The Significance of Image Compression in Plant Phenotyping Applications

Massimo Minervini[A,C], Hanno Scharr[B] and Sotirios A. Tsaftaris[A]

[A]Pattern Recognition and Image Analysis (PRIAn)
IMT Institute for Advanced Studies, Lucca, Italy
[B]Institute of Bio- and Geosciences: Plant Sciences (IBG-2)
Forschungszentrum Jülich GmbH, Jülich, Germany
[C] Corresponding author. Email: m.minervini@imtlucca.it

Abstract

We currently witness an increasingly higher throughput in image-based plant phenotyping experiments. The majority of imaging data are collected based on complex automated procedures, and are then post-processed to extract phenotyping related information. In this article we show that image compression used in such procedures may compromise phenotyping results and needs to be taken into account. We motivate the paper with three illuminating proof of concept experiments which demonstrate that compression (especially in its most common lossy form of JPEG) does affect measurements of plant traits and errors introduced can be high. We further systematically explore how compression affects measurement fidelity, quantified as effects on image quality as well as errors in extracted plant visual traits. To do so we evaluate a variety of image-based phenotyping scenarios, including size and color of shoots, leaf and root growth. To show that even visual impression can be used to assess compression effects we use root system images as examples. Overall, we find that compression has a considerable effect on several types of analyses (albeit visual or quantitative) and that proper care is necessary to ensure that such choice does not affect biological findings. In order to avoid or at least minimize introduced measurement errors, for each scenario we derive recommendations and provide guidelines on how to identify suitable compression options in practice. We also find that certain compression choices can offer beneficial returns, in terms of reducing the amount of data storage without compromising phenotyping results. This may enable even higher throughput experiments in the future.

Additional keywords: computer vision, imaging sensor, coding, lossless, lossy, optical flow, growth analysis, roots.

1. Introduction

In the last decades, image-based plant phenotyping has gained considerable attention due to the availability of high-end commercial solutions as well as open and low-cost approaches (Spalding and Miller, 2013). Vision-based measurements allow recording and monitoring of relevant phenotypes non-invasively, with higher precision, accuracy, and throughput than manual measurement (Shirmohammadi and Ferrero, 2014), at considerably reduced cost and human labor (Furbank and Tester, 2011). The adoption of image-based approaches, favored by the availability of a variety of image processing solutions (Lobet et al., 2013; Spalding and Miller, 2013; Klukas et al., 2014; Sozzani et al., 2014) and robust automation, has increased significantly the throughput of phenotyping experiments, which is key to advancing our understanding of plant structure and function.

However, the design and deployment of such approaches requires a significant multi-disciplinary effort and know-how in a variety of domains such as automation hardware, image acquisition, software engineering, computer vision and image analysis, and of course plant biology (Minervini et al., 2015). Clearly, such know-how can be found in few settings and more often than not plant biologists need to rely on (and cannot control for) choices made by other parties involved (e.g., a manufacturer, a contract provider, or a collaborator). One crucial case we want to highlight here, is the choice of data compression, a procedure by which a file (e.g., an image) can be represented digitally using as few computer storage resources as possible. This process can be done either by downsampling an image (i.e. reducing vertical and horizontal dimensions) which is rather common since it also speeds up analysis performance (due to the smaller image) or using a more sophisticated compression technique. Any choice does affect the fidelity of the available data, and in many cases its presence in the acquisition procedure is unknown to the end user. An inspection of approximately 60 well cited papers in the recent literature among those present in the plant image analysis software database (Lobet et al., 2013) finds that (i) most authors do not report if imaging data were compressed, and that (ii) few authors did use compression with a lossy image format (e.g., JPEG). Both of these findings are worrisome, because in the former case it could be that it is unknown even to the authors if compression was present and in the latter case it is unknown if compression had an effect. These concerns are also shared by others, stating that care in compression choice must be undertaken (Slovak et al., 2014) and that it should be reported (Cwiek et al., 2014).

Clearly, lossy compression (which reduces an image's file size by permanently removing certain information from the original image) must have an effect, but in some scenarios such compression choice is necessary. The constant need to increase experimental scale (e.g., more subjects, higher spatial and temporal resolution, more imaging modalities (Furbank and Tester, 2011; Dhondt et al., 2013; Fiorani and Schurr, 2013)) produces vast amounts of image data (Pieruschka and Poorter, 2012; Cobb et al., 2013; Granier and Vile, 2014). For example, a single experiment with the setup described in (Dhondt et al., 2014), i.e. 10 plants imaged per hour for 19 days, produces approximately 70 gigabytes (GB) of raw image data (equivalent to 15 DVD discs). Using color images and higher resolution camera sensors (e.g., as in (Briese et al., 2013; Knüfer et al., 2013)) would increase that figure even more to 250 GB (equivalent to 53 DVD discs) for the same experiment. State-of-the-art image compression standards are able to compact such data in a way that it would fit in a single DVD disc. While upgrading and ameliorating the e-infrastructure is a key issue (Pieruschka and Poorter, 2012), it is a slowly changing factor and a costly operation, which requires sensible data management strategies and planning. Furthermore, the importance of easy and rapid access to data has been highlighted for

plant phenotyping projects involving institutions and parties geographically distributed (Billiau et al., 2012). Thus, any savings in the amount of data transferred or archived have significant returns to the end user.

In this paper, we first introduce necessary concepts and terminology, relevant error measures allowing to evaluate compression performance in plant phenotyping experiments, as well as image compression standards used here. We then offer three proof of concept experiments to illustrate the effects lossy compression can have: (1) on a simple phenotyping experiment related to measuring growth in a population of 19 *Arabidopsis thaliana* Col-0 individuals; (2) on estimation of local growth rate of an Arabidopsis root tip from a video; and (3) on visual perceptibility of fine roots in high-resolution rhizotron images. For systematic evaluation, we proceed by offering a series of experiments that show, how different choices of compression standards and quality settings affect the extraction of phenotypic information from images and image sequences (of roots, shoots, or leaves) obtained by plant phenotyping experiments.[1] From these systematic evaluations we derive recommendations, such as which compression standards, i.e. which codecs (see below), are suitable for which task using which settings. For image-based plant phenotyping tasks not evaluated here, we derive recommendations, on how to test and decide upon compression choices.

2. Materials and methods

We start by introducing fundamental concepts of digital image and video compression (Sayood, 2012) in order to establish usual terminology. Then, we briefly review the coding standards adopted in the experiments and case studies that follow. Finally, we define the quality and error measures used to compare compression performance in phenotyping experiments.

2.1. Images and image compression

Digital images are two-dimensional grids of picture elements (pixels). For gray scale images each pixel contains a single numerical value indicating the pixel's intensity. Such values are often given as 8-bit integer numbers (thus, ranging between 0 and 255) and therefore uncompressed images use 8 bits per pixel for storage. The most usual color images store 3 values per pixel standing for red, green, and blue (RGB) light intensity, where again each value is given as an 8-bit integer, leading to a storage space of 24 bits per pixel. A digital video is a sequence of gray scale or color images usually called 'frames'.

Typically, raw image data is highly redundant, e.g., in a spatially homogeneous region like a uniform background pixel values do not change when stepping from one pixel to its neighbors. In such cases it is sufficient to store the pixels value once and in addition store for how long this value stays constant, when stepping from pixel to pixel. This is called 'run-length encoding', a base mechanism often used in compression, e.g., in JPEG. Raw image data cannot only be highly redundant in space, but also in time, e.g., when the background remains constant over time; or in color, e.g., when only a fraction of the available color space is used.

Data compression aims to reduce such unwanted redundancy to obtain an as compact digital representation as possible, i.e. a small image file. The smaller the file, the higher is the compression efficiency. It is expressed in terms of bit rate (BR), measured in bits per pixel (bpp):

---

[1] Some of which are carried out and routinely used at the Institute of Bio- and Geosciences: Plant Sciences (IBG-2) of Forschungszentrum Jülich, Germany (`http://www.fz-juelich.de/ibg/ibg-2/EN`).

$$BR = \frac{\text{image file size}}{\text{width} \times \text{height}},\tag{1}$$

where BR denotes the average number of bits required to represent a single image pixel in an image with given width and height.

Current compression standards use a variety of sophisticated techniques to achieve lower and lower bit rates. Such schemes consist of two parts: an encoder converting the original image into a compressed file and a decoder reversing this process, i.e. converting a compressed file into an image. A software or hardware implementation of a compression standard is thus termed 'codec' (coder/decoder). A typical workflow for encoding and decoding is depicted in Figure 1 for background information, but details of the techniques are of no relevance here. However, we need to be aware that there are two general categories of compression.

Lossless compression: Here no loss of information occurs and the decompressed image is a prefect copy of the original, as e.g. in ZIP file compression. Thus, lossless compression does not compromise image quality or results of phenotyping experiments. Their relevance in practice lies in the achieved compression efficiency and the computational effort needed for coding and decoding. We investigate this in the performance analysis section of this paper (Section 4.1).

Lossy compression: Here some information is lost due to compression and the decompressed image is only an approximation of the original. Typically, lossy compression achieves much higher compression efficiency, i.e. smaller file sizes. It has become ubiquitous with the JPEG standard (ITU, 1992). Lossy compression standards are designed to achieve the least mathematical or perceptible (Lee and Ebrahimi, 2012) difference between the original and reconstructed images, with the smallest possible compressed file size. Therefore, applying lossy compression always entails a trade off between smaller file size and better image quality. Another very simple and commonly employed form of lossy compression is the selective downsampling (rescaling) of an image to reduce file size and either performing analyses on the downsampled version or the equal upsampling prior to analysis.

2.2. Metrics for image quality evaluation

The core of this article is to investigate what 'better image quality' really means in a plant phenotyping context. As higher bit rate, BR (Eq. (1)), typically corresponds to less information loss, we evaluate different codecs at various bit rates. We encode the original image I at a given bit rate, reconstruct the image $\hat{I}$ by decompression, and compare it to the unprocessed original I. For this comparison we use several information theoretic or plant science specific metrics, in order to investigate which codec is the best for a given plant phenotyping application.

Most codecs are developed without a specific application in mind. They are therefore usually evaluated against information theoretic measures like execution efficiency, image fidelity, or color distortion. We use these measures for reference and introduce them below. However, these measures are not specific for plant experiments, where the ultimate information of interest is the actual measure describing a plant trait. Therefore, compression performance should be evaluated against how accurately the trait of interest can be measured (Minervini and Tsaftaris, 2013). As quantitative traits we exemplarily investigate image segmentation-based traits and traits based on image sequence analysis. For segmentation-based analyses (i.e. involving the automatic delineation of plant objects in an image) we select two measures, namely Projected Leaf Area (PLA) and a more general segmentation accuracy measure (DSC,

Dice Similarity Coefficient). For image sequence analysis we use Relative Elemental Growth Rate (REGR) as plant-related measure and the Average End-point Error (AEE) as a well-established measure for optical flow accuracy. Finally, we also look at visual fidelity when a human expert evaluates an image. For visual fidelity we show and discuss example images, other measures are defined in the following paragraphs.

Execution efficiency is evaluated in terms of runtime, i.e. the time to encode and decode image data. It is measured in seconds (see Appendix B for further details on execution times).

Image fidelity is expressed in terms of Peak Signal-to-Noise Ratio (PSNR), measured in decibel (dB):

$$\text{PSNR} = 10 \cdot \log_{10} \frac{255^2}{\frac{1}{N} \sum_{i=1}^{N}(I_i - \hat{I}_i)}, \tag{2}$$

where N is the number of image pixels. A higher PSNR indicates higher image fidelity. For videos, we average PSNR values and also bit rate BR (Eq. (1)) across all frames of the sequence.

Color distortion can be quantified using the information theoretic Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951):

$$\text{KL} = \sum_{i=1}^{B} H_i \log_2 \frac{H_i}{\hat{H}_i}, \tag{3}$$

where H and $\hat{H}$ denote normalized histograms of intensity values of a single color channel, computed on the original and reconstructed images, respectively. B is the number of histogram bins. For RGB images, we estimate overall color distortion, $KL^{RGB} = (KL^R + KL^G + KL^B)/3$, as the average between the KL divergence values obtained on the marginal histograms of the RGB color components. KL divergence is a unitless quantity, that should be as close to zero as possible for higher color fidelity.

Projected Leaf Area (PLA) is proportional to the number of plant pixels observed in an image, e.g., a top view of a rosette plant. PLA is frequently used to evaluate shoot development as it correlates well with plant biomass (Walter et al., 2007; Granier et al., 2006). Plant pixels are found via automated segmentation. We quantify the amount of error in plant area estimation when compression is used as the relative change:

$$\text{PLA Error} = \frac{\hat{A} - A}{A}, \tag{4}$$

of area $\hat{A}$ found automatically on the reconstructed image (i.e. the image after the compressed image is decompressed) with respect to the area A found based on the original uncompressed image. We express PLA Error as a percentage, where best possible value is 0%, while positive or negative values indicate an over- or under-estimation of the plant area, respectively.

Segmentation accuracy is more sensitive to segmentation errors than PLA. Suppose a found segment has the correct size, but is distorted or shifted with respect to the ground truth segment, then PLA Error (Eq. (4)) would be 0 despite the segmentation error. A measure capturing such errors is the Dice Similarity Coefficient (DSC) (Dice, 1945):

$$\text{DSC} = \frac{2 \cdot |M \cap \hat{M}|}{|M| + |\hat{M}|}, \tag{5}$$

which estimates the spatial overlap between the binary segmentation masks M and $\widehat{M}$, obtained by segmenting original and reconstructed images, respectively. We express DSC as a percentage, where a DSC value of 100% denotes perfectly matching segmentation masks.

Relative Elemental Growth Rate (REGR): Accurate estimation of local growth rates can be obtained using motion estimation techniques based on optical flow analysis (Schmundt et al., 1998; Walter and Schurr, 2005). We estimate the optical flow from an image sequence using the combined local-global approach in (Bruhn et al., 2005). This allows us to track points through the image sequence. REGR is quantified as the spatial 1D elongation rate between two points on e.g. a root (Peters and Bernstein, 1997; Chavarría-Krauser et al., 2008):

$$\text{REGR} = \frac{1}{T} \ln \frac{l_j(T)}{l_j(0)}, \tag{6}$$

where T is the time duration over which growth is estimated, and $l_j(\cdot)$ is the distance between the points of interest at a given time. REGR is measured in $\%\text{h}^{-1}$, and its calculation relies only on initial (at time $t = 0$) and final (at $t = T$) segment lengths.

Average End-point Error (Otte and Nagel, 1994) is a more general performance measure for optical flow, also applicable e.g. in tracking scenarios. Here, optical flow $\widehat{\boldsymbol{u}}$ calculated on the compressed sequence is compared to the ground-truth flow $\boldsymbol{u}$ calculated on the original sequence using the normalized Average End-point Error (AEE) :

$$\text{AEE} = \frac{\sum_{i=1}^{N} \|\widehat{\boldsymbol{u}}_i - \boldsymbol{u}_i\|_2}{\sum_{i=1}^{N} \|\boldsymbol{u}_i\|_2}, \tag{7}$$

where $\boldsymbol{u}_i$ and $\widehat{\boldsymbol{u}}_i$ denote the displacement estimated on original and reconstructed sequence at the ith pixel, and $\|\cdot\|_2$ denotes the L_2 norm yielding the length of a vector. AEE is expressed as a percentage, with 0% denoting perfectly matching flow fields. We normalize with respect to the ground truth motion vector length, in order to accommodate slow-moving test sequences.

2.3. Employed image and video codecs

We employ a variety of state-of-the-art lossless and lossy image and video coding standards. While these have been developed for multimedia and entertainment applications they are widely used in several other domains.

For lossless image compression we consider: *PNG* (W3C, 2003), *JPEG-LS* (Weinberger et al., 2000), the lossless option of *JPEG 2000* (Skodras et al., 2001), and *WebP*.

For lossy image compression we consider: *JPEG* (ITU, 1992), *JPEG 2000* (Skodras et al., 2001), and *WebP*. We also consider a variant of JPEG 2000, permitting native region-of-interest (ROI) coding (Christopoulos et al., 2000), a feature allowing to encode foreground image regions at a higher quality than background regions.

For video we consider only lossy standards, namely: the royalty free *VP9* (Mukherjee et al., 2013), and the recent *HEVC* (Sullivan et al., 2012). Additional details and parameters used are outlined in Appendix A.

3. Case studies: Effect of compression in typical plant phenotyping applications

In the following, we investigate how lossy image and video compression techniques influence results in three typical plant phenotyping experiments. In Section 3.1, we consider size measurement of *Arabidopsis thaliana* using projected leaf area (PLA), i.e. a typical trait based on automated segmentation of single images. In Section 3.2, we select local growth rate of a root tip, i.e. a trait derived computationally from an image sequence. In Section 3.3, we investigate rhizotron images showing complete root systems in soil in an example where the reference state-of-the-art evaluation still is the human eye.

3.1. Example 1: Size of a rosette plant evaluated by PLA

The purpose of this proof of concept experiment is to demonstrate the type of errors introduced by lossy compression in a typical phenotyping experiment measuring rosette growth over a period of time. We use imaging data of a population of 19 wild-type (Col-0) *Arabidopsis thaliana* plant subjects acquired using off-the-shelf commercial cameras in a controlled environment as described in (Minervini et al., 2014a; Scharr et al., 2014). Twenty (20) observations within an imaging period of 7 consecutive days, 12 days after germination, are obtained for each replicate. The images (width×height: 3108×2324 pixels) are in color and are recorded in the raw, uncompressed camera format.

Two versions of the dataset are considered and are processed individually. One, *uncompressed*, containing the original images, and one *compressed* with the JPEG algorithm at quality factor $q = 27$ (cf. Figure 2a-b). To highlight the subtle compression artifacts, Figure 2c-d shows a zoomed detail of one of the plants (in the blue bounding box in Figure 2a-b). Compression introduces slight discontinuities due to the so-called blocking artifacts but this does not lead to obvious loss in perceived image quality.

Images are analyzed independently to obtain rosette segmentations as described in (Minervini et al., 2014a). Even these slight compression artifacts do affect analysis algorithms: as shown in Figure 2e compression causes changes in the segmentation. We observe leakage (indicated with red pixels) of the plant boundary to non-relevant plant material, in this case moss, where moss extent increases with time for some of the subjects. Compression also tends to slightly affect the delineation of the plant (indicated with blue pixels) almost in its whole periphery and causes also the loss of some of the leaf stems. These segmentation differences directly affect PLA, and PLA Error (Eq. (4)) is 3% in the shown case. We estimate PLA Error for all plants and all time points (cf. Figure 2f). We observe that errors are diverse. PLA is mostly overestimated, up to 12%, a trend increasing with time potentially due to leakage to moss areas becoming more severe, but sometimes also underestimated by up to –4%. Notice that the ordering of dots changes as a function of time and how as time advances larger errors are evident.

These empirical observations are statistically confirmed by an ANOVA. We conduct a one-way repeated measures ANOVA (using Stata version 11, StataCorp LP, College Station, Texas, USA) to investigate significance of effects of time (i.e. within subjects) and of replicate (i.e. between subjects) on PLA error. This indicates if compression affects replicates differently and if the error changes as plants grow. Since PLA error has been normalized by the plant's area before compression, individual growth effects of plants should be minimal, and ANOVA

will be testing a linear effect of the two independent variables (IVs) time and replicate on the error introduced by compression on PLA measurements, via the PLA error dependent variable.

ANOVA results are shown in Table 1. Considering a significance level of < 0.05, time is a relevant factor ($F[4.05, 72.99] = 3.65$, p-value $= 0.008$) with a positive slope and between subject effects are present among the replicates ($F[1,18] = 81.27$, p-value < 0.00001).

We conclude that in this example *visually nearly unnoticeable compression distortion affects rosette growth estimates*.

3.2. Example 2: Quantifying local root growth by REGR

Several image-based plant measurements rely on accurate correspondence analysis, e.g., image-based 3D reconstruction, motion and local growth rate analysis. As an example of this analysis class, we investigate local motion analysis, e.g., used to study growth-related phenotypes such as gravitropic response (Chavarría-Krauser et al., 2008).

This experiment investigates how distortion in an optical flow field due to JPEG compression affects local growth estimation. To this end we adopt the method for REGR estimation in root tips described in (Chavarría-Krauser et al., 2008). We apply it to an image sequence showing a growing Arabidopsis root tip (cf. Figure 3a, original) over 2.5 hours. The sequence consists of 300 images and the processed region of interest is 422×77 pixels.

The method works as follows: the mid-line C of the root (cmp. the green line in Figure 3a, at $t = 0$ h) is given for the first image. C is represented by equidistantly spaced points C_j, one per pixel length. The positions of these points are individually tracked in time by optical flow calculated from the image sequence (using the algorithm given in (Bruhn et al., 2005)). The outcome of this tracking is shown as green points in Figure 3a, original at $t = 2.5$ h. From the distance l_j between a point C_j and its neighbor C_{j+1} at $t = 0$ h and at $t = 2.5$ h Relative Elemental Growth Rate (REGR) is then calculated via Eq. (6).

Figure 3b, black line, shows the so derived growth rate mapped to the mid-line at $t = 0$. In agreement with (Chavarría-Krauser et al., 2008) we observe that the growth rate between quiescence center and growth zone (position x between 0 and approximately 130 μm) is around 5 %h^{-1}. The growth zone starts around position $x = 130$ μm and ends at approximately $x = 450$ μm.

To investigate the effects of compression, we store the sequence in four different JPEG qualities $q \in \{95, 85, 75, 65\}$ and apply the method as before. Figure 3a shows the tracking results for these JPEG qualities and Figure 3b the corresponding REGR curves. Comparing the tracking results for the different qualities we observe, that compression affects tracking of different root zones inconsistently, mainly depending on local image contrast. For example, while the locations of growth maximum and root tip appear stable with respect to compression, width of growth zone decreases up to 20% for higher compression. However, major effects of compression occur in the zone behind the tip, whose maximum growth rate should be constant at approximately 5 %h^{-1} (Chavarría-Krauser et al., 2005, 2008), while already for very high quality JPEG compression ($q = 95$), the observed error in REGR (Eq. (6)) is 21%. For higher

compression ratios (i.e. lower JPEG quality factor q) the error in REGR increases dramatically up to approximately 380% ($q = 65$).

We conclude that in this example *compression affects the estimation of the spatio-temporal pattern of root tip growth, especially in regions with low image contrast*.

3.3. Example 3: Visual evaluation of root images by proxy measures

Below-ground plant organs can be studied non-invasively using the rhizotron (Nagel et al., 2012). Figure 4a shows an example gray scale image (width×height: 4872×3248 pixels) including the root systems of three rapeseed subjects, obtained from root phenotyping experiments at the GROWSCREEN-Rhizo (Nagel et al., 2012). The gold standard for evaluation of such images is still manual delineation of the roots by a human expert.

Lossy compression can alter the appearance of the images, introducing visual distortions or loss of details that may influence the user's capability of accurately delineating the roots. Visually these distortions are evident and in Figure 4c we show a collection of root images at different levels of compression with JPEG and the more advanced JPEG 2000.

At bit rate BR $= 0.1\,$bpp (cf. Eq. (1)), being equivalent to space savings of 98.7%, the root structure is still clearly evident and rich in details and the subsequent delineation by an expert should not be affected. This can be obtained e.g. in JPEG by setting the quality factor at $q = 20$. However, when encoding at even lower bit rates, compression distortion (e.g., blocking artifacts in JPEG or blurring in JPEG 2000) increases substantially, rendering the thin roots increasingly difficult to recognize (even for a trained human observer). At $0.03\,$bpp (i.e. 99.6% space savings), the JPEG image has lost most information and is practically unusable, while in the JPEG 2000 image only the thicker roots are still distinguishable. Since employing experts to manually delineate roots for all possible choices of compression parameters (e.g., bit rate) is extremely tedious, quantitatively measure loss in accuracy is non-trivial. A suitable alternative is to use general image fidelity measures (e.g., PSNR) as proxy. Compression performance of different coding standards is assessed quantitatively in Figure 4b using PSNR (Eq. (2)): JPEG 2000 obtains image fidelity superior to JPEG at any bit rate, while best image quality is achieved by HEVC.

We conclude that *lossy compression is admissible even when subtle structures need to be visually assessed and potentially quantified by a human expert. However, care needs to be taken that compression artifacts remain close to being unnoticeable which can be measured with general image fidelity measures*.

4. Performance analysis of lossless and lossy codecs in plant applications

The case studies above showed that lossy compression can affect results of quantitative evaluation methods. In this section we offer a richer evaluation, including more data sets, metrics, and codecs tested systematically at different compression rates. This allows us to derive recommendations on which codec to use, at which compression rate, and for which task.

We focus on segmentation-based methods for images in Section 4.2 and on growth estimation from videos in Section 4.3. Additionally, we investigate the effectiveness of lossless codecs, which do not compromise image quality. To compare with downsampling as a form of lossy compression, we evaluate the possibly higher fidelity of analyses on the upsampled image, for a fair comparison to more sophisticated compression techniques using the higher resolution as well.

4.1. Lossless coding: space and time savings with no loss in quality

Compression performance obtained using lossless compression approaches on a 16 megapixel gray scale image of rapeseed roots (Figure 4a) is shown in Figure 5. A good balance of bit rate reduction and codec efficiency (cmp. Appendix B) is achieved by JPEG-LS and JPEG 2000, while PNG obtains slightly worse results. Overall, with lossless compression it is possible to reduce file size considerably (to approximately 35% of uncompressed size) with exact reconstruction of the original image and limited computational overhead (in most cases less than one second for decoding, cf. Appendix B). Analogous compression ratios are typically obtained when lossless compression is applied on images composed of more than one component (e.g., 3 for RGB color images, or in general M for hyperspectral data cubes), and scaling appropriately the results in Figure 5 (e.g., ×3 for color, or ×M for hyperspectral) would provide an estimate of expected compression performance. Compared to rescaling image size by half – which is unfortunately a common approach leading to loss in information – lossless compression should always be preferred since it achieves better compression efficiency and no information loss.

To elucidate how data size may reflect to transmission times over a network, we consider an example scenario in which an 18 megapixel color image of a rapeseed shoot (Figure 6b) is acquired in a greenhouse and transmitted to a central processing unit at a different physical location (e.g., for processing or storage). Subsequently, the same image is downloaded from the central repository where it is stored to a user's workstation. We perform this test during working day, to ensure average network traffic conditions, using a workstation and 100 Mbit/s wired network connection. Uploading our test image in uncompressed format (53.7 MB) requires 6 seconds, and downloading the uncompressed image takes 4.9 seconds. On the other hand, encoding the image with JPEG 2000 in lossless mode and transmitting the compressed file (11.3 MB) requires overall only 2 seconds, while downloading and decoding the compressed image locally requires only 1.6 seconds. Image compression leads in this case to 79% space savings and 67% transmission time reduction.

We conclude that *lossless compression does offer significant space savings but for even more savings lossy compression is necessary*.

4.2. Lossy compression in segmentation-based shoot image analysis

Image-based investigations of above-ground plant organs often rely on color images acquired from top or side views. Plant segmentation (i.e. the delineation of the image regions containing a plant object) represents a fundamental step in most image processing pipelines for phenotyping applications (Minervini et al., 2014a), and permits us to calculate a variety of morphological and color features.

The accuracy of plant segmentation affects all subsequent analyses, therefore we investigate compression performance with respect to automated plant segmentation from background. We

adopt three color images of plant shoots acquired from different angles (top or side view), including Arabidopsis (width×height: 2448×2048 pixels), rapeseed (width×height: 5184×3456 pixels), and maize (width×height: 2048×2448 pixels). Original images are shown in Figure 6a-c.

Here, plant segmentation is performed by a pixel-level classifier, which decides if a pixel is foreground (plant) or background. We use a support vector machine (SVM) operating on color values (Briese et al., 2013), trained on labeled image data, where the plant has been delineated manually. The resulting foreground/background classification is refined using morphological operations to fill holes and remove small objects. Figure 6a-c shows example segmentation masks obtained with this method.

We quantify changes in plant segmentation due to compression using three different metrics: (i) Projected Leaf Area Error (PLA Error, Eq. (4)) as a plant related segmentation measure, (ii) Dice Similarity Coefficient (DSC, Eq. (5)) as a well established segmentation measure, and (iii) Kullback-Leibler (KL) divergence (Eq. (3)) on the foreground to see how color information is affected (e.g., used to quantify drought stress tolerance under varying irrigation (Knüfer et al., 2013) or stress (Berger et al., 2010; Sass et al., 2012) conditions).

As shown in Figure 6d-f, using JPEG 2000 and HEVC standards, it is possible to obtain PLA measurements very close to those obtained on the original image, even at low bit rates. The oscillating behavior observed for some codecs at very low bit rates is due to portions of the background that, due to compression artifacts, occasionally appear to the plant segmentation software as belonging to a plant object. Depending on image characteristics and segmentation method, approximation errors due to compression may lead to an over-estimation (e.g., Arabidopsis image, cf. Figure 6d), or an under-estimation (e.g., maize image, cf. Figure 6f) of the plant area.

The accuracy of the segmentation mask (based on which PLA and also several features related to plant morphology can be calculated) is measured by the DSC (see Eq. (5)). As shown in Figure 6g-i, JPEG 2000 + ROI offers best performance, followed by plain JPEG 2000 and HEVC, whereas WebP and JPEG exhibit an erratic behavior. Comparing PLA Error to DSC, we observe that codecs obtaining comparable performance in PLA, e.g. JPEG 2000 + ROI and HEVC, differ in their performance in DSC, i.e. JPEG 2000 + ROI performs better.

Color degradation, quantified by the Kullback-Leibler divergence, is minimized by JPEG 2000 and HEVC (Figure 6j-l). While JPEG systematically introduces higher distortion in a plant's color, performance of WebP depends on the complexity of the image (e.g., cluttered background).

Notice the difference in bit rate ranges among test images, reflecting how much 'compression-friendly' the content is. For equivalent segmentation or color accuracy, the rapeseed and maize images can be compressed at significantly lower bit rates than Arabidopsis, due to the large uniform background regions of the former as opposed to the highly textured background of the latter, i.e. the soil, which is less efficient to encode.

Overall, the plots relative to rapeseed (Figure 6e, h, k) and maize (Figure 6f, i, l) images reveal that for JPEG 2000 + ROI bit rates between 0.1 and 0.2 bpp (i.e. less than 1% of uncompressed 24 bpp image size) are sufficient to adequately encode such data, while for the Arabidopsis test image bit rates higher than 0.5 bpp (i.e. 2% of uncompressed image size) are recommended (cf. Figure 6d, g, j). As we see from this figure, allowing for even higher bit rates does not improve results with respect to the metrics employed here.

For comparison, Figure 6 also shows performance obtained by adopting image rescaling as a form of compression. The Arabidopsis image is scaled down to 430×360 pixels (i.e. the

equivalent of 0.75 bpp), while rapeseed and maize images are scaled down to 480×320 and 187×224 pixels, respectively (equivalent in both cases to 0.2 bpp). Prior to segmentation the images are upscaled to their original size using linear interpolation (as an equivalent of decompression). It is readily apparent from Figure 6 that downscaling is outperformed by all compression approaches (including even JPEG) in all cases. Notice how downscaling leads to underestimation of PLA in the maize image (cf. Figure 6f), because some of the thin leaf structures disappear when the image is resized.

In order to give a visual impression of compression performance, Figure 7 shows example reconstructed images, after compression at 0.05 bpp. A compression factor of 1:480 is applied, to reduce uncompressed image size of 57.7 MB to approximately 112 kB. With JPEG 2000 + ROI the plant appears identical to the original, while plain JPEG 2000, without any prior knowledge on the image regions of interest, is less rich in details and the borders of the segmented plant present small errors. WebP severely over-smooths the image, thus losing the venation patterns in the leaf. Despite the low bit rate, all such images (and corresponding segmentation masks) are visually plausible, as compared to the original image. On the other hand, JPEG (using quality settings of $q = 10$ to achieve file size equivalent to other approaches) exhibits noticeable block artifacts and color degradation, introducing also larger errors (holes) in the segmentation. All of these factors may severely affect accuracy of the phenotypic analyses conducted on JPEG compressed image data.

We conclude that *newer lossy compression standards such as JPEG 2000 + ROI do offer significant benefits in bit rate reduction without degrading results significantly. However, only up to some application-dependent point, since artifacts introduced can severely affect further analysis. Image rescaling to save bits should always be avoided.*

4.3. Local growth estimation of leaves and root tips

Measuring local growth rates in plant tissues by optical flow analysis is a widely applied method (Schmundt et al., 1998; Walter and Schurr, 2005; Dhondt et al., 2013; Pal et al., 2013; Matos et al., 2014). We investigate how compression affects such measurements, adopting two time-lapse sequences (videos) of gray scale images. Example image stills (frames) are shown in Figure 8a and d, respectively, of a growing Arabidopsis leaf (11 frames, width×height: 640×480 pixels) and a growing tobacco root tip (60 frames, width×height: 740×570 pixels). Effects of compression on the optical flow field can be accurately quantified using the Average End-point Error (AEE, Eq. (7)).

Figure 8 shows PSNR and AEE calculated inside a manually defined region of interest, for Arabidopsis leaf and tobacco root tip sequences, encoded at various bit rates. In general, we observe that optical flow calculations are more sensitive to image compression than other applications in previous examples. Thus, in order to keep AEE values reasonably low, we consider the highest range of bit rates (and quality) possible with lossy coding techniques.

Due to the high similarity between consecutive frames of the image sequences, video codecs (VP9, HEVC) provide considerable improvement in PSNR with respect to approaches that compress the frames independently (JPEG, JPEG 2000). Providing the JPEG 2000 encoder with region-of-interest (ROI) information is beneficial only at lower bit rates. Above a certain bit rate (i.e. 1 bpp for the leaf and 0.25 bpp for the root tip), foreground is already encoded at the

best quality possible and further bit budget is spent in background regions. Additionally, at near-lossless coding rates, the underlying effect of encoding the ROI (increased dynamic range of the values to encode, due to so-called bit plane shifting (Christopoulos et al., 2000)) may reduce coding efficiency further.

On the other hand, image fidelity is not strictly correlated with the preservation of the optical flow fields. Surprisingly, on the shorter Arabidopsis leaf sequence (11 frames), JPEG obtains the best AEE performance at several bit rates and is always superior to HEVC (Figure 8c, f), despite opposite PSNR results (Figure 8b, e). For the longer root tip sequence (60 frames), HEVC and VP9 represent the best option for low bit rates, while JPEG is still superior at high bit rates.

We conclude that *if lossy compression is needed, JPEG at highest quality levels should be preferred, but even then additional 2% AEE due to compression should be expected*.

5. Discussion and conclusions

While image-based phenotyping is becoming increasingly important and utilized, several aspects related to dealing with data fidelity and integrity remain unexplored. In this paper we investigate the effects of lossy image compression on phenotyping accuracy and offer guidelines on the proper and guided use and reporting of compression in plant phenotyping experiments.

Our first proof of concept experiment (Section 3.1) illustrates that even in the simple case of measuring rosette plant area, the most popular form of lossy image compression (i.e. images compressed with JPEG) does introduce non-negligible errors in measurements. Compression in this case does not cause visually perceptible distortion, but local loss of image fidelity does affect the outcome of segmentation: the image processing process that lies beneath the measurement of plant rosette area. More importantly, although it appears that with compression PLA is overestimated, the effect of compression on the algorithm is not constant: it is not a systematic error. Unfortunately it is not a completely random error either: it varies as plants grow (as the ANOVA experiment shows), from a time instant to another, and between plants. Although the ANOVA identifies this to be a factor, it is more readily seen in scatter plots: ordering of points changes from time to time, and even more critically the behavior changes among different plants of the same genotype. If compression was a systematic error then this would simply introduce a bias (a change in population means) which would not affect any statistical tests. If compression error was totally random and uncorrelated to the data, then this would simply imply that larger variance attributed to compression is observed and to account for this additional variance a larger sample (more replicates) would merely be necessary, in order to match the statistical power of the data without compression. But also this is not the case.

Compression is a highly influential factor also when growth analysis relies on optical flow fields. Tasks involving the tracking of high contrast structures, e.g. root tip, generally prove robust to higher compression ratios, however, for growth analysis JPEG compression should be limited to very high quality factors ($q \geq 95$). Using more sophisticated compression standards (e.g., JPEG 2000, VP9, HEVC) may not yield better results (cmp. Section 4.3). If the image data shows low contrast in relevant image regions, lossless compression should be adopted to avoid dramatic degradation in accuracy.

These findings are illuminating, since when compression is present without a user's knowledge, the measurements would be affected by the compression. Here, we could observe these errors because the data are acquired in original uncompressed quality. Therefore, users of image-based phenotyping platforms should first identify if compression is used in their system, report it in their papers, and analyze its effects by obtaining some data without compression.

There are several practical reasons that do necessitate the use of compression. The rapid accumulation of data and the need to archive such data for regulatory compliance is the most common. In this case if ample storage is available then without a doubt, as our experiments show, the lossless compression options of JPEG-LS or JPEG 2000 should be considered, since they can still reduce by 60 to 80% (depending on the image characteristics) the amount of data (roughly equivalent to 2-3 bpp per each color channel), while perfectly preserving the original image content. Despite obtaining inferior compression efficiency with respect to other methods, the PNG standard is ubiquitous on the Web and its broad installed codec base eases adoption. The benefits of compression could be realized even in depositing or retrieving data from institutional repositories, where compression will maximize the utilization of the installed e-infrastructure.

If larger compression efficiency is required lossy options are necessary. Depending also on the complexity of the image content at hand (i.e. images with less complicated background), most compression algorithms offer near lossless performance in the 2-3 bpp bit rate range, with no major differences observed among algorithms.

For additional storage savings, below 2 bpp compression efficiency is required. There are several scenarios where such efficiency may be necessary. For example, when images are acquired in a greenhouse facility or even in the field, and are then transmitted to a central location for archival and analysis (e.g., as in the framework proposed by (Minervini and Tsaftaris, 2013), or in the gigapixel time-lapse panoramic imaging system in (Brown et al., 2012)). Another example could be the recent developments towards affordable phenotyping[2] where users in developing countries or in rural remote areas acquire images using affordable and low computational power devices (e.g., mobile phones), and transmit them over wireless communication links (enabled in remote places by long-distance connectivity projects (Murillo et al., 2015) or emerging technologies such as the Brck[3]) and the Internet to cloud services (e.g., the iPlant Collaborative (Goff et al., 2011)), where sophisticated analyses can take place, and results are sent back in response (Minervini and Tsaftaris, 2013; Puhl, 2013). Both of these scenarios involve: a remote sensing device, which does not have the computational power to perform analysis; the use of a limited communication channel, which may not have the capacity to carry many large images; and potentially imaging of plants in non-ideal settings, for example in the field (Andrade-Sanchez et al., 2013; Bucksch et al., 2014) or non-uniformly illuminated conditions, which increase the complexity of the image content. Thus, storage and transmission of the image data represent a (technical and logistic) bottleneck and may reduce overall throughput, rendering image file size a key design aspect.

For compression efficiency below 2 bpp careful evaluation of compression effects and choice of compression practice is necessary. If prior to the final deployment of the system, a set of uncompressed imaging data of a genotype and a fixed image processing pipeline are available, direct phenotyping measurements, such as PLA or model errors, can be used to evaluate compression. Statistical analyses, such as the one in Section 3.1, should be performed

---

[2]http://www.phenotiki.com
[3]http://www.brck.com

14

and the choice of the compression algorithm and its parameters (e.g., JPEG quality factor or bit rate) should be guided accordingly to minimize statistical effects. If group based experiments (of different genotypes) are available, changes in group level differences can be used to identify suitable compression standards and parameters.

When populations are not available but some exemplar images are available instead, then analyses such as the ones reported in Section 4.2 are recommended. If no well tested and validated imaging pipeline is available and human-based image evaluation and annotation will be adopted, visual examination of compressed images still provides a suitable alternative. Furthermore, general fidelity metrics (PSNR) or psycho-visual metrics (e.g., structural similarity index, SSIM (Wang et al., 2004)) can be used to find suitable choices of compression and parameters, as a proxy to estimating the user's capability of performing the analysis (e.g., delineating roots).

When an image analysis pipeline is also available we recommend not only the use of application related measurements (e.g., PLA growth rates, tracking estimates in root tip, and others) but metrics such as segmentation quality and color divergence. They not only offer more sensitive evaluation (cmp. DSC vs. PLA Error in Section 4.2) when compared to general fidelity metrics (e.g., PSNR) but can help address future changes to the analysis pipeline (Soyak et al., 2011; Minervini and Tsaftaris, 2013). This is necessary for example when performing new analyses to isolate new traits and explain behavior not considered during the initial experimental design and data collection.

In general, the performance of the coding procedures in terms of quality (or, alternatively, error) measures, suitable for the application at hand, can be visualized when plotted versus the bit rate achieved by compression. This represents a practical tool to operate lossy compression in applications. When designing a phenotyping setup, the so-called rate-distortion (R-D) curves (Ortega and Ramchandran, 1998) (e.g., those employing PLA Error, DSC, or KL Divergence as distortion measure in Figure 6) allow to select compression approaches and parameter settings that provide the optimal trade off between compression ratio and application accuracy. For a desired level of phenotype extraction fidelity (y-axis), the compression methodology providing the lowest bit rate (x-axis) is the most efficient and should therefore be adopted.

Several compression tools are available and the selection of an appropriate image compression strategy is not trivial. Therefore, in this paper we compare quantitatively a variety of state-of-the-art image and video coding standards, focusing discussion on aspects of practical relevancy: (a) compression efficiency, (b) image fidelity, (c) phenotyping accuracy, and (d) encoding/decoding time efficiency. Until specialized compression algorithms tailored to the problems of plant phenotyping become ubiquitous (Minervini and Tsaftaris, 2013; Minervini et al., 2014b), based on our analysis we recommend the following. For still images:

- JPEG 2000 emerges as the approach achieving the best trade off among all parameters, offering noteworthy (and in several cases top) performance in all experiments but motion estimation. When regions of interests are available, for example after data analysis, JPEG 2000 + ROI offers an exceptional choice to archive data with the highest possible quality and compression efficiency. The limiting factor of JPEG 2000 is the lack of large installed codec base due to its limited popularity. This implies that appropriate software installation on workstations and other computing devices is necessary.

- JPEG should be avoided since it performs poorly in most occasions even though it is the first level of choice among users and is ubiquitous.

- While image rescaling (i.e. downsampling) is effective at reducing processing time for analyses, it should be avoided as a way to reduce file size, since compression with either lossless or lossy approaches outperforms this crude lossy approach.

When image sequences (or videos) are concerned:

- JPEG with high quality settings should be used for short sequences, when high bit rates are available, and motion estimation for growth will be performed.
- The new HEVC video coding standard should be used to achieve high space savings especially for long sequences (or stacks) of images with static background. Furthermore, HEVC is an excellent option for long-term storage of time-lapse sequences for growth estimation or low-resolution video streaming (e.g., transmitting a stream of low-resolution previews acquired at the sensor would allow a user to remotely check system status and adjust parameters to changing conditions, or even operate robotized solutions (Alenya et al., 2013)). Its limitations are the additional computational burden introduced by video coding and not a large installed codec base, however, the latter is changing rapidly as more software and hardware manufacturers will include such codec in their distributions.

While some compression schemes may see their efficiency slightly reduced for very small images, in general this is not a concern in phenotyping applications, where image resolution is typically large (e.g., in the order of megapixels). On the other hand, pixel size (i.e. physical size of the scene portion that is projected onto a single pixel, usually in the order of submillimeter) is an important imaging parameter, because lossy compression will affect more small or thin structures (e.g., roots, stems, leaf veins), or subtle differences among groups (e.g., in PLA). For example, compression artifacts observed at high compression rates (cf. Figure 4c) may render a root whose diameter is close to pixel size indistinguishable from the background, in which case lower compression rates or higher spatial resolution in the imaging setup (i.e. smaller pixel size) should be adopted.

In conclusion, while in recent years, sophisticated computer vision solutions have been proposed to address a variety of problems in image-based plant phenotyping, e.g., plant segmentation (Minervini et al., 2014a), leaf and root growth analysis (Schmundt et al., 1998), 3D reconstruction (Paulus et al., 2014), leaf shape (Rolland-Lagan et al., 2014) and orientation (Dornbusch et al., 2012) analysis, chlorophyll fluorescence analysis (Pieruschka et al., 2012), limited attention to the effects of the accumulation of imaging data has been given. This paper alerts the phenotyping community that compression can be a confounder and suggests best compression strategies for a wide selection of applications, adopting off-the-shelf software libraries.

Acknowledgements

References

Alenya, G., Dellen, B., Foix, S., and Torras, C. (2013). Robotized plant probing: Leaf segmentation utilizing time-of-flight data. *IEEE Robotics & Automation Magazine*, 20(3):50–59.

Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., Salvucci, M. E., and White, J. W. (2013). Development and evaluation of a field-based high-throughput phenotyping platform. *Functional Plant Biology*, 41(1):68–79.

Bankoski, J., Wilkins, P., and Xu, Y. (2011). Technical overview of VP8, an open source video codec for the web. In *International Conference on Multimedia and Expo*, pages 1–6.

Berger, B., Parent, B., and Tester, M. (2010). High-throughput shoot imaging to study drought responses. *Journal of Experimental Botany*, 61(13):3519–3528.

Billiau, K., Sprenger, H., Schudoma, C., Walther, D., and Köhl, K. I. (2012). Data management pipeline for plant phenotyping in a multisite project. *Functional Plant Biology*, 39(11):948–957.

Briese, C., Bodewein, T., and Fiorani, F. (2013). Plant shoot image segmentation using colour features and support vector machines. In *2nd International Workshop on Image Analysis Methods for the Plant Sciences*, Nottingham, England.

Brown, T., Zimmermann, C., Panneton, W., Noah, N., and Borevitz, J. (2012). High-resolution, time-lapse imaging for ecosystem-scale phenotyping in the field. In *High-Throughput Phenotyping in Plants*, volume 918 of *Methods in Molecular Biology*, pages 71–96.

Bruhn, A., Weickert, J., and Schnörr, C. (2005). Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231.

Bucksch, A., Burridge, J., York, L. M., Das, A., Nord, E., Weitz, J. S., and Lynch, J. P. (2014). Image-based high-throughput field phenotyping of crop roots. *Plant Physiology*.

Chavarría-Krauser, A., Jäger, W., and Schurr, U. (2005). Primary root growth: a biophysical model of auxin-related control. *Functional Plant Biology*, 32(9):849–862.

Chavarría-Krauser, A., Nagel, K. A., Palme, K., Schurr, U., Walter, A., and Scharr, H. (2008). Spatio-temporal quantification of differential growth processes in root growth zones based on a novel combination of image sequence processing and refined concepts describing curvature production. *New Phytologist*, 177(3):811–821.

Christopoulos, C., Askelöf, J., and Larsson, M. (2000). Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard. *IEEE Signal Processing Letters*, 7(9):247–249.

Cobb, J., DeClerck, G., Greenberg, A., Clark, R., and McCouch, S. (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*, 126(4):867–887.

Cwiek, H., Pommier, C., Scholz, U., Lange, M., Chen, J., van Dijk, A. D., Nap, J.-P., van Oeveren, J., Kersey, P. J., and Krajewski, P. (2014). On minimum reporting requirements and standard formatting of plant phenotypic data. In *International Plant & Animal Genome XXII Conference*, San Diego, California.

Dhondt, S., Gonzalez, N., Blomme, J., De Milde, L., Van Daele, T., Van Akoleyen, D., Storme, V., Coppens, F., T.S. Beemster, G., and Inzé, D. (2014). High-resolution time-resolved imaging of in vitro arabidopsis rosette growth. *The Plant Journal*, 80(1):172–184.

Dhondt, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. *Trends in Plant Science*, 18(8):428–439.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.

Dornbusch, T., Lorrain, S., Kuznetsov, D., Fortier, A., Liechti, R., Xenarios, I., and Fankhauser, C. (2012). Measuring the diurnal pattern of leaf hyponasty and growth in arabidopsis – a novel phenotyping approach using laser scanning. *Functional Plant Biology*, 39(11):860–869.

Fiorani, F. and Schurr, U. (2013). Future scenarios for plant phenotyping. *Annual Review of Plant Biology*, 64(1):267–291.

Furbank, R. T. and Tester, M. (2011). Phenomics – Technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, 16(12):635–644.

Goff, S. A., Vaughn, M., McKay, S., Lyons, E., Stapleton, A. E., Gessler, D., Matasci, N., Wang, L., Hanlon, M., Lenards, A., Muir, A., Merchant, N., Lowry, S., Mock, S., Helmke, M., Kubach, A., Narro, M., Hopkins, N., Micklos, D., Hilgert, U., Gonzales, M., Jordan, C., Skidmore, E., Dooley, R., Cazes, J., McLay, R., Lu, Z., Pasternak, S., Koesterke, L., Piel, W. H., Grene, R., Noutsos, C., Gendler, K., Feng, X., Tang, C., Lent, M., Kim, S.-j., Kvilekval, K., Manjunath, B., Tannen, V., Stamatakis, A., Sanderson, M., Welch, S. M., Cranston, K., Soltis, P., Soltis, D., O'Meara, B., Ane, C., Brutnell, T., Kleibenstein, D. J., White, J. W., Leebens-Mack, J., Donoghue, M. J., Spalding, E. P., Vision, T. J., Myers, C. R., Lowenthal, D., Enquist, B. J., Boyle, B., Akoglu, A., Andrews, G., Ram, S., Ware, D., Stein, L., and Stanzione, D. (2011). The iPlant collaborative: Cyberinfrastructure for plant biology. *Frontiers in Plant Science*, 2(34).

Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S. J., Dauzat, M., Hamard, P., Thioux, J.-J., Rolland, G., Bouchier-Combaud, S., Lebaudy, A., Muller, B., Simonneau, T., and Tardieu, F. (2006). Phenopsis, an automated platform for reproducible phenotyping of plant responses to soil water deficit in arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytologist*, 169(3):623–635.

Granier, C. and Vile, D. (2014). Phenotyping and beyond: modelling the relationships between traits. *Current Opinion in Plant Biology*, 18:96–102.

ITU (1992). Information technology – Digital compression and coding of continuous-tone still images. ITU-T Recommendation T.81, International Telecommunication Union.

ITU (1995). Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios. ITU-R Recommendation BT.601-5, International Telecommunication Union.

ITU (2013). High efficiency video coding. ITU-T Recommendation H.265, International Telecommunication Union.

Klukas, C., Chen, D., and Pape, J.-M. (2014). Integrated analysis platform: An open-source information system for high-throughput plant phenotyping. *Plant Physiology*, 165(2):506–518.

Knüfer, J., Briese, C., Jansen, M., Kleinen, S., Putz, A., Kumar, S., Stelling, D., Schurr, U., and Wiese-Klinkenberg, A. (2013). A non-invasive imaging procedure to identify drought-stress tolerant oilseed rape plants expressing genes from high altitude plants. In *InterDrought-IV Conference*, Crown Perth, Australia.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.

Lee, J.-S. and Ebrahimi, T. (2012). Perceptual video compression: A survey. *IEEE Journal of Selected Topics in Signal Processing*, 6(6):684–697.

Lobet, G., Draye, X., and Périlleux, C. (2013). An online database for plant image analysis software tools. *Plant Methods*, 9(1):38.

Matos, D. A., Cole, B. J., Whitney, I. P., MacKinnon, K. J.-M., Kay, S. A., and Hazen, S. P. (2014). Daily changes in temperature, not the circadian clock, regulate growth rate in brachypodium distachyon. *PLoS One*, 9(6).

Minervini, M., Abdelsamea, M. M., and Tsaftaris, S. A. (2014a). Image-based plant phenotyping with incremental learning and active contours. *Ecological Informatics*, 23:35–48. Special Issue on Multimedia in Ecology and Environment.

Minervini, M., Rusu, C., and Tsaftaris, S. A. (2014b). Unsupervised and supervised approaches to color space transformation for image coding. In *International Conference on Image Processing*, pages 5576–5580.

Minervini, M., Scharr, H., and Tsaftaris, S. A. (2015). Image analysis: The new bottleneck in plant phenotyping. *IEEE SignalProcessing Magazine*, 32(4).

Minervini, M. and Tsaftaris, S. A. (2013). Application-aware image compression for low cost and distributed plant phenotyping. In *International Conference on Digital Signal Processing*, pages 1–6.

Mukherjee, D., Bankoski, J., Grange, A., Han, J., Koleszar, J., Wilkins, P., Xu, Y., and Bultje, R. (2013). The latest open-source video codec VP9 – An overview and preliminary results. In *Picture Coding Symposium*, pages 390–393.

Murillo, M. J., Paco, J. A., and Wright, D. (2015). Long-distance telecommunication in remote poor areas: From partnerships and implementation to sustainability. *IEEE Technology and Society Magazine*, 34(1):19–30.

Nagel, K. A., Putz, A., Gilmer, F., Heinz, K., Fischbach, A., Pfeifer, J., Faget, M., Blossfeld, S., Ernst, M., Dimaki, C., Kastenholz, B., Kleinert, A.-K., Galinski, A., Scharr, H., Fiorani, F., and Schurr, U. (2012). GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. *Functional Plant Biology*, 39(11):891–904.

Ortega, A. and Ramchandran, K. (1998). Rate-distortion methods for image and video compression. *IEEE Signal Processing Magazine*, 15(6):23–50.

Otte, M. and Nagel, H.-H. (1994). Optical flow estimation: advances and comparisons. In *European Conference on Computer Vision*, volume I, pages 51–60.

Pal, S. K., Liput, M., Piques, M., Ishihara, H., Obata, T., Martins, M. C. M., Sulpice, R., van Dongen, J. T., Fernie, A. R., Yadav, U. P., Lunn, J. E., Usadel, B., and Stitt, M. (2013). Diurnal changes of polysome loading track sucrose content in the rosette of wildtype arabidopsis and the starchless pgm mutant. *Plant Physiology*, 162(3):1246–1265.

Paulus, S., Behmann, J., Mahlein, A.-K., Plümer, L., and Kuhlmann, H. (2014). Low-cost 3D systems: Suitable tools for plant phenotyping. *Sensors*, 14(2):3001–3018.

Peters, W. S. and Bernstein, N. (1997). The determination of relative elemental growth rate profiles from segmental growth rates (a methodological evaluation). *Plant Physiology*, 113(4):1395–1404.

Pieruschka, R., Klimov, D., Berry, J. A., Osmond, C. B., Rascher, U., and Kolber, Z. S. (2012). Remote chlorophyll fluorescence measurements with the laser-induced fluorescence transient approach. In *High-Throughput Phenotyping in Plants*, volume 918 of *Methods in Molecular Biology*, pages 51–59.

Pieruschka, R. and Poorter, H. (2012). Phenotyping plants: genes, phenes and machines. *Functional Plant Biology*, 39(11):813–820.

Puhl, J. (2013). Im Silicon Savannah. *Der Spiegel*, 2013(48):118–122.

Rolland-Lagan, A.-G., Remmler, L., and Girard-Bock, C. (2014). Quantifying shape changes and tissue deformation in leaf development. *Plant Physiology*, 165(2):496–505.

Sass, L., Majer, P., and Hideg, É. (2012). Leaf hue measurements: A high-throughput screening of chlorophyll content. In *High-Throughput Phenotyping in Plants*, volume 918 of *Methods in Molecular Biology*, pages 61–69.

Sayood, K. (2012). *Introduction to Data Compression*. Morgan Kaufmann.

Scharr, H., Minervini, M., Fischbach, A., and Tsaftaris, S. A. (2014). Annotated image datasets of rosette plants. Technical Report FZJ-2014-03837, Forschungszentrum Jülich GmbH, Jülich, Germany.

Schmundt, D., Stitt, M., Jähne, B., and Schurr, U. (1998). Quantitative analysis of the local rates of growth of dicot leaves at a high temporal and spatial resolution, using image sequence analysis. *The Plant Journal*, 16(4):505–514.

Shirmohammadi, S. and Ferrero, A. (2014). Camera as the instrument: the rising trend of vision based measurement. *IEEE Instrumentation Measurement Magazine*, 17(3):41–47.

Skodras, A., Christopoulos, C., and Ebrahimi, T. (2001). The JPEG 2000 still image compression standard. *IEEE Signal Processing Magazine*, 18(5):36–58.

Slovak, R., Göschl, C., Su, X., Shimotani, K., Shiina, T., and Busch, W. (2014). A scalable open-source pipeline for large-scale root phenotyping of arabidopsis. *Plant Cell*, 26(6):2390–2403.

Soyak, E., Tsaftaris, S. A., and Katsaggelos, A. K. (2011). Low-complexity tracking-aware H.264 video compression for transportation surveillance. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(10):1378–1389.

Sozzani, R., Busch, W., Spalding, E. P., and Benfey, P. N. (2014). Advanced imaging techniques for the study of plant growth and development. *Trends in Plant Science*, 19(5):304–310.

Spalding, E. P. and Miller, N. D. (2013). Image analysis is driving a renaissance in growth measurement. *Current Opinion in Plant Biology*, 16(1):100–104.

Sullivan, G. J., Ohm, J., Han, W. J., and Wiegand, T. (2012). Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(12):1649–1668.

W3C (2003). Portable network graphics (PNG) specification. W3C Recommendation ISO/IEC 15948:2003 (E). Available online at http://www.w3.org/TR/PNG/.

Walter, A., Scharr, H., Gilmer, F., Zierer, R., Nagel, K. A., Ernst, M., Wiese, A., Virnich, O., Christ, M. M., Uhlig, B., Jünger, S., and Schurr, U. (2007). Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. *New Phytologist*, 174(2):447–455.

Walter, A. and Schurr, U. (2005). Dynamics of leaf and root growth: Endogenous control versus environmental impact. *Annals of Botany*, 95(6):891–900.

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.

Weinberger, M. J., Seroussi, G., and Sapiro, G. (2000). The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS. *IEEE Transactions on Image Processing*, 9(8):1309–1324.

TABLES

Table 1: One way repeated measurement ANOVA of PLA error measured on 19 Arabidopsis (Col-0) replicate plants over a period of 7 days with 20 repeated measurements of time. 'Time' is the within subjects factor and 'replicate' is the between.

Factor	Sum of Squares	Degrees of Freedom	Mean Square	F	Prob>F[C]
Time (Within)	0.022(0.109)	4.055(72.991)[A]	0.005(0.001)[B]	3.654	**0.00879**
Replicate (Between)	0.178(0.039)	1(18)	0.178(0.002)	81.272	**< 0.00001**

[A] Error term values in parenthesis.
[B] Greenhouse-Geisser corrections are reported to account for deviations from sphericity (Greenhouse-Geisser ε=0.213).
[C] Bold font indicates significance at the $p<0.05$ level.

FIGURES
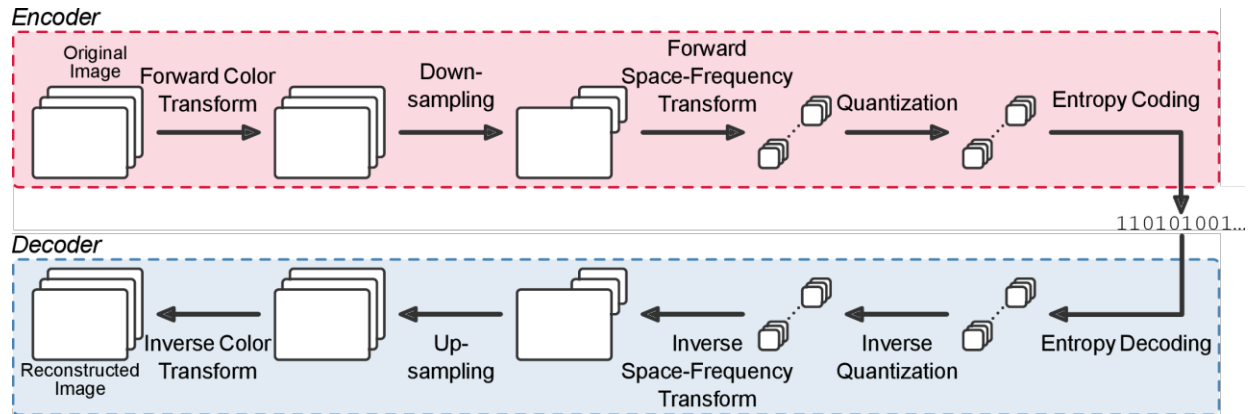
Figure 1: Schematic of a typical encoding and decoding process in lossy image compression, such as JPEG (ITU, 1992). The input image is first converted from the original color space (e.g., RGB) to a representation reducing correlation between color bands (e.g., YC_bC_r (ITU, 1995)). Each color component, possibly after some downsampling, is split into independent coding units (e.g., blocks of 8×8 pixels). Space-frequency transformation permits not only spatial decorrelation but identifies information to be selectively discarded through quantization. The transformed and quantized coefficients are further compressed by an entropy coding stage using lossless approaches. This results in a bit-stream arranged according to a predetermined syntax, that can be decoded performing all previous operations in reverse order, to obtain an approximation of the original image.
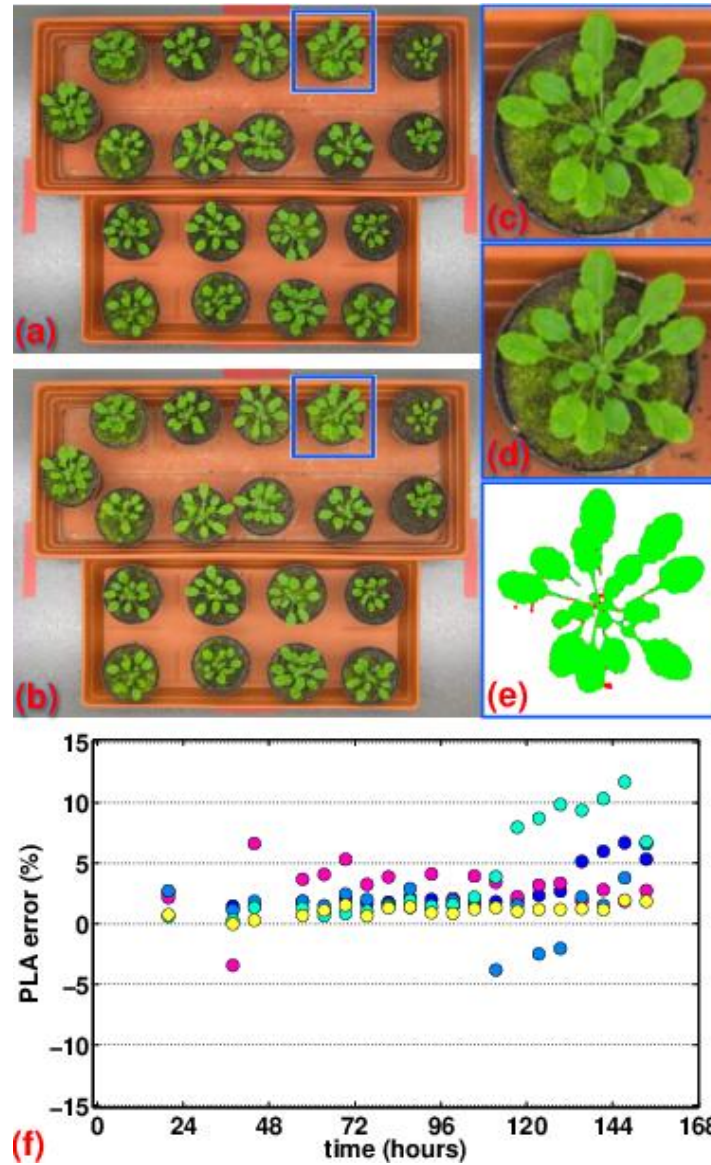
Figure 2: Compression affects growth observations in Arabidopsis. (a) An uncompressed image of 19 individuals of *Arabidopsis thaliana* ecotype Col-0 on day 17 after germination; (b) shows the same image compressed with JPEG; (c) and (d) zoom in detail of the plant in the blue bounding box uncompressed and compressed, respectively; (e) illustrates color coded the segmentation outcome of automatically analyzing this plant using uncompressed or compressed data: green pixels that are identified as plant on both images, red (false positives), and blue (false negatives), as those identified only in the compressed or uncompressed image respectively; (f) plots PLA error (%) of the top 5 plants over 6 days covering the days 12-17 after germination (5 plant measurements are shown only for presentation clarity, and similar trends are observed for all 19 plants). The same colored dot is used for the same plant.
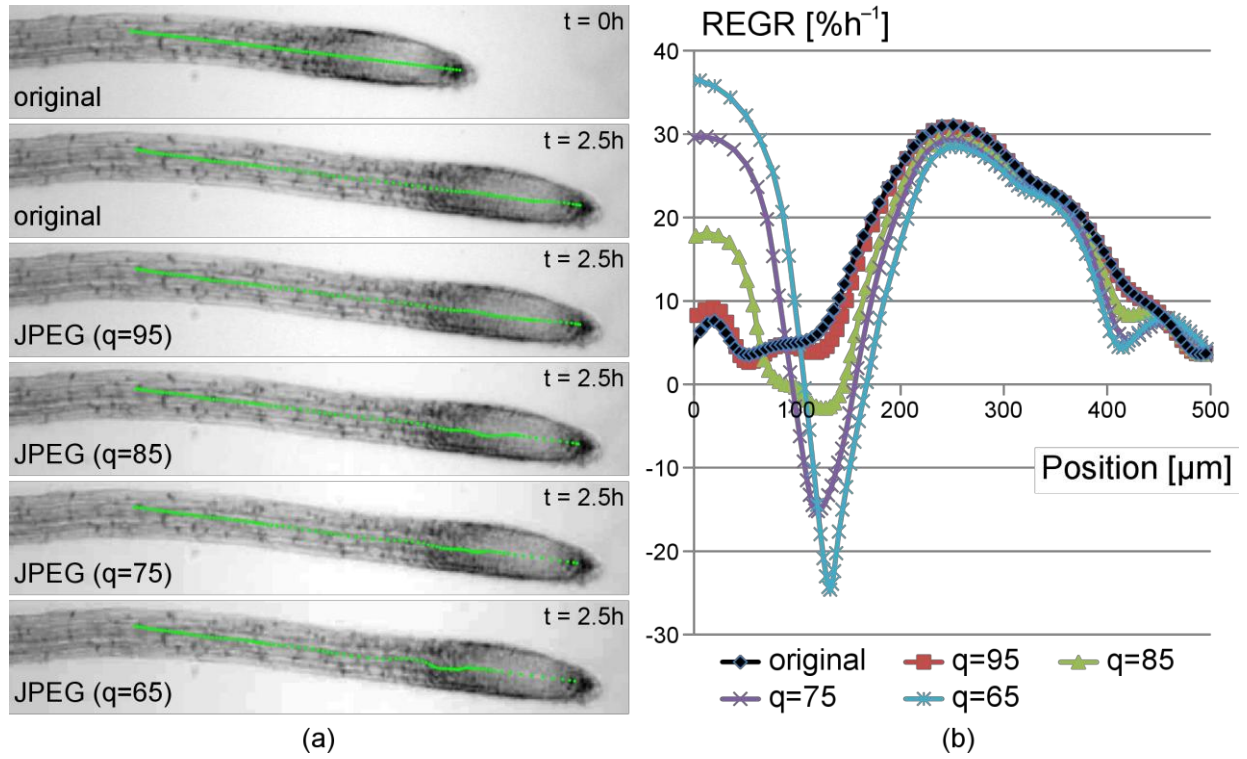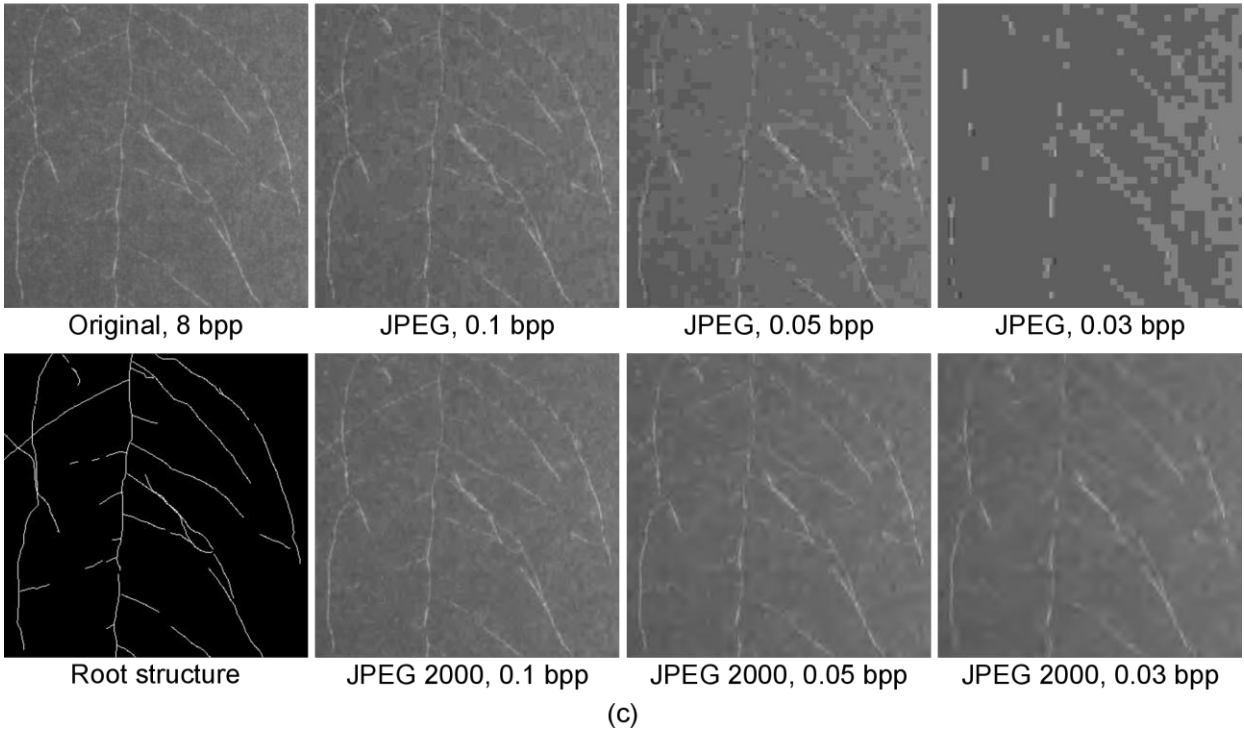
Figure 3: (a) Example images used for root growth analysis. Root tracking results obtained with the method described in (Chavarría-Krauser et al., 2008) are shown as green lines, where distance between points denotes estimated local growth intensity. (b) Spatial growth (REGR) profiles obtained.
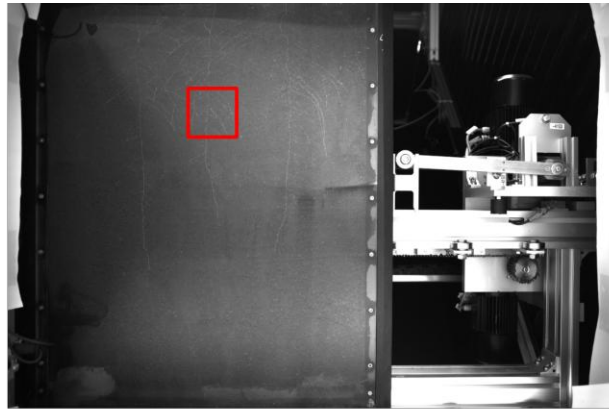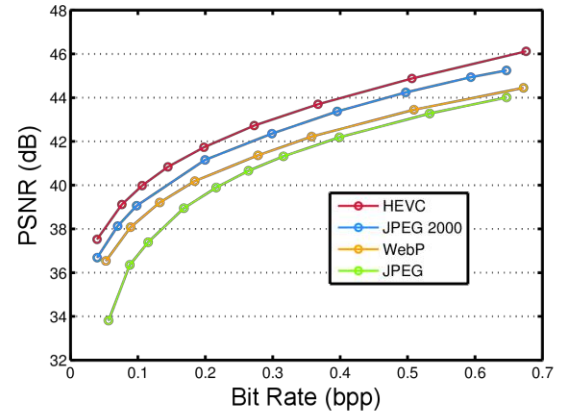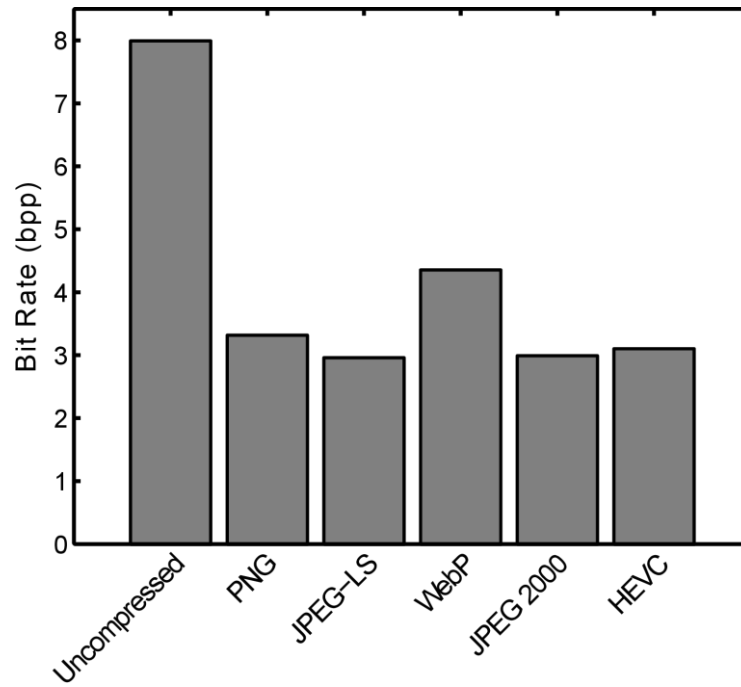
Figure 4: (a) Root systems of three rapeseed subjects imaged at the GROWSCREEN-Rhizo (Nagel et al., 2012). (b) Image fidelity obtained for the image in (a), after compression with different standards. (c) Detail in the red box of (a), compressed at various bit rates using the JPEG and JPEG 2000 standards.

Figure 5: Compression performance obtained by lossless coding standards on the gray scale root image of Figure 4a. Baseline for the comparison is size of uncompressed image (the leftmost bar at 8 bpp).
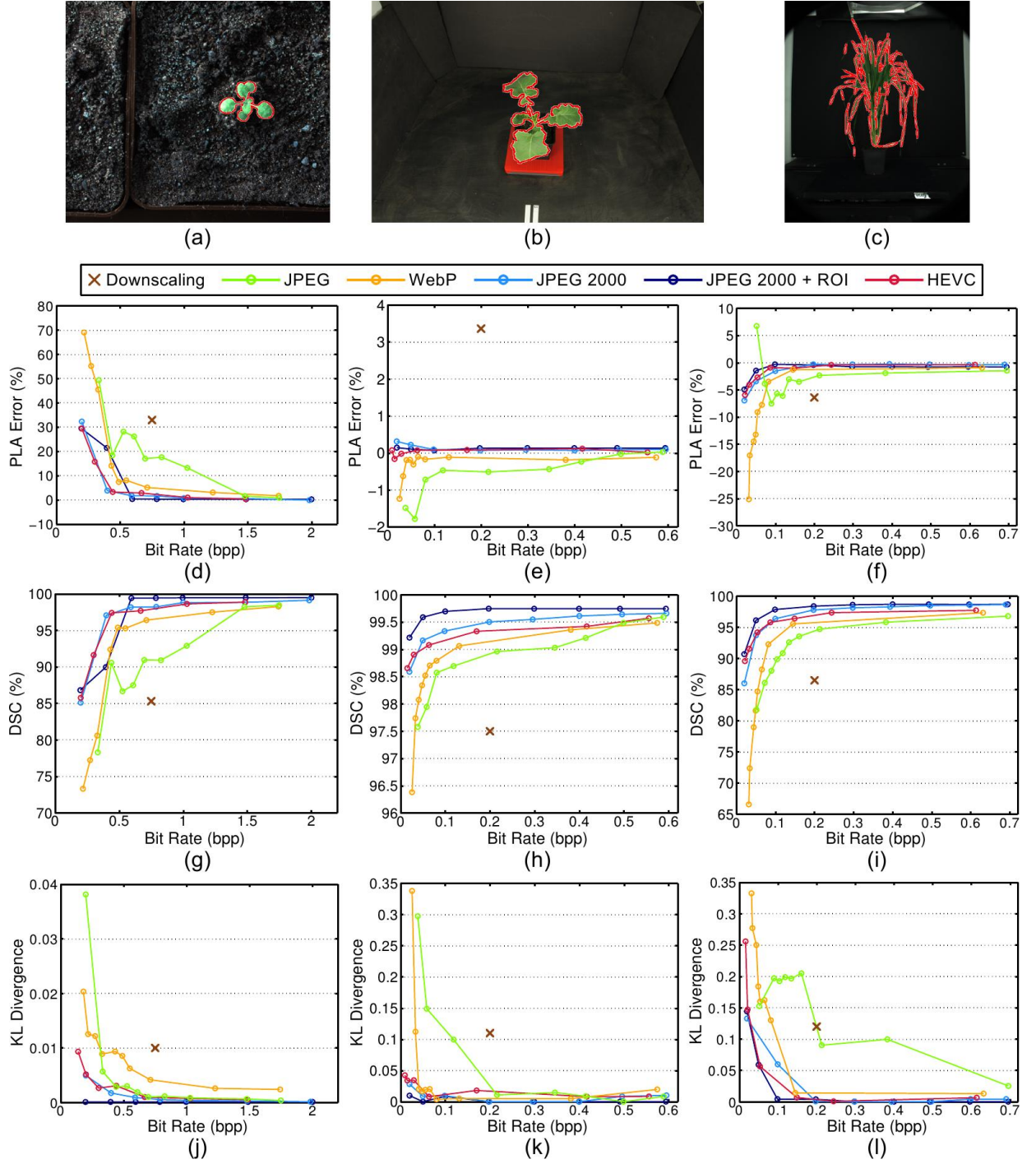
Figure 6: Compression performance obtained with color images of plant shoots: (a) Arabidopsis (b), rapeseed, and (c) maize. Segmentation contour obtained with the method in (Briese et al., 2013) is overlaid in red. (d)-(f) PLA Error and (j)-(l) KL Divergence should be as close to 0 as possible. (g)-(i) Best possible DSC value is 100%.
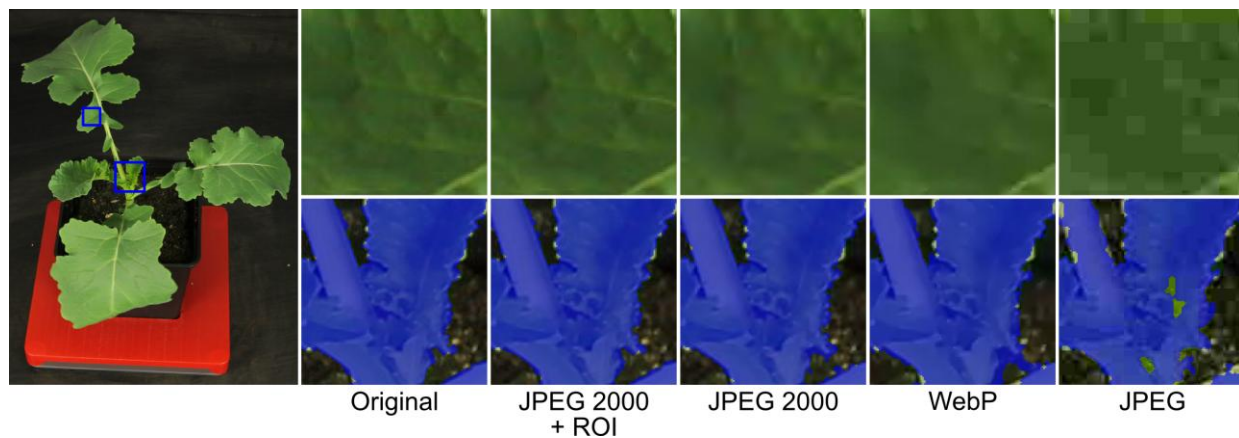
Figure 7: Left: a rapeseed test image; top and bottom rows: zoom in details in the blue bounding boxes, compressed at 0.05 bpp using different compression standards. In the bottom row, the segmentation mask obtained with the method in (Briese et al., 2013) is overlaid in blue.
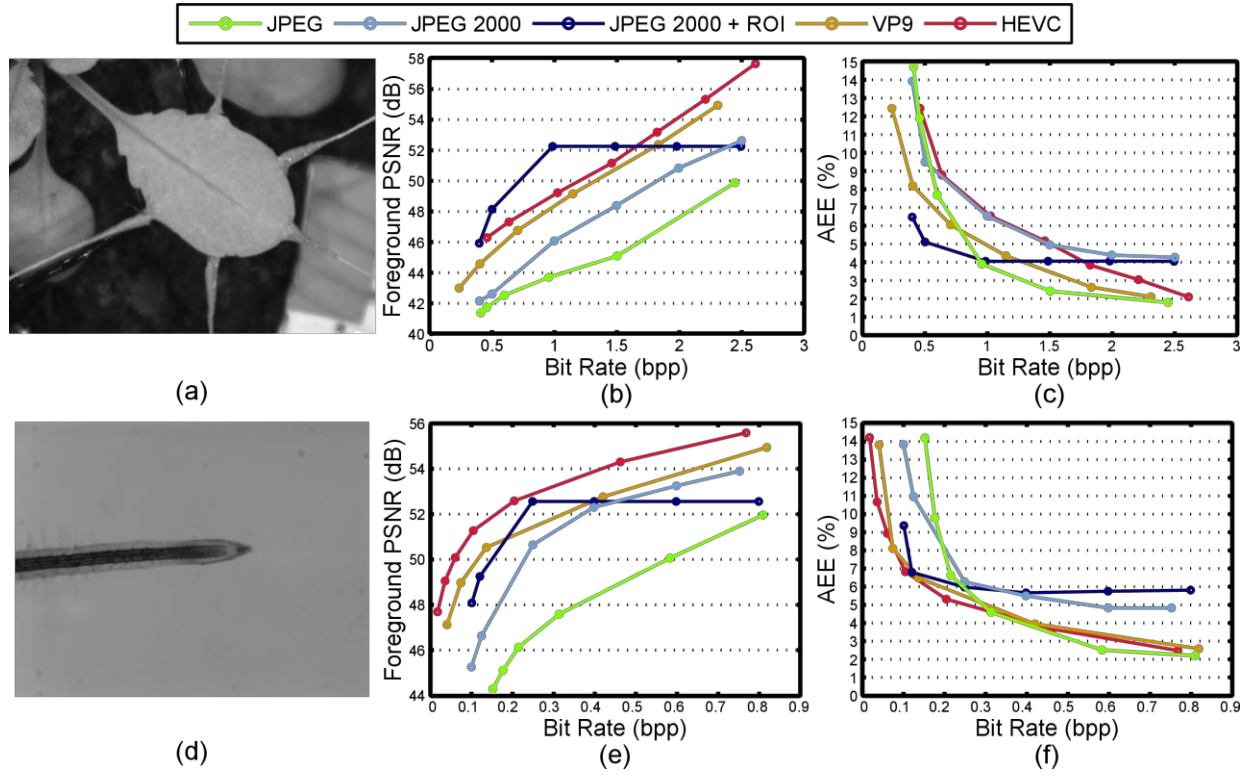
Figure 8: Performance of lossy coding approaches on image data used for optical flow based analysis: (a)-(c) Arabidopsis leaf sequence, (d)-(f) tobacco root tip sequence.

Appendices

A. Compression software and command line options

In this section we expand on the standards and outline parameters used and implementation details.

Portable Network Graphics (PNG) (W3C, 2003) is a lossless compression standard, which uses a filtering function to enable spatial decorrelation and a compression algorithm (Deflate) similar to that of the ZIP file format. It is fast in decoding and can handle both gray scale and color images (RGB).

JPEG-LS (Weinberger et al., 2000) is a lossless and near-lossless compression standard, based on the LOCO-I algorithm (LOw COmplexity LOssless COmpression for Images), and features low computational and memory requirements.

JPEG is a lossy image compression algorithm, based on the discrete cosine transform (ITU, 1992). JPEG is the most widely adopted compression standard for digital photography (it is default format on most commercial-grade cameras) and for image coding on the Web.

JPEG 2000 (Skodras et al., 2001) is a lossless and lossy image coding standard based on the discrete wavelet transform. Notably, JPEG 2000 is capable of native region-of-interest (ROI) coding (Christopoulos et al., 2000), a feature allowing to encode foreground image regions at a higher quality than background regions.

WebP is based on the methodology adopted to compress keyframes in the VP8 video coding standard (Bankoski et al., 2011) for the purpose of royalty-free lossless and lossy image compression.

VP9 (Mukherjee et al., 2013) is a new open and royalty-free library for lossless and lossy video coding.[4] VP9 employs several modern coding tools and is mainly intended for high definition video and targets low decoding complexity.

High Efficiency Video Coding (HEVC) (Sullivan et al., 2012) is the latest generation video coding standard (ITU, 2013). Similar to its predecessors (e.g., H.264 and MPEG-4), HEVC employs sophisticated techniques for intra prediction and motion compensation, in order to address, respectively, spatial and temporal correlation in high definition video signals.

Codec software implementations adopted in the experiments and command line options used to execute the encoders, are listed below.

- *PNG*, libpng v1.6.12 (`http://www.libpng.org`):

  - Lossless: `pnmtopng -compression=9 -comp_mem_level=9 -paeth -comp_window_bits=8 -comp_strategy=filtered`

- *JPEG-LS*, Hewlett-Packard reference encoder v1.0 (`http://www.hpl.hp.com/research/info_theory/loco/`):

  - Lossless: `locoe`

- *JPEG*, libjpeg v9a (`http://www.ijg.org`):

  - Lossy: `cjpeg -dct float -progressive -arithmetic -quality` q

---

[4] `http://www.webmproject.org`

where the quality factor *q* is an integer in the range from 0 (lowest quality, small file) to 100 (best quality, big file).

- *JPEG 2000*, Kakadu v7.4 (`http://www.kakadusoftware.com`):

  o Lossless: `kdu_compress Creversible=yes`

  o Lossy: `kdu_compress -no_weights -rate` *r*

  o Lossy (ROI): `kdu_compress -no_weights Rshift=16 Rlevels=5`
     `-roi` *roifile*`,0.5 -rate` *r*

  where *r* is a float denoting the desired bit rate (bpp), and *roifile* is a PGM file containing the ROI mask.

- *WebP*, libwebp v0.4.1 (`https://developers.google.com/speed/webp/`):

  o Lossless: `cwebp -lossless -m 6 -q 20`

  o Lossy: `cwebp -q` *q*

  where *q* is a quality factor in the range from 0 (lowest quality, small file) to 100 (best quality, big file).

- *VP9*, libvpx v1.3 (`http://www.webmproject.org/vp9/`):

  o Lossy: `vpxenc –codec=vp9 –passes=1 –tune=psnr`
     `–end-usage=cbr-target-bitrate=`*r*

  where *r* is a float denoting target bitrate (kbps).

- *HEVC*, HM v16.0 (`http://hevc.hhi.fraunhofer.de`), and x265, v1.3 (`http://x265.org`):

  o Lossless (gray scale): `TAppEncoderStatic –Profile=main-RExt`
     `–InputChromaFormat=400 –TransquantBypassEnableFlag=1`
      `–CUTransquantBypassFlagValue=1`

  o Lossy: `x265 -qp` *q*

  where *q* is an integer in the range from 0 (best quality, big file) to 51 (lowest quality, small file).

For JPEG-LS and JPEG 2000 we adopt the pre-compiled software libraries provided by the authors, whereas for the others we compile the libraries from source code.

Note that, although HM is the reference implementation of HEVC, x265 achieves superior time performance and is used here for lossy compression. On the other hand, HM is used for lossless compression, since to this day this feature is not supported by x265. To date, lossless compression of color images is not possible with HEVC, due to the chroma sub-sampling strategy mandated by current implementations.

Video codec implementations used in this study accept input data only in the YUV 4:2:0 format (i.e. one luminance component, Y, followed by two chrominance components, U and V, downsampled by a factor of two both horizontally and vertically). Hence, RGB color images are converted to the YC_bC_r color space (ITU, 1995) and chroma sub-sampled prior to encoding with VP9 and HEVC (observe that JPEG and WebP perform analogous operations internally,

as part of their coding strategy, whereas JPEG 2000 does not recommend chroma sub-sampling). Gray scale images are embedded into a YUV formatted byte stream, by augmenting the luminance component with uniform zero-valued chroma components (note that this operation does not affect coding efficiency).

B. Encoding and decoding execution times

To illustrate computational complexity of the codec implementations used in this work, we outline in Tables B.1 and B.2, execution times for encoding/decoding of still images and video sequences, respectively. Computational experiments are conducted on a machine equipped with Intel Core 2 Duo CPU E8200 2.66 GHz and 4 GB memory, running 64-bit GNU/Linux.

For lossless compression (cf. Table B.1), JPEG-LS and JPEG 2000 obtain encoding and decoding times below 1.5 seconds. Fastest decoding is achieved by WebP, which however presents higher encoding time than other still image compression standards. Encoding and decoding times of HEVC (using the HM implementation) are significantly higher than other codecs. Overall, shorter execution times are observed when using lossy compression (cf. Table B.1). JPEG and JPEG 2000 achieve shortest encoding times, while average decoding times of color images remain below one second for all codecs, with fastest decoding obtained by WebP.

For image sequences (cf. Table B.2), still image codecs are generally faster at encoding, with JPEG 2000 requiring on average less than half a second to encode the test sequences. Decoding times are in the same order of magnitude for all codecs (VP9 presents shortest decoding times). The JPEG 2000 + ROI approach results in longer execution times than JPEG 2000, due to the ROI coding feature and no chroma sub-sampling (i.e. more data to process in the entropy coding stage of the encoder). Among video codecs, VP9 is 3 to 4 times slower than HEVC at encoding, but approximately an order of magnitude faster at decoding.

Table B.1: Average execution times for encoding and decoding still images, expressed as mean ± standard deviation. Best results (i.e. less time) are highlighted in bold. For lossy compression, average results are shown for images compressed at a range of bit rates between 0.02 and 2 bpp.

Standard	Gray scale images		Color images	
	Encoding (s)	Decoding (s)	Encoding (s)	Decoding (s)
Lossless				
PNG	3.93±0.02	0.63±0.03	5.79±4.29	0.47±0.32
JPEG-LS	0.97±0.03	1.10±0.01	1.36±0.95	1.52±1.05
JPEG 2000	**0.74±0.02**	0.73±0.08	**0.78±0.40**	0.78±0.35
WebP	8.60±0.10	**0.46±0.00**	7.27±4.22	**0.37±0.26**
HEVC (HM)	96.24±0.14	2.75±0.02	–	–
Lossy				
JPEG	**0.40±0.09**	0.30±0.09	**0.42±0.25**	0.34±0.21
JPEG 2000	**0.40±0.11**	**0.21±0.09**	0.44±0.30	0.27±0.20
JPEG 2000 + ROI	0.67±0.18	**0.21±0.07**	1.13±0.69	0.33±0.24
WebP	3.46±0.27	0.33±0.05	2.26±1.39	**0.20±0.12**
HEVC (x265)	13.77±2.02	1.36±0.24	8.21±5.14	0.75±0.51

Table B.2: Average execution times obtained to encode and decode the sequences for optical flow analysis at a variety of bit rates using lossy coding standards, expressed as mean ± standard deviation. Best results (i.e. less time) are highlighted in bold.

Standard	Arabidopsis leaf		Root tip	
	Encoding (s)	Decoding (s)	Encoding (s)	Decoding (s)
JPEG	0.21±0.06	0.17±0.07	0.96±0.17	0.73±0.21
JPEG 2000	**0.14±0.03**	0.10±0.04	**0.41±0.07**	0.28±0.09
JPEG 2000 + ROI	0.54±0.07	0.21±0.08	2.64±0.09	0.84±0.39
VP9	6.82±1.39	**0.05±0.04**	30.29±7.36	**0.16±0.02**
HEVC (x265)	2.08±0.40	0.47±0.09	6.75±3.51	1.61±0.64