

**A geographical information study on observation  
and analysis methods suitable for capturing  
urban and watershed vegetation: Possibilities of  
360-degree images, ultra high resolution images  
by UAVs and object-based image analysis**

**Ryosuke YAMAMOTO**

Department of Geography,  
Graduate School of Urban Environmental Sciences,  
Tokyo Metropolitan University

March 2022

## **Abstract**

The purpose of this study is to examine the spatial resolution and discrimination accuracy in capturing vegetation using new data and methods such as 360-degree images, ultra-high-resolution images by UAVs (Unmanned Aerial Vehicles), and object-based image analysis, aiming at the social implementation of remote sensing methods that are more efficient and less resource-consuming in urban and watershed planning.

Understanding vegetation is an important issue. In Japan, it is necessary to understand vegetation on a regional scale to develop urban and watershed plans that consider the water cycle function of vegetation. Meanwhile, regarding remote sensing technology for vegetation, the development of UAVs and the improvement of sensor performance have led to extremely high data resolution and an explosive increase in data volume. High-resolution data, however, have the problem of increased cost in terms of time and resources for processing and storage, as well as the problem of reduced accuracy due to the fragmented extraction of objects. Thus, it is important to select an appropriate resolution and analysis method. However, there are no studies that have examined the possibility of using 360-degree images, ultra-high-resolution images by UAVs, and object-based methods in terms of appropriate resolutions and analysis methods for capturing vegetation on a regional scale.

Therefore, Chapter 1 reviewed the previous studies and discussed the above issues. Chapter 2 studied object-based analysis methods that are considered effective for high-resolution images. Chapter 3 examined a method to understand vegetation using a mobile mapping system (MMS) to obtain 360-degree images. Chapter 4 analyzed the relationship between spatial resolution and discrimination accuracy of vegetation. Finally, Chapter 5

discussed the data and analysis methods for understanding vegetation for specific issues related to urban and watershed planning, based on the results of Chapters 2 to 4.

Chapter 2 compared pixel-based and object-based methods for identifying green cover in urban areas, using high-resolution aerial photographs, which are used for municipal vegetation surveys. The study area was Setagaya Ward in Tokyo Metropolis, which has a variety of land covers, and the land cover classification was performed to identify vegetation on a 1 m<sup>2</sup> scale. The results showed that the overall accuracy of the object-based classification was about 6 to 20 percentage points higher than the pixel-based classification. Both methods tended to over-classify bare land and water areas. Shadows of buildings and roads of impervious areas tended to be misclassified as the bare land and the water area in the case. In the pixel-based method, most of the classified areas were small areas of one to several pixels. The join-count statistic, a spatial autocorrelation index, was calculated to evaluate minute areas (salt-and-pepper effect) in the classification. As a result, the values of object-based classification showed a stronger tendency of grouping than the pixel-based classification. Namely, the results quantitatively showed that the object-based classification has fewer minute areas than the pixel-based classification. Furthermore, the green coverage ratio (ratio of area covered by vegetation to the region's total area) was calculated from the classification results and was compared with the survey by Setagaya Ward. The results showed that the object-based classification was closer to the survey by the municipality than the pixel-based classification. These results indicate that the object-based method is effective for capturing vegetation in urban areas using high-resolution digital aerial photographs.

Chapter 3 examined a method for understanding street trees in dense urban areas using 360-degree images obtained from MMS (a vehicle-mounted system for obtaining

geographic information), 3D reconstruction techniques, and object-based analysis. The study area was Suginami Ward in Tokyo Metropolis, a dense urban area with a variety of street trees. The data were obtained while driving on the road by the MMS (IP-S2 Lite, Topcon Corporation). The data were then processed to calculate the 3D position information from the images. Tree height and DBH (Diameter at Breast Height) were measured for 107 trees from the processed images. Measurements using a portable laser rangefinder and a caliper gauge were used as verification data. In addition, green areas in the 360-degree images were extracted using object-based classification, and the green visibility ratio (the ratio of the area of green within the elevational field of view of an image) was calculated. Comparing the measurements, the RMSE (Root Mean Square Error) of the tree height and DBH were 1.07 m and 0.033 m, respectively. The result suggested that the MMS without a laser scanner may be able to measure trees with practical accuracy, considering the measurement accuracy of field surveys. Comparing the tree height and DBH by size, the smaller the tree height, the smaller the error. On the other hand, a tendency of underestimation was found when the tree height was large, and the canopy covered the sky. As for the green visibility ratio, the values calculated from the MMS images were higher at points with many trees with large canopies and lower at points with small canopies and small tree heights. In addition, the green visibility ratio calculated from MMS images showed a positive correlation with tree height, DBH, and NDVI (Normalized Difference Vegetation Index) and the green coverage ratio calculated from aerial photographs. In other words, the index proposed in this study can be used as an indicator of the amount of greenery, suggesting the possibility of estimating the approximate tree height and DBH. These results show that the height and DBH of street trees can be measured by analyzing the 360-degree images of MMS. The results also

show that the overview of greenery can be understood by extracting green areas through object-based analysis.

Chapter 4 analyzed the spatial resolution suitable for capturing forests using UAVs. A cedar forest and a larch forest (at the foot of Mt. Aso in Kumamoto Prefecture and at the foot of Mt. Yatsugatake in Yamanashi Prefecture, respectively), which are important forests for rainwater recharge in watershed planning, were selected as the study areas. Various resolutions ranging from 0.9–100 cm/pixel of DSM (Digital Surface Model) and ortho-mosaic images were generated from stereo pair images taken by UAVs. The results were compared from three perspectives: extraction of the tree crowns, supervised classification, and spatial autocorrelation of images. In the extraction of the tree crowns using DSM, the results resized to a resolution of 5–10 cm/pixel were closest to the reference data. This suggests that the resolution corresponded to the resolution at which the effect of small unevenness in the crown of a single tree, such as leaves, was removed. For supervised classification, the results resized to 10–80 cm/pixel had higher classification accuracy than the original resolution results. The resolutions are thought to correspond to the resolution at which the shadows between trees were pure pixels. In addition, the processing time required for the 3D reconstruction process was reduced to less than 1/20 to 1/300 of the original resolution at 10 cm/pixel. The variogram, a spatial autocorrelation index, was obtained from the images, and the typical indices of nugget, range, and sill were found. The nugget could represent small variations in the observed values, the range could represent the crown size, and the sill could represent the brightness difference between the crown and the gap. The sill is the maximum value of the brightness difference between the crown and the gap, suggesting that the resolution at which each tree crown and gap is a pure pixel is important for tree crown discrimination. These results

suggest that a resolution of 5–10 cm/pixel is sufficient for identifying single trees, and a resolution of 10–50 cm/pixel is sufficient for identifying the approximate structure of a forest.

In conclusion, the results of Chapters 2 through 4 can be expected to be helpful in urban and watershed planning efficiently, such as identifying small-scale vegetation using object-based analysis (Chapter 2), capturing vegetation in dense urban areas through MMS and 360-degree images (Chapter 3), and reducing processing time and costs through efficient operation of UAVs and the use of aerial photographs (Chapter 4). In the future, it is necessary to apply and verify the methods proposed in this study to various types of vegetation, data, and analysis methods toward the social implementation.

## 要旨

論文題目：

都市および流域の植生把握に適した  
観測および解析手法に関する地理情報学的研究  
ーUAVによる超高解像度画像や360度画像および  
オブジェクトベース解析手法の可能性ー

氏名： 山本 遼介

本研究は、都市・流域計画における、より効率的かつ環境負荷の小さい植生リモートセンシング手法の社会実装に向け、UAV による超高解像度画像や 360 度画像およびオブジェクトベース解析などの新しいデータと解析手法を用いた植生の把握における、空間解像度と植生の判別精度を明らかにすることを目的とした。

植生の把握は重要な課題であり、日本においては、植生の水循環機能を考慮した都市計画および流域計画の策定のため、地域スケールの植生把握が必要とされている。一方で植生のリモートセンシング技術に関しては、UAV の発展やセンサーの性能向上などにより、データの解像度が非常に高くなり、データ量が爆発的に増加している。高解像度画像は、対象物の断片的な抽出（微小領域）による精度低下や、処理や保存に要する資源やコストの増大という問題を持つため、過剰に高い解像度や大量のデータを取得するのではなく、適切な解像度および解析手法の選択が重要である。しかしながら、地域スケールの植生把握に対し、適切な解像度および解析手法の観点から、UAV、360 度画像およびオブジェクトベース手法の活用可能性を検討した事例は見られない。

本研究は以上のような問題点を第1章で指摘した上で、高解像度画像の解析に有効とされるオブジェクトベース手法の検証（第2章）、車載型の地理情報取得システム（MMS）を活用した植生把握の検討（第3章）、空間解像度と植生判別精度の関係の分析（第4章）を行った。

第2章では、都市域における緑被の把握手法として、自治体の植生調査などで利用される高解像度の航空写真を用いて、ピクセルベース手法とオブジェクトベース手法を比較した。多様な土地被覆を持つ東京都世田谷区を対象地とし、1 m<sup>2</sup>規模の植生把握のための土地被覆分類を行った。その結果、オブジェクトベース分類の全体精度は、ピクセルベース分類よりも約6~20ポイント（%）高かった。また、空間的自己相関指標のJoin統計量を計算した結果、オブジェクトベース分類はより強い集積傾向を示し、微小領域の発生が少ないことが定量的に示された。さらに、分類結果から算出した緑被率は、オブジェクトベース手法の方が自治体による調査に近い値が得られた。以上の結果から、高解像度の航空写真を用いて1 m<sup>2</sup>規模の植生把握を行う場合に、オブジェクトベース手法は有効であることが明らかになった。

第3章では、密集市街地における樹木の把握手法として、MMSによる360度画像、3次元復元技術、およびオブジェクトベース解析の活用による街路樹の把握を検討した。密集市街地であり多様な街路樹を持つ東京都杉並区を対象地とし、MMS（IP-S2 Lite、株式会社トプコン）により道路上を走行しデータを取得した。その後、画像から3次元位置情報を計算し、107本の街路樹を対象として、樹高および胸高直径を測定した。また、オブジェクトベース解析により360度画像の緑視率（画像内に占める緑地の割合）を計算した。樹高・胸高直径を実踏調査と比較した結果、平方根平均2乗誤差は、樹高が1.07m、胸高直径が0.033mであった。樹高は、樹冠が天空を覆っている場合には過小評価の傾向が見られることが分かった。360度画像から算出した緑視率は、樹高、胸高直径、航空写真による正規化植生指数および緑被率との間に正



の相関が見られた。すなわち、本研究で提案した指標は緑の量を表す指標として利用でき、およその樹高・胸高直径を推定できる可能性が示唆された。

第4章では、樹木の把握における画像の空間解像度と判別精度の関係を分析した。雨水涵養林として流域計画において重要な森林であるスギ林（熊本県阿蘇）およびカラマツ林（山梨県八ヶ岳）を対象に、UAVによって撮影したステレオペア画像から、解像度 0.9~100cm/pixel の DSM およびオルソモザイク画像を生成し、樹冠の抽出、教師付き分類、画像の空間的自己相関の3つの観点から結果を比較した。樹冠の抽出では、5~10cm/pixel の結果が検証データに最も近かった。これは、枝や葉など樹冠内の微小な凹凸の影響が除かれる解像度に対応していた。教師付き分類では、10~80cm/pixel の結果が、元の解像度の結果より分類精度が高かった。これは、樹木間の影がピュアピクセルとなる解像度に対応していると考えられた。3次元復元処理に要する処理時間は、10cm/pixel で元の解像度の 1/20~1/300 未満に減少した。画像から空間的自己相関指標のバリオグラムを求めた結果、ナゲット、レンジ、シルという典型的な指標が得られた。ナゲットは観測値の微小な変動、レンジは樹冠サイズ、シルは樹冠とギャップの明度差を代表する指標となり得ることが示唆された。シルは樹冠とギャップによる明度差の最大値であり、それぞれがピュアピクセルとなる解像度が樹冠の判別に重要であることが示唆された。以上の結果から、樹木の把握には 5~10cm/pixel、森林のおよその構造の把握には 10~50cm/pixel 程度の解像度であれば、十分な判別精度で植生が把握できる可能性が示された。

以上の研究成果は、都市・流域計画において、小規模な植生の把握（第2章）、密集市街地の植生把握（第3章）、UAVの効率的な運用や航空写真の活用による処理時間やコストの低減（第4章）など、植生の効率的な把握に役立つと期待される。社会実装に向けて、様々な種類の植生やデータおよび解析手法に対して、本研究の手法の適用と検証が求められる。