

# NDVI point cloud generator tool using low-cost RGB-D sensor

D. Calero, E. Fernández, M.E. Parés and E. Angelats

**Abstract**— In this manuscript, a NDVI point cloud generator tool based on low-cost active RGB-D sensor is presented. Taking advantage of currently available ROS point cloud generation tools and RGB-D sensor technology (like Microsoft Kinect), that includes an inbuilt active IR camera and a RGB camera, 3D NDVI maps can be quickly and easily generated for vegetation monitoring purposes. When using low-cost sensors for vegetation index estimation, it is necessary to apply a rigorous methodology for extracting reliable information. In this paper, the methodology for NDVI generation using a low-cost sensor as well as experiments to evaluate its performance is presented. The experiments performed show that it is possible to obtain a reliable NDVI point cloud from a Kinect V2.

## I. INTRODUCTION

Monitoring vegetation and extracting vegetation indicators is a common technique for evaluating the vegetation health during its life cycle. These indicators are important for deciding on how to proceed with the plants management goals. In the past, the indicators were derived from site-based measurements but nowadays remote sensing tools are commonly used. These tools involve image acquisition and processing techniques [1].

Light waves can be classified into various frequency bands, in function of their wavelength. When a source of light collides with a material, these frequency bands respond differently depending on properties of the material, such as its reflectivity and absorption. Thus, by measuring the intensity of reflection of each frequency light band, it is possible to characterize the material and its properties. Typical frequency bands used are ultraviolet region (UV) (10 to 380 nm); visible spectra (450 to 750 nm) and Near Infrared Band (NIR) (750 to 1400 nm). Most optical sensors divide the visible spectra in RGB format, which is composed of blue (450 to 495 nm), green (495 to 570 nm), and red (620 to 750 nm) wavelength regions.

The reflectance of light spectra from plants depends on plant species, water content within tissues, and other intrinsic factors [2]. The spectral reflectance data is inversely related to leaf chlorophyll level, which can facilitate to investigate the health state of plants. Thus, it is possible to evaluate quantitative and qualitative parameters of vegetation cover, vigor, and growth dynamics, among other properties by using algorithms involving spectral reflectance [3].

Vegetation Indices (VI) are relations between different light frequency bands and play an important role in monitoring variations in vegetation. Many of the indices make use of the inverse relationship between red and NIR reflectance associated with healthy green vegetation [4]. Among a huge

list of multispectral vegetation indices, such as Difference Vegetation Index (DVI), Simple Ratio Index (SRI), the most common index is the Normalized Difference Vegetation Index (NDVI) [5]. NDVI is derived from remote sensing measurements and determines whether the target being observed contains green vegetation or not. The analysis of NDVI image allows to estimate several biophysical parameters, including chlorophyll density, canopy cover and photosynthetically absorbed active radiations [6] among others.

NDVI images can be collected from different platforms, depending on the use case application. Satellite platforms provide NDVI images with high coverage area, but at low spatial resolution, while their typical ground sample distance (GSD) is in the range of meters per pixel (px), which is not enough for many applications [7]. The use of aircraft or helicopter platforms can increase the GSD and can support heavy payload for the sensing system. They are able to provide high coverage areas and high resolution, but their main drawback is complexity involved in acquisition process and their implementation cost. Another commonly used platforms are Unmanned Aerial Vehicles (UAVs), which can work with multispectral cameras such as Parrot Sequoia multispectral sensor [8] or custom alternatives that combines a NIR camera and a red only or RGB camera. UAVs can provide GSD in the range of cm per px and can cover large areas. The main drawback is the cost, which is in the order of thousands of dollars. Another issue with them is autonomy due to a limited flight of time operation and flying zone restrictions imposed by Government authorities. A non-flying alternative is to use autonomous wheeled robot platforms, that can move through outdoor and indoor environments and can access to smaller places for wide inspection of entire plantation (not only the top vegetation). It is a low-cost solution, that can achieve high-resolution data, cm per px GSD, but within a reduced area.

VI are graphical indicators, which are normally represented in 2D images. Combining VI images with depth information from laser scanner, stereo cameras or RGB-D cameras, it is possible to generate a three-dimensional (3D) point cloud image representation, that provides a capability to visualize the information from different perspectives. An advantage of this representation of plant vegetation is their 3D reconstruction and the ability to provide information about their volumes, which can lead to an easy segmentation and classification of the information.

Nowadays, in the robotics field it is very common to use low-cost RGB-D sensors for 3D mapping applications and for indoor navigation using SLAM techniques [9]. In addition, some recent studies confirm that, it is possible to generate

NDVI images using low-cost RGB-D cameras, like Microsoft Kinect V2 [10]. It has some limitations such as over exposure of light in outdoor environments due to saturation of IR light from the sun radiation.

In this paper, a tool for generating reliable NDVI point cloud using a low cost RGB-D sensor based on an active IR technology is presented. The development of the tool relies on a custom autonomous wheeled robot platform which has incorporated a Microsoft Kinect v2 sensor, but thanks to ROS modularity it can be easily integrated in any other compatible environment. To evaluate its performance, a set of tests are carried out in two different scenarios, an indoor environment and inside a greenhouse for cultivation. The tool can be useful, for example, for monitoring plant health condition in green houses or smart farming applications.

## II. MATERIALS

This section describes the equipment used to develop and test the NDVI point cloud generation tool. Nevertheless, this is a generic approach and it can be used with similar RGB-D sensors based on IR technology.

### A. Sensor - Microsoft Kinect V2

The Microsoft Kinect V2 is a low-cost RGB-D sensor equipped with RGB, NIR camera and an IR emitter [11]. The RGB camera has 1920x1080 px resolution, while the IR camera has 512x424 px resolution. The NIR spectrum covers from 827 to 850nm. The in-built IR emitter produces modulated light pulses at 830nm [12].

The IR camera is used not only to acquire IR images, but also to collect depth images. This is done using Time-Of-Flight (TOF) technology. By measuring the time of flight of the emitted/reflected signal between the camera and the target, it is possible to estimate the object distance.

Kinect V2 can provide up to 30Hz image data, including the processed depth images.

### B. Platform - Autonomous wheeled robot platform

A four wheeled robot platform testbed (Fig. 1) is used for the tests validation [13].



Figure 1. Wheeled robot platform testbed

The robot platform includes an IMU sensor, four quadrature rotary encoders sensors and a GPS. By using these sensors, the robot is able to provide continuous localization; Time Velocity Position and Attitude (TVPA). On top of the robot platform, the Kinect V2 sensor is installed pointing towards the robot. In addition, a microprocessor (Odroid XU4)

is used with Robotic Operating System (ROS) environment for the robot position, images acquisition, processing and point cloud generation.

### C. Target - plant leaves

Two different species of plant leaves were selected as a target for the NDVI point cloud generator evaluation. Fig.2 shows two leaves from the *Alocasia macrorrhizos* (left side) and two leaves of *Monstera deliciosa* (right side). These plants are mostly found in tropical forests, but are also commonly used in indoors for decoration purposes. The main reasons for selecting these leaves are their big size, indoor/outdoor applicability and different health state. Visually, the *Monstera deliciosa* leaves are healthier than the *Alocasia macrorrhizos*.



Figure 2. Plant leaves target

## III. METHODOLOGY

The use of a low-cost sensor implies additional tasks in the traditionally used NDVI point cloud generation methodology. These tasks mainly correspond to the development of rigorous methodology to evaluate and correct the system errors. Thus, the overall NDVI point cloud generation methodology can be split in two steps (Fig. 3). First step involves the procedure for calibration and evaluation of camera sensors While the second step includes the process to generate NDVI point cloud.

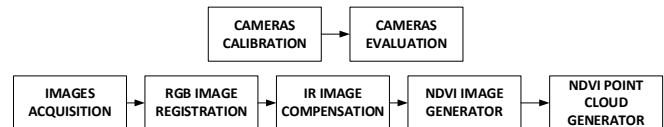


Figure 3. Methodology procedure

### A. RGB and IR cameras calibration

All cameras have a systematic error due to internal camera geometry. These errors lead to the deviation of pixel measurements with respect to space, causing errors in the depth measurement and in registration of images. For this reason, the cameras have to be geometrically calibrated in terms of intrinsic, extrinsic and depth calibration [14]. The intrinsic calibration consists of estimating focal length and optical image center of each camera and is performed using the Conrady-Brown method [15]. The extrinsic calibration refers to establishing geometric relation between RGB and IR camera. The depth calibration is performed by comparing the calibrated stereo vision depth image as a reference to the TOF depth image generated. The output of depth calibration is a matrix with IR image resolution that contains the different distance offsets for each image pixel.

### B. RGB and IR cameras evaluation

The cameras images must be evaluated to determine how the captured pixel intensity values changes depending on the distance of object from the camera and the image region position, while maintaining the environmental illumination conditions.

### C. Data Acquisition

RGB, IR and depth images are acquired and saved into disk. Images are synchronized by using a timestamp as soon as they are acquired. In addition, the robot position and orientation is acquired and stored.

The synchronization between the RGB, IR and the depth images is an important aspect to take into account, whenever these images are acquired using a mobile platform. If the images are not acquired at the same instant of time (not capturing the same scene), an error is introduced at the registration procedure thus, leading to an incorrect NDVI estimation. The magnitude of these errors depends on the relative time between acquisition images and the robot speed.

Two different scenarios are selected for the validation tests. The first environment is an indoor room with constant phosphorescent illumination plus some diffused sun light coming through the windows. The second test scenario is inside a greenhouse with direct sun light.

### D. Image registration

The RGB and IR cameras are slightly separated from each other and have different focal lengths, leading to different image perspectives. The spatial position of captured objects is different, meaning that the images are in different frame coordinates system. The image registration procedure consists of transforming the images into a common coordinate system (same perspective).

Two different methodologies exist to perform the image transformation task. The first one uses feature extraction and matching techniques to find common points between the two images to determine the relative translation, rotation and scale. The second method uses the TOF depth image information to transform one image coordinate to the other. The selected methodology is the TOF image depth technique rectification because it is computationally fast and more robust. Several elements are taken into account such as intrinsic and extrinsic parameters of each camera and TOF calibrated depth image. The image rectification model is based on rotation, scaling, translation, and other affine transformations.

### E. IR image light compensation

A strong and undesired illumination effect is generated from IR light beam of the Kinect creating a non-homogeneous circular illumination. Therefore, the scene is not equally illuminated, with a sharp decay light conditions in the corners of the scene. As the depth and IR images are in the same frame coordinate system, the IR pixel values can be directly compensated due to distance attenuation.

Hence, several images at different distance measurements from the target object are required to estimate the light intensity attenuation model. With the use of the estimated attenuation model and the depth image, the IR image is compensated by adjusting each pixel intensity value.

In addition, a Region Of Interest (ROI) is defined in the IR image due to radial darkening effect (the borders are less illuminated than the centre).

### F. NDVI image generation

Once the IR and the RGB images are in the same frame coordinate system and have the same pixel resolution, the NDVI formula is directly applied to obtain a new NDVI image [5]. First, the red channel is extracted from the RGB image. Then, a new image is created by combining and processing the information from the red channel and the IR channel. The following formula is applied:

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red}) \quad (1)$$

Note that the result of NDVI values varies in-between -1 to 1. This value is then ranged from 0 to 254 to generate a greyscale image using the following formula:

$$\text{NDVI} = (\text{NDVI} + 1) * 127 \quad (2)$$

### G. NDVI Point Cloud generation

The obtained NDVI image and depth image are merged into one single point cloud image. This point cloud contains depth image information (X, Y, Z) associated with NDVI grey image information expressed as RGB (with the same pixel values for all channels).

Once the NDVI point cloud is generated two filtering processes are applied. First, the points containing only depth information, and not having NDVI color information are removed. The second filtering procedure consists of removing all the points outside the valid range. The maximum and minimum range values have been previously described in (section III.B).

### H. Kinematic test

Finally, to validate the point cloud generation tool with dynamics, a kinematic test is performed in an indoor room using the robot wheeled platform equipped with Kinect V2 sensor. During the tests, the robotic platform moves at 0.25m/s speed around *Alocasia macrorrhizos* plant.

## IV. RESULTS

### A. RGB and IR cameras calibration

The IR and RGB cameras have been successfully calibrated geometrically using Kinect2\_calibration node of ROS package Iai\_kinect2. This package uses the OpenCV geometric calibration procedure. Fig. 4 shows the resultant depth calibration matrix representing distance error (in meters) in relation to the pixel XY position coordinates. The depth error mean value obtained is 21cm.

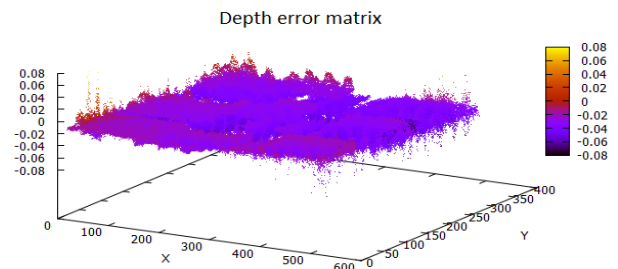


Figure 4. Depth sensor calibration results

### B. RGB and IR cameras evaluation

The analysis of RGB images reveals that neither the distance to the object nor the image position region affects the acquired RGB intensity. For this reason, it has been considered that the RGB camera is only affected from the environmental illumination conditions.

However, the IR images are highly affected from the distance to the target and the image region. This is due to the fact that the measured IR intensity values are closely related with the IR emitter sensor. The reflected IR intensity value variations depend on the object distance to camera due to light attenuation. In addition, the IR emitter has an associated field of view angle projection, meaning that the intensity in the center region of the focus is higher than in the outer region. In addition, the IR images are also affected from external illumination conditions.

Left picture in Fig. 5 has been acquired at 0.5m distance. It can be conceived how the IR camera gets saturated, when the object gets too close to the camera. For this reason, a minimum distance of 0.75m has been determined. By using this minimum distance threshold, it is ensured that the IR image is not saturated (Right picture in Fig. 5).

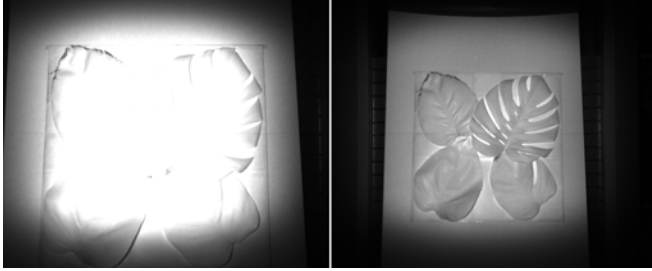


Figure 5. Left: IR image (distance 0.5m); Right: IR image (distance (0.75m))

When the distance to target is greater than a certain distance (considered maximum distance of 1.5m, see Right picture in Fig. 6), the IR light attenuation affects the intensity value measurements from the IR image.

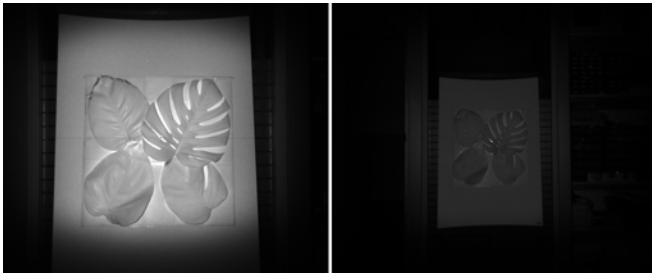


Figure 6. Left: IR image (distance 0.75m); Right: IR image (distance 1.5m)

### C. Data Acquisition

RGB, IR and depth images are simultaneously acquired from the Kinect V2 RGB-D sensor. Fig. 7 shows an IR and a depth image, where it can be observed that both images are in the same frame coordinate system.



Figure 7. Left: IR image; Right: Depth Image

The frame rate achieved during image acquisition is about 1.7Hz, with a high standard deviation of 2.3Hz. This low frame rate is due to Odroid XU4 processor performance and to the involved image processing tasks.

### D. Image registration

As the IR and Depth images are in the same coordinate frame system, thus RGB image is selected for transformation. Fig. 8 shows the original RGB image at the left and the RGB registered image at the right.



Figure 8. Left: original RGB image; Right: RGB registered image

The RGB registered image is scaled to the IR resolution. At the top and at the bottom of the registered RGB image, two black lines can be observed, which are due to scaling process because that area has no RGB pixel information. In addition, some black spots can be observed in the registered image due to the change of perspective, because the original RGB image has no associated pixel data in the new image perspective.

### E. IR image compensation

Several IR images have been acquired at different distances from the target. The intensity value of the central pixel has been measured at each distance. Fig. 9 shows how the pixel intensity value is attenuated, when increasing the distance to object.

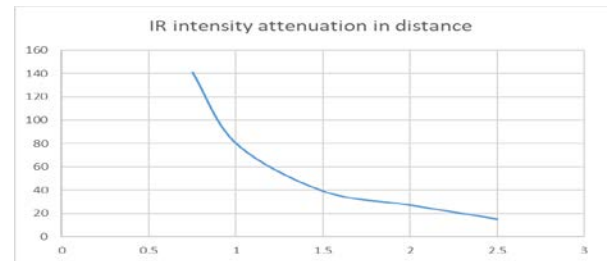


Figure 9. IR attenuation (DN) in distance (meters)

The obtained attenuation has been adjusted with a trend line equation of third order:

$$PV(d) = -67.432d^3 + 387.58d^2 - 745.17d + 508.65 \quad (3)$$



where  $PV(d)$  is the IR pixel value in function of distance and  $d$  is the object distance (expressed in meters).

Once the attenuation profile is characterized, it is possible to compensate the IR pixel values (normalized at 0.75m) using the following compensation formula:

$$IR_c = IR * 140 / PV(d) \quad (4)$$

where  $IR_c$  is the compensated pixel value and  $IR$  is the measured pixel value. Fig. 10 shows the original IR image and the compensated IR image.



Figure 10. Left: Original IR image; Right: Compensated IR image

In addition, a ROI area is defined to eliminate darker pixels at border-side of the image. The original IR image area is 512x424 px and the ROI area defined is 410x340 px centered at the image (Fig. 11).

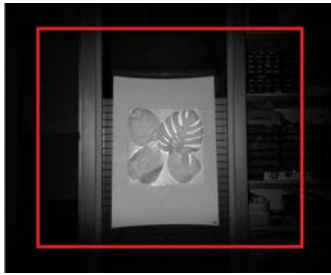


Figure 11. IR image with the ROI area selection

#### F. NDVI image calculation

Once the IR image is compensated and the RGB image is registered, the NDVI image can be generated (Fig. 12).



Figure 12. Left: NDVI greyscale generated image; Right: NDVI colored mapped image. Green color means high NDVI value

After observing the grayscale NDVI image, it can be estimated that the vegetation has higher pixel values than the rest of the image. In addition, the healthier leaves (the right leaves) have higher NDVI index compared to unhealthy leaves (the left leaves).

From the grayscale NDVI image, some white spots can be seen which are the result of the registered RGB image gap

information. These white pixels are needed to be filtered in the point cloud generation filter process.

The NDVI color mapped image has been post processed using the QGIS software. This colored image gives a better interpretation of the NDVI values.

#### G. NDVI Point Cloud generation

Once the NDVI image is generated, the NDVI point cloud (Fig. 13) can be created by fusing the NDVI image with the depth image. The NDVI point cloud was generated using Point Cloud Library (PCL).

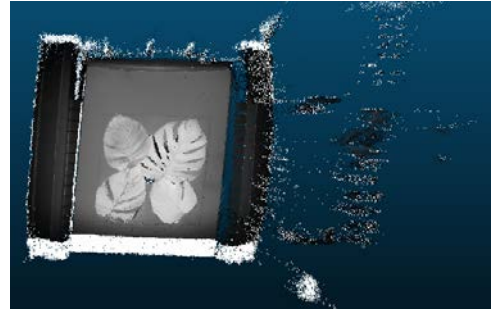


Figure 13. NDVI point cloud

The next process is to filter the NDVI point cloud. The first criteria is to eliminate all white points in the NDVI image (points with no red information). The second filter criteria is to eliminate all the points that fall outside the valid range (0.75 – 1.5m). After the filtering process, a new NDVI point cloud image is obtained (Fig. 14).



Figure 14. Filtered NDVI point cloud

The results show that inside the greenhouse scenario the point cloud obtained is noisier compared to the indoor environment (Fig. 15) and (Fig. 16). This effect is due to direct sun light in the greenhouse environment which affects the IR and depth images quality.



Figure 15. Noisy IR image in the greenhouse scenario

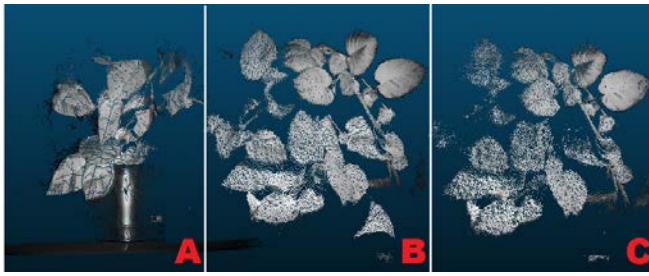


Figure 16. A: Indoor environment; B: GreenHouse with shadow;  
C: Greenhouse without shadows

Nevertheless, the results show that it is possible to generate NDVI point clouds inside greenhouses obtaining coherent NDVI values compared to the indoor scenario.

#### H. Kinematic test

After analyzing the kinematically generated NDVI point clouds, it can be observed that the intensity values at different distances from the plant did not change, thus validating the light compensation procedure. However, a problem with synchronization between RGB and IR images at about 300ms has been identified. This effect leads to an error in the RGB image registration process, affecting the resultant NDVI image. In Fig. 17, left picture shows a static NDVI image correctly registered, while the right picture shows a NDVI image during a rotation of robotic platform. The registration error can be noticed at the right of the leaves.

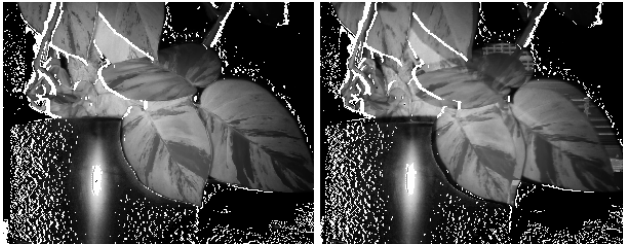


Figure 17. Left: Static NDVI image; Right: Dynamic NDVI image

#### V. CONCLUSIONS

In this paper a tool to generate NDVI point clouds, using a rigorous methodology, has been presented and some preliminary results have been obtained using Kinect V2 sensor data. The IR image illumination has been compensated for a range of distances (providing a compensation model). In addition, a ROI have been defined to minimize the strong beam light effect captured in the IR images. Then, NDVI images and point clouds have been successfully generated and the obtained indices seem to be reliable. The direct sun light affects the quality of the generated point cloud, obtaining better results in indoor environments rather than inside a greenhouse environment. However, quantification has not been achieved due to the lack of reference data. It is worth to mention that the generated NDVI image has low pixel resolution (410x340 pixels). A synchronization issue affects the RGB image registration process, when acquired dynamically. It is important to mention that there is limitation on the operating range from 0.75m to 1.5m due to NIR light effects.

The software developed is available in a ROS compatible package ( [https://github.com/CTTCGeoLab/VI\\_ROS](https://github.com/CTTCGeoLab/VI_ROS) ).

#### VI. OUTLOOK

Further research should include the acquisition and fusion of additional series of point clouds from other perspectives, to increase the point cloud density and to improve 3D NDVI reconstruction. This 3D map would be created by fusing the RGB-D data with the localization of the robot. In addition, the capability to generate point clouds in a temporal frame, would allow monitoring the plant health evolution.

Next steps also include a common timestamp between RGB and depth images to solve the registration issues, the capability to process the grayscale NDVI image into a colored image in real time, the generation of additional vegetation indices such as DVI and SRI, and the optimal placement of the RGB-D sensor in the acquisition platform.

Last but not least, a deep evaluation and validation of the output of the tool, should be done using reference data.

#### REFERENCES

- [1] V.Lawley, M.Lewis, K.Clarke, B.Ostendorf. "Site-Based and remote sensing methods for monitoring indicators of vegetation condition," in *Ecological Indicators*, vol. 60, 2016.
- [2] D.Gates, H.Keegan, J.Schleter, V.Weidner. "Spectral properties of plants," in *Applied optics*, vol.4.
- [3] P.Ceccato, S.Flasse, S.Tarantola, S.Jacquemoud, J.Gregoire. "Detecting vegetation leaf water content using reflectance in the optical domain," in *Remote Sensing of environment*, 2001.
- [4] Sachin D. Khirade, A.B. Patil. "Plant Disease Detection using Image Processing," in *International Conference on Computing Communication Control and Automation*, 2015.
- [5] A.Viña, A.Gitelson, A.Nguy-Robertson, Y.Peng. "Comparison of different vegetation indices for the remote assessment of green leaf area index of crops," in *Remote Sensing of the environment*, 2011.
- [6] Syed A. Hashemi, Setareh K.Chena. "Investigation of NDVI index in Relation to Chlorophyll Content Change and Phenological Event," in *Recent Advances in Environment, Energy Systems and Naval Science*, 2011.
- [7] M.Altaweel. "Measuring Vegetation from Satellite Imagery with NDVI," in *GIS lounge*, 2017.
- [8] C.Moriarty. "Deploying multispectral remote sensing for multitemporal analysis of archaeological crop stress at Ravenshall, Fife," in *Gis research*, University of Edinburgh, 2017.
- [9] F.Endres, J.Hess, J.Sturm, D.Cremers, W.Burgard. "3D mapping with an RGB-D camera," in *IEEE transactions on robotics*, vol.30, 2014.
- [10] AKamal, R. Murtaza, M.Fakhir. "Evaluation of Microsoft Kinect Sensor for Plant Health Monitoring," in *Science Direct*, 2016.
- [11] O.Wasenm, D.Stricker. "Comparison of Kinect v1 and v2 Depth Images in Terms of Accuracy and Precision," in proceedings of *Computer Vision*, 2016.
- [12] L.Valgma. "3D reconstruction using KinectV2 camera," *Bachelor's thesis*, at University of Tartu, Faculty of science and technology, 2016.
- [13] D.Calero, E.Fernandez, M.E.Parés. "Autonomous wheeled robot platform testbed for navigation and mapping using low cost sensors," in proceedings *Low Cost 3D sensors*, 2017.
- [14] M.Chesa, D.Calero. "Obstacle avoidance for an autonomous rover," in *UPC Commons*, 2017.
- [15] C.Brown, Duane. "Decentering distortion on lenses," in *photogrammetric Engineering*, 1966.