Computational Vision in Photogrammetry for Georeferencing: Modern Resources Evaluation for UAV Image Processing

Henrique Lima de Sousa
Universidade Federal Rural do Rio de Janeiro
Henrique.lima@ufrrj.br

DOI: https://dx.doi.org/10.4314/sajg.v14i1.2

Abstract

Environmental Science aims to understand the world; this can be achieved by using geographic information systems (GISs). Also, georeferencing allows for the adjustment and alignment of raster data in combination with other GIS data. Thus, it is possible in terms of these techniques to interpret these types of data and their relationships, patterns, and trends. This study aims to investigate the use of modern engineering procedures, known as computational vision in photogrammetric image processing, as obtained from unarmed aerial vehicles (UAV). This is done via a small camera in the front of an embedded system and combined with proprietary software that uses computational vision resources. Although open-source software was the prioritized choice, the research began with a study on state-of-the-art computational vision algorithms and photogrammetry for drone inspection. The generation, processing, and verification of a set of photographic images were further procedures accompanying the study of the algorithms and photogrammetry which subsequently resulted in a georeferencing system. In fact, proprietary software employing computational vision resources was used in this study at the Universidade de Trás-os-Montes e Alto Douro in Portugal to compare it with the conventional methodology using modern computational resources to determine the benefits achieved. In conclusion, the positional quality of the generated georeferencing system was verified, and satisfactory results were reported. This underscored the potential of these modern computational resources in contemporary photogrammetry.

***Key words**: Computer Vision, Image Processing, Photogrammetry, UAV, GIS*

1. Introduction

It is a stark fact that the use of drones in data acquisition has become more popular worldwide, especially during the last decade. This has been since the associated parts have become less expensive in the quest to obtain cartographic products of a high geometrical quality.

However, it is worth noting that even in its boom days, there are still some limitations regarding drone usage in acquiring data in some situations, such as its instability during flightand the application of simpler sensors. In general, in the latter case, for instance, small non-metric cameras are used. However, a cheaper solution without the correct approach might in fact lead to a solution that might not be entirely successful.

One of the problems mentioned above is called matching. It includes the correlation between relevant variables and the subsequent identification of features between each pair of aerial photographs. According to Galo (2003), all the acquired images through unarmed aerial vehicles relate to conditions of the same rotation, convergence, and difference between scales.

However, Computational vision has in fact already been used to optimize image processing and to yield better cartographic products with more suitable qualities for each application.

Some projects, such as MobileFusion, and Pointless Structure from Motion, which are discussed in this paper, prove that computational vision optimizes images that are processed by mobile phones, According to Ondruska, Kohli & Izadi (2015) and by Nurutdinova, Fitzgibbon (2015).

Most of the photogrammetric processing solution apps are still used as the proprietary software. However, some, such as Opendronemap, are open source and use modern resources. Moutinho (2015) shows the quality of these products via data acquired through drones. Other related research findings could be found, such as those of ISPRS (International Society for Photogrammetry and Remote Sensing).

Reflections in this study about the use of computational vision in the processing photogrammetric concur with the findings of studies that verify the computational gains regarding versus photogrammetric images obtained through remote platforms as opposed to conventional image processing.

2. State of the Art

As mentioned above, computational vision can increase its potential, as well as its optimization of image processing. The two Microsoft projects, known as MobileFusion and Pointless Structure from Motion, are examples to demonstrate the use of modern computer science usage to generate high-quality products, such as digital three-dimensional models, through standard photographs.

Firstly, as shown in Figure 1, MobileFusion is a decoder and image processor that reconstructs real-time volumetric surfaces through images captured from different angles. Ondruska, Kohli & Izadi (2015) show the sweeping of a set of objects, and via a motion sensor, the Kinect, compare them to a baseline The result shows an approximate error of 1.5 cm.

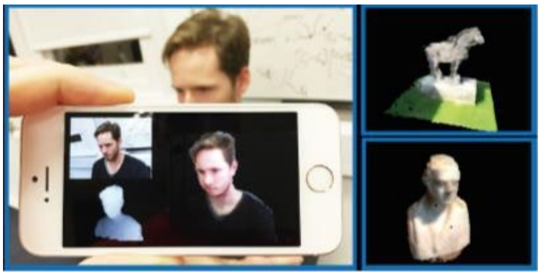

Figure 1: Use of MobileFusion (Ondruska, Kohli & Izadi, 2015)

Secondly, as seen in Figure 2, Pointless Structure from Motion is about a computational vision method to extract high resolution data information from a different range of scales to detect three-dimensional curves. It could be used to better estimate the position of the camera to portray challenging scenarios. Low-texture cameras would be the appropriate type to use in this case (Nurutdinova & Fitzgibbon, 2015)

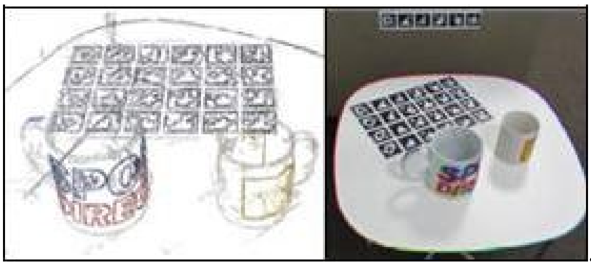

Figure 2: Pointless Structure from Motion (Nurutdinova & Fitzgibbon, 2015)

Once it is possible to obtain an acceptable processed image from low-quality images (e.g., images taken from a mobile phone), the former would not differ that much from drone images.

According to Moutinho (2015), it is possible to obtain cartographic products with outstanding positional precision and accuracy through RBG and NIR processed images. The study, conducted in three different areas in Portugal, made a comparison between two commercial and one open-source software type.

Similarly, Sousa (2016) photogrammetrically processed some of the images used by Moutinho (2015) by using the proprietary software, Agisoft Photoscan, which uses a computational vision algorithm. The comparison between the cartographically generated products that was subsequently made, is presented as follows:

Digital Surface model (DSM), Digital Terrain Model (DTM) and orthomosaic models: They all produced acceptable quality for data storage and environmental monitoring.

Photogrammetry, based on computer vision algorithms is what the remote sensor community envisions as the new trend for embedded sensors that are remotely controlled by drones, for example. Since it is to be expected, an optimization in terms of processing is necessary to overcome the limitations regarding low quality images with a compressed format or even through distortions due to the drone's instability that conventional methodologies generally impose.

Mathematical models for the orientation and generation of conventional surfaces, and their subsequent optimization, are foreseen. Apart from direct georeferencing, better image pattern recognition is possible through these means. For these purposes, several computational resources connect with mathematical models:

Firstly, there is SIFT (scale-invariant feature transformation), which includes a procedure that extracts and detects characteristically relevant spots on images. This makes it possible to obtain descriptors, that do not vary in terms of their parameters, such as scale, rotation, and noise (Sousa, 2016).

According to Lowe (2004), the SIFT algorithm is divided into two parts. The first one, the detector, is based on a Gaussian difference. The descriptor, on the other hand, is based on gradient histograms, which are orientated to describe the local borders of the point of interest. However, this descriptive procedure is conducted in steps: the detection and localization of extreme points, and the detection and localization of key points.

Another simple procedure that could be applied in the computational vision procedure with photogrammetry is to apply the K-means clustering algorithm. According to Dingding *et al*. (2011), this involves conducting a pixelated image in K groups, which initializes parameters, gives attributes to each pixel grouping, and recomputes them.

Similarly, the expectation-maximization clustering algorithm works in three steps. However, it is more generic than the one before, since its function is to find estimations based on parametric models for incomplete data. However, the groupings are represented in terms of probability distributions, and not just through means. In fact, Gaussian distributions are commonly used. For this purpose, clusters are then represented by a mean (μ) and the covariance matrix (Σ) (Dingding *et al*, 2011).

Aiming to optimize the two procedures mentioned above, mean-shift clustering was designed to set several clusters that are theoretically solid. It starts with windows centered on a random tray and to change its center according to its data until the window converges into a histogram mode. Thus, it defines a group, and the process continues in a loop until each bin is associated with a computed mode.

Finally, the interactive image segmentation algorithm provides flexibility and precision for object segmentation in that it incorporates user input. Also, intuitive usage and its speed are two important fundamental aspects. Two approaches were presented by Dingding *et al* (2011): effective image segmentation with an automatic procedure for the refinement of a new limit and rapid image segmentation through descriptive grouping.

As per Dingding *et al.* (2011), the initial approach involves under-segmenting an image into multiple regions. Subsequently, as depicted in Figure 3, for the unlimited portion of the object, the under-segmented areas are combined using the average color of each region in the CIELAB space as a reference.

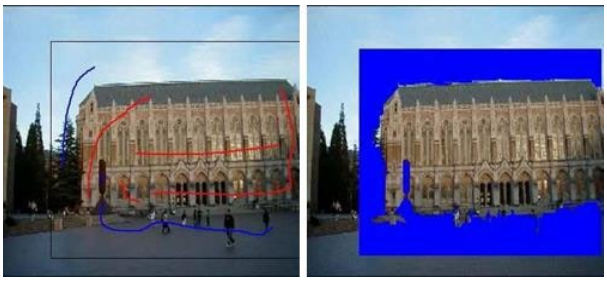

Figure 3: Interactive image segmentation algorithm – the result of an automatic refinement of the procedure (Dingding *et al*, 2011).

As pointed out by Dingding *et al* (2011), the second approach has been investigated by Liu *et al* (2010). As shown in Figure 4, it starts with an initial under-segmentation, using the mean shift algorithm and keeps to clusterization in terms of descriptive and local groupings. It is easy to implement, faster than the graph cuts, and it could be accelerated through an image pyramid and a limits refinement procedure.

Figure 4: Interactive image segmentation algorithm – the result of a descriptive grouping (Dingding *et al*, 2011).

3. Methodology

This study aims to extract cartographic products, such as orthomosaic and digital models of surface and terrain, with the aid of a UAV (Unarmed Aerial Vehicle) with computational vision. The goal is to verify the gains achieved through computational vision in the photogrammetric processing of images, especially through those obtained through sensors mounted on remotely controlled platforms. It stands in contrast to traditional photogrammetry, that does not use modern software solution resources like computational vision.

However, Ondruska, Kohli & Izadi (2015) showed the possibility of rapidly obtaining three-dimensional models using products with high accuracy in the light of the greater input of data in such cases. Therefore, the larger the time interval for exposure, the more advanced the computational methods for image processing needed. All the obtained images through this platform were generally obtained via small UAV cameras, however.

In fact, the issues with small cameras vary according to the scale of mapping and the errors induced by angle variation in terms of the camera's viewings. Compared to manned flights, metric cameras could yield a larger image volume with more extensive overlaps.

To verify the product's quality, extracted by means of software that uses computational vision algorithms, a set of 50 images, obtained through UAV-embedded small cameras, was processed, thereby generating orthomosaic and digital elevation models (DEMs). The 300-hectare land area included three floor buildings and botanical gardens. The embedded sensor used was a Canon IXUS 127 HS, which consists of a RBG non-metric camera with a 16.1 MP resolution, as well as a 4mm focal length.

To achieve this, software developed by the Russian company, Agisoft LCC, and known as Agisoft Photoscan, was used. This software employs computational vision algorithms, such as the Scale-Invariant Feature Transformation (SIFT) to identify corresponding points between images and align them within an arbitrary coordinate system. Consequently, it has the potential to generate a sparse point cloud.

After that, camera auto-calibration was carried out using control points from data acquired and identified in images. Also, as previously described, parameters (e.g., attitude, and control and perspective centers) _were adjusted. Therefore, a geo-referencing system could be made of the image set.

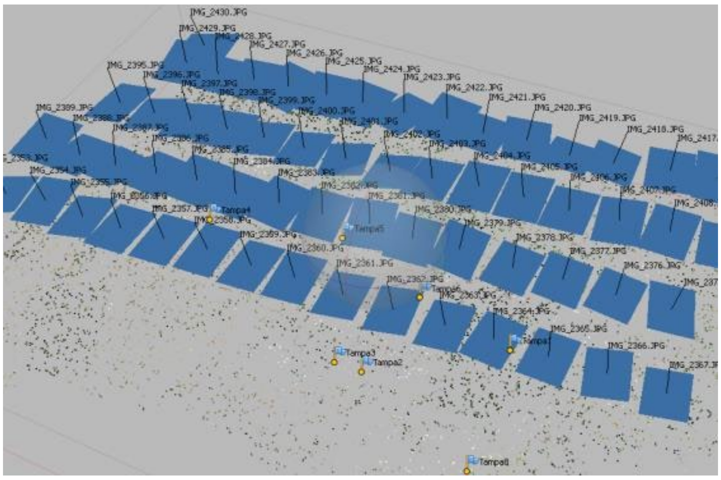

Figure 5: Images aligned and adjusted

The next step consisted of the generation of a cloud of dense points. This was done through the parameters of interior and exterior orientation, as calculated previously. Some image patterns could be automatically identified and a cloud of points created through conventional photogrammetric procedures. At this stage, when compared to the previous one, this cloud of points was already georeferenced with higher density properties.

The points structure constituted the last step. It was created by means of a triangular irregular network (TIN) that was later used to generate a digital elevation model (DEM).

Figure 6: DEM generated with attributed texture

Besides that, also with the same software, an orthomosaic model, as seen in Figure 7, was generated automatically. With a 1:1,000 scale, it consisted of a set of auto-rectified images. , with its pre-defined parameters determined by the image processing software.

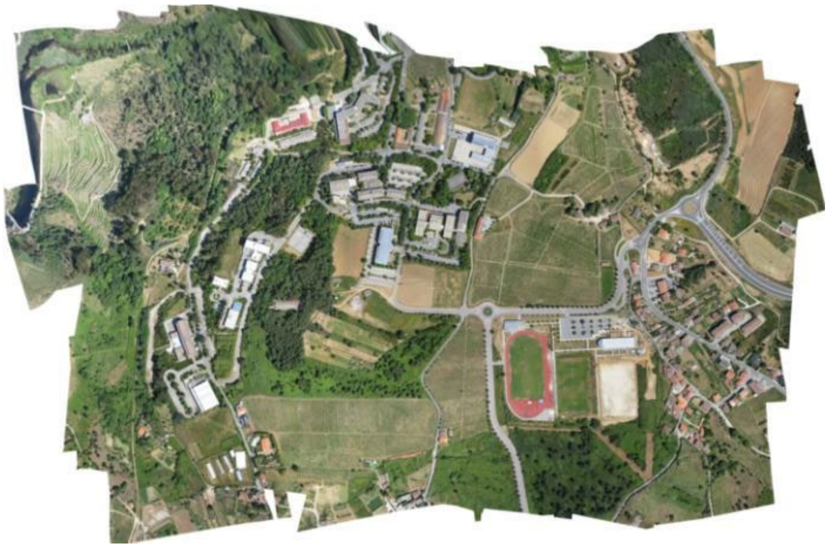

Figure 7: Generated orthomosaic model

The quality analysis exercise which was then undertaken was verified by collecting 38 data verification points via the RTK method (Real-Time Kinematics) using manhole covers on highways, thereby making it possible to identify these points on cartographic products. They further helped to explain the proximity between specific points. It should be noted that this region is considered semi-urban and that there are not many highways criss-crossing it.

Further to the above exercise, the coordinates of the points were extracted and compared. Quality parameters, standardized by Brazilian institutions such as IBGE (Brazilian Geography and Statistics Institute, literally translated), as well as the Brazilian army, were used for collecting statistics and the subsequent analysis. (These standards to ensure cartographic accuracy for digital cartographic products are known as *Padrão de Exatidão Cartográfica dos Produtos Cartográficos Digitais* (PEC-PCD; Cartographic Accuracy Standard for Digital Cartographic Products).

Figure 8: Some of the data verification points used

4. Results and Discussion

The quality analysis used 38 RTK data points from manhole covers along highways, identified on the cartographic products, as can be observed ahead.

Table 1: Tendency Tests for Systematic Errors

Direction	Discrepancies mean	Mean squared error	*t*D	*t* (37, 0,5)	Tendency
E	0.125	0.076	0.078	1.687	No tendency
N	0.108	0.076	0.982	1.687	No tendency
H	0.065	0.248	1.616	1.687	No tendency

Similarly, the precision was verified using the x values considering the standard errors for Class A on a scale of 1:1,000. As seen in the following table, the results were partially satisfying:

Table 2: Positional Precision Analysis

Direction	χ D	χ^2 (37,0.9)	Class A
E	40.00	48.35	Meets the criteria
N	29.95	48.35	Meets the criteria
H	79.01	48.35	Does not meet the criteria

As can be seen in Table 2, the planimetry suits the precision as determined for Class A with a scale of 1:1,000. However, altimetry does not suit the same condition.

Therefore, the same calculation was done for the x axis and coordinate H on the same scale and in respect of Class B. The obtained value on the same scale was χ2H = 20.97. In this case, the condition was suitable, which means that altimetry suits the corresponding scale for Class B. However, because many of the applications do not require a higher quality positional product, it does not disqualify the product,

Based on the determined values, it is possible to affirm that the cartographic products that were generated show a better positional quality of planimetric precision and accuracy. As is evident from figures 9 and 10, it is relatively high and even adequate, depending on the purposes and requirements for environmental monitoring and multipurpose registration, as verified by Sousa (2016).

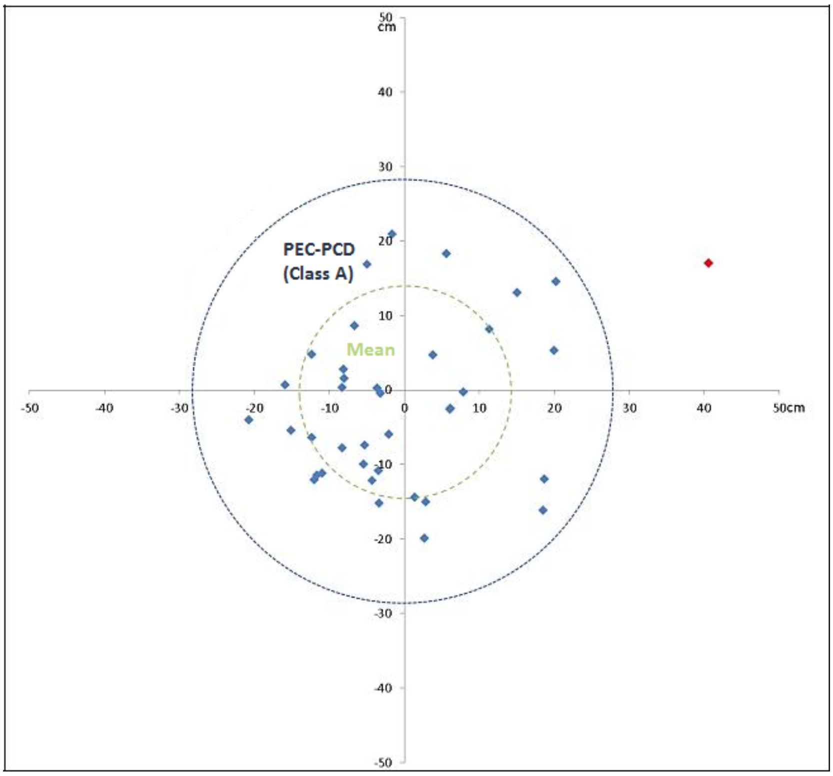

Figure 9: Planimetric accuracy and planimetric precision

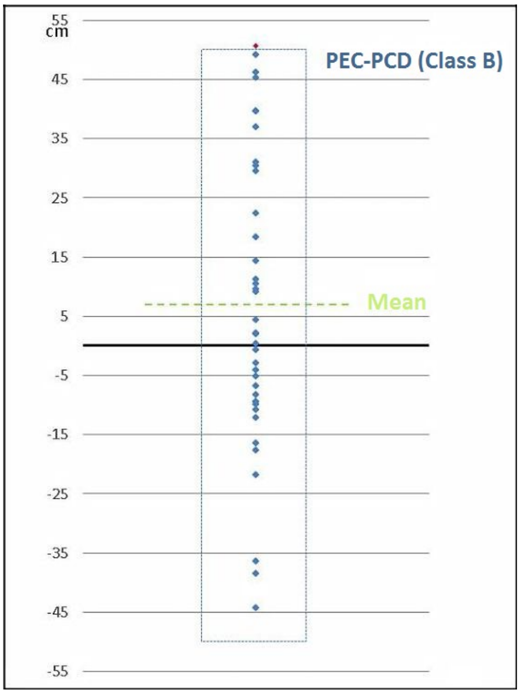

Figure 10: Altimetric accuracy and altimetric precision

5. Conclusion

The results emanating from this paper, although not conclusive, could indicate the benefits of photogrammetry via drones, and the use of geoinformation, in conjunction with modern computational vision. In fact, these computational procedures, ensuring that image processing is suitable, are made possible through embedded sensors in aerial platforms that can create high quality images.

It is expected that this research could inspire the development of new incentives, methodologies and techniques. Besides these, it could assist in the promotion of learning and in finding new avenues for research and its forthcoming solutions, especially in cases where open-source software is used.

Also, it is recommended that in-depth research be conducted to obtain even better results, thereby comparing them with those issuing from different types of software, especially open-source software. Furthermore, it could develop new methodologies to compare with new types of data acquisition such as topographic, geodetic and other cartographic products.

It is possible to state that the use of aerial platforms to obtain geoinformation in conjunction with the photogrammetric procedures discussed above and already used in the literature, could optimize image processing and generation to make high quality cartographic products available.

6. References

Dingding, L., Bilge S., Gregg, P.; Linda S., 2011. A Review of Computer Vision Segmentation Algorithms. University of Washington.

Galo, M., 2013. Automação dos processos de correspondência e orientação relativa em visão estéreo [Automation of corresponding processes and relative orientation in stero vision]. Ph.D. Campinas, SP, Brazil: Universidade Estadual de Campinas.

Liu, D., 2003. Intelligent Interactive Image Segmentation for Camera Phones. Ph.D. Undergraduate, University of Washington, Seattle, WA, United States of America.

Liu, D, Pulli, K, Shapiro, LG & Xiong, Y 2010, Fast interactive image segmentation by discriminative clustering. in *MCMC'10 - Proceedings of the 2010 ACM Workshop on Mobile Cloud Media Computing, Co-located with ACM Multimedia 2010.* pp. 47-52, 2010 ACM Workshop on Mobile Cloud Media Computing, MCMC'10, Co-located with ACM Multimedia 2010, Firenze, Italy, 25-25 October. DOI: 10.1145/1877953.1877967

Lowe, D.G., 2014. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, n 60(2), 91–110.

Moutinho, O, 2015. Evaluation of Photogrammetric Solutions for RPAS: Commercial vs Open Source. MSc. Porto, Portugal: Universidade do Porto.

Nurutdinova, I., Fitzgibbon, A., 2015. Towards Pointless Structure from Motion: 3D Reconstruction and Camera Parameters from General 3D Curves, in: 2015 IEEE International Conference on Computer Vision (ICCV). Presented at the 2015 IEEE International Conference on Computer Vision (ICCV), IEEE. https://doi.org/10.1109/iccv.2015.272

Ondruska, P., Kohli, P., Izadi, S., 2015. MobileFusion: Real-time Volumetric Surface Reconstruction and Dense Tracking on Mobile Phones. IEEE Transactions on Visualization and Computer Graphics 21, 1251–1258. https://doi.org/10.1109/tvcg.2015.2459902

Sousa, H.L., 2016. Veículos aéreos não tripulados (VANT) na gestão territorial: Análise da viabilidade de seus usos em cadastros multifinalitários e controles ambientais [Unarmed aerial vehicles (UAV) for terriotorial management: Viability analysis of its multipurpose register and environmental controls]. Bachelor' Degree. Rio de Janeiro: Universidade do Estado do Rio de Janeiro.