1998-02

Mobile Robot Range Sensing through Visual Looming

Mobile robot range sensing through visual looming

Erol Sahin and Paolo Gaudiano

February 1998

Technical Report CAS/CNS-98-010

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
677 Beacon Street
Boston, MA 02215

Mobile robot range sensing through visual looming

Erol Şahin and Paolo Gaudiano

Boston University Neurobotics Lab
Dept. of Cognitive and Neural Systems
677 Beacon St., Boston, MA 02215 USA
E-mail: {erol,gaudiano}@cns.bu.edu

Submitted to
ISIC/CIRA/ISAS '98
A Joint Conference on the Science and Technology of Intelligent Systems
National Institute of Standards and Technology,
Gaithersbug, MD 20899

This work is supported by DARPA, the Office of Naval Research and the Navy Research Laboratory through grants ONR-00014-96-1-0772 and ONR-00014-95-1-0409.

Abstract

This article describes and evaluates visual looming as a monocular range sensing method for mobile robots. The looming algorithm is based on the relationship between the displacement of a camera relative to an object, and the resulting change in the size of the object's image on the focal plane of the camera. We have carried out systematic experiments to evaluate the ranging accuracy of the looming algorithm using a Pioneer 1 mobile robot equipped with a color camera. We have also performed noise sensitivity for the looming algorithm, obtaining theoretical error bounds on the range estimates for given levels of odometric and visual noise, which were verified through experimental data. Our results suggest that looming can be used as a robust, inexpensive range sensor as a complement to sonar.

1 Introduction

This article describes the *looming algorithm*, a method for extracting range information from a monocular camera mounted on a mobile robot. Visual looming, the expansion of the projection size of an object on the retina, is usually the indication of an approaching object. It is normally perceived as a threat for a possible collision and is known to elicit reactive behaviors in animals (Caviness, Schiff, & Gibson, 1962).

Although its behavioral effects have been studied mainly in psychology, the looming phenomenon has interesting implications for robotics. Several independent studies have reported the use of looming as a method for obstacle avoidance (Joarder & Raviv, 1992) or for extracting the depth of an object (Williams, 1980; Williams & Hanson, 1988; Xu, 1992; Huttenlocher, Leventon, & Rucklidge, 1995). In particular, (Raviv, 1992) has done an excellent quantitative analysis of visual looming.

In this article we propose that looming can be a useful complement to sonar as an inexpensive, robust range finding sensor. The next section summarizes the basic properties of the looming algorithm. Section 3 describes our experimental setup and some initial results in which looming is compared systematically to sonar as a range finding sensor. Section 4 provides a formal analysis of the noise sensitivity of the looming algorithm. The article closes with a discussion and conclusions.

2 The Looming algorithm

The looming algorithm is based on the simple fact that objects look larger as they get closer and smaller as they move away. Fig. 1 depicts a scene in which a camera is viewing an object of size h from two different positions, at distances d_0 and d_1 from the object. Note that it is irrelevant whether the displacement is the result of camera movement or object movement.

Given a constant focal length f, the size of the projection of the object onto the focal plane depends on the distance between the object and the camera. In the case shown in Fig. 1, p_0 and p_1, respectively, represent the size of the projection of the same object at distances d_0 and d_1. Using similar triangles, it can be easily shown that

$$d_1 = -p_0 \frac{\Delta d}{p_1 - p_0}, \quad d_0 = -p_1 \frac{\Delta d}{p_1 - p_0} \tag{1}$$

where $\Delta d = d_1 - d_0$. We refer to Eq. 1 as the *looming equation*.

The looming equation can be shown to hold under either of the following assumptions: Either (1) the object lies anywhere along a plane parallel to the focal plane, and the displacement is in the direction perpendicular to these planes, or (2) the object is not necessarily parallel to the camera's plane, but the camera is aimed toward the center of the object at both positions (in other words, the camera is moving directly toward or directly away from the center of the object).

It should be noticed that when the object is slanted relative to the camera's line of sight, small pointing errors will not be significant: the distance estimated by the above equation will be the average of the distances of the nearest and farthest point of the object projected onto the camera's line of sight, so that as long as the center of the camera's focal plane is somewhere on the object, the distance measured will correspond to a point on the object.

Several observations can be made from Eq. 1.

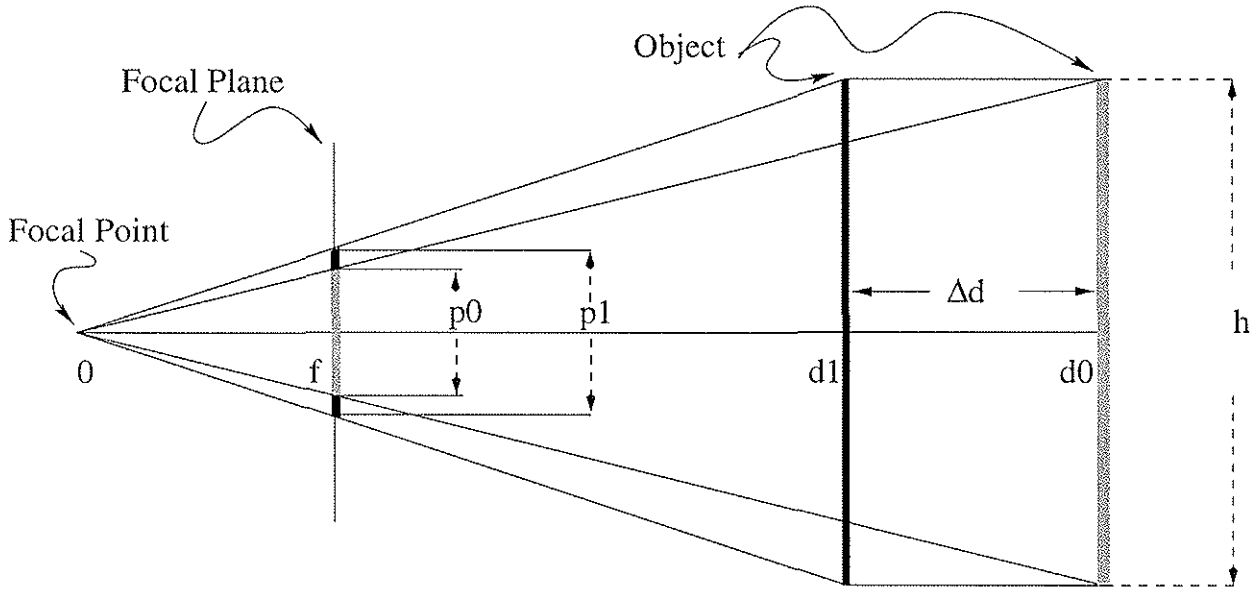

Figure 1: Diagram of the looming relationships in a camera-centered coordinate frame.

1 For 2-D objects, the equation can be applied to each dimension independently, and the two results can then be combined. It is also possible to rewrite the equation in terms of the object's area.

2 The equation is independent of the parameters of the optical system, such as focal length and pixel size. Hence no camera calibration is required. If pixel size or focal length are known, it is possible to estimate the true object size.

3 The equation is independent of the actual object size. The only *caveat* is that the object's projection must fit entirely within the focal plane. In the case of a 2-D object, it is sufficient that one dimension fit entirely within the focal plane.

4 Only two measurements of the object projection separated by a net displacement obeying one of the above conditions are sufficient. No synchronization between projections is required.

The assumptions listed above are not problematic in practical use. If the object is slanted relative to the camera's line of sight, it is a simple matter to aim the camera approximately toward the object's center; furthermore, in typical indoor environments objects are likely to be parallel to the camera's focal plane at least along one dimension (the vertical dimension for objects standing upright or attached to the wall).

Another important issue that should be addressed is the noise sensitivity of the looming algorithm. For small changes in the projection size, the looming equation has the tendency to amplify any noise in the odometric or visual measurements. One way to overcome this problem is to use tracking over objects over multiple frames (Chaney, 1994), as we did, giving time for making larger displacements. This issue will be addressed in detail in Section 4. For now, we will describe the experimental results of the looming algorithm on a real mobile robot.

3 Experimental results with a mobile robot

For our experiments we have used a Pioneer 1 mobile robot (Real World Interface, Inc., Jaffrey, NH). The Pioneer 1 is equipped with a *Chinon CX-062*, a 5mm color camera with 200x250-pixel resolution covering approximately 60 deg of the visual field. The camera is connected to a *Cognachrome 2000 Vision System* (Newton Research Labs, Bellevue, WA), a 68332-based frame grabber that can perform a variety of simple operations at frame rate. Among other things, the Cognachrome can be "trained" to locate blobs of arbitrary colors in the image, returning the horizontal and vertical size, as well as the location, of multiple blobs of the desired color.

The usefulness of the looming algorithm becomes evident when we perform a systematic comparison of looming and sonar under a variety of conditions.

3.1 Comparison of looming and sonar

One profound limitation of sonar as a ranging sensor is that it usually fails for slanted surfaces. This is due to the nature of sonar, which is fairly directional, and is usually reflected off surfaces at an angle that is equal to the angle of incidence (J. Borenstein & Feng, 1996; Leonard & Durrant-Whyte, 1992). Hence a typical sonar sensor will fail to receive an echo from objects that are slanted at more than approximately 15 deg relative to the line of sight of the sonar. This limitation of sonar is true for most types of surface, with the exception of special materials such as styrofoam that tend to reflect sonar in all directions.

In contrast, the looming algorithm functions accurately regardless of an object's slant, as long as the camera is aimed approximately to the center of the object during the robot's displacement. Problems with the looming algorithm may arise if an object is slanted so steeply that its width becomes less than a few pixels. But even in this case looming functions well if the object is not steeply slanted along both dimensions, since the looming algorithm can be applied independently for each dimension.

We have carried out a systematic comparison of the performance of looming and sonar when the object whose distance is being measured is slanted by various amounts relative to the sensor's line of sight. Our experimental setup is shown in Fig. 2. A letter-sized (216×279 mm) piece of colored paper is attached to a wall at approximately the same height as the robot's camera. Radial lines are drawn on the floor, emanating from the point aligned vertically with the the center of the paper, at angles of 22.5-90.0 deg in 11.25 deg increments relative to the wall (90 deg corresponds to a line perpendicular to the wall). Each of these lines is marked at distances of 611, 917, 1223 and 1528 mm.

The Pioneer 1 is placed at each one of the resulting 28 points marked on the floor (four different distances at each of the seven angles), aimed in the direction of the piece of paper. The distance from the Pioneer 1 to the paper is then measured with the robot's forward-facing sonar and with the looming algorithm.

Distance estimation with the looming algorithm requires the robot to move forward or backward in the direction of the piece of paper. In order to make the data collection automatic and more reliable, a reference bar is attached to the floor at each point on the grid, perpendicular to the underlying radial line, to provide a reference distance and orientation. Initially the robot is aligned with the radial line with both wheels touching the bar. At this point the width and height of the paper (as seen through the camera) are used as a first measurement. Then the robot is instructed

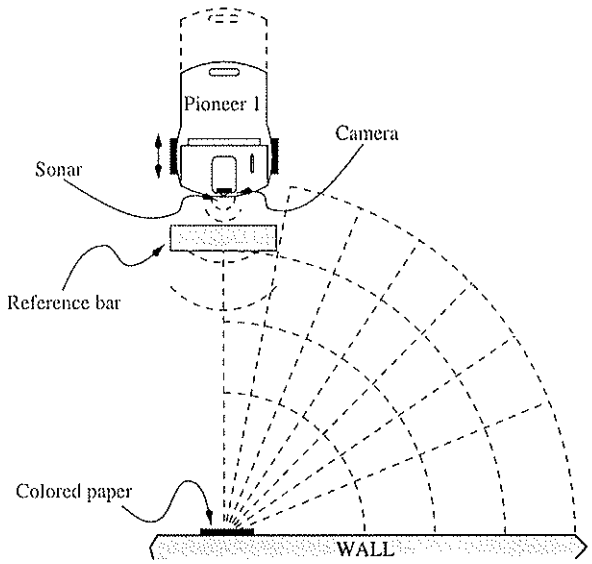

Figure 2: Top view of the experimental setup, showing the Pioneer and the grid of points used for the experiments.

to move backwards by 305 mm and a second measurement is recorded. The looming algorithm uses these two measurements to compute the distance of the paper when it was at the bar. Then the robot moves forward until its wheels hit the bar to ensure that the robot is aligned correctly, though the robot's lateral position might be shifted slightly due to odometry errors. This procedure is then repeated twenty times at each point to derive a meaningful average measurement. After every five measurements, the robot is manually shifted back to the reference point to avoid a bias from positional errors. In any event, the shift accumulated during five cycles was never observed to exceed 40 mm. The sonar reading at each point is also repeated 20 times and averaged.

Figure 3 shows the combined averaged results obtained with sonar (cross symbol), and with the looming algorithm using the paper's width (triangles) or height (squares). Notice that in the figure the vertical line of the grid corresponds to 90 deg, i.e., when the piece of paper is perpendicular to the line of sight of the robot. In **(a)** we include all points from the three range measurements, while in **(b)** we truncate the sonar results to improve the visibility of the results obtained through looming (using width and height independently) for clarity.

The results illustrate that when the paper is slanted by 22.5 deg or more relative to the line of sight of the robot (i.e., 67.5 deg from the wall) the sonar is unable to detect the wall (the large values shown in the figure correspond to the sonar's maximum range, meaning that an echo was not received). In contrast, the looming algorithm performs quite well up to the maximum slant angle.

Some of the biases in the results are worth discussion. Fig. 4 reproduces the looming from height data of Fig. 3 in Cartesian coordinates with error bars. This plot shows clearly that range estimates using looming based on height tend to underestimate the distance with decreasing angle (i.e., increasing slant angle). This is due to an artifact of the object segmentation algorithm that is implemented on the Pioneer's frame grabber. This system uses a bounding box approach to compute the width and the height of the blob being tracked. When the robot is approaching a

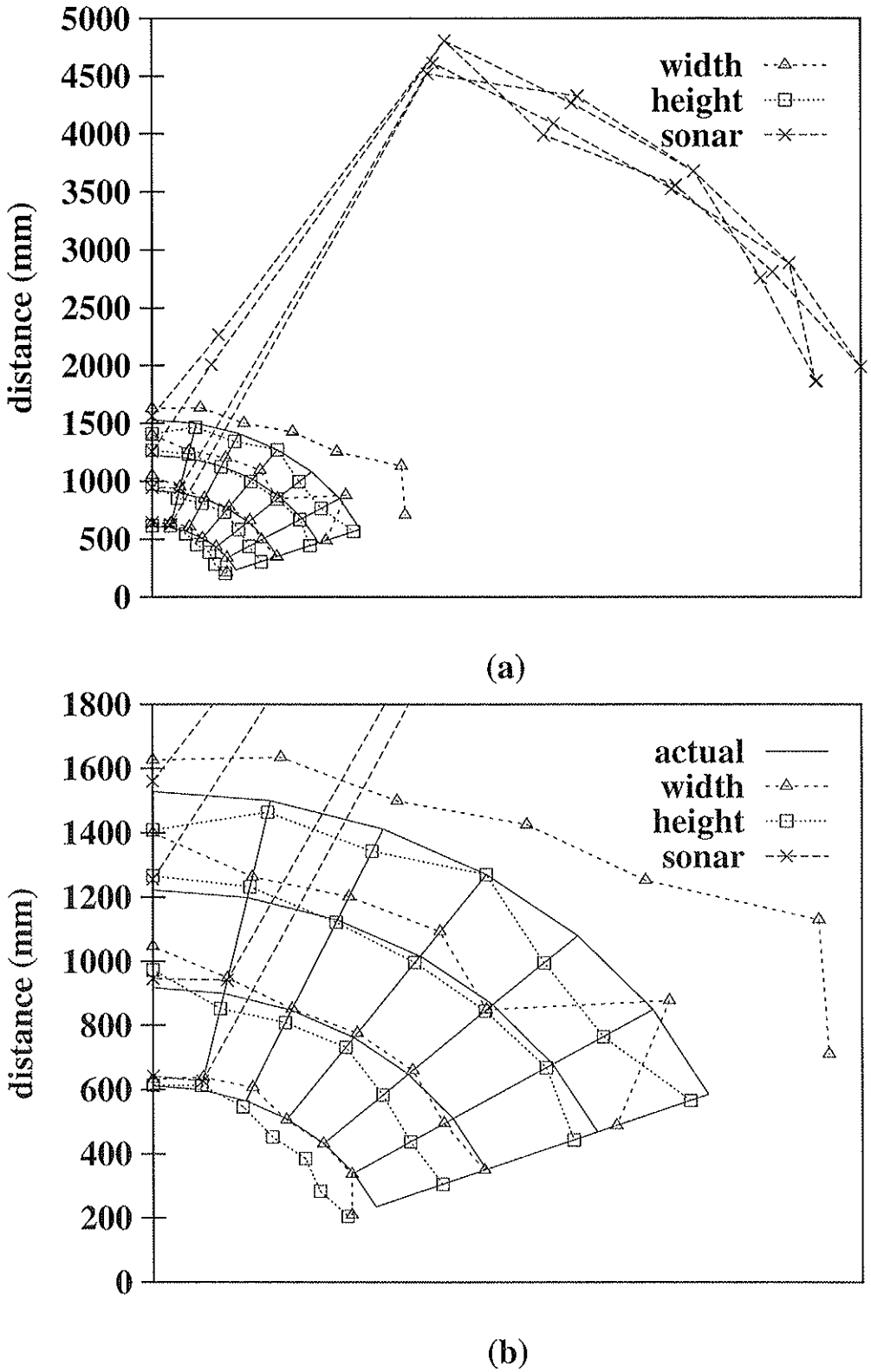

Figure 3: Comparison of sonar (cross symbols) and looming from width (triangles) and height (squares) as range sensors. The data in (**b**) are the same as in (**a**), but truncated to enhance the visibility of the looming data.

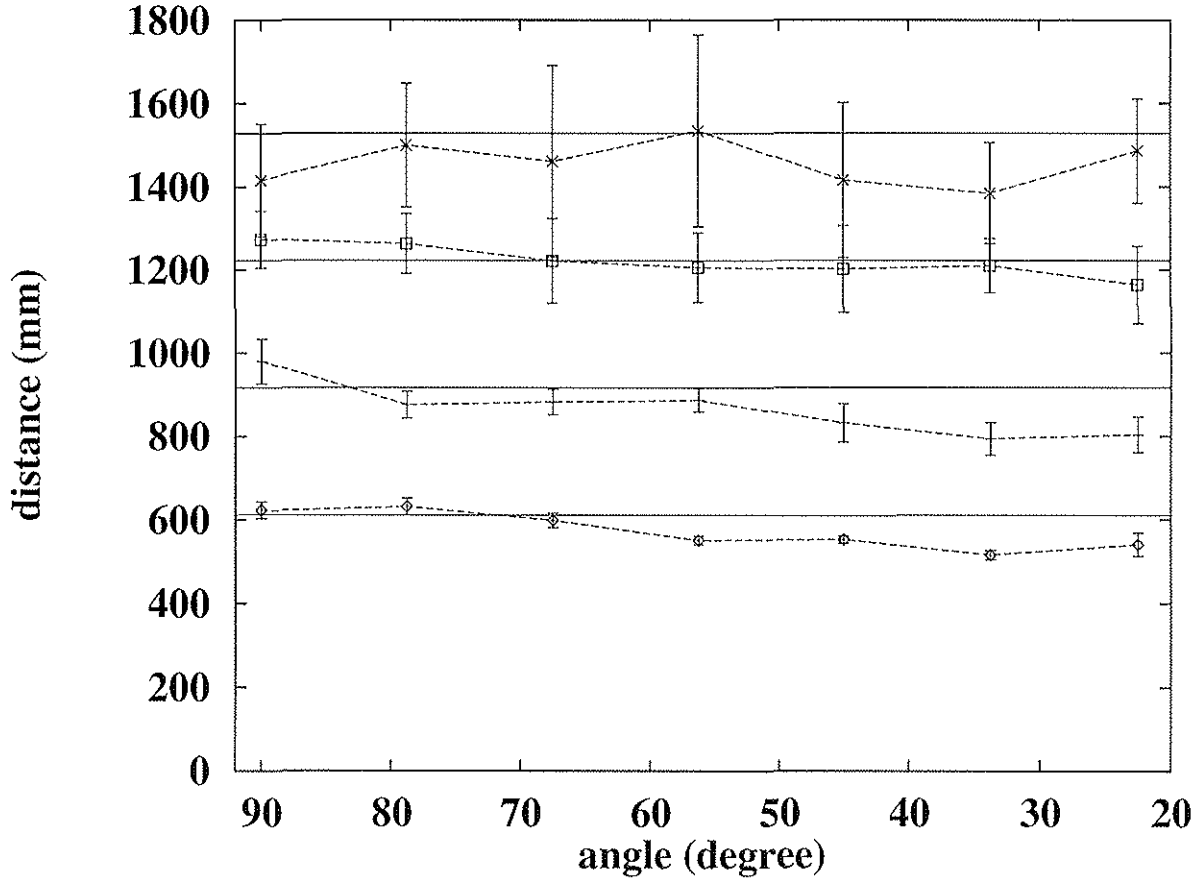

Figure 4: Average distance measured with looming from height as a function of the actual distance and angle to the target. The horizontal lines represent the actual distances at which measurements were taken. Error bars represent +/- one standard deviation of the mean.

slanted object, the closer vertical edge of the object is returned as the height of the blob due to perspective effects. As a result the range value returned by the looming algorithm in this case corresponds to the nearest edge of the object, whereas the 'actual' distance in the plots is measured relative to the center of the object.

Other interesting observations about the results will be discussed in the next section in the context of the noise sensitivity analysis of the looming algorithm.

4 Noise Sensitivity Analysis

An important concern in the practical applicability of the looming algorithm is its noise sensitivity. It can be seen from Eq. 1 that the looming equation is likely to amplify any noise when the change in the projection size is small. Therefore it is important to know how the odometric and visual noise contribute to the range error.

The looming equation has two sources of noise: odometric and visual. Odometric noise ($e_{\Delta d}$) is the error, due to slippage and other factors, in measuring the displacement of the robot. It is proportional to the distance traveled and usually in the order of a few percent of the distance

traveled on most surfaces. Visual noise arises in measuring the projection of the object on the focal plane, and is due mainly to pixel quantization, object segmentation and optical distortion. We distinguish between noise in the projection size (e_p) from noise in the projection size change ($e_{\Delta p}$). This distinction is made in order to avoid making any assumptions about the statistics or possible correlations between independent samples of image projection size. These two sources of error will jointly be referred to as visual error.

The noise sensitivity of the looming algorithm can be analyzed by adding 'noise' to the measured variables, yielding

$$d_1 \pm e_d = (p_0 \pm e_p)\frac{\Delta d(1 \pm e_{\Delta d})}{(p_1 - p_0) \pm e_{\Delta p}}$$

where e_d is the error in the estimated range. Standard error analysis yields the fractional error as

$$\frac{e_d}{d_1} = \sqrt{(\frac{e_p}{p_0})^2 + (\frac{e_{\Delta p}}{p_1 - p_0})^2 + e_{\Delta d}^2}$$

Several observations can be made from this analysis. First, estimated range error is proportional to the distance, which means—other parameters being constant—that estimations will get better as the robot gets closer. Second, in the absence of visual error, the odometric error sets the baseline for the fractional error. Third, the contribution to the fractional error from the projection size noise e_p is inversely proportional to the object's projection size. This formalizes the intuition that the size of an object's projection should be large relative to the visual noise e_p. Finally, for objects whose initial projection size is much larger than the visual error e_p, the fractional error is initially dominated by the noise in the projection size change $e_{\Delta p}$, which decreases as the projection size change increases.

The last point indicates that $e_{\Delta p}$ is likely to dominate for small changes in projection size, and that fractional error should drop quickly with increasingly large changes in projection size. However, the rate at which the projection size changes depends jointly on the distance and size of the object: a short movement towards a nearby small object may yield the same projection size change as a longer movement towards a distant large object. Therefore rather than considering absolute size and distance, it is most sensible to use the ratio of the distance to the size of the object as a dimensionless parameter to measure fractional error.

Figure 5 illustrates the theoretical error bounds and actual estimates of the looming algorithm when the Pioneer 1 starts approaching the paper at 10 and 3 paper height distances. For this figure we ignore the terms e_p and $e_{\Delta d}$, which our analysis shows have a smaller influence than $e_{\Delta p}$. The axes are marked in units of object size to make the analysis invariant with respect to the height of the object or the actual distance to the object. It can be seen that the error bounds shrink rapidly when the initial distance is 3 units. The estimate gets into reasonable bounds after moving half a unit. As expected the error bounds are initially large and shrink more slowly when the initial distance is 10 units. In practice however, we have found that noise in the empirical measurements drops more quickly than the theoretical bounds. This is undoubtedly due to some noise cancellation that occurs in subtracting the two projection size values p_0 and p_1.

In light of this analysis, we also want to discuss some of the biases that can be seen in Figs. 3 and 4. The distance estimates obtained with looming using the height of the paper are generally more stable and accurate than those obtained with the width of the paper. This is to be expected because the increasing slant causes foreshortening of the projection in the width but not in the

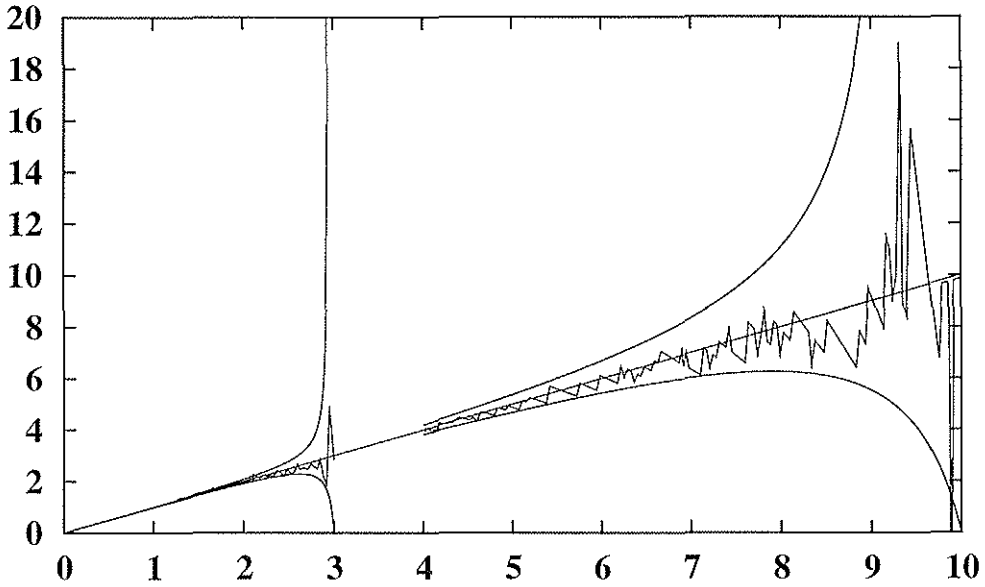

Figure 5: Theoretical and empirical fractional error versus distance. Dimensions along both axes are marked as multiples of the size of the object. For our setup, the focal length of the camera is 287 pixels and $e_{\Delta p}$ is chosen as 2 pixels, neglecting e_p and $e_{\Delta d}$. Note that the upper error bound has a singularity at the point when $e_{\Delta p}$ cancels out the projection size change. The part of the 'upper' error bound which becomes negative after the singularity is dropped for clarity. See text for discussion.

height of the paper. Other things being equal, decreasing width cause the error bounds to increase by increasing the ratio of distance to width. The same effect can also be seen in the distance estimates obtained using the height at different distances: the standard deviation of the estimation increases with increasing distance for a given object.

Another point to be aware of is that if the projection size change has the wrong sign, e.g., if the object appears to be smaller after moving towards it due to visual noise or a moving object, the looming algorithm will estimate negative distances. If we presume that the objects are stationary, these cases can be ruled out easily. We will discuss how they may be handled to get additional information, e.g. whether the object is moving or not, in the next section.

5 Discussion

We have proposed that visual looming can be a complementary range sensor to sonar and presented experimental evidence to support this. However, some obstacles remain in the use of looming as a range sensor. The biggest problem is, without question, the necessity of segmenting objects from their background. In this study we have side-stepped the problem by using pieces of colored paper as targets and a real-time color-tracking system that extracts the size and position of the colored paper during the robot's movements. This type of manipulation would not work in unstructured environments, but several steps could be taken to allow looming to work in less structured environments. One possibility is to develop a fast, robust, general-purpose visual segmentation algorithm. We are not sure whether this kind of segmentation algorithm exists or is even feasible for mobile

robots in completely unstructured environments. An alternative possibility, and one that we are currently pursuing, is to enhance the power of the color tracking algorithm. For instance, rather than using pieces of paper of a given pre-established color, we are developing a tracking algorithm that automatically chooses colors in the image that might correspond to isolated objects.

Another potential limitation of looming is that, at least in theory, it works well only for completely unobstructed objects. However, when treating looming as a sensory-motor algorithm on a mobile robot, this limitation can actually be overcome easily. For instance, suppose that an object is occluded by another object that is placed between the robot and the object being tracked. When the robot moves by a small amount toward the object, the distance estimate obtained through looming will be incorrect: the object will "appear" closer than it is because its projected size has increased more than it should have because of the presence of the occluder[1]. Let's say that the algorithm returns an incorrect estimated distance d_a after the first step. Now if the robot moves again toward the object by an amount Δd, the looming algorithm will again err, and this time the estimated distance will be *smaller* than $d_a - \Delta d$ because even more of the object has been uncovered from behind the occluder. But because the robot *knows* through its odometry that it has only moved by Δd, the algorithm can realize that something is wrong. In addition, with the exception of degenerate cases, the occlusion will not affect both dimensions of the object, so that the algorithm could also notice a discrepancy by comparing the result obtained independently from the two dimensions. Hence the looming algorithm can be extended to include consistency checks across dimensions or even within one dimension across multiple readings, and it can use the results to figure out when there is something wrong with the estimated distances. With a similar reasoning it is also possible to detect moving objects: if two subsequent readings are inconsistent, the robot can stop momentarily to check whether the projection size changes in the absence of robot movements.

6 Conclusions

Range information is essential for many navigation schemes such as evidence grids (Elfes, 1987), vector field histograms (Borenstein & Koren, 1989) and frontier-based navigation (Schultz & Adams, 1996; Yamamuchi, Schultz, & Adams, 1997). Sonar is often used as a sort of 'ray-trace scanner' range sensor within these schemes, but this leads to errors from reflections, false returns and cross-talk. Furthermore, sonar suffers from angular uncertainty, as eloquently summarized by Leonard and Durrant-Whyte (1992): "sonar is in fact a very good sensor... [but the] first step in using the sonar is to accept the realities of the physics of acoustic sensing... Most of the sonar interpretation algorithms have been built as if sonar is a 'ray-trace scanner'."

We believe that the shortcomings of sonar can be remedied by the use of complementary range sensors such as looming. Looming is ideally suited for ranging when objects are slanted relative to the robot's (or sonar's) line of sight. Looming is also well suited for objects that may be located on a wall but above or below the sonar's line of sight. Finally, by combining sensory and "motor" information, looming can be extended to handle complex situations where even a purely visual sensor would not be adequate.

[1] As the robot moves closer to the object, more of the object will be uncovered from behind the occluder, hence the increase in projection size will be greater than it would have been without the occluder.

References

Borenstein, J., & Koren, Y. (1989). A real-time obstacle avoidance for fast mobile robots. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(5), 1179–1187.

Caviness, J. A., Schiff, W., & Gibson, J. J. (1962). Persistent fear responses in rhesus monkeys to the optical stimulus of 'looming'. *Science*, *136*, 982–983.

Chaney, R. D. (1994). Multiple-frame motion estimation using simple region feature. In *Proceedings of SPIE, Intelligent Robots and Computer Vision XIII: 3D Vision, Product Inspection, and Active Vision*, Vol. SPIE-2354, pp. 10–25.

Elfes, A. (1987). Sonar based real-world mapping and navigation. *IEEE Transactions on Robotics and Automation*, *3*(3), 249–265.

Huttenlocher, D. P., Leventon, M. E., & Rucklidge, W. J. (1995). Visually-guided navigation by comparing edge images. In Goldberg, K., Halperin, D., Latombe, J.-C., & Wilson, R. (Eds.), *Algorithmic Foundations of Robotics*, pp. 85–96. A.K. Peters.

J. Borenstein, H. E., & Feng, L. (1996). *Navigating mobile robots*. A.K. Peters, Wellesley, MA.

Joarder, K., & Raviv, D. (1992). Autonomous obstacle avoidance using visual fixation and looming. In *Proceedings of SPIE*, Vol. 1825, pp. 733–744.

Leonard, J. J., & Durrant-Whyte, H. F. (1992). *Directed sonar sensing for mobile robot navigation*. The Kluwer international series in engineering and computer science. Kluwer Academic Publishers, Boston.

Raviv, D. (1992). Visual looming. In *Proceedings of SPIE*, Vol. 1825, pp. 703–713.

Schultz, A. C., & Adams, W. (1996). Continuous localization using evidence grids. Tech. rep. AIC-96-007, NCARAI. Submitted to Autonomous Robots.

Williams, L. R., & Hanson, A. R. (1988). Translating optical flow into token matches and depth from looming. In *Proceedings of the 2nd Intl. Conference on Computer Vision*.

Williams, T. D. (1980). Depth from camera motion in a real world scene. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *2*(6), 511–516.

Xu, S. B. (1992). Qualitative depth from monoscopic cues. In *Proceedings of the International Conference on Image Processing and its Applications*, pp. 437–440.

Yamamuchi, B., Schultz, A. C., & Adams, W. (1997). Integrating exploration and localization for mobile robots. Tech. rep. AIC-97-021, NCARAI. Submitted to Autonomous Robots.