Aerial imagery

1. Payloads
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. LIDAR

11A: Introduction to data products from aerial imagery

Lecture Presented By: Kevin Kochersberger

In this lecture you will learn:

- Typical imaging payloads
- Data products

Imaging systems are the most frequently used payloads on drones

Aerial imagery
1. **Payloads**
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. LIDAR

- Imaging comes in many flavors
  - Passive systems:
    - Visual spectrum (electro-optical) camera
    - Multispectral
    - Hyperspectral
  - Active systems
    - LIDAR
      - Line scanning
      - Imaging
    - Structured light

Parrot Sequoia
Multispectral sensor

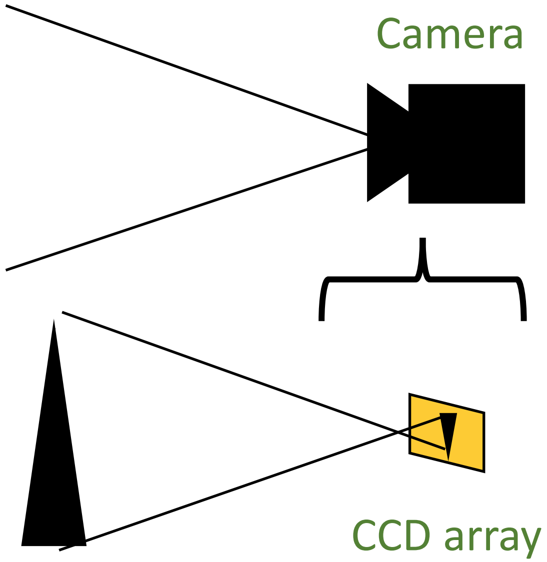
Bayspec hyperspectral imager

Electro-Optical (EO) Cameras

Aerial imagery

1. **Payloads**
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. LIDAR

- Electro-optical cameras work with a CCD (charge coupled device) or CMOS (complementary metal oxide sensor) sensor
  - These sensors consist of an array of semiconductor “pixels” to provide a full frame description of the visual field of view
  - The size of the pixels and the size of the array determines the size of the image, typically expressed in megapixels
  - CCD or CMOS cameras can be either global shutter or rolling shutter
  - CCD cameras tend to be cheaper but CMOS sensors are much more energy efficient



EO camera global vs. rolling shutter design

- Rolling shutter cameras work with a progressive line scan exposure which can cause distortion of moving objects
- Rolling shutter cameras are typically cheaper, with most CMOS-based cameras having a rolling shutter
- The rolling shutter design is usually acceptable for video capture but it is not acceptable if any frame-by-frame analysis is required (such as in 3D reconstruction or anomaly detection)

The rotor of the Yamaha RMAX is not symmetrical in this image due to the rolling shutter

EO camera global vs. rolling shutter design

Aerial imagery

1. **Payloads**
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. LIDAR

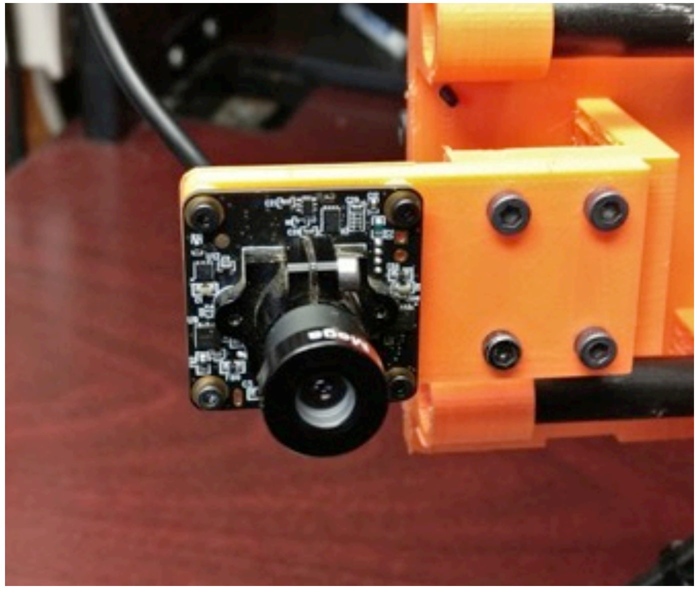
- Global shutter cameras are required in any photogrammetry work or 3D terrain reconstruction when camera motion or motion on the ground can impact image quality
- Global shutter cameras are more expensive than rolling shutter cameras
- If all that is needed is video, or the scene is not dynamic then a rolling shutter camera will work

In this global shutter image of the RMAX, the rotor is symmetrical and shows no asymmetry

Camera installations are frequently at "board level" to save weight

Aerial imagery
1. **Payloads**
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. LIDAR

- In aircraft installations, cameras and other hardware may be considered "original equipment manufacturer" (OEM) equipment
  - These payloads do not have the packaging necessary for stand-alone operation
  - They are intended to be integrated into other systems
  - In the case of drones, this saves a great deal of weight and allows for a customizable design



A camera mounted on PC board that is integrated into a custom drone application

Larger aircraft can accommodate larger, stabilized camera payloads

Aerial imagery
1. **Payloads**
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. LIDAR

- These are more expensive since they have very accurate pointing capability
- They can typically hold a stabilized view from several hundred meters distance
- The gimbal can be controlled by a joystick to change views

VT RMAX carrying a stabilized camera payload on nose

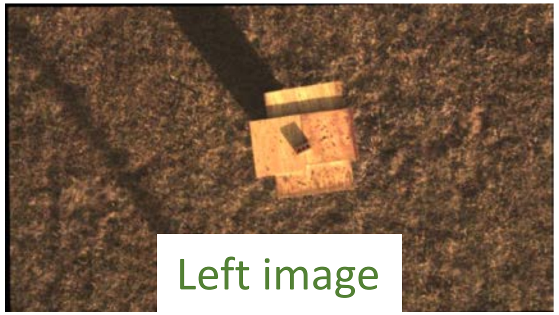
Stereovision systems allow for a rapid 3D reconstruction of the scene

Aerial imagery
1. **Payloads**
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. LIDAR

- A stereovision system matches corresponding features in the left and right cameras, and defines a disparity *d* between the two features in the images
  - This disparity corresponds to a *Z* distance from the cameras to the feature:

$$Z = \frac{fB}{x_1 - x_2} = \frac{fB}{d}$$

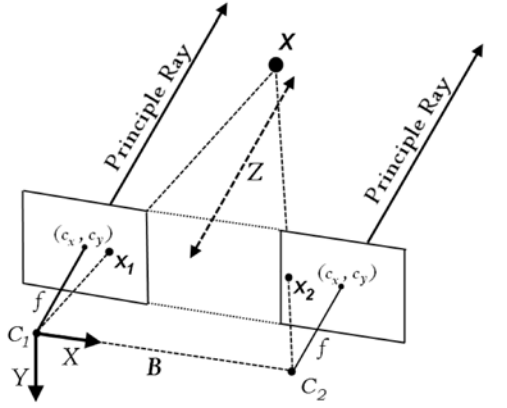
  - The collection of features with a disparity *d* can be represented in a disparity map



Left image

Right image

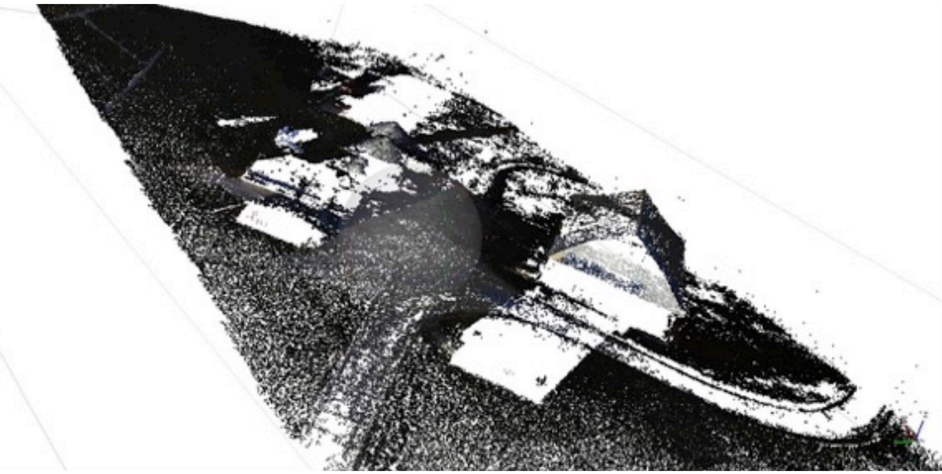
Disparity map



A set of image frames over a scene can be used to reconstruct the 3D environment

Aerial imagery
1. Payloads
2. **3D reconstruction**
3. Multispectral and hyperspectral imagery
4. LIDAR

- Building a 3D terrain model using ”structure from motion” is a process that minimizes error of a set of 3D point locations that are described in multiple image frames
- The computational time to create the 3D terrain will increase with number of image frames, and the processing can take a very long time for large datasets (hundreds of images)
- The first step is to achieve correspondence between image frames

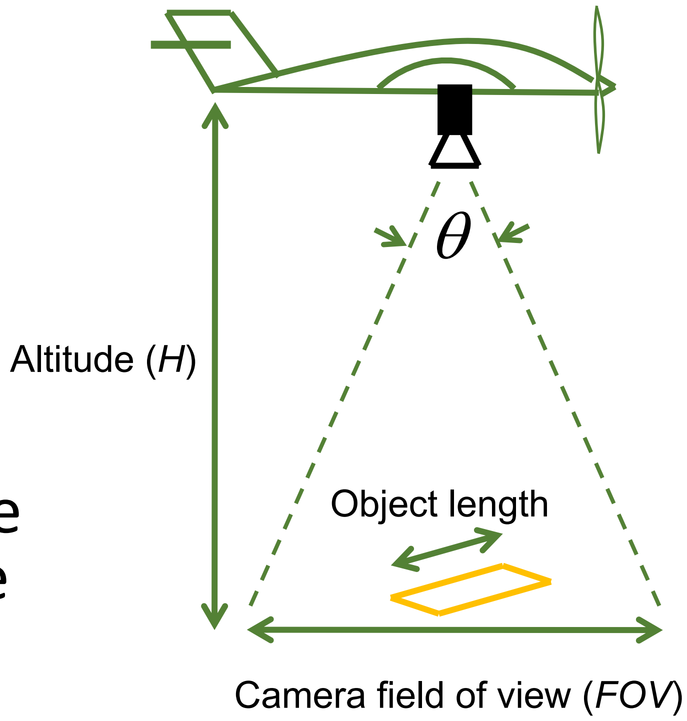


A 3D point cloud generated from a series of images taken from a single drone camera

Actually, the first step in building a 3D reconstruction is collecting the images

Aerial imagery
1. Payloads
2. **3D reconstruction**
3. Multispectral and hyperspectral imagery
4. LIDAR

- It is necessary to achieve a certain % overlap in all directions so that there is correspondence between neighboring images
- The field of view (FOV) of the camera is computed from the imaging angle and height of the aircraft
- For 75% overlap, can you write the relationship between speed of the aircraft and the frame rate of the camera?



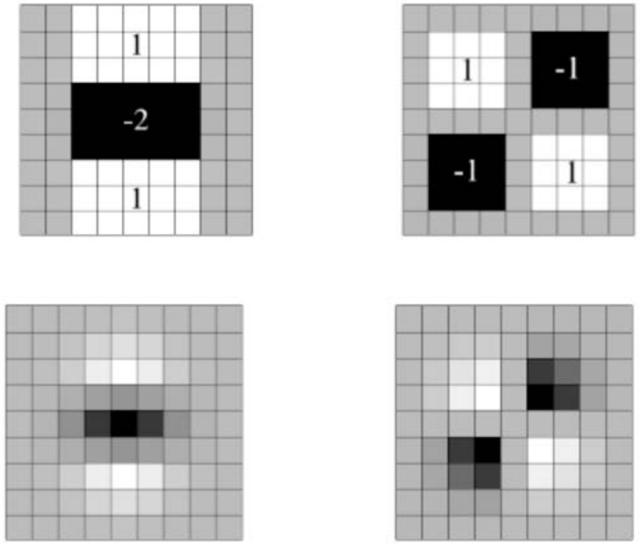
$$\tan\left(\frac{\theta}{2}\right) = \frac{FOV}{2H}$$

Note that a camera imaging angle of 53º corresponds to the FOV = H (a 1:1 relationship)

Aligning image frames

- In order to build a terrain model, registration of points between successive images is required
- There are many ways to do this, but it is common to define **unique features** in the image pixel-space for frame-to-frame matching
  - **SURF** features: Speeded Up Robust Features use a Hessian matrix-based detector and distribution-based descriptor
    - A Gaussian derivative of a binary kernel is used to detect unique features
  - Approximations of point correspondence reduces the space for feature matching

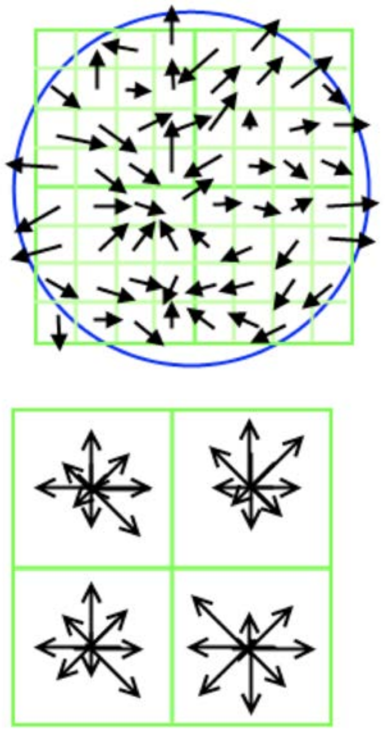
Aerial imagery
1. Payloads
2. **3D reconstruction**
3. Multispectral and hyperspectral imagery
4. LIDAR



The Gaussian kernel is passed over the image to identify unique features in frame-to-frame matching

Aligning image frames

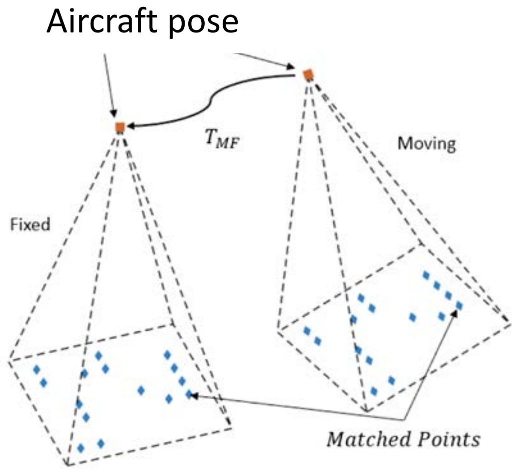
- Another feature commonly used for alignment is the **SIFT** feature – scale invariant feature transform
- SIFT features define an orientation vector from the 2x2 Hessian kernel
  - The vector is based on the max and min response orientation of the feature
  - The descriptor consists a 4x4 array with eight orientations, resulting in a 128-element descriptor, unique to scale



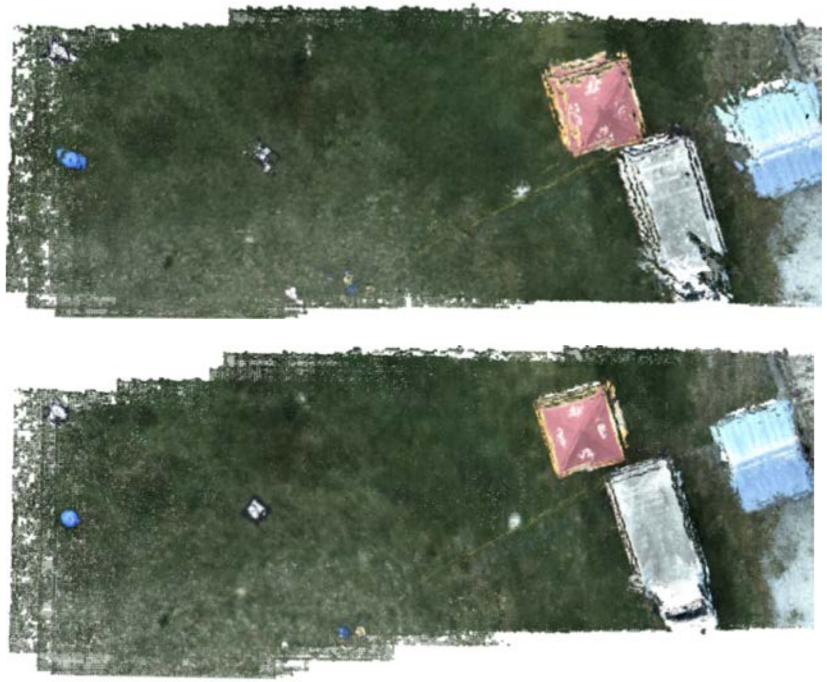
A vector description is scale invariant

Aligning image frames

- Frame-to-frame correspondence is approximately accomplished by using the IMU data from the aircraft
  - The pose of the aircraft is reasonably well defined by the aircraft GPS, accelerometers and gyros
- Final matching of frames is accomplished by feature point detection and matching



Aerial imagery
1. Payloads
2. **3D reconstruction**
3. Multispectral and hyperspectral imagery
4. LIDAR

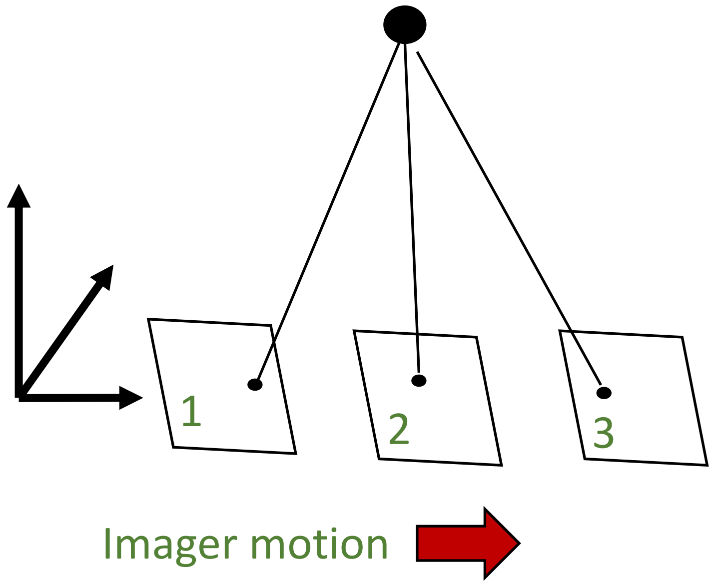


The top image shows a sequence of image frames w/o matching. The bottom image uses feature matching

The 3D terrain is reconstructed by minimizing re-projected error

Aerial imagery
1. Payloads
2. **3D reconstruction**
3. Multispectral and hyperspectral imagery
4. LIDAR

- Minimization of the error in point correspondence using camera pose is called bundle adjustment
- The least error reprojection of an observed point from the image plane over all images requires a non-linear least squares minimization
    - The Levenberg-Marquardt algorithm is frequently used to accomplish this non-linear least squares minimization

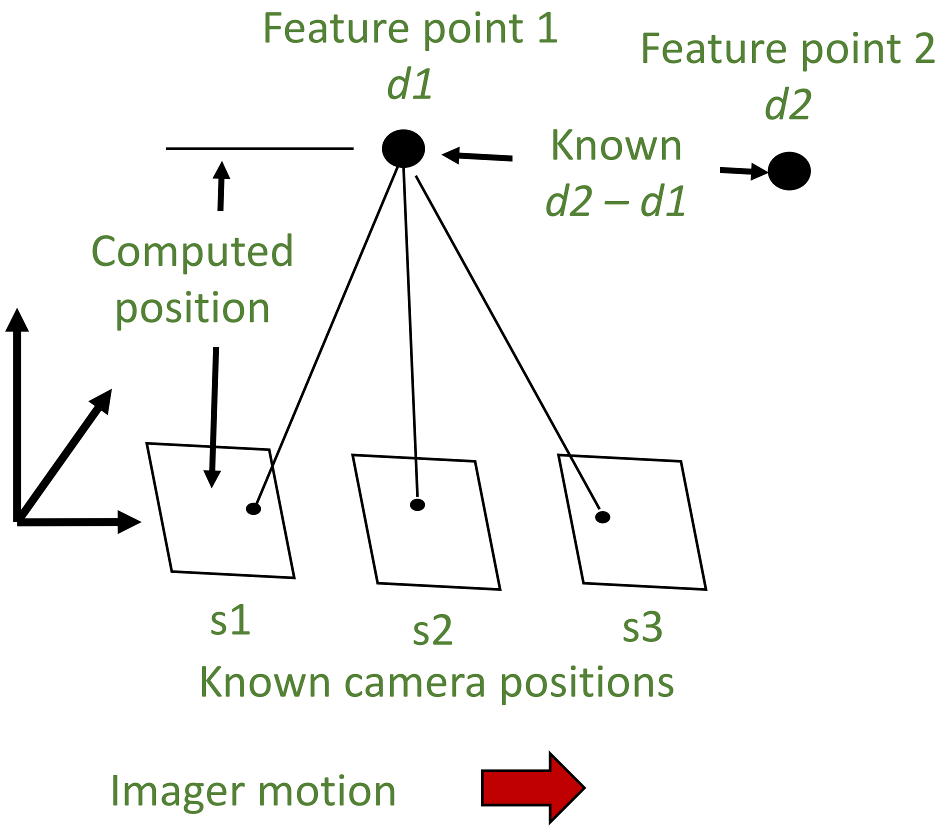


As the camera moves, the image plane point moves as a function of the camera pose relative to the point

The resulting 3D reconstruction will have no scale information unless measurements are available

Aerial imagery
1. Payloads
2. **3D reconstruction**
3. Multispectral and hyperspectral imagery
4. LIDAR

- If the **position of the camera is known** via GPS, then the observed relative motion of the terrain will be scaled properly
  - That will allow an accurate distance from the camera to be computed which results in scale feature point position
- If the **distance between features** in the scene (usually three point positions should be referenced) is known, then the entire scene can be properly scaled



The feature points used in reconstruction are a basis for full scene reconstruction

Aerial imagery
1. Payloads
2. **3D reconstruction**
3. Multispectral and hyperspectral imagery
4. LIDAR

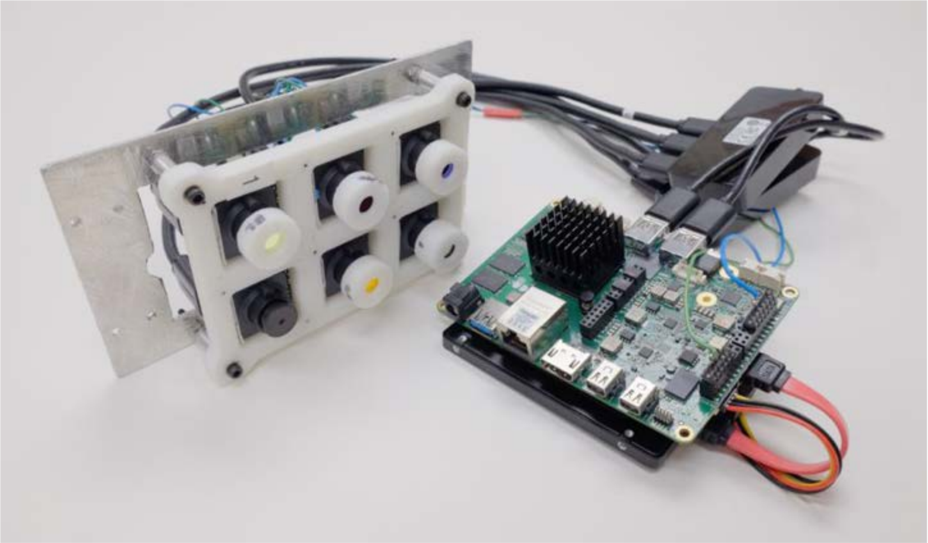
- Feature points used in 3D reconstruction represent good features for matching
  - They do not represent all of the pixels in the scene
- Dense point cloud reconstruction is completed using methods that fill in the remaining 3D description of the environment
  - The result is a 3D position for every pixel that can be related to the feature points
- In some cases, missing feature points caused by poor lighting or poor texture will result in holes in the 3D reconstruction

This scene has missing points due to poor texture in the roof of the building on the left

Multispectral and hyperspectral data products are commonly used to find important environmental characteristics

Aerial imagery
1. Payloads
2. 3D reconstruction
3. **Multispectral and hyperspectral imagery**
4. LIDAR

- A **multispectral camera** is capable of filtering the image to show only certain EM bands of interest
  - These bands provide detailed information about crop health (biomass), moisture (wetlands), and classification for plant identification, etc
- A common index for biomass detection is the **Normalized Difference Vegetative Index (NDVI)** which is based on both near infrared and visual spectral bands



Multispectral camera developed at the USL

$$NDVI = \frac{(R_{NIR} - R_{Red})}{(R_{NIR} + R_{Red})}$$

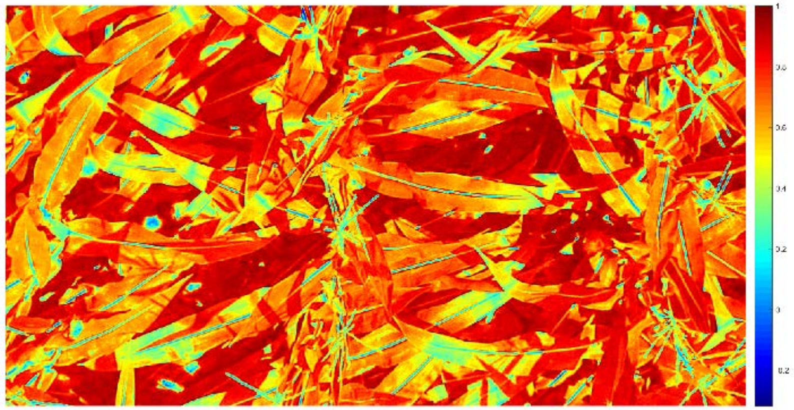
R_{NIR} = NIR reflectance,
R_{Red} = Red reflectance

Multispectral imagery results have been correlated to crop yield data

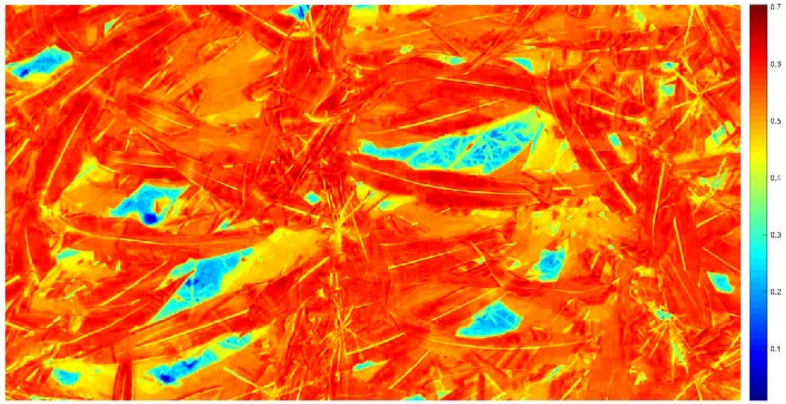
Aerial imagery
1. Payloads
2. 3D reconstruction
3. **Multispectral and hyperspectral imagery**
4. LIDAR

- Image processing should consider subtraction of the non-crop areas to get an accurate assessment of the crop index

Raw image of corn



NDVI of raw image



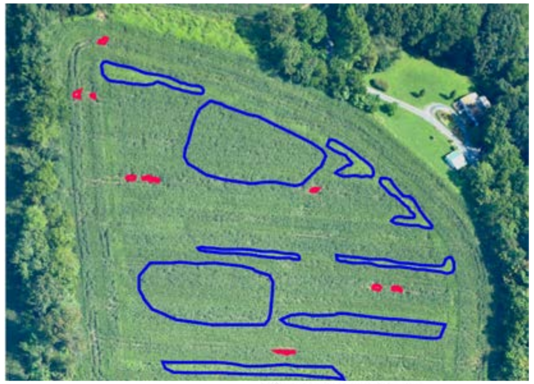
NDVI of image with background taken out

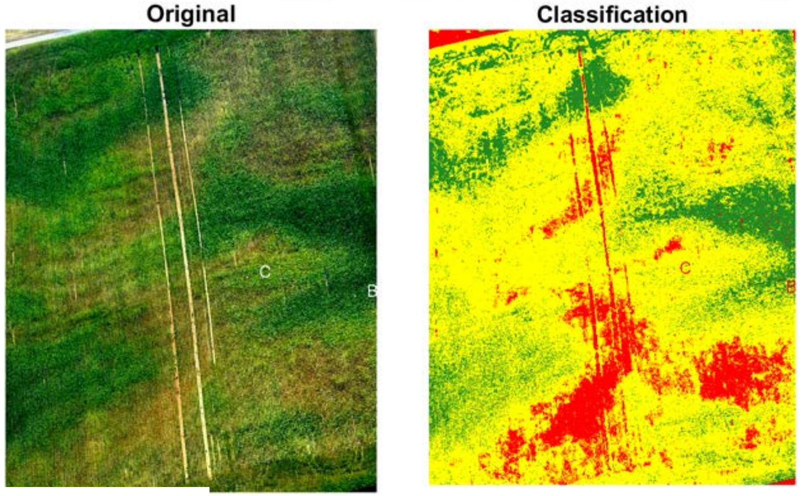
It is possible to use visual spectrum data only to determine crop health or identify other agricultural features

Aerial imagery
1. Payloads
2. 3D reconstruction
3. **Multispectral and hyperspectral imagery**
4. LIDAR

- **Training using annotated ground truth data** creates a successful classifier for crop condition monitoring
- **RGB, HSV,** or **L-a-b** color spaces can be evaluated for best results
- Extracted features from the scene, like geometric shapes that are identified using the Hough transform (for example), can also be investigated in the visual spectrum

Areas that are ground-truthed for a particular condition





Hyperspectral imagery

- **Hyperspectral imagery** uses narrow-band data to extract characteristics in the data that would normally not be identifiable
- Hyperspectral imagery from a typical sensor consists of a range of wavelengths from 600 nm – 1000 nm (Near IR), with 20 nm resolution
- Many hyperspectral imagers are "pushbroom" sensors, meaning that they function as line scanners and the camera must move to create a 3D image

OCI-UAV-2000 hyperspectral camera (not a line scanner but a 256 x 256 imager)

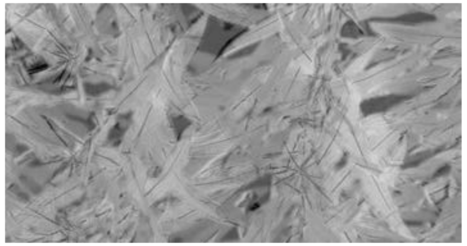
Top: RGB image
Bottom: 850 nm image

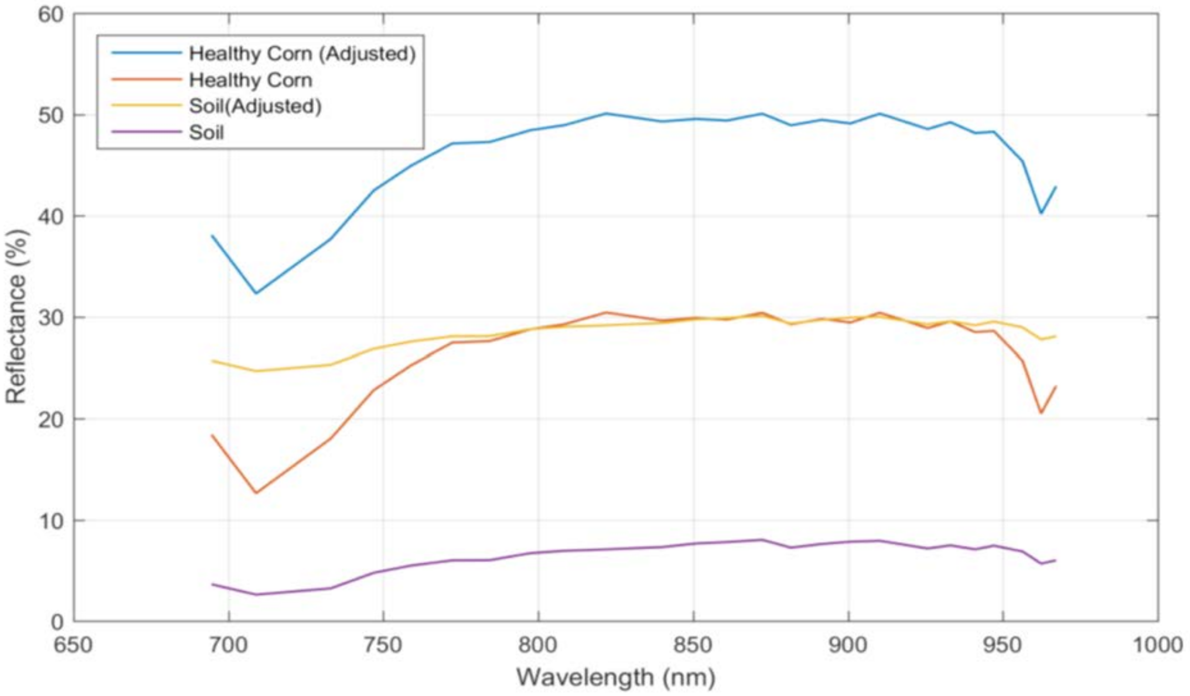
Hyperspectral imagery

- **Creating indices of reflectance at certain bands** can reveal particular vegetative conditions that may allow targeted treatments or responses to disease

$$\text{Normalized difference Red Edge} = \frac{R_{790} - R_{720}}{R_{790} + R_{720}}$$

$$\text{Normalized difference Red Edge} = \frac{48 - 33}{48 + 33} = 0.185$$



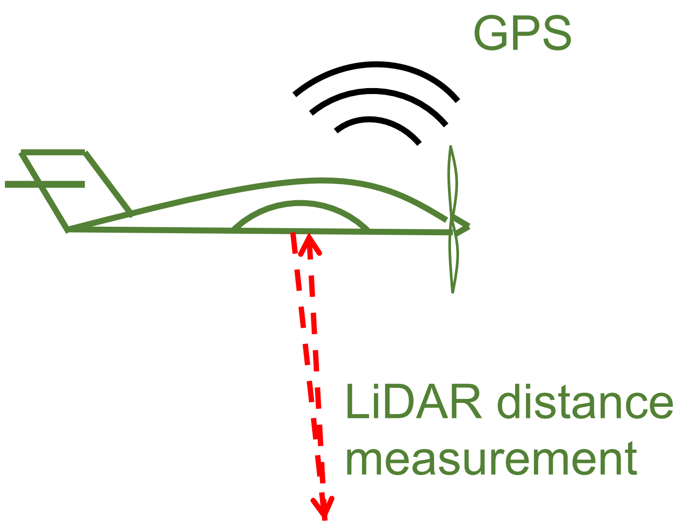


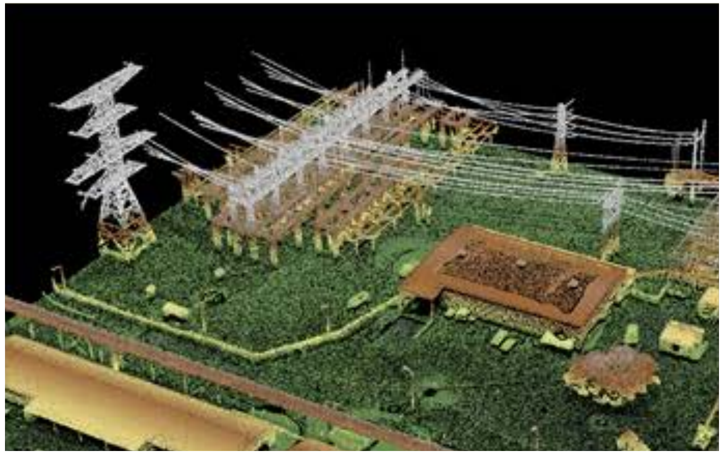
Aerial imagery
1. Payloads
2. 3D reconstruction
3. **Multispectral and hyperspectral imagery**
4. LIDAR

LIDAR = LIght Detection And Ranging

Aerial imagery
1. Payloads
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. **LIDAR**

- A LIDAR system is an active sensor that sends a pulse of infrared energy to the ground, and it is reflected back up to the sensor
- The time of flight of the beam is used to estimate the distance from the sensor to the object that has been interrogated
- Accuracy and reliability tends to be higher than with vision – based systems, but the resolution is usually lower



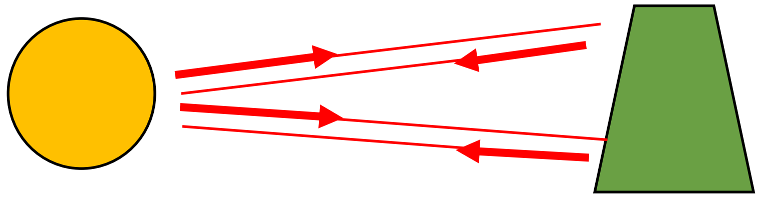


Point cloud image from LIDAR

LIDAR systems come in a variety of form factors which accommodate aircraft of many sizes

Aerial imagery
1. Payloads
2. 3D reconstruction
3. Multispectral and hyperspectral imagery
4. **LIDAR**

- Most line scanners emit a rotating IR laser beam with a timed detection system to measure time of flight
- The resolution is a function of the pulse time and the speed of rotation of a mirror that directs the beam in a 360 degree scan



A Velodyne VLP-16 Puck LITE 16 channel scanner with 100m range – suitable for drones