Low Cost, Low Power 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

ABSTRACT

We evaluate the capabilities of an inexpensive obstacle detection system consisting of a CCD or CMOS optical sensor, synchronously pulsed laser and supporting hardware and software. The goal is to expand the range of feasible autonomous vehicle applications to include those that are currently impractical due to limitations on the price, weight, or power requirements of their sensor suites. This system constitutes an active, mechanically passive sensor, relying on the mechanical activity of its host platform to sweep out samples from its surroundings.

We evaluate sensor configurations in two example host platform designs. The first is a handheld obstacle detector to aid users with vision impairment, while the second is a short range detector used as part of the sensor ensemble for an autonomous ground vehicle.

Tradeoffs for both continuous laser fan and single laser pointer configurations are evaluated. Since the geometric relation between the optical sensor and laser is fixed, we establish effective distance and angle between the laser and sensor given required minimum and maximum ranges, spatial resolution, platform velocity and expected velocities of potential obstacles.

In situations with sufficient ambient light, range data from the laser return is used to speed the computation of well known computer vision techniques for object detection to yield estimates of obstacle positions within the environment.

Pulsing the laser synchronously with a short shutter time on the camera allows operation of the device as an ANSI Z 136 class 1 device since the laser's active duty cycle is highly compressed. This approach renders visible wavelengths effectively invisible to the naked eye.

INTRODUCTION

All autonomous vehicles require some means of obstacle detection for collision avoidance. Many sensor modalities exist to support obstacle detection, including TOF ladars, phase shift ranging cameras, stereo vision systems, sonar, radar and even simple photodiodes used as direction finders for reflected infrared or other illumination [1], [2].

Photodiode sensors standout as being cheap, light weight, and low power, yet they deliver low resolution information, are short range, and thus not sufficient for many navigation tasks required for autonomous navigation.

Structured light based sensors, on the other hand, perform competitively in terms of resolution, range (Fig. 1), and sampling bandwidth while providing the opportunity to realize a low cost, low power, and low weight sensor.

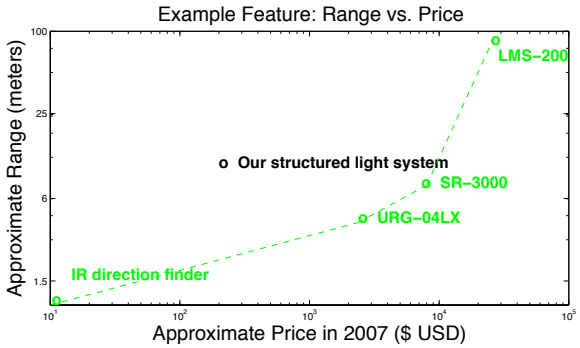

Fig. 1. Price vs. Range comparison for some widely used obstacle detection sensors. Prices current as of early 2008.

Component cost is a barrier for any price sensitive application, whether it is intended for mass consumer markets or targeted to an application where the perceived value does not support a large expenditure. Weight and power constraints are critical for small, close to ground airborne and hovering platforms, hand-held devices, or any battery operated device.

Thus a structured light sensor which is low power, low weight, and low cost provides the opportunity to significantly increase the range of applications where such sensors may be employed — and even create new application areas that are not otherwise possible.

In this paper, we examine the design, calibration and testing of two example structured light sensor applications fitting into this niche. Low price is achieved by using an inexpensive CMOS sensor, LED laser, and a commodity micro-controller. The lack of mechanical actuation, in addition to the use of low-cost commodity parts, ensures low-power and low-weight as well.

For the prototypes considered here, we use devices with comparable sensor performance to evaluate the designs. We use a black and white XVGA (1024×768) camera with a horizontal FOV of about 41 degrees and vertical FOV of about 31 degrees. At full resolution, it is capable of operating at a

frame-rate of 15-fps.

For the two applications discussed, we examine how specialization of the sensor's role in the system motivates positioning of the sensor, the choice of geometry between the camera and laser, and allows for simpler algorithms that support the goal of obstacle detection while remaining within the computational bounds of the sensing system.

HAND HELD RANGING AND OBSTACLE DETECTION

The hand held device considered here is being developed as an assistive technology for vision impaired users. It functions as a detector for hazards such as curbs, drop-offs, stairs, tables, chairs, and similar obstacles to pedestrian travel. The device is intended to be used in a proprioceptive manner, in much the same way as a white cane (albeit one with a much longer reach). The user registers information from each direction into a coherent whole, estimating the device's pose by their innate sense of arm and hand position.

Design Requirements and Resulting Geometry Choices

Precise mapping of the environment is not required, but the device must provide immediate range information as the user points the device and be able to resolve the smallest obstacle that could present a hazard to the user.

The device is held at a nominal height of 1-m (about waist level), with the laser illumination plane having its shortest range of detection when pointed down. Varying range points are attained as the user positions the device. The nominal operation mode has the bore axis of the laser inclined 45-degrees to ground normal (Fig. 2) while the user sweeps the device in a radial fashion.

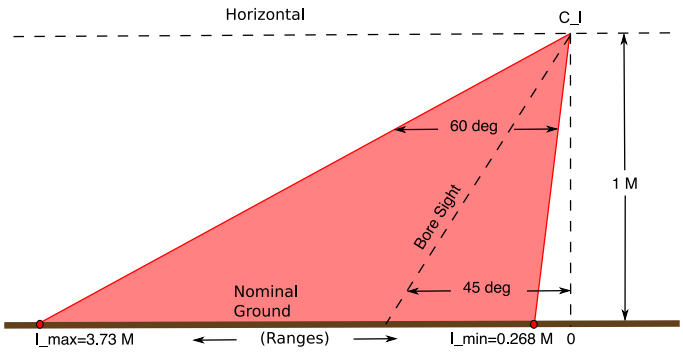

Fig. 2. Ranging of a 60 degree laser fan held with the structured light plane normal to the ground. C_l represents the nominal origin of the laser where the boundaries of fan illumination meet.

The design must meet the following requirements:

- L_{min} (50-cm) Closest detectable ground level object
- L_{max} (3-m) Furthest detectable ground level object
- H_{min} (1-cm) Minimum height resolvable at L_{max}

Variations in the choice of vergence angle and baseline between the camera and structured light source will affect the values of L_{min} and L_{max} achieved. One way to visualize the effects of varying these parameters is indicated in Fig. 3 (see the caption for a detailed description).

We want to achieve the specified L_{min} and L_{max}, stay within a comfortable margin of the useable parameter space incase of small implementation errors, and also use a small baseline to achieve a compact device. These considerations motivate use of a design point near $\beta = 20$-deg, $B = 500$-mm.

The resolving power of the system is ultimately limited by the sensor and lens of the camera. Since we are considering a structured light system, measurement accuracy is limited by the volume of space where the field of view of a particular pixel intersects the structured light source.

In Fig. 4 we represent each such volume of space by the *point* in the nominal structured light plane that joins the optical ray passing through the center of the pixel that views it from the sensor.

To measure the resolution ability of the system at a pixel, we adopt the very conservative method of computing the maximum difference in coordinates of the eight points on the laser plane corresponding to the eight pixel neighborhood surrounding the pixel on the sensor.

In Fig. 4, we use the coordinate of vertical height. Inspection of this figure confirms that we satisfy the H_{min} requirement using the design parameters indicated. Rich information about the resolving power of various pixels on the sensor is readable from this plot.

SIMPLE AV SENSOR FOR SMALL OBSTACLE DETECTION

For an autonomous vehicle, we start with the basic requirements of a forward obstacle detection sensor which returns angle ranges from left to right classified as clear, obstacle, or unknown.

AV Sensor Design Goals

We design a forward obstacle detection sensor to meet the following requirements:

Travel at a maximum forward speed of 40 km/h, being able to detect an obstacle over at least 5 frames, where the obstacle may be moving at up to this maximum speed as well. (Notice that we are only accounting for moving obstacles that are roughly in the FOV of the camera). Almost a meter will be traveled at this speed in a duration of 5 frames. We want to leave a margin of error for events such as missed detection at maximum range and to allow sufficient time to evade an obstacle or stop. With these considerations, a conservative set of specifications are:

L_{min}	1-m	Range to closest detectable ground level object
L_{max}	10-m	Range to furthest nominal ground plane object
H_{min}	5-cm	Minimum height detectable at L_{min}
W	45-cm	Swath width at L_{min} ($\geq$ vehicle width)

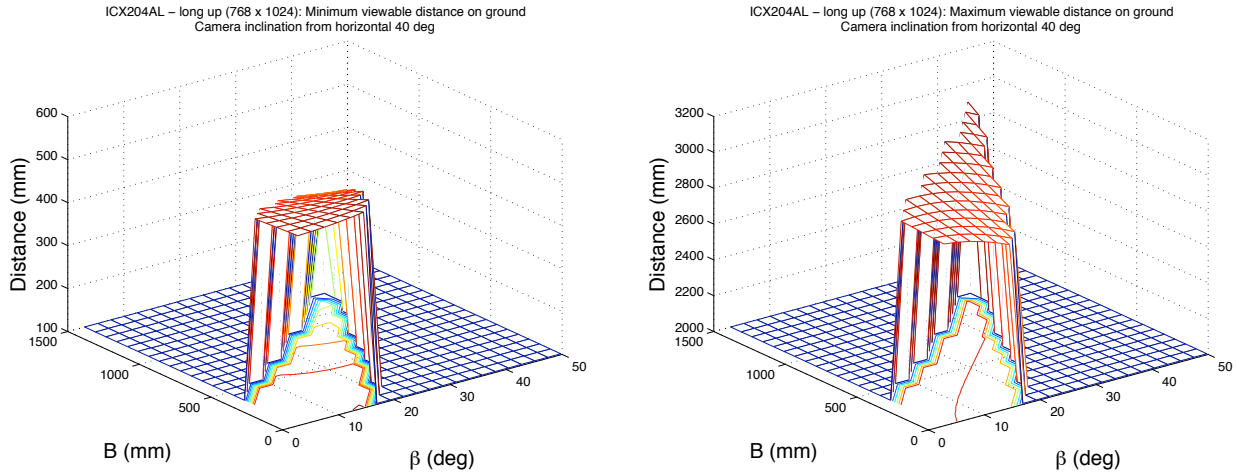

Fig. 3. Minimum and maximum range obtainable at nominal ground level as a function of baseline and vergence angle. For this figure, the camera's principal axis is first tilted 40 degrees below the horizon. The laser's bore sight axis is tilted 45 degrees below the horizon. The vergence angle is then articulated at the camera using the rotation axis normal to the baseline and camera principal axis. The laser plane itself is normal to the ground for these measurements. The drop-off to the flat portion of the plot indicates that the image of the measured range point has reached the corner of the sensor. The constant value in the flat portion of the plot represents unusable parameter choices.

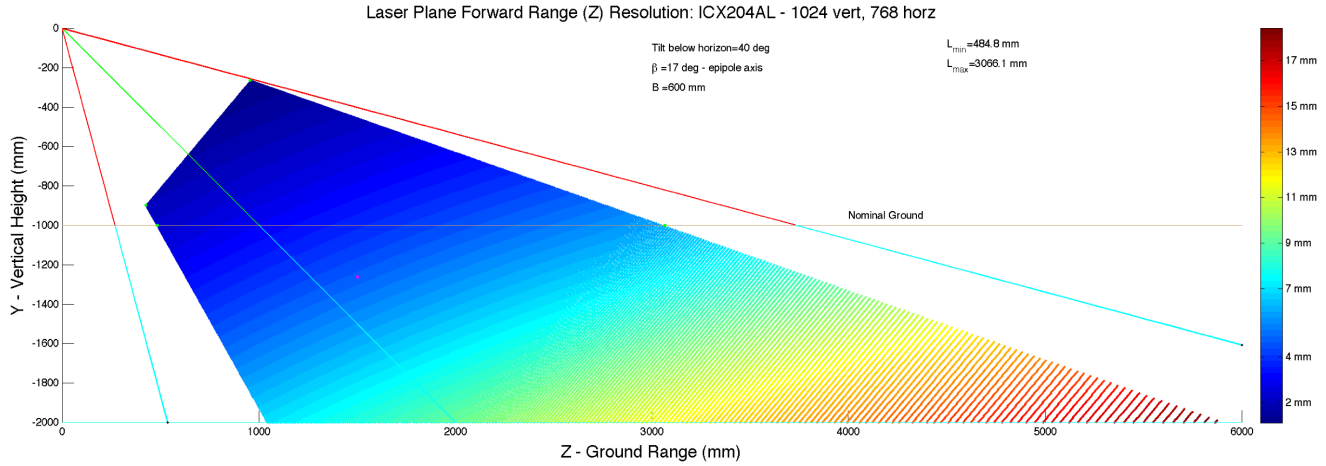

Fig. 4. Conservative calculation for the resolving power of a proposed design. Here we model the worst case error in vertical measurement for a pixel as the the difference between the largest and smallest height measurement in the eight pixel neighborhood around the pixel. The measurement for each pixel is shown color coded, at the location where the optical ray from the center of the pixel intersects the plane of laser illumination. The horizontal line at H =-1000-mm represents the nominal ground plane 1-m below the hand held camera's optical center.

From these, we'll compute:

- H_l Height of laser
- θ_l Inclination of laser plane axis down from horizontal
- H_c Height of camera
- θ_c Inclination of principal axis down from horizontal

If we want to position the camera so that the vertical FOV extends from L_{min} to L_{max}, the choice that places the boundary of the camera view at those locations satisfies Eq. 1 and Eq. 2.

$$H_c = L_{min} \cot(\theta_c - FOV/2) \quad (1)$$

$$H_c = L_{max} \cot(\theta_c + FOV/2) \quad (2)$$

After various manipulations, we arrive at $\theta_c = 38.25$-degrees with a resulting $H_c = 83.8$-cm.

We proceed to compute H_l and θ_l.

For this design we use a low-cost 650-nm red laser diode, with a power rating of < 5-mW, a measured fanning angle α of approximately 23-degrees, and a divergence < 2-mrad.

The requirement for W is satisfied by constraining H_l as in Eq. 3, yielding $H_l \geq 21$-cm.

$$H_l \geq \sqrt{\frac{W}{2} \cot(\alpha/2)^2 - L_{min}^2} \quad (3)$$

Note that if we need to satisfy the requirement of detecting

a 5-cm object at 1-m, while the nominal ground range is 10-m, we already have a requirement for H_l to be about 5.5-cm using similar triangles.

The solution in this case is to realize that the swath requirement is fairly sensitive to L_{min} and we can easily accommodate extending L_{min} a few centimeters. In fact, if L_{min} satisfies Eq. 4, W is wide enough for any choice of H_l.

$$L_{min} \geq \frac{W}{2} cot(\alpha/2) \tag{4}$$

Let us round up $L_{min} = 120$-cm. This results in $H_l = 5.68$-cm and $\theta_l = \arctan(5.68/1000) \approx 1/3$-degree.

Note that we do not need to modify the position of the camera, since it already has the new L_{min} within its field of view.

This design results in a laser that is essentially looking straight ahead, low to the ground. Consider that a small change, setting θ_l to 1 degree will result in a ground range of only about 3-m. Thus the maximum detection range at nominal ground for this design is extremely sensitive to errors that could reasonably be expected to occur during operation.

Yet this design (Fig. 5) is very effective at detecting small objects in front of the vehicle as it travels on fairly flat ground. Weaknesses are poor hole characterization and ambiguity about the height of objects taller than H_{min}. A second laser placed at a higher position can be used to compensate for these shortcomings.

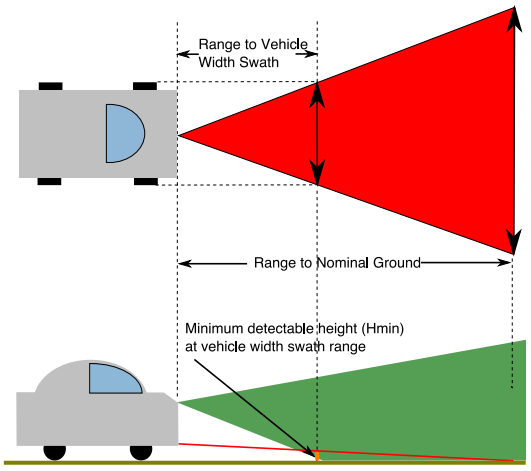

Fig. 5. The laser and camera positions for the small height obstacle detector. Orange represents the laser swath in the top image and the laser plane seen on-edge in the bottom image. The green area represents a possible pose for the camera vertical field of view.

PROTOTYPES

Examples of prototypes are shown in Fig. 6. The prototypes were constructed using the camera mentioned in the introduction and using the laser diode module mentioned in the "AV Sensor Design Goals" above.

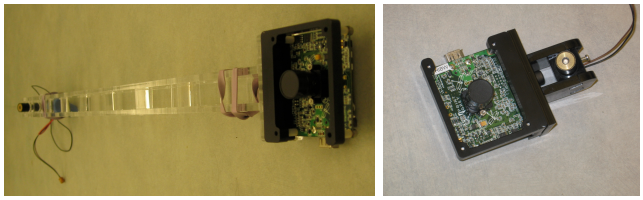

Fig. 6. Prototypes. A 60-cm baseline (left) and 6-cm baseline prototype (right). Both have a 17 degree vergence angle articulated at the laser.

CALIBRATION

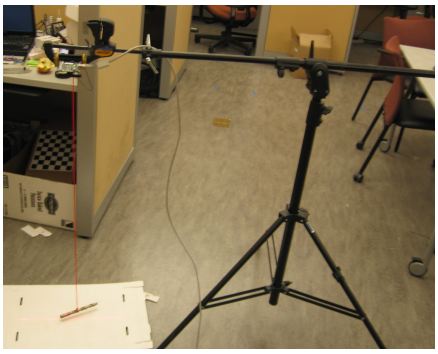

Fig. 7. A simple method of setting the rotation of the laser in its holder for the prototype. The device is held 1-m from the ground with the laser normal to the ground. The laser is adjusted until its image on the ground in the camera roughly matches its design value.

While the obstacle detection type applications do not require precise calibration in order to function, improved accuracy and capability can be obtained with such precision. Applications such as mapping benefit from precise calibration, which can be performed easily and conveniently.

Fig. 7 shows the basic setup for mechanical calibration. Note that only simple tools, spirit levels and plumb bobs are required for mechanical calibration. Greatly improved accuracy is obtained by extending the mechanical calibration with a software calibration scheme, described below.

We use the Camera Calibration Toolbox [3], which is built around Zhang's calibration method [4]. Experimentation has demonstrated that this method is insensitive to the presence of laser return within the image of the calibration object.

We have developed software that allows the user to select points along the laser line projected onto the calibration pattern (Fig. 9). We stipulate that only points lying on the checkerboard pattern plane be selected. This allows us to re-project the 3D position of the selected image point, using the camera calibration results and the knowledge that the selected point lies on the plane of the calibration object.

Once the user has entered all the calibration points sampling the laser return, we estimate the structured light plane parameterized by $\hat{L}_p$ (a point on the plane), and $\hat{L}_n$ (the plane normal).

Let A be the $N \times 3$ matrix of re-projected 3D points. Let A_i

represent row i of A (one of the points). Let $\mathbf{1}$ be the $N \times 1$ matrix whose entries are 1's. The estimate $\hat{L}_p$ is simply the centroid of the 3D points (Eq. 5). The estimate for $\hat{L}_n$ is computed using least squares (Eq. 6).

$$\hat{L}_p = \frac{1}{N} \sum_{i=1}^{N} A_i \tag{5}$$

$$\hat{L}_n = (A^T A)^{-1} A^T \mathbf{1} \tag{6}$$

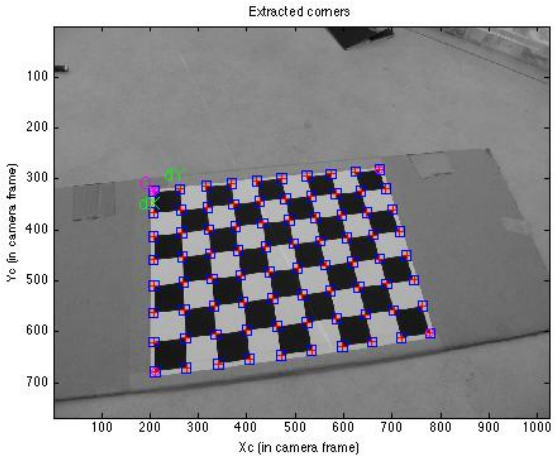

Fig. 8. This is the result of the corner finding procedure in the calibration toolbox [3]. The image of the laser return has not been modified in any way. It does not interfere with the corner finding procedure.

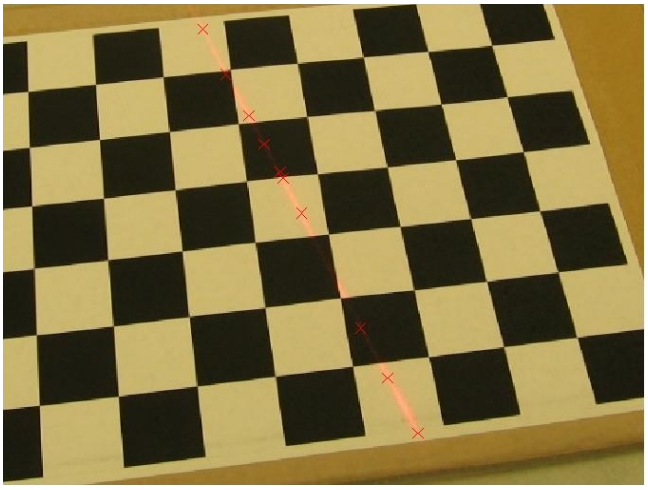

Fig. 9. Here is an example of the user marking points in the image of the structured light source on the calibration object. This is done on the same image as shown in Fig. 8.

LASER RETURN DETECTION

In order to explore the combination of laser ranging information with visual processing of the same image, we must solve the problem of distinguishing laser return from the image scene. A typical method for recognizing visual cues or features in a complex scene is to employ feature recognition algorithms such as Harris corner detection [5] or SIFT features [6]. Here the goal is to employ methods that are less computationally demanding.

Samples of laser return at known distances and shutter speeds on essentially monochromatic surfaces are collected. Fig. 10 shows a subset of these samples, here color tagged to indicate a red or green surface (the camera is black and white). These samples provide us with the largest return image we could expect at the given distance, since the surfaces are normal to the camera and laser. We can see the pixel intensity distribution at a given shutter speed, which allows us to establish a range of useful camera exposure times for a given laser *a priori*. Finally, the samples function as templates for matching in the experiments we describe next.

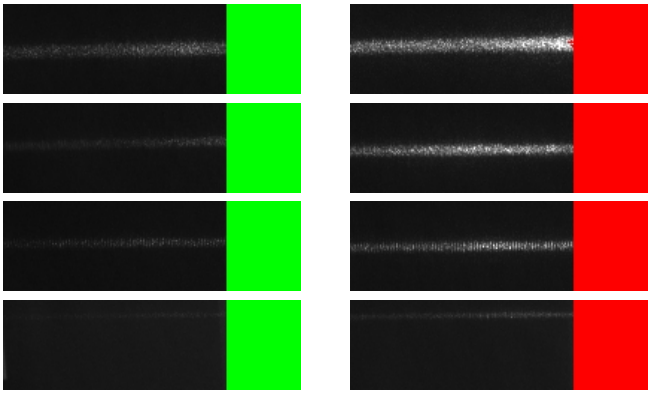

Fig. 10. Black and white images of laser return with color tags added to indicate the surface color. The images are taken at increasing ranges of 50, 75, 100, and 200 cm (top to bottom). The surfaces are oriented normal to the laser and camera.

An extremely basic matching function for a feature is normalized correlation. Let T_{rs} be a $2U+1 \times 2V+1$ template patch, such as a sub-window of our sampled laser returns. Let I_{ij} be the $M \times N$ image. Let W_{ij} represent the $2U+1 \times 2V+1$ sub-window of I centered at i, j.

Let $\mu_{W_{ij}}$ be the mean of W_{ij}, and $\sigma_{W_{ij}}$ be the standard deviation. Let μ_T be the mean of our template T, and σ_T be the standard deviation. Then Eq. 7 defines the correlation of the window centered at i, j.

$$l_{ij} = \sum_{r=-U..+U, s=-V..+V} \frac{(I_{i+r,j+s} - \mu_{W_{ij}})(T_{r,s} - \mu_T)}{\sigma_{W_{ij}} \sigma_T} \tag{7}$$

We compute $\overline{T}_{rs} = (T_{r,s} - \mu_T)/\sigma_T$ ahead of time, so that the computation becomes Eq. 8. Fig. 11 shows the result of this computation with a small sample of laser return. Notice that although the actual laser return has a high degree of correlation with the template, many false positives exist. For a number of these false positives, a large contiguous area matches, which is inconsistent with the very narrow match achieved by actual laser return. This suggests that further

work could establish a reliable criterion for recognizing the laser return, but still less computationally intensive than other methods.

$$C_{ij} = \sum_{r=-U..+U, s=-V..+V} \frac{(I_{i+r,j+s} - \mu_{W_{ij}})\overline{T}_{r,s}}{\sigma_{W_{ij}}} \quad (8)$$

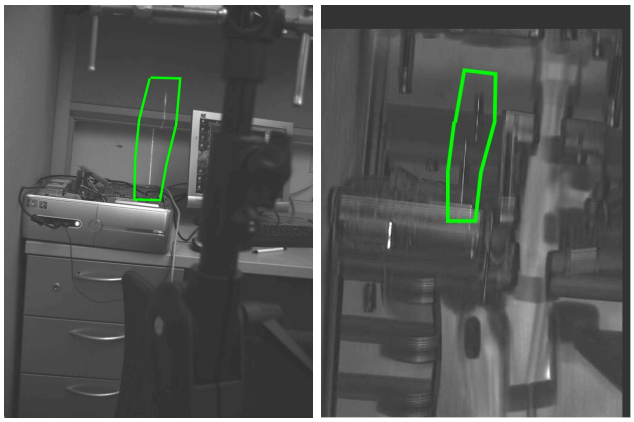

Fig. 11. The left image contains laser return (outlined in green) surrounded by clutter. The results of running normalized cross-correlation with a 21×5 pixel template of laser return are shown on the right. There is a high correlation with the actual laser return in the image. There are numerous false positives as well.

We contrast this situation with the use of a laser pointer in a visually cluttered scene. Fig. 12 shows such a scene acquired using a laser pointer for the structured light source. Fig. 13 show the same sequence with its 8-bit samples thresholded at 240. Most of the noise is masked, but even the remaining noise is easily distinguished from the actual return, since it does not lie on the epipolar line defined by the laser pointer, baseline, and vergence angle.

SIMPLE ALGORITHMS SUPPORTING THE DESIGNS

Clear path discovery for the simple AV obstacle detector

For the AV design considered in section "SIMPLE AV SENSOR FOR SMALL OBSTACLE DETECTION" above, we employ the simplest possible obstacle detection algorithm.

If laser return is detected in the safe range, we cluster each set of contiguous pixels with this property into a set. The range (Z) of the set is determined as the minimum range pixel. The angle (γ) subtended by this set is easily determined, as each column of the sensor corresponds to a direction. If the width of the swath determined by these parameters ($2Z\tan(\gamma/2)$) is large enough for the vehicle to pass through, then the set represents a clear path, and the vehicle has the option to steer through the middle.

Clearly there are cases where this algorithm will fail to detect a clear path when one exists, but it is a simple and fast algorithm that produces useable paths.

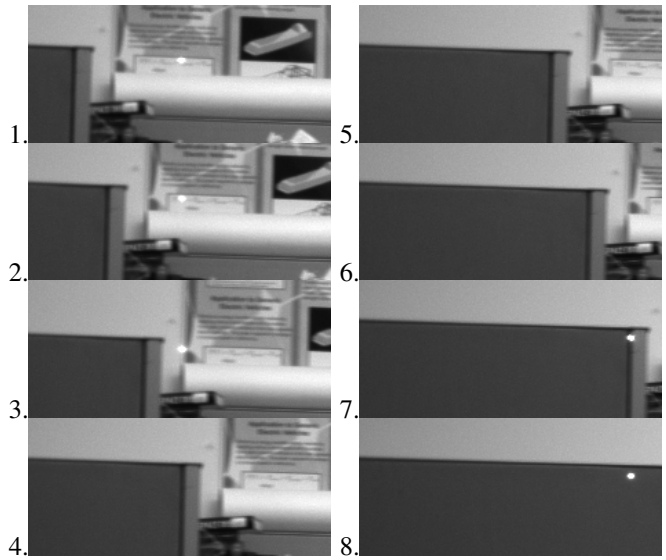

Fig. 12. A sequence of images showing the jump in laser return position that is typical as the return leaves a surface of approximately uniform range and moves to a new surface at a different range. Note the 3 frame gap (images 4-6) where the laser is occluded from the camera by the cubical wall.

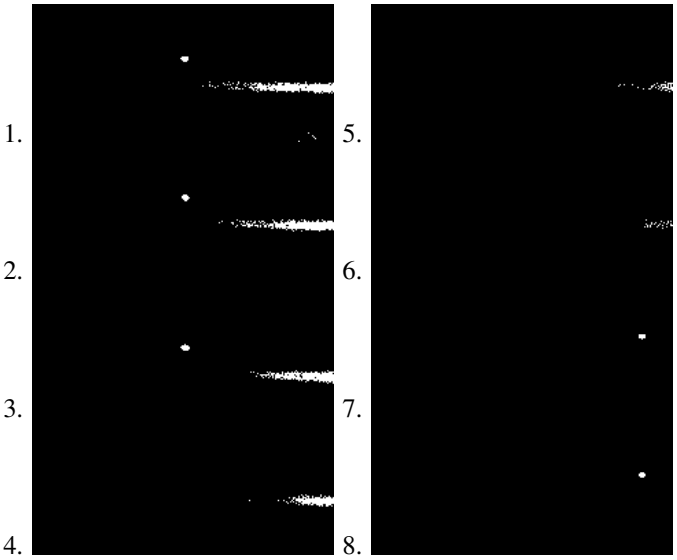

Fig. 13. Simple threshold masking. The 8-bit images are masked at a value of 240. Noise is present, but in these examples it can be ignored as potential laser return since it does not lie on the epipolar line.

Finding real edges of objects using the laser pointer

The sequence of images in Fig. 12 and the corresponding detected laser return in Fig. 13 allow the use of localized edge detection to better characterize object extents.

When the laser return disappears in the fourth image in the series, we can infer that this is due to an occlusion caused by the presence of a nearby object. Thus, if the object presents a visible edge, the edge is close-by in the image.

We run a localized edge detection routine in a window around the pixel centered on the last detected laser return.

If an edge is detected, we attempt to track it from frame to frame until the laser return reappears. The edge separates a local region of the image into two parts. We associate each part with a range measurement that appears within it, thus achieving a rough estimate of the object distance and extent close to the laser return.

ANSI Z 136 CONSIDERATIONS

In [7] synchronous pulsing of the laser to coincide with image exposure time is used to achieve a class 1 device. The laser devices considered in this paper are either class 3a or class 2 when operated *continuously*. None of our designs use continuously operating lasers, but rather pulse them synchronously with the exposure time window of the visual sensor. Here we consider when a class 1 or class 2 device is achieved with this method.

For a frame rate of 15-fps and an activation time of 10-ms, the mean radiant power is scaled by a factor of 0.15. We consider visual spectrum lasers (400-nm - 700-nm). Based on the definitions in section 3.3.3.1 of [8], this immediately transforms a class 3a device into a device with a rating no worse than class 2.

To determine when a class 1 device is achieved, the class 1 AEL is computed. Note that the fanned laser satisfies the definitions of an extended source, but the irradiance is dependent on the distance from which the laser is viewed. The laser pointer does not have this property. The non-extended source computation is the more conservative one for the lasers considered here, due to power dilution with distance. We use the more conservative non-extended computation.

In the following, we consider a 650-nm laser. The computations are similar for visual spectrum lasers at other wavelengths, with the addition of correction factors for some wavelengths. Following the nomenclature of "The American National Standard for Safe Use of Lasers" [8], the appropriate value of T_{max} is the aversion response time of 0.25-s. This is established by section 8.2.2 of the standard. (See also table 4a).

The MPE_{sp} can be read from table 5a as $1.8t^{\cdot 75} \times 10^{-3}$-Jcm^{-2}, where t is the pulse (shutter) time. This is about 5.69×10^{-5}-Jcm^{-2}.

Applying the correction $C_P = n^{-.25}$ (table 6) for multiple pulses during the aversion response time, we get a rule 3 (section 8.2.3) $MPE/Pulse$ of approximately 4.02×10^{-5}-Jcm^{-2}.

To compute the class 1 AEL, we use $D_F = 7$-mm from table 8 with an area of 0.385-cm^2 to arrive at approximately 1.55×10^{-5}-J.

The amount of energy emmited during this time by the laser is less than $5\text{-mW} \cdot 4 \cdot 0.01\text{-s} = 2 \times 10^{-4}$-J.

Based on these calculations, the device is over the class 1 limit. Thus for a pointer light structure, we must settle for a class 2 device, unless the overall frame-rate and shutter time can reduce average power by approximately a factor of ten.

For the fan light structure, we have used very conservative calculations, ignoring the dilution of energy from fanning and assuming the power of the device to be 5-mW, even though a typical device will emmit less power. For the fanned lasers, we expect with careful calculation, we are able place the devices in class 1. A required parameter to establish the extended source computation is the minimum viewing distance, which will vary with this application.

Since the shutter time is a natural parameter to vary with these applications, special care must taken to use the *maximum* duration of the individual pulses when computing MPE_{sp}.

CONCLUSION

We have demonstrated the trade-offs involved in the design space, and built two prototypes based on these designs. Calibration was achieved using simple methods and extending the software calibration toolbox [3] to include the laser fan reflected return. By using the line of laser return on the calibration target, a simple least squares solution defines the laser plane with high precision.

Calculations based on the ANSI Z 136 Standard [8] show that pulsing the laser can reduce the structured light sensor to a Class 2 or even Class 1 device. Further analysis will quantify the amount of power radiated from our sensor.

FUTURE WORK

The analysis of the simple forward obstacle detector for the autonomous vehicle application assumes an essentially level travel surface and ignores vehicle pitch. It would be useful to expand the capabilities of this system to handle both uneven terrain and vehicle pitch, as well as better mitigate the sensitivity of maximum ground range to the inclination angle of the laser. In this context, we will explore use of vehicle pitching as a substitute for actuation of the laser, thus maintaining the mechanically passive nature of the system while gaining more triangulation information from a smaller range of vehicle positions.

The calibration process currently requires manual selection of interest regions from the image. The utility of calibration is greatly enhanced if these steps can be automated, since this allows in-the-field re-calibration. Future efforts will pursue increased automation of the calibration process.

The ability to discern fanned laser return from a complex scene illuminated by ambient light would be of great value to develop further. For the purposes expressed in this paper, such methods would have to have modest computational requirements. We will continue to seek solutions to this problem.

ACKNOWLEDGMENTS

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*. National Institute of Standards and Technology, 2004.

[2] 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.

[3] J.-Y. Bouguet, "Camera calibration toolbox for matlab." [Online]. Available: http://www.vision.caltech.edu/bouguetj

[4] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 11, pp. 1330–1334, 2000.

[5] C. Harris and M. Stephens, "A combined corner and edge detector," in *ALVEY Vision Conference*, 1988, pp. 147–151.

[6] D. Lowe, "Distinctive image features from scale-invariant keypoints," in *International Journal of Computer Vision*, vol. 20, 2003, pp. 91–110. [Online]. Available: citeseer.ist.psu.edu/lowe04distinctive.html

[7] 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.

[8] ANSI, *American National Standard for Safe Use of Lasers*. Laser Institute of America, 2000.