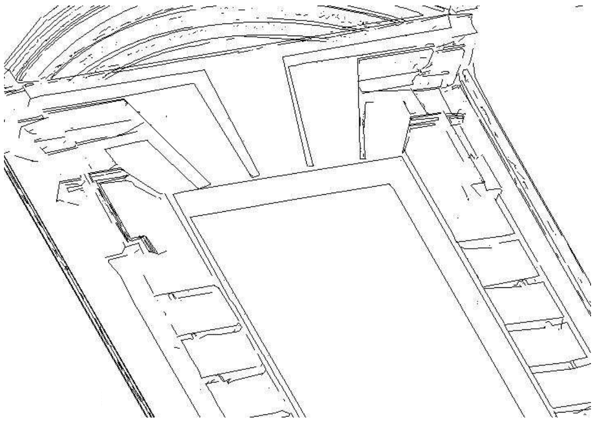

Figure 11. Detail of the smoothed edges.

## 5. CONCLUSIONS AND FUTURE WORKS

In this work, the new integration between data provided by Time-of-Flight (ToF) cameras and a multi-image matching technique has been presented.

The ToF data is usually affected by some systematic errors that can be corrected by using a suitable calibration procedure, achieving complete 3D point clouds comparable with those of traditional LiDAR acquisitions. The resulting point clouds can be registered and oriented in the photogrammetric coordinate system in order to generate an approximate DSM of the object.

The main drawbacks of the ToF cameras are the maximum unambiguous measurement range (tens of meters), that reduces the extension of the area to be surveyed by a single camera position, and the low resolution of actual sensors. On the other hand, range cameras are able to acquire data at video frame rates and they are characterized by low costs and handiness.

The achieved results show that using the ToF DSM, good results in the realization of architectural drawings can be achieved. Using the proposed approach it is possible to drastically reduce both the data acquisition and processing times for 2D and/or 3D rough drawing generation (these drawings were automatically generated using the proposed algorithm).

In order to achieve a complete architectural representation the results need to be manually edited and, if necessary, integrated.

In the future, some improvements have to be performed in order to improve the performances of the integration algorithm and to increase the completeness of the achieved results.

Moreover, further investigations will be performed to assess the geometric accuracy of the final drawings, in order to verify the maximum representation scale that can be achieved in order to have a correct metric representation and to limit the editing time for the final drawing production.

## REFERENCES

Anderson, D., Herman, H., Kelly, A., 2005. Experimental Characterization of commercial flash ladar devices. In: *Proceedings of International Conference on Sensing Technologies*, Palmerston North, New Zealand.

Baltsavias E., 1991. Multiphoto Geometrically Constrained Matching. Phd. dissertation, ETH Zurich, Switzerland.

Canny, J.,1986. A Computational Approach To Edge Detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, n° 8, pp. 679–714.

Chiabrando, F., Chiabrando, R., Piatti, D., Rinaudo, F., 2009. Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera. *Sensors*, n° 9, pp. 10080-10096.

Habib, A. F., Ghanma, M. S., Tait, M., 2004. Integration of LIDAR and photogrammetry for close range applications. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV/B2, Istanbul, Turkey.

Kahlmann, T.; Remondino F.; Ingensand, H., 2006. Calibration for increased accuracy of the range imaging camera Swiss Ranger. *Int. Soc. Photogramm. Remote Sens. 2006,* XXXVI, pp. 136-141.

Lingua, A., Marenchino, D., Nex, F., 2009. Performance Analysis of the SIFT Operator for Automatic Feature Extraction and Matching in Photogrammetric Applications. *Sensors*, n° 9, pp. 3745-3766.

Lichti, D. Self-Calibration of a 3D Range Camera 2008. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, Vol. XXXVII, part B5, pp. 927-932.

Lindner, M.; Kolb, A., 2006. Lateral and depth calibration of PMD-distance sensors. In: *Proceedings of ISVC*, Lake Tahoe, NV, USA, pp. 524-533.

Lingua A., Rinaudo F., 2000. Aerial triangulation data acquisition using a low cost digital photogrammetric system. In: *The International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII/B2, pp. 449-454.

Nex, F., 2010. Multi-Image Matching and LiDAR data new integration approach. PhD Thesis, Politecnico di Torino, Torino, Italy.

Nex, F., Rinaudo, F., 2009. New integration approach of Photogrammetric and LIDAR techniques for architectural surveys. In: *Laserscanning 2009*, Paris, France.

Rapp, H.; Frank, M.; Hamprecht, F.A.; Jähne, B.,2008. A theoretical and experimental investigation of the systematic errors and statistical uncertainties of Time-of-Flight-cameras. *IJISTA*, n° 5, pp. 402-413.

Wallis, R., 1976. An approach to the space variant restoration and enhancement of images, In: *Proceedings of Symposium on Current Mathematical Problems in Image Science*. Monterey CA, USA, pp. 329- 340.

Weyer, C.A.; Bae, K.; Lim, K.; Lichti, D., 2008. Extensive metric performance evaluation of a 3D range camera. *Int. Soc. Photogramm. Remote* Sens. XXXVII, pp. 939-944.

Zhang, L., 2005. Automatic Digital Surface Model (DSM) generation from linear array images. Phd. dissertation, ETH Zurich, Switzerland.

www.mesaimaging.com (accessed on March 2010).