Absolute Range Detection System for STAIR

Giancarlo Garcia and David Jackson – *December 12, 2008*

Introduction

Absolute range detection has been a long-time objective of studies focused on achieving stereo computer vision. Many researchers have attempted to emulate some part of the human stereo system, most commonly motion parallax. These attempts, however, have concluded that motion parallax is not a reliable source of input for systems attempting range measurements and, in fact, often worsens object size and range prediction [1]-[2]. A system with these qualities is needed for applications such as the Stanford AI Robot (STAIR), which must be able to maneuver through dynamic environments in order to carry out its tasks.

Our approach to tackling the problem of range-detection was two-fold. Our first system replaced one of the two cameras in a traditional stereo-vision system with a laser marker. By holding the marker and camera fixed while varying the their distance to an object, a training set was produced from which any new distance could be inferred. Utilizing this method we found that we could achieve reliable (±3%) range estimates at up to 160 feet.

In our second approach, our original objective was to expand upon our first approach to produce a system that could provide absolute range for any pixel in a given camera image, not just one at a time. Through trial and error we eventually we settled on a fringe projection system put forth by Zhang and Yau [3] that allowed us to produce phase maps from which very accurate range information can be extracted.

Part I – Camera-Marker System

A single camera image does not provide enough information on its own for a robot to infer range information. Although it is hard to imagine, this is true as well in humans. While humans utilize several methods of range detection, motion parallax – the ability to see how objects shift with respect to eye position – requires multiple images. By introducing a laser marker in place of a second image capture, we essentially reduce the magnitude of the problem of implementing a motion parallax in machines for full images to one small subset equivalent to a point in an image. While this is not ideal in practice, it helps us to move in the right direction if we find that this method gives promising results.

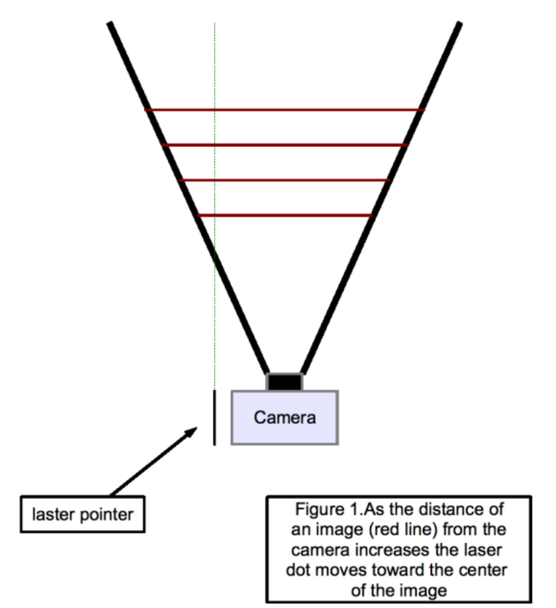


Figure 1

A model for the experiment can be developed by considering what happens when we shine a laser into the field of view (FOV) of a camera. Assuming the laser and camera are aligned, the laser marker at some distance (depth) x enters the camera FOV. Assuming the FOV grows linearly with distance, then we can infer that the ratio of the distance between the marker and center of the and the marker and the edge of the image will go as the inverse of this relationship. Since the image is discretized, the natural interpretation of this ratio is that of pixels per "true" area. Naturally, as a camera captures objects that are far away, less pixels are devoted to objects in the distance than in the foreground. Hence, the distance of the marker to the center of the image can be expressed as:

$$x_{marker} \approx \theta_1 + \theta_2 \frac{1}{d}$$

Here d is the true distance from the camera to the object the laser shines on, and θ_1 and θ_2 are learned parameters to be determined by our learning algorithm.

The experiment we carried out was very simple: by bringing the laser marker into the field of view of the camera, image subtraction can be used on two images of the same scene (laser on/laser off) to determine the approximate pixel-position of the laser in the image. Our setup for this experiment included one camera (Canon sub-SLR, 7 MP, 12x physical zoom) and one (green) laser pointer. Both camera and laser were fixed to a wooden plank and placed on a rolling cart. On each experiment conducted, a fixed target was chosen such that there was ample range of motion away from this target while maintaining equivalent lighting conditions.

Testing was conducted inside for optimal lighting conditions and, once this proved promising, we moved outdoors to produce a training set. In this outdoor

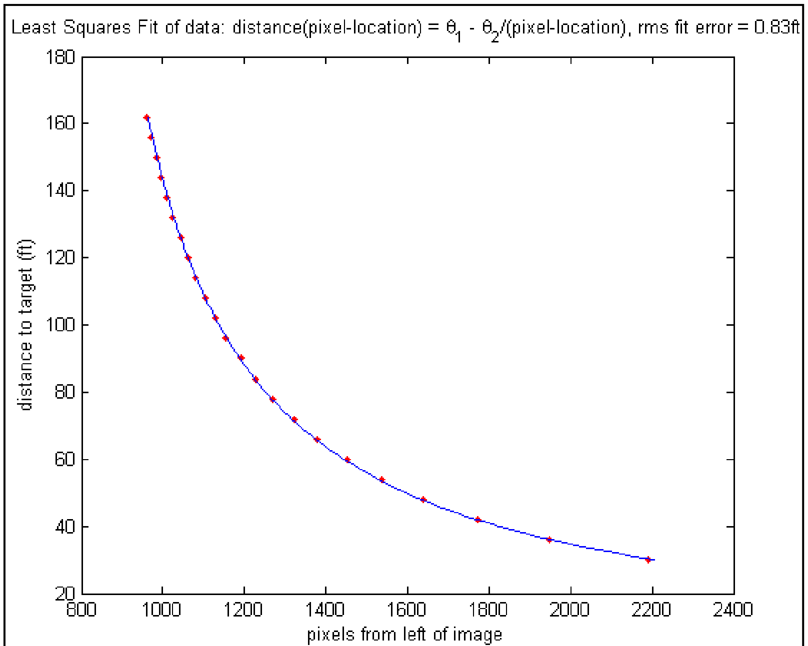


Figure 2

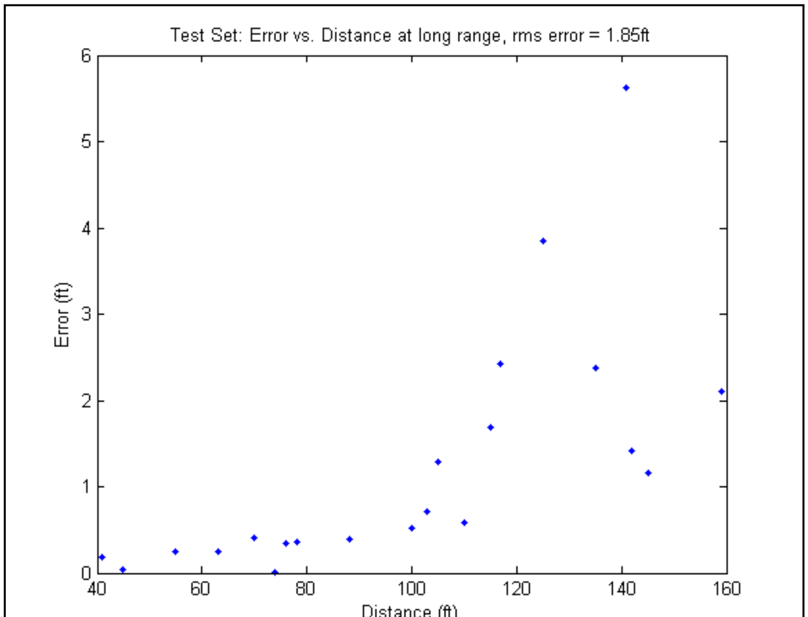


Figure 3

location, the laser pointer was projected onto an office chair placed between 6 and 200 ft away from the camera-laser setup at 3ft intervals (the camera and laser were held fixed and the chair was moved to ensure fixed alignment between the laser and camera). The camera was set at its highest physical zoom level (12x). Figure 2 shows the training set produced using this procedure. The data was fit using the above model utilizing the least-squares method. Least squares guarantees the best fit in the sense that the error will be minimized. Because we could infer the model from our experimental setup, we are confident that this model gives accurate results for regions not explored by the training set.

Several distances were then chosen at random and utilized as a testing set. For the test set several different types of materials were used as targets, with varying texture, color, and reflectivity. The error produced by the test set (figure 3) is very reasonable for values up to 160ft.

Part II – Fringe-Projection System

Our goal for the second part was to develop a system that could determine absolute range on all points in an image at or near real-time speeds. The work done by Zhang provided just what we were looking for, as he has demonstrated a system which maps 2D into 3D data in real time. This system uses a fast camera with synced DLP projector to obtain three pictures of the same image with three sinusoidal fringe patterns projected onto it by the DLP projector. We used three fringe patterns, each offset by 120 degrees:

$$I_1(x,y) = \alpha + \beta\cos(\phi(x,y) - \frac{2\pi}{3})$$

$$I_2(x,y) = \alpha + \beta\cos(\phi(x,y))$$

$$I_3(x,y) = \alpha + \beta\cos(\phi(x,y) + \frac{2\pi}{3})$$

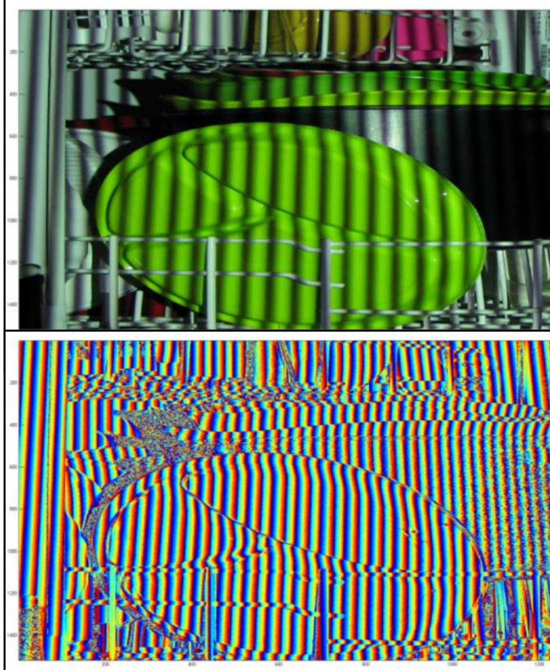
The parameters α and β are fixed coefficients that represent the average intensity and intensity modulation, respectively. Once the fringes-projected images are captured, the effective phase map can be retrieved as:

$$\phi(x,y) = \tan^{-1}(\sqrt{3}\,\frac{I_1 - I_3}{2I_2 - I_1 - I_3})$$

Phase unwrapping is then used to obtain range measurements for each pixel in the image. We were able to reproduce this simple system using a DLP projector and camera. We obtained three fringe-projected images and used these to obtain a phase map (figure 4). From here we knew that phase unwrapping software is readily available to perform the phase-map to range-map transformation.

One drawback that became evident during the testing process is that Zhang et al only use their system in optimal conditions (e.g. a very dark room) so that the light from the DLP projector alone provides a high enough SNR for the fringe patterns to be easily captured on camera. As our main project aim was to produce a range detection system for STAIR, we obviously needed something that could work in sub-optimal lighting conditions (e.g. fluorescent lighting or, at worst, sunlight). We tested the same system in sub-optimal lighting conditions and, as expected, received much worse results (see figure 5a).

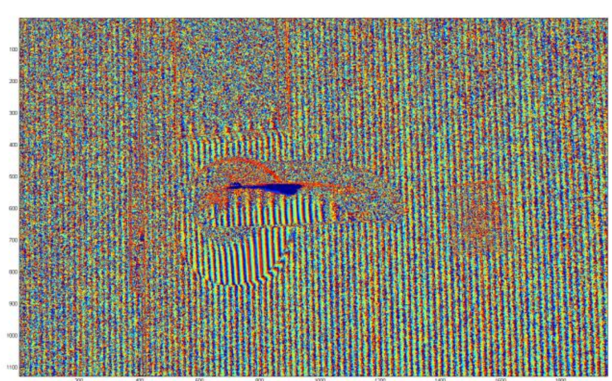
In order to raise the SNR to an adequate level we simply needed to produce more light in order to overcome sources like fluorescent lighting and sunlight. Several methods were discussed, including retrofitting a DLP projector with a stronger light source or moving to the near infrared spectrum (non-overlapping with sunlight). The latter idea received some attention, as Light with at 1800-1900nm wavelength is absorbed by water vapor, so there is almost no background sunlight at this wavelength. Likewise florescent, halogen, and incandescent lights do not produce light in this area of the spectrum. Commercial infrared cameras exist that are sensitive to these wavelengths of light. Conceptually, a system could operate and not disturb surrounding people or have to compete with background ambient lighting. While an interesting vision, exploration and engineering of this concept was not feasible in the time-span allotted by this class.



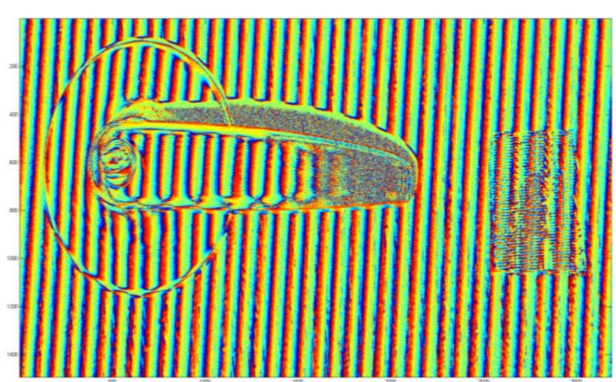
DLP projector in darkness
Figure 4

In short, several of these ideas were pursued, but with no engineering success. After running into difficulty retrofitting a DLP projector with an alternate light source (halogen, HID, xenon, etc.) we decided to try the same idea on an overhead projector, which is much less sophisticated technologically. In order to get the maximum light output we used a Nikon SB-600 camera flash synced with a Nikon D60 DSLR. A custom opening was made for the flash so that it aligned optimally with the mirror beneath the projection screen. A fringe patter was printed on a transparency and moved manually for experiments.

Results using this system were an all-around success (figure 5b). Tests in fluorescent lighting proved comparable to results for the DLP projector trials in absolute darkness. In addition, its power costs are very low compared to the 200W required for a DLP projector, and thus make it suitable as an on-board system for robots such as STAIR.



DLP projector in fluorescent light
Figure 5a



Flash/DSLR in fluorescent light
Figure 5b

Conclusions

We have made some headway into the problem of reliable absolute range detection for autonomous robots. As a first step, we developed a system that reliably predicts the distance of one point in an image up to 160 ft away. Next we worked on ways to extend this ability to capture absolute range for all pixels in an image. Fringe projection provided the basis for this work and we successfully made headway into adapting Zhang and Yau's fringe-projection technique to situations with sub-optimal lighting.

***David and Giancarlo want to thank Morgan and Andrew for their constant support and guidance

References

[1] Luo, Kenyon, Kamper, et al. The Effects of Scene Complexity, Stereovision, and Motion Parallax on SizeConstancy in a Virtual Environment. Virtual Reality Conference, 2007. VR '07. IEEE**.**

[2] Rauschecker, A. M., Solomon, S. G., & Glennerster, A. (2006). Stereo and motion parallax cues in human 3D vision: Can they vanish without a trace? *Journal of Vision*, *6*(12):12, 1471-1485.

[3] Zhang and Yau. High-resolution, real-time 3D absolute coordinate measurement based on a phase-shifting method. Optics Express, Vol. 4, Issue 7, pp 2644-2649.