A Robustness Analysis of Triangulation-Based Robot Self-Positioning

Claus B. Madsen* **& Claus S. Andersen**** **& Jens S. Sørensen**

Laboratory of Image Analysis
Aalborg University
Fr. Bajers Vej 7D1
DK-9220 Aalborg East
Denmark

Abstract. A mobile robot can identify its own position relative to a global environment model by using triangulation based on three landmarks in the environment. It is shown that this procedure may be very sensitive to noise depending on spatial landmark configuration, and relative position between robot and landmarks. A general analysis is presented which permits prediction of the uncertainty in the triangulated position.
In addition an algorithm is presented for automatic selection of optimal landmarks. This algorithm enables a robot to continuously base its position computation on the set of available landmarks, which provides the least noise sensitive position estimate. It is demonstrated that using this algorithm can result in more than one order of magnitude reduction in uncertainty.

1 Introduction

There are basically two approaches to mobile robot navigation: the behaviour based and the model based approaches. Behaviour based, reactive approaches involve little or no global planning and for example enables the robot to move along a corridor essentially balancing the optical flow on both sides of the robot, [9, 5]. The model based approaches involve some level of geometric model of the environment, either built into the system in advance, acquired using sensory information during movement or a combination of both.

Model based mobile robot navigation systems generally consist of three major elements: 1) a path planning module, 2) a (possibly behaviour based) ability to use sensory information for avoiding obstacles not present in the model, and 3) a self-positioning module. The role of the latter is to continuously compute the position of the robot relative to the environment model, in order to circumvent the reliability problems associated with dead-reckoning, i.e., errors in wheel movement feedback due to slippage etc.

There are numerous approaches to self-positioning, and there are different sensor modalities as well (vision, laser range finders, ultrasonic sonars). In this context we shall only consider vision.

If a 3D model of the environment is available pose estimation techniques can be employed to single images in order to compute robot position relative to the model, [4]. Another class of techniques use landmarks. Landmarks are structures or markings in the environment, which can be recognized in images, and for which the positions in the world are known. Carlsson [3] demonstrated that vertical lines in the environment could be used in sets of four to provide cross-ratios, such that permutations of line sets could vote for a robot position.

A more typical use of landmarks involves triangulation, [2, 1]. Given known focal length and a single image of three landmarks it is possible to compute the angular separation between the lines of sight to the landmarks. If the world positions of the landmarks are known the angular separations can be used to compute the robot position and heading relative to a 2D floor map. The simplicity of this approach, and the fact that it does not involve any 3D reconstruction, has made it popular. At least three landmarks are required in order to compute the position, but more can also be used as for example in citegreinerscm96.

This paper studies the use of exactly three landmarks (a landmark triplet) for triangulation. The focus is on *how accurately* the robot position can be estimated using this approach. The sensory input to the

* Email: cbm@vision.auc.dk
** Email: csa@vision.auc.dk

triangulation method is the location in the image of three landmarks. The output is robot position and heading. The paper presents an analysis of how inaccuracies in locating landmarks in images propagate to uncertainties in robot position and heading. Furthermore, an algorithm is presented for automatically selecting an optimal landmark triplet, allowing the robot to minimize the positional uncertainty.

Related work by Crowley presents a thorough review of tools for dealing with uncertainty in mobile robotics, [6]. Crowley presents a Kalman filter framework for up-dating robot position using measured distance and heading to a *single* landmark. The distance and heading can for example be obtained using a stereo camera setup. For use in the Kalman filter Crowley derives formulae for how uncertainty in sensory data influences the position up-date (via the Kalman filter). It must be noted though, that robot position cannot be *computed* from distance and heading to a single landmark, only up-dated.

As described above the present paper studies the use of three landmark headings (no distance information), which is the minimum amount of information allowing for a complete computation of position.

The paper is organized as follows. Section 2 gives an overview of a navigation system context and the motivation for the work presented in this paper, and through an example demonstrates the potential noise sensitivity problems with triangulation. Section 3 outlines the triangulation framework. In section 4 we describe how the uncertainty in recovered robot position is studied and show how much the uncertainty can vary for different landmark triplets and robot positions. Additionally an algorithm for automatic landmark triplet selection is presented. Section 5 then presents some experiments demonstrating the effect of using this selection algorithm as opposed to more naïve choices of landmarks.

2 System context and motivation for work

The physical context for the problem addressed in this paper is a mobile robot moving on a planar surface, equipped with a single camera that can pan freely 360° relative to the robot heading. The pan angle relative to the robot heading is obtained from an encoder on the pan motor. It is assumed that a 2D floor map of the environment is available, and that positions of all landmarks in this floor map are known. Additionally the camera focal length must be known in order to compute the angular separation between landmarks. Lastly, the algorithms and analyses presented in this paper are based on all three landmarks being visible in a single image, so that robot position can be computed from a single image[3].

Landmarks, or beacons as they are sometimes referred to, can be any detectable structure in the physical environment. Some use vertical lines, others use specially designed markers, e.g., crosses or patterns of concentric circles. A popular choice is naturally occurring, planar textures such as door signs, posters or light switches. These patterns can be stored in system memory as sub-images and detected in images using template matching techniques. This is illustrated in figure 1 and 2.

In addition to the floor map containing positions of landmarks this approach requires a representative image of each landmark to be stored in a data base. Given a reasonable estimate of the robot position and heading these model images can then be perspectively transformed or re-mapped to how they should appear in the image. See figure 1. The re-mapped landmark images are then used as templates in a normalized cross-correlation procedure in order to detect the landmarks in the incoming image, figure 2.

As explained above the detected image location of the landmarks can then be used in a triangulation procedure for computing the camera position and heading, which in turn can be converted to robot position and heading.

Naturally, when using real data the detection of landmark image locations will be error prone. What effect do these errors have on the computed position? It is this question this paper attempts to answer. The effect can be visually illustrated using the following simulated example.

Figure 3 shows a simulation of robot self-positioning relative to a floor map. The "diagonal" string of white dots are robot positions on a path. The figure shows a snapshot from a particular position along this path. To simulate the positional uncertainty accruing from inaccuracies in determining landmark image locations, noise has been added to the synthetically generated locations . This noise is equivalent to a standard deviation of 5 pixels. The virtual camera has a resolution of 512×512 pixels, and a field of view of

[3] The results apply directly to situations where the camera may have to pan to obtain visual contact with all three landmarks, but in order for the triangulation to function, the robot will have to stop while "shooting" the directions of landmarks. If all three landmarks are in the field of view, the robot can move while taking images, as long as motion blur is not a problem.

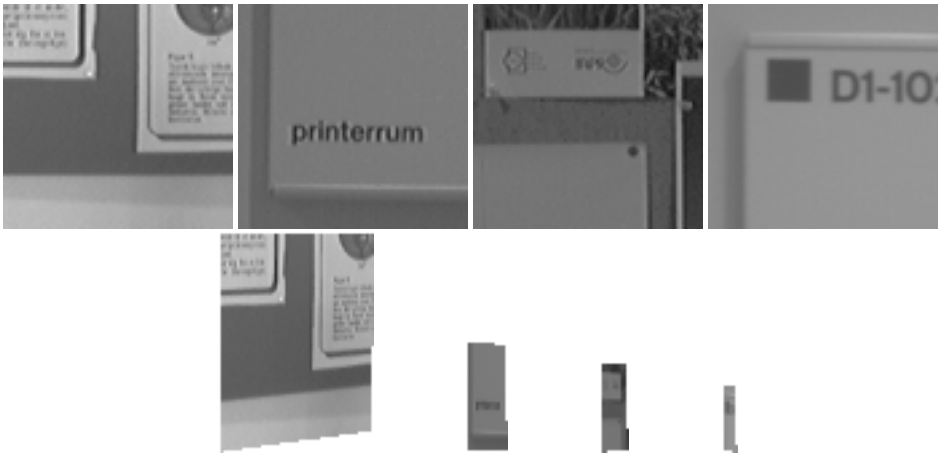

Fig. 1.: **Top row**: stored pixel data of 4 landmarks. **Bottom row**: Landmark templates obtained from perspective re-mapping of the stored pixel data shown in figure 1. The re-mapping is based on an estimated camera position corresponding to the hallway view in figure 2.

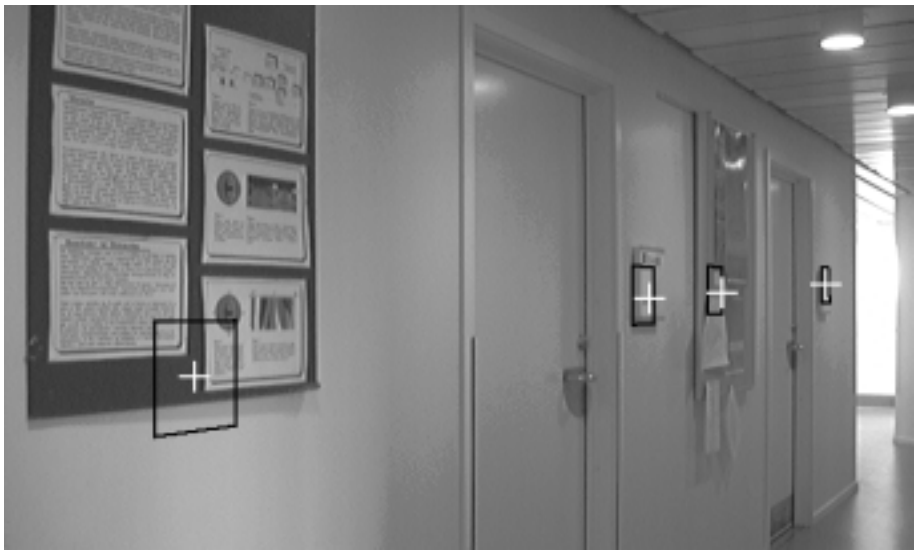

Fig. 2.: Locations of four landmarks detected using normalized cross-correlation after performing perspective re-mapping of the stored landmark pixel data, (figure 1). The white crosses are the loci of maximum correlation coefficient.

$\approx 53°$, (corresponding to a focal length of 6 mm for a 6 mm CCD chip). 1000 robot positions were computed from the noisy data and the positions are shown as the cloud of white dots.

The ellipse around the computed robot positions represents the positional uncertainty as predicted by using the analysis presented in this paper. It is seen that the uncertainty is very high. If the floor map represented a 15m $\times$ 15m room, then the computed positions cover a 2m $\times$ 1m area. Such an uncertainty is prohibitive for navigation in cluttered environments and narrow passages like doorways.

Figure 4 shows that the noise sensitivity varies drastically with the choice of landmarks, though. I.e., subject to the same amount of noise in landmark location detection, the uncertainty in recovered position differs for different landmark triplets. This clearly demonstrates the need for a strategy for selecting good landmark triplets, which in turn requires a technique for evaluating the possible landmark triplets before using them for triangulation.

The presented approach for *predicting* positional uncertainty makes it possible to design a simple strategy for choosing the best landmark triplet, if more triplets are available. That is, at any given position the navigation system can use the floor map to find the set of landmarks that should be visible. These visible

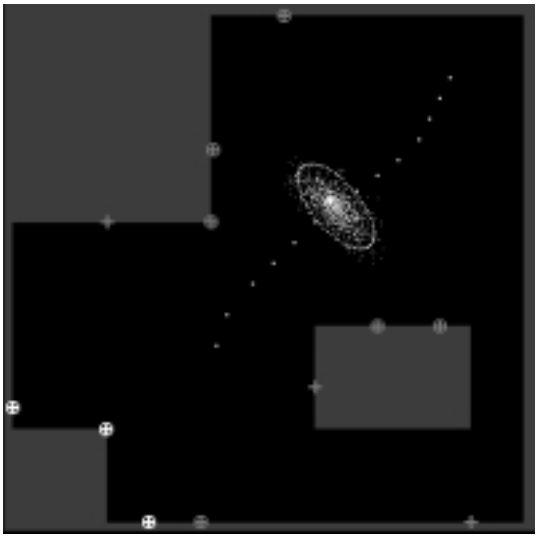

Fig. 3.: 2D floor map; obstacles are shown in grey, all visible landmarks are shown as circles with crosses, non-visible landmarks are shown as crosses. The landmarks used for computing position are shown as white circles with crosses in the bottom left corner. The string of white dots is a trajectory for the robot, and the cloud of white points are robot positions computed with noise added to the sensory measurements (see text). The ellipse is the positional uncertainty predicted by using the analysis presented in this paper.

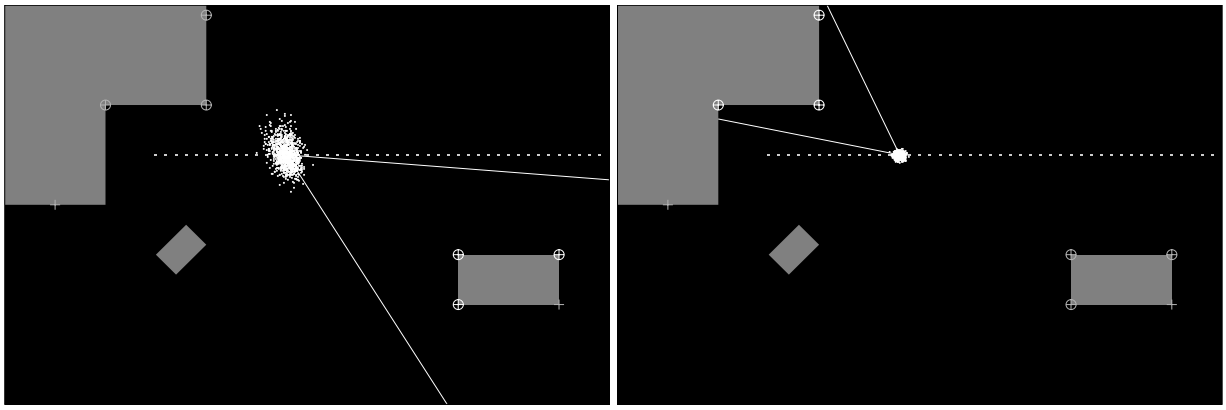

Fig. 4.: Variation in noise sensitivity when choosing different landmark triplets. The right triplet provides a much better position estimate than that in the left figure. The white line emanating from the robot position illustrate the field of view.

landmarks can be combined in various ways to form landmark triplets, and each of these triplets can be evaluated with regard to how severely noise would influence the computed robot position. The system can then choose the landmark triplet which minimizes the noise sensitivity. As will be demonstrated in section 5 this algorithm may result in an uncertainty reduction of more than one order of magnitude.

The noise sensitivity analysis and the triplet selection algorithm will be presented in section 4, but first a brief, slightly more formal introduction to triangulation based self-positioning (and notation) is given.

3 Overview of the triangulation technique

Closed form solutions to robot position and heading given angular separation between three landmarks exist, [1]. Figure 5 illustrate the parameters involved. The triangulation method actually computes the position of the optical center of the camera and the direction of the optical axis relative to the world coordinate system. For notational simplicity the optical center is assumed to coincide with the robot's position. Similarly the direction of the optical axis can be converted to robot heading using knowledge of the camera pan angle, β,

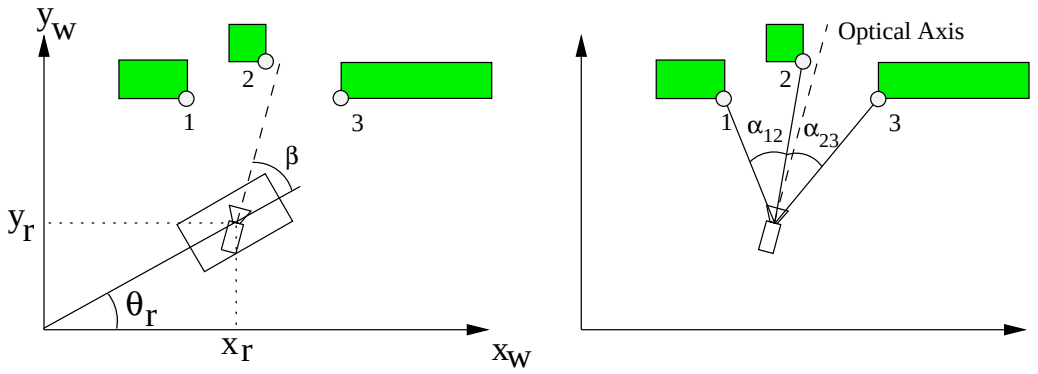

Fig. 5.: Left: robot position and heading relative to a floor map. Right: using image location of three landmarks to compute angular separation.

relative to the robot.

Let for example (u_1, v_1) be the image location of landmark number 1. If assuming the u-axis of the image coordinate system is parallel to the plane of motion of the robot, then the angular separations between pairs of landmarks are functions of the image locations $u_i, i = 1 \ldots 3$ and focal length, f:

$$\alpha_{12} = g(u_1, u_2, f) \qquad \alpha_{23} = g(u_2, u_3, f) \tag{1}$$

The robot position parameters, (x_r, y_r), are then functions of the angular separations and the positions of the landmarks in the world model:

$$x_r = h_x(\alpha_{12}, \alpha_{23}, x_1, y_1, x_2, y_2, x_3, y_3) \qquad y_r = h_y(\alpha_{12}, \alpha_{23}, x_1, y_1, x_2, y_2, x_3, y_3) \tag{2}$$

Similarly, the robot heading is a function of pan angle, the angular separations between landmarks and landmark world positions. Due to limited space the robot heading will not be considered further in this paper[4].

4 Robustness of self-positioning using triangulation

Subsequently the analysis of noise sensitivity is presented followed by the algorithm for purposively selecting landmark triplets.

4.1 Sources of noise and error

As presented in section 3 several parameters influence the result of the triangulation procedure. The camera's focal length, f, for instance must be very accurately calibrated. If not, it will cause a systematic error in the recovered positions. Similarly, if the world positions of landmarks, $(x_i, y_i), i = 1 \ldots 3$, are not absolutely correct in the model of the environment, this will also cause systematic errors. In this paper it will be assumed that focal length and landmark world positions are correct. Rather, the subsequent analysis focuses solely on the effect of uncertainty in the detection of landmark image location. I.e., errors in the $u_i, i = 1 \ldots 3$ parameters.

4.2 Sensitivity analysis

For the purpose of analysis the $u_i, i = 1 \ldots 3$ parameters shall be considered subjected to additive Gaussian noise of zero mean and a variance σ^2_{uu}. The Gaussian model is just a vehicle for a compact analysis. We are interested in the sensitivity towards *errors*, not in the particular distribution model. The algorithm for

[4] It is easy to demonstrate that the uncertainty in robot heading depends linearly on the uncertainty in locating landmarks in images. This paper demonstrates that this is far from true in the case of robot position, making it much more interesting to study this aspect

selection optimal landmark triplets minimizes this sensitivity, regardless of whether the Gaussian model is applicable.

The sensitivity analysis is based on first order covariance propagation, [7, 8], that is, propagating the variance σ_{uu}^2 to the 2×2 covariance matrix for robot position. The propagation is performed in two steps: first a propagation to the covariance matrix for the landmark angular separations, and then from the angular separation to robot position.

Let the covariance matrix of the three landmark image locations, $u_i, i = 1 \ldots 3$, be (assuming no correlation between the three landmarks):

$$\Delta = \begin{bmatrix} \sigma_{uu}^2 & 0 & 0 \\ 0 & \sigma_{uu}^2 & 0 \\ 0 & 0 & \sigma_{uu}^2 \end{bmatrix} \tag{3}$$

We introduce two vectors: $\mathbf{A} = [\alpha_{12}, \alpha_{23}]$ and $\mathbf{U} = [u_1, u_2, u_3]$. Under a first order approximation the covariance matrix, Υ, for the angular separations, α_{12} and α_{23}, becomes:

$$\Upsilon = \frac{\partial \mathbf{A}}{\partial \mathbf{U}} \Delta \frac{\partial \mathbf{A}}{\partial \mathbf{U}}^T \tag{4}$$

where $\frac{\partial \mathbf{A}}{\partial \mathbf{U}}$ is the 2×3 Jacobian matrix of the mapping from the u_i values to the angular separations.

In the second step, the uncertainty in angular separation as expressed by Υ can be propagated to the robot position parameters, $\mathbf{P} = [x_r, y_r]$:

$$\Sigma = \frac{\partial \mathbf{P}}{\partial \mathbf{A}} \Upsilon \frac{\partial \mathbf{P}}{\partial \mathbf{A}}^T \tag{5}$$

In Eq. (5) $\frac{\partial \mathbf{P}}{\partial \mathbf{A}}$ is the 2×2 Jacobian of the mapping from angular separation to robot position.

All 4 entries in Σ can be found analytically, thus for any configuration of 3 landmarks and any robot position it is simple to compute the expected positional uncertainty. The ellipse in figure 3 represents a contour of constant squared Mahalanobis distance from the average position. The particular contour chosen is the region within which 95% of the positional outcomes would fall. As seen the prediction of uncertainty matches the actual 1000 outcomes (the cloud of points) well. The area of this 95% uncertainty ellipse is a simple overall measure of the uncertainty, which shall be used subsequently.

4.3 Variation in noise sensitivity

As demonstrated in section 2 the uncertainty in triangulated position can vary substantially depending on the spatial layout of the landmark triplet and the overall distance from the robot to the landmarks. In order to get an overview of this three simulations have been made, figure 6.

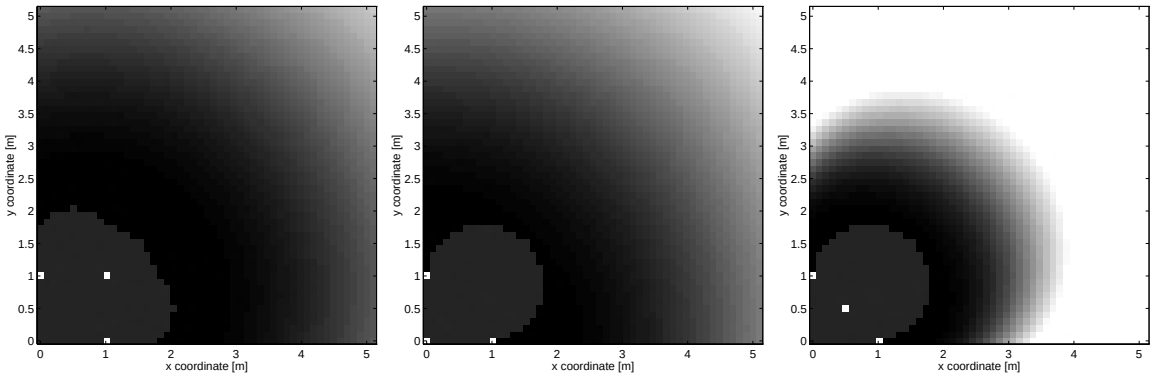

Fig. 6.: Area of uncertainty ellipse mapped to grey level values for three different landmark configurations: topology 1, topology 2, and collinear (left to right). The three white spots in the the lower left corner of each plot are the landmarks. The dark grey blob in the lower left corner is the area where not all three landmarks can fit within field of view and triangulation is not possible.

The three simulations correspond to three different generic spatial landmark layouts: a) topology 1, where the middle landmark is *in front of* the line joining the other two, b) topology 2, where the middle landmark is *behind* the line joining the other two, and finally c) a collinear landmark layout.

From figure 6 it is seen that generally the uncertainty increases with the distance to the landmark triplet. Furthermore, topology 1 triplet result in less uncertainty than topology 2, and collinear landmarks perform worst of the three. As a specific example: if the robot is located at position (2.5 m, 3.0 m), the area of the uncertainty ellipse is 0.2m°, 0.6m°, and 3.5m° respectively. Thus, use of collinear landmarks should generally be avoided, if alternative triplets are available.

The presented ability to predict uncertainty will be exploited in order to design an algorithm for automatic selection of landmark triplets.

4.4 Automatic selection of optimal landmark triplets

Assuming that triangulation based self-positioning is fairly accurate, and that the robot does not move large distances between each time it checks its position, then the last position found by triangulation plus information from odometry (dead reckoning) gives a good estimate of current robot position. From this estimate the navigation system can compute which landmarks in the environment model should be visible from the current position. Using the known focal length (field of view) it is also possible to determine all permutations of landmark triplets that are within field of view for various pan angles.

If a single "goodness measure" can be assigned to any such triplet based on the uncertainty analysis from section 4.2, then a simple algorithm for selecting optimal landmark triplets can be listed as:

1. Determine the set of visible landmarks using the current position estimate
2. Find all possible combinations (not permutations) of three visible landmarks (triplets)
3. Determine the set of landmark triplets that are within the same field of view
4. Compute a goodness measure for all such triplets
5. Rank triplets according to goodness measure and choose best

Many different goodness measures can be relevant for various purposes; a thorough investigation of this is a direction for future research. In the simulation experiments presented in the next section the applied measure is inversely proportional to the *area* of the uncertainty ellipse described previously. This measure simply results in a minimization of the region within which the computed positions will fall, when the landmark image locations are perturbed with noise.

5 Experimental results

The results described in this paper are currently being used in designing a navigation system for a real mobile robot. The system is based on 2D salient regions for landmarks, and using normalized cross correlation for recognizing and locating landmarks in images, as described in section 2. Landmarks are limited to planar texture patches on walls, e.g., light switches, door signs and posters. Thus, using a robot position estimate, known focal length and known position and orientation of the planar landmarks, a perspective distortion of landmark templates can be performed before using normalized cross correlation for detection. Promising results have been obtained so far with this approach, and 30×30 pixel landmark templates can be located in a 30×30 pixel search region in about 0.5 seconds on a regular SGI Indy workstation.

The implementation of this system is not finished so we resort to simulations to demonstrate how use of the presented landmark triplet selection algorithm can drastically reduce positional uncertainty.

5.1 Demonstrating the automatic landmark selection algorithm

The simulation is based on letting the robot move along the path shown in figure 3 and 7. At each position along the path the robot computes its position using triangulation, and uses the presented sensitivity analysis to plot positional uncertainty as ellipses. Two runs of the path were performed. First the camera was always forced to point in the direction of robot motion, and an arbitrary landmark triplet chosen within the resulting field of view. In the second run, the system was allowed to use camera pan to select the optimal landmark triplet for any position along the path. Figure 7 shows the computed uncertainty for a particular path point for each of the two runs, and it is seen how the automatic algorithm really reduces uncertainty.

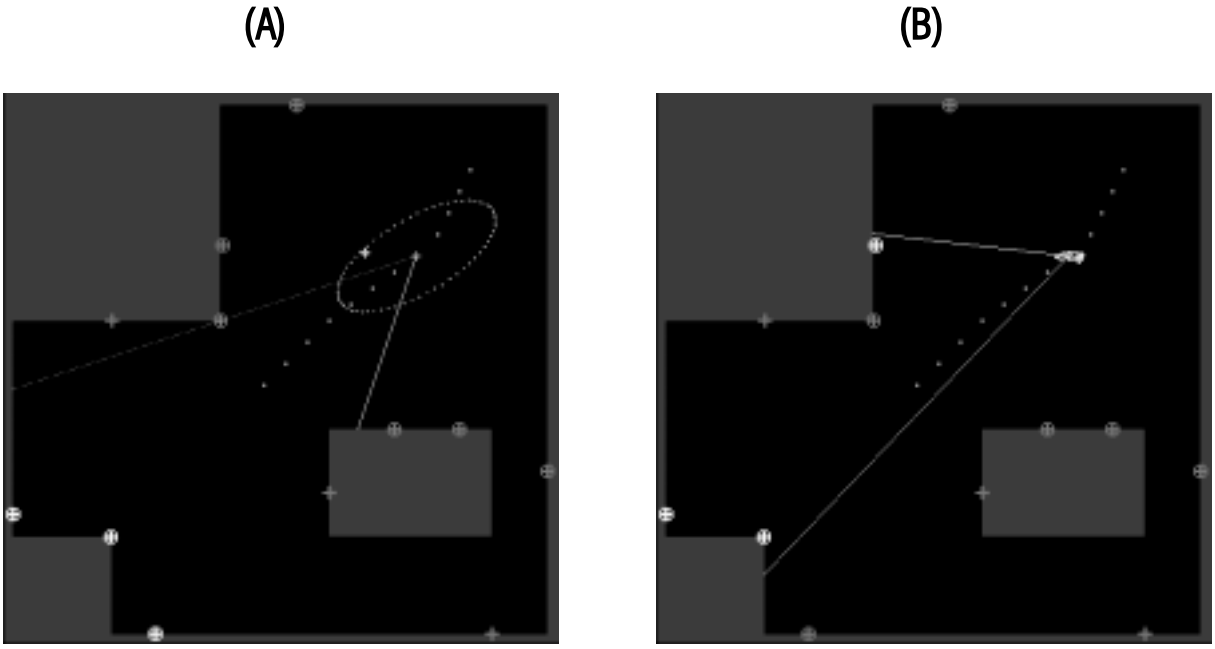

Fig. 7.: Left: position uncertainty when camera is forced to point in the direction of robot movement. Right: position uncertainty when using the automatic selection of optimal landmark triplet. Camera field of view is shown as two white lines emanating from robot position.

For the entire path the automatic landmark selection algorithm results in an average reduction in the area of the uncertainty ellipse of a factor of 6. If the shown floor map were a room of 15m × 15m the average ellipse area for the forced triplet run would be approx. 2m^2, (maximum $8\text{m}^{2!}$). Such an uncertainty makes model based navigation useless, whereas the automatic selection algorithm reduces the average area to approx. 0.3m^2 and the 8m^2 are reduced to 0.12m^2.

In the simulation experiments the standard deviation on landmark image location was set to 5 pixels. This is probably high compared to reality, but not unrealistic. Future experiments on real images aim at determining a suitable value of this standard deviation. Whatever the true standard deviation may be, the automatic landmark selection algorithm will still provide a substantial increase in positional accuracy[5].

5.2 Validity of first order approximation

In the sensitivity analysis we used a first order approximation to the change in robot position accruing from a small change in landmark image location. This is a very common assumption in error interval analysis, and in many situations it is valid. Never the less the formulae relating landmark image positions to robot position in many situations exhibit non-linear behaviour to such an extent that this first order assumption is violated. This is demonstrated in figure 8.

Figure 8 shows a worst case situation. Here the simulation system has been forced to choose the *worst* landmark triplet. The example demonstrates several things. First of all it shows that uncritical selection of landmarks can result in extremely noise sensitive position computations, which is also predicted by the system (the computed uncertainty ellipse areas are 24m° and 37m° respectively). Secondly, the example shows that the first order approximation around points in solution space is not always valid. This is seen from the 1000 noisy position computations, since they do not form an elliptic cloud, but rather lie on a curve. This curve is in fact a part of a horopter curve from one landmark to another, similar to the horopter curve concept used in stereo vision.

Thus the first order approximation does allow the system to predict that a certain landmark triplet will result in very high sensitivity to noise, but the computed uncertainty ellipse should not uncritically be taken as a indication of the distribution of the computed robot positions under noise.

[5] The ellipse area increases with the square of the standard deviation for the $u_i, i = 1 \ldots 3$ parameters. Thus halving the noise in landmark image location results in reducing ellipse area by a factor of four.

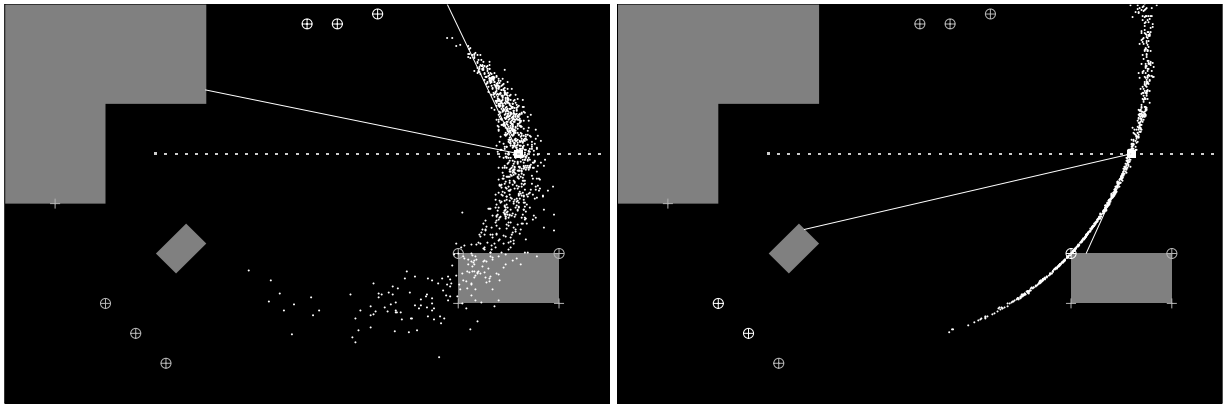

Fig. 8.: Examples showing very poor landmark triplets resulting in triangulation giving basically no knowledge about position. This simulation is subject to the same parameters is all other in this paper, i.e., 53° field of view and a noise std. dev. of 5 pixels.

6 Conclusion

It has been demonstrated how sensitivity analysis can be applied to increasing the robustness of self-positioning using triangulation. The analysis is based on propagating the effect of errors in locating landmarks in images to the resulting errors in computed robot position.

The analysis allowed the design of an algorithm for automatically selecting triplets of landmarks, which minimize positional uncertainty. It was demonstrated that using this approach the positional uncertainty can be reduced with an order of magnitude.

Generally positional uncertainty increases monotonically with the distance from robot to landmark triplet. A specific result of the presented analysis is that landmark triplets, where the landmark positions are collinear provide very noise sensitive position estimates, and are to be avoided.

Future directions of research may include using such robustness analysis in the path planning stage. [10] present a general framework for including sensory uncertainty into the path planning algorithm. By combining this with results from this paper it would be possible to urge the robot the follow paths which allow the most accurate control of position.

7 Acknowledgments

This work was funded in part by the VIRGO research network (EC Contract No. ERBFMRX-CT96-0049) under the TMR Programme, and in part by the SMART-I research network (EC Contract No. ERBCHRX-CT92-0038), under the HCM Programme. This funding is gratefully acknowledged.

References

1. C. S. Andersen and J.G.M. Conçalves. Determining the pose of a mobile robot using triangulation: a vision based approach. Technical Report Tech. Note No. I.95.179, EU Joint Research Center, December 1995.
2. S. Atiya and G. Hager. Real-time vision-based robot localization. *IEEE Transactions on Robotics and Automation*, 9(6):785 – 800, December 1993.
3. S. Carlsson. Relative positioning from model indexing. *Image and Vision Computing*, 12(3):179 – 186, April 1994.
4. H.I. Christensen, N.O. Kirkeby, S. Kristensen, L. Knudsen, and E. Granum. Model-driven vision for in-door navigation. *Robotics and Autonomous Systems*, (12):199 – 207, 1994.
5. D. Coombs and K. Roberts. Centering behavior using peripheral vision. In *Proceedings: IEEE Conference on Computer Vision and Pattern Recognition*, pages 440–445, New York City, New York, USA, June 1993. IEEE, IEEE Press.
6. J. L. Crowley. Mathematical foundations of navigation an perception for an autonomous mobile robot. In L. Dorst, editor, *Reasoning with Uncertainty in Robotics*. Springer Verlag, 1996.

7. R.M. Haralick. Propagating covariance in computer vision. In *Proceedings: 12'th International Conference on Pattern Recognition, Jerusalem, Israel*, pages 493 – 498, October 1994.
8. C. B. Madsen. A comparative study of the robustness of two pose estimation techniques. *Machine Vision and Applications, special issue on performance characterization*, 5/6(9):291 – 303, 1997.
9. J. Santos-Victor, G. Sandini, F. Curotto, and S. Garibaldi. Divergent stereo for robot navigation: Learning from bees. In *Proceedings: IEEE Conference on Computer Vision and Pattern Recognition*, pages 434–439, New York City, New York, USA, June 1993. IEEE, IEEE Press.
10. H. Takeda, C. Facchinetti, and J.-C. Latombe. Planning the motions of a mobile robot in a sensory uncertainty field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(10):1002 – 1017, October 1994.

This article was processed using the TEX macro package with SIRS97 style