On Motion Estimation for a Mobile Robot Navigating in Natural Environment: Matching Laser Range Measurements using the Distance Transform

Per Ljunggren Klöör, Åke Wernersson
Robotics/Autonomous Mechanical Systems
Linköping University, S-581 83 Linköping, Sweden
and FOA 37, S-581 11 Linköping, Sweden

Abstract: *The goal behind this paper is to find a generic method for controlling the motion of a robot relative to an object of an arbitrary shape. In this paper we study;*
- *modelling laser range measurements for different type of objects/surface properties. Outdoor scenes are emphasized.*
- *testing the distance transform on measurements for estimating the motion of a robot relative to an object of arbitrary shape.*

Emphasis in the present study is to get experience of the error mechanism in the distance transform when tested on cluttered laser measurements. Both natural and man-made objects are used and compared in the tests.

Keywords: Range measuring laser, relative positioning, navigation, motion estimation, distance transform, robotics.

1. Introduction

The generic problem addressed is to control the motion of a robot in a cluttered environment. Experiments are made with a pulsed laser having a typical range resolution of 5 cm and a beam width equal to 2 mrad. A 360° scan rotation is obtained by a stepping motor driven mirror. The maximum resolution gives 3200 range measurements for a complete rotation of the beam.

In this papers it is tested, using measurements, how the distance transform can be used to estimate the robot's change in posture (position and orientation) between two measurements. Typical tasks are; controlling the robot to a prescribed position relative to an object, accurate navigation in a cluttered non-stationary scene, detecting and estimating changes in the scene, station keeping, etc.
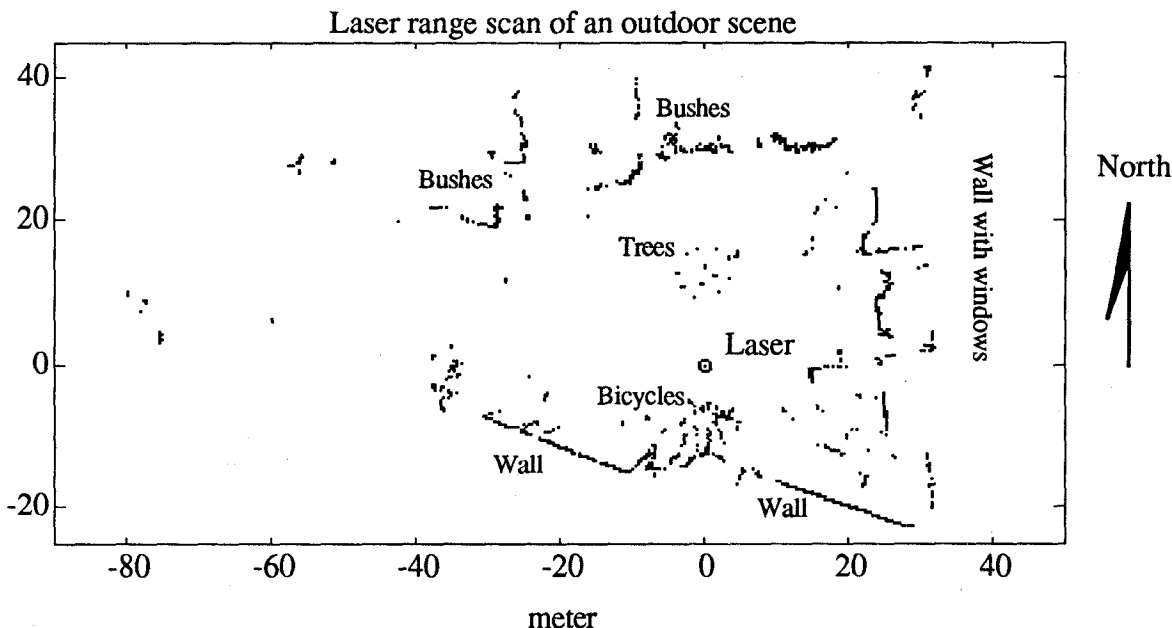

Fig. 1. This range scan with 3200 measurements is from a yard outside a building. The irregular objects at the top are bushes and in the middle of the yard there are 13 small trees in a circle. There are bicycles parked in front of the walls and to the right the laser has measured through the windows. One problem studied is how the distance transform can be used to estimate the robot's change in posture (position and orientation) between two measurements at different locations.

In the accompanying paper [1] a general feedback law is designed for controlling the gripping of an object. The control law was designed to take the robot to a specified posture and within prescribed tolerances. The control law is dependent on the covariance matrix of the error in the posture estimate. For a simple shape it is fairly straight forward to model this estimation error. *If* we can *model* the *estimation error* the distance transform seems to give a *generic method for controlling a robot* using measurements from a range sensor. One goal of the *present study* is to get experience of the error mechanism in the distance transform when *tested on cluttered laser measurements*. In this paper we study the opposite question; How complicated can the objects be and how large can the motion be 'before the distance transform breaks down'?

2. Object Classes and Signal Models

A typical outdoor scene is plotted in fig 1. Compared with the map/drawing the scene is very cluttered due to bicycles, lamps, flag poles, a parked car, a few transparent windows etc. There are also natural objects; the irregular bushes and the trunks of 13 small trees. This outdoor scene is used as a test area in this study.

2.1 Type of objects

In the range scan in fig 1 there are a number of different object types. The model used for signal modelling should take into account these object types and must also cope with the different types of noise that occur in the measurements.

When using range cameras and laser one should consider at least six different types of objects/surface properties. The list describes the geometry and can be made useful for outdoor vehicles as well as for industrial robots and for AGV's:
1. Flat or smooth surfaces (buildings, walls, wood, blasted aluminium).
2. Small isolated objects i.e. a few pixels (tree trunks, poles, wires, legs of chairs, reflective tape).
3. Depth texture (bushes, grass, textiles, wool, foam, ...)
4. Transparent and semitransparent surfaces (windows, plastic, glass)
5. Reflecting surfaces (mirrors, wet smooth surfaces, polished steel)
6. Absorbing objects (mate dark surfaces, smoke, foam rubber,)

The first three types are the most frequent in the outdoor scene in fig 1. Several additional measurements are given in [4,5].

2.2 Models for disturbing objects and noise in the range measurements

Consider a scene with different objects numbered by i. When the robot is at the point (x,y) and looking in the direction ψ (global frame) the distance to object number i is R_i (x,y,ψ). The corresponding measurement, R_o, is the distance to the closest object i.e.

$$R_o(x,y,\psi) = \min_i \{R_i(x,y,\psi)\} \qquad (1)$$

In a real scene there are also objects present not contained in the á priori description. In fig 1 the bicycles in front of the walls and the transparent windows are considered as disturbances giving deviation D in the measured range. At time t_k, consider a robot at location (x_k,y_k) and with heading θ_k. In the coordinates of the robot the laser measures the range $r(\varphi_k,t_k)$ in the direction φ_k - cf. fig 2

$$r(\varphi_k, t_k) = \beta_1 [R_o(x_k, y_k, \varphi_k + \theta_k) + n_k] +$$
$$\beta_2 [R_o(x_k, y_k, \varphi_k + \theta_k) + D(x_k, y_k, \varphi_k + \theta_k) + n_k] +$$
$$\beta_{do} [R_{do}] \qquad (2)$$

where n_k is "ordinary white noise". The random binary (0 or 1) variables β describes if there is a normal measurements $\beta_1=1$, a disturbed measurement $\beta_2=1$ or a drop out $\beta_{do}=1$ with the normalization $\beta_1+ \beta_2 + \beta_{do}=1$. The first line in eq 2 is for normal measurements with 'electronic noise' n. The second line is for measurements on disturbing objects. D is the random size in range of the disturbing object. A bicycle in front of the walls gives D < 0 while a transparent window gives D > 0.
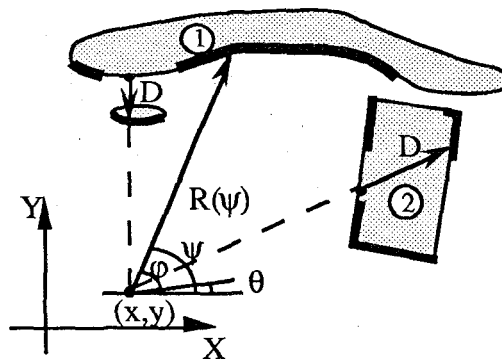

Figure 2. Principles for modeling range measurements in the presence of disturbing objects and transparent surfaces. The modelled range value R(x,y,ψ) in the direction ψ is the distance from (x,y) to the first object in the model class (two objects in this case). When unmodeled objects are present a disturbance D(ψ) is added as an outlier in the model.

Remarks:
- Equation 1 assumes implicitly "well behaved solid objects". For other types of objects, say, mirrors and a collection of bushes, more complicated models are needed.
- Equation 2 is also valid if there is a motion between the individual pulses, compare the navigation in [6].
- Too large distances can be described as R_0 = maximum range of laser or R0 = ∞ depending on the context.
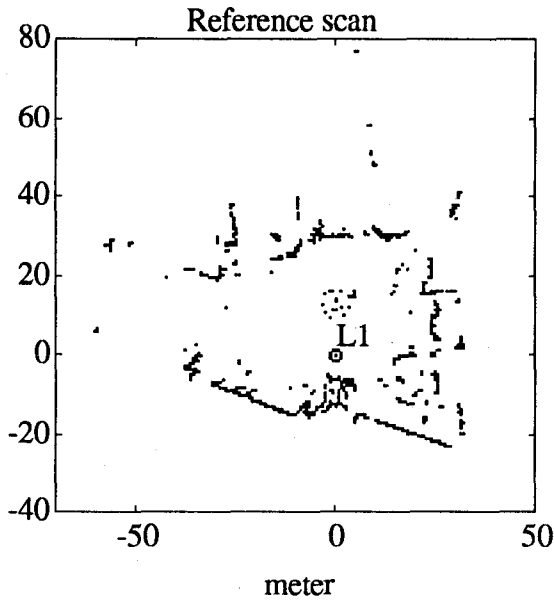
Figure 3A. A range scans with the laser at the initial position L1. This scan is used as a reference when estimating the difference in posture as the laser is moved to another position.
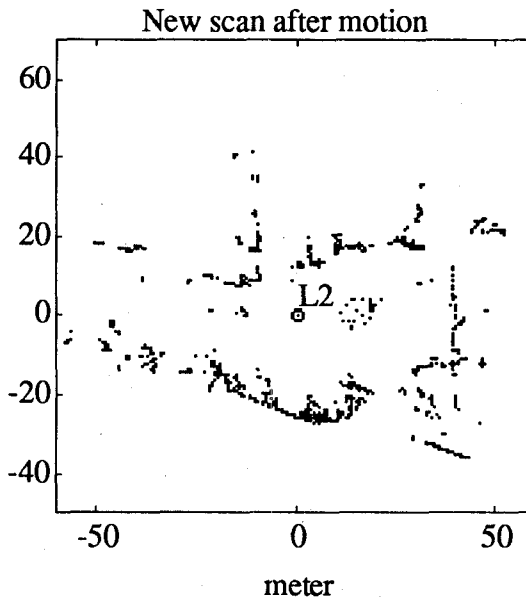
Figure 3B. The new range scans after motion to L2. In this scan a large part of the wall is occluded by the bikes, cf. fig 1. The matching method has to cope with this type of occlusions.

3. Matching Using the Distance Transform

In this section it is shown how the distance transform (DT) can be used for estimation of change of posture between measurements. Further it is shown how spatial filtering can suppress clutter and thus reduce the risk for false matchings. The goal is to estimate the motion of e.g. a mobile robot from range measurements. Here, only the case when there is no motion during the scanning (stop-move-stop) is tested, although the general model in Eq.2 is valid also for the dynamic case.

3.1 The method

In fig 3A,B two range scans taken at two different positions are shown. The measurements are originally one-dimensional in polar format (range as a function of angle), but here it is plotted in cartesian coordinates. In order to estimate the motion one should find the translation and rotation that makes the two scans coincide "the best".

Let (x,y,θ) be the position/orientation of the laser at position 2 expressed in the coordinate system of position 1. Consider point number n in the second scan. Let $d_n(x,y,\theta)$ be the distance between this point and the closest range point in the first (reference) scan. Euclidian distances and the model in eq 2 gives

$$d_n(x,y,\theta)^2 =$$

$$\min_k \{[r_{ref}(\varphi_k) \cos(\varphi_k) - r_{new}(\varphi_n) \cos(\varphi_n+\theta) - x]^2 +$$

$$[r_{ref}(\varphi_k) \sin(\varphi_k) - r_{new}(\varphi_n) \sin(\varphi_n+\theta) - y]^2 \} \quad (3)$$

where r_{ref} and r_{new} are the range measurements from the reference scan and from the new scan respectively. The estimated displacement is then the (x,y,θ) minimizing the criterion function

$$Q(x,y,\theta) = \frac{1}{N} \sum_{n=1}^{N} d_n(x,y,\theta) \quad (4)$$

where N is the total number of range measurements in the second scan.

Transforming the range measurements $r(\varphi_k,t_k)$ to cartesian coordinates $(x_k,y_k)^T$ eq 3 can be written

$$d_n(x,y,\theta) =$$

$$\min_k |[x_k\ y_k]^T_{ref} - Rot(\theta)\ [x_n\ y_n]^T_{new} - [x\ y]^T| \quad (3')$$

where $Rot(\theta)$ is the rotation matrix. The distance d_n can be calculated without minimisation if we precalculate a distance function who's values for an (x,y) in the reference frame is the distance to the closest range point. If we discretize the scans in cartesian coordinates to binary images it is quite straight foreword to calculate this distance function. In fig 4 the distance transform of the reference scene is shown. Fig 5 is the criterion function plotted as a function of x and y when the distances d_n are calculated by making a distance transform (DT) of the reference scan and matching the image of the new scan to the DT of the reference scan [2,3]. Each d_n can then be calculated simply by $d_n=DT(a,b)$, where (a,b) is the image coordinates corresponding to $r_{new}(\varphi_i,t_i)$. The criterion function can thus be written

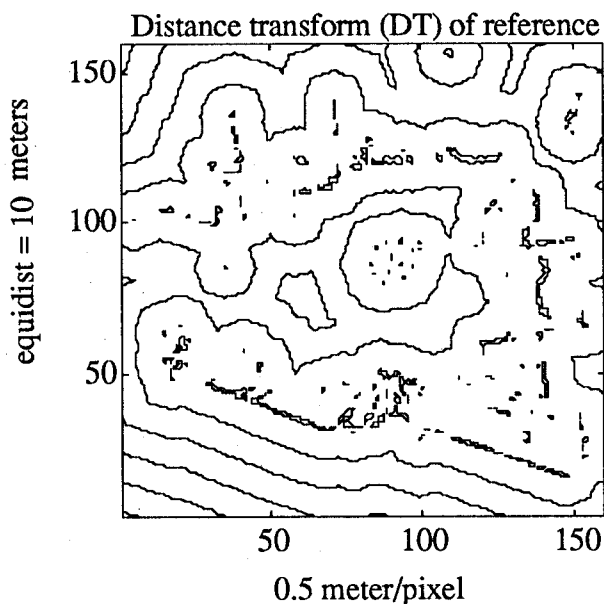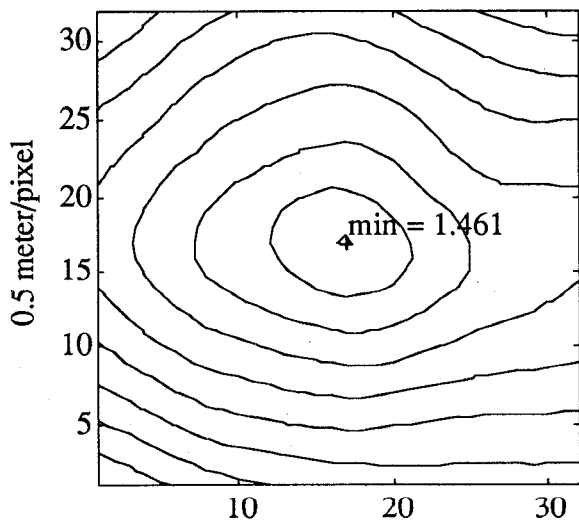**Distance transform (DT) of reference**

0.5 meter/pixel

Figure 4. The distance transform (DT) of the reference scan in fig 3A. In the DT each pixel value is the approximate Euclidian distance to the closest measured range value in the original range scan. The equidistance in the level curves is 10 pixels corresponding to 5 meters.

Crit. function $Q(x,y) = 1/n \sum M(i,j)*DT(i+x,j+y)$


Min at (17,17) => x=-15.7 m, y=11.5 m

Figure 5. Level curves of the criterion function Q(x,y) in the matching. The criterion function Q(x,y) is the average distance between the reference measurement points and the new measurement points. The estimated translation, given by the minimum of Q, is x=-15,7 m and y=11.5 m. Generally Q is a function of both translation and rotation, but here minimisation is only over translation.

$$Q(x,y) = \frac{1}{N} \sum_{j=1}^{m} \sum_{i=1}^{m} M(i,j)\, DT(i+x,j+y) \qquad (5)$$

where in this case N is the number of pixels corresponding to a range measurement in the image M of the second scan. $M(i,j)=1$ if (i,j) corresponds to a range point $r_{new}(\varphi_i, t_i)$ and $M(i,j)=0$ otherwise.

3.2 Test on the entire scene

In Fig 6 the DT of the reference and the new scan after motion are superimposed after having moved the new scan according to the matching results shown in Fig 5. The result is very reasonable - corect within the discretisation.

DT of reference (level curves) & new scan (+)
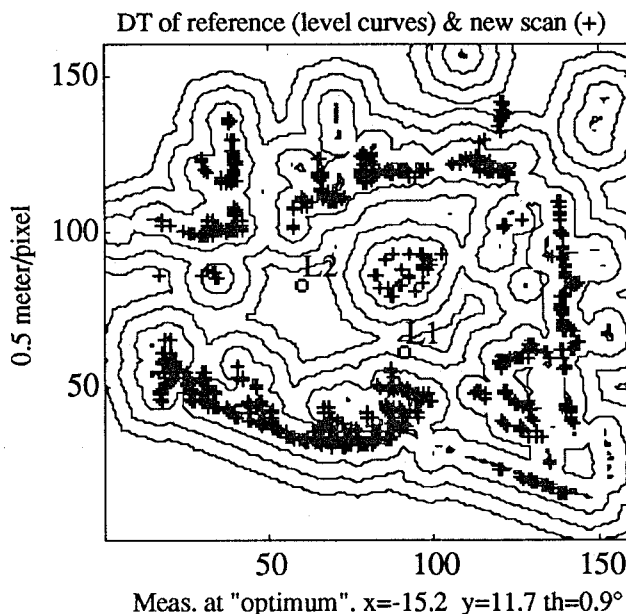
Meas. at "optimum". x=-15.2 y=11.7 th=0.9°

Figure 6. The minimum of the criterion function in fig 5 gives the estimated translation x=-15.7 meter, y=11.5 meter. Here the first scan gives (using the DT) level curves while the measurements of the new translated scan are plotted with + signs. Most of the differences are due to occlusions. The discretization in the images is 0.5 meter/pixel while the discretization in the range data from the laser range finder is 0.1 meter. The positions before and after motion are indicated with the circles at L1 and L2 respectively.

3.3 Clutter in front of man made objects

When the matching region is decreased clutter and disturbances becomes more crucial. As seen in Fig 7B the global minimum of the criterion function of the non-filtered estimation gave a motion estimate offsetted 1.8 meters from the "true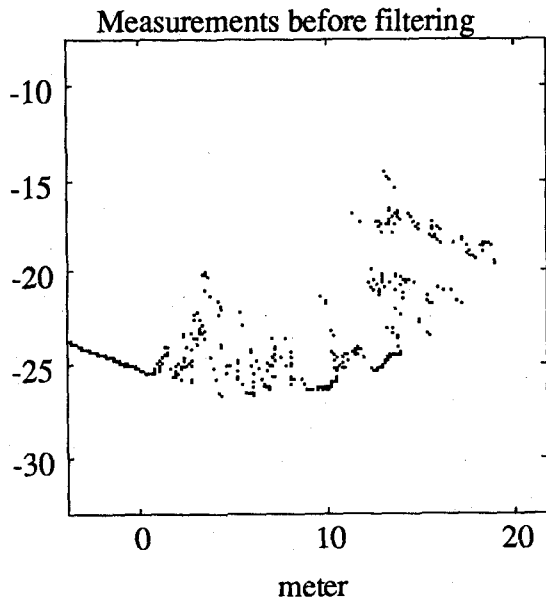 value". For specific object classes a range filtering can be made before the motion estimation. In the example shown in Fig 8 an order statistics filtering was used:

$$\bar{r}(\varphi_k) = Max[\ r(\varphi_{k-m}),\ \dots\ ,\ r(\varphi_k),\ \dots\ ,\ r(\varphi_{k+m})]$$

Both the reference data and the measurement data was filtered before matching, and in this case m=6. Teh estimate is now correct.

Measurements before filtering

Figure 7A. An outcut of fig 3B. The clutter in this area is due to bikes that are parked in front of the building.

Filtered measurements

Figure 8A. The data from fig 7A is filtered in its original polar (range as a fkn of angle) form with a quartile-like filter that enhances the "outer envelop", in this case the wall behind the bikes.

Criterion function Q, no filtering
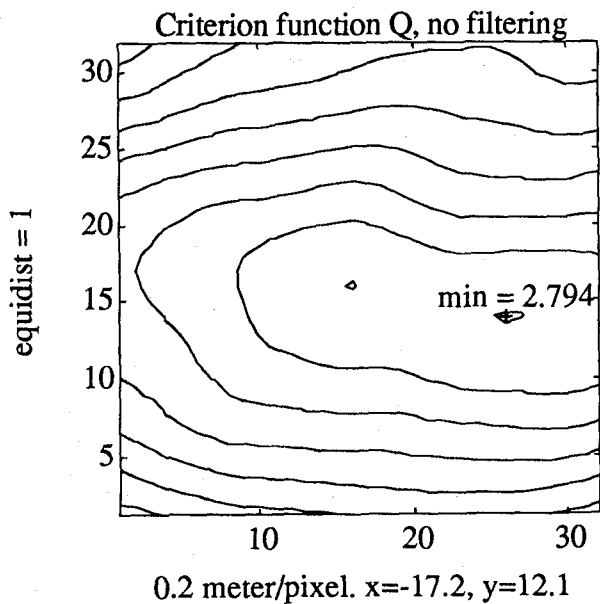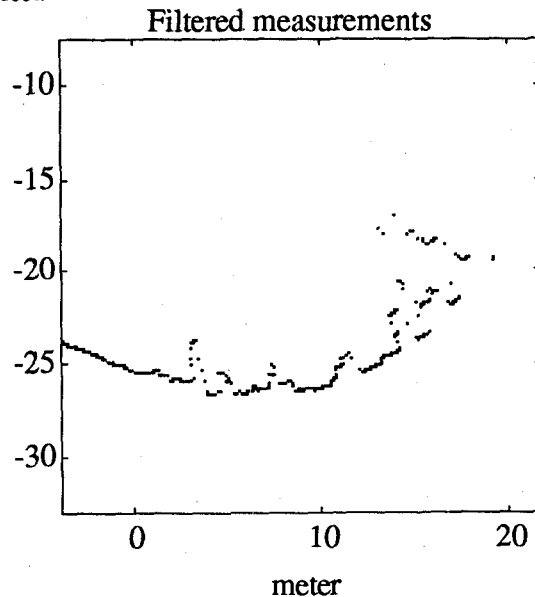

0.2 meter/pixel. x=-17.2, y=12.1

Figure 7B. When estimating the motion the bicycles in front of the the wall should be considered as noise that disturbs the estimate. Here the criterion function for the DT-motion estimate has an erroneous global minimum offsetted 1.8 meter from the "true value".

Criterion function Q, filtered meas.
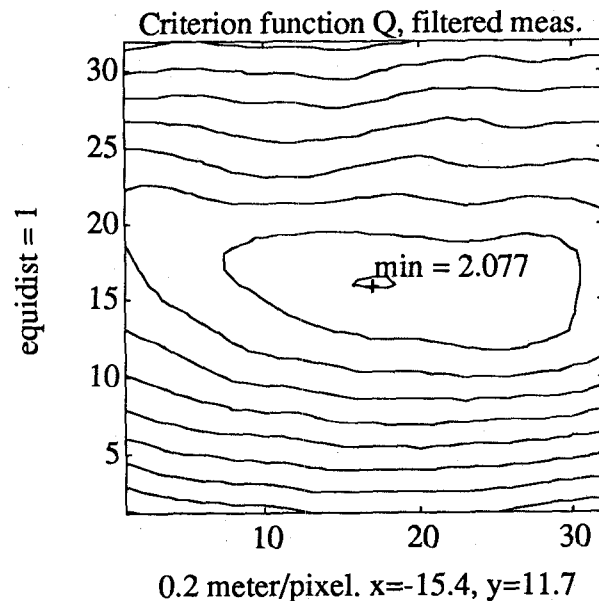

0.2 meter/pixel. x=-15.4, y=11.7

Figure 8B. The result of the estimation/matching using a filtered measurement and a filtered reference. The global minimum is the correct estimate. Also, Q (x,y) turn out to be less wrinkled. It should be observed that the filtering is made in polar coordinates (range as a function of angle) and not in the image.

1425

4 Test on Natural Objects

The estimation of change of posture using the distance transform does not require a parametrization of the objects. Scenes containing only natural objects can also be matched. In fig 9 the data from the trees and bushes are used to do an independent test of the matching method. The result is correct within the discretisation. Observe that no object type specific filtering is made in this heavily cluttered scene .

DT of bushes & new range measurements (+)

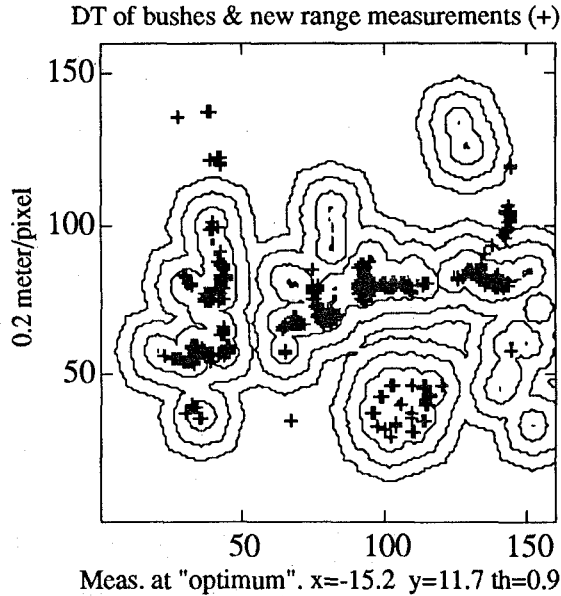Meas. at "optimum". x=-15.2 y=11.7 th=0.9°

Figure 9. Matching on only the part of fig 1 that contains only natural objects. The plotting method is identical to the one in fig 6. The estimate from this part is the same as for the entire scene.

4.1 Testing on the tree trunks

The trunks of the 13 small trees in the middle of fig 1 are almost ideal point objects. Testing the matching method on this part of the scene gives an estimate that is very close to the true displacement. The discrepancies are only due to the noise in the range measuring laser and the size of the tree trunks (deviation from ideal point objects).

The criterion function Q
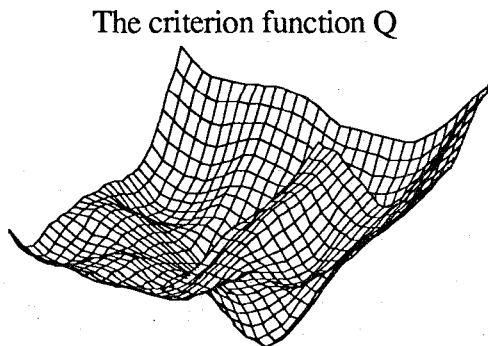
$$Q(x,y) = 1/n \sum M(i,j)*DT(i+x,j+y)$$

Figure 10 A. A 3D plot of Q from matching the trees.

Level curves of Q
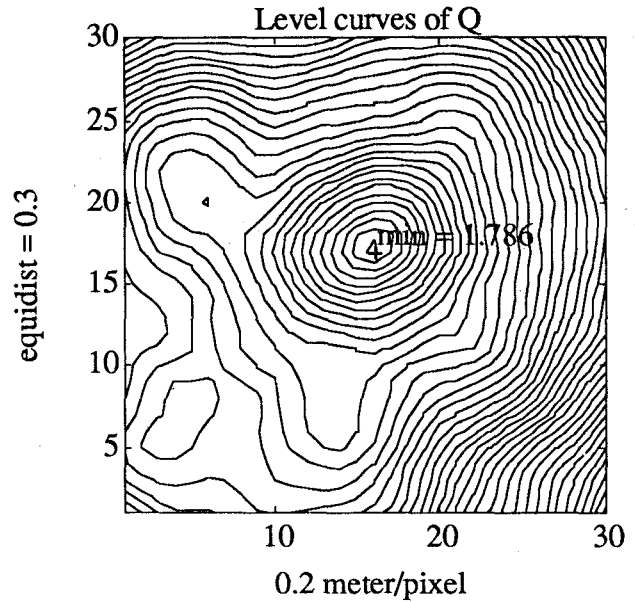
Figure 10 B. The criterion function from matching the trees.

DT of trees (level curves) & new measurements (+)
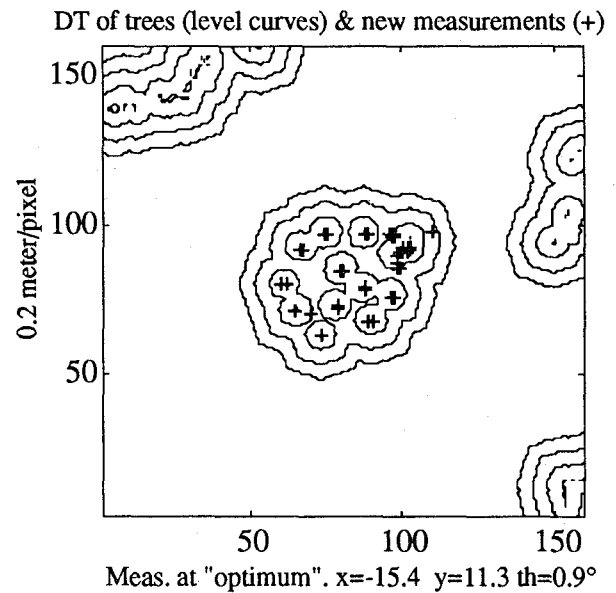
Meas. at "optimum". x=-15.4 y=11.3 th=0.9°

Figure 10 C. Ideally, when matching measurements of a single point object the criterion function is sharp and shaped as a cone. The object in this figure is 13 small trees in a circle. The level curves is the distance transform generated by one set of measurements while the + signs are the new measurements from a different position. The minimum is rounded off by noise. Also there are a number of local minima. The search algorithm to be used must be able to avoid local minima.

1426

4.2 Bushes

One of the stable features of hedge and bushes type of objects (depth texture) is the envelop. The envelop can be extracted by order statistics filtering. It should be observed that this is a spatial filtering in range and *not* an image filtering. In fig 11 the range measurements from the bushes before filtering is matched. Compared to the matching of the filtered measurements, fig 12, the global minimum of the criterion function is offsetted by 0.9 meters.
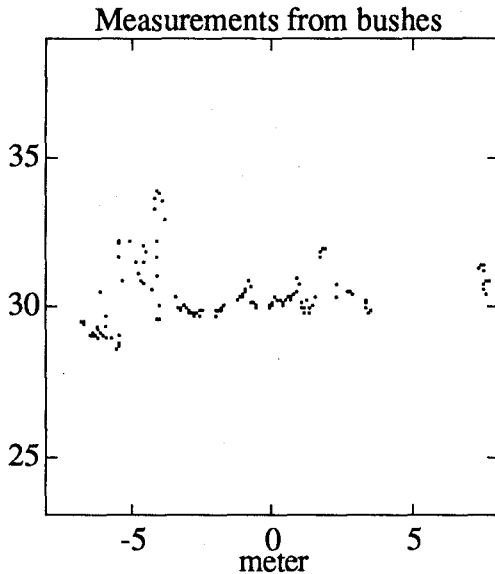
Measurements from bushes

Figure 11 A. An outcut from the scan in fig 3B containing the measurements from the bushes.

Q from matching bushes

0.1 m/pixel x=14.0 y=-12.5

Figure 11 B. Here the criterion function for the DT-motion estimate has an erroneous global minimum offsetted 0.9 meter from the "true value".
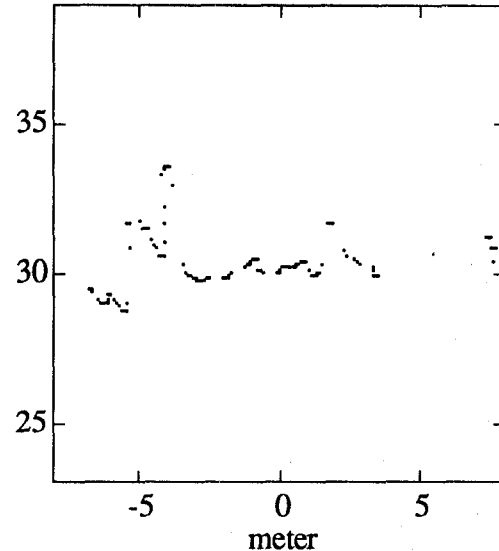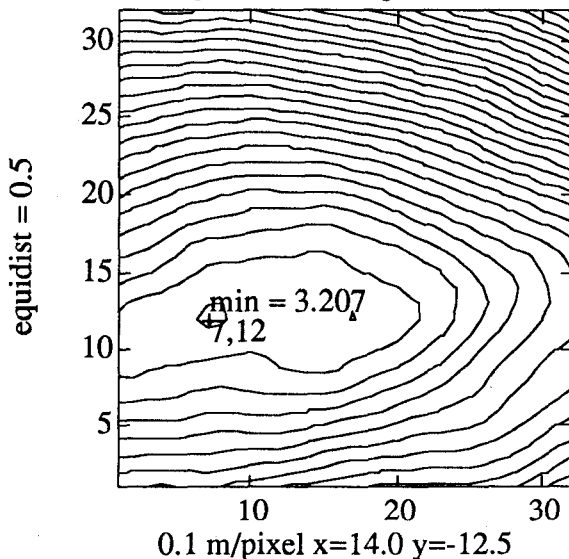
Filtered measurements from bushes

Figure 12 A. The range data from fig 11 A is filtered using order statistics in its original polar (range as a function of angle) format.

Q from matching filtered meas. of bushes
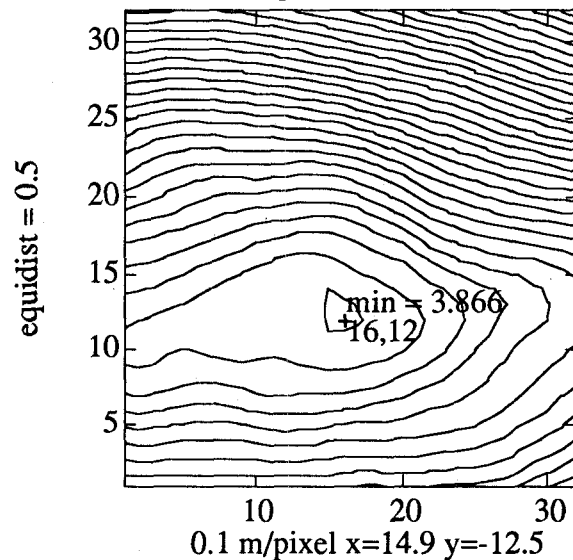

0.1 m/pixel x=14.9 y=-12.5

Figure 12 B. The result of the estimation/matching using a filtered measurement and a filtered reference. The global minimum is close to the correct estimate - cf. fig10. It should be observed that the filtering is made in polar coordinates (range as a function of angle) and not in the image.
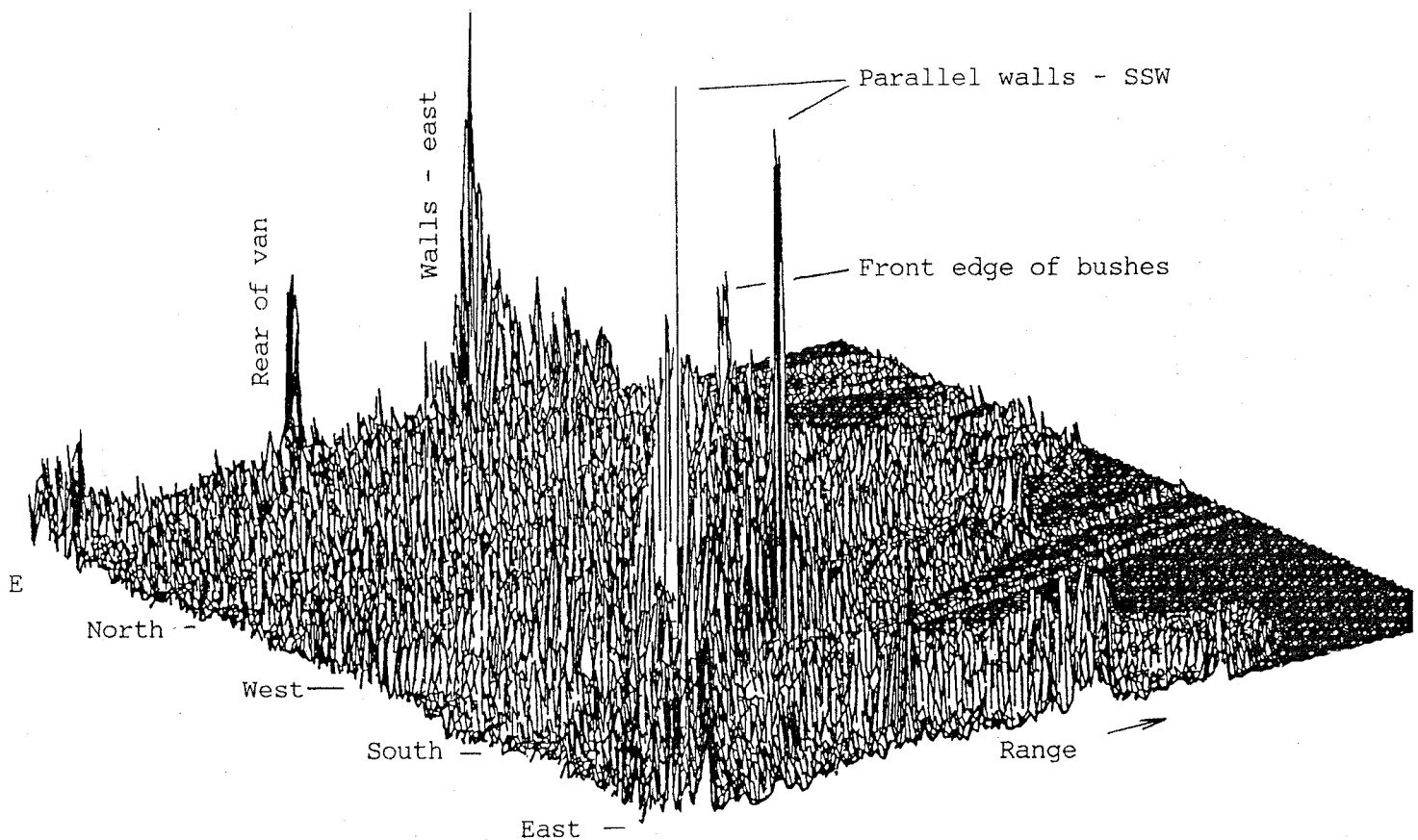
Figure 13. The Hough transform on the entire scene in fig 1. All the straight lines comes out as expected. Notice the large peak from the rear of the van parked close to the laser. Also the front edge of the bushes gives a peak.

5 The Hough Transform

The standard tool for finding straight lines is the Hough transform - fig.13 is a HT of the entire scene. With a range weighting the parallell walls (SSW) will give a smaller peak. The bushes gives a large noise carpet in direction North. It should also be mentioned that the HT have been used for robot navigation in a cluttered environment.

6 Conclusions

The long term goal behind this paper is to find generic methods for positioning a robot relative to objects of arbitrary shape. In this paper the distance transform was tested on laser range measurements in a composite scene with both natural and man-made objects. For small scenes, spatial filtering was found to be advantageous before matching. A preliminary summary of the findings are:

- To structure the problem model classes are introduced (based on geometrical shape, spatial texture and reflective properties).
- For hedge and bushes type of objects front edge envelops are found by order statistics filtering in range.
- Laser range measurements are not images.
- As expected, the Hough transform is efficient for finding straight walls in natural scenes.In this context a range image is 'cleaner' than a camera images.

References

[1] Å. Wernersson, B. Boberg, B. Nilsson, J. Nygårds, T. Rydberg, "On Gripping Using Range Cameras: Sensor Noise, Task Tolerances and Feedback", this conference.

[2] G. Borgefors, "Distance Transformations in Digital Images", in *Computer Vision, Graphics, and Image Processing*, 34, 344-371, 1986.

[3] G. Borgefors, "Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm", in *IEEE Trans. PAMI*, vol.10, no 6, Nov 1988.

[4] P. Klöör, A. Larsson, Å. Wernersson, "Relative Positioning Using a Range Measuring Laser", Symp. on Image Analysis, Stockholm, March 5-7 1991

[5] A. Larsson, "Pulsed Laser for Navigation: an Experimental Prestudy", FOA report C 30598 - 3.4, 1990, in Swedish.

[6] Å. Wernersson, U. Wiklund,U. Andersson, K. Hyyppä, "Vehicle Navigation using 'Image Information': on Association Errors", in Intelligent Autonomous Systems2, Amsterdam, Dec 1989.

[7] Forsberg J. ,Larsson U. ,Åhman P. ,Wernersson Å."On the Hough Transform Inside the Feedback Loop of a Mobile Robot", SSAB, pp 137-140, Symposium on Image Analysis, Uppsala, March 1992.