

COMPUTATIONAL TIME MODELLING FOR TREE CROWN EXTRACTION FROM HIGH RESOLUTION IMAGERY USING GEOGRAPHIC OBJECT-BASED IMAGE ANALYSIS

J.A. Okojie^{1*}, A.E. Effiom², E.N. Tawiah², I.J. Akpejori²

¹ GeoNET Research Initiatives, Abeokuta, Nigeria – okojie35424@alumni.itc.nl

² Faculty of Geoinformation Science and Earth Observation, Enschede, the Netherlands - agboreffiom2004@yahoo.ca, ekown.tawiah@gmail.com, julietakpejori@gmail.com

KEY WORDS: Computational time, spatial resolution, pixel count, photogrammetry, segmentation, regression, sensitivity, modelling

ABSTRACT:

Image segmentation, unlike most conventional approaches, assesses image pixel information in relation to image contextual information to extract congruent image objects. A comprehensive review of literature during the research conceptualization phase revealed that while a lot of work has been done into the optimization of the segmentation computational time, there was a conspicuous research gap in the area of computational time modelling and pre-analytical projections.

The focus of this research was, therefore, to model the computational time requirements of segmentation as a function of variability in image spatial resolutions and pixel counts. The employed methodology integrated structure from motion (SfM), segmentation, and regression modelling into a multiphase workflow to process UAV-acquired imagery into an orthomosaic, which underwent segmentation at multiple resolutions. The segmentation computational time was regressed against the image pixel count and spatial resolution.

The results showed that the image pixel count was the main determinant of the computational time, expressing a statistically strong linear relationship. However, when the regressors were tested individually, the spatial resolution showed an exponential relationship with computational time which, however, exhibited saturation. It can be inferred that image spatial resolution is not the key determinant of segmentation computational time but the pixel count of the image.

This research recommends the adoption of statistically viable models to thoroughly assess computational time requirements for image segmentation and a pre-analytics estimation of the time requirements such that image quality – image processing time adjustments can be made before rather than after analyses have commenced.

1. INTRODUCTION

1.1 Background

Computational time constraint is one of the major factors affecting feasibility and quality of research. The analysis of remotely sensed images from airborne and space-borne platforms provides valuable insight for mapping, environmental monitoring, disaster management and civil and military intelligence (Benz, Hofmann, Willhauck, Lingenfelder, & Heynen, 2004; Turner et al., 2003; Vihervaara et al., 2017). Remote sensing imagery has to be converted into tangible information which can be utilised or integrated with other data sets for the assessment of geographical phenomena (Blaschke, 2010a). Originally, pixel sizes were typically coarser than, or similar in size to the objects of interest and emphasis was placed on both pixel-based and sub-pixel analysis for this conversion. However, recent increases in the availability of remote sensing datasets with higher spatial resolutions triggered a shift from pixel to object-based image analysis.

The object-oriented approach can contribute to powerful automatic and semi-automatic analysis for most remote sensing applications (Benz et al., 2004). GEOgraphic Object-Based Image Analysis (GEOBIA) has been progressively gaining importance in the fields of remote sensing and geographic information science over the last decade, especially for the

processing of very high spatial resolution imagery (Blaschke, 2010b; Drăguț, Csillik, Eisank, & Tiede, 2014a). Creating representative image objects with image segmentation algorithms is an important pre-requisite for classification/feature extraction and further integration in geographical information systems (GIS) analysis. However, in most cases system requirements.

Multi-Resolution Segmentation (MRS) is probably the most common algorithm for these purposes (Baatz & Schäpe, 2000), as it has the dexterity to extract image objects existing at varying scales simultaneously (Baatz & Schäpe, 2000; Okojie, 2017). Since the implementation of MRS in the eCognition@ software (Trimble Geospatial Imaging), this algorithm has quickly become one of the most prominent segmentation algorithms within the GEOBIA domain. MRS relies on a core control, the scale parameter (SP), to split up an image into cogent image objects. The SP controls the internal (spectral) heterogeneity of the image objects and is therefore correlated with the average size of derived image objects, i.e., the higher the SP value, the higher the allowable internal heterogeneity within image objects which consequently, increases the number of pixels per image-object (Baatz and Schäpe, 2000, Benz et al., 2004). Within the field of GEOBIA, a lot of research has been done into the production of analytical pathways optimized for the best possible performance in the shortest possible time

(Csillik, 2017; Drönner et al., 2018; Gonçalves, Pôças, Marcos, Múcher, & Honrado, 2019). There is, however, still a dearth of information pertaining to the exact computational time requirements for Object-based image analysis (OBIA), especially image segmentation. GEOBIA can be time-consuming especially in the extraction of information from very large images due to the fact that most segmentation algorithms make use of the pixel-grid of an image for initial object representation (Csillik, 2017). However, there is little or no readily available information on the relationship between image size (as influenced by spatial resolution) and segmentation computational time. Thus, this research sought to investigate the sensitivity of segmentation computational time to variation in image spatial resolution and pixel count. Precisely, the work will assess the total computational time of segmentation at multiple resolutions for the extraction of forest tree crowns and the sensitivity of segmentation computational time to variation in image spatial resolution and image pixel count.

2. METHODOLOGY

2.1 Flowchart

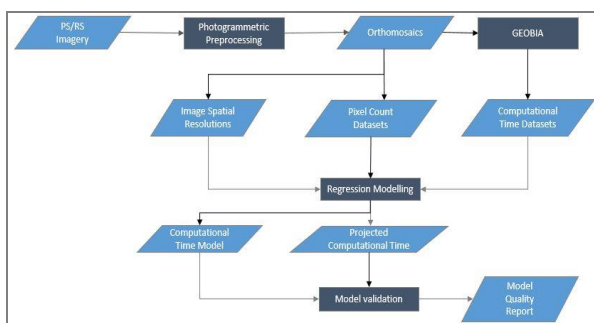


Figure 1. Methodological flowchart

2.2 Datasets and Pre-processing

The raw image datasets used for this research were of a temperate forest tree stand located within Ahaus, a municipality in the district of Borken, North Rhine-Westphalia. The image datasets were captured using a type G180 fixed-wing UAV with a focal length of 18.3mm and the number of recorded pixels was 4928 x 3264 captured in RGB (Red, Green & Blue). These were processed into an orthophoto. The software requirement included Pix4D mapper, ArcMap, eCognition, Microsoft Excel. Also, the ESP2 tool was installed within the eCognition environment and utilized. Datasets used within this study were sourced from the processing of a type G180 UAV-RGB imagery. The sensor on-board the UAV was a GR Lens 18.3 with a focal length of 18.3mm and the number of recorded pixels was 4928 x 3264 captured in RGB (Red, Green & Blue). The captured imagery was processed using a photogrammetric workflow as detailed below;

- **Initial Processing;** involves tie-point detection, description, and matching to give the images relative orientation. Due to the side and forward overlap, consecutive images have similar features. Tie-points refer to points identifying similar features in different images. Once the images are loaded into pix4D, the software identifies similar features in the images. Based on these similar features (points), the images are given an orientation relative to one another in space. With identified tie-points, the images are calibrated. The software then

iteratively uses samples of the tie points to build a model that determines the best orientation of the images. This iterative process is described as Random Sample Consensus (RANSAC). RANSAC reduces the errors associated with giving the photos geolocation on the ground (absolute orientation) (Fischler & Bolles, 1981). Once the report shows sufficient image calibration, ground control points collected during fieldwork are loaded to optimize the calibration process and give the calibrated images geolocation on the ground (absolute orientation). Once the quality is appropriate, as revealed in the quality report, the next phase is initiated.

- **Point Cloud Densification;** this was done to increase the density of the 3D points which were generated from the initial processing using Area-Based Matching (ABM) algorithms. This increases the accuracy of the eventually outputted digital surface model and orthomosaic (Pix4D, 2018).

DSM and Orthomosaic generation;

- **DSM;** The DSM was generated by triangulation of the dense point cloud and was outputted in GeoTIFF format.
- **Orthomosaic;** the orthomosaic was generated from the mosaicking of the orthorectified images (geometrically corrected images possessing a uniform scale). The orthomosaic generated from this process was a true orthophoto as it was generated with the DSM taken into consideration (Rau et al., 2002). It was generated in multiple GeoTIFF tiles which were later merged into a single orthophoto (Pix4D, 2016).
- **Image Resampling;** this was done on an extracted (clipped) portion of the orthomosaic to ensure uniformity in the area and spatial extent of images to be utilized for the multi-resolution segmentation analyses. It was conducted using Bilinear Interpolation. The image resolutions in centimeters were; 11, 15, 20 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105 and 110, adding up to 21 spatial resolutions.

2.3 Analytical Approach

All segmentation analyses were conducted on a single computing system equipped with Windows 10 Home Edition (Operating System) and the processor specifications are Intel © Core™ i5-5200U CPU@ 2.20GHz 2.19GHz, Installed Memory (Random Access Memory); 8GB, System Type; 64-bit Operating System, x64-based processor. This was kept constant all through the image processing and computational time extraction.

2.3.1 Segmentation

The segmentation analyses were focused primarily on the extraction of forest crown image object from the orthomosaic and its time implications. This was done using the ESP2 tool created by Drăguț et al. (2010) for the automated parameterization of multi-scale image segmentation (Drăguț, Tiede, & Levick, 2010). This approach relies on the potential of the local variance (LV) to detect scale transitions in loaded image layers (RGB bands).

The tool functions by detecting the number of image layers added to a project and segmenting them iteratively utilizing a multiresolution segmentation algorithm with a bottom-up approach, wherein the scale factor in the segmentation (scale parameter; SP), increases with a constant lag. The mean LV value of objects in all of the layers is then computed and serves as a pre-condition for stopping the iterations i.e. when a scale level records an LV value equal to or lower than the previous value, iteration ends, and the objects segmented in the previous level are retained (Drăguț, Csillik, Eisank, & Tiede, 2014). The ESP2 tool settings can be viewed in figure 1 below.

- i. Select map; At this step, the loaded image is selected as the 'Main map'
- ii. Use of hierarchy; This was set to 1 indicating that each scale parameter was to be generated within a top-down or bottom-up approach.
- iii. Hierarchy; This was set to 1 indicating that the algorithm should follow a bottom-up approach i.e. analyses starts from the smallest level and merges objects to obtain higher levels.
- iv. 'Starting scale Level 1', 'Starting scale Level 2' and 'Starting scale Level 3'; indicates the minimum scale parameters at which processing starts, for the three levels which are going to be created. All starting scales were set to 1.
- v. 'Step size Level 1', 'Step size Level 2' and 'Step size Level 3'; specifies the increments of scale parameter for the step-wise segmentation processing within the algorithm. For the three levels being created, the values were set at 1, 10 and 100 respectively.
- vi. 'Shape' and 'Compactness'; specifies the composition of the homogeneity criterion as implemented in the multiresolution segmentation. Both of these were set to 0.1 and 0.5 respectively.
- vii. 'Produce LV Graph'; this specifies the production of text files with local variance values stored. The file can be visualized using the ESP_Estimation_Scale_Parameter_Chart.exe tool. This was set to 1 indicating that the text files should be produced.
- viii. 'Number of loops'; stipulates the number of scale levels to be generated if 'Produce LV Graph' was set to 1. Number of scale levels was set to 100.

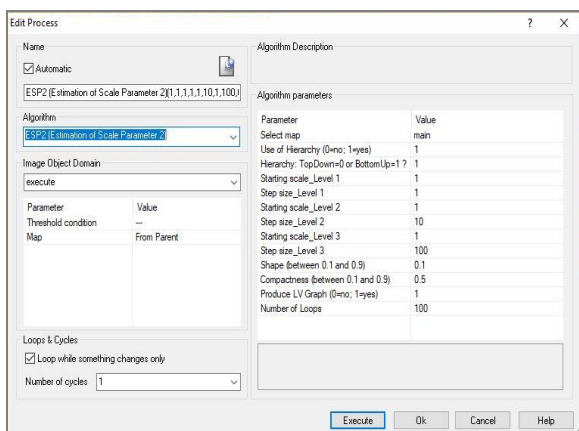


Figure 2. ESP2 Algorithm Description, 2019.

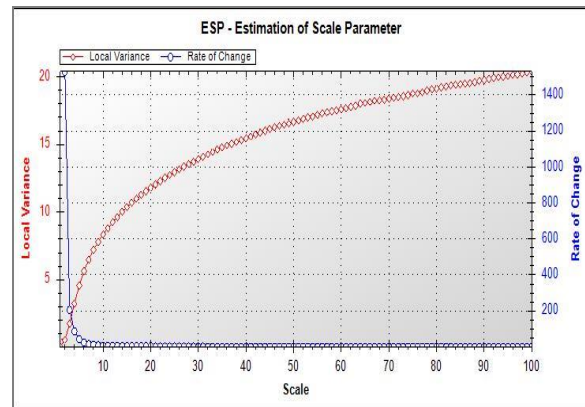


Figure 3. Scale Parameter Estimation Chart at 11cm Resolution, 2019.

2.32 Regression Modelling

From the review of literature, it was inferred that spatial resolution of images was a strong determinant of image pixel count. As such in order to avoid multicollinearity within the regression modelling, some ad hoc tests were conducted including Pearson's Correlation test and Variance Inflation Factor (VIF) computation and assessment.

Computational time costs of segmentation for all specified image resolutions were extracted and regressed against the image spatial resolutions and Image pixel counts as specified in the model below;

$$Y = \beta_0 + \beta_1 X_1 + e \quad (i)$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e \quad (ii)$$

Where, Y = Computational Time in Seconds

β_0 = Slope-Intercept

$\beta_1, 2$ = Regression coefficients of X_1

X_1 = Image Spatial Resolution

X_2 = Image Pixel Count

2.33 Model Validation

Model validation was conducted using an F-test and T-tests. This was done by utilizing created models for projecting the image computational time of the previously segmented images and comparing the modelled time to the actual time used for computation. The F-test and T-test were used to statistically test for a significant difference in variances and means between the actual and modelled computational times.

3. RESULTS

3.1 Correlation and Collinearity

The results of the correlation tests showed a strong inverse relationship existing between image spatial resolution and image pixel count (-0.671). This was significant at 99.9% ($p = 0.00085 < \alpha \leq 0.001$). However, the collinearity test showed there was no multicollinearity among the covariates.

Parameters	Values
Alpha	0.05
Tails	2
Corr. Coeff.	-0.672
Std. Error	0.170
t	-3.952
p-value	0.00085
lower	-1.027
upper	-0.316

Table 1. Correlation Matrix

3.2 Segmentation Output

Table 2 shows the computational time requirements for the segmentation of the UAV-RGB orthomosaics per image spatial resolution. From the results, it can already be inferred that the computational time requirements for segmentation vary directly with both pixel count and image spatial resolution implying that the coarser the spatial resolution of an image, the lower the pixel count and the lower the computational time required to segment the image. However, the variation rate and computational time sensitivity to the change in spatial resolutions and pixel counts are not inferable from this table. To better understand the nature of this trend, regression modelling was conducted and the results are presented in the regression output.

Image Resolution (cm)	Pixel Count (Unit)	Computational Time (Seconds)
11	3989586	822
15	2174655	613
20	1223480	237
25	783144	173
30	543979	126
35	399486	97
40	305870	83
45	241380	75
50	195640	65
55	161385	56
60	135837	51
65	116100	40
70	100111	40
75	87165	39
80	76258	33
85	67768	40
90	60264	34
95	54362	32
100	48910	29
105	44341	31
110	40565	28

Table 2. GEOBIA Resolution/Time Matrix

3.3 Regression Output

3.3.1 Regression Model Fit and Significance: The overall model fit matrix shows that approximately 97.4% of the variation in the computational time values can be explained by the model. This implies that the statistical uncertainty of using this model to estimate the time requirement of segmentation before it is undertaken is approximately 2.6%. The statistical significance of the model was tested at an acceptance threshold of 95%, it, however, proved significant at a statistical confidence level of > 99.9% as can be inferred from the ANOVA table.

ANOVA

ANOVA	df	SS	MS
Regression	2	817599.00	408799.50
Residual	18	21464.58	1192.48
Total	20	839063.58	
Alpha .05 (<i>F</i>)	342.82		
<i>p-value</i>	4.69E-15		
<i>sig</i>	yes		

Table 3. Regression ANOVA Matrix

Multiple R	0.987
R Square	0.974
Adjusted R Square	0.972
Standard Error	34.532
Observations	21

Table 4. Overall Regression Model Fit Matrix

3.3.2 Regression Coefficients and Significance: The regression analysis shows that while there exists an interactive inclination between image resolution, pixel count and computational time (see fig. 3 & 4), only the image pixel count directly and significantly influenced the segmentation computational time. There is a very high sensitivity to the image pixel count, as is evidenced from the regression coefficient which is significant at 99.9% ($p < \alpha \leq 0.001$). As such, it can be inferred that, barring all exogenous factors, the primary determinant of segmentation time is the image pixel count. However, as can be seen from figure 3 below, there exists a very strong exponential negative relationship between the image spatial resolution and the segmentation computational time. This implies that per unit increase in image spatial resolution (decrease in spatial resolution numerical value), a corresponding exponential increase in segmentation computational time occurs. But this relationship saturates, implying that at a particular point, the computational time loses its sensitivity to a spatial resolution change.

Parameters	Resolution		Pixel
	Intercept	(cm)	Count
coefficients	23.31575	-0.04411	0.00021
std err	25.55442	0.33684	0.00001
t stat	0.912	-0.131	19.309
p-value	0.374	0.897	1.77E-13
lower	-30.37	-0.75	0.00
upper	77.00	0.66	0.00
vif		1.82	1.82

Table 5. Coefficients for Computational Time, Image Spatial Resolution and Image Pixel Count

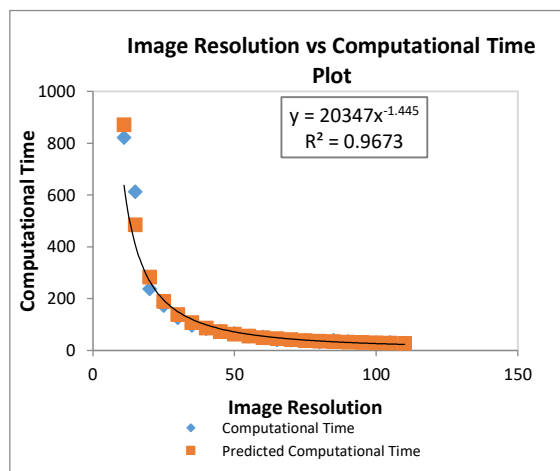


Figure 4; Image Resolution Vs Computational Time Chart

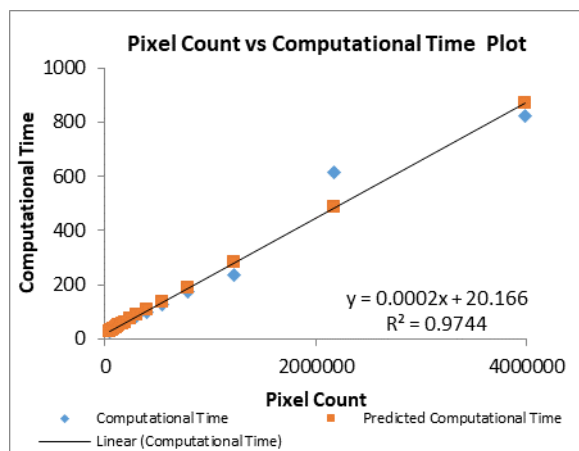


Figure 5; Pixel Count Vs Computational Time Chart

3.4 Model Validation

This was done as explained in the methodology. The results are presented in tables 5 & 6 below. The F-test served as an Ad Hoc test for the statistical determination of the type of t-test to be utilized in the actual comparison of the computed and modelled segmentation time cost. It did this by testing the two variables for equality in variance. The results showed that there was an unequal variance existing between the two variables ($P > \alpha =$

0.05). As such, a two-sample t-test assuming unequal variances was selected for the eventual test of significant difference. From the results, as shown in table 6 below, it can be inferred that there is no significant difference existing between the computed (actual) segmentation time cost and the modelled segmentation time cost ($P > \alpha = 0.05$). This implies that modelled time values for segmentation using the created model will not significantly differ from the actual time values for segmentation of either remote sensing or proximal sensing imagery provided all other variables are kept constant.

	Computed	Modelled
Mean	130.6069	129.1699
Variance	41953.179	39822.83
Observations	21	21
df	20	20
F	1.053	
P(F<=f) one-tail	0.454	
F Critical one-tail	2.124	

Table 6. F-Test Two-Sample for Equality in Variance in Computed and Modelled Analytical Time Cost

	Computed	Modelled
Mean	130.607	129.170
Variance	41953.179	39822.827
Observations	21	21
Hypothesized Mean Difference	0	
df	40	
t Stat	0.023	
P(T<=t) one-tail	0.491	
t Critical one-tail	1.684	
P(T<=t) two-tail	0.982	
t Critical two-tail	2.021	

Table 7. T-Test: Two-Sample Assuming Unequal Variances for Significant Difference in Computed and Modelled Analytical Time Cost

3.5 Conclusion and Recommendation

3.5.1 Pre-knowledge of segmentation time requirements of an image is information that can prove invaluable to a researcher as it can serve as a basis for optimization of the segmentation process. The total segmentation time was found to vary directly with image spatial resolution and the pixel count. This implies that an increase in computational time for assessment can be expected with an increase in spatial resolutions and pixel counts. This is of course contingent on the scene size of the image and the segmentation algorithm in use being kept constant as was done during this research. Also, it has been inferred that image spatial resolution is not the key determinant of segmentation computational time but rather

image spatial resolution in conjunction with the scene size of the image to be analyzed. This is so because these two factors in combination with each other, dictate the pixel count of an image. The computational time sensitivity to variation in pixel counts of images can be considered logical because the segmentation of images is conducted by the assessment of pixel and contextual information for the extraction of meaningful geographic objects.

However, the importance of the image spatial resolution cannot be totally dismissed. Aside from the fact that it influences the pixel count, it exhibits an exponential relationship with computational time. On the basis of this, it can be inferred that per unit increase in spatial resolution triggers an exponential increase in computational time requirements of images. This, in turn, means that computational time can be greatly reduced by splitting imagery into congruent image tiles before performing image segmentation.

As such it is the recommendation of this work is that before segmentation is conducted, the computational time should be modelled either the multiple linear regression (as presented in the result section) or using individually the single indicator regression (as presented in the charts) to gain foreknowledge of the approximate time needed for analysis as this will aid researchers in making proper decisions pertaining to the image based parameters actually required for object-based image analyses. It will also help in deciding on acceptable spatial resolution – computational time trade-offs, due to the foreknowledge of the computational time requirements of imagery to be processed.

REFERENCES

- Baatz, M., & Schäpe, A. (2000). Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII. Beiträge Zum AGIT-Symposium Salzburg 2000, Karlsruhe, Herbert Wichmann Verlag*, 12–23. Retrieved from http://www.ecognition.com/sites/default/files/405_baatz_fp_12.pdf
- Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3–4), 239–258. <http://doi.org/10.1016/j.isprsjprs.2003.10.002>
- Blaschke, T. (2010a). Object-based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2–16. <http://doi.org/10.1016/j.isprsjprs.2009.06.004>
- Blaschke, T. (2010b). Object-based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2–16. <http://doi.org/10.1016/j.isprsjprs.2009.06.004>
- Csillik, O. (2017). Fast Segmentation and Classification of Very High-Resolution Remote Sensing Data Using SLIC Superpixels. *Remote Sensing*, 9(243). <http://doi.org/10.3390/rs9030243>
- Drăguț, L., Csillik, O., Eisank, C., & Tiede, D. (2014a). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88, 119–127. <http://doi.org/10.1016/j.isprsjprs.2013.11.018>
- Drăguț, L., Csillik, O., Eisank, C., & Tiede, D. (2014b). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88, 119–127. <http://doi.org/10.1016/j.isprsjprs.2013.11.018>
- Drăguț, L., Tiede, D., & Levick, S. R. (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, 24(6), 859–871. <http://doi.org/10.1080/13658810903174803>
- Dröchner, J., Korfhage, N., Egli, S., Mühling, M., Thies, B., Bendix, J., ... Seeger, B. (2018). Fast Cloud Segmentation Using Convolutional Neural Networks. *Remote Sensing*, 10(11), 1782. <http://doi.org/10.3390/rs10111782>
- Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395. <http://doi.org/10.1145/358669.358692>
- Gonçalves, J., Pôças, I., Marcos, B., Múcher, C. A., & Honrado, J. P. (2019). SegOptim—A new R package for optimizing object-based image analyses of high-spatial-resolution remotely-sensed data. *International Journal of Applied Earth Observation and Geoinformation*, 76, 218–230. <http://doi.org/10.1016/J.JAG.2018.11.011>
- Okojie, J. (2017). Assessment of Forest Tree Structural Parameter Extractability From Optical Imaging UAV Datasets, in Ahaus. Germany. The University of Twente, Faculty of Geo-information and Earth Observation Science(ITC) MSc thesis. Retrieved from https://webapps.itc.utwente.nl/librarywww/papers_2017/msc/nrm/okojie.pdf
- Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, 18(6), 306–314. [http://doi.org/10.1016/S0169-5347\(03\)00070-3](http://doi.org/10.1016/S0169-5347(03)00070-3)
- Vihervaara, P., Auvinen, A.-P., Mononen, L., Törmä, M., Ahlroth, P., Anttila, S., ... Virkkala, R. (2017). How Essential Biodiversity Variables and remote sensing can help national biodiversity monitoring. *Global Ecology and Conservation*, 10, 43–59. <http://doi.org/10.1016/j.gecco.2017.01.007>