Single Frame Processing for Structured Light Based Obstacle Detection

David Ilstrup *Autonomous Systems Lab*, Computer Science, UC Santa Cruz
Gabriel Hugh Elkaim *Autonomous Systems Lab*, Computer Engineering, UC Santa Cruz

BIOGRAPHY

David Ilstrup is a graduate student at the University of California, Santa Cruz, in the Computer Science department. His research interests include inference and estimation in hybrid systems, sensor fusion, tools and theory for the design, optimization and verification of embedded systems. David received a B.A. degree in Music from UC Berkeley in 1987 and has worked in industry as a programmer and software system designer for 20 years, most recently with Korg R&D in Milpitas, CA. He is currently pursuing his Masters degree.

Gabriel Hugh Elkaim received the B.S. degree in Mechanical/Aerospace Engineering from Princeton University, Princeton, NJ, in 1990, the M.S. and Ph.D. Degrees from Stanford University, Stanford, CA, in Aeronautics and Astronautics, in 1995 and 2002 respectively. In 2003, he joined the faculty of the Computer Engineering department, in the Jack Baskin School of Engineering, at the University of California, Santa Cruz, Santa Cruz, CA, as an Assistant Professor. His research interests include control systems, sensor fusion, GPS, system identification, and autonomous vehicle systems. His research focuses on intelligent autonomous vehicles, with an emphasis on robust guidance, navigation, and control strategies. Specifically, he has founded the Autonomous Systems Lab at UC Santa Cruz, and is currently developing an autonomous wing-sailed marine surface vehicle and off-road autonomous ground vehicles.

ABSTRACT

We explore the capabilities of an inexpensive obstacle detection system consisting of a CCD color sensor, synchronously pulsed laser fan and supporting hardware and software. The novelty is the extreme simplicity of building such a system with commodity hardware, achieving a tiny form factor, low power operation and low cost. The camera, laser fan and supporting software constitute an active sensor that is mechanically passive, relying on the motion of its host platform to probe its surroundings. In this initial investigation, we determine useable range limits as a function of ambient lighting conditions by processing individual frames collected with a prototype system. We examine failure and recovery of the sensor when direct lighting sources come into the field of view. Future work will determine the efficacy of the full design embedded in various host platforms.

Since the geometric relation between the optical sensor and each laser is fixed, we seek optimal parameters for this relationship given the relevant constraints of the chosen system.

In situations with sufficient ambient light, we speed the computation of well known computer vision techniques for object identification to yield estimates of obstacle positions within the environment by incorporating range data obtained from the laser return.

Synchronous pulsing of the laser with a short electronic shutter time on the optical sensor allows operation of the device as an ANSI Z136 class I device since the laser's active duty cycle is highly compressed. This approach renders visible wavelengths invisible to the naked eye under most conditions.

INTRODUCTION

Common to all ground vehicle robotics is the need to localize obstacles within the environment. Many sensor systems are used towards this end, such as sonar, ladar, etc. A sub-set of these sensors use visible and near-visible light to see and map the environment.

A variety of illumination based active sensors are in common use for obstacle avoidance and SLAM (simultaneous localization and mapping) applications. We briefly review trade-offs in features and prices with these systems to motivate choices made for our design.

Amplitude modulated infrared illumination paired with a band-pass filtered IR detector can be effectively used as a direction finder for nearby objects. These devices are extremely inexpensive (less than $10), but work at short ranges, typically less than 1 meter, and have extremely poor spatial resolution.

At the far end of the spectrum, LADAR's, such as the SICK LMS-200, are a sensor of choice for mobile robotics, but come with a high price tag ($25,000 and up). Under ideal conditions, ranges up to 80 meters may be attained. The actual ranges attained under normal operating conditions are less but still quite useable [1]. The field of view of a single LMS unit ranges from 90 to 180 degrees depending on the model. Also, since the range dectection process involves physically scanning a tiny region at any given moment, interference between multiple LADAR's is minimal. LADAR's like the URG-04LX are less expensive (around $2,400), but have shorter range (about 4 meters). See figure 1 for a price vs. range comparison for some these sensors.

The SR-3000 [2] and the Canesta camera [3] illuminate their environment with amplitude modulated IR light at a harmonic frequency whose phase shift on the return will indicate range modulo some value. A typical frequency is 20MHz which repeats phase for the return at intervals of approximately 7.5 meters. A distinctive feature of this technology is the presentation of ranges directly as pixel values. Drawbacks of this technology as currently implemented include severe interference when multiple cameras are used in proximity and a sensitivity to correct gain settings [4].

Multiple structured light sensors *can* be used on a single platform if care is taken to keep the laser return of each sensor out of the FOV of other sensors or account for their presence algorithmically. We will discuss some ideas below that may enable the sensor to reject laser return originating from an unknown laser source as well.

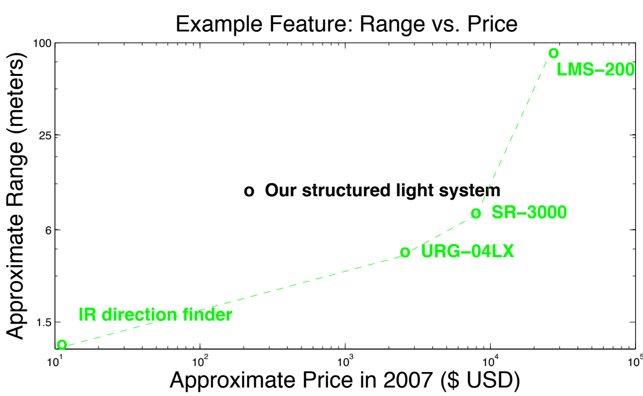

Fig. 1. The price spread for the types of sensors typically used in obstacle detection and SLAM applications

Structured light based sensors use triangulation to detect the ranges of points within their field of view. The same principle is used for stereo and multiview vision, but structured light essentially solves the so-called correspondence problem of stereo vision via the constraints induced by the structure of the light source. Once it is determined that a particular camera pixel contains primary laser return (e.g. not laser light returned from secondary reflections), the range of the reflecting surface viewed in the direction of the pixel is immediately determined to within the resolution capabilities of the system. Thus the correspondence problem is replaced with the computationally much simpler problem of determining which pixels of the sensor are 'seeing' primary laser return. Inter-pixel interpolation is used to improve the estimate of the location of laser return to a better accuracy than that imposed by strict quantization of spatial resolution to the field of view encompassed by the entire pixel.

Bandpass optical filters can be used to block ambient light in daylight conditions so that most of the light energy admitted to the sensor lies in the wavelength of the structured light source. All such filters have the drawback that they attenuate the passband as well as the out-of-band wavelengths. Commonly available [5] red (6̃60-630nm) filters pass around 90% of the desired light at most entrance angles. This filtering approach has previously been used to good effect [6].

Such an approach is best used with a black and white sensor, since color sensors already have color filters in place over every single pixel of the sensor itself. A typical color sensor has these filters arranged in the so-called 'Bayer pattern', seen in figure 2. Since three out of four pixels in this arrangement are filtering out red light, the sensor will clearly be less sensitive to laser return. On the other hand, when there is abundant ambient light, a color sensor can give more information about the scene that may be useful for object or environment classification problems.

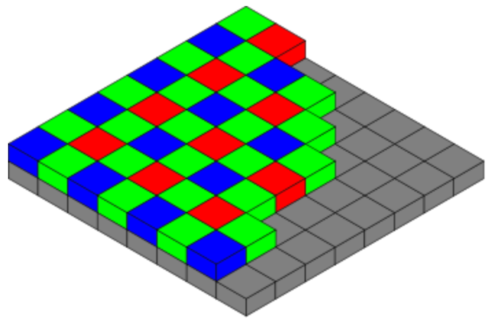

Fig. 2. The Bayer color filter arrangment over pixels in a color CCD or CMOS sensor (courtesy wikipedia.org).

A problem with structured light that does not exist with the other technologies mentioned is eye safety. In many instances, this problem can be avoided (and power saved) by pulsing the laser synchronously with the frame integration performed by the camera sensor, which is a technique others have used to solve the same problem [6].

Given all these considerations, we chose to start by investigating a flexible system that essentially changes modes as ambient lighting conditions vary. We use an un-filtered color camera in order to exploit the camera's full imaging capabilities when adequate ambient light is available, but modify the camera parameters to support the structured light mode when ambient light wanes. Much of the complexity of this initial investigation centers around appropriate ways to modify the gain, shutter and color balance controls of the camera to effectively detect laser return to maximum spatial resolution without sacrificing the color specificity available with a color camera.

Clearly the color sensor will not detect laser return under as wide a range of conditions as a filtered black and white sensor. It remains to be seen if the added flexibility of our hybrid system can be made to compensate for this weakness.

SINGLE FRAME PROCESSING FOR OUR SYSTEM

CALIBRATION

Camera calibration is a well understood area and there are many methods available to calculate both the so-called *intrinsic* parameters of the camera, which model distortions internal to a single camera, as well as the *extrinsic* parameters, which establish the actual geometric relationship between two cameras, or in this case, between the camera and the structured light source. See for example [7] or [8]. The methods described are currently in widespread use and very well understood. While it is possible to utilize a non-calibrated system for obstacle detection, doing so prevents accurate measurements for mapping applications. Design is based on idealized extrinsic parameters, however, the physical implementation deviates from this idealization. Calibration gives us the opportunity to correct this, either via software, physical adjustment or both. Since software is typically a sufficient correction mechanism, the main design concern is to make the mechanical connection between the laser and camera sufficiently robust that system measurements remain useable between calibrations.

OPTIMAL GEOMETRY

Designing the optimal geometry between the camera and structured light source is highly dependent on the application and subject to a large number of constraints of varying critcality. An overly formal approach falsely conveys a sense of complete understanding only to discover some crucial feature of the environment, (e.g. dust above the roadway) makes the design fail. Nevertheless, a formal development of the problem extracts the theoretical maximal utility from inexpensive sensors. Questions such as, “Which pixels on the sensor will ever be able to detect primary laser return with a given geometry?”, have clear answers, which we show below.

A simple example preliminary design is presented in figure 3 for an AGV application. The first constraint is low object detection at full range to maximize collision avoidance time. Under bright lighting conditions, however, we might miss objects at full range where laser return is generally weakest. Thus we place the laser close enough to the ground so that it is able to reflect off of objects of some minimum height (H_{min}) over a range of distances, not just at maximum range. Finally, the swath of view must be at least the width of the vehicle, or objects of concern to the left and right will remain undetected, resulting in an ensuing collision. A more exacting analysis of swath can be made for a particular vehicle by examining its path planning algorithms and handling characteristics.

Define the following variables (the first four being nomininal requirements).

- L_{min} - Range to vehicle width sensing swath
- L_{max} - Range to nominal ground plane sensing
- H_{min} - Minimum height detectable at L_{min}
- W - Swath width at L_{min} ($\geq$ vehicle width)
- H_l - Height of laser (assume range $= 0$)
- H_c - Height of camera (assume range $= 0$)
- β - Inclination of camera principal axis down from horizontal

H_{min},L_{min} and L_{max} taken together completely constrain the height H_l of the laser above the ground.

$$H_l = L_{max} * H_{min}/(L_{max} - L_{min}) \quad (1)$$

The system may be overconstrained since we also require that the swath of the laser at a ground range of L_{min} be at least W. The required angle Γ is equation 2.

$$\Gamma = 2\ atan(W/2\sqrt{H_l^2 + L_{min}^2}) \quad (2)$$

At this point, if a wider laser fan is needed to meet specifications, the designer is faced with the choice of either purchasing a different laser device or relaxing the specifications so that the original device is useable (e.g. by increasing L_{min} so that the swath does infact cover W at that range). Ideally the constraints objectify these choices completely.

Once the laser is positioned, H_c and β must be determined for the camera. Assume that the principal axis of the camera is aligned with the bore sight axis of the laser when viewed from above, the camera has no horizontal roll, and is positioned directly above the laser at zero range. (For a general design, an arbitrary translation and rotation in 3-space can describe the relation between the camera and light source, commonly expressed using the *essential* matrix E [7]).

The minimum height for H_c then depends on the horizontal field of view (FOV_h) of the camera so that the swath of W is achieved at the ground range of L_{min}. The constraint is equation 3.

$$H_c \geq \sqrt{\frac{W^2}{4}\cot(\frac{FOV_h}{2}) - {L_{min}}^2} \quad (3)$$

To determine possible ranges for β, we use the vertical field of view of the camera (FOV_v). The constraints we need to satisfy are equations 4 and 5.

$$H_c\ \cot(\beta + FOV_v/2) \leq L_{min} \quad (4)$$

$$H_c\ \cot(\beta - FOV_v/2) \geq L_{max} \quad (5)$$

If $\beta < FOV_v/2$, then we can ignore the second constraint since the top of the camera will be looking above the horizon.

Notice that typically there is an acceptable range of the final parameter β. Further criterion may constrain β to an exact value.

A very general consideration is effective use of pixels in the sensor. The following general procedure can be used to outline the boundary of pixels that are capable of seeing primary laser return.

Using a pinhole camera model, referring to figure 4, let the point C represents the focal point of the camera. The four arrows leaving C represent the directions associated with the corner boundaries of the field of view of the sensor. (Note that

In the case that an *end point* of I lies inside the sensor boundary, (say point 3), then one of the projections of the laser return boundary meets it. In this case, it would be the segment joining points 1 and 3.

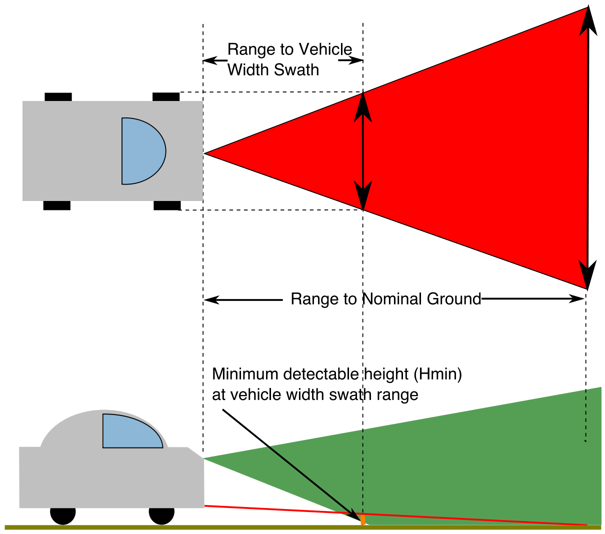

Fig. 3. AGV application where primary constraints are the need to have the camera FOV cover at least the vehicle width from the minimum range where the laser return is capable of doing the same, out to the maximum range. (Red shows laser fan coverage. Green shows vertical camera FOV.)

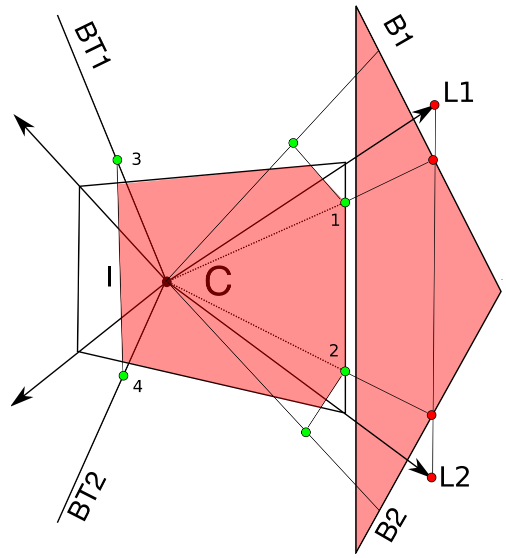

Fig. 4. Calculating the boundary for return on the sensor. The green points lie in the sensor plane. The red points lie in the laser plane.

all green points lie in the image plane and all red points lie in the plane of the laser fan).

A basic idea behind the following procedure is that projections preserve lines. For example, any two points on the ray $B1$ projected onto the image plane through C will define the line containing the image of $B1$ in the image plane.

Start with the direction leaving the camera focal point and passing through the upper right corner of the image plane. This direction intersects the laser plane at $L1$. (The corresponding ray for the lower right corner intersects the laser plane at $L2$). Connecting $L1$ and $L2$, we find where this line segment intersects the boundaries of laser illumination, $B1$ and $B2$. Since this line segment lies in the plane at the right edge of the sensor, we reproject these boundary points back through C and find 2 points (labeled 1 and 2) at the edge of the sensor delimiting the areas of the sensor that can and cannot receive laser return

To complete these boundaries, pick any point along the boundary $B1$ to reproject back through C. We take the intersection point of this reprojection and connect it to the boundary point along the sensor edge. (Note that we show the case where this boundary line will again intersect the edge of the sensor. Another case is mentioned below.)

The other important intercepts involve directions parallel to the laser plane. $BT1$ and $BT2$ are translates of $B1$ and $B2$ respectively. We find the two points where these rays intersect the image plane and connect them. The resulting line segment, I, is the vanishing line of laser light at infinity. If I lies inside the image area, it is another boundary for laser return in the sensor. No return will come back from infinity, so this is an 'outter' bound in practice.

EXPERIMENTAL SETUP

To conduct experiments, we built gimbals to support camera positioning around a good approximation of the no-parallax point [9] of a color webcam (figure 5, top image), an inexpensive 650nm laser fan (< 5mW), and a 366MHz-PII laptop running linux 2.6.x configured with video for linux (v4l) support and the Luc Saillard's pwc camera driver.

Estimates for the field of view on the QC Pro 4000 were obtained from the datasheets [10] for the SONY ICX098BQ sensor used in the camera. With the reported $5.6\mu m$ square pixels, the $6mm$ microlens system used, and the VGA pixel use of the sensor, a pinhole camera model gives estimates of $FOV_h \approx 35.3^o$ and $FOV_v \approx 26.1^o$.

As a test of concept for synchronous pulsing of the laser with the camera, we built a small a current switching circuit based on a MOSFET transistor and a microcontroller breakout board development kit. The small blue wire in the third image of figure 5 brings the shutter signal from the camera off to interrupt pins on the microcontroller to provide the logical signal for switching the laser.

C user-space applications were written to implement the image capture and camera parameter control experiments.

Auto white balance is disabled for all experiments. The results of the onboard color de-mosaicing algorithm in the camera were used for these experiments, so some de-mosaicing artifacts are present in our results. This is typical for a Bayer pattern based sensor. The test camera was only able to provide a raw Bayer image in a proprietary lossy compressed format, so this preferred approach was unavailable.

This setup has processing power comparable to an envisioned embedded system based on a low-power tiny form-factor gumstix computer [11].

GAIN AND SHUTTER CONTROL

The laser return is most useful in low ambient light conditions. With some simplifying assumptions, we can derive a heuristic similar to the well known ‘radar equation’ to put a bound on the amount of returned energy we can measure in a local area of the sensor.

Approximate the distances from the laser to the target and from the target back to the sensor with the same value r. The dropoff in irradiance from a laser fan structure is proportional to $1/r$ due to its fanning angle. (A laser pointer would have essentially no dropoff by this approximation).

Considerable variation in return strength occurs due to the wide range of reflective properties of objects in our environment and the incident angles of irradiance and observation (modeled by the $BRDF$ formalism [12]). We make the simplifying assumption that a typical object will have roughly Lambertian reflectance properties, and thus distribute radiated light fairly uniformly through some solid angle. With this assumption, we can expect a dropoff in return strength proportional to $1/r^2$.

The result is a total dropoff in strength proportional to $1/r^3$ from source to camera. The power of the laser in our experimental setup $< 5mW$, so it is easy to see why bright daylight makes return detection problematic.

Typical situations of low or no ambient light in urban environments and roadways involve many artificial light sources, such as headlights, street lamps, etc. For the sensor to be useful in this type of situation, we need an algorithm that can adjust the shutter time and gain of the camera to pick up any area of the camera image that might have laser return and ignore those areas of localized high light intensity.

Localized ‘excess’ light can in many cases be safely ignored since modern CCD sensors have good anti-blooming characteristics and can simply drain off excess electrons from a pixel cell into the sensor substrate without affecting adjacent pixels. Glare is the limiting factor, since the light collected for the sensor must pass through the lens system of the camera.

Since gain works before AD conversion to digital pixel values in the camera, it can be used to reduce quantization noise if the dynamic range of the signal is small, such as occurs in low light imaging situations. This does not improve the basic SNR of the sensor.

The two ways to improve SNR in the image are increased shutter time and using a larger aperature and/or lens system to capture more light in a given amount of time. Auto-iris lensing systems do exist, but they add complexity, cost and size to the system.

Thus we opt to use increased shutter time as our primary means of increasing system SNR.

Following work done by Nourani-Vatani and Roberts [13] to extract information from images that have great variations in brightness, we are working with image segmentation techniques that essentially block subsets of the image from contributing to the calculation of gain and shutter values. The most useful approach will determine and apply masks dynamically based on the previous location of detected laser return. In initial tests, hand coded masks are being used.

CONCLUSION

We have implemented experimental setups which indicate the feasibility of developing low-cost structured light sensors which feature a cost to detection ratio competitive with other technologies.

Objective functions for the optimization of camera to light-source geometries have been developed for structured light sensing systems.

Digital cameras are complex sensing devices, color cameras especially so. The default algorithms to manage the rich parameter space of color cameras are geared towards producing images pleasing to consumers, not for preserving the fidelity of information useful for computer vision. However, by controlling these parameters with our own applications, we are able to leverage commodity hardware into effective devices.

Techniques to dynamically segment image areas in order to do localized control of gain and shutter parameters have been shown effective in test situations.

FUTURE WORK

As we refine the process of dynamically segmenting brightness regions in the image, we intend to use a higher framerate with oscillating gain/shutter values coordinated with changes in the segmentation mask. This is done in order to see both the area of the image with laser return and areas with extremely bright ambient illumination.

An LED laser with a structured light lens (such as a fan) is non-ideal in terms of uniform beam width, beam divergence, and distribution of light intensity. This deviation does not typically change during operation, thus a promising approach is to ‘finger-print’ the actual light structure of a given unit during a calibration phase and use this finger print to improve signal to noise.

The sensor is intended to be used on a moving platform. This investigation has only addressed issues associated with single frame processing, treating an increase in shutter time as a parameter that is only constrained by ambient lighting conditions. An increase in exposure time directly conflicts with the desire to have an image free of motion blur when the sensor platform and/or objects in the field of view are in motion. Motion blur is usually modeled as a linear shift-invariant system where the convolution kernel is a literal representation of the timed motion of the camera during image capture projected onto the image plane. This makes use of data from an IMU or similar motion sensor already in use an interesting possibility. A low cost solution could be explored using MEMS sensors for short time motion extrapolation which could directly record a convolution kernel for deblurring after image capture. Deconvolution is complicated by the fact that camera rotations and translations have different effects on different parts of the image depending on the corresponding scene's actual distance from the camera, but if laser return is present, we know ahead of time what the ranges are for each part of the image. Conversely, if that portion of the image is not recording the scene at the assumed range, then it should not contain laser return, and reduced fidelity in this region does not significantly alter the system performance. Another possibility is using a point structured laser source, such as a simple laser pointer, to create a visual record of the motion which can then be used to estimate a kernel.

We envision future tests with the sensor using an ambulatory platform as well as tests on an AGV. A more compact experimental setup will be required for ambulatory tests, most likely incorporating a gumstix or similar computing platform. For the AGV, more robust mechanical attachments will be required.

Finally, we plan to explore a system constructed around a black and white sensor using bandpass filtering and compare it to the hybrid color imaging system explored here. We will acquire a color camera capable of delivering full-bandwidth raw image data from the sensor to help avoid color mosaicing artifacts. It will be useful to characterize how much these artifacts do or do not affect system metrics, since camera systems with color processing and mosaicing effects are ubiquitous.

ACKNOWLEDGEMENTS

This work is funded by NSF proposal 0529435, "Exploring the World with a Ray of Light: An Environmental Sensor for the Blind".

REFERENCES

[1] W. C. Stone, M. Juberts, N. Dagalakis, J. Stone, and J. Gorman, "Performance analysis of next-generation ladar for manufacturing, construction, and mobility. nistir 7117," 2004.
[2] "Mesa imaging, ag." [Online]. Available: http://www.mesa-imaging.ch
[3] "Canesta camera." [Online]. Available: http://www.canesta.com
[4] P. Einramhof, S. Olufs, and M. Vincze, "Experimental evaluation of state of the art 3d-sensors for mobile robot navigation," in *International Conference on Robotics and Automation*, 2001.
[5] "Edmundscientific." [Online]. Available: http://www.edmundscientific.com
[6] C. Mertz, J. Kozar, J. Miller, and C. Thorpe, "Eye-safe laser line striper for outside use," in *IEEE Intelligent Vehicle Symposium*, vol. 2, June 2002, pp. 507–512.
[7] O. Faugeras and Q.-T. Luong, *The Geometry of Multiple Images*. Cambridge, MA, USA: The MIT Press, 2004.
[8] J.-Y. Bouguet, "Camera calibration toolbox for matlab." [Online]. Available: http://www.vision.caltech.edu/bouguetj
[9] R. Littlefield, "Theory of the "no-parallax" point in panorama photography," 2006.
[10] "Sony datasheets." [Online]. Available: http://www.sony.co.jp/~semicon/english/90203.html
[11] "Gumstix computers." [Online]. Available: http://www.gumstix.com
[12] B. K. P. Horn, *Robot Vision*. Cambridge, MA, USA: The MIT Press, 1986.
[13] N. Nourani-Vatani and J. Roberts, "Automatic exposure control," navid Nourani-Vatani with Tasmanian ICT Centre, Jonathan Roberts with Autonomous Systems Lab CSIRO ICT Centre, Australia.

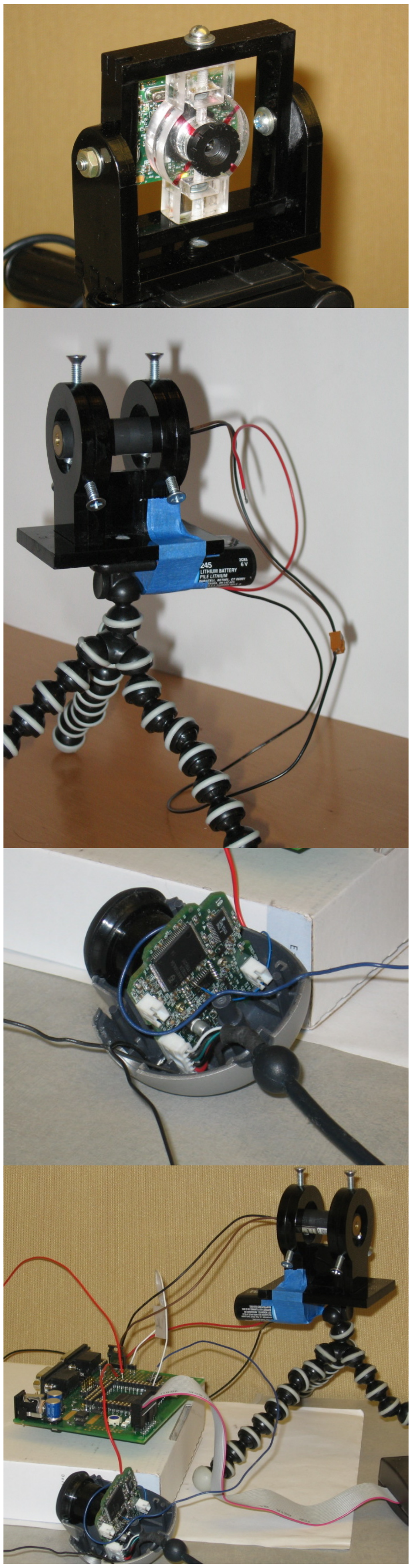

Fig. 5. Experimental setup. The top image shows the camera gimbal mount built to provide arbitrary camera positioning about the no-parallax point of a Logitech QC Pro 4000. The next image shows a portable test mount for the laser fan when used without the synchronous pulsing driver. The third image is a closer look at the camera. The smaller blue wire leaving the chassis to the right carries the shutter information to the microcontroller board. The bottom image shows all the components of the synchronous laser pulsing test circuit. The battery is unused in this configuration.

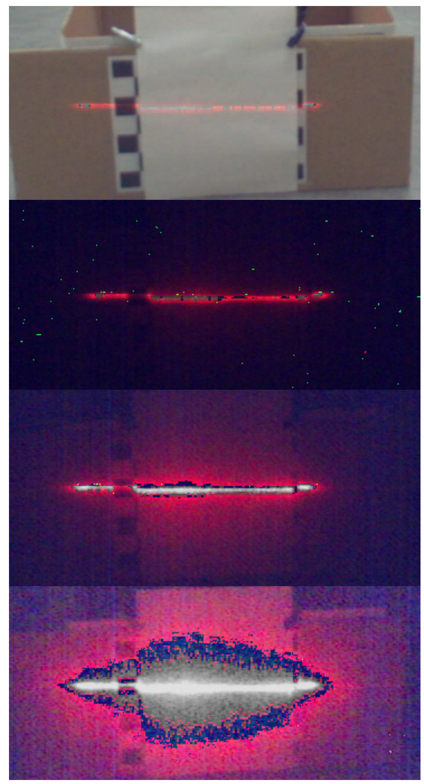

Fig. 6. The inadequacy of the built in auto-gain/shutter control algorithm (AGC). In the top image, (except for the color mosaicing artifacts) we have a good laser return. The hue of the laser is clearly discernable. In the next image, ambient lighting has been removed, and the hue and spatial definition of the laser return is still good. The third image shows the effects of AGC about 1/2 second later. The laser return is completely saturated and no longer discernable as a red light reflection. In the bottom image the AGC has stabilized after about 1 second. Here spatial resolution of the return is ruined as secondary diffuse reflection becomes saturated as well.