

## THE INFLUENCE OF RECORDING OF LONG-SERIES IMAGES TAKEN FROM NON-METRIC CAMERAS ON INCREASING THE QUALITY OF MODELLING OBJECTS WITH THEIR SURROUNDINGS

Bogdan Jankowicz

### Summary

The diversity of methods used both by acquiring data images and later in the creation of a 3D model of an object, gives an opportunity to choose the best possible technique and processing method. It is not always necessary to use the most expensive solutions.

In the terrain analysis and local landscape assessment the use of the modelling method, based on recording of series images taken by non-metric cameras, seems optimal, and is economically attractive.

### Keywords

photographic recording of long-series images • non-metric cameras • 3D modelling

### 1. Introduction

The development of IT technology, including the high development dynamic of image information systems, combined with the Internet as a means of communication and sharing information, is still accelerating, and the newest methods are hard to accept by many researchers. Moreover, in photogrammetry non-metric cameras – which, a short time ago, were there considered controversial – are more and more useful, and as they are constantly improved they may in some situations replace costly and more professional devices. The activity of international organizations, like Open Geospatial Consortium (OGC), and famous firms like Google, Microsoft and Autodesk is also favourable to the development of this field. The new solutions for generating 3D models in the cloud computing environment, in which the data processing is done largely without the participation of the project's author, add a new dimension to the problem of representation, exchange and sharing spatial data over the Internet.

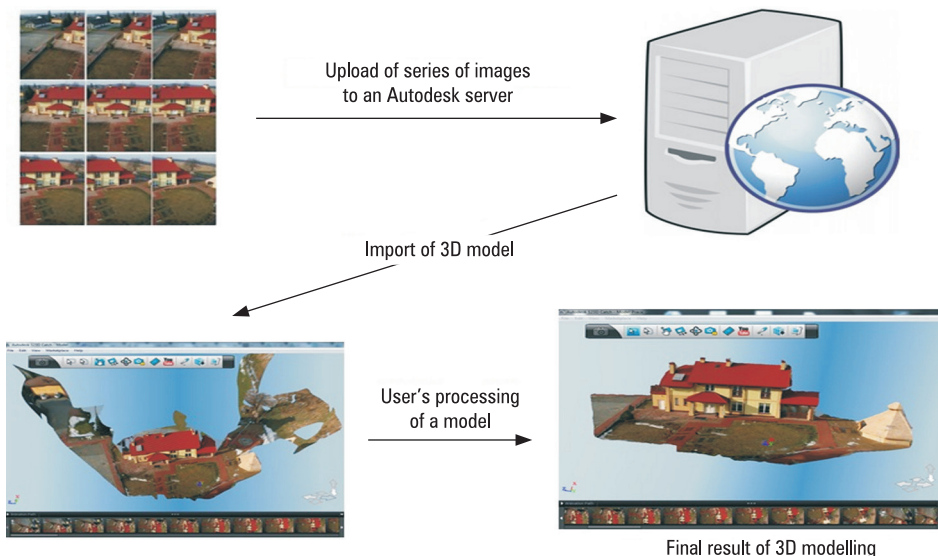
## 2. The research methodology

Taking the above aspects into account, it is possible to acquire visually attractive models of chosen objects or of whole areas in 3D. In case of bigger objects (or objects images with their surroundings), image data can be collected both from terrestrial and above ground stations, where low-altitude images, taken e.g. by unmanned aerial vehicles (UAV), play the special role.

The numerous studies carried out by the author proved that the so-called long-series technique of recording images increased the accuracy of 3D models created with applications based on SIFT algorithm. The studies were conducted in Kraków on many architectural objects but only the following examples will be described below: the monument of Florian Straszewski, a residential building with its surroundings and the interior of the St. Wojciech Church.

To generate 3D models of digital photos of objects or areas the free application 123DCatch (Autodesk), a continuation of Photofly project, has been used.

The application is based on cloud computing, which means that all the operations, including calculations, are carried out over the Internet, and they are performed by means of large information resources, that is, among other things, by network of servers connected by a computational centre. The files are uploaded to a server, processed and stored outside the PC of a project's user or developer, who can remotely access and edit them from any computer via the Internet.



Source: author's study

**Fig. 1.** Cloud computing and 3D modelling

This way one can use the service and be provided with great computational power, with no need to overload the processor of a user's PC, to purchase the licence or install the registered software. Figure 1 shows the schematic representation of a cloud computing and its uses.

During cloud computing, based on many algorithms, the 123D Catch programme generates from images a point cloud, connected in an irregular net of triangles. The net is a graphic equivalent of the studied object, or more precisely, of its outer surface. The net is geometrically delineated, since each point has its spatial coordinates.

The authors' studies proved that a camera's resolution above 3 megapixels has no considerable influence on the result of modelling and the accuracy of observations. The latter can only be improved by using wide-angle lenses, but they are not implemented in simple digital cameras. On the other hand, too large wide-angles can increase the distortion of the images. The suggested camera's features are enough to create 3D visualisations in an easy and fast way, without the need to use professional equipment and software usually requiring the purchase of the license.

Moreover, the measurement stations do not have to be stabilized and images do not need to be taken orthogonally. The angle scale of taking pictures should be 5–15 degrees and the number of images should range from 3 to 50 pictures, although the author's earlier observations indicated that the computation and edition of a model is most efficient when it is generated from more or less 20 pictures [Jankowicz 2014]. The main principle is matching the measurement session with one's own scope of observation, but it should be remembered that the camera should take at least 3 pictures of each point.

It is the technique of taking pictures that has a significant impact on the quality of the model generated in the 123D Catch software. The photographed object should be well lit. The flash lamp should not be used in taking pictures, because it produces shadows. The weather can also be a problem, e.g. when it is too sunny. Taking pictures against the sunlight leads to overexposure. On the other hand when the sky is very cloudy the quality of images worsens, because there is not enough light. When the consecutive pictures are taken in conditions with large differences in illumination of the object, we can convert the images to black and white or to one colour scheme, but other methods of editing images should be avoided. Special attention should be paid to the distance from the photographed object and the focus of the camera, because the pictures need to be clear.

Some elements of object's surroundings, such as repetitive patterns or trees, can cause difficulties in the modelling process, due to the mechanism of features recognition and correlation of images. In that case errors are very likely to occur, because the software, in the matching process, also captures the object's surroundings. Therefore, the application can falsely recognize different leaves and twigs as identical or mutually corresponding. Similar problem can occur with buildings with large number of windows, the order of which is not properly recognized and remembered by the software. This happens especially when open spaces or buildings' elevations located along the streets are photographed – in such cases even 50% of images can be matched incor-

rectly. To overcome that issue the homologous points must be marked. The analogous problem concerns the monochromatic, uniform, smooth and mirror surfaces – which the software sometimes cannot manage. The mirror surfaces reflect the sunlight and the surrounding differently even at the slight change of camera's angle, and that changes their appearance on the pictures.

### 2.1. Correlation of digital images

When finding common features of pictures, finding correlations between them and linking them, the principle of image correlation is applied. The correlation or matching is a process of identification and measurement of homologous points on at least two images. It consists in matching pictures by choosing characteristic points on one photo and finding identical points on other images. In case of digital images the process is usually carried out automatically and is called image matching or automatic stereo matching, and when large number of images is concerned it is called multiple image matching, which is done by using appropriate algorithms.

The practical side of correlation has the following stages: determining the characteristic points of matching, finding their equivalents on other photos, calculating the spatial position of matched elements and defining the accuracy of the matching. The correlation can be used, among other things, in calibration of cameras with software, in interior orientation, relative and absolute, and in the spatial modelling itself.

There are three ways of matching: area based matching (ABM), feature based matching (FBM) and relation matching (RM). The ABM method is based on studying stereograms, the structure of which is clear. The FBM method consists in dividing the pixel image into groups of simple elements and their correlative analysis, whereas the RM is based on finding topographic relations. Each of them uses different computational algorithms.

### 2.2. The conversion of 2D images to 3D model

In the automatic conversion of 2D images to 3D model the traditional method of photogrammetric image correlation is just the first stage of processing the observations. That is why most software used for conversion of ordinary flat digital photos to 3D model is still based on structure from motion (SFM) technology. It can be colloquially described as monitoring and description of points position on the basis of perspective change or the stations settings of the measuring camera.

This method, like in the correlation of a pair of flat photos, consists mainly in identifying homologous points on a few different photos and in finding the function of relation between them by monitoring the change of their position, also their mutual position, on subsequent photos of a series. There should be at least three photos. The computational algorithms automatically check the degree of conformity between discrete image elements, namely pixels. There could be two methods of assigning points, that is narrow and broad bases.

The technique of narrow bases can be used on assumption that the camera station and its orientation are fixed, and pictures do not differ little from one another. In such a case the correlation of pictures can be described as a brightness function of image elements. The intensity of pixels is compared when correlation or the sum of squares of deviations are used. By monitoring the whole sequence of images it is possible to reconstruct the arrangements and parameters of points movements with respect to one another.

The technique of broad bases is used, when distance between cameras stations is large, larger than in a method of narrow bases, where observation points are close to one another. Due to large distances and a certain degree of freedom in setting the stations, there can be considerable differences in the scale of images in one series. There can also occur errors in their position and problems with the accurate representation of an object, especially if the process of transformation of images is automatic. In case of the reconstruction based on non-metric images the required minimal number of homologous points is seven, which usually means large coverage. It can be difficult to achieve in case of large and elongated photographed areas [Roberston and Cipolla 2009].

### 2.3. Computational algorithms

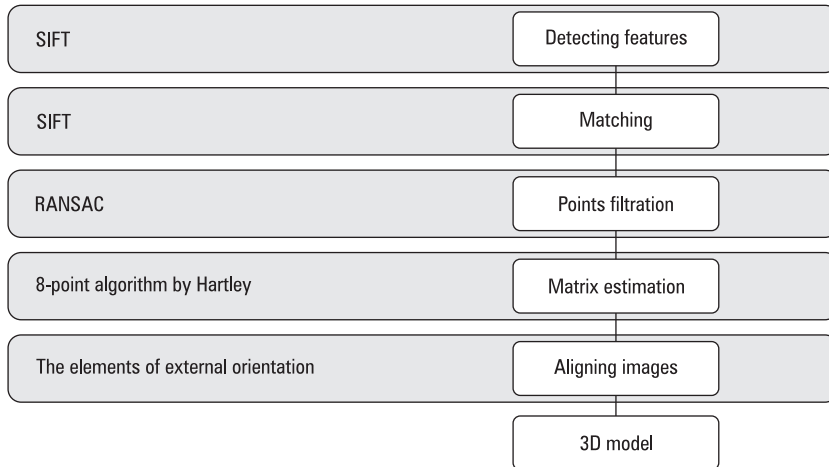
Sequential computational algorithm used in the structure from motion method in developing the collection of non-metric pictures is based on the so-called scale invariant feature transform (SIFT). It was proposed by David Lowe in 1999 and is commonly used in many applications for creating 3D models from pictures. Its principal goal is to identify and describe the key points. Their verification and mutual connection is possible thanks to the computational procedure called random sample consensus (RANSAC), using 8-point algorithm. A general outline of processes and computational algorithms required to identify the correlation between points on the pictures and to generate 3D model are presented in Figure 2.

The whole computation process in 123D Catch software, similarly to other applications for 3D modelling from users' pictures, takes place in the cloud computing environment. This way the software manufactures protect their technological solutions from illegal copying.

The graphic features in a SIFT algorithm are fixed with respect to scaling and rotating an image and to noise and distortions, and partially variable as to exposure or perspective. The large number of features can be singled out from ordinary pictures by means of efficient calculations. Moreover, they are characteristic enough to most likely correctly match a single feature, despite the large collection of analysed data. Thus the shapes and whole objects can be recognized.

As the authors' studies proved, good results in increasing the accuracy in making 3D models in applications based on SIFT algorithm can be achieved by using the technique of long-series images. That is why in a presented example (one of many) the long-series mode has been used to show its functional and accuracy superiority over single

images mode. The advantage spring from the fact that the exposure time of photodetectors is regulated only electronically through a reading, taken with defined frequency, of collected electric charges. Therefore, it does not causes the change of image orientation in continuous shooting mode.



Source: author's study

**Fig. 2.** Algorithm scheme in 2D to 3D processing

The long-series recording of images is a development of continuous shooting mode under the following conditions:

- the pictures are taken in a relatively long series (e.g. a few hundred shots),
- the memory card of the camera must be sufficiently large (it follows from the first condition),
- the regular (with a fixed time interval), sufficiently fast image recording,
- it is recommended that the images are recorded in a series as “frozen”, to avoid blur.

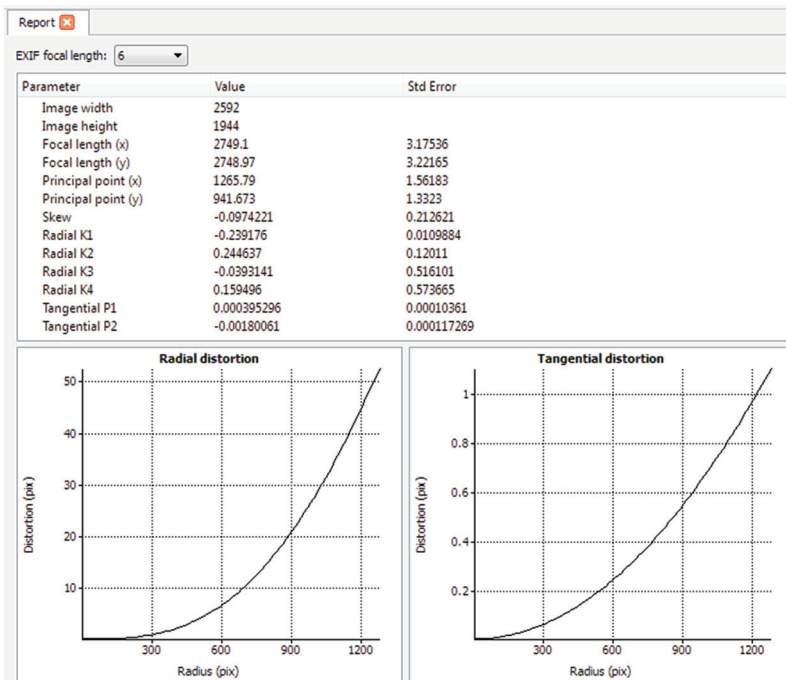
In case of single shooting mode the photo-mechanics of the camera plays an essential role, which decreases the effect of image stability. Among conditions that significantly influence the zoom lens stability of non-metric (digital) compact camera is the movement of zoom lens – sliding and rotary motion in taking single pictures, when the uniqueness of the lens position causes instability of image distance.

The monument of Florian Straszewski, located in Kraków, is one example of author's numerous studies of architectural objects. The model of the object has been measured with the use of pictures taken both in single and continuous shooting modes, and also – to test the results – the measurement were made directly on the object. Similar studies have been carried out on two other objects: residential building with its surroundings and the interior of St. Wojciech Church in Kraków.

2.4. The choice of non-metric camera and its calibration in Agisoft Lens software

According to what the developers of the 123D Catch software believe, and according to author’s studies, the increase of cameras’ resolution above 3 megapixels does not significantly influence the accuracy of measurements obtained from a 3D model, and therefore a non-metric camera Canon PowerShot SX100IS was chosen as a measuring device. In the manufacturer’s specifications the camera has a CCD type optical sensor of 1/2.5”, with a resolution of 8 Mpix and size 5.76 x 4.29 mm. It must be noted that the physical size of the sensor has the greatest influence on the quality of pictures rather the number of pixels. The focal length and lens parameters are adequately 6.0–60.0 mm and F2.8–4.3. The camera can shoot in a continuous mode with 1.3 frames per second or 0.8 frames per second speed with refocusing.

Though there is no need to enter a camera’s parameters into to 123D Catch software, the calibration of the Canon PowerShot camera SX100IS has been made to check the data declared by manufactures and to get to know other important features of the camera, that is the elements of its internal orientation [Litwin and Piech 2013]. The calibration was done with Agisoft Lens software using a computer monitor as a calibration target. Five pictures taken with the right settings, opened in the software, are used for further calculations. The Agisoft Lens calibration report is presented in Figure 3.



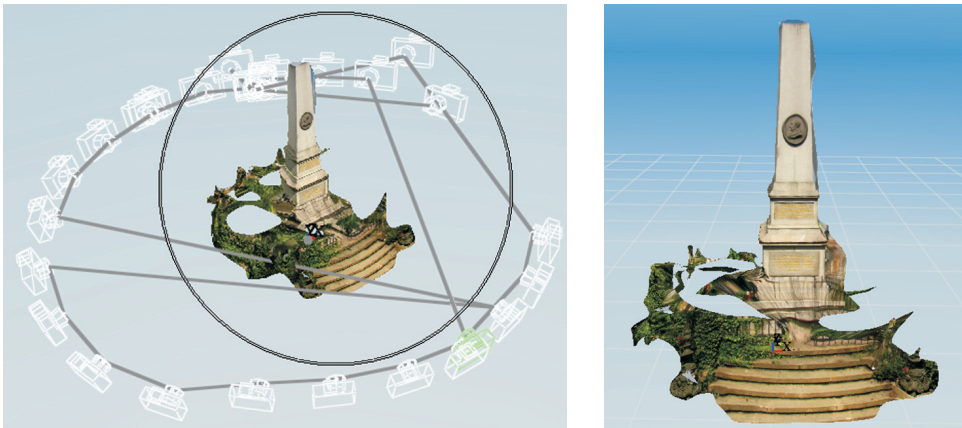
Source: author’s study

**Fig. 3.** Calibration report of Canon PowerShot SX100IS camera in Agisoft Lens software

The calibration report includes the following assessed parameters:

- the width and size of the image: 2592 and 1994 mm respectively,
- horizontal and vertical focal length: 2749.10 [pix] and 2748.97 [pix],
- coordinates of the focal point:  $X = 1265.79$  [pix],  $Y = 941.673$  [pix]
- radial distortion parameter  $K1, K2, K3, K4$  and tangential distortion  $P1, P2$ ,
- the inclination of the axis  $X$  and  $Y$ .

Apart from calculating the parameters of calibration, the scope of error for each value has been determined. And the values of radial and tangential distortions – in pixels, as a relation between the distortion and its radius – have been presented in charts.



Source: author's study

**Fig. 4.** Camera stations as presented in 123D Catch software and the final visualisation of the obelisk

#### 2.5. Comparison of the effects of 3D modelling: long-series technique of recording images versus single pictures technique

The field study and 123D Catch software confirmed that 3D models generated both from images taken by long-series recording method (Figures 5–6) and by single picture technique have high degree of accuracy, but a few differences in their applications should be noted. The differences are corroborated by other studies [Litwin and Pijanowski 2013].

The first visible difference between the models obtained by different image recording methods is the time of photographing of the chosen object. In spite of greater number of pictures in long-series mode, but thanks to an automatic focusing and the speed of taking pictures (around 0.8 frames per second) the chosen method proved to be much faster.

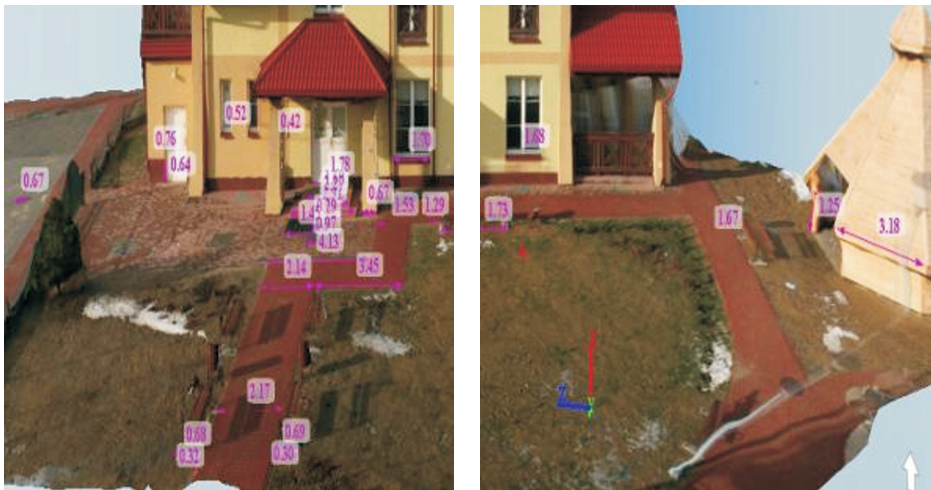


The second difference consisted in visual effect (distortion) of the generated model. The 3D visualisations generated from single pictures, without the manual definition of points that would correspond to one another on various pictures, had more distortions, and the contours were less clear than in the model from images taken in a series.



Source: author's study

**Fig. 5.** Examples of measurements on the monument model generated from long-series images



Source: author's study

**Fig. 6.** Another example of spatial analysis based on 3D modelling (by long-series photos) of the object with its surroundings

Another dissimilarity of results between the two methods of recording images lies in definition of images orientation and stations of the camera. In a long-series method the configuration of images was determined correctly. In the second method the order of

images was incorrect even when the corresponding points on the subsequent pictures have been manually determined.

In addition, though the measurements, carried out in models generated from images taken by both methods of photographing, were highly accurate, in comparison with the field conditions, the three-dimensional visualisation generated in continuous shooting mode had better accuracy parameters. Table 1 presents an accuracy comparison of both models of the above-mentioned monument.

**Table 1.** The accuracy comparison of 3D models generated from single and series images (results are rounded to the nearest centimetre)

Parameter	Long-series method of image recording	Single pictures method
$ \Delta $ max	0.01 m	0.02 m
$ \Delta $ min	below 0.005 m	below 0.005 m
$\Delta$ mean	below 0.005 m	below 0.005 m
Standard deviation $\delta$	below 0.005 m	0.01 m
The number of error-free measurements	70%	30%

The results of the studies summarized in Table 1 show the differences in accuracy of 3D models created with pictures taken by the two methods of photographing described above. Both the long-series method of recording images and single picture method the mean deviation between measurements on 3D visualisation and observations made during a direct measurement of the object was below the mean precision of field measurement, that is 0,5 cm, however the standard deviation was negligibly small in a series mode and 1 cm in the single mode. The measure of maximal deviation between observations made during the measurement of the object and the sections generated on 3D model did not exceed 2 cm. The minimal value of deviation, close to zero, expressed in centimetres, suggested high accuracy of observation. The same results were obtained in 70% of the data set when using the series technique and in 30% when using single images method. In the object presented in Figure 6, the accuracy results – due to the object’s size – have been somewhat worse, but they should be regarded as satisfactory:  $\Delta = 3$  cm. It means that long-series method of recording images in 3D modelling with the use of 123D Catch software has a clear advantage over the single mode of taking pictures.

To confirm the usefulness of the chosen method in survey, another 3D model was generated, the object of which was the interior of St. Wojciech Church, located in the Main Square in Kraków. It was done to test the usefulness of 3D models generated by this method.



Source: author's study

**Fig. 7.** A fragment of a model showing the interior of St. Wojciech Church



Source: author's study

**Fig. 8.** Selected elements of the model were compared with the direct measurement (7–12)

**Table 2.** The results of a measurement carried out on the 3D model and directly on the object [metres]

No. of line	Measurement of a model	Direct measurement	Difference
7	0.95	0.94	0.01
8	1.25	1.25	0.00
9	0.95	0.94	0.01
10	1.25	1.25	0.00
11	1.57	1.58	0.01
12	1.57	1.57	0.00
Mean	0.0152		
Standard deviation	0.0232		

### 3. Conclusions

Following the development of 3D modelling, the technology widely applied in various fields, the study was aimed at assessing its usefulness, while employing pictures from a non-metric camera in a single and long-series mode, in the survey of architectural objects and in gaining new information about the object's adjacent areas and landscape.

The distance measurement has been tested on a 3D model generated from 2D pictures taken by means of a long-series image recording technique and compared with the visualisation based on single pictures method and the direct measurements of the object. The measures described in this paper lead to the conclusion that the use of long-series mode of photographing has a positive influence on the accuracy of 3D model in comparison with the visualisation based on pictures taken in a single mode. The conclusion follows from the presented data.

The first advantage of the long-series method lies in the process of taking pictures. The single shooting mode in photographing the measured object is more time consuming than the long-series recording technique, even if the number of images of the object in these two methods is the same. Moreover, in single shooting technique another inconvenience appears, namely the necessity of setting the focus after each picture. Moreover pictures taken one at a time had more distortions related to the sunlight and it was more likely that the pictures would be overexposed.

Generating a 3D model in the 123D Catch software gave another opportunity to notice the difference between the two modes of taking pictures. In case of single pictures, identifying camera's stations and orientation of pictures was difficult, even after the homologous points were manually marked. The use of long-series method ensured the stability of shooting parameters. Additionally, the aesthetic effect of the visualisation was improved, the reproduction of real contours of edges and cornices of

the monument was better, and the number of the reproduced elements of the object's surroundings – larger. Thanks to it more information about the terrain was gathered and it was more accurately reproduced.

Though both modes of recording images proved to be highly accurate in reproducing the real objects, the long-series mode of taking pictures in creating a 3D visualisation proved to be more accurate: the number of correct measurements of the model was 70% in this technique, with only 30% of them in the single shooting mode. The maximal deviation value for single pictures amounted to 2 cm and was 1 cm larger than the deviation of the long-series mode.

To sum up, the long-series technique of recording images gives better results in the 3D modelling process than the single picture mode method. Moreover the former technique increases the accuracy of the measurements taken in 3D visualisations of the measured object. As it is very likely that this method leads to error-free lengths read on the model, this visualisation could be used in taking measurements that are inaccessible in the field by using traditional methods or in detecting the gross measurement errors.

The use of long-series technique of recording images by a non-metric camera is a way of lowering costs of reproducing and editing the situation in the terrain and makes its update more cost-effective. Furthermore, in connection with a free software, such as 123D Catch, the method is simple, fast and functional. When working in a cloud computing one can use high computational power without charging the computer of the user. This method of taking measurements of the studied object may become an innovative approach to the question of survey of small and large architectural objects with no need to use specialist devices and licensed software. Yet another advantage of the presented method is an easy way of developing and supplementing the methods of gathering information about objects and their surroundings, which is directly related to the question of developing the environment and landscape.

## References

- Hartley R., Zisserman A.** 2003. Multiple view geometry in computer vision. Cambridge University Press, New York.
- Jankowicz B.** 2014. Wielkoseryjne obrazy niskopułapowe w modelowaniu 3D dla analiz przestrzennych terenu. Wydawnictwo Uniwersytetu Rolniczego w Krakowie.
- Robertson D.P., Cipolla R.** 2009. Structure from motion, [In:] M. Varga (ed.), Practical image processing and computer vision. John Wiley, New York.
- Litwin U., Piech I.** 2013. Application of the photogrammetric method in surveying of the palace and park complex building in Mściwojów, Poland, Geomat. Landman. Landsc., 2, 37–44.
- Litwin U., Pijanowski J.M.** 2013. Pilot project in Mściwojów, as an example of activities aimed at the protection and shaping of a landscape based on 3D visualization, Geomat. Landman. Landsc., 2, 45–55.

Dr inż. Bogdan Jankowicz  
Uniwersytet Rolniczy w Krakowie  
Katedra Geodezji Rolnej, Katastru i Fotogrametrii  
31-198 Kraków, ul. Balicka 253a  
e-mail: [rmjankow@cyf-kr.edu.pl](mailto:rmjankow@cyf-kr.edu.pl)