Imaging and image analysis: the next frontier
Thurs 10:30 AM

Nan An, Steve Welch, Jeff White

1. Potential applications of image analysis
    a. Estimation of size or number
        i. Canopy cover
        ii. Flower, spike, or pod number
    b. Canopy architecture
        i. 3D distribution of leaves in canopy
        ii. Leaf area, angle, and position
    c. Shapes, etc.
        i. Leaf shapes
    d. Spatial variation in physiological responses
        i. Canopy architecture
        ii. Canopy reflectance
2. Overview
    a. Image acquisition – Camera spectra; camera/vehicle issues; lighting issues; metadata; throughput; camera body and lens selection; camera positioning
    b. Image analysis – Geometric processing; lens distortion removal; color correction; pixel *vs.* vector methods; canopy *vs.* background
    c. Image pipeline design – Flowcharting; some examples
3. Image acquisition
    a. Camera spectra – three kinds of cameras
        i. RGB cameras involve a chip with light sensitive cells covered by filters with extended transmission ranges which, however, peak in the red, green, and blue regions. A filter blocks all infrared wavelengths.
        ii. A near infrared camera (NIR) can use the same chip but with the infrared filter replaced by one with a narrow pass band A camera designed to sense bands used to compute the normalized difference vegetation index (NDVI) blocks non-peak bands sensed by cells under the green and red Bayer filters and utilizes the NIR wavelengths that are also passed by the blue Bayer filters. NDVI is defined as

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

        Other useful indices can also be calculated from the output of this camera chip.
    b. Various factors must be interrelated in designing an image acquisition system no matter whether the camera is mounted on a ground vehicle, a drone, or a fixed surface. These include: Angle of view, camera height above the target, the ground coverage of the image, the

resulting target resolution, the vehicle speed if the camera is mobile, the imaging frame rate, the overlap between successive images, the exposure time, and the degree of image blurring. Formulas relating these variables are given at the end of these notes.

c. Lighting issues and solutions
    i. The sun angle is constantly changing during the day making shadows quite variable
    ii. Light levels and shadowing are also variable when the sky is partially cloudy
    iii. Shadows greatly complicate separating plants from background because they change the appearance of objects from one part of the image to another.
    iv. Flash lamps can help near the ground but, if used alone, they will be overpowered by the sun. If feasible, blocking the sun from the field of view is the best solution.

d. Metadata needs
    i. The goal of metadata is to make each image and target uniquely identifiable in terms of experimental unit, location, and time.
    ii. It is imperative that metadata systems be redundant so that there is ***ALWAYS*** more than one way to accomplish the goal for any given image. This is because one will always be dealing with anywhere from many dozens to even hundreds of thousands of images – mistakes ***WILL*** be made.
    iii. Some metadata strategies include
        1. Linking GPS locations and image time stamps
        2. Stitching together many images some of which contain markers with known GPS coordinates
        3. Placing decipherable visual codes within images
        4. Code information into file names

e. Image throughput issues
    i. Saving images in raw format preserves information that is highly useful in removing lens distortions and color errors but requires greater storage space and transmission bandwidths.
    ii. As light levels fall, exposure times must increase which reduces image throughput
    iii. Between the sensor and their ultimate residence on a computer hard drive, there are several data stores and transmission bottlenecks through which an image must move. The most rate limiting step is often the USB connection between SIM card storage and the computer.
    iv. It might be that using video cameras may solve this problem but there are metadata and downstream computation issues to which we don't have answers at this time.

f. Camera body and lens selection
    i. Camera body - Pro/semi-pro camera bodies are desirable for

fast shutter speeds and storage on moving platforms. DLRS camera bodies are desirable because they have relatively large sensors. Full frame cameras might be too expensive.

    ii. Lens – Wider angle lenses have larger fields of view requiring fewer images but at the cost of greater amounts of lens distortion to remove. Zoom lenses are unacceptable because a constant focal length is desired during image taking. For the same reason, auto-focusing should be absent or turned off.

  g. While one will naturally want images that look vertically down, stitching and 3D extraction is facilitated if some images are at cross angles. Examples show a field setup that has worked very well, a chamber setup without cross angled images that was analyzed but with difficulty, and a drone. The gimbal on the latter keeps the angle down but lags between its adjustments and changes in plane orientation provide the needed cross-angles.

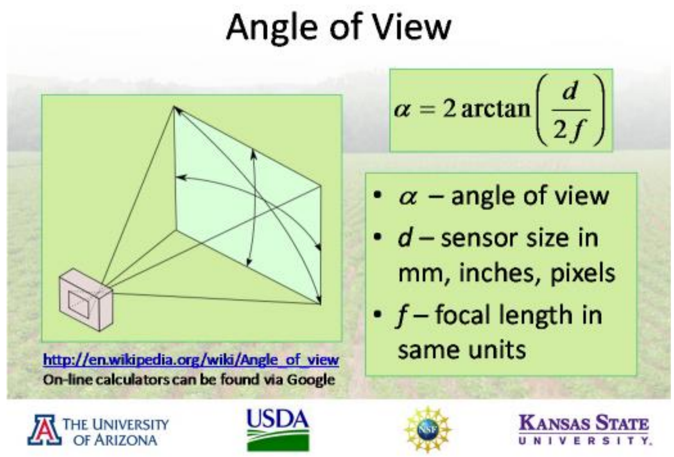
4. Image processing
  a. Geometric processing enables both orthophoto stitching and 3D model generation. An "orthophoto" is a computed image where the line of sight for each pixel is straight down.
  b. Some basic terms used in geometric processing include
    i. Principle point: The intersection of the optical axis and the image plane. It will be close to but generally not exactly at the center of an image. Focal point: The location internal to the camera through which all light rays from the target pass. Image plane: The surface on which the sensed image actually forms.
  c. Quantitative problems to solve:
    i. Lens distortion comprises nonlinear deviations of the image from the shapes that would be formed if all light rays propagated in perfectly straight lines through an ideal focal point. It must be removed but is described by complicated algebraic equations.
    ii. Once done, one needs to calculate:
      1. Interior camera elements – the focal length and the image plane coordinates of the principle point
      2. Exterior camera elements – the (X,Y,Z) coordinates of the principle point and the roll, pitch, and yaw angles of the optical axis
      3. Common points – image plane coordinates at which single points of the target appear. One needs at least hundreds of these per image.
      4. Segmentation – separating target and background pixels.
    iii. Packaged software will typically do all of these for you but not necessarily with equal facility
  d. Lens distortion removal – We have found that the best way to do this

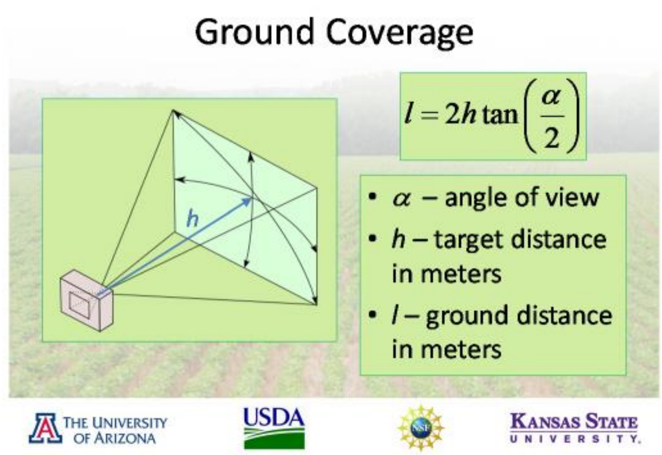
is to use camera vendor software that includes pre-calibrated/calculated databases of corrections that are specific to particular camera/lens combinations. An example is the Canon *Digital Photo Professional* (DPP) freeware. However, the use of such software requires using raw formats so that the camera body- and lens-type/settings metadata are stored in the image.

e. Color correction – Some cameras allow the user to indicate the light environment (sunny, cloudy, etc.) to aid in color correction. Raw format images aid color balancing because the camera is able to store additional useful information in the picture metadata. This metadata is usable by programs like DPP. One can also aid color correction by photographing standard color charts.
f. Common point detection – This is done using an algorithm called "SIFT", Scale Invariant Feature Transform. SIFT works best when there are lots of small irregularities in the images – it does not easily find points in single-leaf close-ups. This is not a problem in chamber situations or field studies using drones. It might be an issue for ground-based field platforms imaging closed canopies.
g. Canopy *vs.* background
    i. Image segmentation is the process of removing unwanted portions of the image, typically the background. Often this is done via some color-based approach. Almost invariably some very small objects will remain (usually in soil) that pass the color filtering criteria but are not part of the desired target. These can be removed based on their size – a process called "despeckling".
    ii. For very small numbers of images extant software will let you perform these operations manually. This can help you find the series of steps that yield the best results. An example is "Fiji", Fiji Is Just ImageJ", an expanded version of a freeware program developed by NIH. At some point you will want to automate these steps which will involve programming plugins in either Java or Python. A set of tools that can help with that is OpenCV.
h. Pixel (raster) *vs.* vector processing – Virtually all image processing is done using algorithms that utilize pixel data because that is the form delivered by cameras. In principle, however, some features such as the detection of shapes might be more easily handled by treating edges as connected vectors. Unfortunately, converting raster (pixel) images to vector form is difficult as several examples show.

5. Image pipeline design – Important steps in designing a system are: (a) identifying the various image acquisition and processing steps and (b) the volumes of data involved and the associated throughput rates. The latter point is essential if the creation of system bottlenecks is to be avoided. To aid in this, flowcharting symbols are offered, each annotated by key variables. Several examples with results are given, ranging from 22 images to over 10^5.

Formulas


Angle of View
$\alpha = 2\arctan\left(\frac{d}{2f}\right)$
• α – angle of view
• d – sensor size in mm, inches, pixels
• f – focal length in same units
http://en.wikipedia.org/wiki/Angle_of_view
On-line calculators can be found via Google
THE UNIVERSITY OF ARIZONA
USDA
KANSAS STATE UNIVERSITY


Ground Coverage
$l = 2h\tan\left(\frac{\alpha}{2}\right)$
h
• α – angle of view
• h – target distance in meters
• l – ground distance in meters
THE UNIVERSITY OF ARIZONA
USDA
KANSAS STATE UNIVERSITY

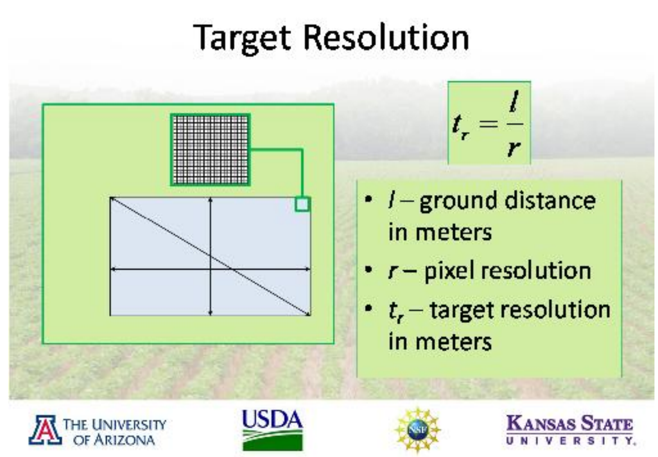
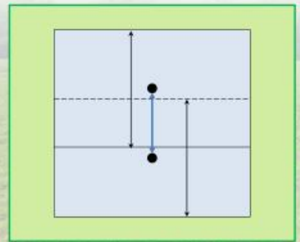

Target Resolution
$t_r = \frac{l}{r}$
• l – ground distance in meters
• r – pixel resolution
• t_r – target resolution in meters
THE UNIVERSITY OF ARIZONA
USDA
KANSAS STATE UNIVERSITY

Image overlap

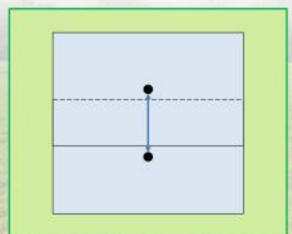


$$O = \max(l - p, 0)$$

- l – ground distance
- p – image principle point spacing
- O – image overlap

All units are meters

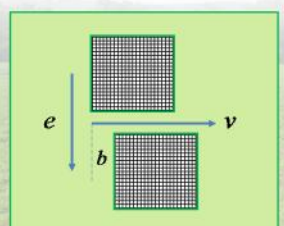
Principle point spacing – vehicles



$$p = \frac{v}{f_r}$$

- v – speed (m/sec)
- f_r – frame rate in images/sec
- p – image principle point spacing (m)

Motion blur – vehicles



$$b = \frac{t_r}{ve}$$

- v – speed (m/sec)
- e – exposure (sec)
- t_r – target resolution in meters
- b – blur in pixels