7th Int. Workshop on Statistics and AI, January 1999, Fort Lauderdale, FL

Geometric Modeling of a Nuclear Environment

J. De Geeter[1], **J. De Schutter**[2], **H. Bruyninckx**[2], **H. Van Brussel**[2], **M. Decréton**[1]

[1]SCK•CEN Research Centre for Nuclear Energy
Teleoperation Project
Boeretang 200, B-2400 Mol, Belgium

[2]K.U.Leuven,
Dept. of Mechanical Engineering, Div. PMA
Celestijnenlaan 300B, 3001 Heverlee, Belgium

email : jan.degeeter@sckcen.be

Abstract

This paper is about the task-directed updating of an incomplete and inaccurate geometric model of a nuclear environment, using only robust radiation-resistant sensors installed on a robot that is remotely controlled by a human operator. In this problem, there are many sources of uncertainty and ambiguity. This paper proposes a probabilistic solution under Gaussian assumptions. Uncertainty is reduced with an estimator based on a Kalman filter. Ambiguity on the measurement-feature association is resolved by running a bank of those estimators in parallel, one for each plausible association. The residual errors of these estimators are used for hypothesis testing and for the calculation of a probability distribution over the remaining hypotheses. The best next sensing action is calculated as a Bayes decision with respect to a loss function that takes into account both the uncertainty on the current estimate, and the variance/precision required by the task.

1 Introduction

Remote handling equipment in nuclear environments is often exposed to intense gamma radiation fields. This radiation influences the properties of many materials, and limits the lifetime of equipment. Cameras in e.g. the future fusion reactor will "live" for about one minute to a few hours. Hence, they will only allow for a quick look, not for continuous live images during task execution. More robust radiation-resistant sensors do exist, but they have simple transducers and remote electronics, such as some types of ultrasonic, force and optical sensors, see Fig. 1. Unfortunately, these sensors return only a limited amount of information, such as the distance to one point on a surface. Therefore, the operator will have to rely on a synthetic image based on a rough geometric model obtained from these "quick-look" images and from CAD models. This geometric model should then be refined, verified and corrected, locally around the robot, with a few carefully chosen measurements of the radiation-resistant sensors. To do this, methods are needed (i) to update a geometric model from partial, uncertain sensor data, (ii) to determine, for each measurement, from which feature of which object it originates, and (iii), to find the best next sensing action.

This paper proposes a probabilistic solution to these three problems, under Gaussian assumptions. A Kalman filter based estimator finds the location of an object, *given* a geometric model of this object and a geometric model of the interaction between the sensor and this object. Objects are represented as collections of features, such as planes and cylinders. Constraints define the relative position of these features. This object representation allows for constructing complicated object models from a limited number of features. Basically, these constraints can be treated as just another source of information, in the same way as a measurement. However, as explained in Section 2, special precautions are needed with nonlinear constraints. Linearisation errors might prevent the Kalman filter to converge. Similarly, also incorrect hypotheses about the object model or about measurement-feature associations prevent the Kalman filter to converge properly. Therefore, Section 3 outlines a multiple-hypothesis algorithm. This algorithm allows for postponing decisions about hypotheses until sufficient evidence is available. New features of this algorithm are that it uses the residual error of the Kalman filter to eliminate hypotheses, and that it does not rely on Bayes' rule to calculate a probability distribution over the remaining hypotheses. The performance of a Kalman filter and of this multiple-hypothesis algorithm is clearly determined by the quality of the raw data, the measurements. Therefore, Section 4 outlines an action selection strategy using Bayesian decision theory. We have proposed a new loss function that takes into account both the precision required by the task and the precision of the current estimate.

The methods outlined in this paper are described more in detail in (De Geeter 1998), and in (De Geeter et al. 1996;

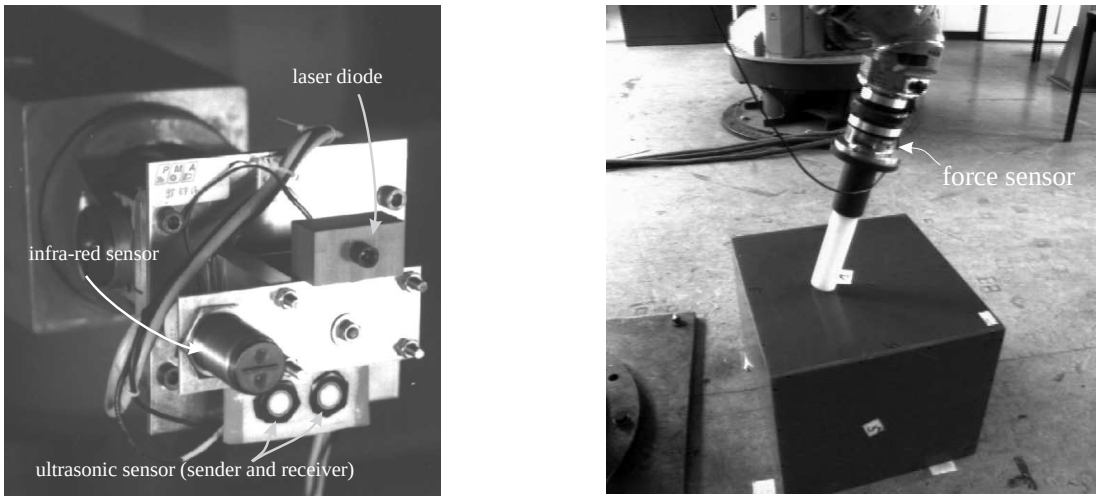

Figure 1: *Robust Radiation-Resistant Sensors.* The left figure shows a robot equipped with an ultrasonic sensor and an infra-red optical detector. The ultrasonic sensor measures the distance to a surface. The infra-red sensor returns an output voltage that is a function of the intensity of the reflected light, and is used to detect edges. The white probe (right figure) is used to explore the environment by touch, very much like the visually impaired. A force sensor measures the forces on this probe in six dimensions (three forces, three moments).

De Geeter et al. 1997; De Geeter et al. 1997; De Geeter et al. 1998).

2 The smoothly constrained Kalman filter

The Kalman filter is a recursive implementation of a weighted least squares estimator (Kalman 1960). Given a linear system model, a linear measurement model, a state estimate and a new measurement, this filter calculates a new estimate that minimizes a quadratic loss function. The system model describes the dynamic behavior of a system, while the measurement model describes the relation between the state of a system and the measurement. In general, these equations are not linear, and need to be linearised around the most recent estimate. In the applications considered in this paper, the system model is trivial since the scenes are static. The measurement model, however, is highly nonlinear. If the sensing action is, as in Fig. 1 (right), to move the force-controlled robot vertically down until the peg touches the object, then the measurement is the position of the robot along this motion path where the robot stops. Hence, this measurement can be predicted, given the state of the object, as the point along the motion path where the top plane of the object, the top plane of the peg and its cylindrical surface intersect in one point, see (De Geeter 1998) for more details.

As long as the amplitude of the measurement noise is large compared to the amplitude of the linearisation errors, the Kalman filter will converge properly. The correct convergence of the filter can easily be verified with the consistency checks, described e.g. in (Bar-Shalom and Li 1993). Usually the Kalman filter converges nicely even with nonlinear measurement equations. However, constraints often cause convergence problems. Similar to a measurement equation, constraints too describe a relation between state variables, the value of which is not determined by a real-world measurement, but by prior knowledge from e.g. CAD models. Hence, there is often no uncertainty at all on this relation. As this relation is perfectly known, the Kalman estimate satisfies this relation perfectly after application. Constraints appear as perfect correlations in the covariance matrix. If this correlation is slightly wrong due to linearisation errors, this prevents the Kalman filter from converging to the true value. To solve this problem, we have proposed the Smoothly Constrained Kalman Filter (SCKF) (De Geeter et al. 1997).

The SCKF artificially weakens constraints, i.e. increases the variance of constraint relations, and applies them several times instead of only once. In this way, the application of a constraint is smoothed over different time steps, with the integration of measurements in between, which ensures the gradual decrease of linearisation errors as the estimate converges to the true value. The SCKF decides when to start and stop applying a constraint in this way, and calculates the weakening variance as a function of the number of times the constraint has been applied. The SCKF outperforms other methods that have been proposed in the past to deal with linearisation errors, such as the iterated extended Kalman filter and the second order Kalman filter, see e.g. (Bar-Shalom and Li 1993), or the pseudo Kalman filter (Viéville and Sander 1992). Wen and Durrant-Whyte (1992) even completely abandon the constrained estimate due to convergence problems, and hence they have to ap-

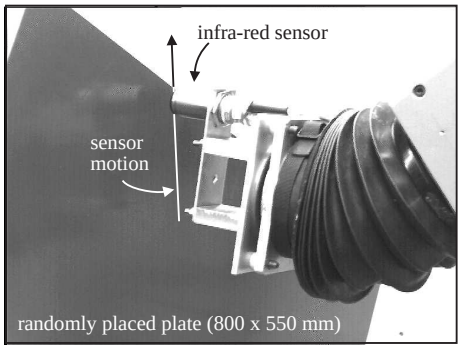

Figure 2: *Detection of the edge-crossing point with an optical sensor.* The measurement is derived from the change in output voltage of the sensor.

ply the constraints again at each time step.

2.1 Example: location of a rectangle with an infra-red sensor

Suppose the task is to locate a rectangular plate, using only an infra-red sensor installed on a six degree-of-freedom robot, Fig. 1 (left). This measurement is the point on the motion path where the sensor crosses the edge of the plate, which appears as a discontinuity in the output voltage of the sensor, Fig. 2. Figure 3 shows the result of a simulation. The rectangle is described, in two dimensions, as a collection of four lines, with five constraints. With the line representation chosen in this example, the three constraints specifying the right angles are linear, while the two constraints specifying the length and width are nonlinear. In this simulation, the measurements are generated according to a uniform distribution over the four edges. Gaussian noise with standard deviation 10mm is added in a direction perpendicular to the edges. Figures 3b-f clearly illustrate the effect of the constraints. Note that the SCKF applies the nonlinear constraint on the length only after the integration of measurement 4. This example shows that the SCKF is able to cope with an initial estimate that deviates substantially from the true value, and hence implies substantial linearisation errors.

3 Multiple-hypothesis tree

An important problem in geometric environment modeling is the *ambiguity* on the origin of measurements, besides the uncertainty (noise) on the value of the measurement. Which feature of which object was observed by sensor s at time step k? It is often hard to find a unique answer to this question. Nonetheless, finding for each measurement the corresponding feature is a prerequisite to the correct estimation of the parameters of the proposed geometric model.

More in general, system state estimation requires the choice of a *system model*, describing the system dynamics, and, for each measurement, a *measurement model*, describing the relation between the system and the measurement.

Instead of taking the risk of a wrong choice of system or measurement model, the multiple-hypothesis tree (MHT) allows for several hypothesised models to coexist until more evidence is available that allows to reject some of them with higher confidence.

The MHT, presented in (De Geeter et al. 1997) solves this problem under two assumptions: 1) all noise distributions are Gaussian, and 2) all measurements originate from a single (though unknown) system with time-invariant dynamics, whose model is one out of a given list of system models. These assumptions may seem very restrictive; their validity can however be verified at all times.

The main innovation of the MHT, compared to existing multiple-hypothesis approaches of e.g. (Cox and Leonard 1994), is the use of the residual error of the Kalman filter to prune the MHT, and to calculate a probability distribution over the remaining hypotheses.

3.1 Definitions

Definition 1 (Hypothesis) *A hypothesis η is the association of a sensing action with (i) a system model, (ii) a measurement model and (iii) a state estimate.*

A collection of hypotheses, one for each sensing action, is called a k-interpretation[1]:

Definition 2 (k-interpretation) *A k-interpretation θ^k consists of a sequence of hypotheses $\eta^j, j = 1 \ldots k$, one for each sensing action up to time step k.*

The number of hypotheses is limited by Assumption 2, which says that all measurements originate from the same system. This means that, if a $(k-1)$-interpretation θ^{k-1} is based on system model m, every k-interpretation θ^k, derived from θ^{k-1} by the addition of one hypothesis, should use the same system model m.

Definition 3 (Multiple-Hypothesis Tree) *The MHT is shown in Fig. 4:*
The top level *in this tree consists of a dummy root node.*
Level 0 *is the* system level, *consisting of one node η_i^0 for each possible system model i.*
Levels $l = 1 \ldots k$ *are* measurement levels. *Each measurement level corresponds to a measurement. The hypotheses on measurement level l are labeled $\eta_1^l \ldots \eta_{N_l}^l$, where N_l is the number of hypotheses on level l.*
An arc $\alpha_{l,n}^{l-1,m}$ *from η_m^{l-1} to η_n^l means that η_n^l borrows from η_m^{l-1} the system model and the state estimate.*

[1] This term is borrowed from (Grimson and Lozano-Pérez 1984)

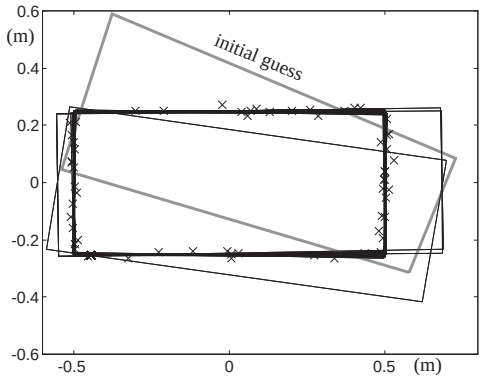

figure a

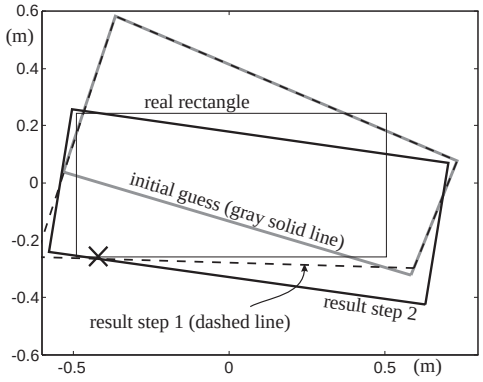

figure b: measurement 1

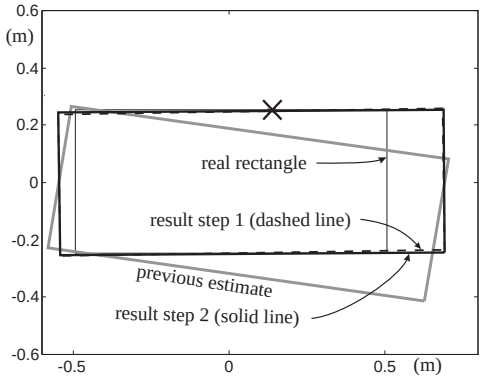

figure c: measurement 2

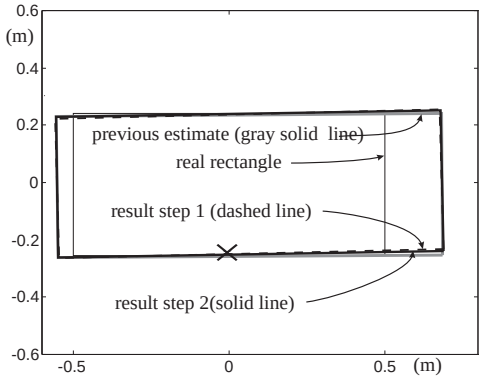

figure d: measurement 3

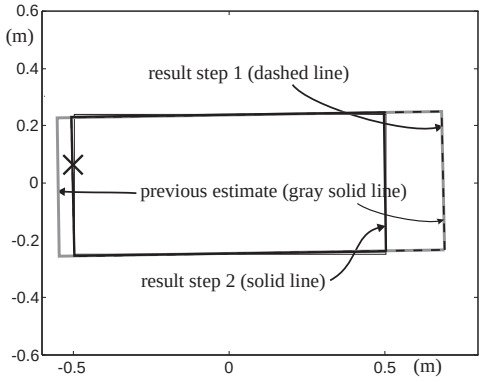

figure e: measurement 4

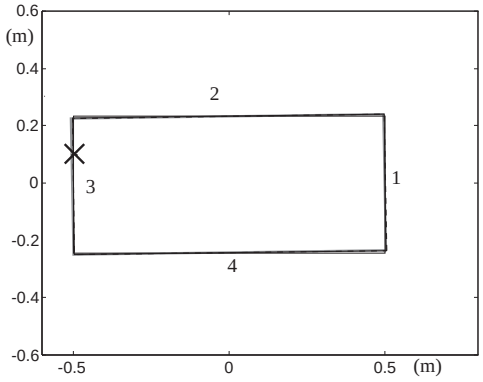

figure f: measurement 5

Figure 3: *Convergence of the SCKF.* Figure a shows the convergence of the estimate towards the true rectangle. Figures b to f show the first five time steps more in detail. The edge labels are defined in Fig. f. In step 1, a KF integrates the new measurement, shown as $\times$, with the previous estimate. Step 2 conditionally applies constraints in the same way as a measurement

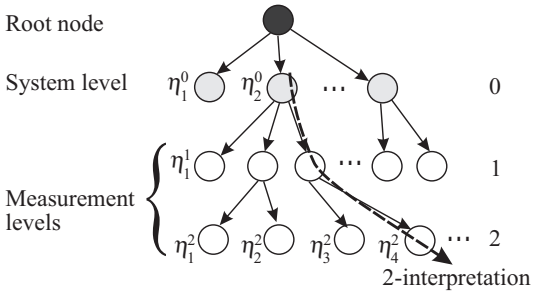

Figure 4: *Definition of the Multiple Hypothesis Tree.*

The set of leaf nodes *corresponds to the hypotheses currently under consideration.*
A path *from a particular leaf node η_j^k to the system level of the MHT corresponds to a* k-interpretation θ_j^k.

3.2 Pruning and calculating probabilities

Pruning of the MHT is done in two steps:

- *A threshold on the innovation.* The innovation is the difference between the actual measurement and the predicted measurement, calculated from the current state estimate and the measurement model. This test eliminates obviously wrong hypotheses, and only serves at avoiding breakdown of calculation accuracy due to geometrically badly conditioned measurement equations. This test also avoids the explosion of the MHT before the test on the residual error, described below, reaches a level of significance high enough to take decisions.
- "Goodness of fit" test on the residual error. The residual error ϵ_k^r of the estimate at time step k can easily be calculated from the Kalman filter equations, for proof see (De Geeter et al. 1997). ϵ_k^r is χ^2-distributed with as many degrees of freedom as statistically independent measurements.

In a Bayesian approach, the choice of the *prior* probabilities, i.e. the probability $P(\theta_j^{k|k-1})$ at time step k using measurements $\boldsymbol{z}_l, l = 1 \ldots k-1$, is a delicate task that requires a lot of domain knowledge. In our approach, these priors are of limited importance, and are chosen equal to the probability of the parent node $\eta_{p_j}^{k-1}$ of η_j^k, divided by the number of child nodes of this parent node. Hence

$$P(\theta_j^{k|k-1}) = P(\theta_{p_j}^{k-1|k-1})/c_{p_j}, \qquad (1)$$

where c_{p_j} is the number of child nodes of node $\eta_{p_j}^{k-1}$.

The posterior probability $P_S(\theta_j^{k|k})$ is chosen to be proportional to $p_{\chi^2}(\epsilon_{k,j}^r)$, the value of the χ^2 probability density function (PDF) at $\epsilon_{k,j}^r$. Normalisation is needed to make sure that the probabilities of all nodes on one level sum to one. Hence

$$P_S(\theta_j^{k|k}) = p_{\chi^2}(\epsilon_{k,j}^r) / \sum_{m=1}^{N_k} p_{\chi^2}(\epsilon_{k,m}^r), \qquad (2)$$

where N_k is the number of nodes on level k (or the number of plausible k-interpretations), and $p_{\chi^2}(.)$ is the χ^2-PDF. Hence, $P_S(\theta_j^{k|k})$ does not use the prior probabilities as calculated above. Hence, a bad choice for these priors does not permanently affect the probability distribution as in a Bayesian approach.

3.3 Example: location of a rectangle with an infra-red sensor

This is the same example as in Section 2.1. Suppose now that the operator is in doubt which of two plates is really present, Fig. 5. Note that the difference between both plate models is only five times the simulated standard deviation of the measurement. The threshold on the innovation is chosen to be 1m, which is the length of the rectangle. Figure 5 shows that during the first few steps, the probability mass spreads out over the hypotheses. The two main branches corresponding to the two system models get assigned approximately equal probability. This is due to the fact that the first three measurements do not provide evidence for one or the other system model. As more measurements are taken, the probability mass concentrates on the correct branch. After a while, the correct interpretation is clearly the most probable. It is clear that with active sensing, attention can be directed towards "interesting" features, such as opposite edges, and the MHT would never blow up as in this example.

4 Active sensing: Tolerance-weighted L-optimal experiment design

Active sensing is the task-directed choice of the controllable parameters of a sensing system (Bajcsy 1988). An active sensor system decides for itself where to look and what sensor to use.

In decision theory or game theory, this decision problem is a game with Nature. The observed system plays the role of Nature that behaves according to a known probability distribution. The optimal Bayes decision (or sensing action) minimises the expected value of a loss function. The loss function expresses the preference of the decision-maker for the possible outcomes of an action, given the task (Luce and Raiffa 1957).

This paper concentrates on "information-gathering" tasks, where no other criteria than "information gain" are relevant. Other gains or losses such as traveled distance, time ... are not considered. The quest for information is univer-

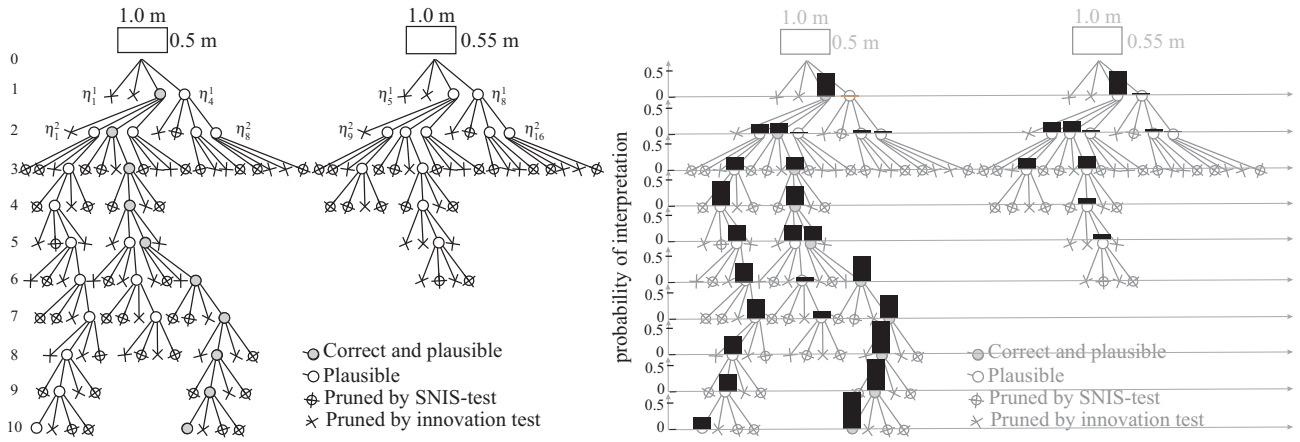

Figure 5: *The MHT, Example.* The right figure shows the same MHT as the left figure, with the probabilities as calculated by Eq. (2) superimposed. SNIS stands for: the Sum of the Normalised Innovations Squared, which is exactly equal to the residual error.

sal to most tasks; the other criteria are usually more task-dependent and can easily be accounted for in the loss function. This paper assumes that all noise distributions are uncorrelated and Gaussian.

The task-specification is similar to the quality specification in factory production: product quality is expressed in terms of *tolerances* $\mathbf{tol}_i$ on the state x_i^t of some *target property* $i,\ \ i = 1 \ldots N_t$ of a system, together with a required fraction M_i of the products that should meet this tolerance. Similarly, the information needs of a task are expressed in terms of tolerances $\mathbf{tol}_i$ on the *estimated* state $\hat{\boldsymbol{x}}_i^t$ of target property i, and a required probability M_i that its true state $\bar{\boldsymbol{x}}_i^t$ is within the tolerance interval. Suppose, for instance, the task is to insert a peg in a hole. The observable system is the object with the hole. The target property is the location of the hole, and the tolerance is the clearance between the peg and the hole. Together, $\mathbf{tol}_i$ and M_i specify a desired standard deviation σ_i on $\hat{\boldsymbol{x}}_i^t$. The *end-condition* for the information-gathering task is $\widehat{\boldsymbol{P}}_{ii,k}^t \leq \sigma_i^2, i,\ i = 1 \ldots N_t$, where $\widehat{\boldsymbol{P}}_{ii,k}^t$ is the variance of the estimated target state $\hat{\boldsymbol{x}}_{i,k}^t$ on time step k.

It is desirable that the success rate of the task, based on the estimated target states $\hat{\boldsymbol{x}}_{i,k}^t$, "smoothly" increases with k. The success rate of the task is to a large extent determined by the most badly estimated target state. Suppose that the information-gathering task is completed for $N_t - 1$ target properties, i.e. the estimated target state variance $\widehat{\boldsymbol{P}}_{ii,k}^t \leq \sigma_i^2$, but that for one target property j, $\widehat{\boldsymbol{P}}_{jj,k}^t >> \sigma_j^2$. Then, it is very likely that the insertion of the peg in the hole fails if sensing would be stopped at this time step k, despite the effort already spent. Therefore, this paper requires that $\sigma_i^2/\widehat{\boldsymbol{P}}_{ii,k}^t \approx \sigma_j^2/\widehat{\boldsymbol{P}}_{jj,k}^t$ for any two target properties i and j throughout the information gathering task. This relation is called the *task-invariant*. Unlike the end-condition, this task-invariant makes the sensing sequence task-dependent.

This decision problem is related to the problem of *optimal experiment design* (Fedorov 1972; Pukelsheim 1993). The purpose of an experiment is to optimally estimate the state of a system, where optimal is to be understood as minimum variance under Gaussian assumptions. Therefore, the optimisation criterion must be a scalar function $l(\widehat{\boldsymbol{P}})$ of the covariance matrix of the estimated state. Since no scalar function can capture all aspects of a matrix, no function suits the needs of every experiment. Pukelsheim (Pukelsheim 1993) derives so-called *information functions*: they possess the properties of nonnegativity, concavity etc. that make them suitable for ranking; cf. *loss functions* in decision theory (Luce and Raiffa 1957). The most popular information functions are the logarithm of the determinant $\log(\det(\widehat{\boldsymbol{P}}))$, the trace (or average variance) $1/n\ \ \mathrm{tr}(\widehat{\boldsymbol{P}})$, the weighted trace $\mathrm{tr}(\boldsymbol{W}\widehat{\boldsymbol{P}})$ and the maximum eigenvalue $\lambda_{\max}(\widehat{\boldsymbol{P}})$, leading to D, A, L and E-optimality respectively (Fedorov 1972; Pukelsheim 1993).

Among these existing optimal experiment designs, the D-optimal design is most popular, see (Caglioti 1994; Swevers et al. 1997; Whaite and Ferrie 1997) for some recent applications. One important reason is the invariance of this experiment design to any transformation with nonsingular Jacobian (Pukelsheim 1993). Unfortunately, a standard D-optimal design does not satisfy the task-invariant: the loss function is not task-dependent. The same remark holds for the A and E-optimal designs. In addition, the latter are not invariant to most transformations such as scaling. Therefore, we proposed a new design that has the property of invariance to transformations with nonsingular Jacobians, and that produces a task-dependent sensing sequence: the tolerance-weighted L-optimal experiment design (De Geeter et al. 1998).

This tolerance-weighted L-optimal design is a special instance of an L-optimal design that minimises $\mathrm{tr}(\boldsymbol{W}\widehat{\boldsymbol{P}})$. The

weighting matrix follows naturally from two transformations: a normalisation transformation that normalises every state variable with the standard deviation required by the task, i.e. the required precision of the estimate of this state variable. The second transformation weighs each state variable with the distance to the goal, which is the difference between the desired state variance in normalised space (equal to 1!) and the actual variance. After some algebra, the transformation matrices may be grouped into one weighting matrix, and the final loss function is

$$l_{\mathrm{tol}}(\widehat{\boldsymbol{P}}_{k+1,a}) = \sum_{i:\, \widehat{\boldsymbol{P}}_{k,ii} > \sigma_i^2} \frac{1}{\sigma_i^4} \left(\widehat{\boldsymbol{P}}_{k,ii} - \sigma_i^2 \right) \widehat{\boldsymbol{P}}_{k+1,ii,a}, \tag{3}$$

where the index a denotes the dependence of the predicted variance $\widehat{\boldsymbol{P}}_{k+1,ii,a}$ on the selected action a.

4.1 Example: location of a rectangle with an infra-red sensor

Suppose now the task is to drill a hole at the location on the plate $\boldsymbol{x}^{\mathrm{hole}} = [0.25m \;\; 0.125m]^T$ (shown as "+" in Fig. 6). The plate location is to be estimated such that there is 95% chance of finding the desired hole location in a rectangle around the estimated hole location, with dimensions $2\mathbf{tol}_1 \times 2\mathbf{tol}_2$, with $\mathbf{tol}_1 = 0.02m$ and $\mathbf{tol}_2 = 0.006m$. Since 95% corresponds to the 2σ limits of a Gaussian distribution, $\sigma_1 = \mathbf{tol}_1/2 = 0.01m$ and $\sigma_2 = \mathbf{tol}_2/2 = 0.003m$.

Fig. 6 shows the resulting measurements. These measurements are taken as close as possible to the desired hole location. Note that if the task were to insert a square peg in a slot, orientation of the plate would be important too, and some measurements would move to the vertices of the rectangle.

The sensing sequence obtained with this tolerance-weighted L-optimal design is shorter than the sequence resulting from a D-optimal design, Fig. 7: the former needs only 14 measurements, while the latter needs 18 measurements.

4.2 Location of a cylinder with an ultrasonic sensor

The task is to drill a hole in the centre point of the top plane of the cylinder. The axis of the hole should coincide with the cylinder axis, Fig. 8. The target state $\boldsymbol{x}^t = [x \;\; y \;\; z \;\; \alpha_c \;\; \beta_c]^T$, where x, y and z are the coordinates of the intersection point and α_c, β_c describe the orientation of the cylinder. The desired standard deviation of each of these target properties is $\sigma_1 = 0.001$m, $\sigma_2 = 0.001$m, $\sigma_3 = 0.01$m, $\sigma_4 = 0.01$rad and $\sigma_5 = 0.01$rad. σ_3 is larger than σ_1 and σ_2, since a force-controlled drilling robot can accommodate for errors on z. The robot uses only the ultrasonic sensor, see Fig. 1. Fig. 9 shows the first 10 sensing actions generated by a tolerance-optimal design. Fig. 10 shows as an illustration the value of the loss function on the top plane of the cylinder, used to plan respectively sensing action 4 and 5. This figure clearly shows the effect of sensing action 4 on the loss function used to plan sensing action 5.

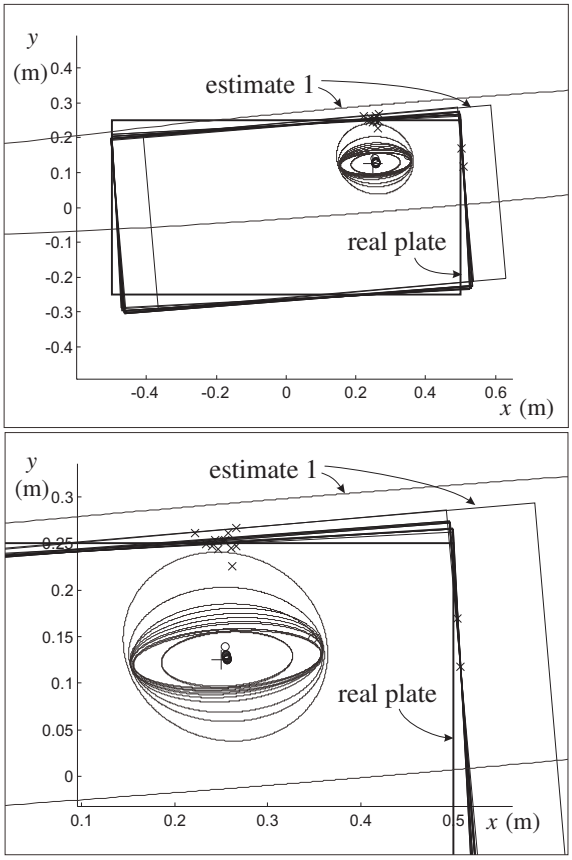

Figure 6: *Tolerance-optimal design.* This figure shows all 14 measurements, the estimated plate location, the estimated hole position, shown as $\circ$, and the 10σ uncertainty ellipse on the hole position. The bottom figure zooms in on part of the top figure.

5 Discussion

3D geometric modeling. This paper only gives examples in 2D for reasons of clarity and conciseness of the paper. For more simulation and experimental results in 3D, the reader is referred to (De Geeter 1998).

Bottom-up vs. top down approach. This work emerged, not only from the obvious need in nuclear applications, but also from the long tradition of research into sensor-based robot control and into the motion of objects in contact, by Bruyninckx, De Schutter and Van Brussel, see e.g. (De Schutter and Van Brussel 1988a; De Schutter and Van Brussel 1988b; Bruyninckx and De Schutter 1996;

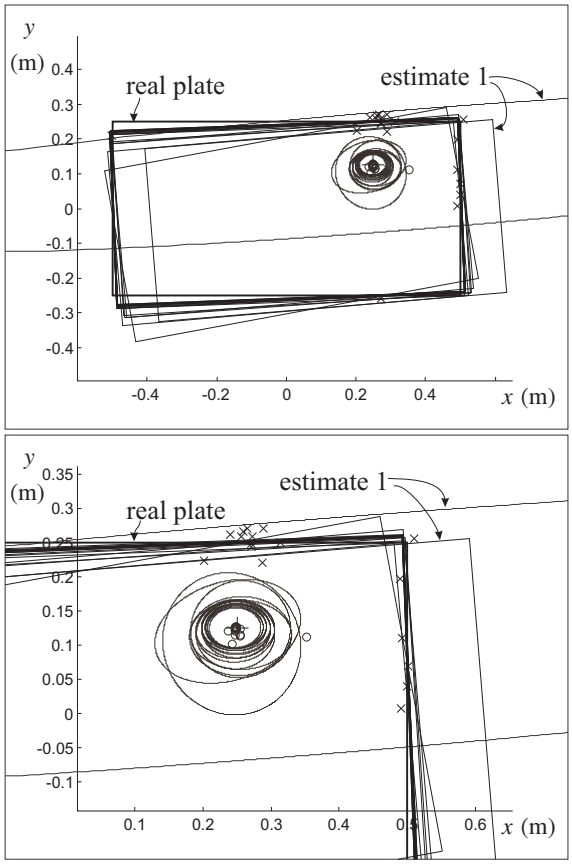

Figure 7: *D-optimal design.* This figure shows all 18 measurements, the estimated plate location, the estimated hole position, shown as $\circ$, and the 10σ uncertainty ellipse on the hole position. The bottom figure zooms in on part of the top figure.

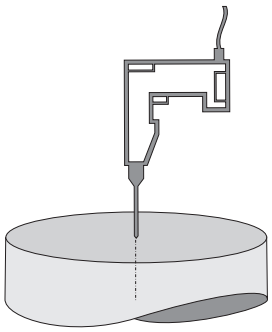

Figure 8: *Locating a cylinder.* The task is to drill a hole in the centre of the top plane of the cylinder, with hole axis aligned with the cylinder axis.

Dutré et al. 1997). The whole purpose of robotics is to modify the world, which requires controlled contact with the world. This is in contrast to the top-down approach often taken by researchers from "AI", mainly in mobile robotics, whose main goal is to avoid contact between the (mobile) robot and the world.

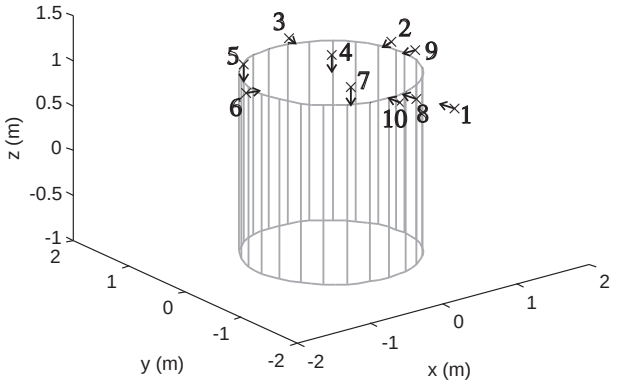

Figure 9: *Locating a cylinder with an ultrasonic sensor: tolerance-optimal design.* This figure shows the first 10 sensing actions. The position and orientation of the US sensor are shown as $\times$ and $\rightarrow$.

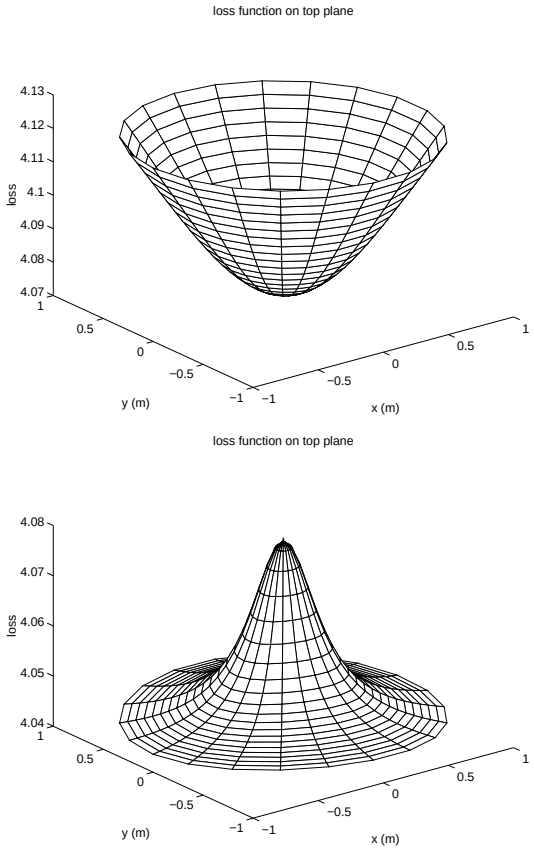

Figure 10: *Locating a cylinder with a US sensor: tolerance-optimal design.* Loss functions $\log_{10}(l_{\mathrm{tol}}(a))$ used to plan sensing action 4 (top) and 5 (bottom), evaluated on the top plane of the cylinder.

Implicit vs. explicit reasoning with uncertainty. Geometric modeling requires one to deal with uncertainty *explicitly*. This is complementary to the implicit way of accommodating for uncertainty in control. Force-controlled compliant robot motion allows to successfully assemble

pieces despite considerable uncertainty on their relative location. However, there are limits to the amount of uncertainty control loops can accommodate for. Recently, there too, the need arose to explicitly model the geometric relation between manipulated objects and the environment, see for instance (Demey 1996; Dutré 1998).

Priors and Bayes' rule. One of the common criticisms on a Bayesian approach is that specification of prior probability distributions is difficult, and requires a lot of domain knowledge. Different "experts" will come up with different probabilities. In this paper, prior probabilities are needed, firstly, to initialise the Kalman filter for each system model. In our experiments, the human operator can see the object to be located, and uses a simple graphical user interface to estimate the position of the object. The performance of the human operator did not differ too much over time and over the position of the object. Hence the variance could be estimated once and for all, and never caused convergence problems for the Kalman filter. Secondly, prior probabilities are needed to predict the probability of an interpretation being correct, see Eq. (1). These are particularly difficult to find, and therefore, their importance is limited in this paper: What is the probability that the next measurement is on edge 4 of the plate, given the current location of the robot, the current estimate of the plate location, the motion command given to the robot, ... ?. This prior probability distribution over the predicted plausible interpretations is only used to select the best one, which is then used to plan the best next sensing action as in Section 4. These priors are not used to calculate the posterior probabilities over the plausible interpretations, since these are calculated at each time step from the residual error of the KF, see Eq. (2).

This paper stops where AI people start getting interested... Many interesting problems remain. For instance, when the MHT rejects an interpretation as being implausible, it is assumed that the large residual error of the KF is due to an incorrect choice of the geometric object or measurement model. However, there might be other reasons why the KF does not converge properly: the sensor might not be working properly due to radiation, the positioning accuracy of the robot is deteriorating, the support of the infrared sensor on the robot might be bent, etc. More elaborate diagnosis would be interesting here, see e.g. (Roumeliotis, Sukhatme, and Bekey 1998). We are currently studying Bayesian networks for this purpose. Also the action planning could be sophisticated further. Its only aim is now to reduce uncertainty. It would also be interesting to take into account the ambiguity (the amount of plausible interpretations), to keep the size of the MHT down, see e.g. (Weinshall and Werman 1997).

χ^2-tests and errors in judgement. A χ^2-variable is used to test whether a value is acceptable as a sample from an n-dimensional Gaussian distribution. χ^2-tests are convenient, as it is lot easier to test on a scalar variable than on an n-dimensional variable. However, as argued already in Section 4 with the design of a loss function, no single scalar variable can capture all the information of a variable in a higher-dimensional space. With the projection of an n-dimensional space onto a one-dimensional space, a lot of information on the original distribution is lost. The question then is whether a χ^2-variable contains the right information for our purpose. Will this not lead to the rejection of correct interpretations? There are probably other scalar variables that allow for easy hypothesis testing, and that are as good as χ^2-variables.

6 Conclusion

This paper briefly reviews a probabilistic solution to dealing with uncertainty when constructing a geometric model of a nuclear environment with only simple but robust sensors. Since sensor information is sparse and expensive, our approach is very different from the approach in computer vision. There, data processing amounts to throwing away most of the data, so as to find the relevant information buried in it. In contrast to this, here our purpose is to actively go and look, so as to find exactly the information that is needed for the task.

The proposed solution consists of three parts. First, the Smoothly Constrained Kalman filter estimates the location of an object, given a geometric object model, given a new measurement and a geometric measurement model for the sensing action. The SCKF is able to integrate nonlinear constraints without convergence problems, better than existing methods. Second, the Multiple Hypothesis Tree manages the ambiguity on the measurement-feature association. The MHT runs a bank of SCKFs in parallel, one for each plausible hypothesis. New features are that the MHT uses the residual error of the SCKF to eliminate unlikely interpretations, and to calculate a probability distribution over the remaining interpretations. This makes the MHT less dependent on "experts", compared to the more classic Bayesian approach. Third, for the best interpretation, the best next sensing action is calculated as a Bayes decision that minimises a loss function. The new loss function that leads to the tolerance-weighted L-optimal experiment design, has the interesting invariance properties of a D-optimal design, and in addition, leads to a shorter task-dependent sensing sequence.

References

Bajcsy, R. (1988, August). Active perception. *Proceedings of the IEEE 76*(8), 996–1005.

Bar-Shalom, Y. and X.-R. Li (1993). *Estimation and Tracking: Principles, Techniques and Software*. Artech House, Norwood.

Bruyninckx, H. and J. De Schutter (1996). Specification

of force-controlled actions in the "Task Frame Formalism": A survey. *IEEE Transactions on Robotics and Automation 12*(5), 581–589.

Caglioti, V. (1994, May). Uncertainty minimisation in the localisation of polyhedral objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence 16*(5), 524–530.

Cox, I. J. and J. J. Leonard (1994). Modeling a dynamic environment using a Bayesian multiple hypothesis approach. *Artificial Intelligence 66*, 311–344.

De Geeter, J. (1998, May). *Constrained system state estimation and task-directed sensing. Application to the local modelling of a structured nuclear environment*. Ph. D. thesis, Dept. of Mechanical Engineering, Katholieke Universiteit Leuven, Heverlee, Belgium.

De Geeter, J., J. De Schutter, H. Bruyninckx, H. Van Brussel, and M. Decréton (1998). Tolerance-weighted L-optimal experiment design for active sensing. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Victoria B.C., Canada, pp. 1670–1675.

De Geeter, J., J. De Schutter, H. Van Brussel, and M. Decréton (1997, October). A smoothly constrained Kalman filter. *IEEE Transactions on Pattern Analysis and Machine Intelligence 19*(10), 1171–1177.

De Geeter, J., H. Van Brussel, J. De Schutter, and M. Decréton (1996). Recognising and locating objects with local sensors. In *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*, Minneapolis, MN, pp. 3478–3483.

De Geeter, J., H. Van Brussel, J. De Schutter, and M. Decréton (1997). Local world modelling for teleoperation in a nuclear environment using a Bayesian multiple hypothesis tree. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Grenoble, France, pp. 1658–1663.

De Schutter, J. and H. Van Brussel (1988a). Compliant robot motion. I. A formalism for specifying compliant motion tasks. *The International Journal of Robotics Research 7*(4), 3–17.

De Schutter, J. and H. Van Brussel (1988b). Compliant robot motion. II. A control approach based on external control loops. *The International Journal of Robotics Research 7*(4), 18–33.

Demey, S. (1996, December). *Shape identification and shape matching for compliant motion based on invariant differential shape descriptions*. Ph. D. thesis, Dept. of Mechanical Engineering, Katholieke Universiteit Leuven, Heverlee, Belgium.

Dutré, S. (1998). *Identification and monitoring of contacts in force controlled robotic manipulation*. Ph. D. thesis, Dept. of Mechanical Engineering, Katholieke Universiteit Leuven, Heverlee, Belgium.

Dutré, S., H. Bruyninckx, S. Demey, and J. De Schutter (1997). Solving contact and grasp uncertainties. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Grenoble, France, pp. 114–119.

Fedorov, V. V. (1972). *Theory of optimal experiments*. New York and London: Academic Press.

Grimson, W. E. L. and T. Lozano-Pérez (1984). Model-based recognition and localisation from sparse range or tactile data. *The International Journal of Robotics Research 3*(3), 3–35.

Kalman, R. E. (1960, March). A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 35–45.

Luce, R. D. and H. Raiffa (1957). *Games and decisions: introduction and critical survey*. New York: John Wiley & Sons.

Pukelsheim, F. (1993). *Optimal design of experiments*. John Wiley & Sons.

Roumeliotis, S. I., G. S. Sukhatme, and G. A. Bekey (1998). Sensor fault detection and identification in a mobile robot. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Victoria B.C., Canada, pp. 1383–1388.

Swevers, J., C. Ganseman, D. Bilgin Tükel, J. De Schutter, and H. Van Brussel (1997, October). Optimal robot excitation and identification. *IEEE Transactions on Robotics and Automation 13*(5), 730–740.

Viéville, T. and P. Sander (1992). Using pseudo Kalman filters in the presence of constraints, application to sensing behaviours. Technical Report 1669, INRIA-Sophia, Valbonne, France.

Weinshall, D. and M. Werman (1997, February). On view likelihood and stability. *IEEE Transactions on Pattern Analysis and Machine Intelligence 19*(2), 97–108.

Wen, W. and H. F. Durrant-Whyte (1992). Model-based multi-sensor data fusion. In *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*, Nice, France, pp. 1720–1726.

Whaite, P. and F. P. Ferrie (1997, March). Autonomous exploration: driven by uncertainty. *IEEE Transactions on Pattern Analysis and Machine Intelligence 19*(3), 193–205.