Investigation of a consumer-grade digital stereo camera

Conference Paper *in* Proceedings of SPIE - The International Society for Optical Engineering · May 2013

DOI: 10.1117/12.2020899

4 authors:

Fabio Menna
Fondazione Bruno Kessler
55 PUBLICATIONS **481** CITATIONS
SEE PROFILE

Erica Nocerino
Fondazione Bruno Kessler
45 PUBLICATIONS **283** CITATIONS
SEE PROFILE

Fabio Remondino
Fondazione Bruno Kessler
282 PUBLICATIONS **3,890** CITATIONS
SEE PROFILE

Mark R Shortis
RMIT University
167 PUBLICATIONS **2,324** CITATIONS
SEE PROFILE

Investigation of a consumer-grade digital stereo camera

Fabio Menna[a*], Erica Nocerino[a,b], Fabio Remondino[a], Mark R. Shortis[c]

[a] 3D Optical Metrology unit, Bruno Kessler Foundation (FBK), 38123 Trento, Italy
Email: <fmenna><nocerino><remondino>@fbk.eu, Web: http://3dom.fbk.eu
[b] Parthenope University of Naples, Dept. of Applied Science, 80143 Naples, Italy
[c] School of Mathematical and Geospatial Sciences, RMIT University, Australia
Email: mark.shortis@rmit.edu.au

ABSTRACT

The paper presents a metric investigation of the Fuji FinePix Real 3D W1 stereo photo-camera. The stereo-camera uses a synchronized Twin Lens-CCD System to acquire simultaneously two images using two Fujinon 3x optical zoom lenses arranged in an aluminum die-cast frame integrated in a very compact body. The nominal baseline is 77 mm and the resolution of the each CCD is 10 megapixels. Given the short baseline and the presence of two optical paths, the investigation aims to evaluate the accuracy of the 3D data that can be produced and the stability of the camera. From a photogrammetric point of view, the interest in this camera is its capability to acquire synchronized image pairs that contain important 3D metric information for many close-range applications (human body parts measurement, rapid prototyping, surveying of archeological artifacts, etc.). Calibration values - for the left and right cameras - at different focal lengths, derived with an in-house software application, are reported together with accuracy analyses. The object coordinates obtained from the bundle adjustment computation for each focal length were compared to reference coordinates of a test range by means of a similarity transformation. Additionally, the article reports on the investigation of the asymmetrical relative orientation between the left and right camera.

Keywords: stereo-camera, calibration, accuracy, 3D modeling

`

INTRODUCTION

In 1838 Sir Charles Wheatstone invented the stereoscope, a device built with two mirrors which showed two separate images taken from two different points of view. For the first time and in a synthetic way, the device invented by Wheatsone permitted the illusion of the depth through a pair of images, known now as a stereoscopic pair. Since that moment, the availability of devices for stereoscopy has required also the realization of instruments able to acquire simultaneously the two images of the stereoscopic pair. Cameras equipped with two lenses at a fixed distance similar to the human inter-pupillary distance were built to take stereoscopic images that could be viewed with the stereoscope or as an anaglyph picture using special glasses. Since its invention, stereoscopy has never been a mass market technology, being limited to technical applications and use by specialists such as photogrammetric stereo-plotting operators. Even with the introduction of computers, stereoscopy has been bound to technical applications such as 3D simulators and for digital photogrammetry (stereo restitution).

Today, thanks to recent technological developments, 3D TVs that use polarized glasses or auto-stereoscopy (eyewear-free) are gradually substituting traditional LCD screens and televisions. At the same time, average consumers are becoming familiar with 3D contents that are often the core of computer animations, videogames, films, etc. Moreover, different low cost 3D digital imaging devices (Figure 1) have been introduced in the market for (i) recording scenes that can be directly viewed in stereo or (ii) producing 3D digital models of simple objects viewable in handheld console and for visualization issues. A well known publicly accessible example is the digital multimedia guide to the Louvre Museum[1]. These low cost devices include 3D mobile phones, cameras and videocameras that use an optical splitter to record simultaneously a stereoscopic image pair on a unique digital sensor or, alternatively, are built as system that comprises two separate sensors and lenses mounted in a single body.

*fmenna@fbk.eu; phone +39 0461 314446; fax +39 0461 314340; http://3dom.fbk.eu/

Videometrics, Range Imaging, and Applications XII; and Automated Visual Inspection, edited by Fabio Remondino, Mark R. Shortis, Jürgen Beyerer, Fernando Puente León, Proc. of SPIE Vol. 8791, 879104
 · doi: 10.1117/12.2020899

Figure 1. Example of low-cost and consumer-grade passive sensors with 3D capabilities: LG Optimus 3D, HTC EVO 3D, Nintendo 3DS, JVC GS-TD1 camcorder, Panasonic SDT750 camcorder (above), Fujifilm FinePix Real 3D W1, PointGrey Bumblebee, (below).

Among many passive imaging devices available on the market, the authors decided to investigate the metric capabilities of a consumer digital stereo-camera: the Fujifilm FinePix Real 3D W1. This camera was announced in July 2009 as the world's first 3D digital imaging system that lets users enjoy "three dimensional" images without the need of wearing special 3D glasses.

The camera received the Technical Image Press Association (TIPA) award of the year 2010 as the best imaging innovation since "it offers an uncomplicated way for consumers to create and perceive 3D images and movies".

In this contribution, a geometric investigation of the metric capabilities of the Fujifilm Finepix Real 3D W1 as a passive optical triangulation device is addressed. The calibration and characterization of consumer-grade active and passive sensors [2-8] is essential in order to evaluate (i) operational performances, (ii) stability and repeatability as well as (iii) spatial/temporal coverage/resolution. In this article firstly a general description of this low cost device is given, the arrangement of its components, the technology behind the functionality as well as theoretical errors involved in the computed 3D data are analyzed and discussed. Laboratory tests performed for geometrical characterization of the optical sensors are reported in the case of static acquisitions. Precision analyses as well as accuracy tests are performed for different objects comparing reference measurements with those obtained from the stereo camera. These tests are inspired by the guidelines suggested by international committees such as the German society VDI/VDE[9].

BACKGROUND AND GENERAL DESCRIPTION OF TESTS

The understanding of basic theories and best practices associated with the use of a device are fundamental to assess important characteristics and realize if the device is suited for a specific application. When devices are used by non-experts, simple and clear protocols of operation should be found to allow a proper use and functioning.

The Finepix Real 3D W1 is mainly designed to allow stereo-vision but its metric potential was immediately clear to the photogrammetric and vision communities. A visible (or artificially illuminated) 3D point is seen from the twin digital cameras arranged along a baseline, therefore the simple geometric principle of optical triangulation could be used for recording and supplying 3D information.

As for all the consumer digital stereo-cameras launched into the market, the Finepix Real 3D W1 features a very small baseline of about 77 mm. This implies that for an object at a distance of only 0.75 m the intersection angle is very narrow (ca. 5 degrees) and therefore non-optimal for photogrammetric measurement. This characteristic should not exclude the camera for accurate measurements but instead only limit its field of applicability on the basis of the required accuracy. In these terms a theoretical formulation on the potential accuracy achievable with the Fuji device should be made, compared to experimental tests and disseminated to final users in the form of guidelines and protocols of data acquisition. Because of the small baseline that characterizes this camera, accurate and reliable procedures for calibrating the optical sensors and determining their relative exterior orientation must be found. In previous publications[10,11], first analyses were performed to assess the accuracy of this stereo-camera in the absence of control points on the object. Distance

measurements on a planar test field located first perpendicular to the optical axes and then with the camera tilted up to approximately 35 degrees were performed. The relative RMSE of 216 measurements of 10 cm long line segments were reported for different focal lengths. Camera calibration and relative orientation parameters are also reported.

The hereafter reported investigation describes on a more extensive test regarding the accuracy and reliability of the Fuji consumer-grade stereo camera.

General description of Fuji FinePix Real 3D W1

The stereo-camera uses a synchronized Twin Lens-CCD System to acquire simultaneously two images that are then viewed on the 3D display on the back of the camera. Two Fujinon 3x optical zoom lenses are mounted in an aluminum die-cast frame that is integrated in a very compact body which measures only (12 x 7 x 3) cm. The nominal baseline is 77 mm and the resolution of the each CCD is 10 megapixels. Technical specifications are shown in Table 1. As is the case for many proprietary digital cameras, the spacing or size of the sensor elements is not given and must be estimated indirectly. From a photogrammetric point of view, the interest in this camera is its capability to acquire synchronized image pairs that contain important 3D metric information for many close range applications (for example human body measurement, rapid prototyping and recording archeological artifacts). Given the short baseline of the Fuji Real 3D W1, our test aims to evaluate the accuracy of 3D data that can be extracted for small and medium size objects.

Table 1. Nominal characteristics of the two cameras of the Fujo Finepix Real 3D W1.

	RGB camera
Sensor type	CCD
Sensor size	1/2.3”
Focal length	6.3-18.9 mm
Pixel size	not available
Raw image format	3648 x 2736 px
FOV H	46 – 13 degrees
FOV V	35 – 10 degrees

CHARACTERIZATION CONCEPTS FOR STEREO-TRIANGULATION SYSTEMS

Theoretical precision

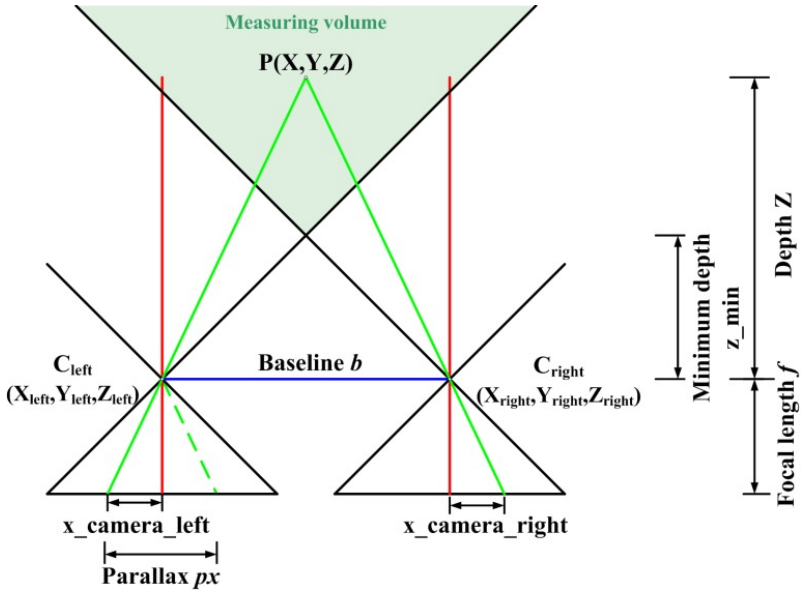


Figure 2. Triangulation system composed of 2 cameras arranged with parallel optical axes.

In a standard stereo-triangulation system composed of two cameras (or a camera and a projector) arranged in a parallel optical axes (Fig. 2), naming the principal distance c and the baseline b, it is well known for aerial and close-range photogrammetry[12] that the depth Z is inversely proportional to the parallax p_x (also known as disparity) according to:

$$Z = \frac{c \cdot b}{p_x} \tag{1}$$

In such a system, the parallax p_x varies within a range [0, *sensor_width*] which corresponds to a depth [$+\infty$, Z_{min}].

The precision (as a measure of uncertainty) of the calculated 3D coordinates of an object point *P(X, Y, Z)* is different in the three dimensions. For points located on the xy plane normal to the optical axis at a distance Z from the imaging sensor, the precision σ_{XY} and σ_Z are defined as follows:

$$\sigma_{XY} \approx \frac{Z}{c} \cdot \sigma_{px} \approx \frac{Z}{c} \cdot \varepsilon \cdot \mu \tag{2a}$$

$$\sigma_Z \approx \frac{Z^2}{c \cdot b} \cdot \sigma_{px} \approx \frac{Z^2}{c \cdot b} \cdot \varepsilon \cdot \mu \tag{2b}$$

where:
- μ is the pixel size of the imaging device;
- σ_{px} is the precision in measuring the parallax p_x which depends on the precision (fraction of pixels ε) of the image measurements.

Figure 2 shows the geometric schema of a triangulation device for which the two equation (2a and 2b) hold. One of the two cameras can be substituted by a projector as in the case of pattern or fringe projector scanners. For the sake of simplicity, the schema in Figure 2 is specific for the normal case where optical axes are parallel. The stereo-camera studied in this investigation can be roughly approximated to this schema, at least for general analysis purposes.

Relative orientation of two cameras

In photogrammetry, the relative orientation between two cameras viewing the same scene is the procedure of finding the five degrees of freedom of one camera relative to the other (one degree of freedom is considered known or fixed). At least five homologous points have to be observed by the two cameras to solve for the five unknowns by means of the co-planarity equations. An alternative method is to determine the absolute orientation (six degrees of freedom) relative to an external reference frame for both cameras independently. Then, a roto-translation is computed to make coincident the external reference frame with one of the two camera reference systems (the left one in general).

In the following equations, the letter *G*, *R, L,* indicate respectively the Global, Right camera and Left camera reference frames. The superscripts specify the reference frame where the quantity is defined, e.g. P^G is the position vector of a generic point P known in the global frame. The notation K_i^j expresses the transformation from the coordinate system $\{i\}$ to the coordinate system $\{j\}$. According to this rule, $\mathrm{M}_{\mathrm{G}}^{\mathrm{R}}$ is the rotation matrix from the global to the right camera frame and $\mathrm{O}_{\mathrm{R}}^{\mathrm{G}}$ represents the origin of right camera frame expressed in the global reference frame. According to this notation, the coordinates of the point P^G known in the global frame can be expressed in the right camera reference frame as function of the exterior orientation parameters of the right camera:

$$P^R = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}^R = M_G^R \cdot \begin{bmatrix} X - X_{oR} \\ Y - Y_{oR} \\ Z - Z_{oR} \end{bmatrix}^G = M_G^R \cdot (P^G - O_R^G) \tag{3}$$

$$M_G^R = \begin{bmatrix} \cos\varphi \cos k & \cos\omega \sin k + \sin\omega \sin\varphi \cos k & \sin\omega \sin k - \cos\omega \sin\varphi \cos k \\ -\cos\varphi \sin k & \cos\omega \cos k - \sin\omega \sin\varphi \sin k & \sin\omega \cos k + \cos\omega \sin\varphi \sin k \\ \sin\varphi & -\sin\omega \cos\varphi & \cos\omega \cos\varphi \end{bmatrix} \tag{4}$$

$$O_R^G = \begin{bmatrix} X_{oR} \\ Y_{oR} \\ Z_{oR} \end{bmatrix}^G \quad (5)$$

where (4) and (5) are respectively the rotation matrix containing the orientation angles and the position vector of the right camera perspective center in the global frame. To perform a relative orientation between the two cameras is equivalent to re-orienting the global reference system to be coincident with the left camera frame. This transformation is expressed as:

$$P^R = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}^R = M_L^R \cdot \begin{bmatrix} X - X_{oR} \\ Y - Y_{oR} \\ Z - Z_{oR} \end{bmatrix}^L = M_L^R \cdot (P^L - O_R^L) \quad (6)$$

$$M_L^R = M_G^R \cdot (M_G^L)^{\mathrm{T}} \quad (7)$$

$$O_R^L = M_G^L \cdot (O_R^G - O_L^G) \quad (8)$$

where (7) and (8) are respectively the rotation matrix (i.e., orientation angles) and coordinates of the right camera in the left camera reference frame. In other words, equation (8) represents the components of the baseline between the two cameras in the reference system centered on the left camera. From the above transformation, the exterior orientation parameters of the chosen camera become null and the exterior orientation of the second camera relative to the first is obtained. This procedure is convenient when the exterior orientation of the two cameras is known, for example from a bundle adjustment process. Typically, a self-calibration with coded targets is performed to recover the interior orientation parameters of the two cameras (or camera and projector) and as a result of the bundle adjustment, the exterior orientation of the cameras are determined. For a stereo-triangulation camera system which is composed of two cameras fixed to each other, the relative orientation in theory should be the same for each camera pose. In practice, the presence of errors in the image measurements or an inappropriate mathematical model (i.e. missing parameters to correct lens distortions) results in different relative orientations for each camera pose. As many camera poses are generally determined in a self-calibration procedure, an average relative orientation and other statistical parameters can be computed.

CHARACTERIZATION OF THE FINEPIX REAL 3D W1

Theoretical precision of the Fujifilm FinePix Real 3D W1 as a passive triangulation device

In Fig. 3b a theoretical diagram of the precision σ_z as a function of object distance Z is reported, considering an optical triangulation system whose characteristics are: b=77 mm, f=6.3 mm, μ=0.0017 mm, ε=1 or 1/10 pixel (for manual or automated image marking measurements, respectively).

Raw data extraction

If used as stereo-camera, the Finepix Real 3D W1 stores the two simultaneous photos in a file format called MPO which contains the two JPEG images taken by the two cameras. Additional information such as the baseline and the convergence angle between the cameras are also stored in the MPO format (Fig. 3c). The single JPEG files for the left and right can be extracted with proprietary and third party software.

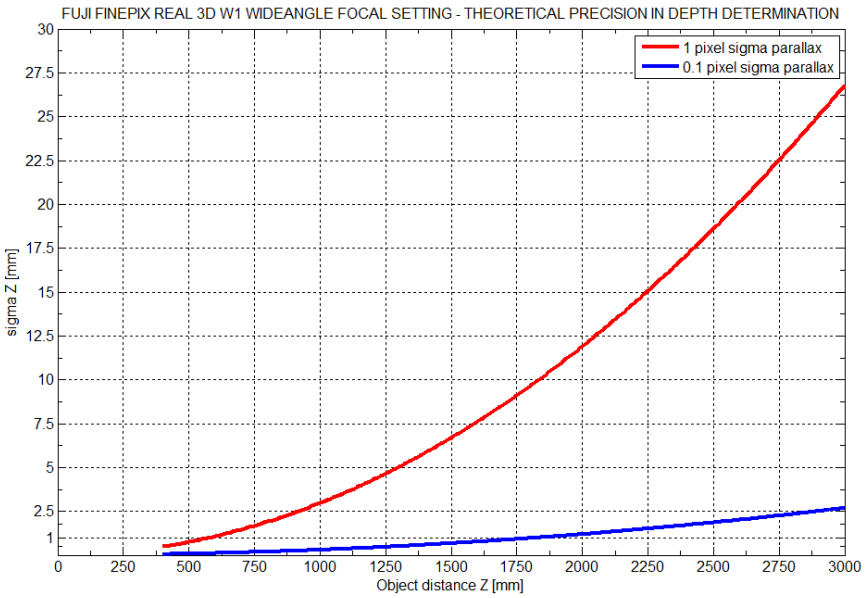
Retrieving technical characteristics and initial parameter approximations

As mentioned previously, the pixel size is not given by the manufacturer. An approximate value can be derived both from the factory specifications for a camera with 1/(2.3) inch sensor optical format and from the information contained in the *exif* tag. A value of 1.7 μm for the pixel size can be assumed, as shown in Table 2.

Camera calibration

The camera has two motorized zoom lenses that are simultaneously controlled by a stick/lever. The focal length can assume ten different values in approximately equally spaced steps. The investigation/characterization was carried out at three different focal lengths: (i) widest setting, (ii) central setting (widest +4 steps), (i) tele setting (widest +9 steps that is the maximum focal length).

a) b)



c)

	---- MPF ----
MPFVersion	0100
NumberOfImages	2
MPImageFlags	Representative image
MPImageFormat	JPEG
MPImageType	Multi-frame Disparity
MPImageLength	5406016
MPImageStart	0
DependentImage1EntryNur	0
DependentImage2EntryNur	0
MPImageFlags	(none)
MPImageFormat	JPEG
MPImageType	Multi-frame Disparity
MPImageLength	4760725
MPImageStart	5406016
DependentImage1EntryNur	0
DependentImage2EntryNur	0
MPIndividualNum	1
BaseViewpointNum	1
ConvergenceAngle	0
BaselineLength	0.077

Figure 3. a) The Fujifilm Finepix Real 3D W1. b)Theoretical behavior of the precision σ_z as a function of the object's distance Z and image measurement precision for the Finepix Real 3D W1 at widest zoom setting. c) Additional information stored in the MPO format.

Table 2. Derivation of approximate values for the pixel size of the Fujifilm Finepix Real 3D W1.

Factory specifications (1/(2.3)" sensor optical format)		***Exif* tag**	
Sensor_x_dimension [mm]	6.17	FocalPlaneResolutionUnit [cm]	1
Sensor_y_dimension [mm]	4.55	-	-
Image_x_dimension [pixel]	3648	FocalPlaneXResolution [pixel]	5952
Image_y_dimension [pixel]	2736	FocalPlaneYResolution [pixel]	5952
pixel_x_dimension [μm]	$\frac{\text{Sensor\_x\_dim}}{\text{Image\_x\_dimension}} = 1.69$	pixel_x_dimension [μm]	$\frac{\text{FocalPlaneResolutionUnit}}{\text{FocalPlaneXResolution}} = 1.68$
pixel_y_dimension [μm]	$\frac{\text{Sensor\_y\_dim}}{\text{Image\_y\_dimension}} = 1.66$	pixel_y_dimension [μm]	$\frac{\text{FocalPlaneResolutionUnit}}{\text{FocalPlaneYResolution}} = 1.68$

Additionally, to investigate the possibility of in-camera distortion corrections or special zoom optical design due to the small space in the camera body (see patent Fujinon bifocal imaging optical system and imaging apparatus[13]), suspected to be only on the widest setting, a fourth calibration was made: (iv) widest +1 step.

A bundle adjustment with self-calibration procedure was used to compute the interior orientation with additional parameters following the Brown model[14]. The test field shown in Figure 4a was imaged from 8 different positions

approximately located on a hemisphere to assure optimal intersections between rays. The images were framed to fill the whole image format. The focus setting was forced to be fixed for a unique distance by pre-focusing on a high contrast target Siemens star located at fixed distance from the camera and re-composing on the test field just before taking the image pair. This was necessary to reduce the principal distance variation during the image acquisition, as the camera cannot lock the focus on a predefined distance. For each station three images, one in landscape and two in portrait position, were taken to guarantee roll diversity that reduces correlations between parameters during self-calibration procedures[6]. The images were acquired in order to guarantee that the whole sensor area was evenly covered by image observations (Fig. 4c). The procedure was repeated for each focal length setting chosen for the calibration. The test field, approximately 800 mm x 800 mm x 100 mm, has 107 circular coded targets, 80 located on the main plane and 37 on 10 raised sections attached to it (ca. 10 cm high). The 3D coordinates of the test field are periodically measured with photogrammetry and geodetic measurements with an accuracy of $\sigma_{xyz} \approx 0.030$ mm where the plane OXY is the main plane of the test field while the Z axis points toward the camera. The coordinates of these points are used as reference for calculating the RMSE of the computed 3D coordinates.

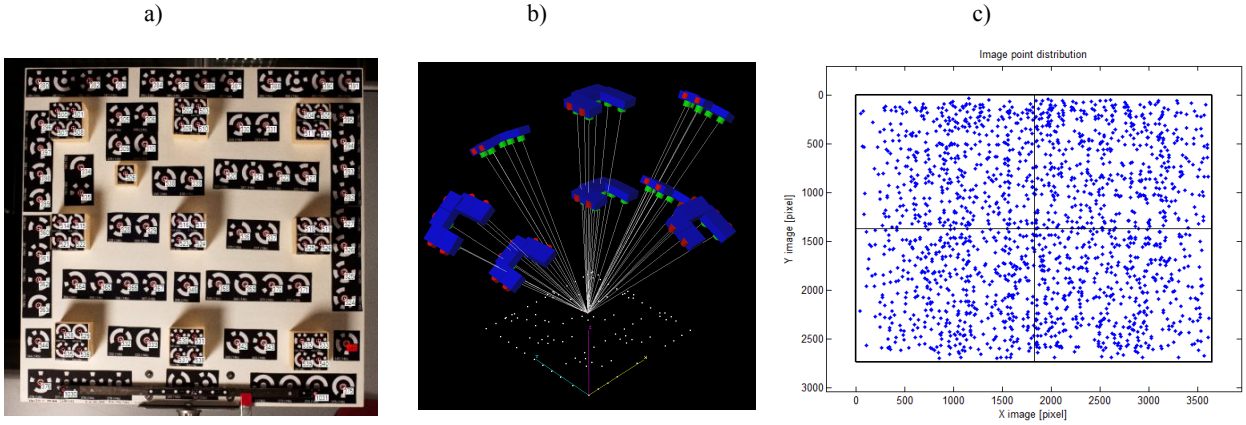


Figure 4. Test field used for self-calibration (a), network geometry (b) and image point distribution within the image frame (c).

Different photogrammetric packages were used to compute the calibration parameters by means of a free network photogrammetric bundle adjustment. From each MPO image file, two separate JPEG files (one for the left and one for right camera) were exported using the proprietary software. The left and right cameras were calibrated separately, then a unique bundle adjustment was performed with both cameras in the network. The results of the network adjustment at each zoom settings are summarized in Table 3. It is worth noting that for the widest focal length, the bundle adjustment did not converge if the affinity factor was not included.

Figure 5 shows the distortion profiles for the two cameras. It is evident from the graphs of the lens distortion that the lenses used in the stereo camera have typical profiles for consumer level digital cameras[15]. The radial lens distortion is consistent for the left and right cameras, which is expected for matched lenses, as radial distortion is primarily a function of lens design. The decentering lens distortion varies between the two cameras, as this is primarily a function of co-axial centering of the lens elements. It is also apparent that the magnitude of the distortion profiles is varying in a consistent manner as the principal distance is changing, which again is consistent with previous studies[16,17]. However what is unusual in these results is the very large magnitude of the decentering distortion profile, particularly at the widest field of view setting, indicating a low quality lens assembly, a special zoom optical design[13] or possibly correlations between parameters in the network solution.

The target coordinates obtained at each focal length from the bundle adjustment computation were compared to the reference coordinates by means of a similarity transformation. The RMSE of the transformation are reported in Table 3 (rows RMSE_XY and RMSE_Z). For most of the configurations, the RMS error is significantly larger than the theoretical prediction for the precision of the XYZ coordinates. The discrepancies are most prominent for the widest and telephoto fields of view. In general it is likely that the differences are caused by a combination of factors such as a parameter correlations, systematic errors and an optimistic estimate of the image measurement precision factor ε.

Table 3. Summary of the results for the network adjustments at each zoom setting.

	Focal Length settings			
Main characteristics of the bundle adjustment	**Widest**	**Widest + 1**	**Central (Widest + 4)**	**Tele (Widest + 9)**
Focusing distance [m]	1.2	1.3	1.5	3
Average GSD [mm]	0.2	0.25	0.2	0.2
Average intersecting angle per point [degrees]	74	84	74	74
Format percentage covered by image observations	95	94	95	94
Image marking residual RMS [pixel]	0.27	0.21	0.18	0.22
Computed precision σXY [mm]	0.006	0.008	0.007	0.004
Computed precision σZ [mm]	0.013	0.024	0.021	0.009
RMSE_XY [mm]	0.045	0.025	0.032	0.039
RMSE_Z [mm]	0.060	0.034	0.057	0.068

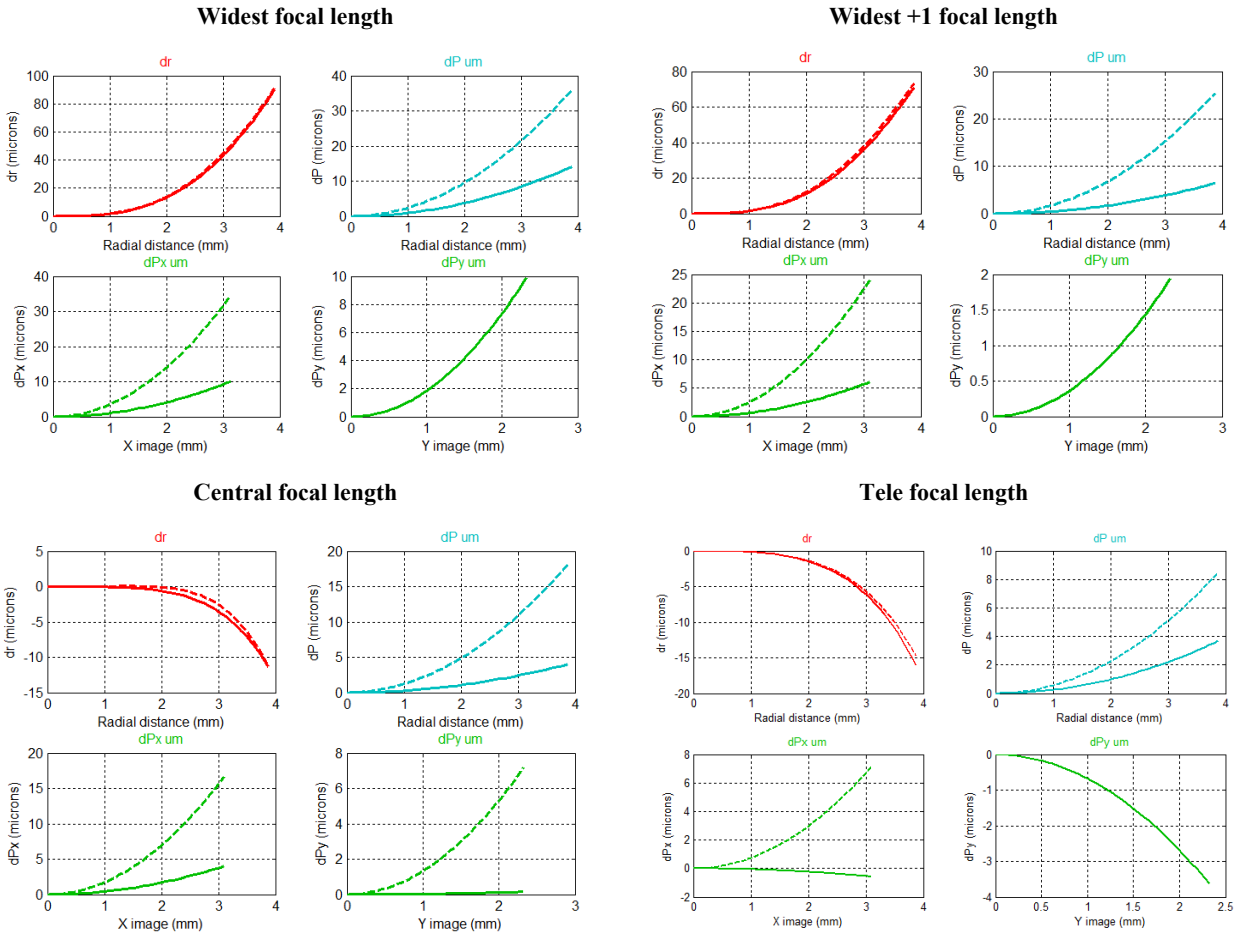


Figure. 5: Distortion curves for left (solid line) and right (dashed line) cameras.

Table 4. Comparison of relative orientation results at each zoom settings.

	RIGHT camera relative orientation							
	Widest		Widest + 1		Central (Widest + 4)		Tele (Widest + 9)	
Asymmetrical relative orientation parameters	**Value**	**Std**	**Value**	**Std**	**Value**	**Std**	**Value**	**Std**
X [mm]	77.36093	0.2811	77.35552	0.2099	77.46489	0.2490	77.73056	0.8565
Y [mm]	-0.83129	0.2595	-0.76463	0.2438	-0.88389	0.2478	-0.77487	0.57080
Z [mm]	-1.15600	0.3874	-1.07008	0.8223	-1.22961	0.7794	-1.87968	2.0964
ω [degrees]	0.235998	0.0211	0.237254	0.0218	0.199681	0.01270	0.114768	0.0143
φ [degrees]	2.768877	0.0161	2.818508	0.0150	2.778874	0.0108	2.781689	0.0149
κ [degrees]	0.147228	0.0069	0.143614	0.0055	0.157231	0.0034	0.147984	0.0088

The relative orientation parameters computed from the bundle adjustment using equations (6-8) are shown in Table 4. With the exception of some of the parameters for the telephoto setting, based on the derived precisions there are no significant variations in these parameters, indicating a high level of consistency in the relative orientation. The precision and reliability of the exterior orientation parameters is poorest for the narrowest field of view setting, so larger variations in the parameters and the RMS errors in the target coordinates are to be expected.

Figure 6 shows the comparison of the calibrated principal distances for the left and right cameras for each zoom setting used. The nominal focal lengths read from the *Exif* tag are also reported. It is clear from these results that there is a significant and consistent scale factor, increasing from the wide to telephoto setting, between the nominal principal distances and the values derived from the photogrammetric solutions.

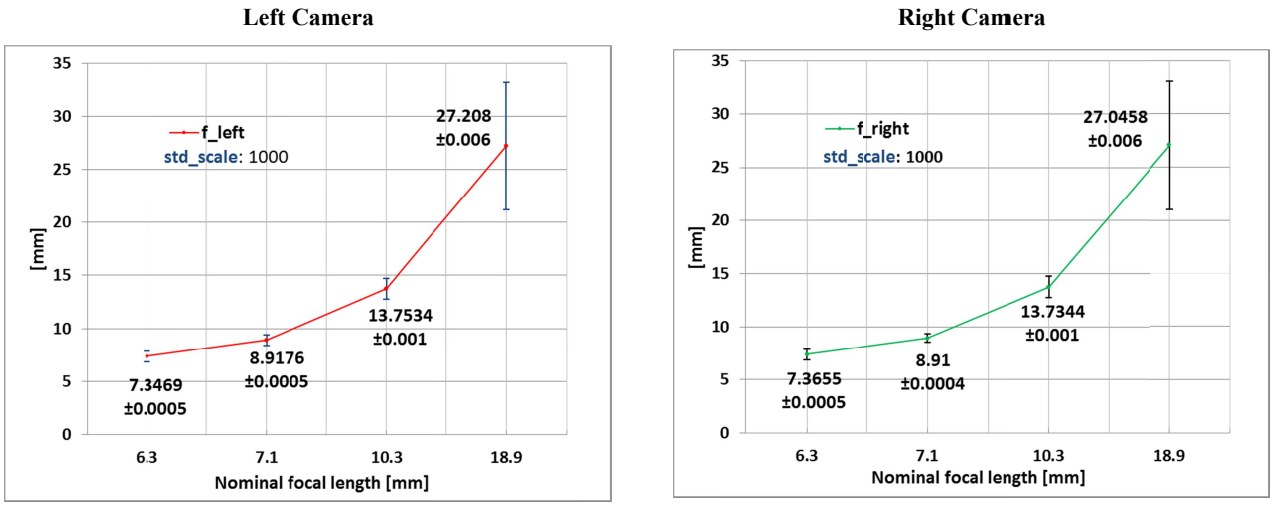


Figure 6: Focal length variation in the left and right cameras.

This scale factor issue may be caused by errors in, or mis-interpretation of, the manufacturer specifications or pixel spacing on the sensors, coupled with correlations between the calibration parameters and exterior orientation parameters such as baseline (Figure 7a) and the affinity term (Figure 7b). The variation in the baseline (Figure 7a) is not statistically significant (Student's t-test) and once more indicates good stability for the relative orientation.

From Figure 7b and Figure 8, it is evident that the affinity factor cannot be neglected especially when using the camera at the widest field of view. As noted above, the large magnitude of the affinity term may be a factor in the scale variations between the nominal and derived principal distances. Figure 9 shows the variation in the principal point location in the X and Y image space coordinates, variation typical of a zoom lens[16,17]. The smooth variation is a function of the change in the positioning of the lens elements and slight non-perpendicularities between the optical axes and the image planes.

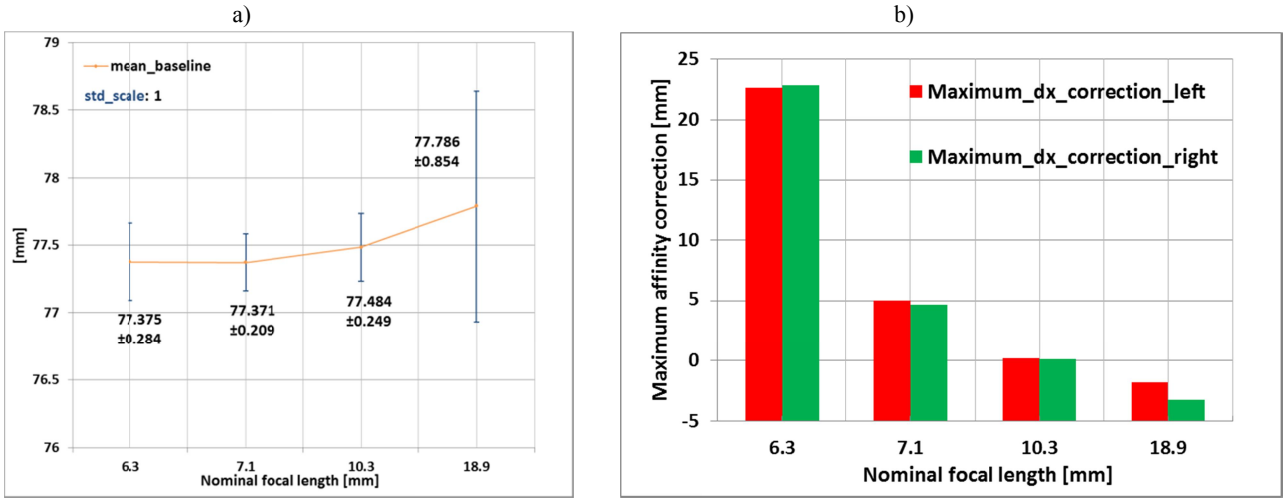


Figure 7: Baseline length variation in mm (a) and affinity factor as correction dx in pixel (b).

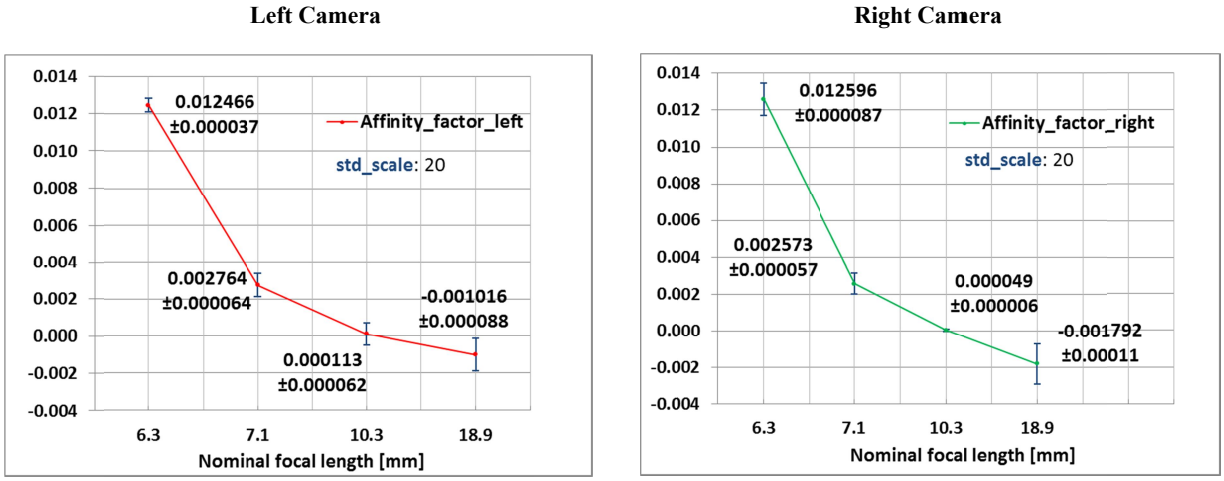


Figure 8. Affinity factor variation in the left and right cameras.

SINGLE PAIR ACCURACY TESTS

To assess the accuracy of the Fujifilm Finepix Real 3D W1 for 3D data measurements, different tests were performed.

Single point accuracy test

The test aims to investigate the accuracy of 3D coordinates using a single image pair with different calibrations and exterior orientation parameters. The test field shown in Figure 4a was imaged with a single image pair acquired at a distance of about 1.2 m. The expected theoretical accuracy for automatic image measurements (ε=0.2 pixel) is σ_{XY} = 0.056 mm, σ_Z = 0.857 mm and σ_{XYZ} = 0.861 mm.

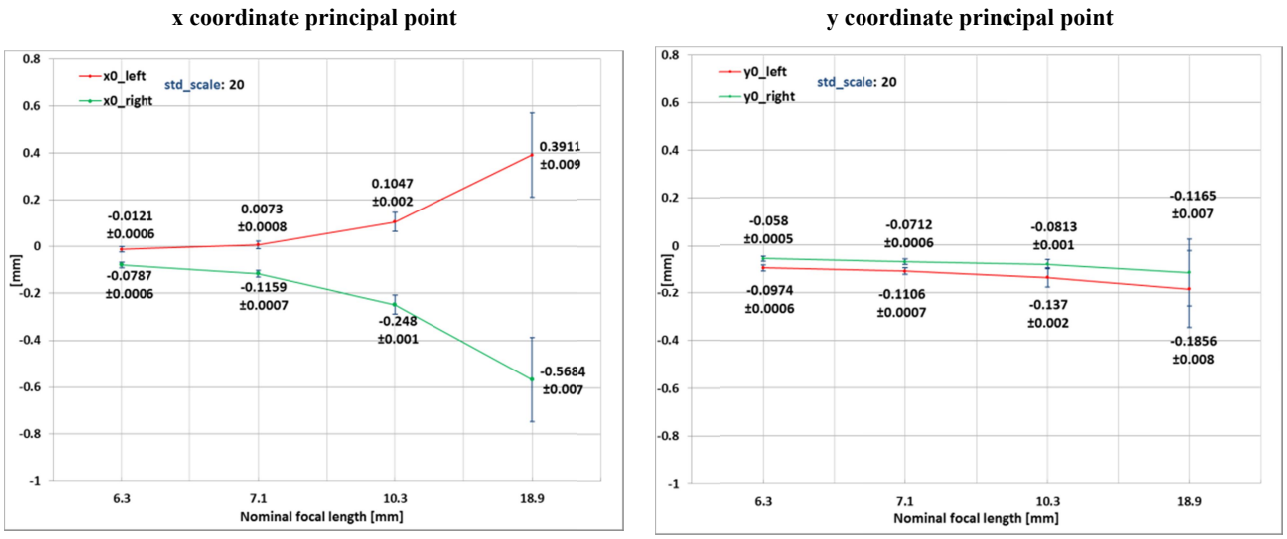


Figure 9: Principal point variation in the left and right cameras.

Different sets of parameters are used to simulate the use of the camera in common operative scenarios:

A) calibrations and mean exterior orientations from the unique bundle adjustment obtained using equations 6-8;

B) calibrations from the unique bundle adjustment, image orientation on 9 points well distributed on the object and scaling based on the distance of the two cameras from the mean baseline obtained using equation 8;

C) non-expert scenario: the MPO is imported in the free software Agisoft Stereoscan. The images are oriented and calibration parameters refined, the scaling is automatically assigned on the basis of the stored value which is 77 mm.

To assess the accuracy of these different scenarios, calibrations and exterior orientation parameters were imported in a photogrammetric software, the circular targets of the test field were marked on the image pair and triangulated with forward intersection through least squares adjustment. The achieved 3D coordinates were compared to the reference points by means of a similarity transformation and then a RMSE was calculated. The test shows the best accuracy achievable with a single pair and automatic image measurements with 0.2 image measurement precision. All the methods do not require reference measurements such as scale bars or control points on the object. The test was repeated over an extended period to assess the stability of the results over time: the camera calibration was performed in November 2011 and the same parameters were used for the successive tests.

Table 5. Summary of the RMS errors between the test field reference target coordinates and the stereo camera measurements.

Test Version	RMSE X [mm]	RMSE Y [mm]	RMSE Z [mm]	RMSE_VECTOR_LENGTH [mm]
A_NOV2011	1.189	1.051	0.757	1.758
B_NOV2011	0.647	0.566	0.888	1.236
C_NOV2011	3.330	1.959	5.027	6.340
A_FEB2012	1.736	1.543	0.902	2.491
B_FEB2012	0.587	0.494	1.278	1.491
C_FEB2012	1.826	3.498	4.889	6.283
A_MAR2013	3.690	3.707	1.478	5.435
B_MAR2013	0.463	0.484	0.855	1.086
C_MAR2013	4.523	4.783	5.315	8.461

For the scenario B, only the image orientation was recomputed while for scenario C the MPO files were imported in Agisoft Stereoscan and the orientation and adjustment of calibration parameters was performed again.

The results are summarized in Table 5 where there are two notable trends. First, method C based on the Agisoft solution, has a consistently poorer RMS error due to the inadequate parameterization of the solution and the use of a nominal baseline. Second, there is a general deterioration in the RMS errors with time, as it could be expected because of the inevitable drift in the calibrations and relative orientation of the cameras. Method B produces the most consistent results, most probably because of the re-orientation of the cameras based on the 9 well distributed points. Only in one or two cases is there reasonable agreement between the actual and predicted precision for the Z coordinates of the targets.

Flatness measurement error of computed point cloud

One of the best potentialities of a stereo-camera is that once the relative exterior orientation of the two cameras is known, 3D measurements can be easily performed. This is particularly suitable for automatic image measurements such as point cloud generation. In this test the accuracy of a point cloud generated using the exterior orientation obtained with the methods B and C described in the previous paragraph is assessed.

In this test the stereo-camera was positioned with the optical axis orthogonal to a planar surface (whose flatness is certified) at a distance of about 1 m. A random texture was projected on it using a DLP projector and an image pair was acquired. The point clouds were computed with image matching algorithm and a *flatness measurement error*[11] was computed as residuals from the best fitting plane evaluated from the point cloud. Figure 10 depicts the flatness measurement errors respectively for the scenarios B on the left and C on the right. From a global analysis of the residuals from the best fitting plane and their position within the captured area, it is evident that some systematic effects can be readily identified for the method C. These systematic errors can arise from an incomplete or inappropriate mathematical model or incorrect calibration parameters. This was likely due to the incomplete or poorly estimated additional parameters computed by Agisoft Stereoscan. For the method B on the left systematic errors are less evident, the error follows a Gaussian distribution and the magnitude of noise is in the order of the one expected from theoretical precision. This test is not able to point out in-plane errors, such as scale effects.

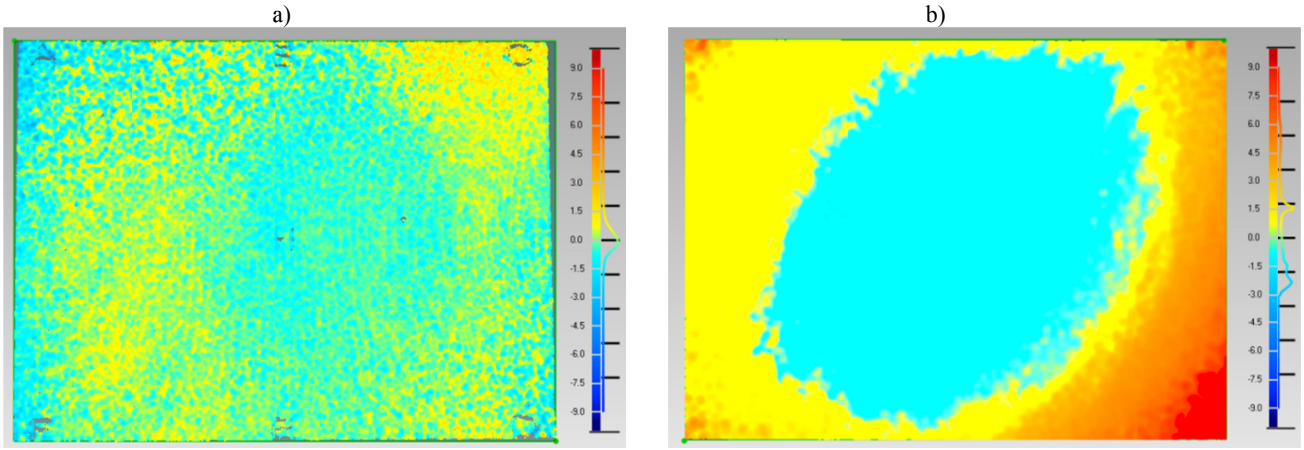


Figure 10. Flatness measurement errors respectively for the scenarios B (a) and C (b).

3D modeling of a small statue

To assess the accuracy for 3D modeling purposes using dense image matching methods, a 35–centimeter high bust in white opaque plastic material was chosen. The Fuji Real 3D W1 was arranged on a tripod on a side of a DLP projector used to project a random pattern on the bust (Figure 11a). The camera was at an average distance of 0.5m from the bust. The ground sample distance on the bust was in this case approximately 0.1 mm. 31 stereo-pairs were acquired rotating the bust at equally spaced angular intervals through 360 degrees.

The collected stereo-pairs were processed following the methods A and C previously reported. For the method A, the interior and exterior orientation were converted for the PMVS2 matching software[18] and a series of point clouds was generated. Each point cloud comprised about 200,000 points. The point clouds were cleaned, triangulated and finally aligned with an ICP algorithm. A final triangular mesh of about 1 million polygons was obtained (Figure 11b).

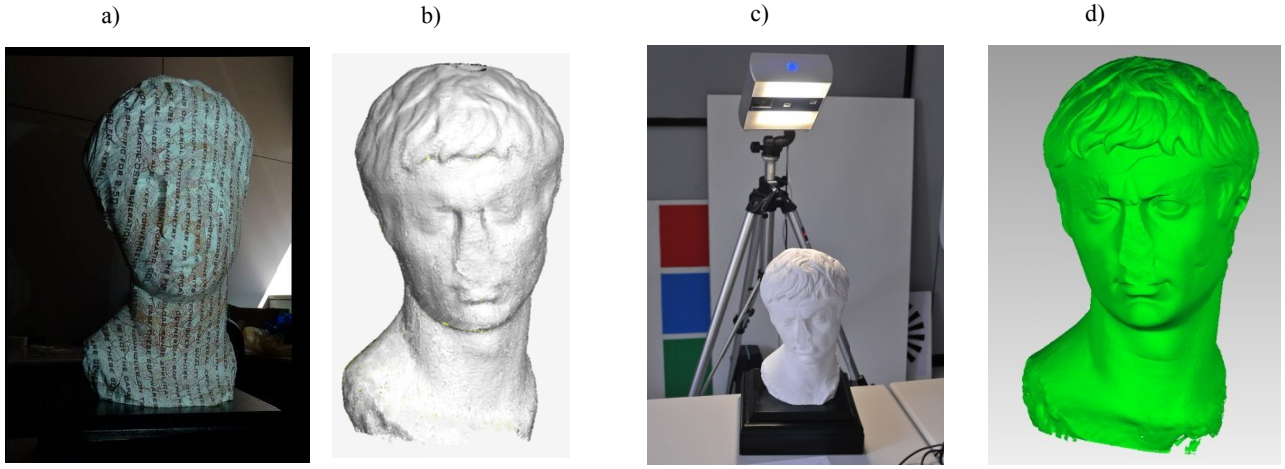


Figure 11. a) Random pattern projected on the 35 cm high statue used to assess the accuracy for 3D modeling purposes of the Fujifilm Real 3D W1. b) 3D model generated the Fujifilm Real 3D W1adopting the method A. c) 0.1mm accurate triangulation-based laser scanner used as reference. d) 3D model generated with the triangulation-based laser scanner.

For the method C, the single triangular meshes were aligned with ICP algorithm but the model did not close correctly, probably because of the systematic deformations described in the previous section.

The 3D model obtained from the method A was compared with the 3D surface model derived using a more accurate triangulation-based laser scanner (about 0.1 mm, Figure 11c and Figure 11d). The model generated with the Fujifilm Finepix Real 3D W1 was aligned using an ICP method to the reference model hence errors were calculated as Euclidean distances of each point of the 3D model to the closest triangle. The RMS error is again commensurate with the theoretical measurement precision (Figure 12).

CONCLUSIONS AND FUTURE DEVELOPMENTS

A rigorous photogrammetric approach has been used to characterize the Fuji Real 3D W1 stereo camera. The theoretical precision has been predicted using standard error propagation, and compared to practical testing in three measurement scenarios.

A number of different methods of determining the relative orientation of the two cameras has also been tested to assess which method is the most accurate and to identify systematic errors. The average relative orientation computed from the bundle adjustment furnishes the most direct estimates of the parameters without the requirement for reference targets on the object. The most accurate procedure requires that some points on the object are visible on both the images and are well distributed over the image format. A non-expert scenario has been investigated using a free computer vision software. Some tests were repeated over a period of more than one year and a good stability of interior and exterior parameters was ascertained for some months whilst only the interior orientation parameters were really stable after one year. For very short distances, less than 1 m, the camera can be used 3D measurement as accurate as 1 mm.

The next phase of this research project is to introduce constraints on the baseline and relative angles into the bundle adjustment. During this study the baseline constraint was tested in the bundle adjustment but the results were inconsistent across the different configurations. Additional development and analysis will be carried out and reported in a future paper.

ACKNOWLEDGMENTS

The authors are thankful to Manfrotto, a Vitec Group company, for their kind support.

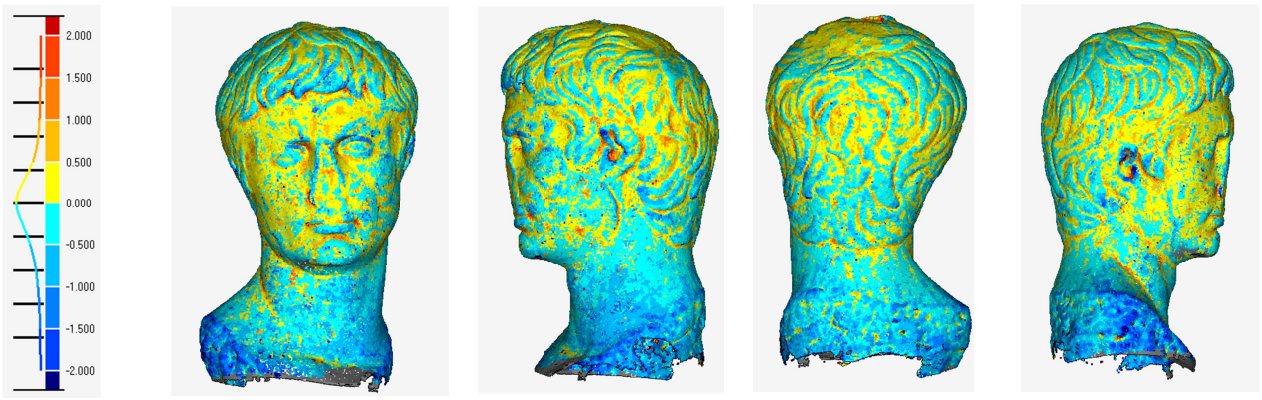


Figure 12.Comparison (in [mm]) between the 3D digital model generated with the Fujifilm Finepix Real 3D W1 and the same model from a more accurate triangulation scanner.

REFERENCES

[1] Louvre, 2013. http://www.louvre.fr/en/museum-audio-guide Accessed 24 March 2013.

[2] Menna, F., Remondino, F., Battisti, R., Nocerino, E., "Geometric investigation of a gaming active device", Videometrics, Range Imaging, and Applications XI, Proc. SPIE Vol. 8085, Remondino/Shortis (Eds), pp. 80850U-80850U-12 (2011).

[3] Akca, D., Gruen, A., "Comparative geometric and radiometric evaluation of mobile phone and still video cameras", The Photogrammetric Record 24(127), 217-245, (2009).

[4] Wackrow, R., Chandler, J. H., Bryan, P., "Geometric consistency and stability of consumer-grade digital cameras for accurate spatial measurement", The Photogrammetric Record 22(118), 121-134, (2007).

[5] Guidi, G., Remondino, F., Morlando, G., Del Mastio, A., Uccheddu, F., Pelagotti, A., "Performance evaluation of a low cost active sensor for cultural heritage documentation", VIII Conference on Optical 3D Measurement Techniques - Gruen/Kahmen (Eds), pp. 59-69, Vol.2, Zurich, Switzerland (2007).

[6] Remondino, F., Fraser, C., "Digital camera calibration methods: considerations and comparisons", Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, Vol. 36(5), 266-272 (2006).

[7] Kahlmann, T., Remondino, F., Ingensand., H., "Calibration for increased accuracy of the range imaging camera Swissranger™", Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, Vol. 36(5), pp. 136-141 (2006).

[8] Chandler, J. H., Fryer, J. G., Jack, A., "Metric Capabilities of low-cost digital cameras for close range surface measurement", The Photogrammetric Record, Vol. 20(109), pp. 12-26 (2005).

[9] VDI/VDE, "Optical 3D measuring systems / Optical systems based on area scanning", in VDI/VDE manual on Metrology II – VDI/VDE Guidelines, Beuth, Berlin, Germany, (2002).

[10] Takahashi, G., Matsuoka, R., "Accuracy of measurement using a pair of stereo images acquired by Finepix Real 3D W1 without controls", Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38(5), pp. 565-570 (2010).

[11] Matsuoka, R. Takahashi, G., Asonuma, K., "Calibration of low-cost measurement system by using a consumer digital stereo camera", Videometrics, Range Imaging, and Applications XI, Proc. SPIE Vol. 8085, Remondino/Shortis (Eds), pp. 80850U-80850U-12 (2011).

[12] Wong, K. W., "Mathematical formulation and digital analysis in close range photogrammetry", Photogrammetric Engineering and Remote Sensing, Vol. 41(11), pp. 1355-1371 (1975).

[13] Sato, K., "Bifol imaging optical system and imaging apparatus", United States Patent, US 7,724,445 B2 (2010).

[14] Brown, D.C., "Close-range camera calibration". PE&RS, Vol. 37(8), pp. 855-866, (1971).

[15] Chikatsu, H., Takahashi, Y., "Comparative evaluation between consumer grade cameras and mobile phone cameras for close range photogrammetry", Videometrics, Range Imaging and Applications X, Proc. SPIE Vol. 7447, Remondino/Shortis/El-Hakim (Eds.), paper 74470H (2009).

[16] Wiley, A.G. and Wong, K. W., “Geometric calibration of zoom lenses for computer vision metrology”, Photogrammetric Engineering and Remote Sensing, Vol. 61(1), pp. 69-74 (1995).
[17] Fraser, C. S., and S. Al-Ajlouni, "Zoom-dependent camera calibration in digital close-range photogrammetry." Photogrammetric Engineering and Remote Sensing, Vol. 72 (9), pp. 1017-1026 (2006).
[18] Furukawa, Y., Ponce, J., “Accurate, dense, and robust multi-view stereopsis”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32(8), pp. 1362-1376 (2010).