Black Spot: a platform for automated and rapid estimation of leaf area from scanned images

Varun Varma[1] **and Anand M. Osuri**[1,2]

(1) National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bellary Road, Bangalore, 560065, Karnataka, India

(2) Nature Conservation Foundation, 3076/5, IV Cross, Gokulam Park, Mysore, 570002, Karnataka, India

Corresponding author: Varun Varma

Email: varunvarma17@gmail.com

Note: This is an author-generated copy. The final publication is available at http://link.springer.com/article/10.1007%2Fs11258-013-0273-z/fulltext.html

Software, scripts and user support: http://www.ncbs.res.in/blackspot.html

Citation: Varma, V & Osuri, A.M. (2013). Black Spot: a platform for automated and rapid estimation of leaf area from scanned images. Plant Ecology. DOI 10.1007/s11258-013-0273-z (Refer to http://link.springer.com/article/10.1007%2Fs11258-013-0273-z/fulltext.html for up to date citation)

Abstract

Leaf area and its derivatives (e.g. specific leaf area) are widely used in ecological assessments, especially in the fields of plant–animal interactions, plant community assembly, ecosystem functioning and global change. Estimating leaf area is highly time-consuming, even when using specialized software to process scanned leaf images, because manual inputs are invariably required for scale detection and leaf surface digitisation. We introduce Black Spot Leaf Area Calculator (hereafter, Black Spot), a technique and stand-alone software package for rapid and automated leaf area assessment from images of leaves taken with standard flatbed scanners. Black Spot operates on comprehensive rule-sets for colour band ratios to carry out pixel-based classification which isolates leaf surfaces from the image background. Importantly, the software extracts information from associated image meta-data to detect image scale, thereby eliminating the need for time-consuming manual scale calibration. Black Spot's output provides the user with estimates of leaf area as well as classified images for error checking. We tested this method and software combination on a set of 100 leaves of 51 different plant species collected from the field. Leaf area estimates generated using Black Spot and by manual processing of the images using an image editing software generated statistically identical results. Mean error rate in leaf area estimates from Black Spot relative to manual processing was −0.4 % (SD = 0.76). The key advantage of Black Spot is the ability to rapidly batch process multi-species datasets with minimal user effort and at low cost, thus making it a valuable tool for field ecologists.

Introduction

The surface area of individual leaves is a property of considerable interest and utility in plant sciences (O'Neal et al. 2002). Leaf area is closely associated with a number of physiological processes relating plants to their environment and resources, with widespread applications in ecology, global change biology and agriculture (Anyia and Herzog 2004; Corre-Hellou et al. 2006; Kleiman and Aarssen 2007; Reich et al. 1998). Leaf area is further used to derive metrics that are extensively used in ecosystems and functional ecology. For instance, the plant functional trait specific leaf area (SLA = leaf area/leaf dry weight) is closely linked to plant strategy schemes and defines the position of plants along the fundamental resource capture–resource storage axis (Wright et al. 2004). Traits such as SLA can be quite variable, necessitating large sample sizes for robust estimates (Wilson et al. 1999).

While the need to estimate leaf area for a variety of ecological assessments is clear, data collection (Cornelissen et al. 2003) requires high user time and effort. While leaf outlines were manually traced onto graph paper in the pre-computer era, more modern methods either utilise specialized equipment (e.g. LI-COR LI-3000C Portable Leaf Area Meter), or digital image processing software coupled with low-budget hardware such as desktop scanners to estimate leaf area. The underlying operating procedure for most existing software involves (1) processing the leaf image to separate the leaf from the image background, (2) creating a calibration scale using an image of an object of known dimensions (e.g. a ruler), and (3) estimating leaf area using leaf pixel counts and the calibration scale. In carrying out these steps, available software differ in the nature and amount of user intervention required during the leaf digitization and calibration steps. Extraction of leaf pixels from the image background requires either complete manual processing (e.g. Igathinathane et al.2006), or manual assignment of pixel intensity thresholds (e.g. ImageJ; Rasband 2011), or is semi-automated using standard image processing software (O'Neal et al. 2002).

More sophisticated and specialized applications automate the extraction of leaf foreground using algorithms based on the spectral properties of leaves, such as low values of blue band reflectance (e.g. LAMINA: Bylesjö et al. 2008; SIR: Femat-Diaz et al. 2011), and segmentation to detect leaf bounding polygons (Bylesjö et al. 2008). However, these softwares still require a manual step to assign image scale, and further, apply spectral rules which are often calibrated to work accurately on single to small pools of species. As a result, batch-processing large, multi-species datasets is not possible.

We introduce Black Spot, a free stand-alone software platform and method for rapid and automated leaf image processing and leaf area estimation from scanned images of leaves. The name Black Spot recalls an early version of the software which required users to scan leaves against a background containing a black square of known area, which served as a reference for scale calibration (in the current version, scale calibration does not require a black square in the background). The objectives of developing Black Spot were to (1) provide field ecologists an efficient, yet cost-effective batch-processing tool to rapidly process large multi-species datasets of leaf images, (2) to minimise the amount of user input and, therefore, user time required during estimation, (3) to output data in a readily usable format for statistical analysis, and (4) to output adequate information to the user for error checking.

Development of Black Spot

Black Spot is written in Python (version 2.7; http://www.python.org) and utilises tools within Python Imaging Library (PIL; http://www.pythonware.com/products/pil), NumPy (version 1.6; http://www.numpy.org) and Scipy (version 0.11; http://www.scipy.org) modules, and is compiled to run as a windows executable application using the py2exe (http://www.py2exe.org) module. The software and a detailed user manual are available as a 12.4 MB download at http://sourceforge.net/projects/blackspot. At the time of writing, Black Spot is available for Windows operating systems and can also be run on Linux distributions and Mac OSX using a Windows emulator.

Black Spot requires users to scan images of leaves using a plain white paper sheet as a background. The software classifies each image into leaf (foreground) and non-leaf (background) components by pixel-based classification using spectral rule-sets operating on image band ratios as opposed to pixel intensities. The underlying principles of the rule-sets are as follows: (a) the white background has an equal reflectance across the red, green and blue bands. Therefore, the background's band ratio will tend towards one; (b) fresh leaves absorb a portion of incident light (e.g. high red absorbance in green leaves), resulting in a strong deviation from one in the reflected band ratios. Importantly, the use of band ratios also allows the software to reduce errors due to the presence of shadows in the image. Here the assumption is that shadows result in a lowering of pixel intensities without much alteration in the relative intensities between bands. Hence, shadows of leaf falling on the white paper, though of lower intensity, will share the same band ratio as that of paper, while shadows of leaf on leaf (which is a common problem when scanning leaves with undulating surfaces) will deviate from one.

The band ratio-based spectral classification rule-sets were developed and calibrated using a multi-species image database of 800 images. At the time of writing, Black Spot has inbuilt spectral rule-sets for Canon and Hewlett Packard scanners (tested on the following models: HP Scanjet G3110, HP Deskjet F4488, Canon LiDE 110 and Canon LiDE 25). During the development of these rules, we observed consistency across models within scanner brands in spectral properties of leaf images, which may be due to similar image capture and processing technologies. Therefore, we believe that the spectral rule-sets developed can be applied across models within brands.

Running Black Spot

Stage 1: capturing leaf images

The procedure to use Black Spot involves scanning fresh leaves using a flatbed scanner in .jpg format (Fig. 1a, b). During this scanning step, it is recommended that the user paste a plain white sheet to the inner lid of the scanner to prevent permanent staining of the scanner lid. This white sheet additionally serves as an easily identifiable background and should, therefore, be free of any annotations or marks. Depending on the resolution of data required, leaf samples may either comprise single leaves, or a set of leaves from a plant that might comprise one replicate in species level comparisons (Garnier et al. 2001). All scanned images must be stored in a single folder for processing with Black Spot. The Black Spot user manual provides detailed guidelines for scanning leaves in a variety of species cases, such as large leaves, leaves with particularly thick petioles and rachis, and compound leaves.

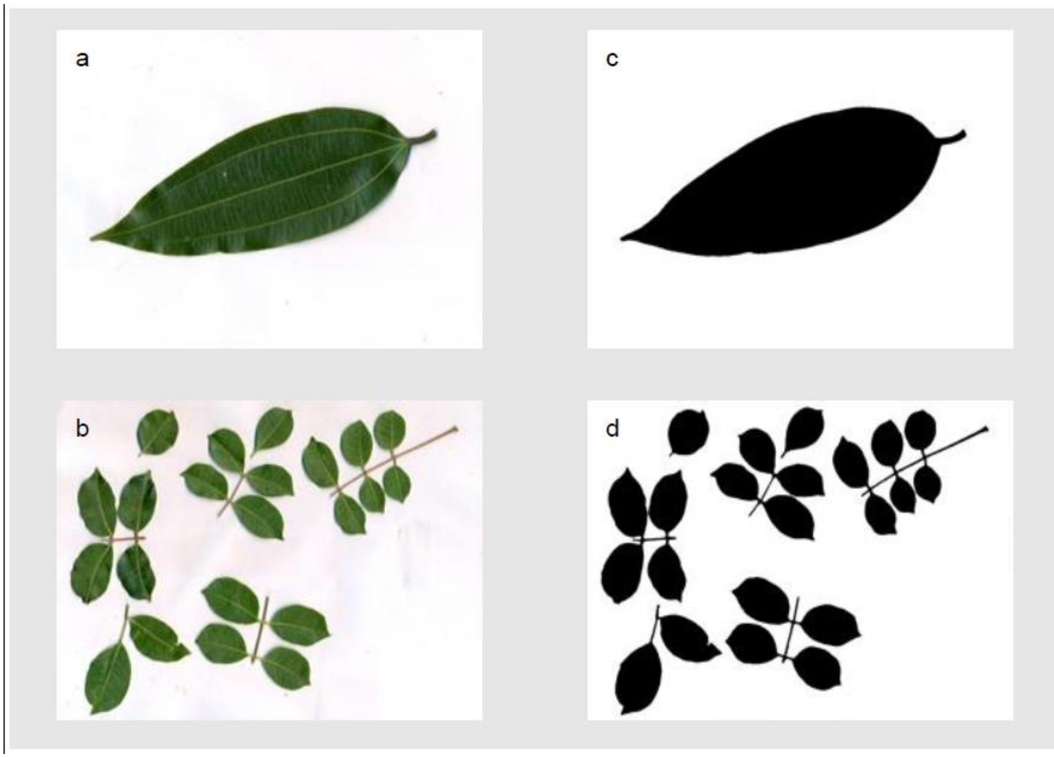

Fig. 1*: Examples of input scanned images for a simple (Cinnamomum malabatrum) and b compound (Sclerocarya birrea) leaves.Processing with Black Spot results in output leaf mask images (c) and (d), which the user can utilise for error checking*

Stage 2: running Black Spot and initial user input

Once scanning is complete, the Black Spot software is run to estimate leaf area from all captured images. The software is launched using an executable file which opens up a shell window. Prior to image processing, users are required to input information on the (1) path to the folder containing the leaf images, (2) desired name for the output folder to save results, (3) desired size of a smoothening filter (either 3 × 3, 5 × 5, 7 × 7 or 9 × 9 pixels), and (4) desired breadth of the image margin to exclude from the analysis. The smoothening filter (step 3) serves to correct for small artefacts which may give rise to speckling. For a 300-dpi image, the side of the moving window ranges in size from 0.25 mm for a 3 × 3 window to 0.8 mm for a 9 × 9 window. The choice of window sizes provided to users is deliberately constrained to odd numbers (for symmetry around the focal pixel) and numbers no greater than nine, as larger window sizes may reduce the accuracy of leaf area estimates. The exclusion of image margins from analysis (step 4) is required to correct for a common source of error during image capture, which arises from light leakage into the scan area, usually manifested as darkening around the image margins. Finally, the user is given the option to resize input images, which might be required in case of very large input file sizes. Following these steps, Black Spot begins image processing.

Stage 3: image processing

Figure 2 contains a schematic representing the work flow of Black Spot. Each image in the input folder is processed sequentially. First, based on user input (stage 2), images are resized and margins excluded. Then, for each image, the exchangeable image file format (exif) information is extracted. The exif information contains the scanner brand used to capture the leaf image, as well as the pixel resolution during image capture. The latter is recorded as dots per inch (dpi) and in flat-bed scanners directly translates to image scale. Based on the scanner brand identified, the software applies the appropriate spectral rule-set to identify leaf from background.

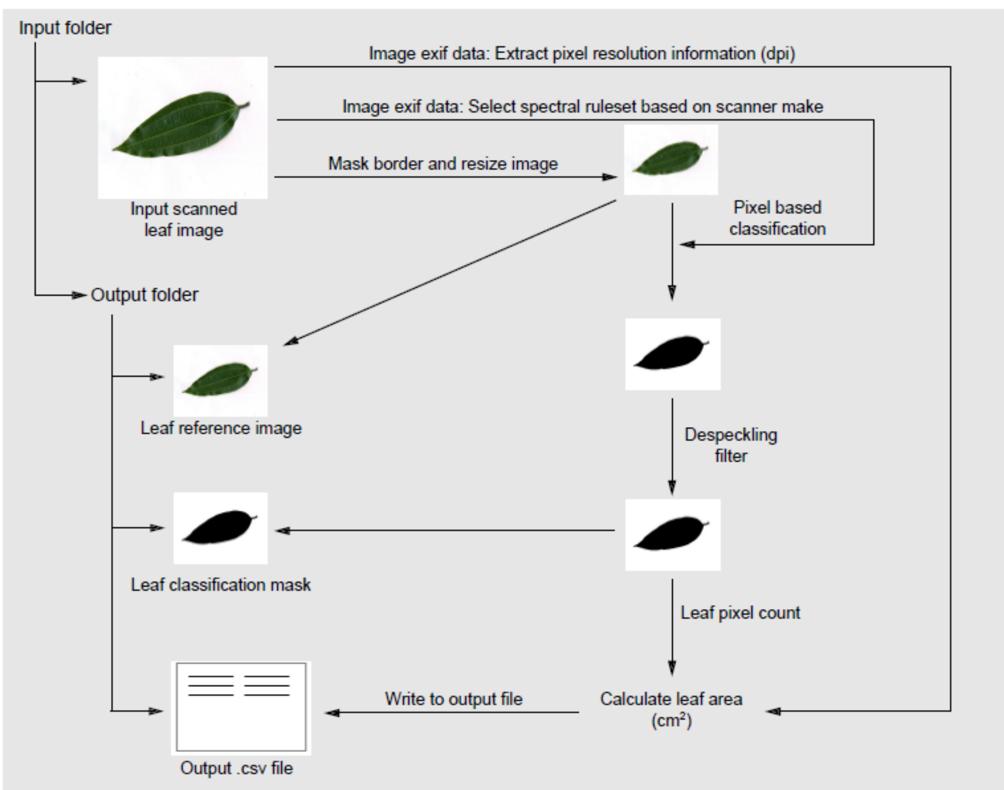

Fig. 2 *Schematic representation of Black Spot work flow*

A first pass of the pixel-based classification is followed by the application of a de-speckling filter over the classified image, the size of which is specified by the user (stage 2). The classification and filter processes result in a binary image (leaf and non-leaf) where the number of leaf pixels is determined. Leaf area is then calculated in square centimetres (cm^2) by utilising pixel resolution values extracted from the image exif data.

Stage 4: output and error checking

As processing of each image is completed, an output comma separated value (csv) file is updated with one row of data containing the image name, a flag reflecting whether image exif data was found, scanner brand, pixel resolution, number of leaf pixels detected, estimated leaf area and an error flag indicating any discrepancies during processing. This csv file is stored in an output folder which Black Spot creates based on user input (stage 2). In a situation where exif information is not found, Black Spot will only report a count of leaf pixels, not estimate leaf area, and will flag the sample in the error flag field.

To allow for error checking, Black Spot also stores each leaf image and its classification image (leaf mask; Figs. 1c, d and 2) in the output folder. The user can access these images at a later time to visually check for errors in leaf classification. Images where errors are deemed too large by the user may be edited in image-processing software and reprocessed in Black Spot. A detailed description of commonly encountered errors, and editing required to correct these prior to reprocessing is provided in the Black Spot user manual.

Performance

To check the accuracy of leaf area estimates from Black Spot, we compared the results of the software to manual processing of 100 leaves collected from the field and processed using an image-editing software (GNU Image Manipulation Program; GIMP 2.8; http://www.gimp.org). Test images comprised 51 species spanning rainforest trees in southern India, woody plants and forbes from the Indian trans-Himalaya, and savannah trees from southern Africa. All images in the test dataset were exclusive of the calibration dataset. Manual processing involved using the image-editing software's foreground and colour selection tools to manually select the colour range for the leaf by sampling leaf pixels from the original image. This was followed by manually correcting boundaries of leaves in the selection by adding or subtracting selections interactively. The leaf area estimated by manual processing was compared to that obtained using Black Spot for the same leaf images.

A paired t test revealed no significant differences in the number of pixels classified as leaf between the two methods ($t = 0.0336$, $df = 198$, $p = 0.97$). On average, Black Spot's estimates deviated from manual estimates by −0.4 % (SD = 0.76). Importantly, while manual processing of the 100 test leaf images took a total of 220 min and 18 s, Black Spot completed the task in 20 min and 12 s (images were not resized; Black Spot was run on a PC running Windows XP SP3, Intel Core 2 Duo 2.93 GHz and 3 GB RAM).

Conclusions

The requirements for multi-species datasets (e.g. Kattge et al. 2011) are growing in ecology and vegetation science, particularly in the realm of plant functional traits research. Trait values derived from leaf area measurements (e.g. SLA) play an important role in understanding ecosystem responses to environmental perturbations such as global change. However, data collection poses considerable challenges due to a high variability in leaf type, size, shape and colour. These along with the requirement of calibrating image scale makes both image capture and processing to provide leaf area estimates demanding on user time and effort.

The key features of Black Spot include an automated detection of image scale and an inbuilt set of spectral rules to differentiate leaf from image background. These features allow for batch-processing datasets which contain images of leaves across multiple species to produce highly accurate estimates of leaf areas, while considerably reducing user input. Further, Black Spot serves to standardize data quality when data are collected by large teams or comparisons are made across datasets and regions. Importantly, Black Spot is freely downloadable, is not dependant on expensive hardware or fast processors, and requires a very basic knowledge of computers to use. While Black Spot is presently calibrated to work best with Canon and Hewlett Packard scanners, we expect feedback from users using different scanner brands to further expand the scope of the software.

Acknowledgments

We acknowledge Yadugiri VT, Chengappa SK, Vijay Kumar S, Siddharth Iyengar, Rutuja Dhamale, Priyanka Runwal, Atul Joshi and Harinandan PV for volunteering time to collect field samples, manually process images and test out versions of the software. We are grateful to Mahesh Sankaran, Jayashree Ratnam and Fiona Savory for several useful discussions. We are grateful to two anonymous reviewers for suggestions that have improved the quality of this manuscript. We thank the Rufford Small Grants Foundation and National Centre for Biological Sciences (NCBS) for funding fellowships, fieldwork and equipment. We acknowledge NCBS, and in particular Prasanta Baruah, for helping create the Black Spot web page and user support systems.

References

Anyia AO, Herzog H (2004) Water-use efficiency, leaf area and leaf gas exchange of cowpeas under mid-season drought. Eur J Agron 20:327–339

Bylesjö M, Segura V, Soolanayakanahally RY, Rae AM, Trygg J et al (2008) LAMINA: a tool for rapid quantification of leaf size and shape parameters. BMC Plant Biol 8:82

Cornelissen JHC, Lavorel S, Garnier E, Diaz S, Buchmann N et al (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Aust J Bot 51:335–380

Corre-Hellou G, Fustec J, Crozat Y (2006) Interspecific competition for soil N and its interaction with N2 fixation, leaf expansion and crop growth in pea–barley intercrops. Plant Soil 282:195–208

Femat-Diaz A, Vargas-Vazquez D, Huerta-Manzanilla E, Rico-Garcia E, Herrera-Ruiz G (2011) Scanner image methodology (SIM) to measure dimensions of leaves for agronomical applications. Afr J Biotechnol 10:1840–1847

Garnier E, Shipley B, Roumet C, Laurent G (2001) A standardized protocol for the determination of specific leaf area and leaf dry matter content. Funct Ecol 15:688–695

Igathinathane C, Prakash VSS, Padma U, Babu GR, Womac AR (2006) Interactive computer software development for leaf area measurement. Comput Electron Agric 51:1–16

Kattge J, Diaz S, Lavorel S, Prentice IC, Leadley P et al (2011) TRY—a global database of plant traits. Glob Change Biol 17:2905–2935

Kleiman D, Aarssen LW (2007) The leaf size/number trade-off in trees. J Ecol 95:376–382

O'Neal ME, Landis DA, Isaacs R (2002) An inexpensive, accurate method for measuring leaf area and defoliation through digital image analysis. J Econ Entomol 95:1190–1194

Rasband WS (2011) ImageJ. US National Institutes of Health, Bethesda

Reich PB, Ellsworth DS, Walters MB (1998) Leaf structure (specific leaf area) modulates photosynthesis-nitrogen relations: evidence from within and across species and functional groups. Funct Ecol 12:948–958

Wilson PJ, Thompson KEN, Hodgson JG (1999) Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. New Phytol 143:155–162

Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z et al (2004) The worldwide leaf economics spectrum. Nature 428:821–827