

# USING DEMS TO REGISTER OBLIQUE PHOTOGRAPHS AND WEB CAMERAS

**Peter L. Guth**, Professor  
Department of Oceanography  
United States Naval Academy  
572C Holloway Road  
Annapolis, MD 24102  
[pguth@usna.edu](mailto:pguth@usna.edu)

**Tom Craven**, Vice President  
GeodesyBase  
PO Box 1132  
Sisters, Oregon 97759  
[tcraven@aol.com](mailto:tcraven@aol.com)

## ABSTRACT

We have developed algorithms and software to use digital elevation models (DEMs) to register oblique photographs and web cameras. After registration, we can compute the coordinates of every ground pixel in the image and use the image for coordinate digitization and area measurement. Registration requires the camera location (latitude and longitude and elevation above ground, most easily obtained from a GPS unit), and three angles: the azimuthal heading of the camera, the upward or downward pitch of the camera, and a sideways tilt. The camera angular orientation proves the most difficult to accurately obtain with current cameras and orientation sensors. From these six camera parameters we create a synthetic perspective view from the DEM, with the same field of view and pixel size as the photograph, which we compare to the photograph. Using alpha blending of the DEM view and photograph, we iteratively adjust the camera position to best match the DEM view with the photograph. We can also drape imagery or vector overlays on the DEM in the perspective view, and use road intersections or other features to compare the photograph with the synthetic view and improve the estimated camera position. While DEM resolution, camera resolution, and lens optics limit the accuracy that can be achieved with this method, the procedure works well for many monitoring and surveillance applications. Improved sensors for camera orientation will greatly extend the potential uses of the algorithms.

## INTRODUCTION

A photograph or webcam image captures a depiction of the terrain, based on the geometry of the camera's orientation and the characteristics of the lens. If we can model that geometry, it would be possible to make measurements of position from the photograph or webcam image. Digital elevation models (DEMs) provide high resolution, three dimensional representations of the earth's topography, and can create synthetic images to match with a camera's depiction. We report here on work to use DEMs to register imagery and make measurements on the imagery.

We envision a range of applications for this technique. At one extreme, close in work with very high resolution DEMs like 1 m LIDAR (Fowler, 2001) would provide locations and monitoring in urban or other complex terrain. At the other extreme, regional work could use 10-30 m (1/3 to 1 arc second) DEMs and could provide positional information out to ranges of 10's of kilometers. Table 1 provides the details of the three examples shown in this paper. The first return LIDAR DEM includes vegetation and buildings, the most appropriate choice for applications like this (Guth, 2005).

**Table 1. Camera registration with DEMs used in this paper.**

Camera	DEM	Other data shown
Webcam of Half Dome from Sentinel Dome in Yosemite National Park. View extends 32 km. ( <a href="http://www.yosemite.org/vryos/sentinelcam.htm">http://www.yosemite.org/vryos/sentinelcam.htm</a> ),	10 m USGS SDTS	TIGER
Photograph from the fire tower on Black Butte, OR, View extends 28 km.	1 second NED	TIGER and Landsat Thematic Mapper
Photograph from Room 318, Nimitz Library, US Naval Academy. View extends 1.8 km.	1 m LIDAR	0.75 m USGS Urban Area Orthoimagery

## METHODS

Table 2 shows the seven parameters required to register an oblique photo or web cam image. Figure 1 provides a flow chart with the sequence of steps to register an image. With an actual or assumed value for all seven parameters except the roll, we create a synthetic perspective view from the DEM with all parameters to match the camera. This view will have the same dimensions (pixel size) as the photograph, and the vertical field of view is defined by the horizontal field of view and the aspect ratio of the image. We use the roll angle to rotate the photograph or webcam image to horizontal, which is easier than roll-rotating the DEM generated view but creates the same result. We then compare the synthetic view to the photograph by superimposing the two images and using alpha blending to create a merge which makes comparison easy and fast. We iterate by revising the parameters until we obtain a best visual match. While we generally follow the sequence shown in Figure 1, we may adjust parameters at any time if the initial estimate was so far off that it precludes getting a reasonable match.

**Table 2. Parameters required for DEM image registration**

Variable	Definition	Determined From
Latitude		GPS, map, aerial photography
Longitude		GPS, map, aerial photography
Elevation	Camera height above the terrain	Knowledge of camera location above the DEM elevation
Azimuth	Direction to the center of the view	Compass integrated with camera, or comparison with DEM generated view
Pitch	Up/down tilt angle of the camera, about a horizontal axis in the focal plane of the camera	Orientation sensor integrated with camera, or comparison with DEM generated view
Roll	Sideways tilt angle of the camera, about a horizontal axis perpendicular to the focal plane of the camera	Orientation sensor integrated with camera, or comparison with DEM generated view
Horizontal field of view	Angular width, depending on camera focal length	Camera specifications, EXIF information in JPEG file, or comparison with DEM generated view

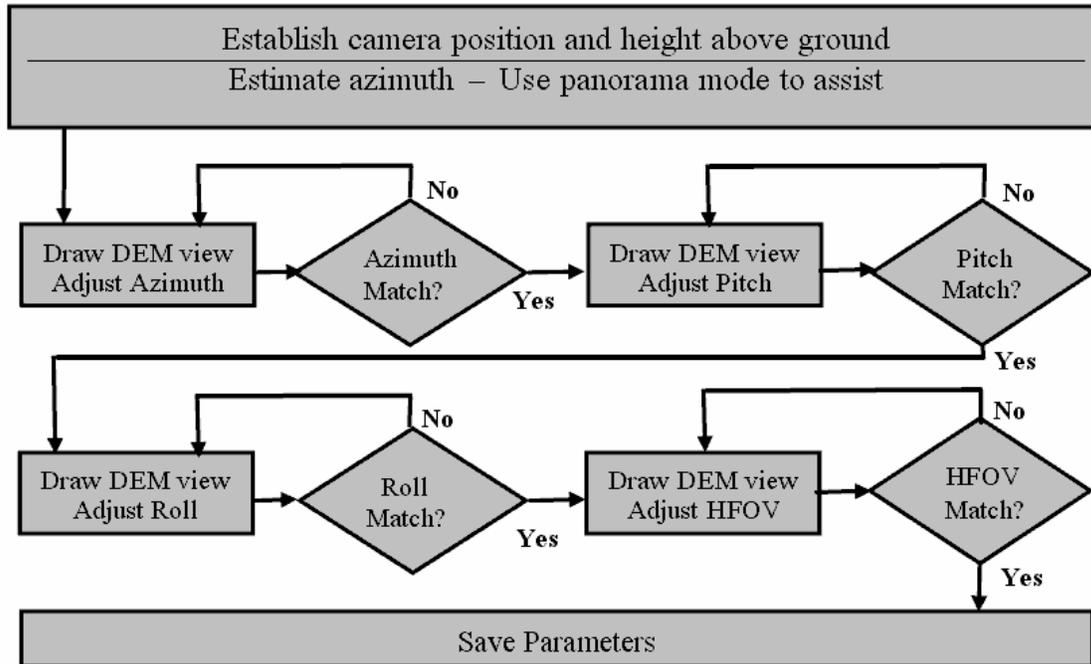
For photographs we take and webcams we control directly, we get the latitude and longitude from a hand-held GPS and the elevation from an estimate of the height of the camera above the ground. For the accuracy we need for the imagery and ranges we intend to use, this provides location information within a few meters and these parameters can be considered fixed for our registration needs. For webcams we do not control directly, this information may or may not be available on the website, and its absence complicates the registration process. We have used Terraserver imagery ([www.terraserver.com](http://www.terraserver.com)) to try to locate structures on which the webcam might be mounted.

We can use a DEM reflectance map, or a satellite/air photo drape, and can also overlay vector data such as TIGER roads and streams, to improve the registration process by providing more points to compare in the photograph and the DEM view. We can highlight ridge crests in the synthetic view, which can help in interpreting the photograph. At one time we overlaid the horizon computed from the DEM on the photograph, but the alpha blending provides the horizon and so many other points for comparison that we now use only the alpha blending. A panorama view, taken drawn over 360° from the camera location, can greatly assist in the initial parameter selection.

If the JPEG file uses the EXIF standard (JEITA, 2002), it might be possible to compute the horizontal field of view

from the camera focal length data stored in the file. We have not seen EXIF metadata with any webcams, and have found it simplest to treat the HFOV as a variable. We compute the vertical field of view (VFOV) from the HFOV and the aspect ratio of the image.

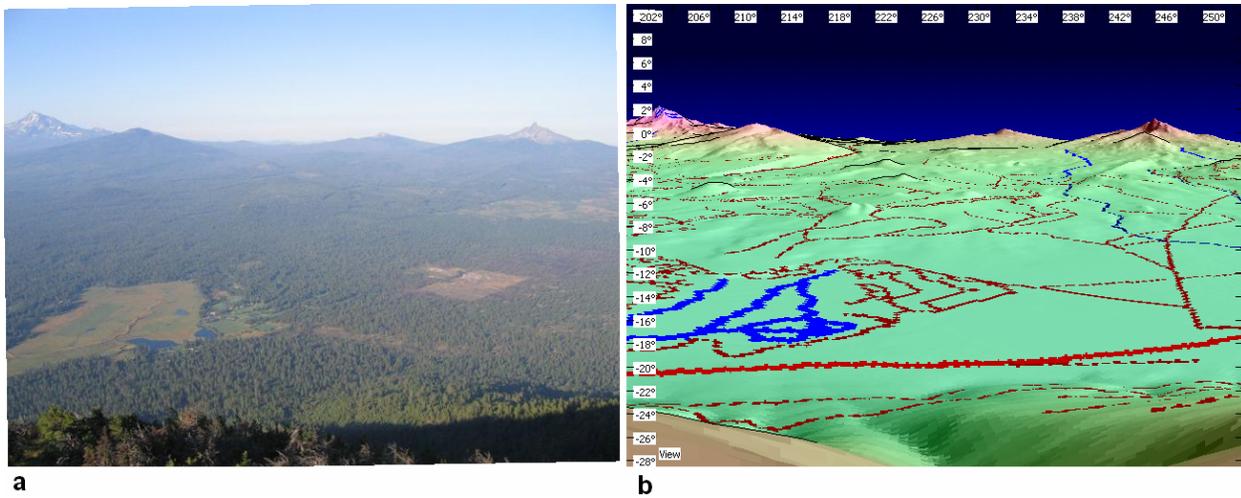
Large camera images present challenges in the registration process unless a very high resolution monitor can be used, and megapixel camera resolution is increasing faster than monitor resolution. Using scroll bars to navigate a large image does not work well with our current software implementation for many operations (such as alpha blending), and we have resolved this by subsampling the photos. This leads to decreased accuracy in our registrations, as will be discussed below, and is one of the improvements that we will be working on.



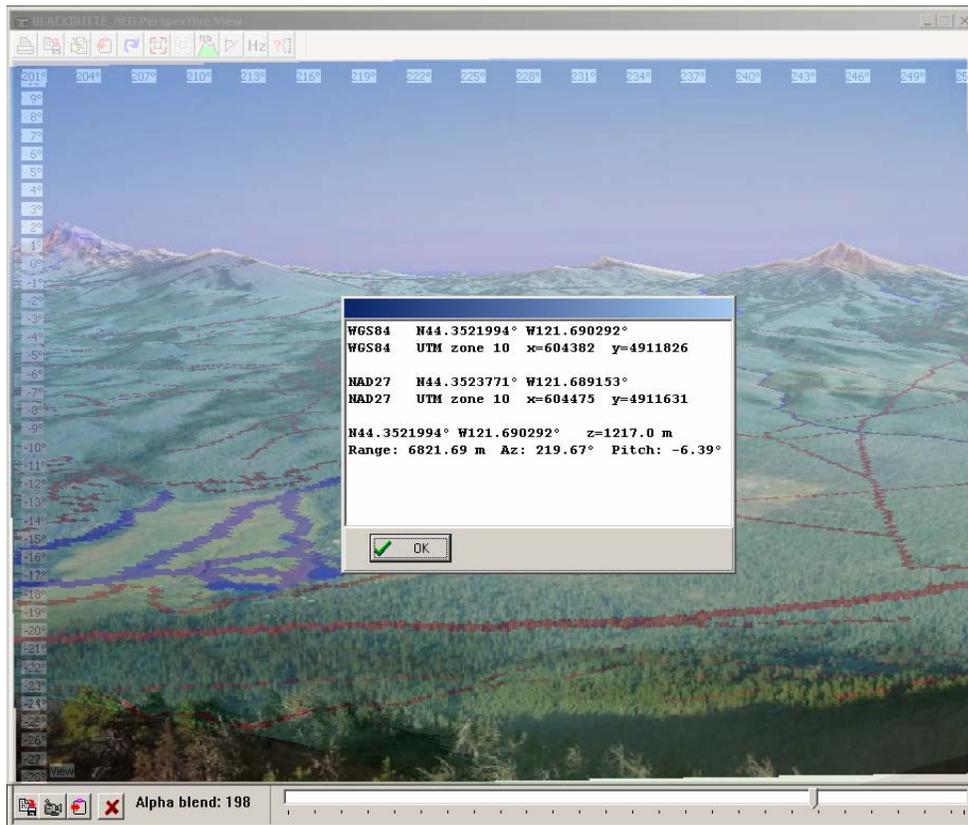
**Figure 1.** Flow chart of the registration process for a photograph taken with a camera from a known location.

## RESULTS

Figure 2 shows the results of registering a photograph taken with a digital camera from the fire lookout tower atop Black Butte near Sisters, Oregon. The camera view on the left matches the shaded reflectance view on the right, which also overlays the roads and streams from the Census Bureau's TIGER data. In this heavily forested area, the roads do not show up clearly in the photograph, although some of the stream valleys can be seen. Figure 3 shows an alpha blending of the two images, and the ability to measure coordinates, range, bearing, and pitch to any location on the photograph.



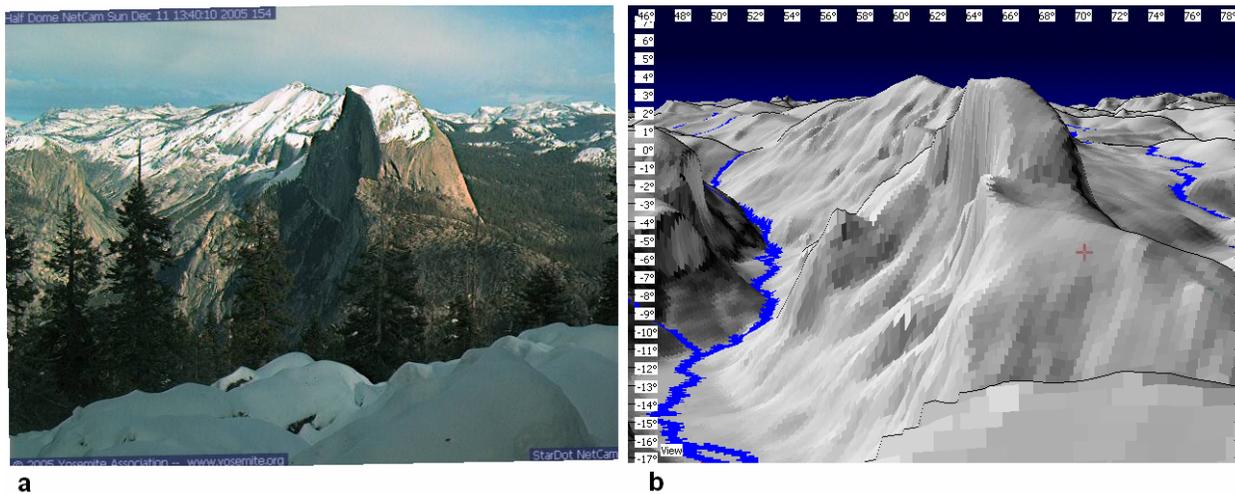
**Figure 2.** View from Black Butte, Oregon. (a) Shot from a digital camera, and (b) DEM view that matches the same view and overlays the roads and streams from the Census Bureau's TIGER data base.



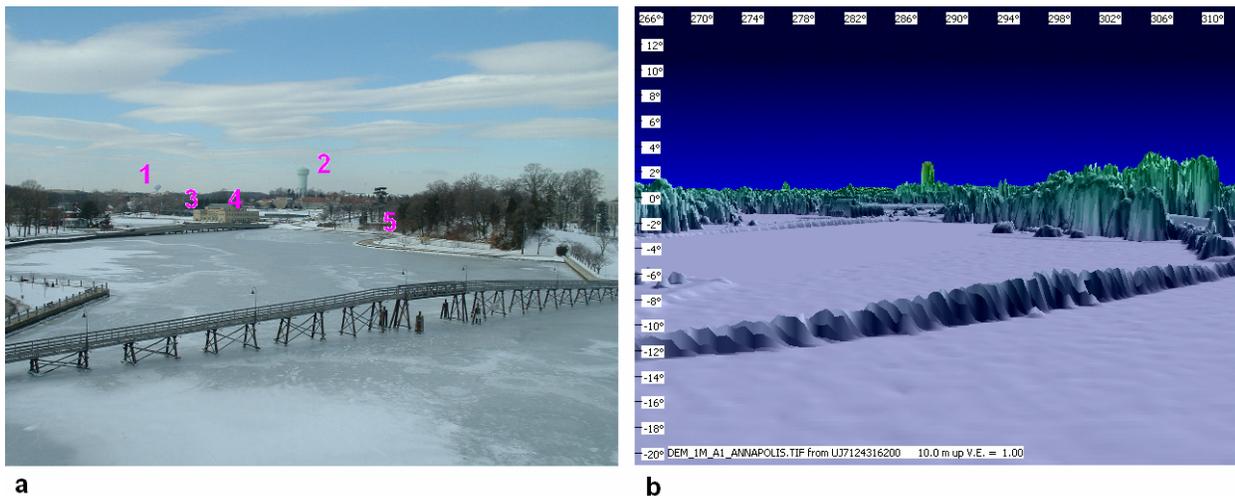
**Figure 3.** Alpha blend of the view from Black Butte, OR, in Figure 2, showing the window controls and the positional information displayed when clicking on a location within the image. Location, range, azimuth, and pitch computations also appear in a status bar of the main program window while roaming over the image with the mouse. Note the slider bar in the bottom right, which allows varying the alpha blending to smoothly transition from the photograph to the DEM view.

Figure 4 shows the registration of a webcam image found on the world wide web. This image had only very general location information (Sentinel Dome, Yosemite National Park) and required iteration involving all 7 parameters to achieve the registration shown. The TIGER data has no roads that show up in the view, and the two streams have little expression because of the tree cover.

Figure 5 shows a photograph of College Creek in Annapolis, Maryland. This covers a much smaller region than the large panoramas shown in Figures 2 and 4, but we have a high resolution (1 m grid postings) LIDAR DEM which shows individual trees and buildings. The photo was taken during winter, and the LIDAR first return DEM in summer with leaves on the trees, which created some challenges for the registration process.



**Figure 4.** (a) Webcam view of Half Dome from Sentinel Dome in Yosemite National Park on the left (<http://www.yosemite.org/vryos/sentinelcam.htm>), and (b) synthetic view on the right constructed from USGS 10 m DEMs. We started with no knowledge of the camera location, and successfully registered the view by iteratively adjusting all 7 parameters.

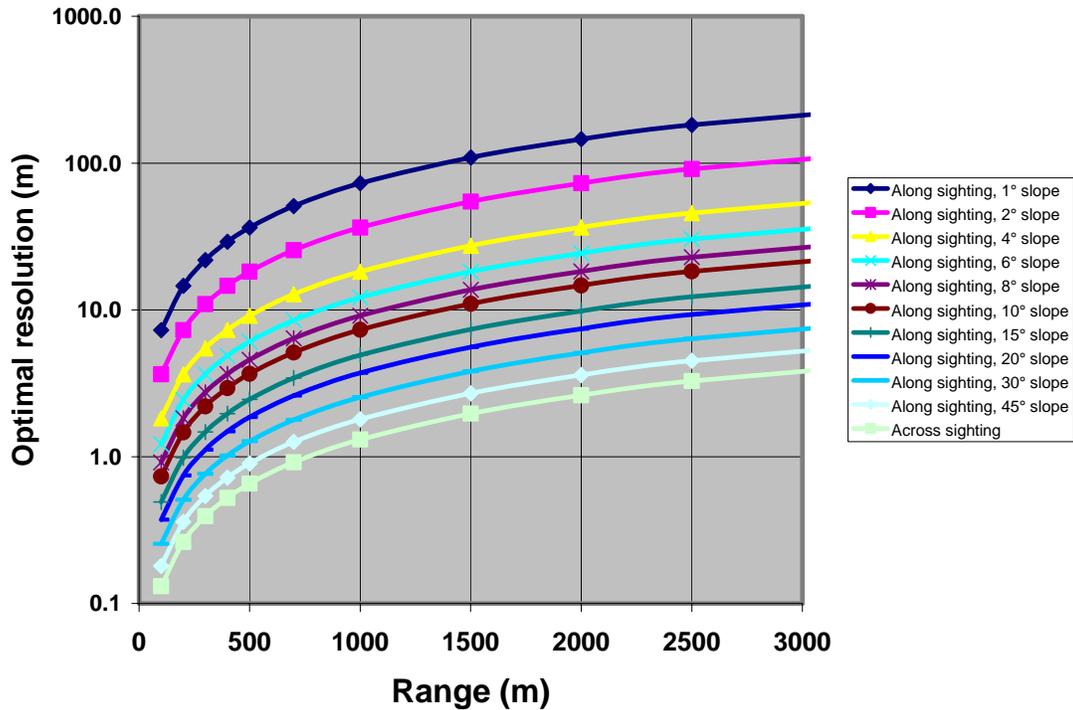


**Figure 5.** View from Nimitz Library, US Naval Academy, taken with a Fuji FinePix 2800Z digital camera. Synthetic view on the right created with a 1 m LIDAR DEM. The numbers on the photograph correspond to the target features shown in Table 2. The water tower at #1 is about 1800 m from the observer.

## ERROR ANALYSIS

The accuracy achievable with this process depends on the pixel size of the photograph, the field of view from the lens, since most digital and web cameras provide a zoom function, and the geometry of the terrain. As in our analysis using DEMs to locate targets from an Osborne firefinder (Guth et al., 2005), we can create a spreadsheet showing the expected resolution, and then graph the results. Figure 6 shows the results for the camera geometry depicted in Figure 5. We divide the best resolution we can achieve (or the potential error) into along sighting and across sighting components. Across sighting error is a function of the horizontal field of view and the width of the image in pixels. As an angular measure (degrees per pixel), it increases with range from the camera. Along sighting error depends on the viewing geometry, and will be best when the sight line is perpendicular to topography either because of an elevated camera or steep topography. Along sighting error also increase as the range from the camera increases. Lens distortion is beyond the scope of this analysis, but our work with multiple web cameras and consumer grade digital cameras suggests that this procedure provides a reasonable first approximation to location determination.

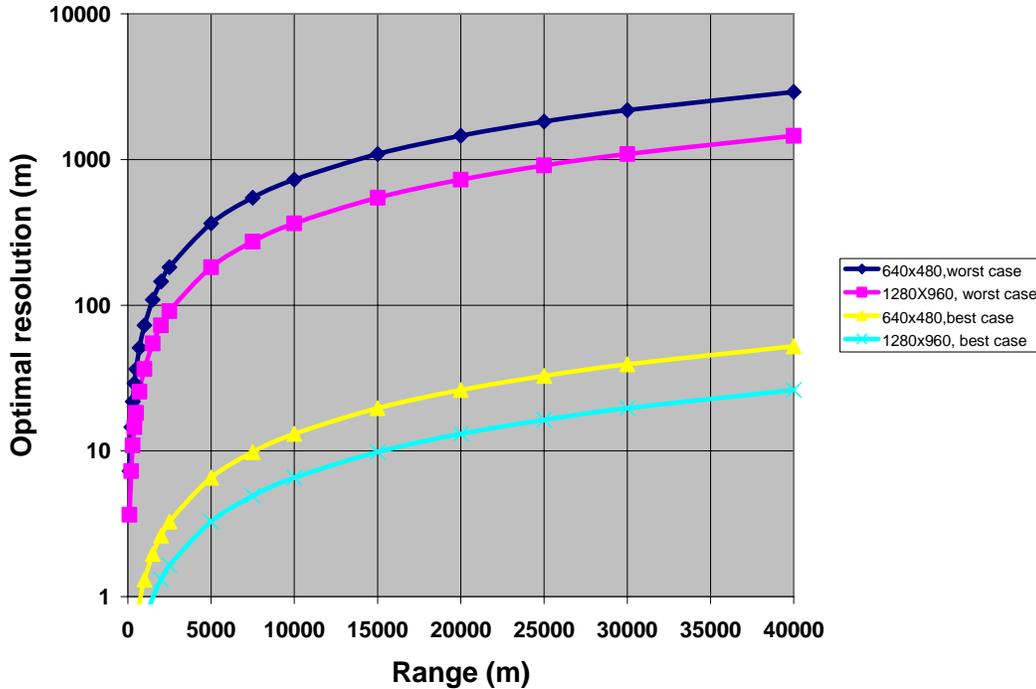
Figure 7 shows the effect of camera resolution on the possible accuracy achieved by DEM registration. The best case is the across sighting resolution, which will also be the along sighting resolution if the sight line intersects the terrain at a 90° angle. The worst case is the along sighting resolution for an intersection of 1° between the sight line and the terrain. The two camera resolutions shown differ by a factor of two in the registration resolution they can achieve, as pixels twice as big double the uncertainty in locations.



**Figure 6.** Optimal resolution for the camera and DEM view geometry, which depends on image size in pixels and the zoom setting of the camera lens, which determines horizontal field of angle. This graph is for a 640x480 image with an HFOV of 48°. The slope angle is the angle between the camera’s sight line and topography, and for an elevated camera could approach 90° which would provide the same resolution as the across sighting values. Note the logarithmic scale on the vertical axis, and that the graphs in fact show a linear relationship with range.

Table 3 shows the results of testing the location error using the DEM registration of the scene shown in Figure 5. We registered a digital camera image, taken in the winter, to a DEM view created with a LIDAR DEM. We then located five points on the image, and determined their locations from the registered DEM view. Our “ground truth” for the locations came from the LIDAR DEM, a 1 m USGS color infrared DOQQ (USGS SDDS, 2006), and 0.75 m

USGS Urban Area Orthoimagery (USGS SDDS, 2006). Except for the tops of the two water towers, which suffered uncorrected radial displacement in both images, the coordinates for the registration points agreed in all three data sets. Figure 8 shows the locations of the 5 test points on the LIDAR DEM, and the locations calculated from the registered photograph.



**Figure 7.** Optimal resolution for the same camera HFOV of 48°, two different pixel resolutions, and best and worst case DEM view geometry.

The results for the first four control points in Table 3 and Figure 8 show that we can location points within 15 m for points 500 m to almost 2000 m away from the camera. These results used a photograph with resolution of 640 by 480, and this accuracy is consistent with that suggested in Figure 7. The errors in Table 3 combine the across sighting and along sighting errors in Figure 6 in a single error value. Two points, the closest to the camera location, had much larger errors that suggest limitations to the method. We located target 4 twice, clicking high on the building and low on the building. The location derived low on the building proved to be off by 100 m, because in the DEM view the line of sight intersected the bridge. Either the registration was off slightly or the mouse selection was off by a pixel. In either case, this suggests that selecting points in the center of a target, or using the highest resolution possible, will minimize the impact of such effects. Target 5, the tree at the SW end of the columbarium, shows the challenges of comparing leaf-on and leaf-off imagery in LIDAR DEMs. Matching a bare tree to its LIDAR image is difficult, and in this case the ground truth location may also be in error in addition to locating the tree on the leaf-off picture.

Figure 9 shows the coverage on the DEM map of the photograph in Figure 5. Each red dot on the map shows the location of a pixel on the photograph, and the patterns clearly show the along track errors. The camera in this case is on the fourth floor of a large library, about 30 m above the water in the foreground of the photograph. This minimizes the along track potential for two reasons: the intersection of the radial viewing lines with the topography will be a relatively large angle close the camera, and the range will be small. Close to the camera location, the map shows a solid red pattern. Farther away, the red lines form discrete arcs with increasing spacing. A point on the photo will be either on one arc or its adjacent neighbor when we use the DEM view for location determination, and the distance between arcs is the along sighting error. Additionally, the pattern shows the effects of buildings and vegetation in creating shadow or masked zones that cannot be seen from the library.

**Table 3. Location error from tests at the US Naval Academy**

TARGET FEATURE	GROUND LOCATION		PHOTO LOCATION		LOCATION ERROR (m)
	LAT	LONG	LAT2	LONG2	
1--Stadium Water Tower, range=1780 m	38.9872363°	-76.5070383°	38.9872436°	-76.5069457°	8.0
2--Perry Circle Water Tower, range=1140 m	38.9884974°	-76.4990743°	38.9885035°	-76.4989393°	11.7
3--SW Corner Hubbard Hall, range=590 m	38.9862237°	-76.493401°	38.986227°	-76.4932485°	13.2
4--SE Corner Hubbard Hall (high), range=550 m	38.986385°	-76.4928921°	38.9864023°	-76.4929228°	3.3
4--SE Corner Hubbard Hall (low), range=550 m	38.986385°	-76.4928921°	38.9862079°	-76.4917722°	99.0
5--Tree SW end of columbarium, range=330 m	38.9865618°	-76.4901526°	38.9863872°	-76.4895809°	53.2

## DISCUSSION

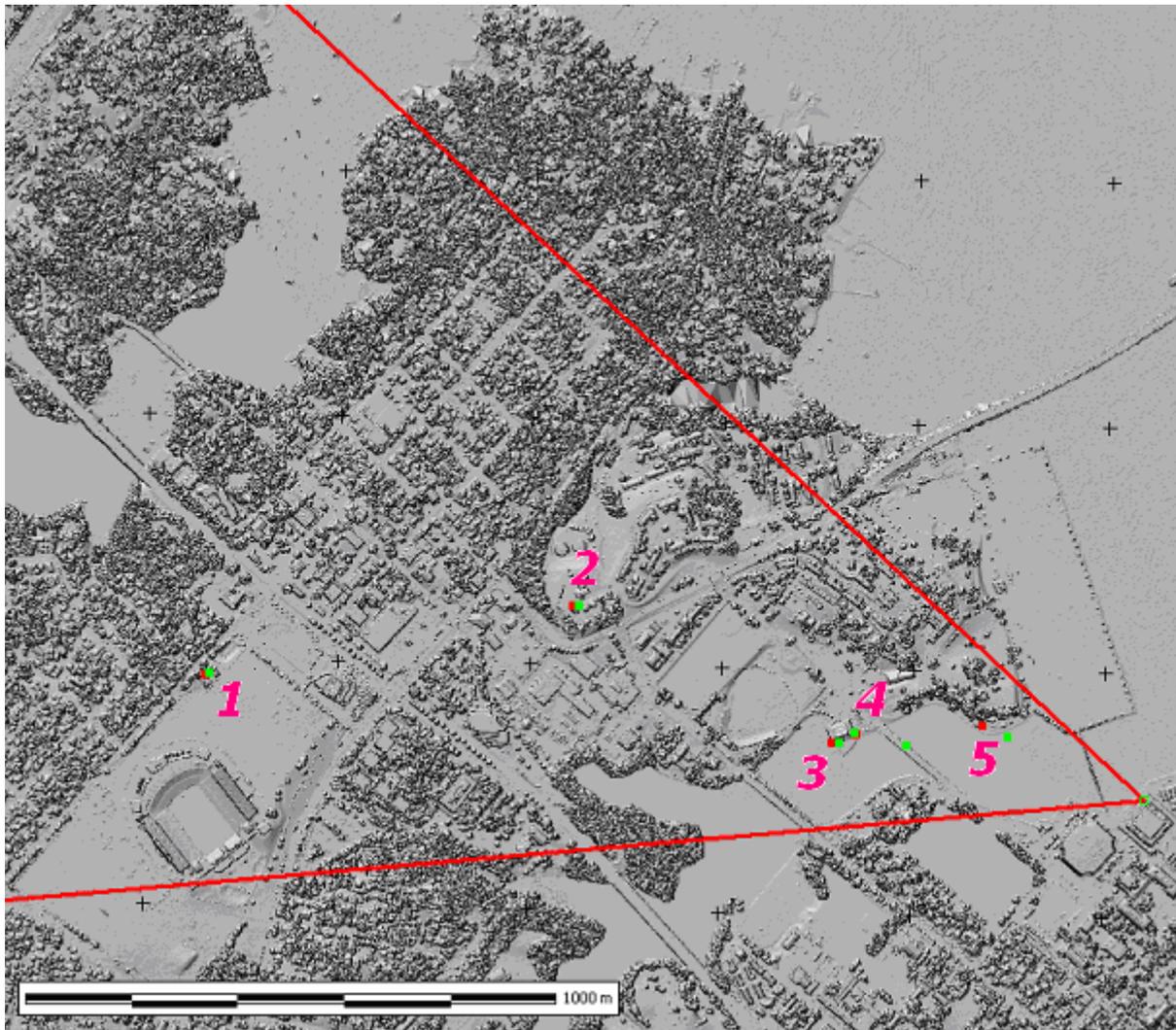
The images presented so far have used DEMs to provide the base map and 3D view. These include grayscale reflectance maps (Figures 8 and 9), grayscale 3d perspective views (Figure 4b), color tinted reflectance views (Figures 2b and 5b), and draping of vector data such as TIGER line data on the maps (Figures 2b and 4b). In addition, we can drape imagery such as Landsat TM (Figure 10b) or the USGS high resolution orthoimagery rapidly becoming available (Figure 10c and 10d). In urban areas, the results of image draping will probably resemble Figure 10d, where we will have to accept the lack of coverage on building sides, but in many cases the imagery will help to locate points in the photograph, such as the golf course in Figure 10 b.

A major challenge in improving our results will be in software design and user interface issues with the large images and map that must be open simultaneously. Image registration without full camera orientation and location requires having a DEM map, an image map, the photograph, and the 3D DEM view open simultaneously. Digital photographs can easily fill the entire screen, and the user must have a reasonable set of windows and controls available.

We have experimented with sensors integrated with the camera to record the GPS location and camera orientation into the EXIF data that can be embedded into a JPEG image (JEITA, 2002). We have not found a satisfactory system including a good camera and complete orientation information, but the potential exists for this tight integration, which would decrease the required effort to fully register imagery.

## CONCLUSIONS

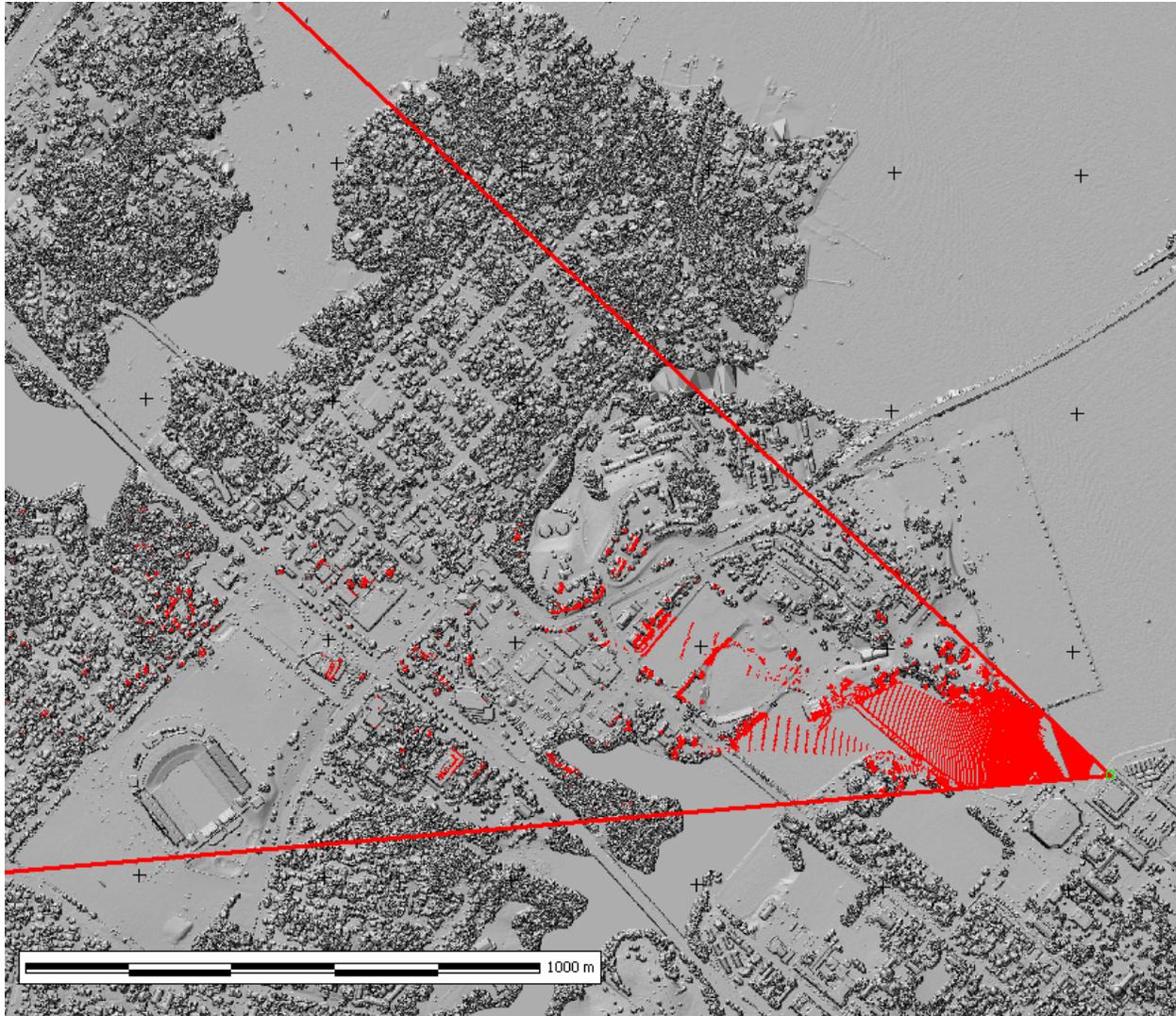
These results show the potential of using DEMs to provide location data in conjunction with other observation. This builds on our work with the Osborne firefinder located in many fire towers (Guth et al., 2005), but could also be extended to unmanned locations. Webcams on peaks with solar power and wireless communications have already been deployed to monitor fires (National Center for Landscape Fire Analysis, 2006), and in that case the 3D location will be known to high accuracy. Our experience shows that the resulting webcam image can be registered relatively easily, and that we could monitor the area covered by the camera and accurately locate new fires. Integrating DEMs and other GIS data directory with field sensors and observations promises to greatly increase our ability to monitor the environment and accurately locate positions.



**Figure 8.** Coverage for the photograph and DEM view shown in Figure 5. The five targets shown in Table 3 are indicated by the red squares, and the computed positions using the DEM registered photograph are show by the green squares. This image has approximately 4 m pixels, and does not capture the full information present in the LIDAR DEM.

### ACKNOWLEDGEMENTS

PLG is grateful to have been able to spend several weeks each winter for the past three years in Nimitz Library Room 318 with its view of the changing moods of College Creek, and thanks Steve Sarigianis for alerting him to the availability of the LIDAR DEM. Combining JPEG images from Nimitz 318 and the LIDAR DEM provided a good start to documenting the accuracy of the DEM registration process.



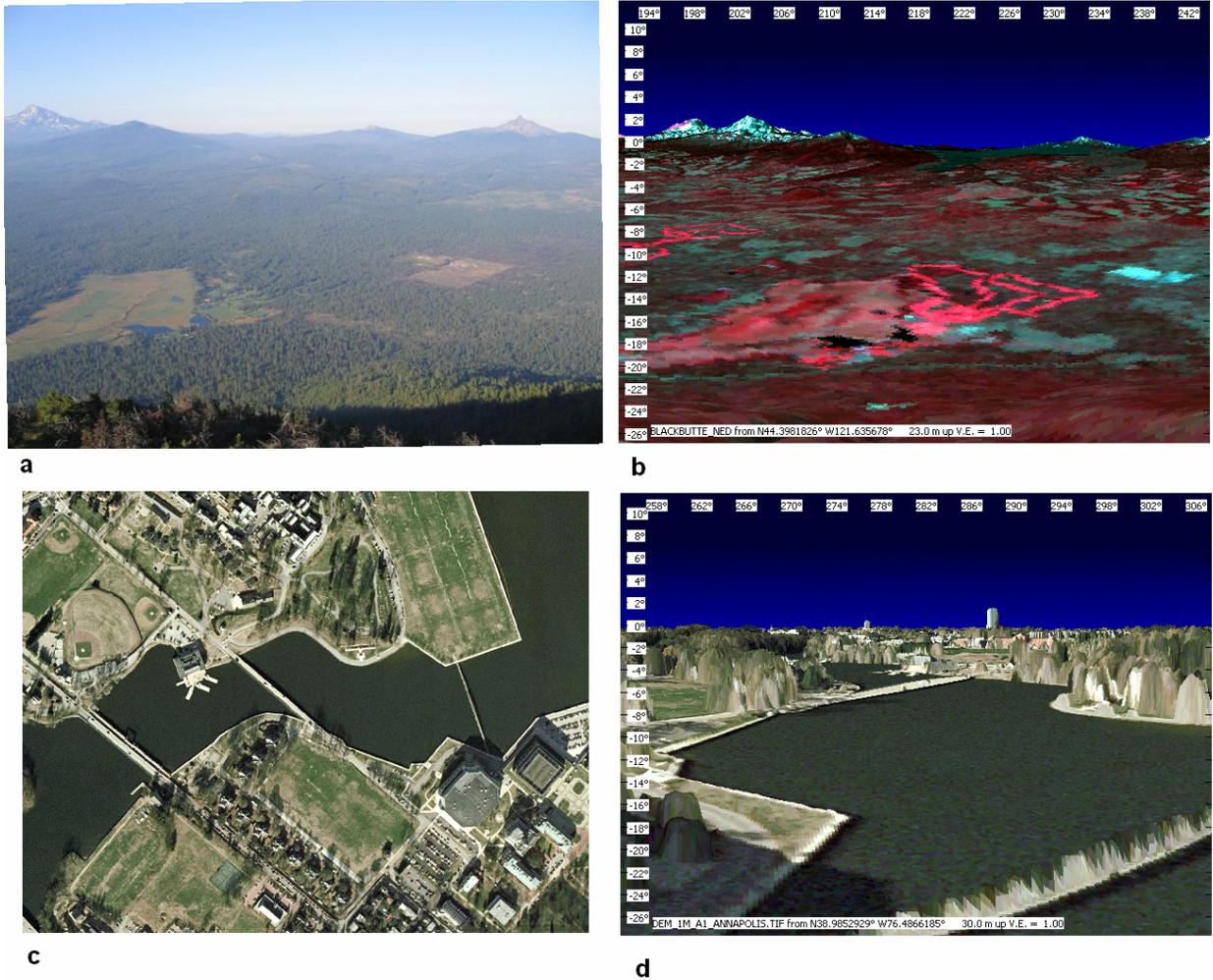
**Figure 9.** Computed position of every pixel in the photograph in Figure 5 indicated by a red point. Because of the scene geometry and the summer foliage in the DEM, this photograph provides good positional accuracy in the foreground, and only limited coverage in the background. Because the comparison photograph was a winter shot, many more building appear in it that are masked by foliage in the DEM view.

## REFERENCES

- Fowler, R. (2001). Topographic lidar. In Maune, D.F., ed., *Digital elevation model technologies and applications: the DEM users manual*. The American Society for Photogrammetry and Remote Sensing, Bethesda, pp 207-236
- Guth, P.L. (2005). Challenges for Military Application of High Resolution DEMs: in MilGeo2005, Nathanail, C.P, and Mangaroo, A., eds., *Proceedings, Sixth International Conference on Military Geology and Geography*, 19 - 22 June 2005, School of Geography, University of Nottingham, United Kingdom, 14 p.
- Guth, P., Craven, T., Chester, T., O'Leary, Z., and Shotwell, J. (2005). Fire location from a single Osborne firefinder and a DEM: Proceedings ASPRS 2005 Annual Conference, Baltimore, MD, March 7-11, 2005, 10 page paper on CD-ROM.
- JEITA. (2002) Exchangeable image file format for digital still cameras: Exif Version 2.2, <http://www.exif.org/Exif2-2.PDF>, last accessed 6 Feb 2006.

National Center for Landscape Fire Analysis. (2006). Fire Intelligence Module: Assignment to the Freezeout Fire: <http://firecenter.forestry.umt.edu/firecenter/default.asp?content=FireMod&display=freezeout>, last accessed 6 Feb 2006.

USGS SDDS. (2006). Seamless Data Distribution System, <http://seamless.usgs.gov/>, last accessed 5 Feb 2006.



**Figure 10.** (a) Photograph from Black Butte shown in Figure 2. (b) Landsat TM imagery (bands 4-3-2) draped on 10 m USGS DEM to match the photograph. The bright red regions in the center foreground are the fairways of a golf course, which do not show up nearly as clearly in the photograph. (c) High resolution orthoimagery (0.75 m) from the USGS Seamless Server, covering the foreground area of the photograph in Figure 5. (d) DEM view with the imagery draped on the LIDAR, showing the lack of coverage on the sides of buildings, although the blue water tower in the center of the views shows up well.