The 11th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2010)
A. Artusi, M. Joly-Parvex, G. Lucet, A. Ribes, and D. Pitzalis (Editors) *Short and Project Papers*

Dense point cloud acquisition via stereo matching applied to: the Kilwa archaeological site and the Gallo-Roman theatre of Mandeure

J. F. Hullo[†1] , P. Grussenmeyer[1] , P. Assali[1] , E. Smigiel[1]

[1] The Image Sciences, Computer Sciences and Remote Sensing Laboratory UMR LSIIT 7005, INSA Strasbourg, France

Abstract

In this paper, we present a methodology for Dense Stereo Matching (DStM) acquisition and processing. Two main applications are detailed. In our application sites, many constraints led us to choose DStM. In the archaeological site of the Gallo-Roman Theatre of Mandeure (France), a initial terrestrial laser scanning campaign had to be completed with patches acquired from DStM because of the site geometric complexity. In the epigraphic and archaeological site of the region of Kilwa (Saudi Arabia), we applied DStM for both prospecting and excavating. We distinguished two kinds of applications in the paper: 1 to 5 stereopairs and 5 to 30 stereopairs. We inserted the data of both kind of cases into larger datasets. These datasets are currently used for archaeological reconstruction purposes of as-built buildings.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Digitizing and scanning—

1. Introduction

Even though the theory of Dense Stereo Matching (called DStM in our paper) from images has been known for a while, hardware of a standard workstation did not easily allow this kind of process. Recently, both camera and computer developments led to the creation of new tools. Computer vision and photogrammetry are working in parallel but with different objectives.

The Computer Vision community is rather working with low resolution images for 3D reconstitution, with non calibrated cameras and high level of automation. Accuracy is not the main aim. Literature on this subject has really increased in the last few years. Algorithms based on projective geometry are now straightforward and deal with a lot of different configurations. The photogrammetric approach is more mainly interested in precision and structured measurements than in automation.

[SvHvG*08] and [FP07] have shown that DStM can be used accurately with terrestrial high resolution images even if the camera's intrinsic parameters are not known. But very often, a photogrammetrist is able to calibrate his camera and can thus consider it as constant, reducing the correlation process time without a loss of accuracy.

Several matching algorithms based on epipolar constraints are compared in [SCD*06]. [Fas07] compares DStM and laser scanner technologies and it is now clear that DStM offers more than an alternative. Lately, several commercial solutions have been developed and initial experiments have been done for Cultural Heritage, e.g. [RP08].

In this paper we present an approach for the application of DStM on archaeological sites excavations as well as epigraphic investigation. The first part gives references to understand our approach and methodology, based on a previous theoretical study. It then leads to our applications and results. We conclude with an analysis of our results in line with our objectives and compared to other techniques. For the applications, PhotoModeler Scanner (PMSc) software from EOS Systems has been used. The camera used for image acquisition is a Canon EOS 5D mark I with three fixed optics (20 mm, 50 mm and 85mm). This solution for the acquisition of dense point clouds (hardware and software) is low cost in comparison to many others: about 5000 euros.

† Corresponding Author. jf[at]hullo.fr

DOI: 10.2312/PE/VAST/VAST10S/067-070

2. Theory and frameworks

Understanding the different steps of the algorithm has enabled to list the kind of the nature of errors, and their origins. Afterwards, it is possible to define a methodology of acquisition and processing trying to minimize them. The main idea of DStM is to use stereoscopic relationships and pixel correlation process. The theoretical aspects of DStM are detailed in [KE01] and [HZ04]. Distortions and calibration are studied in [Zha96] and [TB04]. A synthesis and analysis has been presented in [HGF09].

Many error sources are taken into account when defining the framework: optical distortions, weak geometries (for both bundle block adjustment and stereo correlation, recovery), unfavourable radiometry or lightning conditions, blurry images. This leads to precisely define bad cases for base over height ratio, network geometry, object illumination, focal length, aperture. One of the main advantage of PMSc is that a project can be driven from the camera calibration to the export of textured mesh. Four steps are considered: data acquisition, orientation of the images, Dense Stereo Matching, post-processing. The acquisition and processing methodology are presented and explained in [HGF09].

3. Experiments

3.1. Simple applications based on a limited number of stereopairs

First of all, upstream of the image acquisition on the site is calibrated the camera (body and optics). The knowledge of the intrinsic parameters of the camera (sensor size and resolution, principal point coordinates, focal length, optical distortions) is directly linked with the accuracy of the final results. In PMSc, the distortion model is based on a two polynomial model respectively for radial and tangential distortions. It is also recommended to acquire always 3 images, in order to choose 2 out of 3, to select the optimal stereo or to compare light.

Figure 1: *Example of a stereopair of an arabic inscription (Kilwa, Saudi Arabia).*

Then, the first step of the computation is the orientation of the couple. The best couple for DStM computation is chosen from the position and orientation of the cameras. In this step, epipolar geometry is (implicitly) calculated. Thereafter, we will prepare the data set for matching. Firstly, we create two new images, free of distortions, by applying the inverse distortion equations. Once we create those new « idealized » images, and corresponding camera, epipolar lines are straights, but not parallel. Consequently, the epipolar rectification is the next step. In PMSc, this epipolar rectification is made during the DStM computation, after the setting of the sampling rate. Two bitmap images are created in the project folder. Then, after adjusting parameters (depth range above and below global model, sub pixel factor, matching region radius and texture type), matching is processed. Both images are used as reference in order to minimize wrong matching and two disparity maps are computed as bitmap files in the project folder too. Then 3D points are triangulated and the point cloud is created.

Top left: meshes are partially shown.

Figure 2: *Final 3D textured model of the inscription (Kilwa, Saudi Arabia).*

Acquisition time for a single stereopair is less than one minute. The time required for the process is about one hour, from data acquisition to export, for a single stereopair processed on recent workstations or laptops; the operator only has to work around fifteen minutes for the creation of the model. The average generated number of points for a stereopair was 150 000 per stereopair. Thereafter, the point cloud can be denoised, segmented, meshed and textured. 3D data can be exported in many standard 3D formats. In this project, we used both OBJ and VRML. OBJ data has been converted into a 3D PDF (*.u3d) that allows measuring and navigation, faster and easier than VRML explorers (such as Cortona for instance).

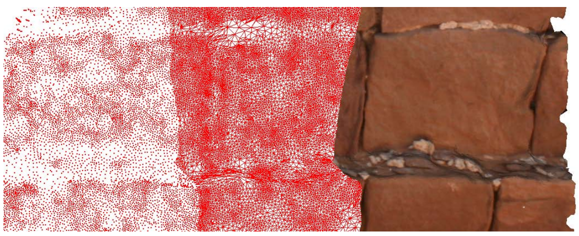


Figure 3: *Dots, wireframe and textured model of a part of a wall.*

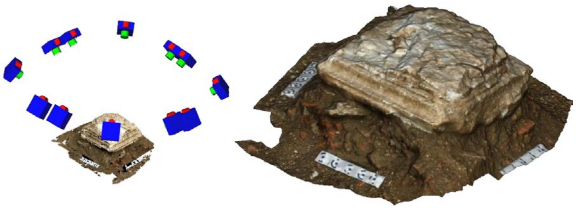


Figure 4: *Camera positions around the object and resulting textured model.*

3.2. Complex application based on a large set of stereopairs

Over 10 stereopairs, the bundle adjusment may need additional information to reach the position required. In this instance, two solutions are suggested: either to insert control points surveyed by total station, or to register local generated point clouds on laser scanner data. The acquisition method must be fitted in order to ensure data fusion.

For the buildings of Kilwa, a single point of view does not ensure a total coverage of the object. Thus, stereopairs need to be multiplied in order to model the details of the whole building (Figure 5). But to compute a bundle adjustment of the block, additional images are needed. We followed the 3x3 CIPA rules `http://cipa.icomos.org`. Many targets were placed all around the building (for the automation of the image orientation), and finally 50 images were used in the project. To scale the block, 20 control points measured by total station were used. The acquisition time was about 4 hours. 2 millions of points were generated. The local resolution was in the range of 2 to 6 mm. The tolerance after global registration was about 10 to 30 mm. The size of the building is around 7m by 7m, with wall elevations of 2m. The estimated time for processing by a single operator is 30 hours.

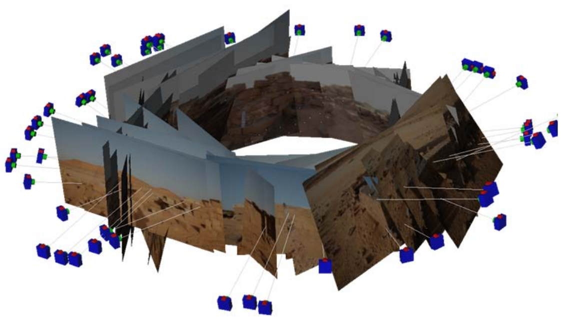


Figure 5: *Photogrammetric block of the external part of the building.*

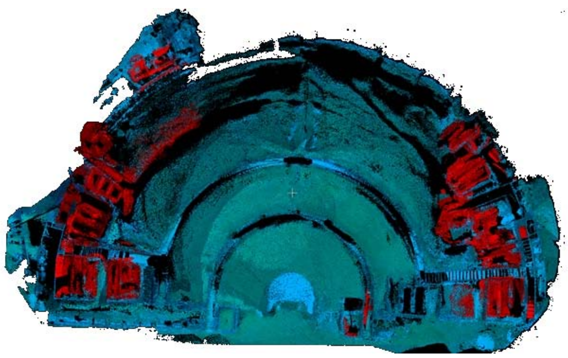
After the block orientation, the best stereopairs were chosen to apply DStM separately for each of them. Parameters were implemented differently for each couple, depending on

Figure 6: *3D model of the building (Kilwa, Saudi Arabia).*

the size, the illumination and a few other criteria. After denoising the whole model, we merged all the point clouds via the Merge function of PMSc. Thereafter, we processed the merged point cloud classically: meshing, filling holes in meshes and smoothing in PMSc. It is an important fact, because in most 3D works, several programs must be used for the different parts of the process, increasing cost, but also increasing complexity of the work flow.

After the export, we firstly converted the model to a 3D PDF file, and merged the 3D model with a Digital Elevation Model of the site. Because of the buildings function (settling tanks), we wanted to model the hydrometric network, first at the site scale, and more precisely at the building scale. Since water resources remain the main problem in desert areas, this work reveals a lot about the organisation of the monastery and its environment.

For the archaeological site of the theatre of Mandeure (France), the aim was to complete holes from the terrestrial laser scanner data set. Many parts of the theatre were indeed not accessible to the laser scanner. Therefore, we improved the 3D model with point clouds from DStM.



Points clouds (in red) based on DStM are combined with TLS data (green-blue).

Figure 7: *Complete 3D point cloud of the theater of Mandeure (France).*

450 images were used for the DStM part. The total acquisition time was about 20 hours. 14 millions of points were generated by DStM and 7 millions by laser scanner. The local resolution was about 5 to 10 mm. The tolerance after global registration was about 20 to 45 mm. The estimated time for processing by a single operator is 80 hours. Each

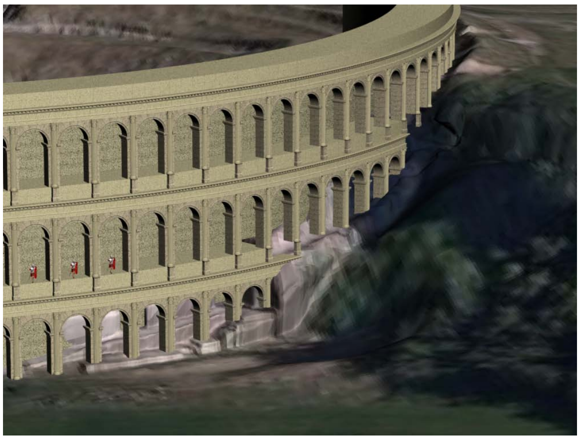


Figure 8: *view of a hypothesis of as-built reconstruction of the Gallo-Roman theater of Mandeure on the Digital Elevation Model surveyed from DStM and laser scanning.*

stereopair was generated through PMSc, but the point clouds were registered in another software.

The generated data is currently used by architects and archaeologists for historical investigation. Figure 8 shows a guess-work based on the point cloud of the theatre. The numbers of places of the theatre has also been estimated as 18 000.

4. Conclusion

Even though commercial solutions for Dense Stereo Matching only appeared a few years ago, they offer surprisingly good results on various points. Firstly, it is possible to reach high precision for both geometric and rendering aspects. Secondly, photogrammetry preserves through DStM all the flexibility and cost advantages, with dense point clouds as output. Laser scanner techniques only offered these advantages recently. Finally, data acquisition for the single stereopair case does not require a high level of knowledge in photogrammetry; with a few recommendations on good practice, many people such as archaeologists and epigraphists would be able to take photographs in the right way. Nevertheless, due to the texture properties of the object, this passive stereo solution cannot yet be applied to particularly self coloured or reflective materials. Passive stereo techniques (pattern or grid projection on the object) make this kind of objects measurable but are difficult to implement outdoor. For complex cases, the data fusion could not be done into PMSc. Image acquisition and processing for large projects (more than 5 stereopairs) requires good knowledges in photogrammetry, surveying and geometry.

A comparison with other acquisition techniques has also been realized. Firstly, the man-made sketches of archaeological materials is now not required for each object, since 3D high-quality textured models are generated. Secondly, laser scanner has still many disadvantages for many archaeological missions. The drawbacks are mainly the cost, the complexity of transport, of use and post-processing, and the fragility. Moreover, laser scanner are scale dependent whereas DStM can measure small objects as well as large environments. Users interested in both geometrical and visual properties of an object will find in DStM an easy, cheap, fast and accurate solution for documentation.

The many advantages and developments will certainly help the Cultural Heritage community to integrate this technology, and to take two parallel images (at least) instead of one to get the 3^{rd} dimension more easily.

Acknowledgements

We would like to thank Saba Fares and Jean-Yves Marc for permission to work on respectively Kilwa and Mandeure sites.

References

[Fas07] FASSI F.: 3d modeling of complex architecture integrating different techniques - a critical overview. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI-5/W47* (2007), 11p. 1

[FP07] FURUKAWA Y., PONCE J.: Accurate, dense and robust multi view stereopsis. *In Proc. Int'l Conf. on Computer Vision and Pattern Recognition* (2007), 1–8. 1

[HGF09] HULLO J.-F., GRUSSENMEYER P., FARES S.: Photogrammetry and dense stereo matching approach applied to the documentation of the cultural heritage site of kilwa, saudi arabia. *CIPA International Archives for Documentation of Cultural Heritage XXII* (October 2009), 6. (CIPA'09) `http://cipa.icomos.org/textfiles/KYOTO/132-1.pdf`. 2

[HZ04] HARTLEY R., ZISSERMAN A.: *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004. ISBN 0521-54051-8. 2

[KE01] KASSER M., EGELS Y.: *Digital Photogrammetry*. Taylor & Francis, 2001. ISBN 0-7484-0945-9. 2

[RP08] REZNICEK J., PAVELKA K.: New low-cost 3d scanning techniques for cultural heritage documentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVII-B Part B5* (2008), 237–240. 1

[SCD*06] SEITZ M., CURLESS B., DIEBEL J., SCHARTEIN D., SZELISKI R.: A comparison and evaluation of multi view stereo reconstruction algorithms. *Proc. Int'l Conf. on Computer Vision and Pattern Recognition 1* (2006), 519–528. 1

[SvHvG*08] STRECHA C., VON HAUSEN W., VAN GOOL J., FUA P., THOENNESSEN U.: On benchmarking camera calibration and multi-view stereo for high resolution imagery. *Proc. IntŠ1 Conf. on Computer Vision and Pattern Recognition* (2008), 1–8. 1

[TB04] TEBOURBI R., BELHADJ Z.: Rectification épipolaire d'images stéréoscopiques. *International Symposium on Image/Video Communications over Fixed and Mobile Networks, ISIVC'04* (2004). 2

[Zha96] ZHANG Z.: On the epipolar geometry between two images with lens distortion. *Proc. Int'l Conf. Pattern Recognition 1* (1996), 407–411. 2