MULTI-IMAGE MATCHING: AN “OLD AND NEW” PHOTOGRAMMETRIC ANSWER TO LIDAR TECHNIQUES

F. Nex [a] *, F. Rinaudo [a]

[a] Dept. of Land, Environment and Geo-engineering (DITAG), Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy, (francesco.nex, fulvio.rinaudo)@polito.it

Commission V, Working Group V/3

KEY WORDS: Close Range Photogrammetry, Digital, Matching, Segmentation, Software, Accuracy, Cultural Heritage

ABSTRACT:

Over the last decade, LIDAR techniques have replaced traditional photogrammetric techniques in many applications because of their speed in point cloud generation. However, these laser scanning techniques have non-negligible limits and, for this reason, many researchers have decided to focus on improving the performances of matching technique in order to generate dense point clouds from images.The first tests carried out at the Politecnico di Torino on the first fully-automated multi-image matching commercial software, the ZScan Menci Software, are described in this paper. This instrument was first devised to allow inexperienced users to generate very dense point clouds from image triplets; a customized calibrated bar (0,90 m length) is used for image acquisition. Recently a new version of *ZScan* has been created in order to elaborate triplets of oriented aerial images and generate DSM: the first results obtained in this way are presented in this paper. Several tests have been performed on the *ZScan* performances analysing different geometrical configurations (base-to-height ratio) and textures. The evaluation of the geometric precision obtained by this software in point cloud generation may help to understand which performances can be achieved with a fully automated multi-image matching. The evaluation concerns what the most critical aspects of these techniques are and what improvements will be possible in the future. Furthermore a possible new research project is described which has the aim of transferring useful information about breakline location from images to point clouds in order to derive automatically the segmentation algorithms.

1. INTRODUCTION

Over the last decade, LIDAR techniques have replaced traditional photogrammetric techniques in many applications, because of their speed in point cloud generation and the percentage of acceptable points; traditional image matching, based on the use of stereopairs, can acquire no more than 80% of the possible points (the holes have to be manually completed).

However, these laser scanning techniques have non-negligible limits due to the impossibility of directly obtaining radiometric information and the exact position of object breaklines. For this reason, most LIDAR surveys are integrated by digital image acquisition. Digital photogrammetry directly associates a radiometric information to the acquired points and the use of a stereopair allows a manual survey of the breaklines when automatic algorithms fail. Then, other limits for LIDAR terrestrial applications are the weight and size of the instruments, the limited range of applications (especially for environmental applications) and last but not least the cost.

Finally, when point clouds are produced, segmentation and classification procedures have to be applied in order to correctly interpret and model the surveyed object. The acquired experience shows that automatic algorithms are not ready to offer correct solutions without direct human help or without an image interpretation.

For these reasons, many researchers have decided to focus on improving the performances of matching techniques in order to generate dense point clouds from only images.

2. MATCHING ALGORITHMS

As is well known, these techniques were devised more than twenty years ago and are nowadays subdivided into Area Based Matching (*ABM*) and Feature Based Matching (*FBM*).

Since then, the improvement of these techniques has been ongoing with research into new algorithms that are capable of improving their efficiency and precision. This research has led to matching algorithms which have become more complete (Gruen, 1985) and improved their effectiveness, imposing new constraints, such as collinearity conditions (Gruen, Baltsavias, 1988), and integrating the surface reconstruction into the process. Furthermore, these techniques now allow more than two images to be handled and simultaneous determination of all matches of a point: these features help overcome the problems of mismatches thus increasing the precision achieved (Forstner, 1998).

These techniques have recently been improved by other authors who have adapted the algorithms to aerial (Gruen, Zhang, 2004) and terrestrial photogrammetry (El-Hakim, et al.2007; Remondino, 2007), partly overcoming some of the matching problems (wide baselines between images, illumination changes, repeated pattern and occlusions). An important aspect of these

---

[*] Corresponding author.

new algorithms is the integration of photogrammetric equations with other devices commonly used in the Computer Vision application field such as rectification (Zhang et al., 2003; Du, et al., 2004) and tensorial notations (Roncella et al., 2005; Hartley, et al., 2000). The practicality of these instruments is essentially that they simplify and speed up the photogrammetric process or provide traditional techniques with information obtained using new techniques such as the shape from shading (Fassold et al., 2004) where the texture is not sufficient for matching techniques.

Thanks to these algorithms, over the last decade some 3D modelling software packages have been created: these new instruments can be divided into semi-automated and automated software.

Other papers (El-Hakim et al., 2007) have already described the performances of human interactive (semi-automated) software and demonstrated the great potential of these techniques.

The main aim of this paper is to describe the tests that have been carried at the Politecnico di Torino on the new Menci Software ZScan, one of the first fully automated multi-image matching commercial software programmes. This software was first devised to allow inexperienced users to generate very dense point clouds from image triplets; a customized calibrated bar (0,90 m length) is used for image acquisition. Recently, a new version of ZScan has been created to elaborate triplets of oriented aerial images and generate DSM.

Some tests are described in the following sections with the aim of demonstrating the advantages of using a multi-image matching approach compared with traditional stereopair management, using ZScan software both for terrestrial and aerial applications.

The principal goal of these tests was to understand whether a multi-image approach could produce a point cloud with the same precision and density as LIDAR point clouds.

3. POINT CLOUDS MANAGEMENT

Once a point cloud has been acquired, many automated and manual interventions have to be applied in order to segment, classify and model the surveyed points. The main topic is the extraction of the breaklines.

Different ways of solving this problem have already been proposed in scientific literature. In photogrammetry one way of solving this problem is represented by the extraction of edges from images and then their matching in the space using different algorithms (Hauel, et al., 2001; Zhang, 2005).

A segmentation in LIDAR applications has instead only been performed using the point cloud information as curvature (Beinat et al., 2007) or static moment (Roggero, 2002). However practical experience has shown that the use of images can be of help for breakline extraction, so all LIDAR surveys are usually integrated by digital images recorded during LIDAR acquisition. If the images are oriented in the LIDAR coordinate system, it is possible to integrate LIDAR point clouds and derive automatic segmentation algorithms to find the breaklines. If point clouds are generated by means of a photogrammetric approach the breaklines can be directly plotted from stereomodels.

A possible research project is described in the last section with the aim of transferring useful information about breakline location from images to point clouds in order to derive the segmentation algorithms automatically.

4. ZSCAN SYSTEM

The ZScan system was originally thought up to allow inexperienced users to easily generate point clouds from image triplets, especially in architectural and cultural heritage surveys. In order to reach this goal, ZScan Software needs to have three images acquired using a 0.9 m long calibrated bar. The maximum baseline between adjacent images is 0.45 m.

The image acquisition is performed acquiring each image and then translating the camera to the following position (figure 1). Finally, the acquired images are processed by the ZScan Software, just defining the area of interest on the images and the dense matching step. The computational times are usually moderate as ZScan takes about 30-45 minutes per image triplet with a matching step of 3 pixels.

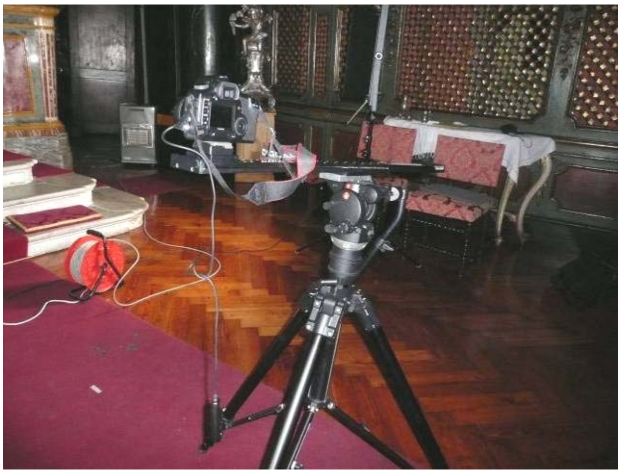

Figure 1. The ZScan calibrated bar

In our tests, a Canon EOS 5D camera was used as it guarantees good image definition. The technical characteristics of this camera are reported in table 1.

The ZScan software has been modified to also manage aerial images and, in a more general approach, each kind of normalized triplet of images. The absence of a y-parallaxes constitutes a constraint for the correct use of the software. It is well known that the relative orientation and the normalization of a set of images is an autonomous procedure in digital photogrammetry.

Camera	**f [mm]**	**Pixel dim. [mm]**	**Image dim. [pixel x pixel]**
Canon EOS 5D	25,10	0,0082	4368 x 2912

Table 1. Technical characteristics of the camera

5. POINT CLOUD PHOTOGRAMMETRY SURVEY: TERRESTRIAL CASE

It is well known that image matching gives different results, depending on the texture of the analysed object. For this reason, a first series of tests was performed analysing building facades characterised by different textures, in order to evaluate the strength of point cloud generation even on poorly textured facades.

Four different facades, made of different materials, were analysed in particular. In order to make an easy comparison between the texture of the facades possible, a coefficient was calculated for each facade. This coefficient was achieved considering the grey levels image and computing the standard deviation of 9X9 neighbouring pixels around each pixel, as shown in the following formula:

$$\sigma_{x,y} = \sqrt{\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(x_{ij}-\bar{x}\right)^2} \tag{1}$$

where N = dimension of the template: 9 pixels

$\sigma_{x,y}$= standard deviation of the central template pixel: N/2+0.5, N/2+0.5
x_{ij}= grey level value of the pixel
$\bar{x}$= mean grey level value of the 9X9 template

In this way a standard deviation value has been computed for each image pixel. The texture coefficient was finally obtained computing the mean of these values on the image. This value was considered representative of the minimum template dimensions in an Area Based Matching (ABM) approach.
The texture coefficient values of the tested facades are reported in the following table.

Façade type	Texture coefficient	
Coloured marble	15.195	
Painted wall	3.424	
Brick wall	8.126	
Rock wall	7.557	

Table 2. Texture coefficients and image examples

According to scientific literature (Kraus, 1993), the base-to-distance ratio greatly influences the precision of the generated points: in particular, considering the used camera and the maximum base achievable by the ZScan System, it is possible to define the precision that can be achieved at a certain distance. In order to test the point cloud geometric accuracy, several tests were performed varying this ratio from 1/4 to 1/18.

The main advantage of the ZScan System is given by the use of three images in DSM generation at the same time instead of two, as in other commercial software. In order to quantify the improvement of multi-image techniques, using the same images, a comparison was made between a ZScan point cloud and a DSM generated by LPS (Leica Photogrammetry Suite), using only two images (e.g. a traditional photogrammetric approach).
Another kind of test considered different rotations between the image plane (defined by the acquisition bar) and the facades. Finally, different matching steps were considered in order to compare the geometrical precision that could be achieved changing this parameter.

In order to define the precision of the ZScan System each point cloud generated during the tests was compared with reference surfaces acquired using a traditional laser scanner. In particular, a *Riegl LMS-Z420* laser scanner was used whose precision is of about ± 5 mm in range measures. This comparison was performed using a best fitting approach.

5.1 Results

Texture traditionally represents one of the most difficult issues in image-matching technologies. Furthermore, the use of more than two images has not appreciably improved the results: less-textured areas are difficult to model and are affected by noise.
The previously defined effect is clearly related to the texture coefficient: image regions characterized by a low texture coefficient value show large noisy areas, and vice-versa.

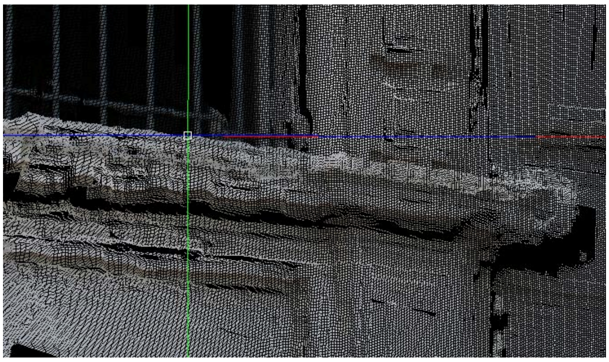

Figure 2. Noisy area on a painted wall

If the image triplet has a medium texture coefficient (that is to say from 6 to 10) the quality of the point cloud is already good. A generated point cloud is shown in figure 3.

The image triplet was acquired at a distance of 12 meters. The modelled façade has a texture coefficient of only about 8, but the geometrical details are correctly represented.

In order to define the change and the loss of geometrical precision in point cloud generation, several tests were carried out over this façade, increasing the taking distance by a metre at a time and using the maximum baseline on the calibrated bar. Images were acquired positioning the bar almost parallel to the surveyed façade. The generated point clouds were compared to laser scanner data in order to quantify the precision at each distance. The comparison, made at three different distances which were considered as representative for near, middle and far distances are described in table 3.

Figure 3. An example of a ZScan point cloud

As can be seen, the mean differences between the ZScan and the laser scanner data are approximately null, confirming the lack of systematic errors in the photogrammetric point clouds. Furthermore, the ZScan standard deviation is very similar to a theoretical RMS defined using the following formula (Kraus, 1993):

$$\sigma_y = \frac{Y^2}{c \cdot B} \sigma_{pi} \tag{2}$$

where c = focal length
Y = taking distance
B = base between the external images
σ_y= RMS in depth
σ_{pi} = RMS in image point position

A σ_{pi} equal to half a pixel was in particular considered.

Taking Distance [m]	**Δ mean [m]**	**Δ St. dev. [m]**	**Theoretical RMS [m]**
5	0.006	0.016	0.005
12	0.001	0.031	0.026
15	0.009	0.049	0.041

Table 3. Comparison between the ZScan and laser scanner data

Focusing on the residual distribution it is possible to notice that greatest differences are concentrated around the edges and where there are small details (figure 4). In particular, small façade details are smoothed by the matching process and corners are rounded off.

This problem can be explained in two different ways. The first one concerns the matching step which is lower in ZScan point clouds than in laser scanner data (resolution of 0.050 gon), in particular when the taking distance is increased. As an example, at a distance of 12 meters and with a matching step equal to 3 pixels, the points are 1.2 cm far from each other. It is easy to understand that small details are hard to model.

ZScan is also a area-based matching software. According to literature (Gruen, 1985), this technique considers the template and the search images as flat areas: consequently an area at the corner is considered as a flat part and the correct matching results is approximated with a smooth surface (figure 5). The same effects are present in a LIDAR point cloud as well.

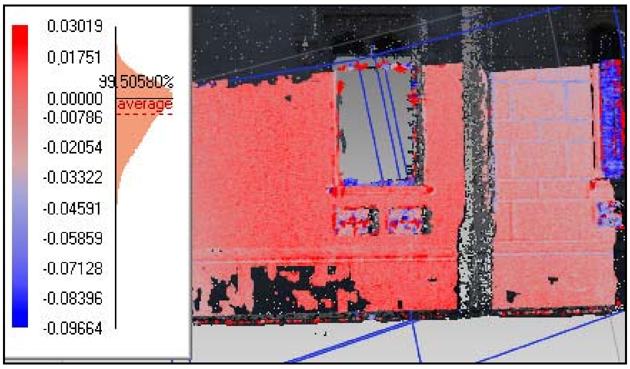

Figure 4. Residual distribution at a distance of 5 meters

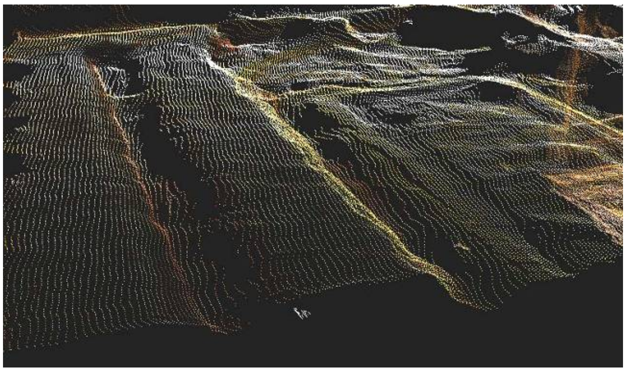

Figure 5. Smoothing in a point cloud

A traditional DSM generation has been performed using two of the triplet images (the two with the biggest baseline) and utilizing LPS digital photogrammetry software. Comparing the two DSMs (the one from two images and the one extracted using ZScan multi-image approach), it is possible to appreciate that a multi-image approach gives a better result, in terms of density and precision, than the traditional stereopair approach.

DSM generated	**# images**	**# extracted points**	**Δ mean [m]**	**Δ St. dev. [m]**
ZScan	3	723724	0.016	0.005
LPS	2	9718	0.022	0.613

Table 4. Comparison between the ZScan and LPS DSM

Table 4 shows the results of the comparison with the LIDAR point cloud for the images acquired with a 1/6 base-distance ratio.

6. POINT CLOUD PHOTOGRAMMETRIC SURVEY: AERIAL CASE

The ZScan System was originally realized for terrestrial applications, using a calibrated bar. The first test results using aerial images are presented in this paper.

An image triplet acquired over the city of Turin (Italy) was used. These images were acquired using a Leica RC 30 camera at a 1:8000 nominal scale in 2004 and they were later scanned at a 1200 dpi resolution. The achieved ground sample distance (GSD) was about 20 cm. The image overlap was 80% and the external orientation parameters were known from previous works: in some ways, these images can be considered as having been acquired by an extremely long calibrated bar.

The adopted matching step was of 5 pixel, that is, 1 pt/m^2 of resolution.

In order to define the achieved precision, the generated point cloud was compared with a DSM whose precision is known (± 10 cm). This DSM considers roofs as flat surfaces and does not describe any tree or vegetation, so an error in correspondence to the roof tops and trees can be expected.

An overview of the ZScan generated point cloud is presented in figure 6 and a zoom of the Olympic Stadium in Turin is shown in the top right panel. The point cloud seems to be complete and blank areas are limited to hidden regions.

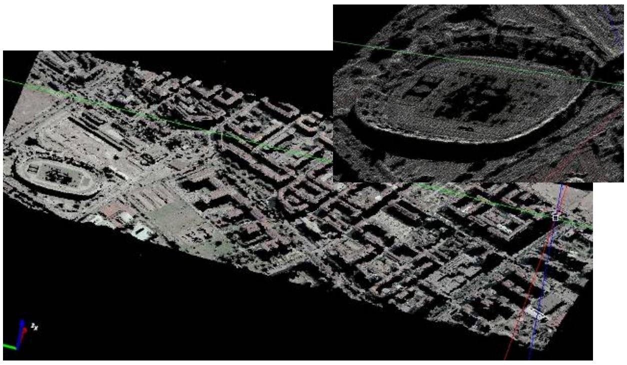

Figure 6. Point cloud from aerial images

Furthermore, a DSM was generated using two images a time in LPS software in order to estimate the improvement due to multi-image algorithms.

6.1 Results

In order to compare the generated point cloud and the reference DSM, the RSI ENVI commercial software was used. The coordinate system adopted in the comparison was the UTM WGS84 reference system.

The comparison results are shown in figure 7. As can be seen the ZScan systems offer a good precision over flat areas, as the differences are usually less than 1 m (green area). Furthermore, differences of between 1 m to 5 m (yellow area) are concentrated on roofs and on wooden areas but, as known, the reference DSM considers roofs as being flat and does not describe trees or vegetation.
As an example, the mean differences in the violet area are 0.31 m and standard deviation is 0.49 m: in other words, residuals have the same order of magnitude as GSD and they are comparable to the reference DSM precision.

However there are still prominent differences around buildings (red area), in correspondence to breaklines where the differences are usually above 5 m and holes in the point cloud are concentrated. The blank areas can only partially be explained by the shadows around buildings. The black panel (figure 6) shows residuals along blue lines; from this panel it is possible to clearly see the largest negative differences over building corners and small residuals in correspondence to pitched roofs. Negative differences mean that, in the same planimetric position, the point cloud defines points at the roof heights while, according to the reference DSM, they should be on the ground. In other words, the ZScan System has a sort of systematic error in correspondence to the breaklines (border of roofs): this problem is probably due to the lack of an efficient algorithm for the segmentation of the images.

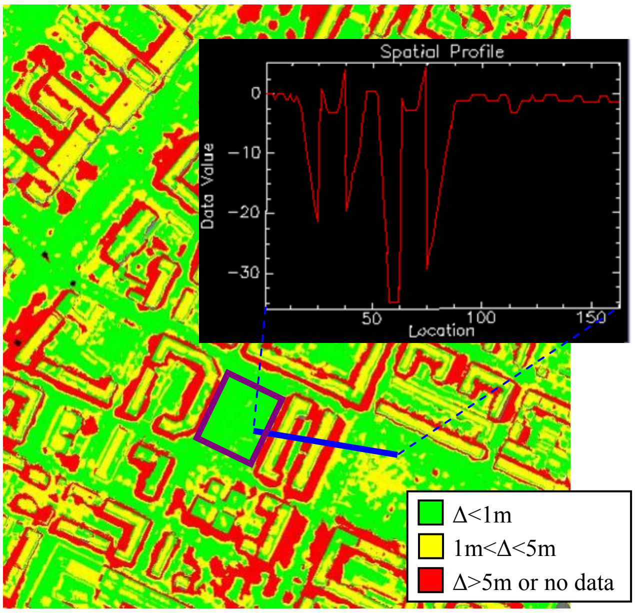

Figure 7. Comparison between the ZScan and DSM data

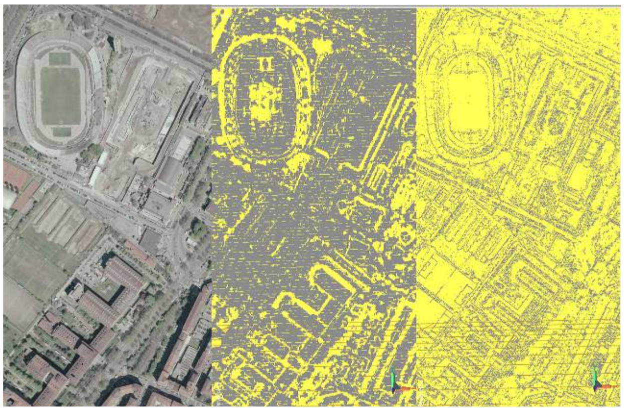

Figure 8. Source image, ZScan DSM and 2-image generated DSM

The DSM generated using only two images gave approximately the same results, in terms of geometrical precision. As in the ZScan point cloud building, the borders are hard to detect, but flat areas and roofs are correctly surveyed.
In spite of this, the number of detected points roughly decrease when using only a stereopair: there were 1138783 detected points in the ZScan triplet while, for the same area, the stereopair DSM detected only 432599 points. This difference

can clearly be seen in figure 8 where the points for the Olympic Stadium are compared.

7. CONCLUSIONS

Recalling the first aim of this paper (see paragraph 2), we can say that multi-image matching can produce point clouds which have the same quality, in term of density and precision, as the point clouds acquired by LIDAR acquisition units.

The result is valid both for terrestrial and aerial applications. It can therefore be stated that the decision to use photogrammetric or LIDAR systems to produce a point cloud is not influenced by precision, density nor by the time required to obtain the final product. The photogrammetric steps are completely automatic and can be run just after the image recording, thanks to the direct exterior orientation parameter measurements (by using GPS/IMU systems such as for LIDAR instruments).

As far as the generation pointed out in paragraph 3 is concerned, it is easy to see that photogrammetric point cloud generation offers a special advantage over the LIDAR approach. The same oriented images used to generate the point clouds can, in fact, be used to assist and improve the subsequent segmentation phase. LIDAR requires an “extra” image acquisition and orientation in order to obtain the same data as photogrammetry. This need for “extra” data obviously means higher costs, but we cannot overcome the problem that arise from the merging of data of from different sources.

Finally it can be stated that a multi-image matching can replace LIDAR techniques in DSM generation in many cases.

The presence of oriented and overlapping images and a point cloud suggests some future research topics. Breaklines are usually homogeneous radiometric region borders on images: an automatic edge extraction on oriented images can therefore be projected onto a point cloud in order to carry out the automatic segmentation approach in a more realistic and affordable way. In the same way automatic segmentation results can be verified and corrected by projecting them onto oriented images.

The Politecnico di Torino research group has started to investigate this original sharing of information between images and point clouds in order to reach a possible automatic, coherent and precise segmentation of a point cloud, which could represent a new step towards automatic and autonomous photogrammetric plotting.

REFERENCES

Beinat, A., Crosilla, F., Visintini, D., Sepic, F., 2007. Automatic non parametric procedures for terrestrial laser point clouds processing. In: *Proceedings of PIA07 Photogrammetric Image Analysis*, Munich, Germany September 19-21, 2007

Du, Li, Gu 2004. A Simple rectification method for linear multi-baseline stereovision system. In: *Journal of Zhejiang University Science*, ISSN 1009-3095, 2004 5(5): pp. 567-571.

El-Hakim, S., Remondino, F., Gonzo, L., Voltolini, F., 2007. Effective High resolution 3D geometric reconstruction of heritage and archeological sites from images. In *Proceedings of CAA 2007 Conference*, Berlin, Germany, April 2007.

Fassold, H., Danzl R., Schindler K., Bischof, H., 2004. Reconstruction of Archaeological Finds using Shape from Stereo and Shape from Shading In: *CVWW (Computer Vision Winter Workshop) 2004*, ISBN 961-90901-1-X, pp. 21-30, Piran, Slowenien, 4.-6.2.2004

Forstner, W., 1998. On the theoretical accuracy of multi-image matching, restoration and triangulation. In: *Festschrift zum 65. Geburtstag von Prof. Dr.-Ing. mult. G. Konecny*, 1998, Institut fur Photogrammetrie, Universitat Hannover

Gruen, A. W., 1985. Adaptive least square correlation: a powerful image matching technique. In: *South African Journal of Photogrammetry, Remote Sensing and Cartography*, Vol.14(3), pp.175-187.

Gruen A., Baltsavias E., 1988. Geometrically Constrained Multiphoto Matching, *Photogrammetric Engineering and Remote Sensing*, Vol.54, No.5, May 1988, pp.633-641.

Hartley R., Zisserman A., 2000, *Multiple view Geometry in Computer Vision,* Cambridge Press, UK, pp. 365-396.

Heuel, S., Forstner, W., 2001. Matching, Reconstructing and Grouping 3D Lines From Multiple Views Using Uncertain Projective Geometry. In: *Computer Society Conference on Computer Vision and Pattern Recognition 2001*, Kauai, HI, USA, ISBN 0-7695-1272-0

Kraus, K., 1993. *Photogrammetry*. Ferd Dummler Verlag, Bonn, Germany.

Remondino, F., 2007. Detailed image-based 3D geometric reconstruction of heritage objects. DGPF Tagunsband 16, pp. 483-492.

Roggero, M., 2002. Object segmentation with region growing and principal component analysis. In: *Proceedings ISPRS Commission III Symposium*, Graz, September 9-11 2002

Roncella, R., Forlani, G., Varini, A., 2005. Georeferenziazione fotogrammetrica automatica in applicazioni Mobile Mapping. In: *Bollettino Sifet*, Art.5, N°1/2005

Zhang H., Cech J., Sara R., Wu F., Zhanyi H, 2003. A linear Trinocular Rectification Method for Accurate Stereoscopic Matching. In: *BMVA proceedings*, 2003.

Zhang, L., Gruen, A., 2004. Automatic DSM extraction from linear array imagery data. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 35, Part B3, pp. 128-133, XXth ISPRS Congress, Istanbul, Turkey,12-23 July 2004.

Zhang, L., 2005. Automatic Digital Surface Model (DSM) generation from linear array images.Thesis Diss. ETH No. 16078, Technische Wissenschaften ETH Zurich, 2005, IGP Mitteilung N. 90

ACKNOWLEDGE MENTS

The authors would like to thank Menci Software s.r.l. for providing the assistance during tests.