GPGPU accelerated large scale 3D object reconstruction application for fixed and rotary wing unmanned aerial vehicles

András Molnár, Dániel Stojcsics

***Abstract*—**GPGPU (general-purpose computing on graphics processing unit) cards compared to a standard CPU (central processor unit) has many times more cores. The advantage of this technology is the possibility of using multiple GPGPU cards in a single PC making the parallelization opportunity to a next level. Large scale 3D object reconstructions usually made by human piloted airplanes with massive sensor arrays (high resolution RGB camera or camera array, LiDAR or laser scanner). With the advance of small size fixed or rotary wing unmanned aerial vehicles (UAVs) and a relatively cheap camera this procedure can be made with much less effort. No need for using public airports, airspace and human piloted expensive aircrafts. UAV operator can make the measurements within a few hundred meters from the desired object. Using GPGPU accelerated photogrammetry methods large scale 3D object reconstruction can be made from zero to 3cm/pixel resolution 3D reconstructed and textured object within hours. The paper presents the procedure step by step, showing the technology critical parts of the method.

***Keywords*—**GPGPU, photogrammetry, 3D reconstruction, multicopter, drone, UAV, APM, Sensefly

I. The Basis of Taking Pictures and Its Technical Background

The basics of 3D reconstruction are the same as the procedures of making orthophotos [1][2][3]. Pictures of proper resolution that overlap must be made of the area or the object to be surveyed. Afterwards, these photos must be fitted. Due to the huge amount of data this process can only take place efficiently if automated. Of course, automation does not mean that the 3D model can be created without the interference of the operator. After the preliminary check of some parameters human interference is required at certain stages of processing [4][5][6].

Matching the mosaic parts (images) occurs by means of the point pairs typical of the picture. To find these point pairs and to have enough point pairs to fit, at least 40-50% overlapping per page is required [7][8].

A. Molnar is with the Faculty of Economy, J. Selye University 3322 Bratislavská cesta, 945 01, Komárno, SLOVAKIA (e-mail: molnara@selyeuni.sk).

D. Stojcsics is with the Faculty of Economy, J. Selye University 3322 Bratislavská cesta, 945 01, Komárno, SLOVAKIA (e-mail: stojcsicsd@selyeuni.sk).

To ensure proper overlapping an aviation programme must be determined so that the pictures taken should meet the basic requirements of further processing by taking the special features of the camera and manouvering into account. The basic pictures of an area are taken in accordance with Fig. 1. It is more practical to make pictures of objects as shown by Fig. 2.

After the flight the interface software looks for typical point pairs on the images (they can even amount to several thousand per picture), which are matched afterwards. The widely used processes (SIFT or SURF) [9] are then optimised on modern video cards (GPGPU) so these devices help accelerate the process which is then ten times faster than with a traditional processor [10][11].

After the software finished matching the point pairs, a colourful point cloud is created. There is still an opportunity for deleting faulty or not appropriate images and point pairs. Afterwards, a properly filtered 3D surface is created from the point cloud on which a low resolution texture is made by montaging the original pictures. Geo referencing is possible on both it and the original pictures on the basis of the typical objects on the site or the surveyed points.

This procedure is not sensitive to slewing the picture and the camera's position on a stabilised platform is not a criterion, either. That is, the camera should not exactly look downwards vertically or right to the centre of the object to be photographed at all times.

Reconstructing a large scale object e.g. a quarry, dam or a reservoir needs mission planning solution. Commercial and open source softwares are available. Market leading Sensfly company has the 'eMotion 2' (Fig. 4.) mission planner which grants easy area coverage with polygon based area selection while open source softwares has also similar technologies e.g. 'APM Mission Planner' (Fig. 5.). Both capable of resolution (typically 3-10cm / pixel) based flight path design which is desired for large scale aerial 3D reconstruction.

II. 3D Reconstruction of Surface Topography

Elevation models (DTM) can be made in addition to pictures whose objective is the creation of an orthophoto of the area by processing the raw images. Classic methods used LiDAR for this purpose [12].

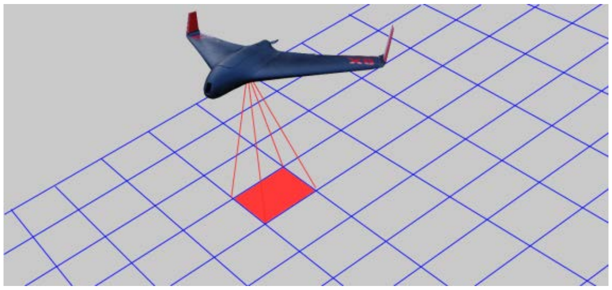

Fig. 1: The theoretic draft of taking pictures of the surface

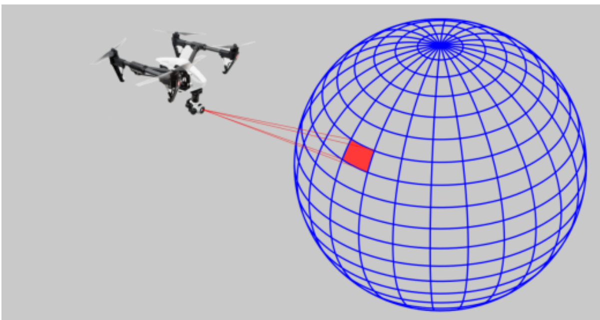

Fig. 2: The theoretic draft of taking pictures of objects

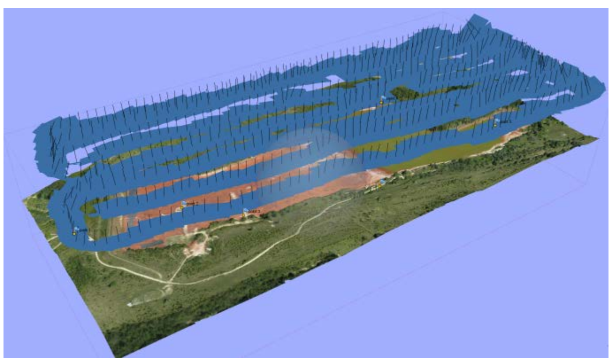

Fig. 3: The spatial positions and orientations of the images of the surveyed area (Agisoft Photoscan)

In general, the size of the examined area requires the use of such an aviation vehicle that can autonomously on a predetermined route while taking pictures. It is necessary as the borders cannot be flown over safely by the operator, which, by all means, calls for telemetric data that help check the exact position of the flying object on the map.

In general, although telemetric information is inevitable to take pictures as the immediate interference in aviation can be ensured when necessary, flights are more practical to be made autonomously.

The advantage of autonomous flight is that flying over the area to be surveyed can be planned before. Knowing the type of the camera used such a route can be designated which ensures the minimal overlapping necessary between the recorded images. Usually aircrafts with rigid wings can effectively be applied for taking pictures of the surface regarding their consumption, load capacity and speed.

Taking off and landing are minor disadvantages that require a plain or an area free from other flying objects and trees.

Figure 4: Autonomous mission plan in “Sensefly eMotion 2”

Figure 5: Autonomous mission plan in “APM Misson Planner”

Fig. 6: The orthophoto of the red mud reservoir

Fig. 7: Part of the red mud reservoir by the free Google Earth programme

Fig. 8: Part of the red mud reservoir on Figure 5 recorded by UAV

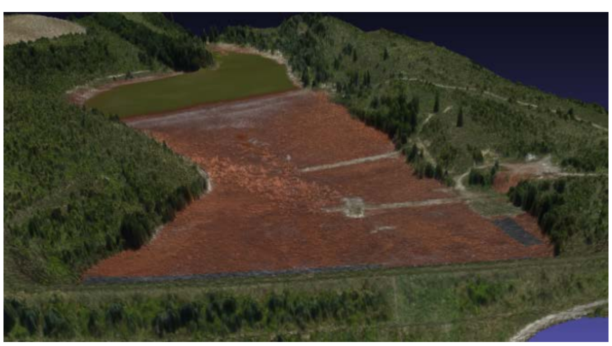

Fig. 9: The 3D model of the red mud reservoir (MeshLab)

During the experiment a carrier with wings all over was used whose total take-off weight did not exceed 3 kg. This could also be launched from the hand but also launching by a rubber cord was also suitable. Due to its light weight, landing was also easy.

The first step of surveying the area was planning the flight route. The route should be planned in a way so that at least 40% overlapping between the images is ensured in parallel flights. Further criteria include setting the altitude in a way so that the horizontal overlapping of the images should also be at least 40% depending on the speed of the aircraft and the optics of the camera used. The camera in the experiments was a Canon Powershot A2600 where a special firmware made automated shooting possible.

The travelling speed of the carrier was 15 m/s and the altitude exceeded 150 m above the launch site. During the fight 853 pictures were made amounting to 3.4 GB data.

The designated area was approximately 1600 m long and 600 m wide. Fig. 3 illustrates the real flight route and the exact position of taking certain pictures. As each image has overlapping with 4 or at least 3 neighbouring images in general, the spatial position and orientation of recording the images can be determined.

Consequently, the position of the relevant image points in the picture (those that make up pairs found in at least two pictures) can also be determined. In the first steps of processing an orthophoto is made. After georeferencing, it can be exported to, among others, *.kmz format managed by the Google Earth programme.

Due to georeferencing, Google Earth can read the exact geographical position of the area. The created photo can serve as the basis for measurements and calculations (distances, areas etc.).

The advantage of the photography over the satellite picture of Google Earth is its updated content and resolution. The resolution of the orthophoto made with the parameters above was 3.0 pixel/cm (Fig. 6).

Fig. 7 and 8 illustrates the resolution of the orthophoto made up by mosaics. Figure 6 shows part of red mud by Google Earth programme while Fig. 8 presents part of the same area in the same size. It can be seen that the amount of details of the photo made by UAVs significantly exceed those of the free satellite images.

In addition to its high resolution, a further advantage of the UAV picture is its updatedness. On the area recorded regularly with so many details, such changes can be noticed that are of vital importance for specialists such as leakage of the reservoir. The high resolution picture is also advantageous to create the texture that overlaps the 3D model. The more detailed this texture is, the more lifelike the results are.

The point clouds gained while processing the images are linked by triangles in the programme. These triangles make up the surface which is the 3D model of the surveyed area, of course, the number of polygons or any other ways of detailing the area can also be displayed. The 3D model can also be observed from a discretionary point of view after being rotated. The model only includes surface, i.e. no thickness of the polygons covering the surface is given. In another step of processing a texture can be fitted to the 3D surface. The programme creates the texture with the help of the orthophotos prepared in the previous steps.

Basically two models of fitting the texture can be selected. A simpler but more practical model in the case of surface models is when the texture, as a rubber film, is stretched on the surface model. This process results in the fact that the vertical parts (if any on the surveyed area) are coloured by the colour of the pixels at the edges. This method cannot be applied to buildings or vertical as well as concave surfaces.

While exporting the 3D relief map the 3D surface model itself is created and also an image file that contains one or more textures (depending on the size and the parameters of the model). All these files are necessary to create the image of Fig. 9 which can be ensured by several programmes. For example, MeshLab programme can be downloaded and used free of charge.

III. 3D Reconstruction of Objects

Due to the relatively exact data gained by this method the spatial reconstruction of some objects is possible. In this case it is more sensible to use rotorcrafts or other aerial vehicles that can float instead of aircrafts with rigid wings. In the experiments the multirotor flying objects proved to be suitable. With the help of the camera fixed on board the selected object can freely be photographed from all directions. In this case the common parts, i.e. overlapping between the pictures are a must. A further condition is that the object must be photographed from different points and angles. It is extremely important if there are certain uncovered parts of the object from a certain point such as part of a building under the eaves or parts of the roof that uncover the walls of the building or its other parts. In such cases photographs must be taken by a camera facing upwards while flying low by multicopter.

The success of the procedure greatly depends on the richness of the details of the certain pictures. To this end, high resolution cameras of at least 10-12 megpixel, free from distortion must be used.

Fig. 10: The aerial photography of a chapel from which 3D reconstruction was made

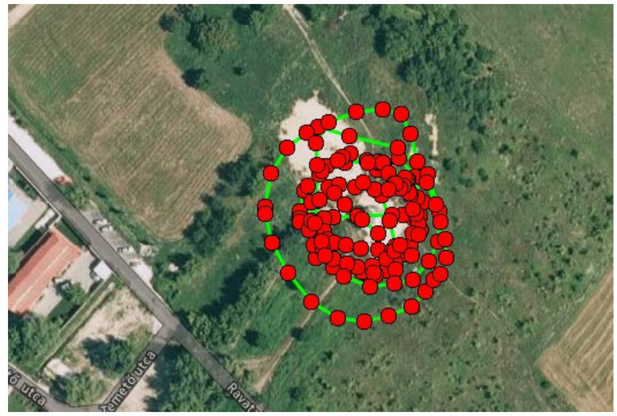

Fig. 11: The position of the basic picture of the chapel (Pix4Dmapper)

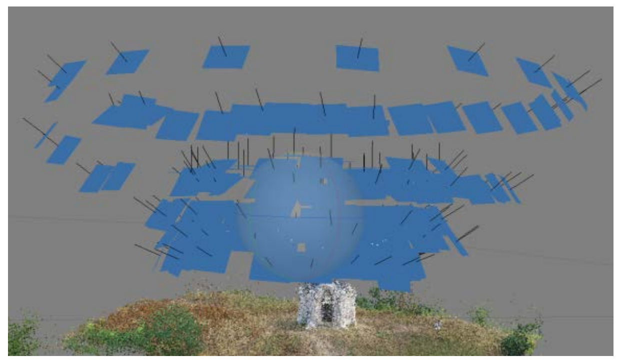

Fig. 12: The postion and orientation of the raw images of the chapel (Agisoft Photoscan)

Fig. 13: Picture of the 3D model of the chapel – Meshlab

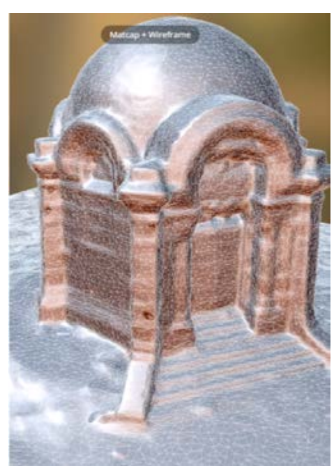

Fig. 14: 3D reconstructed object (”chapel”) visible object faces - Sketchfab

Even with that resolution it is necessary to take as up-close pictures of the object as possible. Again, based on our experience a distance of 30 m or closer is also necessary while taking the pictures to ensure enough details for further reconstruction.

If the objective is the visibility of the surroundings of the recorded object on the 3D model, in addition to the up-close pictures also others taken from a distance is possible, where, of course, the object to be photographed must also be discernible.

Taking pictures can occur by autonomous or remote controlled flight. In the case of the autonomous one the preciseness and reliability of the robot system are essential. Following the route with an absolute ±3 meter at least is a must due to the proximity of the object. In the case of GPS-based systems this preciseness can hardly be ensured especially when flying low when the object to be photographed uncovers a significant part of the sky (1/4-1/3) in GPS reception.

Fortunately, pictures can relatively be taken easily by remote control so the risks described above can be minimised. At the same time, while flying it must also be considered that during the possible flight around the object the operator cannot see the multicopter at all times. In that case controlling and driving the multicopter is based on the live broadcast but pasting the object vertically is also an option. In the latter case the operator can make the series of pictures vertically and repeat the procedure from a new position by turning round the object till he manages to go around the entire object by multicopter.

Fig. 10 presents the aerophoto of a chapel. It can be seen that the chapel stands in the open air, i.e. there is nothing that would prevent flying or taking pictures such as other objects or trees so the object is ideal for carrying out a 3D reconstruction experiment. The chapel is rich enough in motifs but does not have tapering parts whose reconstruction is complicated. A multicopter equipped with a 12 Mpixel camera free from distortion was selected for the experiment. Images were made automatically every 5 second while the picture of the camera was available from the land real time during the whole flight. While taking these pictures the multicopter was remote controlled.

The system did not cater for overlapping between the pictures so the operator of the multicopter carried it out while directing the machine. During the flight 150 pictures were made amounting to more than 700 MB data.

Two programmes were tested for 3D reconstruction. The demo version of ’Pix4Dmapper’ can be downloaded after registration and can be used without almost any restrictions.

The only barrier is that the work created cannot be saved. Fig. 11 presents the position of the images read by Pix4Dmapper on a map. It can be seen that the movement of the camera was not regular and also the spatial density of the images also differs to a great extent. All this can be due to manual control but it is not relevant, either, for 3D reconstruction.

It can also be seen that a lot of pictures were taken near the chapel but series were made from a distance, as well, to make its surroundings also visible.

The ’Agisoft Photoscan’ programme does not only provide information about the position of the camera but also its spatial orientation. The optic axis of the camera are illustrated by the thin sticks in the picture (Fig. 12). Of course, these data were also calculated by the Pix4Dmapper programme but the display is slightly different. Fig. 11 presents the sparse point cloud primarily calculated in addition to providing information about the place of taking the pictures.

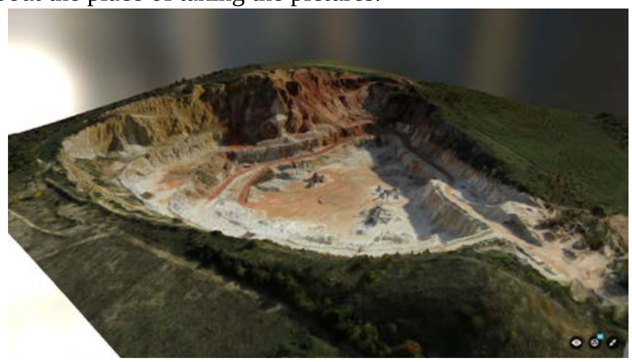

Fig. 15: Reconstructed quarry textured model

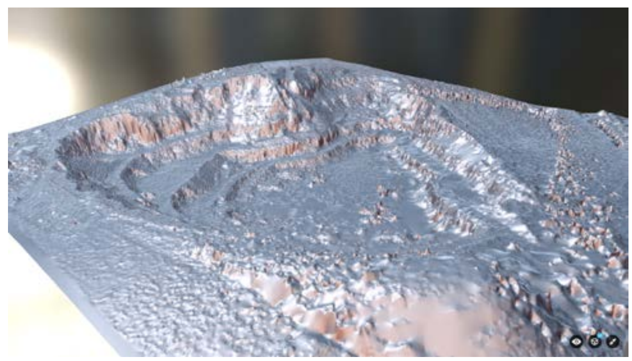

Fig. 16: Reconstructed quarry shaded model, without texture

Fig. 17: FDM printed 3D reconstructed objects

As the resolution of the sources is quite high and have a lot of details, the weak point cloud act as if the 3D reconstruction has already been made. However, it lacks the surface model consisting of the polygons that link the points and the texture stretched on it.

Processing the images took several hours for both programmes. Approximately, the duration of processing is 20 hours on a computer with an 8-core i7 2.4 GHz, 64 bit processor and 8 GB memory. There was no other application running on the computer while processing the same data on a similar computer with 32 GB memory and nVidia Titan X GPGPU card took 2 hours. The duration of processing greatly depends on the individual settings of software. The more details the reconstruction has, the longer time processing takes.

Processing takes place in several steps for the programmes. Although batched processing can be set in both cases, it is worth stopping after each stage. It is possible to correct the point clouds manually. It is also worth keeping only the area to be reconstructed and delete the points farther away together with the points that stand apart (far from the surface). These can make the next step faster and also can improve the quality of reconstruction.

After processing the images the 3D surface model is created with one or more textures. Together they make up the reconstructed object and its surroundings. The 3D model can be seen via several other free programmes (Fig. 13 and Fig. 14) such as 'MeshLab' in addition to the modelling software.

IV. Conclusion

One of the biggest advantages of the 3D models based on aerophotography is that they do not require special or expensive telemetrical instruments. In addition to the attractive display the surface elevation models also provide a lot of additional information to specialists such as planning the automated cultivation of agricultural areas where the exact knowledge about the elevation is a mandatory. Other advantages are the e.g. area and volume measurements for mining (Fig. 15 and Fig. 16) and real 3D representation of the reconstructed objects (Fig. 17.).

Using GPGPU card computation speed can be usually 10 times faster even in a desktop PC. For example the shown 'Chapel' reconstruction times are:

- 8-core Intel i7 2.4 GHz, 64 bit processor, no GPGPU: 20hrs
- 8-core Intel i7 2.4 GHz, 64 bit processor, with nVidia Titan X GPGPU enabled: 2hrs

The 3D reconstruction of buildings and other objects is not widespread yet. There are only initiatives and experiments whose implementation will be the topic of the future. However, they are already on display in virtual museums when introducing different styles in architecture or geomorphological formations.

References

[1] K. Douterloigne, S. Gautama, W. Philips: On the accuracy of 3D landscapes from UAV image data, (2010) International Geoscience and Remote Sensing Symposium (IGARSS), art. no. 5651391, pp. 589-592.

[2] P.J. Zarco-Tejada, R. Diaz-Varela, V. Angileri, P. Loudjani: Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods (2014) European Journal of Agronomy, 55, pp. 89-99.

[3] C. Martínez, I.F. Mondragón, M.A. Olivares-Méndez, P. Campoy: On-board and ground visual pose estimation techniques for UAV control, (2011) Journal of Intelligent and Robotic Systems: Theory and Applications, 61 (1-4), pp. 301-320.

[4] J. Sládek, M. Rusnák: Low-cost micro UAV technologies in geography (a new method of spatial data collection) [Nizkonákladove mikro-UAV techológie v geografii (nová metóda zberu priestorových dát)],(2013) Geograficky Casopis, 65 (3), pp. 269-285.

[5] F. Bachmann, R. Herbst, R. Gebbers, V.V., Hafner: Micro UAV based georeferenced orthophoto generation in VIS+NIR for precision agriculture, (2013) International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 40 (1W2), pp. 11-16.

[6] A. Lucieer, S.M. Jong, D. Turner: Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography, (2014) Progress in Physical Geography, 38 (1), pp. 97-116.

[7] T. Berteška, B. Ruzgiene: Photogrammetric mapping based on UAV imagery, (2013) Geodesy and Cartography, 39 (4), pp. 158-163.

[8] H. Eisenbeiss, M. Sauerbier: Investigation of UAV systems and flight modes for photogrammetric applications, (2011) Photogrammetric Record, 26 (136), pp. 400-421.

[9] L. Barazzetti, F. Remondino, M. Scaioni, R. Brumana: Fully automatic UAV image-based sensor orientation, (2010) International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 38

[10] S. Szénási, Z. Vámossy: Evolutionary algorithm for optimizing parameters of GPGPU-based image segmentation (2013) Acta Polytechnica Hungarica, 10 (5), pp. 7-28.

[11] G. Kertész, S. Szénási, Z. Vámossy: Parallelization Methods of the Template Matching Method on Graphics Accelerators, (2015) 16th IEEE International Symposium on Computational Intelligence and Informatics – CINTI, pp. 161-164. ISBN:978-1-4673-8520-6

[12] C. Flener et. al.: Seamless mapping of river channels at high resolution using mobile liDAR and UAV-photography, (2013) Remote Sensing, 5 (12), pp. 6382-6407

A. Molnar was born in Moscow, USSR on 5 May 1967, and received his MSc degree at University of Szeged Faculty of Science and Informatics, Hungary in 2001. He received his PhD degree in 2006 at the Zrinyi Miklos National Defence University, Hungary. He is an Associate Professor at J. Selye University, Slovakia. His main research areas are autonomous control of Unmanned Aerial Vehicles and related applications.

D. Stojcsics was born in Budapest, Hungary on 21 February 1984, and received his MSc degree at Obuda University John von Neumann Faculty of Informatics, Budapest, Hungary in 2010. He received his PhD degree in 2012 at the Obuda University, Doctoral School of Applied Informatics. He is a Lecturer at J. Selye University, Slovakia. His main research area is guidance methods of fixed wing Unmanned Aerial Vehicles.