1 INTRODUCTION

Software based on the technology of computer vision for the processing of digital images is now widely applied for aerial surveys, the creation of Digital Model Reliefs (DMR), orthophotos through the automated creation of textured 3D models of any objects (for the purposes of reconstruction, documentation or presentation), and the highly accurate non-selective measurements of deformations (Pavelka et al., 2013). The indisputable advantage of this treatment method is the relatively low cost of the equipment (software and hardware). The offer of software allowing fully automated processing of images into 3D models is already very varied on the market; it ranges from freely provided web applications (Hypr3D, Autodesk 123D Catch, ARC3D Webservice, Microsoft Photosynth, etc.), via software distributed under General Public License (GPL) such as VisualSFM, to commercial software such as Agisoft PhotoScan or PhotoModeler Scanner 2012. In addition to the software, a quality digital camera (most digital SLR) and a sufficiently powerful computer are needed. The cost of the above equipment often represents only a fraction of the cost of special 3D laser or triangulation scanners (Hanzalová and Pavelka, 2012), while the time-consumption of the process mainly depends on the performance of the computer used (especially its CPU and RAM).

2 STRUCTURE FROM MOTION

Structure from Motion (SFM) is generally a process that is looking for the three-dimensional structure of an object by analyzing local motion signals across time. It is also part of the fields of computer vision and visual perception. In computer vision, SFM is especially used in a situation when the observed scene is stable, and only the position of the camera varies around the object. By analyzing the changes between individual frames, finding the significant points and their pairing, it is possible to create a three-dimensional reconstruction of a scene. In this processing algorithms such as Scale Invariant Feature Transform (SIFT) or Random Sample Consensus (RANSAC) can be involved as subtasks, because the input for SFM are 2D measurements containing noise and gross errors. The two-dimensional location of key points in an image depends on the spatial coordinates, the 3D relative motion between the camera and scene elements, and the internal orientation of the camera.
The geometric theory of SFM allows for the calculation of projection matrices parallel with the 3D points using only the corresponding points in different views. Formally, if we are given \( n \) points \( u_{ij}, i \in \{1 \ldots m\}, j \in \{1 \ldots n\} \) projected on \( m \) images, the task is to find the projection matrix \( P_1 \ldots P_m \) and also a consistent structure \( X_1 \ldots X_n \).

With regard to the search for correspondences between significant points (features), there are two basic approaches, i.e., either the above-mentioned feature-based approach (SIFT + RANSAC) or a pixel-based approach (registration based on patches with the Lucas-Kanade algorithm), which is more suitable for video sequences (Cippola, 2008).

3 AGISOFT PHOTOSCAN

Developers from the Russian company Agisoft have created a very user-friendly software called “PhotoScan”, which produces high quality results usable in nearly every sector from archeology through aerial surveys to reverse engineering. With the creation of this software, they used already existing works in the field of image processing (Bundler, SFM), but they overwrote the algorithms used again, added many of their own improvements, and achieved both a speed up of the calculation process and an increased accuracy of the results (Semyonov, 2011). PhotoScan is able to process any images, even in the absence of EXIF data, so it can be used, for example, for the processing of scanned original analog images or frames extracted from a movie.

The orientation of images, the calibration of cameras, and the generation of 3D models and textures are fully automated, but they are divided into several adjustable steps, thus giving the user more freedom in decision-making. The resulting 3D model can then be transformed into any coordinate reference system using Ground Control Points (GCP) or simply defining the scale over 2 points (Semyonov, 2011). The software is constantly being improved, and has a wide supporting community in the form of forums. Although it has not yet reached version 1.0, at Intergeo 2012 in Hannover it was recommended by the majority of companies selling Unmanned Aerial Vehicle (UAV) systems for the processing of aerial photographs. Numerous studies (Verhoeven, 2011; Doneus, 2011; Marčiš, 2012, etc.) confirm that from the point of view of experts in the field of photogrammetry, the software is very simple in terms of control and allows for the use of photogrammetry for the creation of 3D models for users without a photogrammetric education; at the same time, it produces very good results compared to other professional software (e.g., PhotoModeler Scanner).

4 TESTING THE AGISOFT PHOTOSCAN SYSTEM

Solving practical applications using the PhotoScan software has created several fundamental questions regarding the appropriate configuration of camera positions and the level of detail that can be captured, because software based on image correlations tends to smooth the sharp edges in the 3D model produced. Since PhotoScan software allows for hanging the settings of the final quality of the calculations and since no available resources have yet published the numerical characteristics under which it would be possible to estimate the accuracy of each method, the following subchapters deal with testing the PhotoScan software with respect not only to the above-mentioned tasks, but also to their time requirements and the expected degree of accuracy when using different settings.

Fig. 1 Camera positions during testing of h/y ratio.
4.1 Base to distance ratio

Software running on the principles of image correlation are based on searching for matches in image-pairs, so it is possible to talk about stereophotogrammetry, despite the fact stereophotogrammetry uses parallax measurements. In this method, the resulting accuracy of the specified point depends on the ratio between the photogrammetry baselines \( b \) (the distance between two camera positions) and the average object distance \( y \). The axis of the cameras may converge slightly, given the need to achieve as many overlaps between the neighboring images as possible. Drawing on the literature (Gal, 1967), the base to distance ratio (\( b/y \)) for stereophotogrammetry is normally in the range of 1:4 to 1:20 (0.25 to 0.05). The creators of the PhotoModeler Scanner software (www.photomodeler.com, 2012) recommend the \( b/y \) ratio in a range of 1:2 to 1:10 (0.5 to 0.1).

For applications using image correlation, it should be taken into account to maintain a sufficient similarity between the images to be able to find a correlation between adjacent images, which means shorter baselines. At the same time, however, the baselines should not be too short to achieve a higher degree of accuracy in the direction of the image axis (accuracy in depth). This theoretical accuracy can be based on digital stereophotogrammetry calculated according to the formula:

\[
m_{y} = 0.5 \cdot M_{s} \cdot P \cdot \frac{y}{b},
\]

where \( M_{s} \) is the scale factor of the image, \( P \) is the size of the 1 pixel on the CCD sensor, \( y \) is the distance from the object, and \( b \) is the baseline between the two cameras.

In order to test the software’s ability in extreme situations and find a suitable \( b/y \) ratio, a test was created with the following positions of the Nikon D5100 camera (focal length 35 mm) on the test object (Fig. 1).

Twenty-four images around the object provided a semicircle with an average object distance of 1.37 meters, while the neighboring camera positions were shifted against each other by about 12 cm.

All the images were oriented together in one project, and only specific pairs of images (16 variants) with different ratios of \( b/y \), which ranged from 0.08 to 1.78, were always selected to generate a 3D model. The quality of the 3D model generation was set to high.

The test object consisted of two printed, textured A4 papers glued on a glass sheet with 12-bit coded targets added for easily transforming all the models into a common coordinate system. One sample was a flat surface; a second sample contained a spatial unit, i.e., sharp protruding edges (Fig. 2).

The individual 3D models were then compared in the Geomagic Qualify Lite software with the reference model generated from 5 shots from a distance of approximately 0.7 m (more than twice as small Ground Sample Distance (GSD) as in the twenty-four images case). The resulting standard deviations of the flat surface and the sharp detail, depending on the \( b/y \) ratio, can be seen in Fig. 3. The deviation values are based directly on the numerical outputs from the comparison in the Geomagic Qualify Lite software and calculated only in a direction perpendicular to the surface of the test object.

The graph in Fig. 3 shows that the software does not have any problem modeling flat surfaces by the processing of significantly convergent images (in this case, the maximum convergence angle between the axes of the images were up to 125°) and large \( b/y \) ratios (up to 1.78). When the object surface is irregular, there is a loss of detail due to the very low degree of similarity between the images of such striking angles. Reducing the \( b/y \) ratio provides for an increase in the accuracy of the details, and their completeness, since the irregularity of the surfaces causes occluded areas. The improvement, however, occurs only at a \( b/y \) ratio of about 1:5 (0.2), since for shorter baselines, the model already has a strong noise. In practice it is therefore possible to recommend the use of a \( b/y \) ratio of 1:2 (0.5) to 1:5 (0.2) as a compromise between the desired accuracy of the depth and the capturing of the sharp details on the surface modeled.

In exceptional cases, of course, it is not possible to ensure an optimum \( b/y \) ratio because of obstacles in the field. Increasing the number of images at different angles can always be helpful. In principle there should always be a combination of parallel and convergent images; the algorithms in PhotoScan are already designed to choose the most suitable pair of images for the modeling of specific areas. The resulting surface is then generated from a combination of several overlapping point clouds.

Fig. 2 Test object with stone texture and coded targets.

Fig. 3 The standard deviations and their dependence on the \( b/y \) ratio.

The test object with stone texture and coded targets.

Fig. 2
4.2 Angle between the axis of the imaging and the surface

The quality of the 3D model generated depends not only on the suitability of the \(b/y\) ratio, but also on the angle between the direction of the image axis and the photographed surface. In practice, an effort is always made to shoot images with axes perpendicular to the surface of the object. Sharper angles produce deformed model surfaces (wavy, noisy) (Fig. 4).

To study the impact of the angle on the resulting surface modeled, a new test with a different configuration of the camera positions than before was created (Fig. 5). Again, the camera used was the Nikon D5100; the imaging distance was 1.2 m. To secure the same angle on the test surface for each pair of images, the camera was on a tripod with vertical scrolling for each position at two different heights (with baselines of \(b/y = 0.25\)). The test area was covered with a sandy texture (color laser printed again, this time to A3). The calculation of the 3D model was processed only for each image-pair located one above the other.

The 12 resulting 3D models were successively compared with the reference model that was created using a middle-format Phase One 645D camera with a digital back Leaf Aptus II-7. The following chart in Fig. 6 is based on differential models from the Geomagic Qualify Lite software. The standard deviations are based on numerical outputs from a comparison of the 3D models in the Geomagic software in a direction perpendicular to the surface of the test object.

From the graph it can be seen that relatively good results can be achieved by imaging the surface at an angle in the range of 35-90°. This of course again only applies to some planar or flat surfaces, because under sharp angles, and in the presence of sharp edges and details, occluded areas are created. It should also be said that the models created from images at an angle of 30° (generating a 3D model in a high quality) showed a large number of holes. The model corresponding to an angle of 22° consisted almost solely of 166,000 triangles, which is 5% compared to the model produced from images at an angle of 89°, and the area was significantly incomplete and noisy. The angles close to 20° can already be seen as a critical value at which the PhotoScan software is able to generate any surface.
QUALITY OF 3D MODELS GENERATED BY SFM TECHNOLOGY

4.3 Processing sharp edges

In addressing the previous experiments and practical applications, the PhotoScan repeatedly manifested smoothing of sharp edges in protruding and nested details. In the literature one can find many different examples where a high degree of accuracy has been demonstrated with the modeling based on the image correlation (Beliš, 2009; Junghaus, 2010), and although they were also carried out with respect to a different type of texture, there were always mostly planar or flat surfaces. Therefore, a test was created, this time focusing on finding the relationship between changes in GSD (an increasing distance), the type of texture, and increasing the smoothing of sharp edges. In addition, a change in accuracy was investigated on the influence of variable focus lengths in a range of 18-270 mm (by the same GSD). During both tests all previous findings on the appropriate position of the camera on the object were taken into account.

The test object consisted of a 820 x 820 mm glass plate, on which paper samples of six different types of natural textures were glued (Fig. 9): 1 - brick wall, 2 - stone wall, 3 – plaster facade with coarse grain size, 4 - plaster facade with fine grit, 5 - wood and 6 - standard interior plaster.

Each type of texture consisted of two basic kinds of detail - protruding and nested; the angle of the sharp edge of every detail accounted for 90°. In addition to these two basic samples for each texture, a faded detail, a 90° sharp edge were created, the base of which steadily dwindled from 25 mm to zero. Based on this detail the critical value at which it ceases to exist could later be determined, and no detail in the 3D model is visually identifiable. The test object was attached to a wall with an irregular random texture (wood paneling) for minimizing complications in the orientation of distant images.

The positions of the camera were determined by keeping the ratio \( b/y = 0.25 \). For each object-distance \( y \), 5 images with a Nikon D5100 camera with the support of a photographic tripod were created. On the test object and the wall behind coded targets were glued for processing with the PhotoModeler Scanner 2012 software; also, RAD targets were applied on the wall. On the object itself 12-bit coded targets used in PhotoScan were applied. The scale was determined by measuring the lengths between the RAD-targets in different directions, i.e. vertically and horizontally; for the definition of the scale a vertical pair of points closest to the object (to avoid the impact of the excess of the measuring tape) was selected. The PhotoModeler Scanner 2012 was only used for defining the coordinate system (scaling, rotation, translation); by processing convergent images from the 33 Mpx PhaseOne 645D camera, the overall mean error on the object points of 0.05 mm was achieved. These 12 points subsequently served as fitting points for transforming the models from the PhotoScan into a common coordinate system. The models were generated in the PhotoScan in an ultra high quality in order to achieve as much detail as possible, since the ultra high setting includes each pixel of the image in the calculations.
As a reference 3D model the output from the Comet LED structure-light scanner was used, with a declared accuracy of 0.05 mm. The comparison was done using Geomagic Studio software by creating sections with regular spacing across the different samples (Fig. 10).

In each section the peak height of the sharp edge for each texture sample was measured, i.e., the minimum value for detail “A” and the maximum for detail “B”. For each type of detail (A and B) 10 values were obtained, which were used for calculating the standard deviation from the reference model based on the formula:

\[ \sigma = \sqrt{\frac{\sum \epsilon^2}{n}}, \]  

where \( \epsilon \) is the difference between the y-coordinate of the sharp edge on the test model and the reference model, \( n = 10 \). The y axis is directed perpendicularly on the surface of the test object.

In detail C the standard deviation in the cuts was not calculated, but the length of the detail has been measured, from its wide baseline towards the point where it was not possible to visually identify it on the model. For each texture type 3 such lengths were obtained.

The height \( dh \) of the detail, which had not been modeled, was then calculated from the similarity of the triangles (Fig. 11). The resulting value of the faded detail is based on the simple arithmetic average of the three values.

### 4.3.1 Case with the same focal length and variable GSD

Taking the images was performed with a Nikon D5100 DSLR with the focus length set to 35 mm. The object distance \( y \) was progressively prolonged with regular spacing (2, 4, 6, 8, 10, 12, and 14 m), which changed the GSD (from 0.27 to 1.92 mm).

From the results it is possible to pick out not only the impact of the GSD on the size of the deviation, but also on the quality and type of the texture, which varies with the increasing distance. The flat textures (4 and 6) showed the highest noise level, which corresponded with the largest irregularity in the change in the standard deviation, depending on the GSD. The other types of textures (1, 2, 3, and 5) showed a standard deviation of an almost linear shape with the size of the GSD, but the results gradually deteriorated nonlinearly due to the deteriorating quality of the textures (Fig. 12).
At an object distance of 14 m, the worst results were obtained for texture No. 6 and brick texture No. 1. The reasons for the poor quality of the results for the brick texture can mainly be found in the inhomogeneity caused by the dark places between the bricks. Conversely, the best results were achieved with grainy texture Nos. 2 and 3, which not only had a soft, but also a distinct granularity from the larger object distances. Taking into account the average standard deviations corresponding to all types of the textures except for type No. 6, it is possible to approximate the course of the power function of the form $Y = \alpha \cdot X^\beta$, where $Y$ is the amount of the deviation, and $X$ corresponds to the value of the GSD. With the approximate coefficients $A = 2.5$ and $B = 1.3$, it is possible to render the curve as approximating empirically obtained values (Fig. 13 - curve “Equation 1”).

It is noteworthy that the a priori depth accuracy based on stereophotogrammetry (relation (1)) coincides with the empirical values of the standard deviation of the smoothness of the detail in the initial values of the GSD, in which the texture was of the highest and clearest quality. The shifting of the standard deviations away from the linear a priori accuracy may be related to the already-mentioned effects of the gradual degradation of the texture; the a priori relationship (1) therefore applies only in the case of the quality of the irregular random texture.

In this part of the test an empirical relationship for the level of the smoothness of the sharp edges was derived:

$$\sigma_{\text{detail}} = 2.5 \cdot \text{GSD}^{1.3}. \quad (2)$$

Testing the critical value of the faded detail C produced the results seen in Fig. 14. In the graph in this case the size of the detail that had not already been modeled is shown.

If we take an average from all the newly obtained values and plot it on a graph, it is possible to approximate its course with a power function, but with the parameters $A = 4.1$ and $B = 1.3$.

On the graph in Fig. 15, it can be observed that the detail is lost before the estimated standard deviation of the sharp edges reaches the size of the edge. A sharp edge is at some level so low that the software considers it no longer to be a deformation of the surface, but only as a textural character of a flat surface. On a practical level, it is therefore to be expected that once the standard deviation for the sharp edges reaches about 60% of its overall size, the detail in the PhotoScan software will be unmodeled at the highest quality setting of the mesh generation.

The critical height of the sharp edge can then be based on test results computed with the empirical formula:

$$\sigma_{\text{crit.det}} = 4.1 \cdot \text{GSD}^{1.3}. \quad (4)$$

The constant $A = 4.1$ may of course vary slightly with respect to the actual type of texture; it is also the same by calculating the estimated standard deviation ($A = 2.5$).

At the end of this subchapter it may be said that the worst results of flat texture No. 6 were achieved by the standard deviation of the sharp edges up to 4 pixels. The faded detail with texture types 4 and 6 reached a maximum size of 6-7 pixels.
4.3.2 Analysis of flat regions

As the test samples also consisted of flat regions, the individual textures were not analyzed only at the sharp edges. The standard deviations of the comparison with the reference model from the Comet scanner were obtained with the Geomagic Studio software. In this section only the progress of the average standard deviations of all kinds of the textures is shown as a preview. Fig. 16 shows the relationship between the variations in the sharp edges and flat areas of the building; the difference is almost 5 times greater than the average values in the flat areas.

The accuracy of the evaluation with the PhotoScan software thus achieves good results, especially in predominantly macroscopically flat areas; the average accuracy of the depth for the appropriate types of textures (1, 2, 3 and 5) is around 1/3 of the GSD (only 0.3 pixels). With texture No. 6, the worst results were always achieved, the standard deviation of the planar areas ranged up to 2 pixels in depth.

4.3.2.1 The correlation between the standard deviation of the sharp edge and the type of texture

For finding the correlation between the type of texture applied and any other parameter, it is first necessary to quantify the particular types of texture and describe them according to certain characteristics; in other words, a textural classification must be done. Since the quality of the generated 3D model depends primarily on the homogeneity of the texture, it seems appropriate to classify the texture by the so-called energy or Angular Second Moment (ASM). Energy expresses the homogeneity of the texture; the texture is homogeneous (flatter) if the ASM reaches a higher value. The ASM calculation is based on the formula:

$$ASM = \sum_{a,b} P_{\Phi,d}(a,b).$$ (5)

The relationship is based on the so-called matrices of repeated appearance - the appearance of a gray level configuration can be described with a matrix of relative frequencies, describing how often two pixels with gray values $a, b$ appear in a window at a distance $d$ in the direction $\Phi$ (Fíčnik, 2010).

The ASM values can be obtained, for example, by using ImageJ software with additional plugins. ImageJ is a freely available software solution using the Java programming language for processing and image analysis. The program has a wide supportive community, whose members produce various useful plugins that are specialized in different disciplines, i.e., from medicine, biology and geology through the environment. The ASM values were obtained for the purpose of this work using the GLCM texture plugin, which also allows for calculating the contrast, correlation, inverse differential moment and entropy.

<table>
<thead>
<tr>
<th>Texture / ASM</th>
<th>0.27</th>
<th>0.55</th>
<th>0.82</th>
<th>1.09</th>
<th>1.37</th>
<th>1.64</th>
<th>1.92</th>
<th>ASM corre-</th>
<th>GSD corre-</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSD (mm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 ASM</td>
<td>0.00028</td>
<td>0.00027</td>
<td>0.00026</td>
<td>0.00034</td>
<td>0.00046</td>
<td>0.00060</td>
<td>0.00064</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>1 σ (mm)</td>
<td>0.58</td>
<td>1.29</td>
<td>2.28</td>
<td>3.30</td>
<td>4.32</td>
<td>5.57</td>
<td>6.86</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>2 ASM</td>
<td>0.00042</td>
<td>0.00048</td>
<td>0.00063</td>
<td>0.00093</td>
<td>0.00125</td>
<td>0.00200</td>
<td>0.00200</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td>2 σ (mm)</td>
<td>0.51</td>
<td>0.90</td>
<td>1.51</td>
<td>2.77</td>
<td>3.48</td>
<td>4.33</td>
<td>4.89</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td>3 ASM</td>
<td>0.00020</td>
<td>0.00021</td>
<td>0.00028</td>
<td>0.00037</td>
<td>0.00054</td>
<td>0.00076</td>
<td>0.00083</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td>3 σ (mm)</td>
<td>0.44</td>
<td>1.12</td>
<td>1.87</td>
<td>2.41</td>
<td>3.37</td>
<td>4.01</td>
<td>4.79</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td>4 ASM</td>
<td>0.00200</td>
<td>0.00300</td>
<td>0.00300</td>
<td>0.00400</td>
<td>0.00550</td>
<td>0.00700</td>
<td>0.00700</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>4 σ (mm)</td>
<td>0.43</td>
<td>1.07</td>
<td>1.62</td>
<td>2.47</td>
<td>3.64</td>
<td>4.84</td>
<td>5.86</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>5 ASM</td>
<td>0.00150</td>
<td>0.00088</td>
<td>0.00150</td>
<td>0.00225</td>
<td>0.00250</td>
<td>0.00325</td>
<td>0.00250</td>
<td>0.84</td>
<td>1.00</td>
</tr>
<tr>
<td>5 σ (mm)</td>
<td>0.52</td>
<td>1.19</td>
<td>1.86</td>
<td>2.93</td>
<td>4.11</td>
<td>4.75</td>
<td>6.00</td>
<td>0.84</td>
<td>1.00</td>
</tr>
<tr>
<td>6 ASM</td>
<td>0.00450</td>
<td>0.00800</td>
<td>0.01125</td>
<td>0.01550</td>
<td>0.01725</td>
<td>0.01650</td>
<td>0.02050</td>
<td>0.92</td>
<td>0.98</td>
</tr>
<tr>
<td>6 σ (mm)</td>
<td>0.58</td>
<td>1.26</td>
<td>3.17</td>
<td>4.32</td>
<td>4.11</td>
<td>7.03</td>
<td>7.80</td>
<td>0.92</td>
<td>0.98</td>
</tr>
</tbody>
</table>
Since the ASM gains different values for the selected type of texture depending on the direction chosen, for each type of texture, 4 ASM values were obtained (for the predefined angles of 0°, 90°, 180° and 270°). The mean values of the ASM for each type of different texture and GSD values are listed in Tab. 1.

From Tab. 1 it is possible to see that the highest values of the ASM were achieved, especially for the flat textures 4 and 6; both of these textures showed the highest levels of noise during processing in the PhotoScan. The lowest values of the ASM were reached for the contrasting textures 1, 2 and 3 with fine and coarser grain sizes; these 3 textures also achieved the smallest values of the standard deviations of the sharp edges. In Fig. 12, texture No. 1 is indeed an exception; however, as already mentioned, it is related to the large and irregular changes in texture (wide black gaps between the bricks) which, when calculating the step for the ASM, was set to 1 pixel, and of course, could not be taken into account. In Tab. 1, the correlation coefficients can also be found between the change in the standard deviations of the sharp edges and the changing ASM (ASM correl. column) and between the correlation coefficients in the standard deviations of the sharp edges and the changing GSD (GSD correl. column). The correlation coefficient between the two data sets was calculated based on the following formula, whose values can range from 0 to ±1:

\[
\text{Correl}(U, V) = \frac{\sum(u-u)(v-v)}{\sqrt{\sum(u-u)^2 \sum(v-v)^2}}.
\]  (6)

The correlation coefficients in Tab. 1 show that wooden texture # 5 has a lower level of correlation between the ASM and the variation of the deviation in the sharp edges than any other, which is by the nature of the texture perfectly logical; pitch dark bands generated in the two perpendicular directions of the ASM calculation twofold show different values. The result would of course be influenced by changing the direction of calculating the ASM by 45°; the selected GLCM texture plugin, however, in its current version (December 2012) did not allow this. Based on the results a significant correlation is undeniable between the change in the standard deviations of the sharp edges and the changes in the ASM and GSD.

4.3.2.2 Accuracy of generating a 3D model in various settings

To illustrate the impact of changing the quality of the generating mesh, images from the object distance of 2 m were processed not only with ultra high settings, but also with lowest, low, medium and high settings.

The average standard deviations for all the types of textures can be seen in Fig. 17.

From the ratio between the neighboring quality settings, it is possible to see that the ultra high setting is about 1.5 times better than the high one; in the other settings, it is an approximately 2.5-fold increase in accuracy. In relation to the medium, the ultra high is therefore about 4 times more accurate, versus the low, around 10 times; and the lowest, 25 times more accurate. So it is only up to the user what level of precision of processing he requires, since it is necessary to also consider the time costs of the calculations.

The different degrees of accuracy and speed of processing result from the fact that the ultra high setting considers each pixel (based on the original data), while at the high setting, the resolution is reduced by down sampling the images so that the 3D position is given only for every 2 pixels, in the medium for every 4, etc.

4.3.2.3 Time costs of generating the 3D model in various settings

The effect of changing the quality settings of the mesh generation on the processing time requirements can be seen in Fig. 18. From the graph it is possible to see a considerable discrepancy between the time consumption and an increase in accuracy, particularly regarding the use of the high and ultra high quality settings. The increase in the accuracy in this case was only about 1.5-fold, but its time consumption is 6-fold. In the other settings, there were approximately 3 - 4-fold time increases in the neighboring lower quality.

The above results indicate that for most practical applications, it is sufficient to use the medium and high settings if images with a sufficiently high resolution are entering the processing. A 1.5-fold
increase in accuracy is not justified by a 6-fold longer time needed for processing at the ultra high setting, so it makes sense to apply this setting only in special cases.

4.3.3 A case with a variable focal length with the same GSD

For the purposes of digital photogrammetry, focal lengths with a length of about 30-50 mm are most commonly used, whereas with very short focus lengths, the lens distortion reach very high values and, vice versa, with very long focus lengths, the bundle of determining rays becomes unstable. In the field, however, a situation may arise where, because of the complicated availability of an object or confined spaces, it is required to use a wider range of focal lengths, from so-called fisheyes to telephoto lens. In the next test a full range of Tamron 18-270 mm telephoto lenses was used, while the distance from the test object changed in accordance with maintaining approximately the same GSD (that object always occupied approximately the same area in the visual field) of the Nikon D5100 digital SLR camera. The \( b/y \) ratio during the imaging was always 1:4 (0.25) with the support of a photographic tripod. With the gradual increase in focal length, the aperture of the lens always reduces and decreases the quality of the textural information (contrast, intensity). The image quality can be expressed through the Modulation Transfer Function (MTF), which is defined as the ratio of the image contrast in the image plane due to the contrast of the object in the object plane of a sinusoidal test pattern. At the same time it should be noted that despite the use of a tripod, imaging with very long focal lengths and a lack of light in the optical system can result in a motion blur of the image, since the whole system is more sensitive to subtle vibrations. This effect was evident even when taking shots using a Tamron 18-270 mm, especially at focal lengths approaching 200 mm and longer. Altogether, 8 sets of data were generated for the focal lengths 18, 35, 55, 70, 100, 170, 250 and 270 mm.

An assessment is again shown for the planar areas and sharp edges, and it is the average of all types of the textures. From the graph a multifold difference in the accuracy of the modeling of the sharp edges against the predominantly planar areas can again be read. It also clearly shows that prolonging the focal length increases the level of uncertainty to generate a 3D model, especially in terms of sharp edges and details. The good thing is that despite the recommendation of the PhotoScan developers to use regular wide angle lenses and not to go to extremes such as using telephoto lenses, quite good results can be achieved even in such exceptional cases, since the standard deviations at very long focal lengths over 200 mm also ranged in value to 4.5 pixels.

4.4 Summary of findings from the individual trials

From the knowledge gained during the detailed testing and solving of the practical applications, the following conclusions and recommendations for field work and processing using the Agisoft PhotoScan software can be drawn:

– The \( b/y \) ratio should be in a range of 1:2 to 1:5 (0.5 to 0.2) with respect to the irregularity of the photographed surface. A smaller \( b/y \) ratio produces greater uncertainty in the depth (noise); in contrast, a greater ratio causes a loss of sharp detail. A \( b/y \)-ratio of 1:4 (0.25) can be regarded as an appropriate compromise.

– The axis of the camera to the imaging surface should be as perpendicular as possible. PhotoScan is able to produce relatively high quality results in a range of 35° to 140°. However, this only applies to predominantly planar surfaces; in the case of a rugged surface, sharp details can be lost. From about 20°, PhotoScan is not able to generate almost any 3D model.

– A 3D model is noisier if the surface has a flat texture. The suitability of textures can be determined by calculating the ASM. The smaller the value of the ASM, the more the texture is suitable for optical scanning based on natural textures. Shaded areas (joints, holes, cracks) in the software are usually evaluated only as a textural characteristic and are smoothed. It is therefore necessary to consider the provision of suitable lighting conditions with additional lighting; a flash is not recommended due to the change in the direction of the shadows and the strong reflections on glossy surfaces.

– PhotoScan generally achieves good results in modeling planar surfaces (0.5-1 pixel in depth). The level of the erasure of sharp edges depends both on the GSD and also on the quality settings to
generate a 3D model. In the case of the quality of sharp textures, the a priori accuracy of the depth can be calculated based on the stereophotogrammetry relationship (1). With an increasing distance, the quality of the textures is reduced, and the deviations from the sharp edges in the 3D model are increased, which can be estimated using equation (2). The detail, however, ceases as soon as the estimated deviation reaches about 60% of its size.

– During imaging it is appropriate to use focal lengths of about 30-50 mm, depending on the sensor’s size. Its lengthening or shortening produces extra noise in the 3D model. But in extreme cases it can easily be usable, as well as with very long focal lengths of 200 mm or more. Accuracy at a depth of a 4.5 pixel on sharp edges and a 0.5 pixel in the flat areas can be achieved (in case of compliance, the b/y ratio is 1:4).

– To generate a 3D model it is sufficient in most applications to use settings of high and medium, as the ultra high setting delivers only a 1.5-fold improvement in accuracy compared to the high, but a 6-fold extension of the computational time. The time consumption mainly depends on the number and resolution of the images to process, but increasing the number of images by taking shots from short distances and handling them in the high setting can produce greater savings in time, than to use fewer shots from a longer distance and use the ultra high setting, all while achieving the same degree of accuracy.

– Selecting the most appropriate camera position during imaging in terms of efficiency is of course necessary, but trying to excessively minimize the number of images does not make practical sense. The developers of the PhotoScan software recommend that more shots are always better than fewer. The hardware is becoming more and more powerful with a corresponding decrease in time costs. The selection of appropriate images can be done at any time in an office, and the unnecessary images can be excluded from the processing. Their density is of course suitable for choosing according to the GSD and b/y ratio according to the actual situation.

– Never forget to define the scale. If no option is available for a geodetic survey of the control points, it is sufficient to measure the length between clearly identifiable points by a tape. Two to three control lengths should be a part of every survey.

5 PRACTICAL APPLICATION

The accuracy of the optical scanning can also be checked on an actual heritage object - in this case it was a replica of Michelangelo’s statue of David (its head was approximately 0.2 m high). The statue was scanned by a KONICA MINOLTA VI-91 3D image-based scanner, the declared accuracy of which is 0.05 mm in optimal conditions for TELE lenses. The shots were taken with different cameras - a Nikon D5100 digital SLR (16 Mpx) and a Pentax D 645 middle format digital camera (40 Mpx). About 50 shots of the head of this statue were processed.

The 3D model was generated in medium and high qualities of mesh generation. In Fig. 22 it is possible to see the models from both cameras used by setting the quality on high.

The standard deviations from the reference model in both cases reached a maximum of 0.1 mm, but mostly 0.05 mm, which is close to the declared accuracy of the KONICA scanner. With both cameras a GSD of 0.08 mm for an average object distance of 0.6 m was achieved (Nikon D5100 with f = 35 mm and Pentax 645D with f = 45 mm). The greatest values of the deviations were reached in areas with streaks (in the hair), and with respect to the GSD, represented about 1.3 pixels, which correspond to the results obtained during the testing in the previous chapters.

6 CONCLUSION

The automation of processing image data into a 3D model is currently a very topical issue, as it is widely used in a large variety of applications, from ecology through geology, geodesy and architecture to civil engineering (Pavelka and Řezníček, 2011). In the documentation of a cultural heritage, the main focus of these methods is to provide a safe, precise and non-contact way of obtaining 3D data about the objects (Musílek et al., 2001). The financial costs of the technologies based on computer vision are often only a fraction of the costs of specialized optical or laser scanners and achieve a comparable degree of accuracy. Time consumption is thus dependent only on the power of the computer, and the price / performance ratio is lower from year to year.
These results are part of a dissertation on the conservation and restoration of a cultural heritage using digital photogrammetry methods. The conclusions drawn from testing different processing settings for different camera configurations can help the users of these systems to improve the process of modeling 3D objects from optical records. The study also presents the time impact of different quality settings of 3D model generation with the PhotoScan software with respect to the accuracy of the output, which enables the user to find an effective processing configuration for any specific application.

REFERENCES

Belíš, M. - Nemčík, M. (2009) Analýza merania plôch fotogrametrickým skenerom (Analysis of surface measurement with a photogrammetrical scanner). Študentská vedecká konferencia (Student scientific conference), Slovak University of Technology in Bratislave, Faculty of Civil Engineering.


Doneus, M. - Verhoeven, G. - Fera, M. - Briese, CH. - Kucera, M. - Neubauer, W. (2011) From deposit to point cloud: a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. GEOINFORMATICS (Faculty of Civil Engineering, Czech Technical University in Prague), 6., pp. 81-88, ISBN 978-80-010-4885-6


Gál, P. et al. (1967) Fotogrametrická dokumentácia stavebných pamiatok (Photogrammetrical documentation of architectural monuments). Záverečná správa úlohy 02.12 f (Final report from task No. 02.12 f.). Scientific Laboratory of Photogrammetry, Faculty of Civil Engineering, Slovak University of Technology in Bratislava.


