Monocular Vision, Optical Blur and Visual Looming as a Basis for Mobile Robot Obstacle Avoidance

Iain J. Brookshaw* and Dr. Tobias Low
University of Southern Queensland, Australia
*Corresponding: w0086292@umail.usq.edu.au

Abstract—**Avoidance of obstacles is a fundamental problem for mobile robots. In recent times, a rapid increase in computing power combined with an equally steep decrease in size, electrical requirements and cost in computers has led to the increase in the use of cameras and vision as a realistic means of achieving real time avoidance.**

In this paper, we investigate a means by which a machine vision system could utilise optical blur as an avoidance indicator. The methods used are intended for monocular systems and employ the blur recovery methods of Hu and de Haan to find optical blur and optical blur and the looming method described by Raviv and Joarder and Sahin and Gaudiano to relate this to object approach. It was intended that this system be relatively simple in hardware and software implementation. To verify the success of the design, we conducted tests in a controlled environment. It was found that obstacle approach could reliably be computed through this method, but its success depended on the camera lens properties.

Keywords—**mobile robots, obstacle avoidance, looming, optical blur**

1 INTRODUCTION

Recently, machine vision has been an increasingly popular sensor choice for obstacle avoidance. While there are a number of differing means for achieving this, we feel that they could be augmented with a slightly different approach. In this paper, we investigate the feasibility of using optical blur as an avoidance cue. We aim to provide the basis of an avoidance method that is robust, simple and not tied to explicitly known camera properties. By combining a robust blur recovery mechanism with the concept of visual looming, we illustrate that the optical blur of an object in an image can be used to produce viable navigation cues in a fashion that is suitable for mobile robot avoidance.

Classically, avoidance has been achieved by stereo vision, the use two cameras has been popular in various depth perception schemes and details occur in almost every vision text (eg: [1, 2] etc.). However true stereo systems have their own complexities, not the least of which is their need for two cameras [3].

In monocular vision, a stereo approximation can be created through use of the ‘stereo-through-motion’ concept, which uses the discrepancy between frames caused by robot motion to approximate true stereo. However this method is limited by the fact that no depth information can be recovered for objects on the optical axis [1]. For a forward moving robot these are the very objects most likely to cause collision.

Aside from stereo and related systems, there exist several other paradigms for visual obstacle detection. Various methods of optical flow are recounted in [1, 2]. However these sources also point out that such methods suffer from a number of qualifiers and problems, such as the aperture problem.

The method we chose to investigate is based on the Looming method described in [4] and [5]. This is based on the change of various image properties over time. These two sources contain most of the key information relating to this method. The work of [5] demonstrates the use of projected object area, using the increasing area of approaching objects in a controlled environment as a basis for avoidance. This work is supported by the lengthy paper of [5], where other cues, such as texture and irradiance are discussed. The work of [5] also discusses the use of optical blur in some detail, although no results are presented for this approach. Additional sources for this method are rare, typically the work of [6], which make reference to the looming approach, but prefer a method augmented by other sensors.

2 KEY COMPONENTS

2.1 Visual Looming

Visual looming is based on the simple fact that “objects look larger as they get closer and smaller as they move away” [4]. While this definition was originally used in relation to objects' projected area, the same principle can be applied to several image object properties, such as texture, irradiance and optical blur [5]. Whichever quantity is used, the basic mathematical description is the same. As described in [5] A quantity L, the scalar looming value of units $[time^{-1}]$, is defined by the following equation:

$$L = \frac{\frac{dR}{dt}}{R} \qquad (1)$$

Where R is the range between the object and the camera plane and the negative sign is to ensure that decreasing range creates increasing (positive) looming value [5]. At first glance this is not very useful, as range (R) is unknown. However, if R could be replaced by some other quantity g, proportional to R, then the value L could be established. If the value g were computed in real-time then an evolving estimate of would be available [5]. Also if g is directly proportional to R, then it would be unnecessary to know any camera properties, thus negating the need for camera calibration. This estimate of looming could be combined with a threshold value of *L* computed by used of (1).

The value of *R* in that calculation would be the desired minimum range for safe operation and *dR/dt* would be the current speed. By comparing this threshold to the variable value of *L* computed from some image property *g*, it becomes possible to establish a mechanism for imminent collision warning, without the range *R* being explicitly known at all times. As pointed out in [5] there are a number of image quantities that could be considered candidates for *g*. They mention object projection area (used to considerable effect in [4]), conduct extensive tests with object irradiance and texture and discuss blur at length. However no data for blur as a looming cue is presented. Using the optical model described below, it can be seen that blur radius, like projected area, is directly related to the distance of the object from the image plane. It is also preferable to projected area as a measurement quantity, because the area approach assumes that an object is entirely within the image frame in at least one direction [6]. Blur has the advantage that it can be measured at the peripheral of the object regardless of how much of the object is in the frame.

2.2 Optical Blur and Blur Recovery

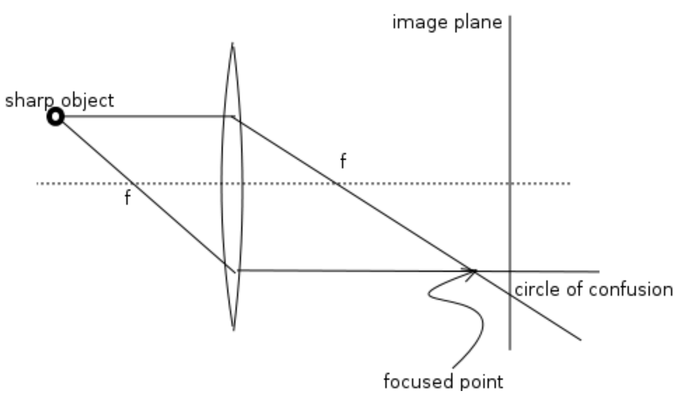

Fig. 1: Illustration of the circle of confusion and blur optics.

Beginning with the thin lens model (Thin lenses are the standard model as described in [7, 8] and most other sources.) depicted in Fig. 1, it can be seen that a sharp object will only appear sharp in the image if the image plane is at a certain point. In actual fact this 'point' is a narrow range, the width of which is dependent on the lens specifics and sensor resolution. This range is known as the depth of field and is closely related to the lens focal length [6]. If the image plane is not at this point (in this range) then the sharp point becomes a blurred circle, referred to as the circle of confusion [9].If this circle could be quantified it could be used in the looming calculations to compute L, as the radius of the circle of confusion (blur) will be related to distance from the image plane to the object, assuming the camera and lens specifics are constant [8].

Many sources, often beginning with the work of [10], model the blur circle as a two dimensional, symmetric Gaussian function. Thus radius of blur becomes the σ value of the Gaussian (actually, it is proportional to this, but σ is usually held to be a good measure of blur). The method of Hu and de Haan recovers this value by re-blurring the original image twice with a Gaussian blur, using increasing values of σ each time [11]. They then illustrate that using a ratio of differences between the blurred images and assuming that the re-blur values of σ are much larger than the original, one can recover a good approximation of the original σ at the edges of objects. The blur values are only fully accurate at the edges of image objects. This is because the mathematical mechanism developed by Hu and de Haan computes blur based on the maximum difference between the original and re-blurred frames, which occurs at object edges (see [11] for details). When implementing their mechanism on a two dimensional image, they divided the image into a small grid and used the largest difference for that area to compute blur. A similar approach was taken here, it is important to observe that as one moves away from an edge, the recovered blur values become less accurate. It is important to note that the original blur must be relatively small compared to the re-blur values [11]. As there is a practical limit to how large these values can be made (or computation becomes intractable), one must assume that there exists a point for which blur is too large to be reliably computed.

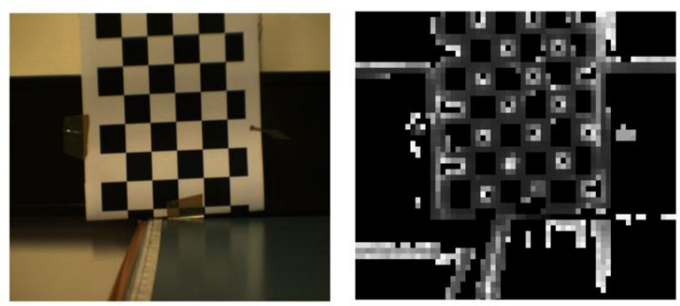

Fig. 2: The experimental setup as seen by the camera, showing the input image on the left and the blur image on the right, light squares represent more blur. Average blur was computed in a central window. These photos show image 0 of the 4.3mm lens sequence.

3 EXPERIMENTAL SETUP

In order to demonstrate the feasibility of avoidance through blur and looming, a series of images were taken in a controlled environment and processed using the ideas outlined above. Inspired by the experiments in [5] and [4], an idealised planar object was fixed to a wall. In this case the object was a checker-board with clear, arbitrarily sized squares, sufficiently large to be easily distinguished at a distance. A tape measure was then fastened perpendicular to the image on a bench. The camera was placed at the maximum distance from the image object and focused on that object by hand. This hand focusing was relatively inexact, relying on decreasing the returned value of blur for the target image until a minimum had been reached. This was as we wished to evaluate the system without resorting to detailed camera calibration, believing that the method should be sufficiently robust to withstand small errors from focusing. For purposes of comparison, three lenses were used, a 25mm telephoto, an 8mm telephoto and a 4.3mm ‘standard’ lens. To illustrate the consistency of the results, two tests were undertaken for each lens.

Raviv and Joarder [5] suggest that the camera should be focused at infinity for blur based looming. Clearly, this is impractical as the blur for any object in view would be too large for the approximations described in section 2.2 and thus too large to compute. Thus the camera was focused at the maximum test range (1m). To compensate for this, the true looming curve was also ‘zeroed’ at this point. Following the setup stage, a series of images were taken, each approaching closer to the image object each frame by a known distance step. Blur and looming were then found from this sequence. In all tests the maximum (starting) distance was 1m and the step 20mm. Distance steps and maximum distance was expressly known so that a true looming could be computed for comparison, they were not used in the blur looming computations.

For these tests blur was computed by taking a rectangle in the centre of the camera image and computing a blur average for this area. The change in blur average for the sequence of frames provided the Δr value in (1). For these tests, it was assumed that $\Delta t = 1$. This was considered acceptable as ‘speed’ could be any constant, or indeed a variable control input.

The averaging approach for blur finding has disadvantages, in that it is simplistic and assumes that all objects in the area are on the same plane and orientation relative to the camera. However, for the controlled test environment this was considered an acceptable assumption, the averaging approach providing consistent results (see below). Although such a method could not be used in a real mobile application (the assumption that the area contains one at object would almost certainly be violated), the overall method of looming-through-blur could be implemented by substituting this area averaging technique for a simple object tracking method. The critical point is to find the previous frame history of blur radius r for a given object. In the experiments we were able to assume that the area of interest contained the same object in each frame. The advantages of this experimental set up are that, while it is not real-time, it is controlled and produces clear data, illustrating the efficacy of looming-through-blur. While modifications needed to implement this concept in real time avoidance for mobile robots they were considered relatively minor.

4 RESULTS

The test as described above was conducted for three lenses, a 25mm and 8mm telephoto lenses as well as for a 4.3mm standard lens. The returned values of blur are illustrated in Fig. 3. They show clearly that blur is related to distance. However they also show that the blur rate of change is specific to each lens.

As the initial Looming results were very noisy, two forms of filter were applied to the data, a moving average filter and a Kalman filter. The moving average had a generous window of 15 data points while the Kalman parameters were generous estimates. Some experimentation showed that the filter output was not very sensitive to these parameters. As this style of filter was used largely to illustrate the difference between the moving average and this more classic approach, this was not considered very important.

Fig. 5 shows the unfiltered looming values for all lenses and tests. Fig. 6 shows the looming data filtered with a moving average filter, while Fig. 7 shows the data filtered using a Kalman filter. The errors for these three looming results are given in Fig. 8 through 10.

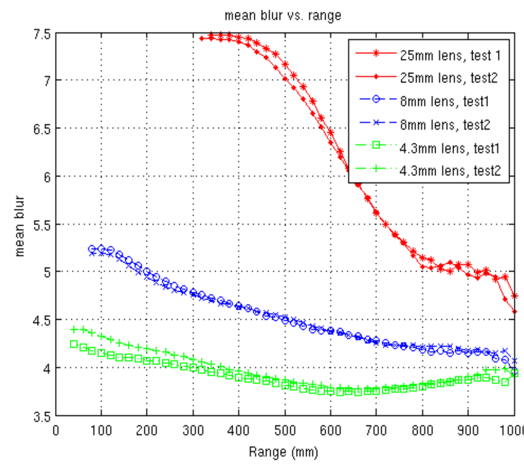

Fig. 3: The average blur radius vs. range for all lenses (25mm, 8mm, and 4.3mm) and tests.

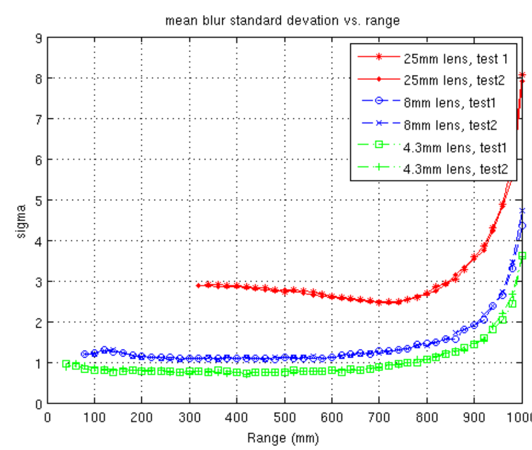

Fig. 4: The standard deviation vs. range for all lenses and tests

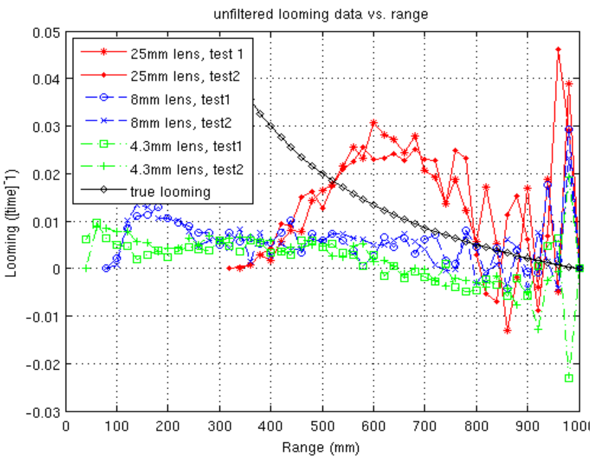

Fig. 5: Unfiltered Looming data vs. Range for all tests.

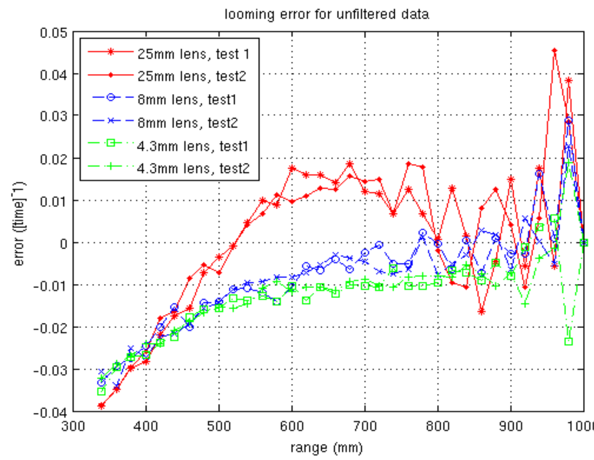

Fig. 8: Unfiltered looming error vs. Range for all tests.

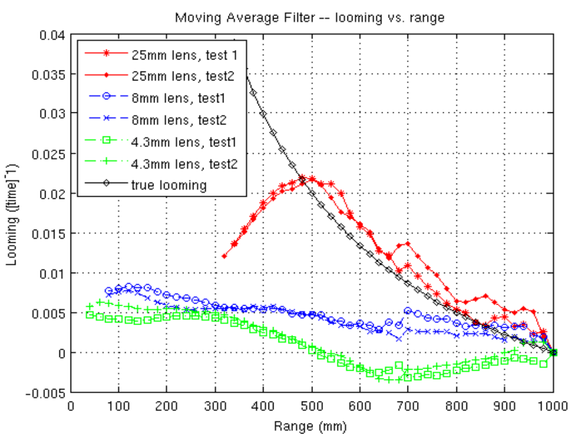

Fig. 6: Looming with a Moving Average Filter vs. Range for all tests.

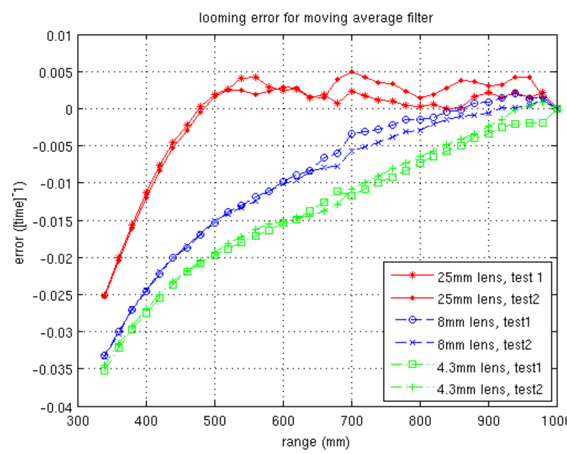

Fig. 9: Looming Error for moving average filtered data vs. Range.

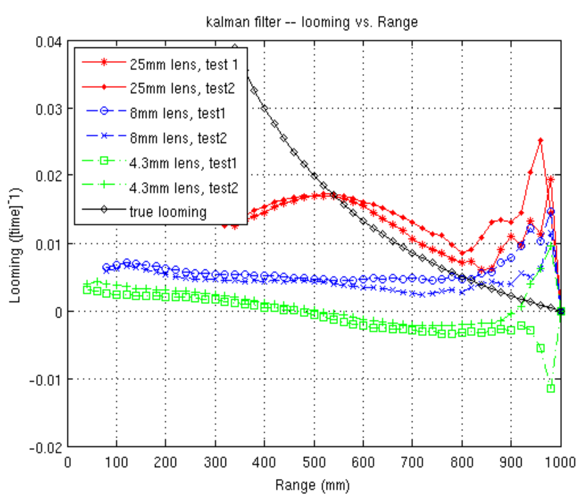

Fig. 7: Looming data vs. Range for all tests. Filtered using a Kalman filter.

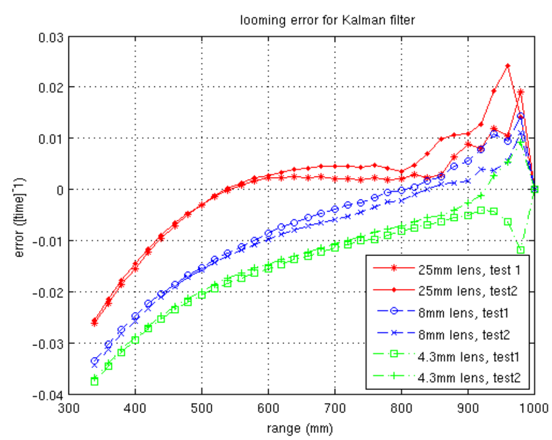

Fig. 10: Looming Error for Kalman filtered data vs. Range.

5 DISSCUSSION

5.1. Blur Mean and Variance

As can be seen in Fig. 3, the blur mean for all lenses advances in a smooth fashion. However the variance is significant, especially when the image is sharpest. We believe that at sharp points in the image the blur recovery is strongly affected by noise, with small noise points being erroneously reported as ‘sharp’ objects. While the variance decays rapidly, its final state is still quite high. Qualitative examination of the blur recovered illustrates that a ‘penumbra’ of less accurate blur exists around each edge in the image, as the method of Hu and de Haan is most accurate on the edge [11]. These increasingly inaccurate blur values could account for the relatively wide spread of the mean. However, this was not considered a serious problem due to the smooth evolution of the mean and its clear relationship with range. With a more advanced tracking system, this might not be an issue.

5.2. Lenses and Depth of Field

It is clear that there is a great difference between the three lenses in the looming results. The 25mm telephoto shows much better looming results than the two shorter focus length cameras. We hypothesize that this is because the longer lens has a much shorter depth of field due to its longer focal length, producing much more distinct ‘layers’ of blur than the other two lenses. This is supported by Fig. 3, notice how the blur mean for the longer lens advances much faster relative to distance, thus the Δr value is much larger in this lens, resulting in a correspondingly larger value of looming. This same figure shows that, although the lenses have a very smooth blur vs. range curve, they advance at very different rates. It is not known at this stage if other long focal length lenses would be able to produce the same curve, or if the blur would advance even faster, resulting in an overestimated looming curve as the 4.3mm lens was underestimated. If this occurred it would be necessary to produce some means of scaling.

Additionally, Fig. 3 shows that the shorter the focal length, the greater the difficulty in focusing the image to begin with. Notice that the 4.3mm lens in particular does not start at its lowest blur. This indicates that the minimum blur is difficult to find, further illustrating that there is little clear distinction between different blur states for this lens.

5.3. Looming

Despite the clear difference in looming output for the three lenses, it is also clear that they are exhibiting a measurable looming increase as range diminishes. However, it is doubtful that the shorter lenses could be directly useful as their error is clearly not linearly related to range (see Fig. 8-10). From the results in Fig. 6 and Fig. 9 the long focus lens could definitely provide an accurate estimate of the looming value; however it is not known if this good estimate will be the same for all lenses possessing a suitably short depth of field, or if a longer focus lens (say a 50mm) would give looming values that rose much faster than the true value. If the latter case, some calibration could be necessary.

Additionally, the calculated looming values are clearly very noisy, some form of filtering is necessary to provide coherent output. Of the two filters tried, the moving average gave the best results. As discussed, the worst noise is located near the sharpest image.

Finally, the looming values show a marked decrease as range decreases past a certain point. This is not as marked in the shorter lenses, but is clear in the 25mm lens. This is believed to be a result of the continuing increase in blur, after this point the blur is becoming too large to compute accurately.

In comparison to previous looming efforts, such as [4, 5] the results are promising. In the experimental results for area looming in [5] the looming error over 320mm (starting at 1300mm and advancing in steps of 20mm for 16 steps) was on approximate order of 0.01, while their irradiance and texture tests showed similar orders of error (ranging from about 0.01 for area and texture to about 0.025 for irradiance). While our results were more erratic (their data was shown without smoothing), Fig. 9 shows the error for the 25mm lens to relatively good by comparison, around 0.005 for a moving average filter (the results in [4] are of a different form). It should be noted that the values given for error from [5] may not be exact, as their results are presented in graphical form. Also they conducted tests at various angles, but only zero angle tests were compared to this work.

6 FUTURE WORK: MOBILE AVOIDANCE

Clearly, the next stage is the implementation of this method on a mobile platform. The key obstacle to this is the accurate computation of the change in blur from frame to frame. The averaging approach used above would be unlikely to produce accurate results in an un-controlled environment and would need to be replaced with some form of object tracking. While a discussion on the various means of achieving this is beyond the scope of this paper, we envisage a less distributed value for blur could be obtained by using only the object edges.

Upon actual implementation it would also be necessary to consider looming as a control signal. In this paper we did not consider looming for variable speeds, the looming value that would cue avoidance is based on the change in range (see (1)). Thus this rate of change (speed) would either have to be known *a priori* or used as a control input. Alternatively, looming could be used to find range directly, as in [4].

7 CONCLUSIONS

The results clearly show that there is a measurable increase in looming values as range decreases. For lenses with a short focal length (long depth of focus), this is not closely related to the true value of looming as computed from range. However for a 25mm lens, this (when filtered) closely adheres to the theoretical curve. We believe that this is due to the much reduced depth of focus in this lens, although it is not known if all long focal length lenses would produce the same result.

These values were obtained without performing camera calibration and by manually focusing the camera lens at a known distance. The results, although noisy, illustrate that optical blur could be used as an effective avoidance cue for mobile machines.

REFERENCES

[1] E. R. Davie, *Machine vision, theory, algorithms, practicalities*. San Francisco, Ca.: Elsevier, 3rd edition, 2006, pp. 445-482, 505-516.

[2] G. Bradski and A. Kaehler, *Learning Open CV, computer vision with the Open CV library*. O'Reilly Publishing, 2008, pp. 322-337.

[3] R. Jain, R. Kasturi, B. G. Schunck, *Machine vision*. Singapore: McGraw-Hill Press, 1995, pp. 251-253, 289-305.

[4] E. Sahin and P. Gaudiano, Proceedings of intelligent control (ISIC), held jointly with IEEE International symposium on computational intelligence in robotics and automation (CIRA), intelligent systems and semiotics (ISAS), Gaithersburg, MD, USA, 1998, pp. 370-375.

[5] D. Raviv, K. Joarder, "The visual looming navigation cue: a unified approach", *Computer Vision and Image Understanding*, vol. 79, no. 3, pp. 331-363, 2000.

[6] A. Tsalatsanis, K. Valavanis, N. Tsourveloudis, "Mobile robot navigation using sonar and range measurements from uncalibrated cameras", *Journal of Intelligent & Robotic Systems*, vol. 48, no. 2, pp. 253-284, 2007.

[7] V. Kumar, M. Gavrilova, C. Tan, P. L. Ecuyer (Eds.), Computational Science and Its Applications, Springer, Berlin/Heidelberg, Germany, 2003, vol. 2669, p. 982.

[8] P. Mouroulis and J. Macdonald. *Geometrical optics and optical design*. Oxford Series in Optical and Imaging Sciences. New York, NY: Oxford University Press, 1997, pp. 94-108.

[9] E. Hecht and A. Zajac, *Optics*. World Student, Addison-Wesley Press, 1974, pp. 99-175.

[10] M. Subbarao and N. Gurumoorthy, Proceedings of Computer Society Conference on Computer Vision and Pattern Recognition, Ann Arbor, Michigan, USA, 1988, pp. 498 -503.

[11] H. Hu and G. D. Haan, Proceedings of the IEEE International Conference on Image Processing, Atlanta, Georgia, USA, 2006, pp. 617-620.