

UAV downwash dynamic texture features for terrain classification on autonomous navigation

J. P. Matos-Carvalho, José M. Fonseca, André Mora

Computational Intelligence Group of CTS/UNINOVA, FCT, University NOVA of Lisbon

Email: jp.carvalho@uninova.pt, jmf@uninova.pt, atm@uninova.pt

Abstract—The information generated by a computer vision system capable of labelling a land surface as water, vegetation, soil or other type, can be used for mapping and decision making. For example, an unmanned aerial vehicle (UAV) can use it to find a suitable landing position or to cooperate with other robots to navigate across an unknown region. Previous works on terrain classification from RGB images taken onboard of UAVs shown that only static pixel-based features were tested with a considerable classification error. This paper proposes a robust and efficient computer vision algorithm capable of classifying the terrain from RGB images with improved accuracy. The algorithm complement the static image features with dynamic texture patterns produced by UAVs rotors downwash effect (visible at lower altitudes) and machine learning methods to classify the underlying terrain. The system is validated using videos acquired onboard of a UAV.

Keywords—Image processing, Texture, Machine Learning, Terrain Classification, UAV

I. INTRODUCTION

Nowadays, due to UAVs' higher availability and capabilities, there is a research trend to explore innovative applications of UAVs useful to the society. They are having a major impact on search and rescue missions, in logistics, in precision agriculture, among other applications. Key issues are to provide a safe and reliable operation and to perceptionate the surrounding area. This latter, within this paper, will be to identify the underlying terrain. Terrain classification is a crucial functionality for a wide range of autonomous vehicles [1]: either for ground vehicles to avoid water bodies, aerial vehicles to determine suitable landing areas, or surface vehicles to detect safe passageways. As further explained in section II, several approaches have been used for terrain classification. However, there is still margin for improving accuracy by extracting more complex image features. When at lower altitudes, UAV's rotors downwash effect create singular image texture patterns depending on the type of terrain, which can be used to differentiate them.

The main goal of this paper is to propose a computer vision algorithm that using RGB images captured by a camera onboard of a UAV is capable of classifying a terrain by analysing static image features (colour and texture) and rotors downwash effect on the underlying surface. There are several issues that must be addressed in order to achieve this goal, namely: Which terrains can be more accurately classified using the downwash effect? Which are the texture and motion patterns of each terrain (water movement for example)? Which static and dynamic image features can be extracted to classify the terrain? To address these challenges, new optimization

procedures and techniques will be proposed in this paper, aiming the best possible performance.

This paper is structured with six sections starting with an introductory section and followed by a presentation of related works. In the experimental setup section the system background, namely the hardware and the terrain types, are described. On the Terrain Classification Method the system architecture, the static and dynamic texture features and the machine learning classifier will be presented. The article finishes with the experimental results and drawn conclusions.

II. RELATED WORK

UAVs (Unmanned Aerial Vehicle) play an important role on the new generation of information technology and is predicted to have a major impact in the human life in the near future [2]. One of the areas is in computer vision, where it is possible to acquire, process, analyse and understand aerial images. Many researchers have proposed terrain classification systems based on features derived from colour information [3], texture patterns [4], [5], [6] and from additional sensors, as is the case of laser scan systems [7], [8], [9]. Although many of these algorithms are for terrestrial unmanned ground vehicles, currently there is a shift towards UAVs, where the visual features have wider importance.

One of the most recent works of terrain detection and classification is presented in [10]. The authors use the concept of optical flow to detect the water texture direction in images acquired by an RGB camera onboard of a UAV. From the directions of the textural features, the algorithm determines if the terrain, where the UAV is flying over, is water or non-water. One of the problems identified is that the UAV must be stable over the target while identifying the type of terrain, which, in the best case, takes four seconds to execute. Another reason that requires the UAV to be stand still during calculations is that the computer vision algorithm does not compensate the UAV movement. Thus, when the directions of the features are calculated, the results do not represent the reality.

A classification method using colour features was proposed in [3]. The proposed method converts a RGB image into an image entitled "normal RGB", where each pixel is divided by the square root of the three colour channels. Thus, each terrain will emphasize the colour that represents it (for example, green for vegetation). The proposed method was limited due to the fact that it varies significantly with illumination.

Laser scanners have proven to be important to distinguishing between land and water as presented in [7], [8] and [9].

However, in low water depths the laser sensor produces incorrect results, due to the fact that it captures reflections from the seabed and misclassifies it as non-water terrain. Therefore, this laser scan approach, by itself, reveals to be insufficient and requires additional equipment.

III. EXPERIMENTAL SETUP

The dynamic of different terrains when exposed to wind provoke singular texture patterns that can be used in their identification. In this paper we study the importance of static image features, such as colour and texture, when compared with the dynamic features exhibited by the downwash effect, for terrain classification.

In this work three different terrain types (water, vegetation and sand), which can benefit from the downwash effect for their identification (Figure 1) were identified. It can be seen that the downwash effect produces: on water a circular dynamic texture; on vegetation a linear spread from inside outwards; and on sand it is almost stable or it moves outwards.

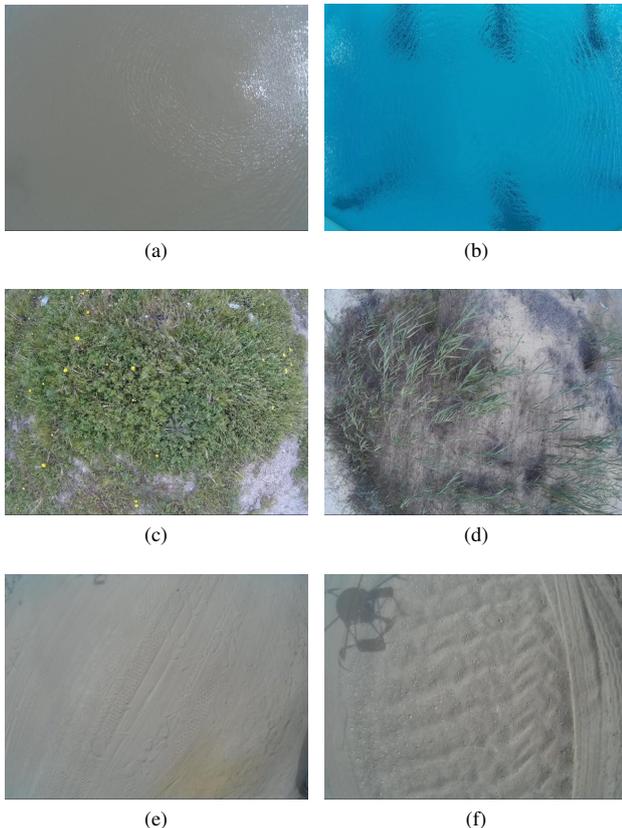


Fig. 1. Examples of terrain types: water (a)(b); vegetation (c)(d); and sand (e)(f).

IV. TERRAIN CLASSIFICATION METHOD

If different types of terrain behave differently when exposed to UAV rotors downwash effect, then it should be possible to obtain unique information for their identification. Based in this research hypothesis, it is possible to obtain some conclusions. When exposed to the downwash effect, water

particles' movement is always greater than in vegetation and sand terrains. Also, regarding static texture, usually vegetation has a more rough texture than sand or water terrains; water only presents roughness when exposed to wind and downwash effect; and sand (fine grains) has a lower roughness. It can also be seen that sand depends on the patterns already in the terrain, showing usually a more irregular texture (figures 1.e and 1.f) when compared with water that shows a unique signature and regular texture when exposed to wind (figures 1.a and 1.b).

A. System Architecture

The proposed system architecture to classify the terrain using texture information is shown in Figure 2. As previously identified in sections I and IV, two texture features are proposed to classify the terrain, namely, static and dynamic textures. At this stage it were also assessed the features that can be computed in parallel, in order to speedup the system execution time.

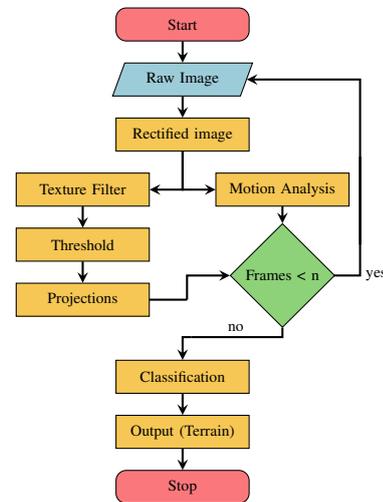


Fig. 2. Proposed system architecture.

Five main processes were identified in the architecture (figure 2), namely:

- **Rectified Image:** Performs lens geometrical corrections;
- **Texture Filter:** Extracts terrain's static textural information using Gabor filters;
- **Threshold:** A thresholding is applied to the static texture image to highlight the terrain roughness;
- **Projections:** Vertical and horizontal projections were applied to the thresholded image, extracting unique features that help differentiate the different types of terrains;
- **Motion Analysis:** Extracts information from dynamic textures. Optical flow and thresholding techniques are used to identify the moving parts;
- **Classification:** The extracted features are used as inputs of an automatic classified to identify the type of terrain. Machine learning techniques already proved to be efficient for terrain classification [11], [12], [13].

B. Static Textures

This section presents the proposed method for extracting terrain's static textures, based on the Gabor filter to be able to choose multiple texture directions. This filter is the impulse response formed by a multiplication of a sinusoidal signal with a Gaussian envelope function and can be computed using the following complex equation:

$$G(x, y, \lambda, \theta, \psi, \sigma, \gamma) = e^{\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} e^{i\left(2\pi\frac{x'}{\lambda} + \psi\right)} \quad (1)$$

Its real and an imaginary components can be obtained by equations 2 and 3, respectively:

$$G(x, y, \lambda, \theta, \psi, \sigma, \gamma) = e^{\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} \cos\left(2\pi\frac{x'}{\lambda} + \psi\right) \quad (2)$$

$$G(x, y, \lambda, \theta, \psi, \sigma, \gamma) = e^{\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} \sin\left(2\pi\frac{x'}{\lambda} + \psi\right) \quad (3)$$

where:

$$x' = x \cos(\theta) + y \sin(\theta) \quad (4)$$

$$y' = -x \sin(\theta) + y \cos(\theta) \quad (5)$$

These equations (1, 2 and 3) require as input parameters:

- **x and y:** Filter coordinates, where x represents the columns and y the rows;
- **Lambda (λ):** Represents the sinusoid's wavelength;
- **Theta (θ):** Defines the Gaussian envelope orientation;
- **Psi (ψ):** Symbolizes the phase offset;
- **Sigma (σ):** Describes the Gaussian envelope size;
- **Gamma (γ):** Reflects the shape of the ellipse in the gabor filter space.

In this work we used only the real component of the Gabor function (equation 2). After obtaining the multiplication of a Gaussian with a sinusoidal function, i.e. the kernel of the filter, it will be convolved with the original image (equation 6). The result of the Gabor filter applied over a water surface is presented in Figure 3.

$$f[x, y] * g[x, y] = \sum_{-n1}^{n1} \sum_{-n2}^{n2} f[n1, n2] \cdot g[x - n1, y - n2] \quad (6)$$

As can be seen in figure 3, it is possible to obtain the texture of a water-type terrain when it is affected by the downwash effect of the UAV. From the binarized image, a vertical projection was made to see the singular features of this terrain type (Figure 4).

From the observed vertical projection of water type terrain (figure 4) it can be seen that it produces an undulatory effect with a local minimum in the centre of the downwash. This effect in water type terrains is due to the lower roughness

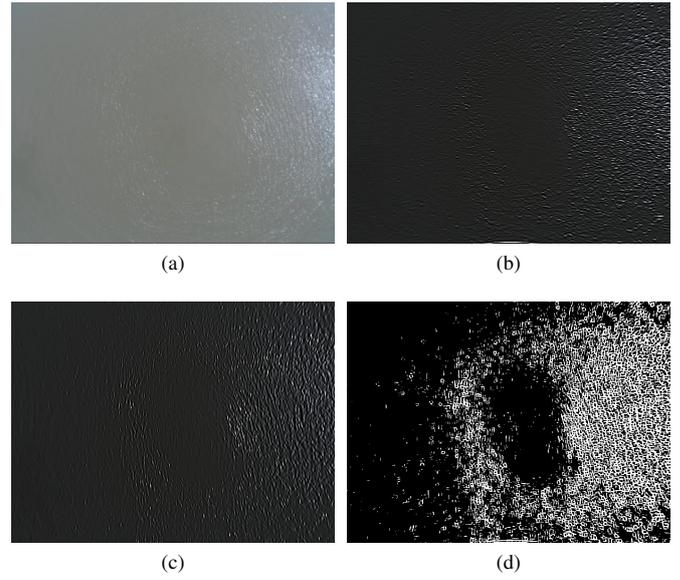


Fig. 3. Example of a static texture extraction: a) Raw image; b) c) Convolution with the Gabor filter $\theta=0$ degrees (b) and $\theta=90$ degrees (c); d) Sum of images b) and c) after thresholding.

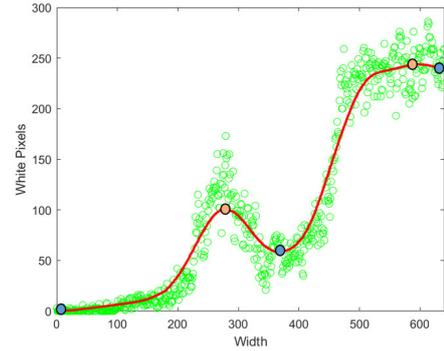


Fig. 4. Vertical projection of the example in Figure 3.d.

in the centre of the downwash. However, due to the water movement, around the centre a higher roughness is observed (white pixels in the binarized image in figure 3.d). The next step was to translate this observed feature into a computational model.

By calculating the local maxima and minima of the vertical projection in figure 3.d, it is possible to calculate a line (red line in Figure 4) that most closely approximates these points. A polynomial regression was used.

After obtaining this smoothed projection, new local minima and maxima are calculated and used to obtain two features: Area measured between the local minimum and its respective two local maxima; and Integral between the local minimum and its two respective local maxima. The first has the advantage of being relative to minima and maxima values, while the integral gives an absolute value and will vary for lower and higher roughness.

C. Dynamic Textures

This section presents the proposed method for extracting dynamic terrain textures.

As mentioned in section IV, water-type terrain only exhibit dynamic texture when exposed to the downwash effect. However, in spite having a dynamic texture, when analysing the optical flow it is never stronger than the dynamic observed for sand and vegetation. As referred in section III, the optical flow method can calculate the distance travelled by block matching features in a given frame sequence. In this paper, the Farneback algorithm [14] was used to detect the movement of these features. One of the advantages to using the Farneback algorithm is the direct flow, F_d , return of features between two frames.

With the obtained flow is then used to calculate the distance travelled (trajectory) by each feature in a sequence of frames:

$$Travel_{distance} = \sum_{i=2}^n \sqrt{A_x(i) + B_y(i)} \quad (7)$$

where:

$$A_x(i) = [x_1 - x_{i-1} + F_{dx}]^2 \quad (8)$$

$$B_y(i) = [y_1 - y_{i-1} + F_{dy}]^2 \quad (9)$$

and x_i and y_i are the positions in x and y in the most recent frame (n), x_1 and y_1 are the initial positions ($n = 1$) and F_{dx} and F_{dy} are the flow displacements between frames n and $n - 1$. We used normalized x and y coordinates for the calculations.

To eliminate features that did not move or were almost static in a sequence of frames, we filtered those not exceeding a pre-defined empiric threshold (1%). Then, knowing the maximum number of features, we calculate the percentage of dynamic features that appear in the image (equation 10). An example is shown in Figure 5.

$$Dynamic_{feature} = \frac{filtered\ features}{Total\ features} \cdot 100\% \quad (10)$$

D. Classification

To increase certainty and automate the classification of the type of terrain, a machine learning technique was used, namely a feed-forward neural network (NN). The architecture of the designed neural network, was composed by two layers, a hidden layer with 10 neurons and an output layer with 3 neurons (water, vegetation and sand). A sigmoidal function was used as activation function and the final classification was derived from the output neuron with highest activation value.

The training dataset was composed by 251 samples, from which 70% were for training, 15% for testing and 15% for validation. After training the NN, it was obtained 92.9% accuracy with the training set and 93.8% with the test dataset.

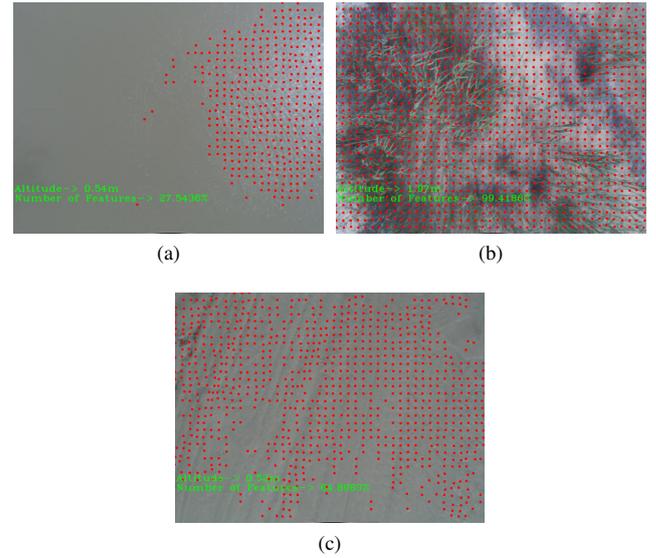


Fig. 5. Dynamic textures detection by Farneback algorithm and distance travelled calculation. a) water; c) vegetation and d) sand.

V. EXPERIMENTAL RESULTS

To validate the proposed static and dynamic texture features for terrain classification, a total 251 frames from several types of terrains were used to validate the proposed system. From these 90 frames were for water, 88 frames for vegetation and 73 frames for sand.

Regarding the static texture feature the area and integral were calculated and displayed in Figure 6. It is possible to observe a clear separation between water, vegetation and sand, even with some outliers. In water type terrain, the three clusters that can be noticed for the integral feature, were mainly due to different water environments (lake and pool).

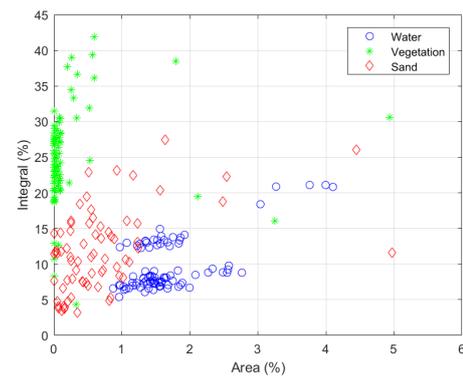


Fig. 6. Static Texture - Relation in area with respect to the integral of minimum and maxima locals.

To validate the dynamic texture feature it was calculated in a three frame period ($n = 3$) and plotted against the area feature from the static texture. This feature shows the same discriminant level to separate the different terrains. From Figure 7, it can be seen that water type terrain obtained a lower dynamic texture value ($< 45\%$), which can due to a higher

concentration of these dynamic features in the downwash centre and outside hasn't exceed the threshold. Sand and vegetation shown a more uniform pattern, obtaining a higher number of features. On average, sand presents a percentage between 55% to 90%. Finally, vegetation with a percentage of features between 90% and 100%, is the terrain with highest dynamic texture, i.e., moving features.

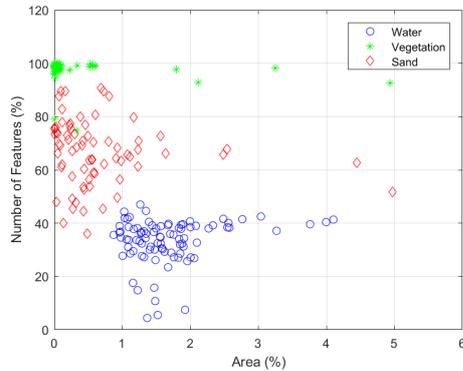


Fig. 7. Dynamic Texture - Relation in number of features with respect to the integral of minimum and maxima locals.

Finally, these features were extracted from the figures 1.a-f and shown to the neural network classifier, which outputted the automated terrain classification. The extracted features and the classification result is shown in Table I. As expected the proposed features and classification method, shown good results by classifying correctly all the six examples, reinforcing the idea that a combination of static and dynamic texture can be used to automatically extract terrain type from RGB images.

TABLE I. EXPERIMENTAL RESULTS

| Figure | Static Texture | | Dynamic Texture | Classification |
|--------|----------------|--------------|------------------------|----------------|
| | Area (%) | Integral (%) | Number of Features (%) | |
| 1 | 1.44 | 8.32 | 32.88 | water |
| 2 | 1.55 | 7.59 | 36.70 | water |
| 3 | 0.01 | 27.42 | 98.32 | vegetation |
| 4 | 0.05 | 23.71 | 98.69 | vegetation |
| 5 | 0.24 | 14.67 | 63.74 | sand |
| 6 | 0.07 | 4.07 | 59.30 | sand |

VI. CONCLUSIONS

The main objective of this paper was to design a computer vision system capable of extracting static and dynamic image features, such as optical flow and texture features, to identify the type of terrain with improved accuracy by taking advantage of the the UAV's rotors downwash pattern effect. For this, it was necessary to conduct a research into detection methods already implemented and of interest to this work.

Texture features, such as Gabor filtering (static textures) and optical flow (dynamic textures), were studied to improve terrain classification aiming the best possible performance.

We emphasize that by implementing the static textures filter, vegetation-like terrains were found to have a higher texture than sand and water type terrains. On the other hand, water-type terrain, also presents a singular characteristic due to the downwash effect provoked by the UAV, which can be decisive to different it from other terrain types.

ACKNOWLEDGMENT

This work was partially funded by FCT Strategic Program UID/EEA/00066/203 of the Center of Technologies and System (CTS) of UNINOVA – Institute for the Development of new Technologies. In last, this work was not possible without the support and commitment of several fellow colleagues and friends, namely: Ricardo Mendonça, Francisco Marques, André Lourenço, Eduardo Pinto and José Barata.

REFERENCES

- [1] Pinto, E., Marques, F., Mendonca, R., Lourenco, A., Santana, P., & Barata, J. (2014). An autonomous surface-aerial marsupial robotic team for riverine environmental monitoring: Benefiting from coordinated aerial, underwater, and surface level perception. 2014 IEEE International Conference on Robotics and Biomimetics, IEEE ROBOT 2014, 443–450. <https://doi.org/10.1109/ROBOT.2014.7090371>
- [2] Bestaoui Sebbane, Y. (2018). Intelligent Autonomy of UAVs : Advanced Missions and Future Use. CRC Press.
- [3] Ebadi, F., & Norouzi, M. (2017). Road Terrain detection and Classification algorithm based on the Color Feature extraction. In 2017 Artificial Intelligence and Robotics (IRANOPEN) (pp. 139–146). IEEE. <https://doi.org/10.1109/RIOS.2017.7956457>
- [4] Feng, Q., Liu, J., & Gong, J. (2015). UAV Remote sensing for urban vegetation mapping using random forest and texture analysis. Remote Sensing, 7(1), 1074–1094. <https://doi.org/10.3390/rs70101074>
- [5] Khan, Y. N., Komma, P., Bohlmann, K., & Zell, A. (2011). Grid-based visual terrain classification for outdoor robots using local features. IEEE SSCI 2011: Symposium Series on Computational Intelligence - CIVTS 2011: 2011 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems, 16–22. <https://doi.org/10.1109/CIVTS.2011.5949534>
- [6] Pietikäinen, Matti, Abdenour Hadid, Guoying Zhao, and Timo Ahoenen. 2011. Computer Vision Using Local Binary Patterns. Vol. 40. Computational Imaging and Vision. London: Springer London. <https://doi.org/10.1007/978-0-85729-748-8>.
- [7] Yan, W. Y., Shaker, A., & El-Ashmary, N. (2015). Urban land cover classification using airborne LiDAR data: A review. Remote Sensing of Environment, 158, 295–310. <https://doi.org/10.1016/j.rse.2014.11.001>
- [8] Wallace, L., Lucieer, A., Malenovsky, Z., Turner, D., & Vopěnka, P. (2016). Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and structure from motion (SfM) point clouds. Forests, 7(3), 1–16. <https://doi.org/10.3390/f7030062>
- [9] Gruszczynski, Wojciech, Matwij, Wojciech, Cwiakała, P. (2017). Comparison of low-altitude UAV photogrammetry with terrestrial laser scanning as data-source methods for terrain covered in low vegetation. ISPRS Journal of Photogrammetry and Remote Sensing, 126, 168–179. <https://doi.org/10.1016/j.isprsjprs.2017.02.015>
- [10] Pombeiro, R., Mendonca, R., Rodrigues, P., Marques, F., Lourenco, A., Pinto, E., Barata, J. (2015). Water detection from downwash-induced optical flow for a multirotor UAV. In OCEANS 2015 - MTS/IEEE Washington (pp. 1–6). IEEE. <https://doi.org/10.23919/OCEANS.2015.7404458>
- [11] Mora, A.; Santos, T.M.A.; Łukasik, S.; Silva, J.M.N.; Falcão, A.J.; Fonseca, J.M.; Ribeiro, R.A. (2017) Land Cover Classification from Multispectral Data Using Computational Intelligence Tools: A Comparative Study. Information 2017, 8, 147. <https://doi.org/10.3390/info8040147>
- [12] Heung, Brandon, Hung Chak Ho, Jin Zhang, Anders Knudby, Chuck E. Bulmer, and Margaret G. Schmidt. 2016. "An Overview and Comparison of Machine-Learning Techniques for Classification Purposes in Digital Soil Mapping." Geoderma 265 (March): 62–77. <https://doi.org/10.1016/j.geoderma.2015.11.014>
- [13] Giusti, Alessandro, Jerome Guzzi, Dan C. Ciresan, Fang-Lin He, Juan P. Rodriguez, Flavio Fontana, Matthias Faessler, et al. 2016. "A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots." IEEE Robotics and Automation Letters 1 (2): 661–67. <https://doi.org/10.1109/LRA.2015.2509024>.
- [14] Farneback, G. (2003). Two-Frame Motion Estimation Based on Polynomial Expansion. Lecture Notes in Computer Science, 2749(1), 363–370. https://doi.org/10.1007/3-540-45103-X_50