RECTIFICATION, GEOREFERENCING, AND MOSAICKING OF IMAGES ACQUIRED WITH REMOTELY OPERATED AERIAL PLATFORMS

Gabriel B. Ladd, Graduate Student (Marine Estuarine and Environmental Sciences)
Abhijit Nagchaudhuri, Professor (Engineering and Aviation Sciences)
Tracy J. Earl, GIS Program Manager
Madhumi Mitra, Assistant Professor (Biology and Environmental Sciences)
University of Maryland Eastern Shore
Princess Anne, MD 21853

Geoffrey L. Bland, Scientist
NASA Wallops Flight Facility
Wallops Island, VA 23337

ABSTRACT
Remote sensing efforts are being employed as part of the recently initiated “Environmentally Conscious Precision Agriculture” project at University of Maryland Eastern Shore. Autonomous and remote controlled aerial platforms with nadir view cameras are being utilized in conjunction with precision farming equipment. Initial efforts have involved development of project specific tools utilizing commercial software environments for rectification, georeferencing, and mosaicking of image frames acquired using small remotely operated airborne platforms. This paper will delineate image acquisition, rectification, georeferencing, and mosaicking of image frames to generate a composite image of the entire field. Yield data collected by a yield monitoring system is compared with information obtained from the images.

INTRODUCTION
The University of Maryland Eastern Shore (UMES) Environmentally Conscious Precision Agriculture (ECPA) project is developing infrastructure on campus to provide a platform for education and research in this emerging field (Nagchaudhuri, Mitra et al., 2005). This paper focuses on the utilization of photogrammetric techniques for data mining of imagery captured by small remotely operated aerial platforms. The paper begins with a summary of photogrammetric techniques and follows with an account of how they were used as part of the project. The paper concludes with an evaluation of the usefulness of the techniques and the future plans.

BACKGROUND
The agronomic revolution of the last seventy five years is an evolution towards detailed management of intra field variations. This trend began back in the late 1920’s and has continued to this day (Sudduth, 1999). This revolution has developed into what is today known as precision agriculture or site specific management (Pfister, 1998).

Precision agriculture is a combination of technologies that work together to give the farmer a detailed picture of the variations that exist within a given agricultural field. The three most common technologies encountered are the mechanized combined harvester or combine, a global positioning system (GPS) receiver, and a yield monitoring system (Pfister, 1998). The combine is a mechanical harvester that both harvests and shells or shucks plants. A GPS receiver is an electronic device that uses signals broadcast from a constellation of satellites to determine its location (Steede-Terry, 2000). A yield monitor is a system of sensors installed within a combine that measure the grain weight, grain flow rate, travel speed and harvest area (Batchelor, 1997; Colvin, 1999; Shearer, Fulton et al., 1999; Grisso, Alley et al., 2003; Technology, 2003). The yield monitor system also tracks and stores the global locations calculated by the GPS receiver (Shearer, Fulton et al., 1999; Grisso, Alley et al., 2003). The data generated by the combine, GPS receiver, and yield monitor are then brought into a Geographic Information System (GIS) to produce a yield map. GIS is a spatially distributed database that uses basic shapes and images to visualize data and create
The purpose of a yield map is to detail how the crop yield varied throughout the field (Batchelor, 1997; Colvin, 1999; Doraiswamy, Moulin et al., 2003; Grisso, Alley et al., 2003).

Knowing how the yield varies throughout the field allows the farmer to implement management policies that maximize yield while minimizing input. Many of these management policies require grid mapping/sampling of other factors present in the fields. Grid mapping/sampling is the process of sampling factors within a field in such a way that a map can be extrapolated based on a select number of sample locations. It is so called because most often the field is divided into a grid pattern and one sample is taken from each grid (Grisso, Alley et al., 2003; Mylavarapu and Lee, 2006). The grid sampled data can then be mapped and used in conjunction with the yield data to determine the most efficient inputs of fertilizers, pesticides, lime, etc. for the particular field. The use of highly detailed and nearly continuously varying input data into management models produces smoothly varying output recommendations. To meet these models Variable Rate Technology (VRT) is used (Clark and McGuckin, 1996). Computer controlled sprayers, spreaders and seeders etc. have been developed to allow farmers to precisely control and vary the amount being applied on the field at any given time.

One of the more recent technologies to be employed in the agricultural fields is remote sensing from small Uninhabited Aerial Vehicles (UAVs). Imaging from an aerial platform is not new. The earliest known aerial photograph dates back to 1850 and was taken by a Frenchman from a hot air balloon. The advent of aircraft provided the camera with yet another platform (Lillesand and Kiefer, 2000). The space age ushered in the era of satellite photography. Traditionally, there has been a large cost and resolution discrepancy between satellite and manned aircraft photography (Hatch, Brooks et al., 1999). The solution to this problem is the use of small UAVs. UAVs tend to be more cost effective and are small enough to be flown at the low altitudes to generate high resolutions. Currently, systems are under development for applications ranging from pest mapping to Normalized Difference Vegetation Index (NDVI) (Hunt, Everitt et al., 2003). The UAVs used for agriculture work tend to be either small rugged fixed wing aircraft or small computer controlled/stabilized helicopters (Sullivan, 2005). Use of small lightweight aircraft produces a unique set of photogrammetry problems.

**UAV AERIAL IMAGE PROCESSING BASICS**

Photogrammetry is defined as: “The art or process of surveying or measuring, as in map making, by taking photographs, especially aerial photographs” (Webster's, 1999). The photogrammetric techniques discussed in this paper are: scale, resolution, coverage, rectification, georeferencing, and mosaicking.

Scale is the ratio of a distance found on a representation, to the actual distance between the two objects. For example scale of an aerial image looking straight down at an airport can easily be measured by calculating the ratio of the runway length in the photograph divided by actual runway length. Scale is represented as an equivalence, fraction, or ratio. In all cases the reference distance (map, photo, etc) is listed first or as the numerator. Then the actual distance is listed as the right side of the ratio or as the denominator (Wolf and Dewitt, 2000).

When imaging from an aircraft, scale will change with distance from an object. The higher the camera flies the smaller the runway will appear and thus the smaller the numerator will become. Scale can be calculated as a function of distance above the ground (H) and camera focal length f (see Figure 1 for more details). Scale is the ratio of DE to BC. Using similar triangles it quickly becomes clear the scale is also the ratio f/H.

![Figure 1A Side view](image1.png)  ![Figure 1B 3D View](image2.png)

**Figure 1.** Side and 3D view of an aerial image taken over flat ground.
Resolution is a measure of how much area is represented by one reactive agent in an imaging device. For film that reactive agent is the grain. In digital imagers it is one CCD cell which translates into one pixel of the stored image. In Figure 1 as H increases the distance between B and C increases but the distance between D and E does not change. Effectively the camera is able to see more area and is trying to store it to the same amount of image area. The larger H becomes the more area on the ground is represented with a single grain or pixel. This means as a camera is flown higher, the scale increases and the resolution decreases. The image resolution is calculated by dividing the image area by the ground area represented by the image or vice-versa. For digital images this is usually expressed in area/pixel.

The ground area represented by the image is referred to as the coverage area. The coverage area is an important consideration when setting up an aerial imaging program. The three critical factors are: knowing what kind of terrain, knowing how much of it will be mapped, and knowing what resolution is needed. These three factors dictate the choice of both aerial platform, and imaging device.

Image rectification is the process of flattening tilted images so that their normal vector is parallel to the normal vector of the ground plane they represent (Milkhail, Bethel et al., 2001). This technique can be particularly difficult when working with slow moving and light weight aerial platforms. Their low speed and mass make them much more susceptible to the effects of wind and other atmospheric perturbations (Anderson, 1999; Raymer, 1999; Corke, 2003).

Wind gusts and thermals among other things can cause an aircraft to move unexpectedly and even erratically. The motion of an aircraft is divided into three distinct types: pitch, roll and yaw. Pitch is the motion of nose and tail of the aircraft about the center of gravity in the vertical direction. Roll is the rotation of the aircraft about an axis that runs from the nose through the tail of the aircraft. Yaw is the rotation of the body of an aircraft in the horizontal plane (Etkin and Reid, 1996; Anderson, 2001). The pitch, roll and yaw of an aircraft directly affect how the camera is tilted in relation to the ground below.

Image tilt must be understood to produce a rectified image. To this effect a convention of angles is used (Wolf and Dewitt, 2000). In computational applications the angles of rotation of the image are used. These angles are defined about a fixed coordinate system orthogonal to the flat ground below and the image coordinate system that has its origin at the sensor medium (see Figure 2). The angles use rotation about the image orthogonal axis $x'$, $y'$, $z'$ which are denoted by $\phi$, $\omega$, $\kappa$ respectively. These rotations are modeled as being sequentially added to the orthogonal axes of the image in $\phi$, $\omega$, $\kappa$ order. This will produce the tilted image as illustrated in Figure 2 (Wolf and Dewitt, 2000).

![Figure 2. Image coordinate system and image true projection to the ground.](image)

The amount of rotation about each axis must be determined before the image can be rectified. Rotation is determined by the comparison of points within the image. The image locations and the spatial locations in the real world are known. The difference in their locations relative to each other in the image is compared to the corresponding difference in relative locations in the real world. This difference can be attributed to the rotation of the image coordinate system $x$, $y$, $z$ with respect to the orthogonal coordinate system $x'$, $y'$, $z'$.
The points of known location in both the image and the real world are referred to as ground control points (GCPs). GCPs can be anything from geographic features to a marker flag. The only requirement is that their position be known (Simonett, 1983). This usually entails either finding an accurate map with their location or mapping them. The precision and especially, the accuracy of their known positions is extremely important. The control point precision will limit the precision of the rectified image. The uncertainty will propagate through the calculations with the potential of making the rectified image useless.

Rectification is the process of subtracting the angles of rotation from the image to leave the image aligned with the image orthogonal axis. The simplest method of rectification is a two dimensional projective transform. Essentially the projective transform creates linear functions to map the picture coordinates into ground coordinates. In Equation 1 below, \( X_w \) and \( Y_w \) are world coordinates while \( x \) and \( y \) are the coordinates for the same location in the image, \( c, d, \) and \( e \) are constants (Wolf and Dewitt, 2000).

\[
\begin{align*}
X_w &= \frac{d_1 x + e_1 y + c_1}{d_3 x + e_3 y + 1} \\
Y_w &= \frac{d_2 x + e_2 y + c_2}{d_3 x + e_3 y + 1}
\end{align*}
\]

All computer based image rectification processes resample the images. This means they work with the individual pixels to move and recombine them into the rectified image (Wolf and Dewitt, 2000). Depending on the degree of tilt, the image resolution, and scale the rectification and re-sampling process can lead to severe image distortion. This distortion is measured by comparing the location of a point in XYZ coordinate frame to where the rectified image coordinate frame says the point is. The magnitude of the error in the rectified coordinate frame is called a residual and indicates how well the rectification approximated reality (Wolf and Dewitt, 2000). The smaller the \( , , \) angles the less correction will be required and the smaller the residuals.

To georeference is to align something to an earth centered coordinate system. With aerial images this is the alignment of a rectified image to an earth based coordinate system, for example latitude and longitude. The alignment process is a simple rotation and translation of the rectified image based upon GCPs visible in the image. Many programs actually incorporate rectification and georeferencing into one command. This allows the user to both rectify and georeference an image based upon one set of GCPs.

One popular way to georeference an image is to use an orthographic base image. An orthographic base image is a pre-existing image that has already been rectified and georeferenced to a known precision. These are available through state and federal sources.

The idea behind mosaic images is that a larger picture can be made from a series of small pictures. This is done by aligning two images based upon shared control points (CPs). These CPs are not to be confused with GCPs. CPs are inter picture control points with no known correlation to an outside coordinate system. CP can be GCPs but this is not required to combine two images. To align and overlay the two images, all that is needed are three or more points that appear in both images. These points can then be aligned based upon the plane created by the three points. The best quality mosaics are usually generated from previously rectified images. These images are usually rectified using GCPs (Milkhail, Bethel et al., 2001).

There are several possible sequences of georeferencing or rectifying, and mosaicking multiple images together to generate a large georeferenced image. A basic choice is to start by mosaicking all the images together first. The resulting composite image would have the possibility of more control points. Another option would be to georeference one image and then mosaic others to that one image. Yet another option is to georeference each individual image and then mosaic all the images together based upon their coordinates. This is the ideal solution because it is easiest to track errors and it allows for very high resolution, precision, and accuracy throughout the mosaic.

**APPLICATIONS OF PHOTOGRAMMETRY IN UMES AIRSPACES PROJECT**

The AIRSPACES project utilized two types of fixed wing aircraft. An AeroBird Extreme R/C Aircraft (HobbieZone, 2005). The AeroBird has been modified by fitting an Aiptek MegaCam (Aiptek, 2006) to the aircraft’s belly. This aircraft is hand launched with the camera running in video recording mode. The camera records
250MBs of low resolution video footage. This is a recording time running from one to three minutes depending on what is being imaged. Post flight the camera is connected to a laptop and the footage is downloaded.

![AeroBird Extreme](image1.png)

**Figure 3.** AeroBird Extreme.

The second aircraft utilized by this project is a highly modified spectra electric sail plane (GreatPlanes, 2002). The aircraft is equipped with a nadir view Sony mini color CCD camera. The video feed from this camera is converted to a 2.4GHz video feed and transmitted down to receiver on the ground. The video composite feed is then recorded on MiniDV and VHS. Post flight the MiniDV tape is digitally imported into the computer via an IEEE 1394 fire wire connection (Gowan, 1999).

![Modified Spectra](image2.png)

**Figure 4.** Modified Spectra.

When working with small light weight aerial platforms the image processing begins with image selection. This means finding images with minimum pitch, roll, and yaw or small , , angles. Small , , angles are needed to minimize the image distortion created during the rectification portion of the georeferencing process. This is done by examining the digitized flight footage frame by frame. One instance of this time consuming process produced two hundred and fifty frames from a seventeen minute flight.

Figure 5 shows examples of frames captured from aerial platforms used in this project. The major problem with the image in Figure 5A is that there are no GCPs to georeference the image to our orthographic base images. Figure 5A is also lacking the CPs to mosaic it with another image for later georeferencing. Figure 5B is one of our better images, having enough GCPs to be georeferenced and showing barren spots in the image which should be identifiable in the yield map. Figure 5C is an example of the kind of unusable image frames that may be generated from small light weight aerial platforms used in the initial stage of the project. In Figure 5C the aircraft has been rolled to the left by a gust of wind so far that the horizon is visible.

Ideally, there will be one frame to cover each portion of the area of interest. Invariably, this is not the case with small light weight aerial platforms. Usually there are gaps between images and in corners of the area of interest. For complete coverage the solution is to use multiple images that significantly overlap.
There are several ways to assemble unreferenced aerial image into a map. The first option is to work with them on an individual basis. Rectifying and then georeferencing each image based upon a set of GCPs in each individual image. The images can then be displayed on a map based upon their locations. Figure 6 is a good example of one image before and after rectification and georeferencing. The distortion in Figure 6B becomes worse in left half of the image where there are no GCPs to use in the image rectification.

If there are insufficient GCPs in an image that image can be mosaicked with another or even multiple images to create a larger image that possesses enough GCPs to be rectified and georeferenced. In this project we used a MATLAB based code to mosaic images together (Serrano, 2005). The code utilized inter image control points. These control points are common points between the two images and not necessarily GCPs. Figure 7 is an example of an image created using this code. There are some inherent problems using this method. The images in the mosaic are rectified to one single image producing an error that is then compounded when the image is rectified to a ground based coordinate system.
In ArcGIS 9 an orthographic base image is loaded and then unreferenced images are loaded. Next step is to locate a GCP on one unreferenced image and then locate the same location on the orthographic image. As the GCPs are located liking the two images together the unreferenced image is overlaid and transformed with each successive GCP. The user can see instantly the effect of each GCP and also adjust the transformation equation on the fly. The on the fly processing produces very useful maps like the one seen in Figure 8.

![Figure 8](image)

**Figure 8.** A number of aerial images georeferenced to orthographic base image.

There are advantages to both MATLAB and ArcGIS for georeferencing images. ArcGIS has a fast and intuitive GUI (Graphical User Interface) based method. Unfortunately, it does not allow the level of customization and repeatability needed for this project. MATLAB based code has a very controlled and reproducible process that can be used to export the referenced images for later use. The fact that the process is based upon changeable computer code it will allow easy expansion for future applications.

**RESULTS**

Figure 9 illustrates the UMES agricultural fields. The yield data in this paper only covers the Bozman fields. These fields are used to farm corn, wheat, and soy beans in a standard crop rotation.

![Figure 9](image)

**Figure 9.** UMES agriculture fields.

ASPRS 2006 Annual Conference
Reno, Nevada • May 1-5, 2006
To create the corn harvest mosaic of the Bozman field, two hundred and fifty still images were selected from seventeen minutes of video footage. From those two hundred and fifty images twenty five images were selected for being closest to normal with ground, covering the area of interest, and having usable control points. Of those images fifteen were used to produce a full field mosaic of the corn immediately before harvest in the Bozman field as shown in Figure 10A. In these images the green sections within the fields are weeds.

The full field mosaic image can now be compared to the yield data gathered by the yield monitor. The yield monitor utilizes load, moisture, conductivity, and rotation sensors to develop, a near instantaneous picture of crop yield as it is being harvested (Technology, 2003). These sensors are calibrated before each harvest to reduce vibration induced error from the data. During the harvest several combine loads are sent to be measured separately. The yield and moisture data is then input into yield monitor to calibrate it (Colvin, 1999). The yield monitor data in Figure 10B shows some obvious qualitative similarities with the composite image of Figure 10A. The low yield section circled is clear in both. This section is a low section of the field that is regularly water logged. The worst yield in this section is 2.5 standard deviations below the mean yield of 112 (bu/acr) or approximately 25 (bu/acr). The boxed image is a row that has been heavily stressed and shows up in Figure 10A, Figure 10B as well as in Figure 5A. Figure 5A is an image of the same region earlier in the season. The yield in the boxed region is mostly below the mean by 1.5 standard deviations. This means the yield in this section is 59-60 (bu/acr).

Eventually, the weed sections of the georeferenced mosaic can be isolated by color and mapped to produce an overlay that can be correlated with yield maps similar to Figure 10B. We have had several problems with this to date. These problems become evident upon closer inspection of Figure 10A. First, the mosaic color has not been normalized resulting in different shades of green for each image. Second, many of the images are over exposed and much of the color range has been lost.

Third, the aerial images taken by both aircraft are too low in resolution. The color and exposure issues can be resolved by a combination of lens filters and computer based image processing. The resolution problem is in the cameras themselves. The cameras do not have enough pixels in the CCD arrays. Changing camera can easily solve this. Future efforts will address these problems with a series of camera and airframe upgrades.

CONCLUSION

This project has succeeded in its goals of developing a basic working infrastructure and skill set in aerial imaging and remote sensing at UMES. This is the beginning of one of the first small UAV remote sensing programs. The next steps will concentrate upon developing improved MATLAB codes for image georeferencing and

Figure 10. Corn harvest maps.

ASPRS 2006 Annual Conference
Reno, Nevada • May 1-5, 2006
developing software methods for image color normalization (Jensen, 1996; Finlayson, Schiele et al., 1998). Efforts are already underway to acquire better cameras and aerial platforms for improved results.

ACKNOWLEDGMENT AND COLLABORATIVE ARRANGEMENTS

The work presented in this paper is supported by ASTI, ECPA and AIRSPACES projects at UMES. ASTI (the Airborne Science and Technology Institute) is a collaboration among UMES, NASA, and L3/BAI Aerospace. The purpose of the institute is to educate students in the operation and use of Uninhabited Aerial Vehicles (UAVs) for earth science applications. ASTI project has enabled training ECPA project pilots. In return, the project has been able to evaluate the ASTI aircrafts as imaging platforms. Planning and implementation of improvements that will benefit both ASTI and the ECPA project are already underway.

This project has also been supported by a grant from the Maryland Space Grant Consortium (MDSGC). The MDSGC grant in support of the Aerial Imaging and Remote Sensing for Precision Agriculture and Environmental Stewardship (AIRSPACES) project has directly facilitated the work upon which this paper is based.

REFERENCES


