Proceedings of the 10th International Workshop on Simulation for Energy,Sustainable Development&Environment,006
19th International Multidisciplinary Modeling & Simulation Multiconference

2724-0029 © 2022 The Authors.
doi: 10.46354/i3m.2022.sesde.006

Visual Change Detection in Multi-Temporal Vegetation Transects of Alpine Plants

Sebastian Pritz[1,2,*], Christoph Praschl[1], Roland Kaiser[3,4] and Gerald Adam Zwettler[1,5]

[1]Research Group Advanced Information Systems and Technology, Research and Development Department, University of Applied Sciences Upper Austria, Softwarepark 11, Hagenberg, 4232, Austria
[2]Department for Medical and Bioinformatics, School of Informatics, Communications and Media, University of Applied Sciences Upper Austria, Softwarepark 11, Hagenberg, 4232, Austria
[3]ENNACON environment nature consulting KG, Altheim 13, 5143 Feldkirchen bei Mattighofen, Austria
[4]Department Environment and Biodiversity, Paris-Lodron University Salzburg, Hellbrunner Str. 34, 5020 Salzburg, Austria
[5]Department of Software Engineering, School of Informatics, Communications and Media, University of Applied Sciences Upper Austria, Softwarepark 11, Hagenberg, 4232, Austria

*Corresponding author. Email address: sebastian.pritz@fh-hagenberg.at

Abstract

Due to the apparent effects of climate change on the Earth's ecosystems, it is more important than ever to monitor flora and fauna in affected regions, e.g. mountain areas above the tree line. In the alpine ecosystem, and not just there, Vegetation plays a fundamental role and is the subject of this study. The work aims to develop algorithms for recognising small stature alpine plants from close range top view images. Ideally, automated assessment algorithms of the plant cover should objectively help scientists observe and interpret the state of the plant ecosystem over a long time series. Therefore, the aim in this respect was to derive visualisations that accurately describe plant growth and displacement (translocation). Additionally, recording changes in biodiversity was an intent. This work uses multi-temporal data comprising RGB images and multi-label masks to accomplish the aforementioned task. The evaluated methods involve mask comparison, optical flow estimation, detection of individual plants, and descriptive statistical analysis of image feature properties. Tests on the given data set show that all methods but the optical flow estimation have great potential. The mask comparison method captured plant growth and translocation most satisfactory. Individual plant detection and statistical analysis further helped to evaluate changes in biodiversity. When combined, the proposed methods give an immediate overview about relevant changes in the multi-temporal transects, which has not been done before for close-distance images of alpine plants.

Keywords: Visual Change Monitoring; Alpine Plants; Optical Flow; Structural Plant Analysis; Change Localization and Visualization

1. Introduction

Concerning the already impending climate catastrophe, as reported by the Intergovernmental Panel on Climate Change (IPCC, cp. Masson-Delmotte et al. 2021; Shukla et al. 2019), the integrity of natural habitats such as forests, seas, meadows, mountains, and so on is increasingly affected. For assumed effects in the alpine belt see Körner and Hiltbrunner (2021). Balanced ecosystems are the foundation of life on our planet and are, among other factors,

based on the biodiversity within habitats (Sandifer et al., 2015).The hypothesised process of shrinking certain habitat types of the alpine flora concerns local populations and may alter complex interrelationships with their habitats.

The conservation of this biodiversity requires seamless monitoring of the individual species across habitats. Such monitoring can be done by detecting visual changes as described by Rensink (2002) using three partial tasks: (I) the detection that something has changed, (II) the identification of what has changed, and (III) the localization, where the change happened. Based on these definitions, the present work introduces visual methods for change detection within multi-temporal, geo-referenced images showing plants of alpine habitats. This represents a novel approach for this area of application.

2. Related Work

Remote sensing images taken from satellites nowadays provide means for visual change detection. Consequently, this field witnessed significant research, e.g. in the detection of changes after disaster events such as flooding (Schlaffer et al., 2015; Dellepiane and Angiati, 2012; Tomowski et al., 2010). Likewise, the evaluation of land cover changes (Partsinevelos et al., 2015), but also general saliency detection (Feng et al., 2018; Du et al., 2013), has been of interest. The present application differs significantly from the previous research mainly regarding the image modalities used – close range RGB images captured by consumer cameras instead of hyperspectral satellite recordings. Also, the size of the area of interest (the scale of the scene) is different; e.g. a tree crown (10 m diameter) corresponds roughly to a rosette of a small stature herbaceous plant (1 – 5 cm). Nevertheless, the mentioned publications partially use methods comparable with this work to visualise changes using colour overlays and difference masks.

Yu et al. (2013) presents a methodology for change detection of growth stages of maise. Like in the present work, the authors visualise temporal differences of plants in geo-referenced sections and also use skeleton-based graph representations to identify and cluster leaves to form, e.g. a plant rosette.

In evaluating geometric changes in segmented spatio-temporal frames, different strategies from the field of object tracking are attractive. A very general way is the application of optical flow to the intensity or colour profiles of image pairs, as presented, e.g. by Farnebäck (2003). Furthermore, SIFT-based object detection allows tracking regions by their edges using multi-resolution scale-spaces (see Zhou et al., 2009).

With pre-fragmented regions available from instance segmentation, conventional shape-interpolation strategies such as the Poisson shape interpolation elucidated by Xu et al. (2005) are applicable if regions overlap partially. Shape interpolation can also be established based on AI-powered interpolation of intermediate frames as doable with DAIN (cp. Bao et al., 2019). Familiar shape interpolation strategies tend to fail whenever the neighbouring segmentation masks are fragile, and the spatio-temporal distances get large. In this case, which is highly relevant in the medical domain (cf. brain MRI in axial view with thin gyrus and sulcus courses), the correct identification of congruent regions is not trivial. However, the method developed by Rajagopalan et al. (2003) might solve it.

Finally, Varghese et al. (2018) presented a methodology based on convolutional neural networks for change detection. The advocated model can detect changes of added/removed objects in scenarios with partially deviating camera positions. In contrast, the present work focuses on visual changes of individuals in multi-temporal and statically geo-referenced views, focusing not on the appearance/absence of objects but their transformations.

3. Material

The methodology introduced in this work is based on multi-temporal images of 14 geo-referenced image mosaics (image tiles arranged along a transect) obtained from the National Park Hohe Tauern (Austria, for details see Körner et al. 2022). A single image tile depicts an orthogonal (nadir) view to the ground of a 50×50 cm (3000×3000 px) section of the floor surface. Images are taken standardised along a sampling raster 1 m wide and 8 m long. In total a transect is made up of 32 images tiles; 1 m^2 equals four 50 cm image tiles (cp. Eberl and Kaiser, 2019). In the image acquisition task, the transect is split using geo-referenced markers. Every single image is taken using a shading curtain in combination with two diffused flashes; colour charts and grey cards guarantee the colourfastness of the images. Altogether this technique ensures uniform illumination of the motif that can then be analysed in a multi-temporal manner without external influences, like the prevailing weather condition at the shooting. The analysed data set consists of a random sample of image cut-outs of the 14 transects showing an area of 10×10 cm (600×600 px) in 2017 and 2021.

For the comparison of the transformation of the plants, multi-class labels have been created for the following set of 11 plant species: *Euphrasia minima*, *Gnaphalium supinum**, *Leucanthemopsis alpina**, *Potentilla aurea** *Primula glutinosa**, *Primula minima**, *Salix herbacea**, *Scorzoneroides helvetica*, *Soldanella pusilla**, *Vaccinium gaultheroides*, and *Polytrichum sexangulare*. Many of the mentioned species show clonal growth (marked with an asterisk) through rhizomes (cp. Klimeš et al., 1997). For this reason, the position of renewal buds can change from year to year due to directional growth of the rhizome. One species has an annual life cycle, one is a prostrate shrub, and one species of moss is included.

4. Methodology

This work aims to assist in the automated image-based change detection regarding the growth, translocation, and biodiversity of selected alpine plants over long periods.

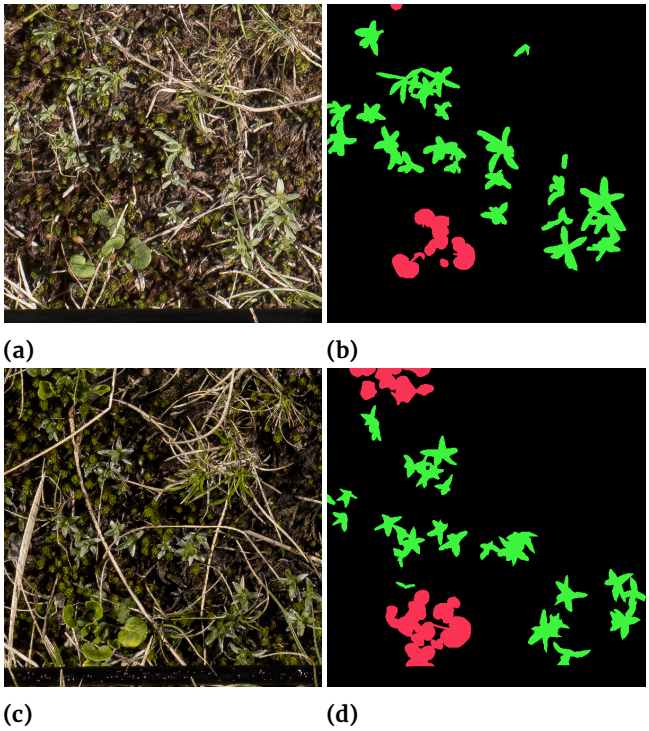


Figure 1. A sample geo-referenced image cut-out (10 × 10 cm) is shown in (a) for 2017 and in (c) for 2021. Next to the source images, the associated multi-label masks are shown in (b) again for 2017 and (d) for 2021. The masks depict two sample plant species with green for *Gnaphalium supinum* and red for *Soldanella pusilla*.

The reference images and the multi-label masks mentioned in the previous section form the data basis for this paper. An additional file encodes the labels and the specific colours attributed to a segmented species.

In order to solve the presented problem, we consider several methods: (I) raw mask comparison, (II) optical flow, (III) detection of individual plants, and (IV) descriptive statistics. An overview of all applied image-based methods is offered in Figure 2 and 3. After elaborating on basic requirements and preprocessing, the four methods above get described in further detail in subsequent sections.

4.1. Basic Requirements

Temporal data is needed to visualise changes in plant growth and position. This work conducts a pairwise comparison of time-separated images and their corresponding multi-label masks. Therefore, every proposed method requires five different kinds of input data, namely:

1. Base image img_{base}: showing the initial image cut-out
2. Reference image img_{ref}: showing the same image detail as the base image, but with a temporal offset
3. Multi-label mask of base image
4. Multi-label mask of reference image
5. Label file

Based on these files, the following reprocessing operations are evaluated. All changes made to create suitable visualisations are directly applied to the base image.

4.2. Preprocessing

In order to create easy to follow visualisations, the multi-label masks are split up by plant species at the start of an algorithm. The splitting is done by extracting the pixels with the relevant colours stated in the label file and creating new temporary binary masks for each species. Additionally, the background of the base image, which does not contain relevant plant species, is darkened and converted to grayscale to highlight plants of interest. This step ensures that visual clutter in result images is kept at a minimum.

4.3. Raw mask comparison

One first attempt to compare plant growth and translocation is to compare the masks for each plant species directly. This can be done using (I) contour overlay or (II) difference overlay. The proposed approaches will be discussed in further detail below.

4.3.1. *Contour overlay*

This variant that can be seen in Figure 2a is based on edge detection and image overlay. First, a Canny edge detection algorithm (c.f. Canny, 1986) is applied to the split mask so that only the edges of an object remain. After that, the resulting edge mask is morphologically dilated to thicken the lines outward for increased visibility in the final image. This altered mask can then be used to set the corresponding pixels in the source image to a specific colour.

The advantages of this method are that the contours do not obstruct the view of the masked areas while still providing information about which plants are affected and where they are located in the given images and time period. Since the changes are applied to the base image, it is immediately evident how a specific region has changed over time.

4.3.2. *Difference overlay*

Similar to the contour overlay, overlaying the entire masked regions is possible. Instead of applying an edge detector, the difference between the two masks is calculated. Like the contour overlay variant, the mask is then used to colour the corresponding regions in the reference image with the desired colour. These regions can be shaded in (I) static colours or (II) in a distance-based manner.

While static colours give a basic idea about which areas have changed between the two images, they do not give information about how much growth or movement has happened. This problem can be solved by calculating distance-maps of both masks, so that the intensities of each pixel match the distance to the nearest masked pixel in the other mask. By combining these newly generated masks with chosen colours, the distances directly corre-

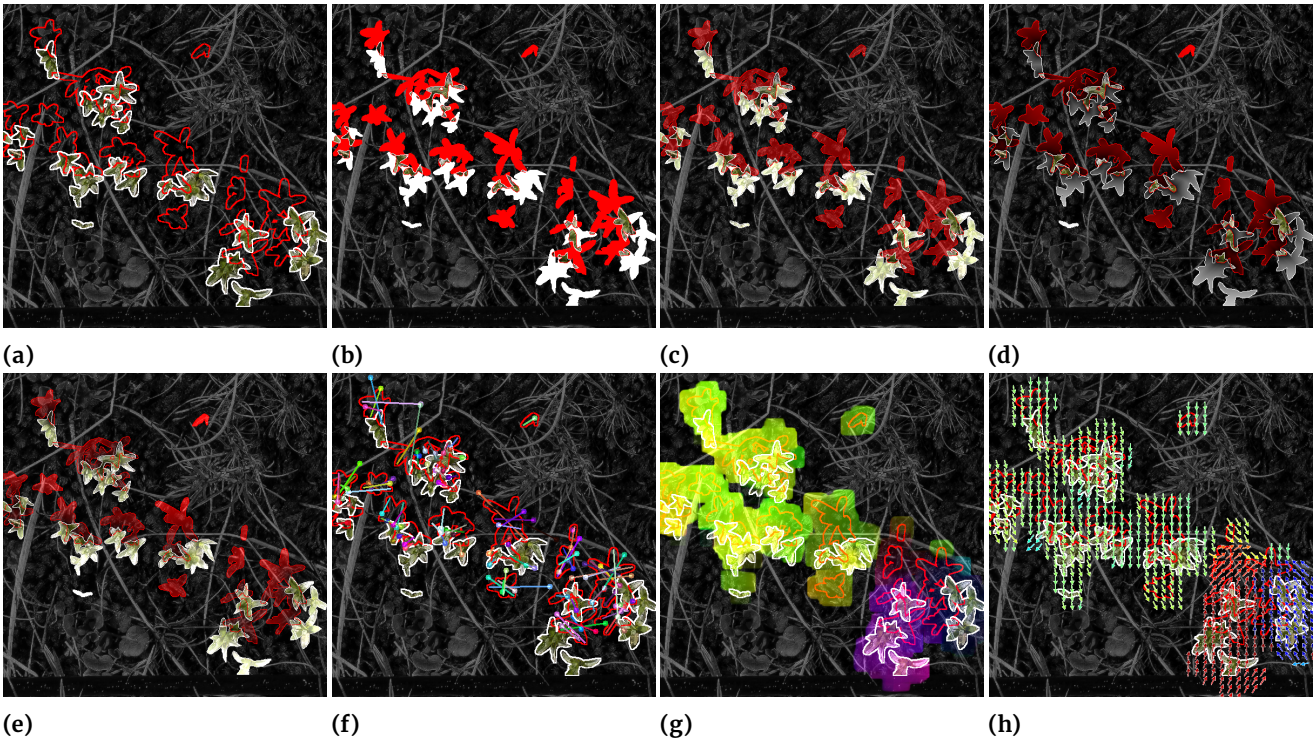


Figure 2. A visual presentation of all implemented methods based on the example image cut-out depicted in Figure 1. Figures (a) to (e) are based on raw mask comparison. Starting from the upper left, the visualisations display the contour overlay (a), the difference overlay using static colours (b), blended static colours (c), distance-based intensities (d), and blended distance-based intensities (e). Figures (f) to (h) show optical flow estimation of the given images in the order of sparse optical flow (f), dense optical flow in HSV representation (g), and dense optical flow using quiver plots (h).

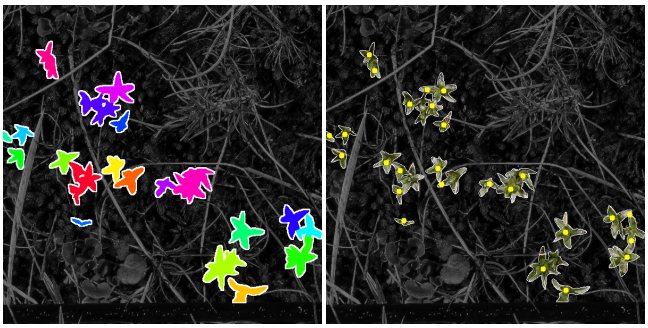


Figure 3. The separation of individual plants computed for the transects visible in Figure 1 using watershed segmentation (a) and graph-based differentiation (b).

late with the pixel's brightness, resulting in bright colours where significant growth or translocation has occurred, c.f. Figure 2b-2e.

Alternatively, the resulting overlay can be blended to preserve detail in plant texture while also applying the visualisation technique. This task is achieved by morphing the images using linear blending, as presented by Szeliski (2021) and shown in Equation 1,

$$I = \alpha * p1_{i,j} + \beta * p2_{i,j} + \gamma, \quad (1)$$

where $p_{i,j}$ represents a given pixel value at position (i,j), α and β describe the weights for the pixel's colour values, and γ is the gamma correction.

4.4. Optical flow

Instead of using the supplied masks directly, the optical flow of the images can be calculated. As proposed by Fleet and Weiss (2006), optical flow is applied to display patterns of motion, an estimation of the direction and speed of movement also called the *2d motion field*. Its estimation generally works by using time-varying image data to approximate said field. This property is relevant to the presented problem since it might give a more detailed answer to where, how much and in which direction growth or translocation has occurred and acts as an alternative to the distance-map based overlay approach. Both dense and sparse optical flow was considered; details are as follows.

4.4.1. *Sparse optical flow*

Sparse optical flow is based on detecting features in images and calculating motion for all pixels in a small region around the features simultaneously (Spruyt et al., 2013). This visualisation technique is depicted in Figure 2f. Here, Shi-Tomasi corner detection (Shi and Tomasi, 1994) is executed to acquire image features which are then used by the Pyramid Lucas & Kanade algorithm (Lucas and Kanade, 1981) to calculate sparse optical flow (see Bouguet, 2000).

4.4.2. *Dense optical flow*

In contrast to sparse optical flow, dense optical flow calculates the displacement vector for every pixel in the image and does not require feature detection or tracking. To compute dense optical flow, the polynomial expansion algorithm, proposed by Farnebäck (2003), is used.

However, instead of using the default representation of mapping colour to direction, the motion field is displayed as a quiver plot, in which a grid of directional arrows shows motion. Each arrow indicates the predicted movement direction at the arrow's position in the image. In addition, the arrows are colour-coded, similar to the standard version, to distinguish the different directions better. A comparison between the two methods is shown in Figure 2g and 2h.

4.5. Detection of individual plants

The given multi-label masks prove great when differentiating between plant types. However, when multiple plants of the same type grow too close together, it is impossible to deduce the amount and location of individual plants from the mask alone since plants of the same species share the same colour coding.

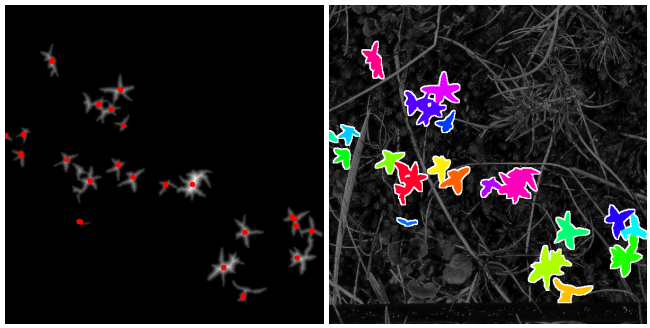
Knowing how many individuals of a species exist and being able to monitor the change over the years would allow for a more sophisticated evaluation in this regard. Additional techniques are required to remedy this situation. The problem can be solved by applying (I) watershed segmentation or (II) graph-based clustering. The different implementations will be described in more detail in the following two sections.

4.5.1. *Watershed segmentation*

As explained by Preim and Botha (2013), watershed segmentation interprets the entire image as a topographic landscape comprised of valleys and mountains. The basic idea is that the intensity values of a grayscale image represent elevations like a digital terrain model. Water will accumulate in this 'landscape' at the local minima of the virtual topography, the so-called catchment basins. When taking the path of the steepest descent, every point on the topography that ends in a local minimum is assigned to a catchment basin. After completing this process for the entire image, every pixel gets assigned to a specific catchment basin. The ridges they are surrounded by then represent the segmentation boundaries. Since the base algorithm is highly susceptible to over-segmentation, results usually need to be followed up by basin merging.

Instead of relying solely on plain elevation data, a marker-based watershed estimation is an alternative. In this method, the user can actively supply markers from which the flooding originates, rather than starting the flooding at the local minima. Usually, markers for the target structure (foreground, include markers) and ones for the background (exclude markers) are defined.

The marker-based watershed algorithm aids to segment the individual plants in this work. The supplied markers are determined by creating a distance map of the existing segmentation masks and extracting local maxima. A maximum filter is applied beforehand to combine spatially close local maxima to avoid problems due to noise. The process and its results can be seen in Figure 4.



(a) (b)

Figure 4. Usage of marker-based watershed segmentation for plant separation. A distance-map (a) is being used to calculate the input markers (red), which are then used to compute the segmented image (b).

4.5.2. *Graph-based differentiation*

In addition to separating individual plants, detecting plant bud and stem locations proved to be of interest. Since most plant species in question converge and grow from a single vegetative point and form a star pattern when skeletonized (rosette plants), graph-based analysis is introduced to unravel the problem of finding a plant's center point.

In order to disentangle the graph problem, the masks are first converted to graph networks. In this way, the individual masks created during preprocessing are thinned to a skeleton, as described by Zhang and Suen (1984) and subsequently converted to a graph. This conversion is performed by implementing a line tracer which searches for graph leaves in the thinned images and then recursively follows the line and visits all branches until every connected pixel is visited. The coordinates of all graph leaves and bifurcations are saved during the process, and a distance matrix representing the graph is created for each connected region.

After calculating the graph, clustering can be applied to split up plants that are spatially too close to each other and estimate bud and stem positions. For this purpose, several methods were investigated: (I) density-based clustering, (II) bifurcation-oriented weight thresholding, and (III) a novel leaf-oriented hierarchical clustering algorithm. The explanation of each method follows in the subsequent paragraphs.

General rules for location estimation. Finding bud or stem locations does not always require clustering. If the initial graph is small enough, no further processing is necessary. This condition is only fulfilled if the graph has one or no

branches. Hence, if the graph consists solely of leaves, their coordinates are averaged to estimate the plant's centre (vegetative bud or flowering stem origin). If the graph has precisely one bifurcation, its coordinates are considered the bud or stem location. Graphs with two or more bifurcations are handled by the means described below.

Density based clustering. Clustering using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) allows clustering of nodes based on their distance to each other and their spatial density. DBSCAN is an algorithm proposed by Ester et al. (1996) allowing the computation of clusters without defining a cluster number beforehand. The latter is especially important since the final plant count (and therefore the number of clusters) is unknown at the beginning. This prerequisite is also the reason why other methods, like k-Means clustering (MacQueen, 1967), are unfeasible, as they require the user to supply a cluster number.

Bifurcation-based weight thresholding. With an alternative approach, the problem can be solved without the need for clustering. In doing so the weight of each bifurcation node is calculated by adding up the distances from the node itself to all adjacent nodes in the graph and only keeping the nodes that meet a certain weight threshold. Based on the assumption that bud and stem locations are usually located in the centre of individual plants, they are expected to show a maximum number of links to other points in the graph. The threshold needed by this method is defined as 40% of the maximum node weight; thus, all bifurcation nodes with a weight above this threshold are considered as bud locations.

Leaf-oriented hierarchical clustering. This algorithm uses a greedy hierarchical clustering approach to solve the proposed problem. To optimize results, clustering constraints have been set so that clustering will always start at the leaves of the graph. Additionally, the growth of clusters is regulated so that after passing a certain cluster weight threshold, it becomes harder to merge a cluster with others. Here, the sum of edge lengths (distances between nodes in the sub-graphs) is used as cluster weight. The formula applied to obtain a threshold value is deduced empirically. It is set to twice the average edge length using the formula $T_w = \sum d_{i,j}/N * 2$, where $d_{i,j}$ is the edge length between nodes i and j, and N is the total amount of nodes in the cluster.

When starting to merge clusters, of which at least one exceeds the weight threshold, the single-linkage distance between the clusters is considered. The distance needs to be below a given threshold to merge clusters successfully. The threshold is set to ten times the ratio between the given clusters during empirical analysis. The formula to calculate a threshold value is $T_d = 10 * weight(C1)/weight(C2)$, where $C1$ is the bigger cluster, and $C2$ is the smaller one.

Therefore, merging clusters becomes easier the more prominent the disparity between their size is. This set of rules is applied to receive clusters of similar size and prevent excessive merging.

In order to calculate a cluster-based region fragmentation, each node gets assigned to its own cluster at the beginning. The first step of the algorithm is to cluster the leaves by iterating through the distance matrix and searching for the adjacent bifurcation nodes of the leaves to merge their clusters.

Following this initial clustering, the algorithm seen in Figure 5 attempts to merge clusters until a termination criterion is reached continuously. The algorithm meets the termination criteria if only one cluster remains or if no further merging of clusters is possible. The latter happens only if the remaining clusters have exceeded the weight threshold and are too far away from each other. The order in which the algorithm starts to merge clusters is determined by the distance so that closer clusters get evaluated and merged first.

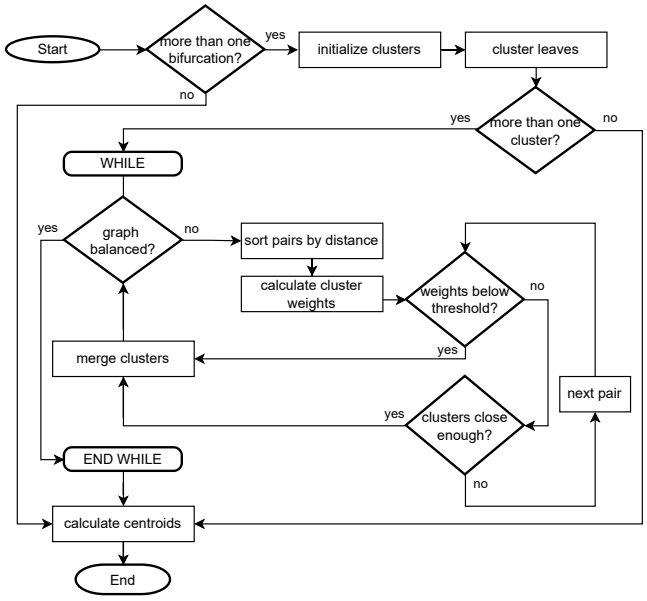
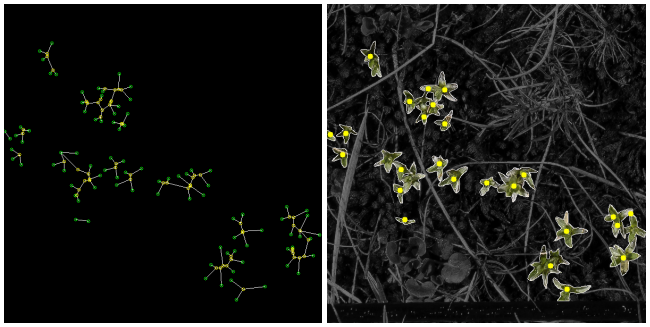


Figure 5. The leaf-oriented hierarchical clustering algorithm implemented for plant separation. The algorithm's core revolves around continuously sorting and merging clusters until no more changes are possible.

4.6. Statistical data

Another method for evaluating images of the given kind is to create descriptive statistics by calculating texture features and other numerical characteristics. In particular, computed feature properties included the following: (I) circularity, (II) vegetation density, (III) colour variance and (IV) co-occurrence features. Their meaning and calculation is detailed below.



(a) (b)

Figure 6. The procedure of graph construction followed by leaf-oriented hierarchical clustering. The skeleton of the binary mask is being traced to obtain all leaf- and bifurcation nodes (a). These nodes are then clustered using the algorithm in Figure 5 to obtain the bud points (c).

4.6.1. Vegetation density

Vegetation density (d) is the ratio of how many pixels in the target mask M are part of vegetation compared to how many pixels the mask has in total. This process is done separately for each plant species and is calculated by counting all foreground pixels in the target mask and dividing this amount by the total pixel count as in Equation 2.

$$d = \frac{A(M > 0)}{A(M)} \tag{2}$$

4.6.2. Circularity

Circularity (c) is the ratio of a region's area $A(R)$ to its perimeter $P(R)$ and is calculated according to Equation 3.

$$c = 4\pi \frac{A(R)}{P^2(R)} \tag{3}$$

The area $A(R)$ is calculated by counting the connected pixels of the given region R in a specific mask. To compute the perimeter, a modified line tracer, similar to what is presented in section 4.5.2, is implemented. As discussed in the named section, the mask is split further so that each connected region of each plant species is evaluated individually. However, in this case, the outer edges of the region R are traced instead and the individual euclidean distances between pixels are summed up.

4.6.3. Colour variance

Colour variance (C), or colourfulness, is a metric to describe how colourful an image is (c.f. Hasler and Suesstrunk, 2003). As explained below, it is calculated by determining the mean values and standard deviations of specific colour channels. The first step consists of calculating the opponent colour space representation using the values of the initial red (R), green (G) and blue (B) channels. Based on these values, the red-green (rg) and yellow-blue (yb) channel are calculated using Equation 4 and 5, respectively.

Feature	Definition	Description
Entropy	$-\sum_{i=1}^{m}\sum_{j=1}^{n} p_{i,j} \ln p_{i,j}$	Information content
Energy	$\sum_{i=1}^{m}\sum_{j=1}^{n} p_{i,j}^2$	Intensity
Variance	$\sum_{i=1}^{m}\sum_{j=1}^{n} (i-j)^2 p_{i,j}$	Intensity level
Homogeneity	$\sum_{i=1}^{m}\sum_{j=1}^{n} \frac{p_{i,j}}{1+\lvert i-j\rvert}$	Homogeneity

Table 1. The definitions and descriptions for common co-occurrence features, where $p_{i,j}$ represents the intensity value of a given pixel at the position (i,j) in an image of size $n \times m$.

$$rg = R - G \tag{4}$$

$$yb = \frac{1}{2}(R + G) - B \tag{5}$$

Subsequently, using this colour representation, the collective mean value and standard deviation for both channels is computed using the formulas given in Equation 6 and 7.

$$\mu_{rgyb} = \sqrt{\mu_{rg}^2 + \mu_{yb}^2} \tag{6}$$

$$\sigma_{rgyb} = \sqrt{\sigma_{rg}^2 + \sigma_{yb}^2} \tag{7}$$

Finally, using μ and σ values, the colour variance (C) can be calculated with Equation 8.

$$C = \sigma_{rgyb} + 0.3 * \mu_{rgyb} \tag{8}$$

4.6.4. Co-occurrence features

Additional texture features can be calculated using a grey-level co-occurrence matrix (GLCM). This matrix can be calculated by converting a given image to grayscale, narrowing the intensity range through normalization and counting how often a pair of co-occurring intensities exist within a given distance, as specified by Haralick (1990). In this case, a GLCM with a size of 64×64 is used, which is computed using foreground pixels only. The co-occurrence features, which are used in this work, are denoted in Table 1.

5. Results

In this section, the proposed methods are evaluated in terms of suitability, robustness and usability.

5.1. Results of the contour overlay approach

The implemented contour overlay shows excellent results in all plant types in the tested data set (Figure 2a). However, it also yields the smallest information gain compared to the other methods.

5.2. Results of the difference overlay approach

Like the contour overlay, the difference overlay does not face any strict requirements and can therefore be applied to all images in the test data set. Interestingly, compared to the contour overlay, difference overlay imposes an improvement giving a better idea of the scale of how much change has happened, see Figure 2b-2e.

5.3. Results of the optical flow approach

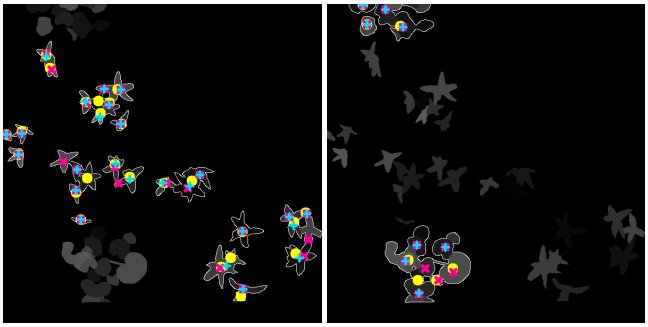
Optical flow, while attractive in theory, does not adapt well to the presented task in most cases because the change between images is often too big. The apparent transformation is most likely due to plants being replaced by new ones or new emerging shoots from the same clonal individual nearby when the base image (img_{base}, as explained in subsection 4.1) is taken. This is a burden since the proposed algorithms are based on the assumption that the intensity-structure of images stays constant between the compared images. The required condition is usually only fulfilled for short periods (cp. Beauchemin and Barron, 1995). Since the plants themselves might be different, the intensity values may change accordingly, which leads to bad results when computing the motion field, as can be seen in Figure 2f and 2h. If it were successful, the visualisations would show movements from regions of the reference image (red) to regions of the base image (white).

5.4. Results of the watershed segmentation approach

Acceptable results could be achieved by applying marker-based watershed segmentation to the given masks. It became apparent that how well this method works depends on the initial mask. Judging by the evaluated transects, segmentation worked great for plants with short, wide leaves and circular masks and failed for plants with long and narrow leaves. This behaviour restricts the method to specific plant architecture. Examples for the application of this method are shown in Figure 3a. Its results are shown in Table 2 and are further compared with the graph-based approaches in Table 3.

Plant species	Split ratio	Merge ratio
Gnaphalium supinum	1.58	1.36
Salix herbacea	1.14	1.41
Scorzoneroides helvetica	2.09	2.00
Soldanella pusilla	1.00	2.20
Leucanthemopsis alpina	3.22	1.00

Table 2. Evaluation of error in watershed segmentation by manually comparing annotated plants with segmented plants. A split ratio of 5 indicates that watershed segmentation falsely splits a single plant into five on average. The merge ratio does the opposite, where a value of 5 means that 5 plants are merged and incorrectly treated as one on average.



(a) (b)

Figure 7. A example image cut-out containing plants of species *Gnaphalium supinum* (a) and *Soldanella pusilla* (b) evaluated using the three proposed methods. Each method is displayed in a different colour over a manually segmented mask. DBSCAN (yellow), bifurcation-oriented weight thresholding (pink), and leaf-oriented hierarchical clustering (blue).

5.5. Results of the graph-based differentiation approach

Three different graph-based methods were applied to separate individual plants. A comparison of the applied methods can be seen in Figure 7.

While the separation is not perfect, compared to the initial mask it offers a better picture of how many individuals actually exist in the image and where their center is.

Plant species	GT	I	II	III	IV
Gnaphalium supinum	26	23/25	23/26	23/25	20/20
Soldanella pusilla	25	7/7	11/11	8/8	13/13

Table 3. Comparing ground truth plant count (GT) to the ones calculated by the algorithms (I: DBSCAN, II: Bifurcation-oriented weight thresholding, III: leaf-based hierarchical clustering, IV: watershed segmentation). For each method, the true-positive and total amount of predictions (true-positive + false-positive) is denoted.

Judging by the results in Table 3, watershed segmentation and bifurcation-oriented weight thresholding proved to be the most robust, detecting a total of 33 and 34 plants (sum of true-positives over both species), respectively, compared to 51 found in manual annotations. Contrary to the other algorithms, the two best performing algorithms are better at separating larger regions made up of several smaller plants while excelling for regions with few plants.

5.6. Results of the statistical data approach

Most of the statistics in question are evaluated for individual plants, by cropping the region of interest, and calculating them for this particular region only. Only selected measures, like vegetation density and leaf count, are calculated for the entire image. In order to compare the results, all features are calculated for all plant species present in a set of 100 images, as illustrated in Figure 8. These values are additionally normalized per feature and over all plants with $feature_{norm} = {feature - min(feature)}/{max(feature) - min(feature)}$.

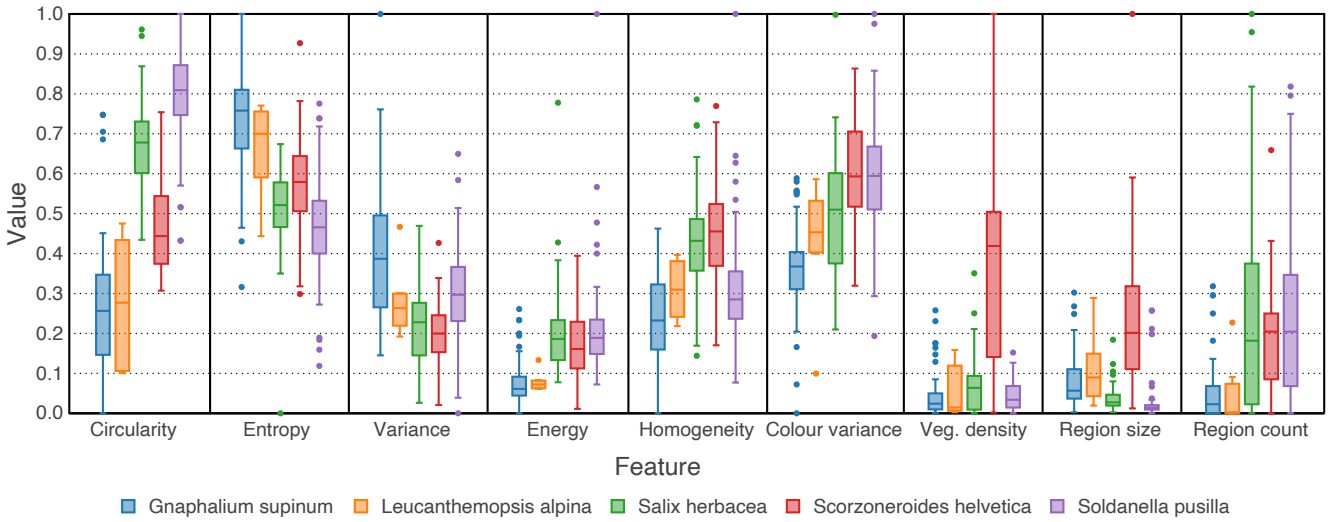


Figure 8. The proposed texture features circularity, entropy, variance, energy, homogeneity, colour variance, vegetation density, average region size and region count calculated and normalized over a set of each 50 images in the years 2017 and 2021 for the plant species within the evaluated transect displayed as box-plots.

6. Conclusion and Outlook

This work aimed to present methods that assist in automatic change detection in multi-temporal images of alpine vegetation and its constituent plant species. The main focus was visualising the change in individual plant growth and translocation. In order to achieve this goal, several methods were tested and compared to available manually annotated data of select alpine plants. The coloured overlays offered a good impression of the plant size and location change in testing. By comparing each plant individual separately, assumptions on plant performance can be drawn more effectively. The latter statement gets enriched by adopting statistical data, like vegetation density or plant count. The presented plant separation techniques are helpful in order to obtain accurate figures of plant count (population density), which is a fundamental property in population biology. In addition, the proposed methods may be applied in tandem in order to achieve the best results.

In the future, the presented methods could be beneficial in post-processing images for which there are automatically generated plant segmentation masks, e.g. as a result of machine learning algorithms such as neural networks. Instead of manually labelling and separating the plants, the user could predict the labels using the given plant separation algorithms. Fortunately, the results of the manual labels mimic the results of the watershed segmentation.

As a refinement, advanced image registration methods should be applied to align the images more accurately, as slight deviations are still possible. This alignment could be performed by finding time-invariant landmarks like stones embedded in the ground and using them to superimpose the images.

Finally, the proposed methods need to be evaluated by experts in the field in order to obtain qualitative feedback on whether the implemented methods bring an easement to the daily work routine.

Acknowledgment

Our thanks to the province of Upper Austria for facilitating the project AlpinIO with the easy2innovate funding program.

References

Bao, W., Lai, W.-S., Ma, C., Zhang, X., Gao, Z., and Yang, M.-H. (2019). Depth-aware video frame interpolation. In *IEEE Conf. on Computer Vision and Pattern Recognition*.

Beauchemin, S. S. and Barron, J. L. (1995). The computation of optical flow. *ACM Comput. Surv.*, 27:433–467.

Bouguet, J.-Y. (2000). Pyramidal implementation of the lucas kanade feature tracker description of the algorithm. *OpenCV Document, Intel, Microprocessor Research Labs*, 1.

Canny, J. (1986). A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8:679 – 698.

Dellepiane, S. G. and Angiati, E. (2012). A new method for cross-normalization and multitemporal visualization of sar images for the detection of flooded areas. *IEEE Trans. on Geoscience and Remote Sensing*, 50:2765–2779.

Du, P., Liu, S., Xia, J., and Zhao, Y. (2013). Information fusion techniques for change detection from multitemporal remote sensing images. *Information Fusion*, 14(1):19–27.

Eberl, T. and Kaiser, R. (2019). Langzeitmonitoring von ökosystemprozessen im nationalpark hohe tauern. modul 02: Botanisch / vegetationskundliche analysen. methoden-handbuch.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996).

A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, volume 96, pages 226–231. AAAI Press.

Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. In *Image Analysis*, volume 2749, pages 363–370. Springer Berlin Heidelberg.

Feng, W., Sui, H., Tu, J., Huang, W., and Sun, K. (2018). A novel change detection approach based on visual saliency and random forest from multi-temporal high-resolution remote-sensing images. *International Journal of Remote Sensing*, 39(22):7998–8021.

Fleet, D. and Weiss, Y. (2006). *Optical Flow Estimation*. Springer US, Boston, MA.

Haralick, R. (1990). Texture features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics - TSMC*.

Hasler, D. and Suesstrunk, S. (2003). Measuring colourfulness in natural images. *Proceedings of SPIE - The International Society for Optical Engineering*, 5007:87–95.

Klimeš, L., Klimešová, J., Hendriks, R., and van Groenendael, J. (1997). *The ecology and evolution of clonal plants*, chapter Clonal plant architectures: a comparative analysis of form and function, pages 1–29. Backhuys Publishers.

Körner, C., Berninger, U., Daim, A., Eberl, T., Fernández Mendoza, F., Füreder, L., Grube, M., Hainzer, E., Kaiser, R., Meyer, E., Newesely, C., Niedrist, G., Petermann, J., Seeber, J., Tappeiner, U., and Wickham, S. (in press 2022). Long-term monitoring of high elevation terrestrial and 2 aquatic ecosystems in the alps – a five-year synthesis. *ecomont*, 14:44–65.

Körner, C. and Hiltbrunner, E. (2021). Why is the alpine flora comparatively robust against climatic warming? *Diversity*, 13:1–15.

Lucas, B. D. and Kanade, T. (1981). *An Iterative Image Registration Technique with an Application to Stereo Vision*. IJCAI'81. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *In 5-th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297.

Masson-Delmotte, V., P. Zhai, A. P. a. C., C. Péan, S. B., Chen, N. C. Y., Goldfarb, L., Gomis, M., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J., Maycock, T., Waterfield, T., Yelekçi, O., Yu, R., and (eds.), B. Z. (2021). Ipcc, 2021: Climate change 2021-the physical science basis. *Press*.

Partsinevelos, P., Nikolakaki, N., Psillakis, P., Miliaresis, G., and Xanthakis, M. (2015). Landcover change modeling through visualization and classification enhancement of multi-temporal imagery. *Global NEST (under publication)*.

Preim, B. and Botha, C. (2013). Visual computing for medicine: Theory, algorithms, and applications: Second edition. *Visual Computing for Medicine: Theory, Algorithms, and Applications: Second Edition*, pages 1–812.

Rajagopalan, S., Karwoski, R., and Robb, R. (2003). Shape-based interpolation of porous and tortuous binary objects. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, volume 2879, pages 957–958. Springer Berlin Heidelberg.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, 53(1):245–277. PMID: 11752486.

Sandifer, P. A., Sutton-Grier, A. E., and Ward, B. P. (2015). Exploring connections among nature, biodiversity, ecosystem services, and human health and well-being: Opportunities to enhance health and biodiversity conservation. *Ecosystem services*, 12:1–15.

Schlaffer, S., Matgen, P., Hollaus, M., and Wagner, W. (2015). Flood detection from multi-temporal sar data using harmonic analysis and change detection. *International Journal of Applied Earth Observation and Geoinformation*, 38:15–24.

Shi, J. and Tomasi (1994). Good features to track. In *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600.

Shukla, P., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H., Roberts, D., Zhai, P., Slade, R., Connors, S., Van Diemen, R., et al. (2019). Ipcc, 2019: Climate change and land: an ipcc special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. *Press*.

Spruyt, V., Ledda, A., and Philips, W. (2013). Sparse optical flow regularization for real-time visual tracking. In *2013 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6.

Szeliski, R. (2021). *Computer Vision: Algorithms and Applications*. Springer.

Tomowski, D., Klonus, S., Ehlers, M., Michel, U., and Reinartz, P. (2010). Change visualization through a texture-based analysis approach for disaster applications. In *ISPRS Proceedings*, pages 1–6.

Varghese, A., Gubbi, J., Ramaswamy, A., and Balamuralidhar, P. (2018). Changenet: A deep learning architecture for visual change detection. In *Proc. of the European Conf. on Computer Vision (ECCV) Workshops*.

Xu, D., Zhang, H., Wang, Q., and Bao, H. (2005). Poisson shape interpolation. In *Proceedings of the 2005 ACM Symposium on Solid and Physical Modeling*, SPM '05, page 267–274, New York, NY, USA. Association for Computing Machinery.

Yu, Z., Cao, Z.-G., Wu, X., Bai, X., Qin, Y., Zhuo, W., Xiao, Y., Zhang, X., and Xue, H. (2013). Automatic image-based detection technology for two critical growth stages of maize: Emergence and three-leaf stage. *Agricultural and Forest Meteorology*, s 174–175:65–84.

Zhang, T. and Suen, C. (1984). A fast parallel algorithm for thinning digital patterns. *Commun. ACM*, 27:236–239.

Zhou, H., Yuan, Y., and Shi, C. (2009). Object tracking using sift features and mean shift. *Computer Vision and Image Understanding*, 113:345–352.