PFG 2012 / 6, 0727–0740
Stuttgart, December 2012

Report

Automatic 3D Object Reconstruction from Multiple Images for Architectural, Cultural Heritage and Archaeological Applications Using Open-Source Software and Web Services

THOMAS P. KERSTEN & MAREN LINDSTAEDT, Hamburg

Keywords: 3D, automation, comparison, image matching, modelling, point cloud, reconstruction

Summary: Constant improvements in the performance of internet and computer technologies combined with rapid advancements in computer vision algorithms now make it possible to efficiently and flexibly reconstruct the 3D geometry of objects. Objects of different sizes can be modelled using image sequences from commercial digital cameras that are processed by web services and freely available software packages, forming low-cost systems for numerous applications (restoration, historical care of monuments, visualization, analysis of the state of construction and the damage, etc.). In this contribution various cultural objects (historical buildings, statues/figures, archaeological finds, etc.) have been reconstructed in order to investigate the potential of this technology which enables the automatic generation of 3D point clouds or surface models (as 3D polygons) with photo-realistic texture from image data. These so-called low-cost systems represent an efficient alternative to expensive terrestrial laser scanning systems for the as-built documentation of 3D objects in architecture, cultural heritage and archaeology. The accuracy of the automatically generated 3D models is assessed by comparison with results from terrestrial laser scanning.

Zusammenfassung: *Automatische 3D-Objektrekonstruktion aus digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie durch open-source Software und Webservices.* Durch die stetig zunehmende Leistungsfähigkeit des Internets und der Computertechnologie sowie der raschen Weiterentwicklung von Computer Vision Algorithmen ist es heute möglich, die 3D-Geometrie von Objekten unterschiedlicher Größe mit handelsüblichen digitalen Kameras als Low-Cost-Systeme für zahlreiche Anwendungen (Restaurierung, historische Denkmalpflege, Visualisierung, Analyse des Bauzustandes und der Beschädigung, etc.) effizient und flexibel in Bildsequenzen zu erfassen. Anhand von diversen Kulturobjekten (historische Gebäude, Statuen/Figuren, archäologische Fundstücke, etc.) wird in diesem Beitrag das Potential von Webservices und frei verfügbaren Softwarepaketen aufgezeigt, mit denen 3D-Punktwolken oder Oberflächenmodelle (als 3D-Polygone) mit fotorealistischer Textur automatisch aus Bilddaten erzeugt werden. Diese so genannten Low-Cost-Systeme stellen heute für die As-Built-Dokumentation von 3D-Objekten in Architektur, Denkmalpflege und Archäologie eine effiziente Alternative zu teuren terrestrischen Laserscanningsystemen dar. Die Genauigkeit der automatisch erzeugten 3D-Modelle wird durch den Vergleich mit Ergebnissen des terrestrischen Laserscannings aufgezeigt.

1 Introduction

As state-of-the-art geodetic measuring methods photogrammetric multi-image techniques and, increasingly, terrestrial laser scanning, as a standalone system or in combination with other methods, are used for precise 3D data acquisition of complex objects. Requirements for the generation of 3D models are often very high with respect to level-of-detail, complete-

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0152 1432-8364/12/0152 $ 4.00

ness, reliability, accuracy (geometrical and visual quality), efficiency, data volume, costs and operational aspects, but the priority order depends upon the object to be recorded.

However, in recent years, real alternatives to classical systems and methods are presented by the large number of digital cameras on the market, which can be efficiently and successfully used as passive low-cost sensors when combined with appropriate algorithms such as structure-from-motion (SfM) and/or dense image matching for different 3D applications (object reconstruction, navigation, mapping, tracking, recognition, gaming, etc.). Due to the very low costs and current approval for open-source methods such systems (sensors in combination with appropriate algorithms) are very popular in many application fields. Nevertheless, the metrological aspect should not be neglected, if these systems are to be acknowledged as serious measuring and modelling procedures. Therefore, clear statements about the accuracy potential and efficiency of such systems must be empirically investigated through appropriate testing. In this context the 3D modelling results must be also analysed and compared with respect to reference data.

Practical examples of image-based modelling have been reported by Kersten et al. (2004), Kersten (2006) and Remondino & Menna (2008), while Remondino et al. (2008) generate comparable results for the 3D documentation of cultural monuments using image-based and range-based procedures in comparison. Barazzetti et al. (2009) present the combined use of photogrammetric and computer vision procedures for automatic and exact 3D modelling of terrestrial objects. They also show that similar results can be achieved with image-based and range-based recording systems. Barthelsen et al. (2012) present an approach for detailed and precise automatic dense 3D reconstruction of urban scenes using possibly unordered image sets from consumer cameras on a small unmanned aerial system.

In this paper the potential of web services and freely available software packages is investigated on the basis of practical examples where 3D point clouds or surface models (as 3D polygons) with photo-realistic texture are automatically derived from image data. After a brief introduction to the applied software in section 2 the entire workflow for the image-based low-cost 3D reconstruction procedure (section 3) is outlined. Practical results and 3D comparisons with reference data are summarized in section 4.

2 Applied Software

For investigation of the automatic generation of 3D point clouds and 3D surface models from image data the following software packages and/or web services were used: Bundler/PMVS2 and VisualSFM (open-source software), Microsoft Photosynth (web service), and Autodesk Photofly and/or 123D Catch Beta (web service).

2.1 *Bundler/PMVS2*

Bundler (Snavely et al. 2008) and PMVS2 (patch based multi view stereo software, Furukawa & Ponce 2010) were developed at the University of Washington in Seattle (USA) in C and C++ under the GNU General Public License as freely available software. Bundler works as a structure-from-motion (SfM) procedure for arbitrarily arranged imagery and

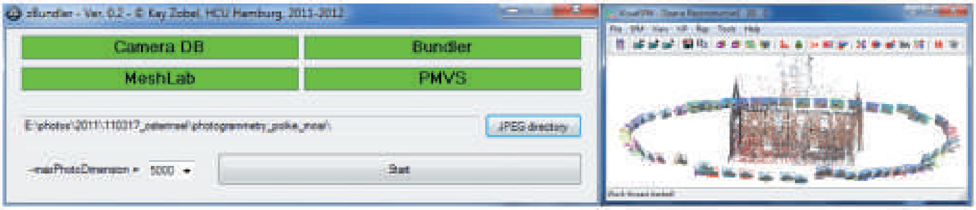

Fig. 1: Left: HCU graphical user interface (GUI) for the automatic workflow of Bundler/PMVS2; right: GUI for accelerated data processing with VisualSFM.

was developed for Microsoft's Photo Tourism Project (Snavely et al. 2006). Feature extraction in the images is performed by the SIFT (scale invariant feature transform) algorithm from Lowe (2004). The software supports camera calibration data (focal length f from EXIF data, two radial distortion parameters k1 and k2), image orientations and a thin 3D point cloud (scene geometry) as results for any image blocks using a modified bundle block adjustment from Lourakis & Argyros (2004). The results of Bundler are used in PMVS2 in order to generate a denser point cloud of non-moving objects by dense image matching. As well as the 3D coordinate each point additionally receives the colour value of the object taken from the images.

For use at HCU Hamburg a graphical user interface (Fig. 1, left) was developed which provides automation in the workflow between necessary software elements. As an example, after input of the images Bundler and PMVS2 are executed automatically and the result is finally presented in MeshLab. MeshLab is an open-source, portable, and extensible system for the processing and editing of unstructured 3D triangular meshes developed at the Visual Computing Lab, which is an institute of the National Research Council of Italy in Pisa (Cignoni et al. 2008).

2.2 *VisualSFM*

VisualSFM is a GUI application of multicore-accelerated SfM, which was developed at the University of Washington (Wu 2007). The software is a re-implementation of the SfM system of the Photo Tourism Project and it includes improvements by integrating both SIFT on the graphics processing unit (SiftGPU) and multicore bundle adjustment (Wu 2011). The camera parameters are defined as follows. The focal lengths (in pixels) of the camera are automatically calculated according to EXIF data. The principal point for each image is assumed to be at image centre except when using a single fixed calibration and the software uses only one radial distortion value. Dense reconstruction can also be performed through VisualSFM using PMVS/CMVS (patch or cluster based multi view stereo software, Furukawa & Ponce 2010).

2.3 *Microsoft Photosynth*

Photosynth has been developed for the Photo Tourism Project in co-operation between Microsoft Live Labs and the University of Washington (Snavely et al. 2006). The program Bundler forms the basis for automatic generation of 3D point clouds by free data processing using an external web service. For the use of Microsoft Photosynth a program for uploading the photos to a server and a Windows Live ID, e.g. email address, is initially necessary (Photosynth 2012). Depending upon the number of photos a result can be viewed in some seconds or a few minutes later online in all usual browsers; however no access to the data is possible. Only using the external program SynthExport from Hausner (2010), the computed 3D points and the camera parameters can be exported. However, the results correspond to those from Bundler. Nevertheless, Pomaska (2009) demonstrated how a photorealistic 3D model with true scale can be generated from imagery based on a low level polygon mesh and UV texture mapping using Photosynth for point cloud generation.

2.4 *Autodesk Photofly/123D Catch Beta*

In summer 2010 Autodesk introduced the project Photofly, a free web service, which allows the user to derive a meshed 3D model automatically from at least five overlapping photos of the recorded object (Abate et al. 2011). The basis of Photofly is the program smart3Dcapture from the French company acute3D in Nice. Photofly uses algorithms from computer vision and photogrammetry and the performance of cloud computing in order to provide 3D models efficiently from 2D imagery. The fundamental algorithms of this software are described by Courchay et al. (2010), while detailed information about the algorithms used are not available from Autodesk.

Photofly uses the "Autodesk Photo Scene Editor", which must be installed on the user's computer as a communication platform between users and server. The very clear user interface of this software transfers the selected photos to the server. Depending upon the object complexity a 3D model will be provided for download within a short time period, i.e. usually in some minutes. This can be processed further within the software. Important functions in the software include selecting parts of the triangle meshes, navigation options, selecting of points and the definition of a reference distance for the absolute scaling of the model. Individual photos can be inserted into the model after the initial modelling by measurement of identical points (tie points). For the generation of 3D models three quality levels, mobile, standard and maximum (optimal result), are available. The results can be exported to different formats, e.g. OBJ or LAS. In November 2011 Photofly was replaced by 123D Catch Beta after the company acute3D has presented its software smart3D-capture to the public in October 2011. Detailed information about 123D Catch and related tutorials are available under 123D (2012).

3 Workflow

The general workflow for image-based 3D reconstruction using low-cost systems is illustrated in Fig. 2 to document the degree of automation of the individual procedures which is symbolised in colour (red = manual, yellow = interactive and green = automatic).

For photogrammetric object recording multiple photos are taken of the object from different positions, whereby coverage of common object parts should be available from at least three but preferably five photographs from different camera positions. After import of the images into the respective processing software the parameters for camera calibration (interior orientation) and (exterior) image orientations are automatically computed. The subsequent generation of 3D point clouds or 3D surface models is also carried out in full automatic mode. Only for the 3D transformation of the point cloud or the meshed model into a superordinate coordinate system must the user measure control points interactively. The derived 3D model is automatically textured using the original image data so that video sequences, e.g. in 123D Catch Beta, can be generated using these models. If a CAD model needs to be constructed from a transformed and geo-referenced, coloured 3D point cloud, manual processing has to be carried out in a CAD program, e.g. AutoCAD. This 3D CAD model can later be manually textured in visualisation software, e.g. 3D studio, Cinema4D, Maya, etc., using the digital photographs in order to produce visualisations and/or video sequences.

4 Results and 3D Comparisons

In this section the results of the applied software packages Microsoft Photosynth, Bundler/PMVS2, VisualSFM and Autodesk Photofly and/or 123D Catch Beta, respectively,

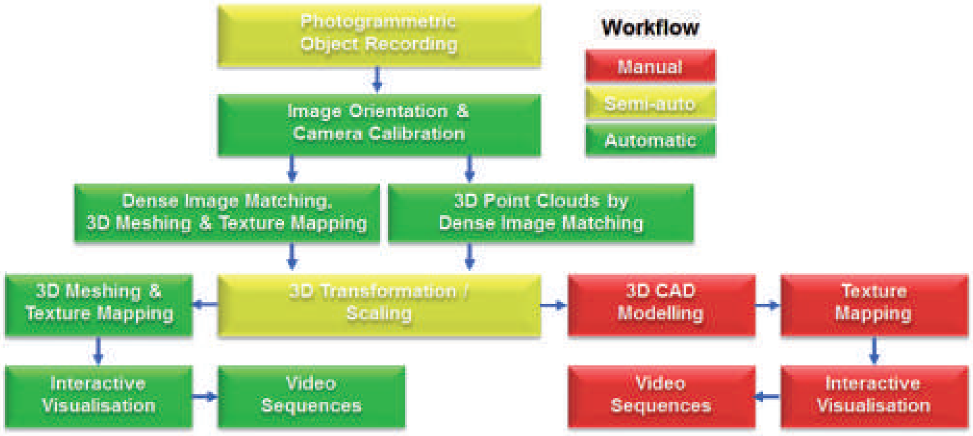

Fig. 2: Workflow for image-based low-cost 3D object reconstruction procedures.

Tab. 1: Statistics on different generated 3D objects using image-based systems, Pho. = photos.

Object	Camera/lens (mm)	Pixel	#Pho.	Software/service	# Points	# Triangle
Town house	Nikon D90 / 20	4288x2848	19	Photosynth	20,237	-
Town house	Nikon D90 / 20	4288x2848	19	Photofly	272,350	515,442
Town house	Nikon D90 / 20	4288x2848	19	Bundler/PMVS2	1,016,874	895,986
Fire House	Nikon D90 / 18	4288x2848	66	123D Catch	176,919	352,091
Fire House	Nikon D90 / 18	4288x2848	66	Bundler/PMVS2	1,541,924	-
Fire House	Nikon D90 / 18	4288x2848	66	VisualSFM	1,167,906	-
Zwinger	Nikon D90 / 28	4288x2848	15	Photosynth	18,553	-
Zwinger	Nikon D90 / 28	4288x2848	15	Photofly	155,697	285,669
Zwinger	Nikon D90 / 28	4288x2848	15	Bundler/PMVS2	917,965	-
Lion	Nikon D90 / 20	4288x2848	39	123D Catch	344,679	686,285
Lion	Nikon D90 / 20	4288x2848	39	Bundler/PMVS2	1,373,712	2,669,244
Moai Poike	Nikon D70 / 35	3008x2000	27	123D Catch	85,092	169,131
Moai Poike	Nikon D70 / 35	3008x2000	27	Bundler/PMVS2	629,644	-
Stone	Optio X, D70/80	diverse	48	123D Catch	-	291,613
Stone	Optio X, D70/80	diverse	48	Bundler/PMVS2	214,940	-
Pottery	Nikon D40 / 34	3008x2000	80	Bundler/PMVS2	-	323,402
Fragment	Nikon D90 / 24	4288x2848	58	VisualSFM	-	1,135,284

using the standard parameter values of each software package are presented for applications in architecture, cultural heritage and archaeology (Tab. 1), whereby individually generated datasets were compared with reference data from Zoller + Fröhlich's IMAGER 5006h and IMAGER 5010 terrestrial laser scanners (Tab. 2).

4.1 *Applications in Architecture*

Fig. 3 shows the results for the front façade of the Old Segeberger Town House (Bürgerhaus) – one of the oldest existing buildings in Schleswig-Holstein dated from the year 1539 with the front façade from 1606 – which are generated by Photosynth, Bundler/PMVS2 and Photofly. The 19 photographs used were acquired with a Nikon D90 (Nikkor 20 mm lens) from different positions as part of a student project at the HCU Hamburg. The result from Photosynth (20,237 points) is not suitable for façade modelling, while Photofly with 272,350 points and 515,442 triangles supplied the most attractive visual result. However, at the sides of the front of the façade and at the roof edges so-called virtual points, which do not exist in reality (Fig. 3, right), were meshed with Photofly. With Bundler/PMVS2 a dense point cloud of 1,016,874 points was generated (Fig. 3, centre right).

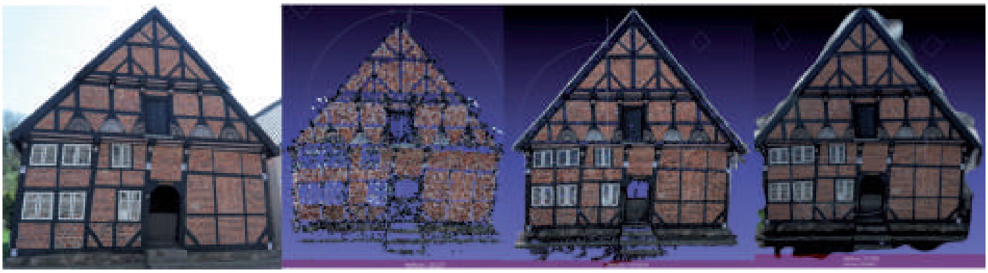

Fig. 3: Front façade of the old Town House (Bürgerhaus) in Bad Segeberg (19 photos from Nikon D90 with 20 mm lens); left: original photo; centre left: point cloud from Photosynth; centre right: Bundler/PMVS2; right: 3D meshing from Photofly.

The 3D meshing for the data from Bundler/PMVS2 of the front façade was carried out using Geo-magic Studio 12 and 2012. Geomagic Studio transforms 3D point clouds from range- or image-based systems and polygon meshes into accurate, usable 3D digital models for advanced product design, reverse engineering, custom manufacturing, CAD and analysis (Geomagic 2012). The meshed 3D front façade of Bundler/PMVS2 (from 5 mm grid spacing) and Photofly is visualized in Fig. 4 in comparison with the reference data of the IMAGER 5006h. According to the technical specifications of the system manufacturer the range noise of this scanner (1 sigma) for unfiltered raw data at 25 m distance is 2.6 mm for a reflectivity of 10% (black), 1.5 mm for a reflectivity of 20% (dark grey), and 0.7 mm for a reflectivity of 100% (white) using a scanning rate of 127,000 pixel/s (high power mode). Some areas at the black timber framework could not be measured with Bundler/PMVS2 or with the scanner thus small gaps are visible (Fig. 4, left). The triangle meshing of Photofly shows a noisy front with distinctive artefacts at the edges (Fig. 4, centre) while gaps were simply closed.

The absolute scaling of the data from Bundler/PMVS2 and Photofly was performed using measurements of two photogrammetric control points on the façade and the distance between them. Prior to 3D comparison both datasets were registered to the laser scanning data by an iterative closest point algorithm (ICP) (Besl & McKay 1992) in Geomagic. In the 3D comparison of the meshes from image-

Fig. 4: 3D meshing of the front façade from the Town House in Bad Segeberg; left: Bundler/PMVS2; centre: Photofly; right: IMAGER 5006h.

Fig. 5: Registration to the reference data of IMAGER 5006h (green < 3 mm); left: Bundler/PMVS2; right: Photofly.

Fig. 6: Old Fire Brigade House located in the village Kirkeby of the Danish island Rømø; left: Meshed 3D model from 123D Catch Beta; centre: point cloud from VisualSFM; right: Bundler/PMVS2.

and range-based systems Bundler/PMVS2 clearly shows a better result than Photofly (Fig. 5) because most differences are smaller than 3 mm (see green areas). Due to the significantly higher number of 3D points obtained a geometrical better result was expected for Bundler/PMVS2 than for Photofly which is also documented in Tab. 2 through the lower number for the average deviation and the standard deviation.

In order to compare computation times an image block of 66 photographs with the Nikon D90 (Nikkor 18 mm lens) of the Old Fire Brigade House located in the village Kirkeby of the Danish island Rømø (Fig. 6) was computed with Bundler/PMVS2 and VisualSFM. The notebook used for processing consisted of an Intel Core i7 CPU Q740 processor with 1.73 GHz, an internal memory of 16 GB RAM running on the operating system Windows 7 Enterprise (64 bit), and a NVIDIA graphic card GeForce GT 445M. Computation time for Bundler/PMVS2 was 24 hours supplying 1,541,924 points, while the computing time of VisualSFM was only 42 minutes, which corresponds to a factor of 33 times faster than Bundler. However, approximately 374,000 points (24%) fewer were measured and these were particularly evident in the gaps within the shadow and roof areas (Fig. 6). Similar results (factor 24 faster, 16% fewer points) were obtained with an image block of 44 images (Nikon D90, 20 mm lens) from the former Swedish toll house in the old harbour of the city of Wismar (Kersten et al. 2012). In a further test the meshed model of 123D Catch Beta was computed in 16 min using the web service. This result looked visually better than the other two but the edges are smoothed and there are some virtual (geometrically incorrect) points at the antennas and on the top roof areas.

4.2 *Applications in Cultural Heritage*

As an example of applications in cultural heritage documentation the results of a figure from the Zwinger palace in Dresden are illustrated in Fig. 7. The figure was acquired from two different heights in 15 photographs using a Nikon D90 with a Nikkor zoom lens (focal length 28 mm). Photofly could generate a visually-appealing, near-complete mesh with 285,669 triangles from 155,697 points (Fig. 7, left), while with Bundler/PMVS2, despite measuring 917,965 points, gaps were visible (Fig. 7, centre) and in Photosynth a very small point density was measured with only 18,553 points (Fig. 7, right). Unfortunately, no reference data was available for this figure meaning that no geometrical analysis could be carried out.

The second example for applications in cultural heritage documentation is the 3D comparison of different results for the lion figure at the entrance of the imperial church (Kaiserdom) in Königlutter as illustrated in Figs. 9 and 10. The lion was acquired by image and range-based methods, i.e. in 39 photographs from different heights around the object using a Nikon D90 with a Nikkor lens (focal length 20 mm) and from two close scan stations using the laser scanner IMAGER 5006h. 123D Catch Beta could generate a visually appealing, almost complete meshed model with 686,285 triangles from 344,679 points (Fig. 8, top left), while with Bundler/PMVS2 despite 2,433,077 points, computed in 9 hours and 51 minutes (15 min per photo), some gaps are visible (Fig. 9, bottom left). The two registered clouds of the laser scanner which do not represent full coverage of the lion, were reduced to 660,000 points in total with a grid spacing of 2 mm due to the huge data volumes, which yielded a meshed model of 1,314,603 triangles. The 3D comparison of the meshed models from Bundler/PMVS2, also reduced to 2 mm grid spacing, and from 123D Catch Beta with the reference data was carried out in Geomagic after ICP registration of each model to the reference data. The results in Figs. 9 and 10 indicate deviations of ±2 mm to the reference for the most parts of the meshed models of the figure, while the average deviations are in the range of ±20 mm (Tab. 2) due to higher deviations at the wall and on the ground. The 3D comparison of 123D Catch Beta vs. IMAGER 5006h shows higher deviations especially in areas with higher curvature and edges due to smoothing effects, while the comparison of Bundler/PMVS2 vs. IMAGER 5006h demonstrates some systematic deviations on the front part of the lion.

4.3 *Applications in Archaeology*

Photofly and/or 123D Catch Beta could generate 3D models that are quite attractive visually when the archaeological objects are small and relatively rotund. A moai at the volcano crater Poike on Easter Island (Chile) was acquired in 27 surrounding photos with a Nikon D70 (Nikkor zoom lens with 35 mm). The result from 123D Catch Beta provided an almost-complete, textured 3D model (Fig. 10) with 169,131 triangles (from 85,092 points) while with Bundler/PMVS2 despite measurement of nearly 630,000 points, some gaps at the neck and at the bottom of the figure were visible.

Fig. 7: Figure in the Zwinger palace of Dresden; left: Autodesk Photofly; centre: Bundler/PMVS2; right: Microsoft Photosynth.

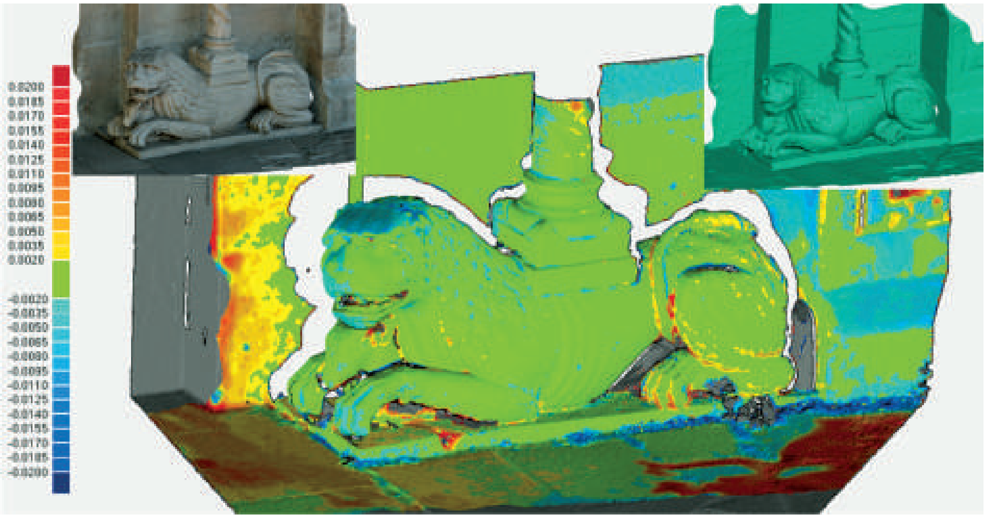

Fig. 8: 3D comparison of the lion figure at the imperial church in Königslutter between the 3D models from 123D Catch Beta and IMAGER 5006; top left: 3D model with texture; top right: without texture from 123D Catch Beta (green < 2 mm).

Fig. 9: 3D comparison of the lion figure at the imperial church in Königslutter between the 3D models from Bundler/PMVS2 and IMAGER 5006; bottom left: point cloud from Bundler/PMVS2; bottom centre: IMAGER 5006h; bottom right: meshed 3D model from the IMAGER 5006h (green < 2 mm).

A geometrical accuracy analysis of this data could not be carried out due to the absence of reference data.

A 3D model of the rear of a moai eye, carved from Obsidian stone, could only be successfully generated using 123D Catch Beta, because both laser scanner IMAGER 5006i and Bundler/PMVS2 could only measure very noisy point clouds with each including some gaps (Fig. 11). For measurements

Fig. 10: Small moai (height 0.7 m) at Poike on Easter Island (Chile); left: measured points; centre left: 3D meshing; centre right: textured 3D model in each case from 123D Catch Beta; right: dense coloured point cloud measured with Bundler/PMVS2 identifying some gaps (white ellipses).

with 123D Catch Beta and Bundler/PMVS2 image data from the following cameras were used: a) 27 photos from the Pentax Optio X (2560 x 1920 pixel) with 10 mm focal length, b) ten photos from a Nikon D70 (3008 x 2000 pixel) with a focal length of approx. 40 mm, and c) eleven photos from a Nikon D80 (3872 x 2592 pixel) with a focal length of 50 mm. The textured 3D model of the moai eye generated with 123D Catch Beta consists of 291,613 triangles (Fig. 11, right), while in total only 214,940 points were measured with Bundler/PMVS2 and only 95,820 points were scanned with the IMAGER 5006i. Although 3D comparison with reference data could not be conducted for accuracy analysis, subjective comparison of the provided 3D model suggests that it corresponds very well to the original. The model could not be scaled as no metric information was available.

For the next two examples (pottery/ceramic and architectural fragment from Yeha/Ethiopia) reference data were available. The pottery (ceramic(s)) was acquired by camera and laser scanner in order to compare image-based and range-based object recording. The pottery (Fig. 12) was photographed in the upper part with 50 photos and in the lower part with 30 photos using a Nikon D40 (Nikkor zoom lens, focal length 34 mm). With the terrestrial laser scanner IMAGER 5006h nine scans were required to achieve complete recording of the object. The triangle meshing of the pottery (approx. 160,590 triangles) was computed using the registered scans in Geomagic. The image data was processed with Bundler/PMVS2 using a reduced image resolution of 2400 pixels, i.e. a point cloud was separately generated for the upper and lower object part (Fig. 12). Using the measuring scale (10 cm), which was additionally placed into object space for the photographs, the scaling of the two datasets could be carried out. Firstly, the upper and lower 3D point clouds were registered using ICP algorithm. Secondly the 3D model was generated using the common point cloud with 0.5 mm grid spacing. Finally, a 3D comparison between model and reference was computed in Geomagic. This showed that differences larger than 1 mm only occurred at the boundary regions and in those areas with large curvatures (Fig. 12). These differences can be explained by the higher resolution of the imagery compared to the more smoothed scan data. The average deviation between image-based and range-based 3D model is 1 mm, while the standard deviation is 1.4 mm (Tab. 2).

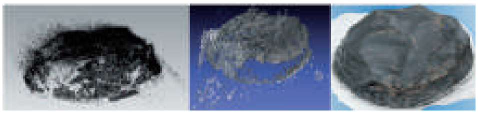

Fig. 11: Black Obsidian stone as the rear of a moai eye at Easter Island; left: point cloud from IMAGER 5006i; centre: from Bundler/PMVS2; right: meshed 3D model from 123D Catch Beta.

Another archaeological finds from the excavation in Yeha (Ethiopia) was an architectural fragment which was photographed in two image blocks using a Nikon D90 (Nikkor zoom lens 24 mm): 33 photos (back side) and 25 photos (front side). A small scale bar was placed into object space for image acquisition thus facilitating scaling of the two datasets during processing. The image data was processed with VisualSFM using the full image resolution of the photographs for each image block producing two point clouds (595,933 points in total with grid spacing of 1 mm). After registration of the two point clouds with ICP algorithm a combined meshed 3D model (approximately 1,135,284 triangles in total) was computed (Fig. 13 left in each case).

As a reference, the object was scanned with the IMAGER 5010 from nine scan stations. According to the technical specifications of the system manufacturer the IMAGER 5010 has improved range noise behaviour. Range noise of this scanner (1 sigma) for unfiltered

Fig. 12: Left: Photograph of the pottery from Yeha (Ethiopia) with the measuring stick for object scaling; centre left: composing the pottery from Bundler/PMVS2; centre right: meshed and smoothed 3D model of the pottery from IMAGER 5006h; right: result of 3D comparison in Geomagic Studio 12 (green < 1 mm).

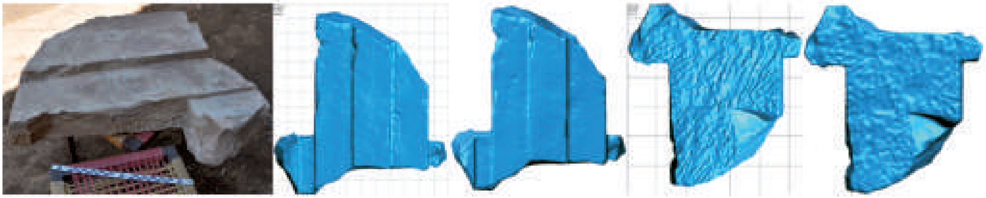

Fig. 13: Architectural fragment (archaeological find) in Yeha (Ethiopia): from left to right: photograph with scale bar; meshed 3D model of front (VisualSFM); meshed 3D model of front (IMAGER 5010); rear (VisualSFM); rear (IMAGER 5010).

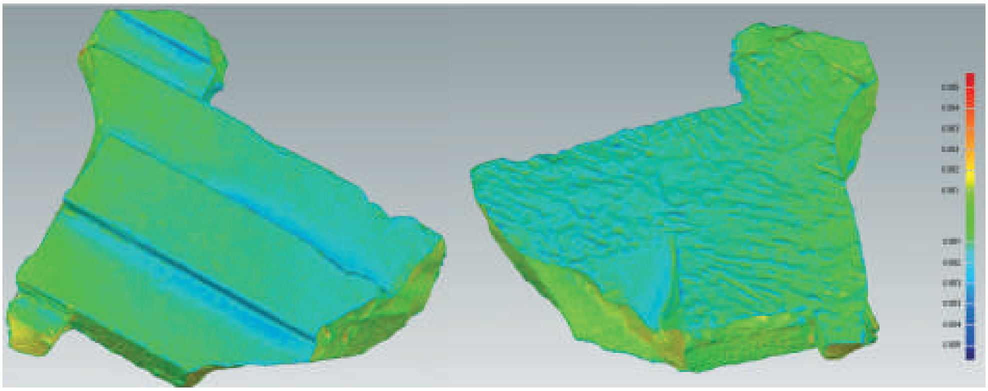

Fig. 14: Deviation between 3D model of the architectural fragment to reference data of the IMAGER 5010 illustrated in Geomagic Studio 2012 (green < 1 mm).

raw data at 10 m distance is 0.5 mm for a reflectivity of 14% (black), 0.4 mm for a reflectivity of 37% (dark grey), and 0.3 mm for a reflectivity of 80% (white). The registered point cloud (2,072,781 points) was filtered and reduced in Geomagic to a regular point distance of 2 mm, resulting in a mesh of approximately 1,120,580 triangles (Fig. 13). First the scan data was slightly smoothed using a noise filter. After ICP registration of the meshed model from image-based data with the meshed model from laser scanning a 3D comparison was performed in Geomagic producing the results of the front and rear sides of the fragment presented in Fig. 14. Differences from the image-based data to the reference, which are below 1 mm in the average (Tab. 2), are depicted in green. This comparison demonstrated that the dataset corresponds most closely to the result from laser scanning in the range of ±1 mm. Since the image data obviously supplied more details with a higher resolution than the smoothed scanning data in this case study (Fig. 14), these comparisons should primarily

Tab. 2: Comparison of image-based measurement (Geomagic Studio) and terrestrial laser scanning measurement (reference data); dev. = deviation, av. = average, std. = standard; (mm).

Object	Software/Criterion	# meshes	Max. dev.	Av. dev.+	Av. dev.-	Std. dev.
Town house	Photofly	504,449	530.0	36.6	-17.9	77.1
Town house	Bundler/PMVS2	895,986	517.0	7.5	-4.2	18.0
Lion	123D Catch Beta	686,285	153.0	18.7	-18.4	38.3
Lion	Bundler/PMVS2	2,669,244	153.0	14.9	-13.2	32.2
Pottery	Bundler/PMVS2	323,402	8.0	0.9	-1.0	1.4
Fragment	VisualSFM	1,135,284	−4.7	0.6	-0.8	0.8

detect gross errors in the models, which obviously did not arise here. A fringe projection (structured light) system could supply more precise reference data with higher resolution, but such a system was not available in Yeha/Ethiopia at that time. However, an average deviation of better than 0.3 mm in the 3D comparison between a 3D model of a similar fragment generated from images and scanned with a fringe projection system Breuckmann OptoTOP-HE confirmed the good results of these image-based systems (Kersten & Lindstaedt 2012).

5 Conclusion and Outlook

In this contribution economical, image-based recording and modelling procedures, which are able to generate precise and detailed 3D surface models from terrestrial photographs for applications in architecture, cultural heritage and archaeology, were presented. The quality of results from image-based systems depends on the number of images used, the image resolution, photo scale, illumination conditions and the parameter settings of the software applied. The results from some software packages, e.g. Bundler/PMVS2, using images from SLR digital cameras and standard parameter settings in the software are comparable with results from expensive terrestrial laser scanners. Object recording with cameras is simple, very fast, very flexible and economical. The entire procedure is to a large extent automated and works even without targets as control points. Objects can be scaled through the inclusion of a single scale bar in object space. The results presented here show that the open-source software Bundler/PMVS2, VisualSFM and the web service Photofly and/or 123D Catch Beta can generate equivalent 3D models when compared with terrestrial laser scanners, although the exact results depend upon the size and shape of the objects. However, the reliability of the image-based systems requires improvement because some of the achieved results were geometrically unusable. Depending upon object material and lighting conditions, noisy point clouds were sometimes produced with Bundler/PMVS2. Photofly/123D Catch Beta showed results that were visually very attractive (smoothed – whereby small holes are automatically filled by the software) when used on small and roundish objects. These results could fulfill the requirements of many users for visualisation in the Internet. However, the results from Photosynth are not useful for 3D modelling since the point density is too small. Nevertheless, further investigations with other more precise recording procedures such as fringe projection (structured light) or close-range scanners are necessary for small objects in order to be able to provide verified statements about the geometrical quality. Furthermore, CAD modelling with image-based 3D point clouds such as presented by Kersten et al. (2004), Kersten (2006) and Kersten (2007) should be investigated in the future to evaluate the potential of this automatically generated data.

These investigations have shown that Bundler/PMVS2 was more efficient than Photofly/123D Catch Beta for larger objects although the computation times were substantially longer. The significantly shorter processing time with the web service implies that Autodesk makes plentifully computing resources available for this service. Due to the optional use of these web services the resources of the user's own computer could be preserved, and a time-saving is achieved relative to the computation on a local PC. However, the availability of a fast Internet connection is not guaranteed everywhere in the world, particularly at the locations of many archaeological expeditions, meaning that the operation of open-source software on a user's own PC has an advantage. Furthermore, issues relating to data privacy should not be neglected when using web services.

Due to the sensible combination of computer vision algorithms and photogrammetric procedures the workflow from object recording, through 3D modelling to visualization has become increasingly automated without significantly neglecting geometrical accuracy. These combined procedures are only at the beginning of their development, since the speed of such algorithms can be significantly increased by future implementation of the software on graphics processor units (GPU) and because the mutual integration of both procedures can still be substantially optimized. A large number of algorithms for pixel-based matching in stereo or multi-view photographs with different performance potential from the field of computer vision are also available (Scharstein & Szeliski 2009) and have yet to be applied to this field. However, the level of automation is so high that many solutions are black-boxes with poor repeatability and low reliability. Remondino et al. (2012) present an investigation of automated image orientation packages in order to clarify potentialities and performances (including for camera calibration) when dealing with large and complex datasets. Their results demonstrate that, in case of complex and long sequences, SfM methods suffer of reliability and repeatability.

The dominant market position for the last 10 years of airborne and terrestrial laser scanning as tools for extensive data acquisition is being challenged by efficient photogrammetric procedures supported by computer vision algorithms and improved computer technology.

Acknowledgements

The authors would like to gratefully acknowledge the support of Kay Zobel (HafenCity University Hamburg) for developing the customised graphical user interface of Bundler/PMVS2 and MeshLab (Fig. 1), which provided an automatic workflow for the software.

References

123d, 2012: http://www.123dapp.com/catch/ (3.9.2012).

Abate, D., Furini, G., Migliori, S. & Pierattini, S., 2011: Project Photofly: New 3D Modelling Online WEB Service. – Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (5/W16), http://www.isprs.org/proceedings/XXXVIII/5-W16/pdf/abate_etal.pdf (4.9.2012).

Barazzetti, L., Remondino, F. & Scaioni, M., 2009: Combined use of photogrammetric and computer vision techniques for fully automated and accurate 3D modeling of terrestrial objects. – SPIE Optics+Photonics, 7447, 2–3 August, San Diego, CA, USA.

Barthelsen, J., Mayer, H., Hirschmüller, H., Kuhn, A. & Michelini, M., 2012: Orientation and Dense Reconstruction from Unordered Wide Baseline Image Sets. – PFG **2012** (4): 421–432.

Besl, P. & McKay, N., 1992: A Method for Registration of 3-D Shapes. – IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) **14** (2): 239–256.

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. & Ranzuglia, G., 2008: MeshLab: an open-source mesh processing tool. – Scarano, V., De Chiara, R. & Erra, U. (eds.): Eurographics Italian Chapter Conference, The Eurographics Association: 129–136.

Courchay, J., Pons, J.-P., Monasse, P. & Keriven, R., 2010: Dense and Accurate Spatio-temporal Multi-view Stereovision. – Computer Vision – ACCV 2009, Lecture Notes in Computer Science **5995:** 11–22.

Furukawa, Y. & Ponce, J., 2010: Accurate, Dense, and Robust Multi-View Stereopsis. – IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (8): 1362–1376.

Geomagic, 2012: http://www.geomagic.com (3.9.2012).

Hausner, C., 2010: SynthExport. http://synthexport.codeplex.com/ (4.1.2012).

Kersten, Th., 2006: Combination and Comparison of Digital Photogrammetry and Terrestrial Laser Scanning for the Generation of Virtual Models in Cultural Heritage Applications. – Ioannides, M., Arnold, D., Niccolucci, F. & Mania, K. (eds.): 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST **2006:** 207–214.

Kersten, Th., 2007: Virtual Reality Model of the Northern Sluice of the Ancient Dam in Marib/Yemen by Combination of Digital Photogrammetry and Terrestrial Laser Scanning for Archaeological Applications. – International Journal of Architectural Computing, Special Focus on Cultural Heritage **02** (05): 339–354, Multiscience.

Kersten, Th., Acevedo Pardo, C. & Lindstaedt, M., 2004: 3D Acquisition, Modelling and Visualization of north German Castles by Digital Architectural Photogrammetry. – Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (5): 126–132.

Kersten, Th. & Lindstaedt, M., 2012: Image-Based Low Cost Systems for Automatic 3D Recording and Modelling of Archaeological Finds and Objects. – EuroMed 2012 – Int. Conference on Cultural Heritage. – Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F. & Caffo, R. (eds.). – Lecture Notes in Computer Science (LNCS) **7616**: 1–10, Springer.

Kersten, Th., Lindstaedt, M., Mechelke, K. & Zobel, K., 2012: Automatische 3D-Objektrekonstruktion aus unstrukturierten digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie. – Seyfert, E. (Hrsg.): Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V. **21:** 137–148, DGPF, Potsdam, CD-ROM.

Lourakis, M.I.A. & Argyros, A.A., 2004: Design and Implementation of a Generic Sparse Bundle Adjustment Software Package Based on the Levenberg-Marquardt Algorithm. – Institute of Computer Science, FORTH-ICS, Tech. Rep. **340**, Heraklion, Greece, http://www.ics.forth.gr/~lourakis/sba (4.9.2012).

Lowe, D.G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. – International Journal of Computer Vision **60** (2): 91–110.

Photosynth, 2012: http://photosynth.net/ (4.9.2012).

Pomaska, G., 2009: Utilization of Photosynth Point Clouds for 3D Object Reconstruction. – **22nd** CIPA Symposium, Kyoto, Japan, http://cipa.icomos.org/fileadmin/template/doc/KYOTO/34.pdf (4.9.2012).

Remondino, F. & Menna, F., 2008: Image-based surface measurement for close-range heritage documentation. – Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B5-1): 199–206.

Remondino, F., El-Hakim, S.F., Grün, A. & Zhang, L., 2008: Turning images into 3-D models. – IEEE Signal Processing Magazine **25** (4): 55–65.

Remondino, F., Del Pizzo, S., Kersten, Th. & Troisi, S., 2012: Low-cost and Open-Source Solutions for Automated Image Orientation – a Critical Overview. – EuroMed 2012 – Int. Conference on Cultural Heritage. – Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F. & Caffo, R. (eds.). – Lecture Notes in Computer Science (LNCS) **7616**: 40–54, Springer.

Scharstein, D. & Szeliski, R., 2009: Middlebury Stereo Vision Page. http://vision.middlebury.edu/stereo/ (7.1.2012).

Snavely, N., Seitz, S.M. & Szeliski, R., 2006: Photo tourism: exploring photo collections in 3D. – ACM Transactions on Graphics **25** (3): 835–846.

Snavely, N., Seitz, S.M. & Szeliski, R., 2008: Modeling the World from Internet Photo Collections. – International Journal of Computer Vision **80** (2): 189–210.

Wu, C., 2007: SiftGPU: A GPU implementation of Scale Invariant Feature Transform (SIFT). http://cs.unc.edu/~ccwu/siftgpu (4.9.2012).

Wu, C., 2011: VisualSFM: A Visual Structure from Motion System. http://www.cs.washington.edu/homes/ccwu/vsfm/ (4.9.2012).

Address of the authors:

Thomas P. Kersten and Maren Lindstaedt, HafenCity Universität Hamburg, Labor für Photogrammetrie & Laserscanning, Hebebrandstraße 1, D-22297 Hamburg, Tel.: +49-40-42827-5343, Fax: +49-40-42827-5399, e-mail: {thomas.kersten}{maren.lindstaedt}@hcu-hamburg.de

Manuskript eingereicht: April 2012
Angenommen: August 2012