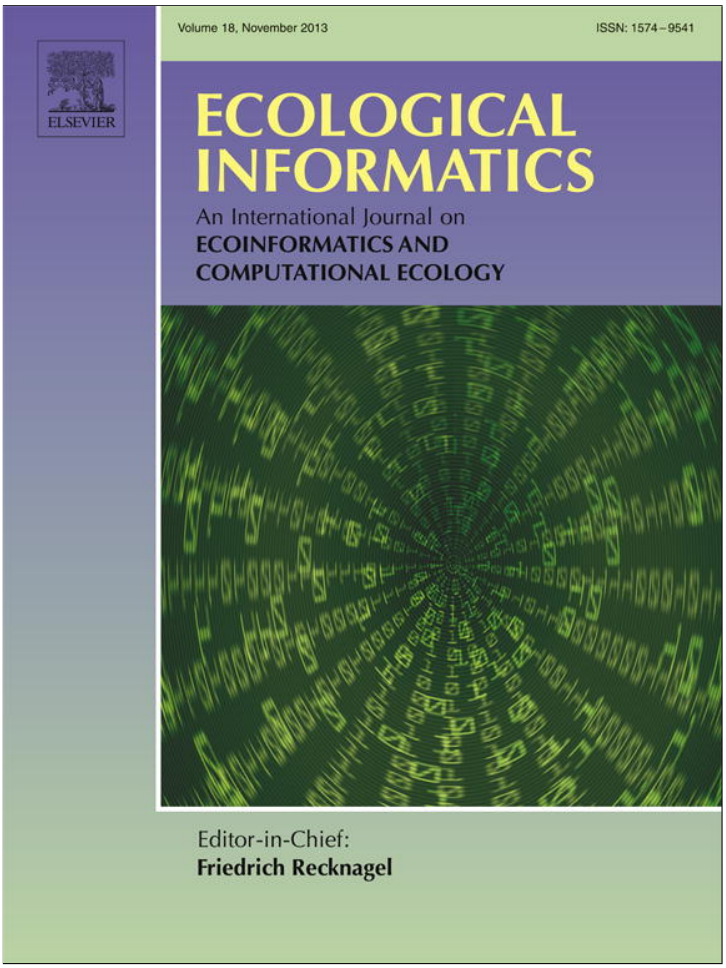

Ecological Informatics 18 (2013) 194–202

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

EcoIS: An image serialization library for plot-based plant flowering phenology

Joel A. Granados [a,*], Philippe Bonnet [a], Lars H. Hansen [b], Niels M. Schmidt [b]

[a] IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen, Denmark
[b] Department of Bioscience, Arctic Research Center, Frederiksborgvej 399, 4000 Roskilde, Denmark

ARTICLE INFO

Article history:
Received 28 June 2013
Received in revised form 20 September 2013
Accepted 25 September 2013
Available online 18 October 2013

Keywords:
Image series
Ecological monitoring
Phenology
Arctic
Computer vision
Photo-plot

ABSTRACT

Image series are increasingly being used to output ecological indicators because they provide the ability to reanalyze data that has already been collected and they improve temporal as well as spatial resolution. We address both the increased utilization and the need to diversify the way they are produced by introducing an open source Python (www.python.org) library called EcoIS that creates image series from unaligned pictures of specially equipped plots. We use EcoIS to sample flowering phenology plots in a high arctic environment and create image series that later generate phenophase counts and automatically estimate phenological dates of interest. Our results exhibit one day difference between EcoIS estimations of local indicators and the ones calculated with the established field-based process. We show that EcoIS' error is similar to the one of image series generated with fixed camera setups. We see that EcoIS processes an image in 3.8 s and show how it is equipped to handle data intensive scenarios. We additionally identify in-camera automatic image formatting and image acquiring oversight as contributing factors for increasing the overall error. Our main conclusion is that EcoIS creates usable image series that maintain the spatiotemporal qualities of the original images and can successfully be utilized to generate ecological indicators. EcoIS is relevant as a replacement for traditional image series infrastructure where the cost of deploying EcoIS is smaller or less intrusive to the ecosystem.

1. Introduction

Arctic ecosystems are changing rapidly and dramatically along with observed changes in climate (Gilg et al., 2012; Post et al., 2009). Quantifying the interaction between climate and ecosystems is complex and requires extended concurrent data collection from multiple compartments of an ecosystem (Meltofte et al., 2008a). Our ability to understand and predict the effects of climate change on ecosystems depends on coordinated long-term monitoring programs (Schmidt et al., 2012b) that establish the difference between effects inherent to ecosystems and those caused by environmental perturbations like climate change (Meltofte et al., 2008a), as well as provide context for interpreting experiments conducive to designing, implementing and evaluating environmental policy (Lovett et al., 2007).

Technology has long aided long-term monitoring by providing solutions like tape recorders that facilitate documentation, personal desktop assistants (PDA) that automate data transmission and loggers that collect data for long periods of time (Michener et al., 2011). Recent developments in digital photography have broadened this scope by facilitating projects that range from phenological event detection (Richardson et al., 2007) to demographics (Bolger et al., 2012).

Image series (ISeries) are of special interest because they provide the ability to reanalyze data that has already been collected and can improve spatial and temporal resolution of long-term monitoring variables while at the same time reduce labor (Ide and Oguma, 2010). Plant phenology which is an observable trait impacted by climatic variations (Badeck et al., 2004; Høye et al., 2007) vital for understanding species responses, ecosystem function and the effects of climate (Wright et al., 1999) has been detailed by ISeries (Graham et al., 2009; Richardson et al., 2007) and has been related to measurements such as carbon dioxide uptake (Mizunuma et al., 2013) and gross primary production (Saitoh et al., 2012).

In general, ISeries are generated from cameras placed in housing platforms designed to provide stability, power and protection. Housing platforms that have been used to fix cameras close to the ground in order to measure leaf area index (Ryu et al., 2012), have enclosed cameras that generated simple field estimations of photosyntheses (Graham et al., 2006) and have positioned cameras near and above forest canopies (Sonnentag et al., 2012; Zhao et al., 2012). They provide translational (Graham et al., 2009) as well as rotational (Granados et al., 2013) movement increasing the amount of possible arrangements. A housing platform is what aligns all images by providing the same view point.

But what if a housing platform cannot be deployed? As we move monitoring efforts away from well known infrastructure (electricity grids and communication networks) into remote areas where powering and maintaining equipment are expensive and resource intensive, we

* Corresponding author.

1574-9541/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ecoinf.2013.09.006

encounter situations where the cost of a camera housing platform is prohibitively expensive. Moreover, the inherent invasive nature of housing platforms might prevent their deployment in research stations where invasive structures are prohibited inside denominated undisturbed areas (Meltofte et al., 2008b). It is in these situations where we need an alternative way to create ISeries.

In these cases we replace the invasive housing platform with a more versatile approach that uses autonomous mobile entities, like humans or unmanned autonomous vehicles (UAV) as vessels that transport and actuate a camera. Because of their mobile nature, they would generate unaligned images which would not be suitable for generating ISeries. Projects have addressed this by manually aligning the images (Liang et al., 2012), but this becomes unmanageable as the amount of images increases.

Inspired by the need to generate ISeries from unaligned images we have created EcoIS (Granados, 2010a), a Python (http://www.python.org) library that automates the alignment process and suggests an alternative way of gathering data based on ISeries. We give a detailed description of its inner workings and outline how it was used in an existing long-term phenological monitoring program located in a high arctic research station. We characterize changes in the arctic monitoring work-flows by describing their established processes and comparing them to new ones brought on by the use of EcoIS. We see how ecological indicators (from established work-flows) can be created and demonstrate how to generate other spatiotemporal measurements that can only proceed from ISeries. Finally, we see how errors intrinsic to camera systems and image transformations affect ISeries and the indicators that they generate.

2. Materials and methods

2.1. EcoIS

EcoIS (Granados, 2010a) is an open source Python library that creates ISeries from images taken from regions of interest delineated by special markers. It aligns images giving them a unique virtual view point which is similar to creating orthophotos (Duhaime et al., 1997). It creates ISeries which are the foundation for spatiotemporal analysis used in ecological indicator calculations.

2.1.1. Photo-plot layout

Photo-plot layout (Fig. 1B) is square and enclosed by three spherical (Fig. 2A) and one chessboard (Fig. 2B) markers. Spheres mark three of the plot corners (center of chessboards marks the fourth) while chessboards contain information used for plot identification (Samples and ID sections, Fig. 2B) as well as sphere detection (Sphere section, Fig. 2B). All markers are fastened to the ground for the duration of deployments.

The Samples section contains squares depicting six possible colors that appear in the ID section (Fig. 2B) two of which were not used in our implementation but were left in the chessboard for future use. Each square from the ID section encodes two bits and contains part of the plot ID representation (Fig. 3). The amount of squares in the ID section depends on the number of plots being identified (greater number → more ID squares). For example, Fig. 2B shows three squares in the ID section which encode to 35 (see Fig. 3 for encoding calculation) and have the potential to identify 64 plots.

2.1.2. Serialization algorithm

EcoIS begins searching for a chessboard (Fig. 3) by using OpenCV's *findChessboardCorners* function (Bradski and Daebler, 2008a) which operates in five steps (Rufli et al., 2008): 1) Images are converted to grayscale. 2) They are then segmented by applying adaptive thresholding which binarizes the images (Bradski and Daebler, 2008b). 3) The binary image is then eroded (Jähne and Haußecker, 2000) with a 3×3 rectangular kernel which is gradually increased when quadrangles (squares) are difficult to detect. 4) Closed contours (Bradski and Daebler, 2008c) are then calculated which the algorithm uses to fit into quadrangles. 5) Finally, every successfully fitted quadrangle is linked to adjacent ones (Rufli et al., 2008). If a chessboard is not found after these steps, the image is discarded and EcoIS continues with the next image.

After the chessboard is found, the plot number contained in the ID section (Fig. 3) is calculated. First, the HSV color space (Smith, 1978) is segmented into six consecutive compartments based on mean hue

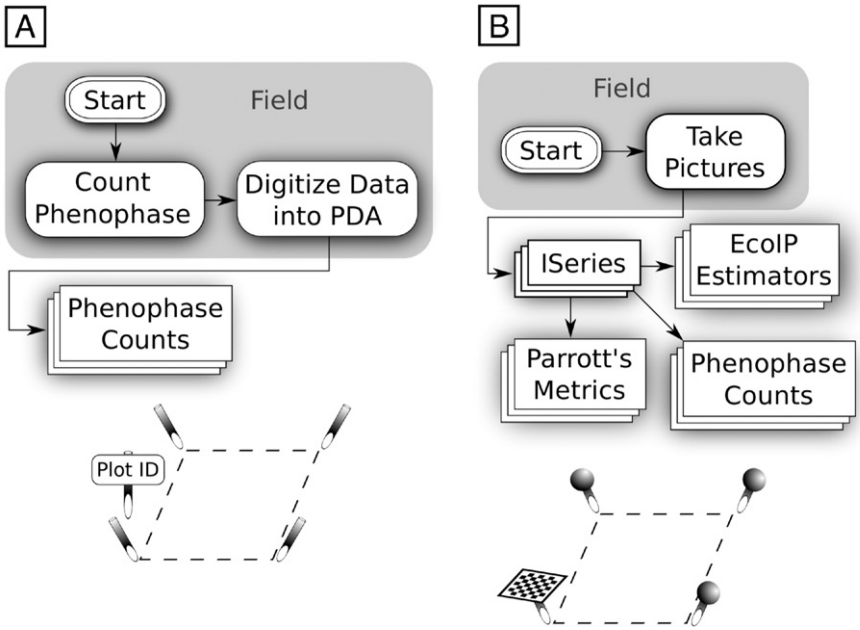

Fig. 1. Established and photo-plot work-flows. A) Established work-flow. Counting and digitizing of phenophases are manually done in the field at a predefined frequency. Plot layout is rectangular of varying size marked by stakes driven into the ground at each corner. Next to each plot there is a sign containing plot ID. B) Photo-plot work-flow. Digital images are taken from several view points at a predefined frequency. ISeries are automatically created by EcoIS from field images which in turn are used for phenophase counts, EcoIP estimators and Parrott's metrics. Plot layout is rectangular enclosed by three spherical and one chessboard markers which are all driven into the ground. Chessboard marker contains plot ID.

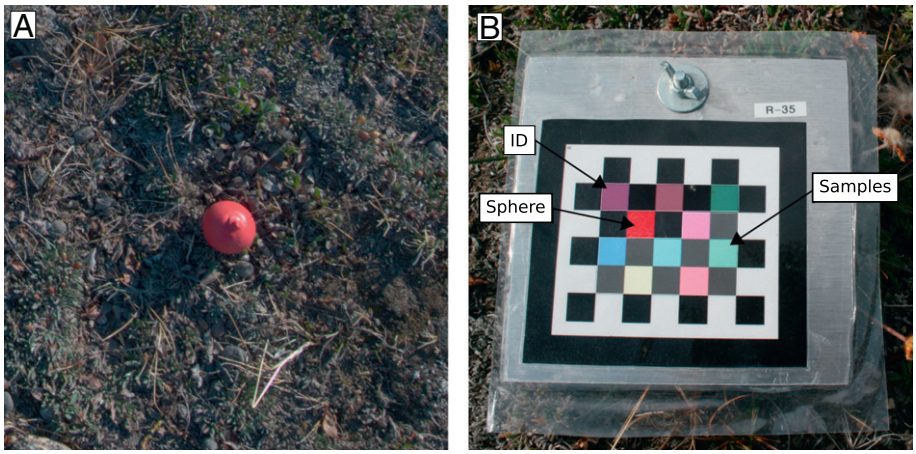

Fig. 2. Closeup of markers in the field. A) Sphere marker painted with water proof paint. B) Chessboard marker is printed on paper, placed on an aluminum base and laminated to make it waterproof. This is a marker for plot 35 containing red spheres. Samples section is used to model squares containing encoding colors. Sphere section holds sphere color. ID section has encoding squares detected by models from Samples section.

values of the Samples squares (Fig. 4A). We then use this segmentation to count the hue values of the ID squares and choose the color based on the segment where the count is greater (Fig. 4B). Sampling of colors prior to detecting is done for every image in order to adjust for pixel value variability in the ID section caused by lighting differences and automatic camera adjustments.

To find the sphere centers (Fig. 3) we first sample the Sphere square (Fig. 2) and calculate mean ($\bar{c}$) and standard deviation (σ_c) for color components in CIE L*a*b* (CIE (Commission Internationale de l'Eclairage), 1986) color space. We then create a range containing the sphere color ($\bar{c} \pm \sigma_c$) and use it to binarize the image. We continue by applying consecutive *open* and *close* operations (Jähne and Haußecker, 2000) with a circle shaped kernel of varying sizes that reduces noise and makes the spheres more prominent. Finally, we use OpenCV's *HoughCircles* function which uses Hough circle transform (Bradski and Daebler, 2008d) to approximate the circle centers. If EcoIS does not find exactly three spheres, the image is discarded.

The final step in our algorithm consists in moving all the pixels in the image in such a way that the corners of the plots (spheres and chessboard) are in the image corners (Fig. 3). This involves multiplying all coordinates by a transformation matrix and then performing

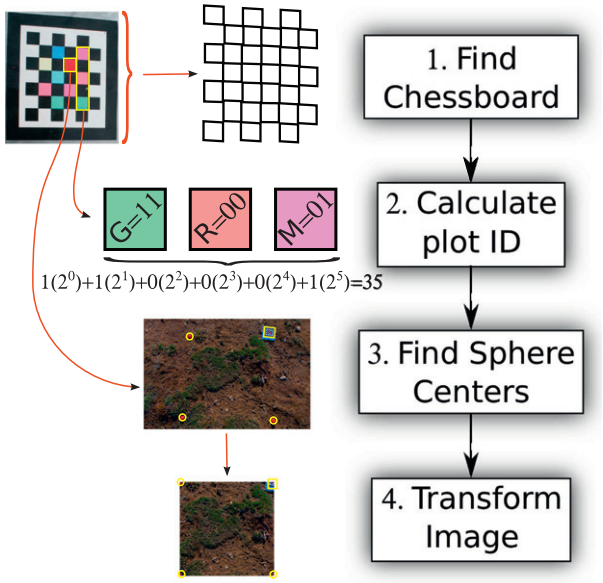

Fig. 3. Serialization algorithm. First step: Find a chessboard in each image. Second step: Calculate ID from ID section. Each color encodes two bits: green encodes 11, red 00 and magenta 01. They each represent a power of 2 and sum to 32. Third step: With color sample from Sphere section, find sphere centers. Fourth step: Transform image so all centers are placed in the image corners.

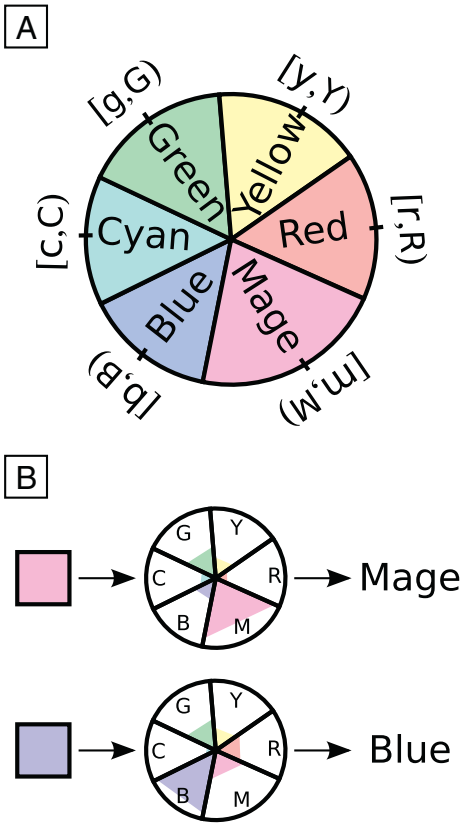

Fig. 4. Chessboard color detection. A) For each of the Samples squares we calculated min value (lower case), mean (midrange cue) and max value (upper case) effectively segmenting the hue dimension of the HSV color space (Smith, 1978). The "Mage" abbreviation means magenta. B) Each ID square is binned into ranges and the one with greater value is selected. The first is detected as magenta and the second as blue.

pixel value interpolation on the result. The matrix is calculated by solving $X_i^{'} = M \cdot X_i; i = 0, 1, 2, 3$ where M is the transformation matrix, X_i is the i^{th} plot coordinate and $X_i^{'}$ is the i^{th} destination coordinate. For our particular case the plot coordinates are the image coordinates where the spheres and chessboard are located; the destination coordinates are the re-projected image corners (coordinates for a 5000×5000 pixel image are {(5000,0); (5000,5000); (0,5000); (0,0)}). Once the matrix is calculated we use OpenCV's *warpPerspective* function (Bradski and Daebler, 2008d) to actually move the pixels and perform the interpolation.

These four steps (Fig. 3) take the view point of the original image and re-project it into a virtual one which is shared by all re-projections of the plot (Fig. 5). The new image is scaled to a resolution of 5000×5000 pixels and indexed into a directory using its plot ID. Additional sanity checks are executed to avoid images where corner detection errors occurred.

2.1.3. Photo-plot work-flow

Plots are created by arranging three spherical markers and a chessboard in a square containing elements of interest (Fig. 1B). Spheres are painted with the same color as the Sphere section (Fig. 2B) which should contrast with the elements in the plots to avoid interference. Images are taken from only four markers, if a fifth (sphere or chessboard) is visible in the frame, it might interfere with correct detection and could contribute to inaccuracies. Moreover, plots should occupy as much of the frame as possible in order to maximize the amount of pixel information available for markers and for elements of interest.

Plots are photo sampled (photographed) at a predetermined frequency taking care of acquiring several photographs from different view points per photo sample in order to increase the possibility of at least one image getting serialized (Fig. 1B). When we have finished collecting images, EcoIS is executed in order to create an ISeries for each plot which can subsequently be used to explore the behavior of the elements of interest. This work-flow is denominated photo-plot work-flow (Fig. 1B).

EcoIS requires an image based storage system which has its own challenges, yet having the data in an ISeries representation creates a new potential in terms of data usability. ISeries not only provide the possibility for re-measuring plots (for corroboration), it also allows scientists to ask and answer new questions given sufficient image quality. We demonstrate this in an arctic deployment by using ISeries to create three types of ecological indicators (Fig. 1B).

2.2. Field deployment at Zackenberg

We deployed in 2012 at Zackenberg station in northeast Greenland (74°30′N, 20°30′W), a high arctic research station run by the Department of Bioscience at Aarhus University in Denmark. There, we focused on the established plant flowering phenology work-flow which has kept track of phenological phases of interest (phenophases) by doing seasonal counts (Fig. 1A). The process is mostly done in the field where technicians count individuals in a plot and then digitize the information into a PDA (Fig. 1A), they do this once every week for the duration of the season. This process is carried out for 28 permanent monitoring plots located in the valley lowland and produces files containing yearly plot phenophase counts (Fig. 1A).

2.2.1. Photo-plot layout

We implemented the EcoIS photo-plot layout (Section 2.1.1) with spheres of 40 mm in diameter (Fig. 2A) and chessboards that measured 110×90 mm (Fig. 2B) We used water resistant paint on the spheres and laminated the chessboard marker with transparent plastic (Fig. 2B) to prevent weather damage. We used a seven by six chessboard that allowed us to track 64 plots (we deployed nine) and placed it on top of an aluminum plate to prevent it from deforming. Stakes were screwed on all markers and driven into the ground to hold them in place. Plots had dimensions of 80×80 cm and were captured from an approximate height of 125 cm with an average focal length of 26.0 mm (35 mm equiv).

2.2.2. Photo-plot work-flow

We applied EcoIS' photo-plot work-flow (Fig. 1B) by visiting (weekly) nine plots containing Mountain Avens (*Dryas octopetala*/*integrifolia*; hereafter referred to as *Dryas*) from Day Of Year (DOY) 167 to 219 which generated a total of 60 photo samples with an average of four images per sample. Here it is important to distinguish between photo samples and individual images: the first is a series of consecutive images taken of a plot from different view points on one specific date, the second is just one picture. Notice that there are redundant images per photo sample to compensate for EcoIS' raw error.

A total of 265 images were taken with three cameras (Table 1) which produced raw files that were formatted into Joint Photographic Experts Group (JPEG; ITU, 1992) with a raw image processing application (RIPA) called Rawtherapee (www.rawtherapee.com). The Sony camera generated raw and JPEG versions which allowed us to add 108 Sony generated JPEGs bringing the total to 375.[1] Plots were visited between ten in the morning and seven in the afternoon. Images were taken by avoiding shadows on markers, avoiding positions where chessboards reflected the sun and including only four markers.

EcoIS automatically created all the ISeries on a Lenovo W500 (Model W500, Lenovo, Morrisville, North Carolina) laptop with 8 GB of memory and an Intel® Core™ Duo (2.66 GHz) processor. After analyzing all the JPEGs, EcoIS had effectively put all images into their respective ISeries and put all discarded images into an error directory. The photo-plot work-flow ended by counting the phenophases of interest on the created ISeries in an office back in Denmark.

2.2.3. Calculating error

We calculated two types of error related to discarded images: raw error and ISeries error. The first refers to the total number of discarded images which points to how often EcoIS fails but does not reflect the proportion of missing photo samples. The second refers to the total amount of missing photo samples and increases when EcoIS fails to serialize all the images of a photo sample. The ISeries error expresses missing data that cannot be reclaimed by the photo sample redundancy and gives us an idea of the impact of taking multiple images. Additionally we looked at image quality by measuring the virtual movement related to OpenCV's *warpPerspective* function (Bradski and Daebler, 2008d) by following specific elements throughout an ISeries and calculating their Euclidean distance from image to image.

We calculated missing and movement values from ISeries created with a pan–tilt–zoom (PTZ) camera (Model VB-C50iR, Canon U.S.A., Inc., Lake Success, New York) placed 30 m above ground and compared them to the ones from EcoIS ISeries. Given that the camera configurations were different we normalized the focal length, distance to objects and crop factor using a simple pinhole camera model (Bradski and Daebler, 2008a) in order to compare the EcoIS images with the PTZ ones (model not shown). For the movement comparisons we only considered inanimate objects that did not grow such as rocks, markers or pebbles. Finally we characterized the impact of using camera formatted JPEGs in the photo-plot work-flow by comparing the success rate of images formatted using a RIPA with images formatted using the Sony camera. We report the number of rejected images in both of these cases.

2.2.4. Using image series

Parrott's three dimensional metrics (Parrott et al., 2008) are based on a stack of successive spatial images sampled at uniform intervals

[1] 216 Sony images: 108 formatted by Rawtherapee and 108 formatted by the camera.

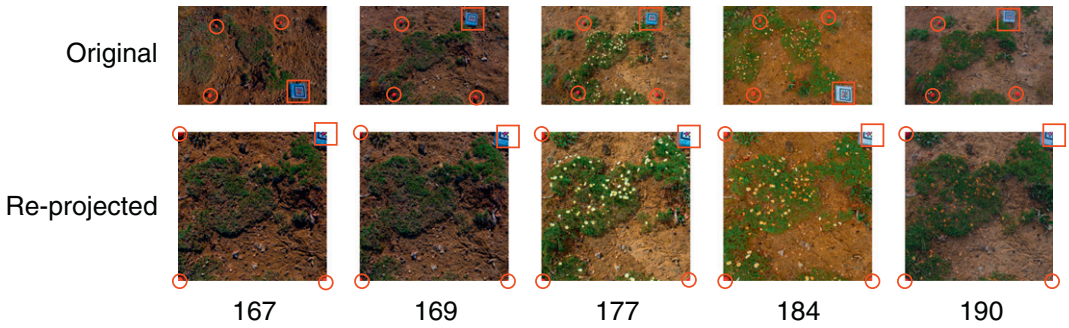

Fig. 5. Plot image originals and their transformations. The first row contains original images taken for each DOY of plot four. The second is the transformed images using EcoIS algorithm (Fig. 3). Notice that there are no shadows on the markers in the original images. DOY 167 and 169 contain *Dryas* buds (too small to see), DOY 177 shows the plot flowering and DOY 184 and 190 show the senescent stage of this species.

called *space-time cubes*. The smallest constitutional units are voxels which are regular pixels with an added temporal dimension that, when put together, make up three dimensional blobs (fig. 1b in Parrott et al., 2008). The two ways of making blobs are: with 26-voxel (Moore's system) and with 6-voxel (Von Neumann's system) neighborhoods. These constructs are used to characterize three-dimensional data sets by calculating metrics designed to represent a property. Among the ones described in (Parrott et al., 2008) there is shape complexity defined as the ratio of blob volume to bounding box volume which ranges from zero (complex shapes) to one (simple rectangular shapes). There is also contagion measuring the dispersion of blob types and ranges from zero (random mix of voxel types) to one (contiguous landscape with little change). Finally, there is spatiotemporal complexity which measures how one type of blob occupies the *cube* and also ranges from zero (uniform blob shapes) to one (complex and random shapes).

To demonstrate EcoIS' usability we applied Parrott's metrics to see how well they describe the resulting ISeries. We created an intermediate representation of our ISeries data (*space-time cube*) by stacking binary versions of individual images into a three dimensional *cube* (we used Moore's system). We then took this *space-time cube* and used it to run Parrott's code which ultimately gave us values for each of the metrics.

EcoIP (Granados et al., 2013) is a toolkit used for characterizing phenophases of different species which bear distinct colors, like the tones of leaves in the fall. It involves creating a statistical model with a set of training images and applying it to an ISeries to produce a representative signal which is then analyzed by a semiautomatic process that ends with the estimation of beginning and ending dates. It is based on a Naive Bayesian model of color values applied to ISeries images that produces temporal estimators by finding the inflection points of fitted sigmoid signals. Among its outstanding features is the ability to use a variety of color transformations to adjust the accuracy of the estimations.

To further demonstrate the quality of the generated ISeries, we estimated flowering periods with EcoIP (Granados et al., 2013). Of all the deployed photo-plots we chose plot four because it had a high sample number and a missing value of zero (Table 2). We trained a model with the characteristic yellow of the flowering *Dryas* and used that to generate the representative signal. The semiautomatic procedure described in (Granados et al., 2013) then estimated the beginning and ending dates of flowering periods.

Finally, to demonstrate that no functionality was lost with the photo-plot work-flow (Fig. 1B), we generated season counts by identifying phenophases of interest (buds, flowers and senescent) in ISeries with an annotation tool (Granados, 2010b). These phenophase counts (new work-flow, Fig. 1B) were validated by two trained field ecologists and were compared with field counts by looking at the actual counts (Table 3) as well as 50% flowering and senescence onset estimators (Table 4; Høye et al., 2007, 2013; Iler et al., 2013).

3. Results

3.1. General

EcoIS spent a total of 962 min analyzing a subset of 248 images. On average, it spent 3.8 min per image which meant that we had to leave it overnight to process all the 373 images. It spent more time on error images as the algorithm was designed to try different detection configurations before discarding an image. Given the approximate

Table 1
Cameras in Zackenberg. Characteristics of cameras used in our Zackenberg deployment. Resolution is given in pixels, focal length is given in millimeters and is normalized to 35 mm equivalence, exposure is given in seconds and "Imgs" gives the number of images taken with each camera. Formats refer to how the images were formatted in camera; only the Sony camera kept each image in raw and JPEG (ITU, 1992) formats.

	Sony Nex-3	Nikon D700	Nikon D300
Lens	Sony E 18–55mm *f*/3.5–5.6	AF-S Nikkor 14–24mm *f*/2.8G ED	AF-S Nikkor 14–24mm *f*/2.8G ED
Resolution	4608 × 3072	4288 × 2844	4352 × 2868
Aperture	*f*/9.0–22.0	*f*/10.0–11.0	*f*/8.0–11.0
ISO	200	200–500	200–640
Exposure	1/320–1/8	1/500–1/125	1/500–1/160
Focal length	27	24	25–27
Imgs	108	56	101
Flash	No	No	No
Formats	JPEG/raw	Raw	Raw

Table 2
Missing samples. EcoIS refers to Zackenberg ISeries. PTZ refers to ISeries gathered in the James Reserve by a pan–tilt–zoom camera. ID is the plot ID for EcoIS and the position ID for PTZ. Sample is the number of intended samples. Missing is the number of samples that went missing. The last row is the error calculated as the number of discarded images proportional to the number of intended ones.

EcoIS			PTZ		
ID	Sample	Missing	ID	Sample	Missing
4	8	0	0	365	88
35	8	0	6	365	82
59	4	0	1	365	35
0	8	1	2	365	25
16	4	1	3	365	44
20	4	1	4	365	34
24	8	1	5	365	76
47	8	2	7	365	89
63	8	2	8	365	82
Error: 13.33%			Error: 16.86%		

Table 3
Dryas flowering phenology counts. Values are from plot four. DOY is Day Of Year. F is the field counts. I is the count done on ISeries minus field counts where a plus (+) is for larger ISeries , a minus (−) is for smaller and (0) when there is no change. Buds are flowers that are not yet open, Flowers are open *Dryas* giving access to their reproductive organs and Senescence is when all the petals turn brown or are missing (Schmidt et al., 2012a). Total is the sum of all the elements and AF is the average absolute difference between the F and I that varied.

DOY	Buds		Flowers		Senescence		Total	
	F	I	F	I	F	I	F	I
167	75	+24	0	0	0	0	75	+24
169	93	+31	2	−2	0	0	95	+29
177	3	+22	108	−12	1	−1	112	+22
184	0	0	22	+14	91	+5	113	+19
190	0	0	1	0	70	+57	71	+57
197	0	0	0	0	105	+10	105	+10
204	0	0	0	0	101	+14	101	+14
211	0	0	0	0	105	−7	105	−7
AF	25.6		7		15.6		22.75	

view point height of 125 cm, we calculated an average distance to the markers of 150 cm which produced images containing spheres of 100 pixels in diameter and chessboard markers with dimensions of 170 × 200 pixels.

3.2. EcoIS error

Raw error was the number of discarded images (180) proportional to the total number of images (373) and was calculated to be 48.25% (Table 5). The error for the Sony images formatted by the Sony camera was 67.59% (Table 5). Image taken with the Sony camera formatted by the raw image processing application (RIPA) had error of 42.59% while the ones taken with the Nikon cameras had an error of 38.85% (Table 5). The error of the images formatted only with the RIPA was 40.37% (Table 5).

Of the 60 photo samples that should have produced 60 images for the ISeries, EcoIS produced 52 (eight missing) which represented an ISeries error of 13.33% (Table 2). The other 86.66% were correctly identified, re-projected and placed in a directory as an ISeries image. Virtual movement for the EcoIS generated ISeries was, on average, 1.85% of the image size which represented a distance of 130.85 pixels. PTZ objects, on the other hand, moved an average distance of 0.31 of the image size.

3.3. Image series

We calculated the number of blobs, shape complexity, contagion and spatiotemporal complexity (Table 6) with Parrott's metrics (Parrott et al., 2008). We also applied EcoIP's semiautomatic process to the data of plot four which resulted in estimations of the beginning and ending dates of the *Dryas* flowering period (Fig. 6). We further used the field and ISeries counts to calculate and compare 50% flowering onset and 50% senescence onset where the duration of the *Dryas* flowering period was the same (8 days) with both count types (Table 4). Finally, we compared field counts and ISeries counts from plot four and calculated the average differences in numbers to be 25.6, 7 and 15.6 for the bud, flowers and senescence counts respectively which averaged to 22.75 for the whole plot (Table 3).

Table 4
50% phenophase onset counts. Timing of phenological events in *Dryas* using field counts and images series counts, respectively. Values represent DOY. Field numbers come from counts done in the field. ISeries numbers come from counts done on the ISeries. Flowering onset is the DOY when 50% of the plot flowered. Senescence onset is the DOY when 50% of the plot became senescent. Duration is (senescence onset) − (flowering onset).

	Field	ISeries
Flowering onset	173	174
Senescence onset	181	182
Duration	8	8

Table 5
Raw error. Img is total number of images, Discarded is the amount of discarded images and Err is Discarded/Img. The row labeled Sony by Sony contains the JPEG files that were taken and formatted by the Sony camera. The Sony by RIPA and Nikon by RIPA are image taken by the Sony and Nikon cameras and formatted by the raw image processing application (RIPA) Rawtherapee (www.rawtherapee.com). The fourth row contains all RIPA formatted JPEGs and the last row consolidates the first three rows to display the total number of analyzed JPEGs.

	Img	Discarded	Err
Sony by Sony	108	73	67.59%
Sony by RIPA	108	46	42.59%
Nikon by RIPA	157	61	38.85%
Sony & Nikon by RIPA	256	107	40.37%
Total raw error	373	180	48.25%

4. Discussion

4.1. Applicability of image series

4.1.1. Phenophase counts

There was a tendency to undercount buds in the field (Table 3) due to the difficulty of seeing them at an early stage (buds are small). On the other hand, detection was facilitated on ISeries because technicians could identify the position and state of a bud by referring to images in the past and in the future. In other words, technicians had information in images from different dates which hinted at the location of very small elements; if one image in an ISeries contained an element in one location, the other images probably had the same element in the same location.

The average difference between the field and ISeries counts dropped from 25.6 for the buds to 7 for the flowers. This was because the white flowers were easier to detect on a dark green background in both the field and the ISeries scenarios. This difference went back up to 15.6 for the senescent counts which were more difficult to spot in the field as the *Dryas* turned brown and blended with the background.

Despite the differences in the counts for all *Dryas* phenophases, 50% flowering onset and 50% senescence onset for both the field and the ISeries counts coincided well with a difference of just one day (Table 4) and the duration of the flowering period (from flowering onset to senescence onset) was identical in both cases (Table 4). This gives credibility to results generated with EcoIS ISeries and suggests that the ecological indicators are contained in the ISeries themselves.

4.1.2. Parrott's algorithm

Parrott's algorithm (Parrott et al., 2008) generated metrics that accurately described the spatiotemporal form of the Zackenberg plots. The shape complexity mean value (Table 6) was greater than 0.6 and points to simple blob complexities where the blob shape tends to fill

Table 6
Parrott's three dimensional metrics (Parrott et al., 2008). All values except Number of blobs range from 0 to 1. Number of blobs is the amount of blobs in the *space-time cube*. Shape complexity ranges from 0 (complex ratio) to 1 (simple ratio). Contagion ranges from 0 (random blobs) to 1 (continuous blobs). Spatiotemporal complexity (STC) ranges from 0 (uniform shapes) to 1 (complex shapes).

	Value
Number of blobs	261
Shape complexity mean	0.664
Contagion	0.656
STC	0.282

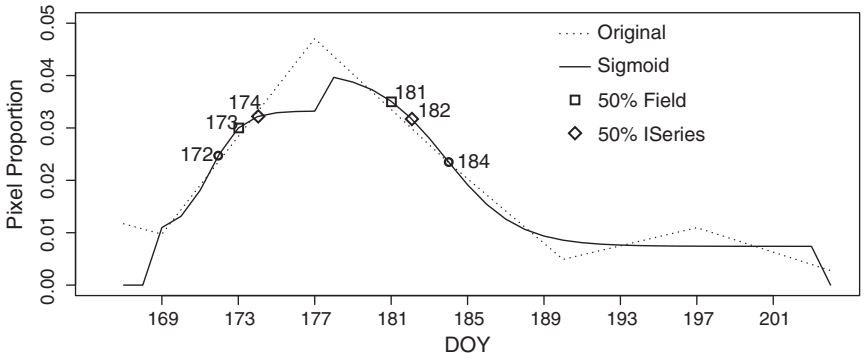

Fig. 6. Phenophase estimations. Sigmoid fit (solid line) to the original signal (dotted line) was calculated by EcoIP (Granados et al., 2013) of plot four containing *Dryas*. Training was done with images from other plots of the 2012 season. Circles are estimated from the inflection points of the sigmoid and represent beginning and ending days of year (DOY). Squares show 50% onset flowering and senescence calculated from field counts. Diamonds show 50% onset flowering and senescence calculated from ISeries counts.

its bounding box. This agrees with elements in the Zackenberg plots moving little throughout their three stages (bud, flower and senescent) effectively creating a cylinder in the *space-time cube*. A contagion value of over 0.6 was expected as values that tend to one are considered to originate from continuous blobs which should be formed when we stack our images one on top of the other. Though we would expect a value closer to one due to the cylindrical nature of the *Dryas* elements in the *space-time cube*, the result tended towards 0.5 due to the virtual movement caused by the *warpPerspective* function (Bradski and Daebler, 2008d) which separates blobs that should otherwise be together. This inadvertent separation also caused the number of blobs (261) to be greater than the maximum amount of elements present in the plot (134, Table 3); these two values should be closer as each three dimensional blob is supposed to represent a *Dryas* element in the plot. Finally we have the spatiotemporal value of 0.28 which agrees with the uniform shape characteristic seen in the *Dryas* plots. In other words, instead of having complex patterns in the *space-time cube* we see long rectangle shapes which are the longest in the time axis.

4.1.3. EcoIP

EcoIS was able to provide EcoIP (Granados et al., 2013) with ISeries that produced an estimation of the beginning and ending flowering dates for plot four (Fig. 6). This shows that EcoIS generated ISeries maintain pixel positional coherency and pixel color values even after going through OpenCV's *warpPerspective* transform (Bradski and Daebler, 2008d). The estimated onset date fell between the day where most of the *Dryas* were still buds (DOY 169) and the day where most of them had bloomed (DOY 177) which was the same date for the field and the ISeries counts (Table 3). The estimated ending date fell between the day were most of the *Dryas* were flowering (DOY 177) and the date where there were almost no flowers because they were mostly senescent (DOY 190) which again was equal for the field and ISeries counts (Table 3). Finally there were at most four days of difference between the 50% onset field and ISeries values and the ones calculated by EcoIP (Fig. 6) which supports the notion of EcoIS producing usable ISeries that contain spatiotemporal information fit for estimations.

4.2. Accuracy

4.2.1. Serialization error

On average, close to half (48.25%) of the JPEGs analyzed by EcoIS were discarded because of lack of information (Table 5) and that value only dropped to 40.37% when we ignored the Sony created files which had a negative effect on the process (Table 5). If the same proportion of dates was missing from an ISeries, it would have been useless; which is the reason we had multiple images per photo sample. Though we managed to go from the raw error of 48.25% (Table 5) to the ISeries error of 13.33% (Table 2), we also increased the amount of images being analyzed which, in turn, increased the amount of analysis time. This suggests that the raw error is important because its reduction directly translates into the reduction of the execution time and is relevant for the user experience of EcoIS.

An ISeries error of 13.33% (Table 2) meant that 86.66% of the dates were correctly serialized which was enough to characterize the season using EcoIP (Fig. 6), Parrott's three-dimensional metrics (Table 6) and the 50% onset values (Table 4). This error (13.33%) was lower than the PTZ deployment error (Table 2) of 16.86% and though they did not have the same cause, they could be compared as they both represented missing images. This comparison was relevant because it showed that an ISeries is still useful despite a 13.33% ISeries error and showed that EcoIS generated ISeries could be used for ecological analysis in the same way as the ISeries generated with PTZ cameras in Granados et al. (2013).

We experienced additional serialization error related to automatic camera formatting of variables like white balance, saturation and contrast (Table 5). We could clearly visualize the effect that the Sony camera had on the error by comparing the first two rows of Table 5 where both represent JPEGs generated from the same raw files yet have very different behavior. This points towards the use of raw formats for automating phenology as the better choice over camera generated JPEGs which have great variability because of diversity in manufacturers, formatting variables and user customizations. The use of Rawtherapee (www.rawtherapee.com) allowed the standardization of raw image formatting variables reducing the error in EcoIS but also restricted the type of cameras that we could use to those that could produce raw files.

Finally, EcoIS' ability to analyze images depends on the lighting on the chessboard being similar to the one on the spheres and although we tried to avoid it, 24 images were excluded due to the photographer casting shadow on at least one marker. None of these were correctly serialized and represented 13.33% of the total error. This implies that we can potentially reduce the error by 13.33% if we are more careful when acquiring the images.

4.2.2. Virtual movement

As expected from the outset, the amount of movement in ISeries generated with the PTZ camera was less than the ones generated by EcoIS. Though it has three degrees of freedom, the PTZ camera was able to return to a predefined position which resulted in a virtual movement of 0.31%. EcoIS, on the other hand, was less accurate (Fig. 7) with a virtual movement value (1.85%) that was six times larger. This greater variability (Fig. 7) did not overly affect the phenophase counts as individual elements could still be identified based on their

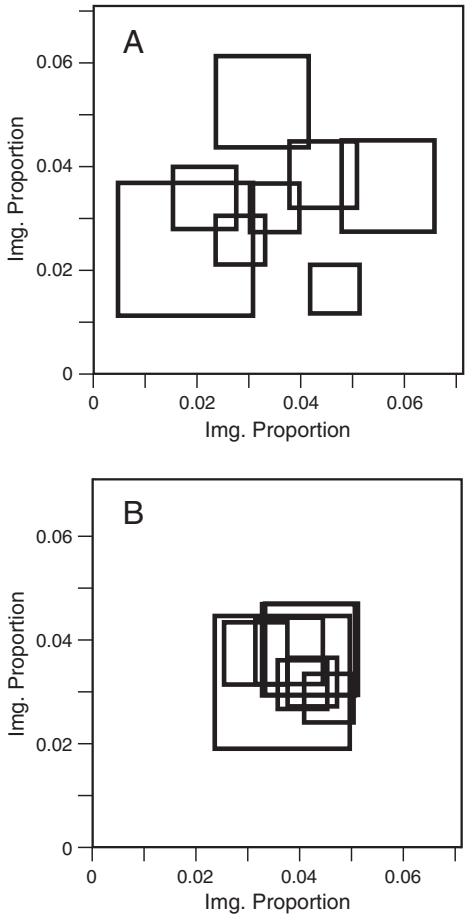

Fig. 7. Virtual movement. A) Squares represent the position of the same flower throughout an EcoIS generated ISeries. B) It is the movement in A calculated with the movement values of the PTZ cameras. Both plots are a subset of a complete image and the axis is given proportional to the total size of the original image. This figure shows the difference in type of movement between EcoIS and PTZ cameras.

locations in prior and posterior images. Moreover an average difference of 130 pixels (1.85%) is not significant when compared to the resolution of the original images (Table 1) or the generated re-projections (5000 × 5000 pixels).

The analysis done with EcoIP (Granados et al., 2013) was also mostly unaffected by virtual movement (130 pixels) as its temporal estimations are based on complete images instead of regions of interest. In contrast, Parrott's (Parrott et al., 2008) numbers were affected where estimators like contagion were expected to be closer to one but ended up being 0.656 (Table 6) due to blob separation caused by virtual movement. This was addressed by using Moore's neighborhood (Parrott et al., 2008) which reduced the chance of blobs getting separated due to the increased number of adjacent voxels.

4.3. Deploying EcoIS

4.3.1. In the field

We looked at the workload already in place (Fig. 1A) and revised it in order to increase automation while at the same time reduced the number of steps at the field (Fig. 1B). We replaced the two steps needed to service a plot (counting and digitizing, Fig. 1A) with just one (imaging, Fig. 1B) and in the same way replaced the hardware needed for the old work-flow (PDA and mechanical counters) with one camera. And though this reduction is relevant, it still remains to be seen if EcoIS' overhead is optimal (in terms of time and cost).

4.3.2. Time of analysis

Our experiments show that the algorithm spent 3.8 min (on average) analyzing images taken in eight weeks from nine plots (one photo sample per week per plot). But how would EcoIS behave with more data? How long would it take EcoIS to serialize a hypothetical year-round deployment that measured once every day on nine plots? Having four pictures per photo sample of nine plots we would have 36 pictures taken per day. This would give us a total of 13,140 images per year which would need 49,932 min to be serialized (34.67 days). We can reduce this hypothetical month if we segment the totality of the images and analyze each in a separate process. If we had 16 processors, we would reduce the 34.67 days to just 2.16 and as machines get faster this time will be reduced even more. Additionally these are not man-hours and are times where researchers can do other tasks.

4.3.3. EcoIS scope

Dryas are a few centimeters across when fully bloomed and flower stems usually grow to a height of 5 to 7 cm. This is a pattern that repeats itself across the species at Zackenberg and is a very convenient characteristic for EcoIS because virtual movement is minimized for elements that are close to the ground. If we measured taller species (e.g. shrubs in the low Arctic) the virtual movement caused by OpenCV's *warpPerspective* function (Bradski and Daebler, 2008d) would be too much to follow the plant through the ISeries. Depending on the height it could even block the markers, rendering serialization impossible. Our approach fits comfortably with species found at Zackenberg as well as cases where the studied elements have manageable height as in Graham et al. (2006).

5. Conclusions

We have introduced EcoIS, a toolkit that creates image series by re-projecting and identifying images taken from specially marked plots. We have successfully fitted EcoIS into an established work-flow in a high arctic monitoring station and reduced the amount of steps that were needed to sample field plots. We have shown that the phenophase count differences between our photo-plot work-flow and the established Zackenberg procedure do not affect the 50% onset event interpolation values (flowering onset and senescence onset) used as an ecological indicator. We have demonstrated that in addition to procuring phenological counts, ISeries produced by EcoIS can be used to calculate spatiotemporal metrics (Parrott et al., 2008) and estimate beginning and ending phenophase dates (Granados et al., 2013).

We found that the number of EcoIS discarded images is similar to more traditional PTZ camera setups and documented how camera formatted images can increase this number. We showed that despite the presence of missing images, temporal information and spatial information were maintained in ISeries. We demonstrated that the virtual movement of ISeries created with EcoIS, which was greater than the ones created with PTZ setups, does not impede analysis based on visual inspection nor analysis based on automatic processes. Finally we show how EcoIS can be used in a data intensive scenario by spreading the load among several processors to adjust for the extended time of execution.

Acknowledgments

The authors thank Palle Smedegaard Nielsen and Jannik Hansen for their assistance in the field as well as Aarhus University for providing access to Zackenberg. We are also thankful for receiving funding from the European Union's Seventh Framework Program [FP7/2007–2013] under grant agreement number 262693 [INTERACT].

References

Badeck, F.-W., Bondeau, A., Bottcher, K., Doktor, D., Lucht, W., Schaber, J., Sitch, S., 2004. Responses of spring phenology to climate change. New Phytol. 162 (2), 295–309. http://dx.doi.org/10.1111/j.1469-8137.2004.01059.x (May).

Jähne, Bernd, Haußecker, Horst, 2000. Computer vision and applications, a guide for students and practitioners. Morphol. Oper. 483–515 (Ch.).

Bolger, D.T., Morrison, T.A., Vance, B., Lee, D., Farid, H., 2012. A computer-assisted system for photographic mark-recapture analysis. Methods Ecol. Evol. 3 (5), 813–822. http://dx.doi.org/10.1111/j.2041-210X.2012.00212.x (Oct.).

Bradski, G., Daebler, A., 2008a. Learning OpenCV. Computer vision with OpenCV library. Camera Models Calibration 370–404 (Ch.).

Bradski, G., Daebler, A., 2008b. Learning OpenCV. Computer vision with OpenCV library. Image Process. 109–143 (Ch.).

Bradski, G., Daebler, A., 2008c. Learning OpenCV. Computer vision with OpenCV library. Contours 222–264 (Ch.).

Bradski, G., Daebler, A., 2008d. Learning OpenCV. Computer vision with OpenCV library. Image Transforms 144–192 (Ch.).

CIE (Commission Internationale de l'Eclairage), 1986. Colorimetry, Tech. Rep., Central Bureau of the Commission Internationale de l'Eclairage, Vienna, 15.22nd ed. CIE Publication.

Duhaime, R.J., August, P.V., Wright, W.R., 1997. Automated vegetation mapping using digital orthophotography. Photogramm. Eng. Remote Sens. 63 (11), 1295–1302.

Gilg, O., Kovacs, K.M., Aars, J., Fort, J., Gauthier, G., Grémillet, D., Ims, R.A., Meltofte, H., Moreau, J., Post, E., Schmidt, N.M., Yannic, G., Bollache, L., 2012. Climate change and the ecology and evolution of Arctic vertebrates. Ann. N. Y. Acad. Sci. 1249, 166–190 (Feb., URL http://www.ncbi.nlm.nih.gov/pubmed/22329928).

Graham, E.A., Hamilton, M.P., Mishler, B.D., Rundel, P.W., Hansen, M.H., 2006. Use of a networked digital camera to estimate net CO_2 uptake of a desiccationtolerant moss. Int. J. Plant Sci. 167 (4), 751–758 (Jul., URL http://www.jstor.org/stable/10.1086/503786).

Graham, E.A., Yuen, E.M., Robertson, G.F., Kaiser, W.J., Hamilton, M.P., Rundel, P.W., 2009. Budburst and leaf area expansion measured with a novel mobile camera system and simple color thresholding. Environ. Exp. Bot. 65 (2-3), 238–244 (Mar.,URL http://linkinghub.elsevier.com/retrieve/pii/S0098847208001214).

Granados, J.A., 2010a. Ecological Image Serilizer (EcoIS). https://github.com/Joelgranados/EcoIS (Online; accessed 09-Apr-2013).

Granados, J.A., 2010b. Image Annotation Tool. https://github.com/Joelgranados/EcoAN (Online; accessed 11-Apr-2013).

Granados, J.A., Graham, E.A., Bonnet, P., Yuen, E.M., Hamilton, M., 2013. EcoIP: an open source image analysis toolkit to identify different stages of plant phenology for multiple species with pan tilt zoom cameras. Ecol. Inf. 15, 58–65.

Høye, T.T., Post, E., Meltofte, H., Schmidt, N.M., Forchhammer, M.C., 2007. Rapid advancement of spring in the high Arctic. Curr. Biol. 17 (12), R449–R451 (Jun., URL http://www.ncbi.nlm.nih.gov/pubmed/17580070).

Høye, T.T., Post, E., Schmidt, N.M., Trøjelsgaard, K., Forchhammer, M.C., 2013. Shorter flowering seasons and declining abundance of flower visitors in a warmer Arctic. Nat. Clim. Chang. 3 (6), 1–5 (Jun., URL http://www.nature.com/doifinder/10.1038/nclimate1909).

Ide, R., Oguma, H., 2010. Use of digital cameras for phenological observations. Ecol. Inf. 5 (5), 339–347 (Sep., URL http://linkinghub.elsevier.com/retrieve/pii/S1574954110000762).

Iler, A.M., Hye, T.T., Inouye, D.W., Schmidt, N.M., 2013. Long-term trends mask variation in the direction and magnitude of short-term phenological shifts. Am. J. Bot. 100 (7), 1398–1406 (URL http://www.amjbot.org/content/early/2013/05/08/ajb.1200490.abstract).

ITU, 1992. Information technology — digital compression and coding of continuous-tone still images — requirements and guidelines. Tech. rep.International Telecommunication Union (ITU), Genve, Switzerland.

Liang, L., Schwartz, M.D., Fei, S., 2012. Photographic assessment of temperate forest understory phenology in relation to springtime meteorological drivers. Int. J. Biometeorol. 56 (2), 343–355 (Mar., URL http://www.ncbi.nlm.nih.gov/pubmed/21557038).

Lovett, G.M., Burns, D.A., Driscoll, C.T., Jenkins, J.C., Mitchell, M.J., Rustad, L., Shanley, J.B., Likens, G.E., Haeuber, R., 2007. Who needs environmental monitoring. Front. Ecol. Environ. 5 (5), 253–260. http://dx.doi.org/10.1890/1540-9295(2007)5[253:WNEM]2.0.CO;2 Jun., URL.

Meltofte, H., Christensen, T.R., Elberling, B., Forchhammer, M.C., Rasch, M., 2008a. High-Arctic ecosystem dynamics in a changing climate. Introduction. Elsevier, London, pp. 1–12 (Ch.).

Meltofte, H., Christensen, T.R., Elberling, B., Forchhammer, M.C., Rasch, M., 2008b. High-Arctic ecosystem dynamics in a changing climate. Ch. The study area at Zackenberg. Elsevier, pp. 101–110.

Michener, W.K., Porter, J., Servilla, M., Vanderbilt, K., 2011. Long term ecological research and information management. Ecol. Inf. 1, 13–24 (Jan., URL http://linkinghub.elsevier.com/retrieve/pii/S1574954110001159).

Mizunuma, T., Wilkinson, M., Eaton, E.L., Mencuccini, M., Morison, J.I.L., Grace, J., 2013. The relationship between carbon dioxide uptake and canopy colour from two camera systems in a deciduous forest in southern England. Funct. Ecol. 27 (1), 196–207. http://dx.doi.org/10.1111/1365-2435.12026 (Feb., URL).

Parrott, L., Proulx, R., Thibert-Plante, X., 2008. Three-dimensional metrics for the analysis of spatiotemporal data in ecology. Ecol. Inf. 3 (6), 343–353 (Dec., URL http://linkinghub.elsevier.com/retrieve/pii/S157495410800037X).

Post, E., Forchhammer, M.C., Bret-Harte, M.S., Callaghan, T.V., Christensen, T.R., Elberling, B., Fox, A.D., Gilg, O., Hik, D.S., Høye, T.T., Ims, R.A., Jeppesen, E., Klein, D.R., Madsen, J., McGuire, A.D., Rysgaard, S., Schindler, D.E., Stirling, I., Tamstorf, M.P., Tyler, N.J.C., van der Wal, R., Welker, J., Wookey, P.A., Schmidt, N.M., Aastrup, P., 2009. Ecological dynamics across the Arctic associated with recent climate change. Science (New York, N.Y.) 325 (5946), 1355–1358 (Sep., URL http://www.ncbi.nlm.nih.gov/pubmed/19745143).

Richardson, A.D., Jenkins, J.P., Braswell, B.H., Hollinger, D.Y., Ollinger, S.V., Smith, M.-L., 2007. Use of digital webcam images to track spring green-up in a deciduous broadleaf fores. Oecologia 152 (2), 323–334 (May, URL http://www.ncbi.nlm.nih.gov/pubmed/17342508).

Rufli, M., Scaramuzza, D., Siegwart, R., 2008. Automatic detection of checkerboards on blurred and distorted images. International Conference on Intelligent Robots and Systems 3121–3126 (Sep., URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4650703).

Ryu, Y., Verfaillie, J., Macfarlane, C., Kobayashi, H., Sonnentag, O., Vargas, R., Ma, S., Baldocchi, D.D., 2012. Continuous observation of tree leaf area index at ecosystem scale using upward-pointing digital cameras. Remote Sens. Environ. 126, 116–125 (Nov., URL http://linkinghub.elsevier.com/retrieve/pii/S003442571200346X).

Saitoh, T.M., Nagai, S., Saigusa, N., Kobayashi, H., Suzuki, R., Nasahara, K.N., Muraoka, H., 2012. Assessing the use of camera-based indices for characterizing canopy phenology in relation to gross primary production in a deciduous broad-leaved and an evergreen coniferous forest in Japan. Ecol. Inf. 11, 45–54 (Sep., URL http://linkinghub.elsevier.com/retrieve/pii/S1574954112000416).

Schmidt, N.M., Hansen, L.H., Meltofte, H., Hansen, J., Berg, T.B., 2012a. BioBasis manual. Tech. Rep. (Aarhus University, Roskilde).

Schmidt, N.M., Ims, R.A., Høye, T.T., Gilg, O., Hansen, L.H., Hansen, J., Lund, M., Fuglei, E., Forchhammer, M.C., Sittler, B., 2012b. Response of an arctic predator guild to collapsing lemming cycles. Proc. R. Soc. B Biol. Sci. 279 (1746), 4417–4422 (Nov., URL http://www.ncbi.nlm.nih.gov/pubmed/22977153).

Smith, A.R., 1978. Color Gamut Transformation Pairs. In: Beatty, J., Kellogg, S. (Eds.), IEEE Comput Society Press, pp. 12–19.

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., Milliman, T., OKeefe, J., Richardson, A.D., 2012. Digital repeat photography for phenological research in forest ecosystems. Agr. Forest. Meteorol. 152, 159–177 (Jan., URL http://linkinghub.elsevier.com/retrieve/pii/S0168192311002851).

Wright, S.J., Carrasco, C., Calderón, O., Paton, S., 1999. The el niño southern oscillation, variable fruit production, and famine in a tropical forest. Ecology 80 (5), 1632–1647.

Zhao, J., Zhang, Y., Tan, Z., Song, Q., Liang, N., Yu, L., Zhao, J., 2012. Using digital cameras for comparative phenological monitoring in an evergreen broad-leaved forest and a seasonal rain forest. Ecol. Inf. 10, 65–72 (Jul., URL http://linkinghub.elsevier.com/retrieve/pii/S1574954112000192).