Using Tactile-Exploration with the Unscented Kalman Filter for High Precision On-line Shape and Pose Estimation of a 3D Workpiece

Isaac Monteath and Raymond Sheh
Intelligent Robotics Group, Department of Computing,
Curtin University, Western Australia
isaac.monteath@graduate.curtin.edu.au, raymond.sheh@curtin.edu.au

Daniel Fick
Australian Institute of Robotic Orthopaedics
and The Joint Studio, Western Australia

Riaz JK Khan
Australian Institute of Robotic Orthopaedics
The Joint Studio
University of Western Australia
Notre Dame University, Western Australia

Brett W Robertson
Australian Institute of Robotic Orthopaedics, Western Australia

Abstract

A common problem faced in robotic manipulation is developing techniques for accurate localisation and mapping of 3D objects. Many techniques already exist to aid in estimating the structure of the world using information from a robot's sensors such as stereo cameras, time-of-flight or structured light. These sensors and techniques used for modelling can often be made accurate enough for most practical applications (such as picking-up an object). However, some applications require a higher degree of accuracy (sub-millimeter) that is difficult to achieve with the information available from these sensors. This paper proposes the use of tactile exploration to incrementally improve the accuracy of a prior 3D object model as the robot touches different parts of a workpiece. A modified Unscented Kalman Filter (UKF) has been developed to fuse the touch probe data with the existing model and refine it over time. The approach presented in this paper is intended for applications that require a high degree of accuracy and reliability (such as medical procedures) and as such, focuses on three primary requirements—accuracy, robustness and practicality.

1 Introduction

The task is to develop a technique for modelling a 3D workpiece, using tactile exploration to incrementally refine a prior model until particular areas of interest reach a specified level of accuracy. This technique is intended to be used in applications that require a high degree of accuracy (~0.4mm). It is applicable to many fields, including medical robotics, quality control and high accuracy assembly. The techniques presented will be used to refine an approximate model of the workpiece, as well as estimate its pose relative to the tool being held by the robot. In addition, this technique will allow the robot to adapt its model to accurately reflect any distortions of the workpiece. The ability of the robot to generate a highly accurate model, and to adapt when the environment changes, is an important step towards achieving the broader goal of autonomously performing high precision workpiece operations.

The novelty presented in this paper is the application of a UKF to the problem of localising a workpiece from sparse data points. The Kalman Filter family of algorithms is most commonly used for tracking applications,

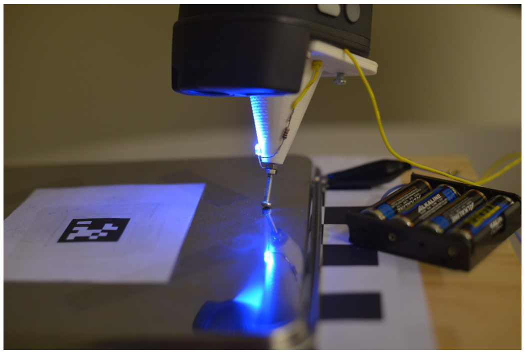

Figure 1: Touch probe equipped to the end-effector of a Baxter Research Robot

such as missile, submarine or spacecraft tracking. [Ristic *et al.*, 2004] The key difference between these applications and the presented problem, is that the position of the tracked target (the workpiece) is static. The pose and shape of the workpiece are the key estimates being generated. In applications such as spacecraft tracking, pose may be estimated using additional on-board sensors such as gyroscopes. By contrast, this application only uses the touch probe 3D coordinates as observations. In addition, this paper proposes novel parameters and optimisations to the standard Unscented Kalman Filter that improve performance for this specific application.

1.1 Performance Requirements

Three primary requirements were considered when undertaking this research. The first of these is that the system must be highly accurate. The system should be capable of estimating the shape and pose of the model to within approximately 0.4mm. This figure is based on the accuracy found in sensitive medical applications, which demand significant accuracy and repeatability. A tactile sensing probe, mounted to a fixed point on the tool, was chosen for this task due to its simplicity and close proximity to the workpiece the system is modelling. This reduces the possibility of error as the measurement relies only on the position of the end-effector, and the calibrated position of the touch-probe relative to it. By contrast, time-of-flight laser or camera-based systems are more complex and rely on many more factors such as the round-trip error between the workpiece, sensor and the tool. This results in increased error. In addition, tactile sensing is often relied upon heavily by humans in high precision applications, such as in surgery, jewellery-making and laboratory work. In these cases, vision is often used only for approximating pose, while touch is used to refine and improve the estimate. For this reason, we believe that the use of a tactile sensor is suitable for this type of problem.

The second requirement is that the system must be robust and tolerant to spurious measurements. To address this issue, the UKF state-estimation technique is used to combine multiple (potentially noisy) measurements from the touch probe into a single estimate of the workpiece shape and pose. Provided that the measurement error distribution in the touch probe is modelled correctly, the UKF will factor it into the estimate. The final requirement is practicality. The disadvantage of tactile exploration compared to other techniques is that the data points are relatively sparse compared to other 3D sensors. It can also be impractical to touch certain parts of the workpiece, leading to an uneven distribution of data points. By leveraging a-priori information, the algorithm should be capable of providing an accurate estimate, even if there is a lack of data points in some areas. It should also be able to provide an accurate estimate within a minimal number of observations in the area of interest.

1.2 Discussion Topics

This paper will cover the research process taken in the development of this system so far (both conceptually and practically). This includes general background information into where similar work has been done, as well as the fields where it could be deployed. This paper will then compare and discuss other, similar state-estimation techniques, and justify why this particular UKF configuration was chosen as the most appropriate solution for the problem. Section 5 will provide details on the current state of the system and how it was implemented. Some preliminary results will then be presented and the performance of the current implementation will be evaluated. At the end of the paper, there will be a brief discussion on how the system could be improved, and planned future work in the area.

2 Background

Research into the effective use of tactile sensors in robotics is far less common than that of other 3D sensors (such as cameras and lasers), particularly in the context of localisation and mapping. This is often due to the density of the data generated by tactile sensors not being high enough to perform effective Simultaneous Localisation and Mapping (SLAM). More commonly, tactile sensing is applied to grippers for sensitively interacting with objects. However, the use of tactile sensing is becoming increasingly popular in other areas, with new techniques such as Tactile SLAM becoming viable due to faster data refresh rates from tactile sensors. A common method of quickly measuring geometrical characteristics of physical objects is through the use of a Coordinate Measuring Machine (CMM). These dedicated machines are useful in performing this particular task, but lack the flexibility of a general-purpose robotic arm. To a certain extent, the aim of this research is to develop a general-purpose robot that can mimic the functionality of a CMM.

An interesting application of Tactile SLAM is the Whiskered Robot. [Fox *et al.*, 2012] This is a small robot on wheels, equipped with whiskers capable of sensing touch. The robot attempts to map the environment around it in a grid using only tactile sensing. The mapping part of the problem addressed by Fox et al. [2012] has similarities to the problem presented in this paper. Both problems require the mapping of objects around the robot using tactile sensing, as well as require some prior knowledge about the state of the objects. However, the techniques presented by Fox et al. [2012] do not have a focus on mapping individual objects in the environment with a high degree of accuracy. Instead, they focus on building an approximate map of the environment for simple collision-avoidance and navigation.

Another study, undertaken by Strub et al. [2014] uses tactile exploration to estimate the shape of a physical object. The paper describes their use of a neuro-dynamical model to represent a discovered shape. It also shows experimental results of a robotic hand manipulating a polygon-shaped object and using the tactile data to generate a 2D representation of that object. A key difference between our research and the research by Strub et al. [2014] is that our research makes the assumption that the general shape of the object is known as a prior. As a result of this, their research focuses on the problem of learning a shape's representation rather than generating complex, accurate models. Their experimental results show a standard deviation of approximately 1mm for a 2D 6-sided polygon, and higher for more sides. This is already larger than the 0.4mm accuracy requirement of this project, and is likely to significantly increase again with the addition of another dimension.

In a paper by Meir et al. [2011], a similar problem is addressed by using multiple Kalman Filters and a k-d tree to reconstruct the surfaces of a shape from a point cloud of touch point data. Experiments from this paper show an error between 0mm and 1mm from the ground truth depending on the object. However, this approach only generates small patches on the surface and focuses on using the estimates to classify objects, rather than model them.

Part of the problem involves finding a generic solution for representing the prior 3D model. This representation must capture the general shape of the object while exposing key parameters that can be manipulated (by the UKF in this case) to deform the shape. This is a common requirement in many applications such as computer graphics and facial recognition [Ansari *et al.*, 2006]. One method of representation suggests the use of multiple parametric superellipses to model a 2D shape. [Gong *et al.*, 2004] The parameters of the superellipses can then be modified to deform the overall shape. The paper states that the benefit of using a superellipse is that it can be used to represent many different objects such as ellipses, rectangles, parallelograms and pinched diamonds. The 3D generalisation of the superellipse, the superellipsoid, also has these capabilities and will be a consideration for representing complex objects in future work. Another promising technique suggests the use of an implicit shape potential function (ISP) to implicitly define the 3D shape. [Dragiev *et al.*, 2011] According to the paper, the representation was developed specifically to be used with probabilistic sensor fusion techniques. At the time of writing this paper, the implementation has not yet been tested with complex objects requiring the use of these representations. However, in future they will be considered as candidates for representing much more complex objects.

The type of touch sensor is another topic of interest in this research. In the context of this problem, a touch input is just an observation of a single point on the surface of the workpiece. This theoretically allows any sensor capable of observing points on a surface to simulate a touch probe. One such alternative that has been used in many applications is a laser. In a study, Liu [1997] compared the use of tactile sensors against laser scanning systems for reverse-engineering parts. The study notes that laser systems can achieve much higher data collection rates than tactile sensors, but generally at the cost of accuracy. Factors that affect accuracy in laser scanners include the properties of the surface being scanned, limited field-of-view, background lighting and the probing angle. It was also found that while tactile sensors are generally more accurate, they suffer from problems such as tip calibration, curve measurement and sparseness of data points. [Liu, 1997] It is clear that both types of sensor could be applied to different variations of the problem in this paper. The techniques presented in this paper will be capable of handling any type of sensor capable of providing "touch" observations. However, due to the focus on accuracy, a tactile sensor is proposed as the preferred data source.

3 System Overview

The system requires several inputs in order to function correctly. It is tolerant to some measurement error in these inputs. However, the accuracy of the final estimate will partially depend on the degree of this error, as well how accurately the error is modelled. The system will take a prior, parametric model of the workpiece as an input. The parameters of this model will be estimated by the UKF. Currently, this model is defined manually for testing purposes. However, in the future it could come from other sources such as 3D modelling

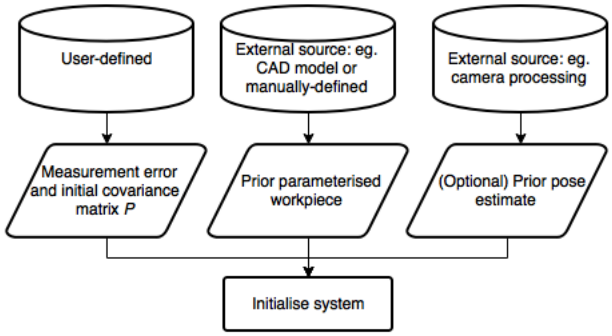

Figure 2: High-level flow-diagram showing initial inputs to the system (time $k = 0$)

programs, MRI scans or 3D scanning tools (as shown in Figure 2). A prior approximation of the workpiece position can also be used to bootstrap the system to further improve performance (this could be generated by other sensors, such as a camera). The next input to the system is the chosen coordinate-frame to operate in. For workpieces with a fixed-position, this can simply be any fixed coordinate-frame (such as the base of the robot or the world). However, for moving workpieces, the coordinate-frame can be tracked using fiducial markers or constellations that are fixed to the workpiece. The pose of the workpiece will then be estimated relative to the chosen coordinate-frame.

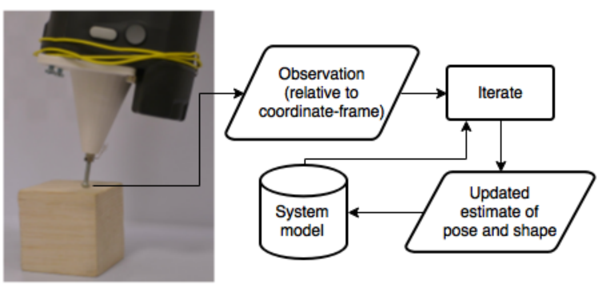

Figure 3: High-level flow-diagram showing a single observation and iteration of the filter

As shown in Figure 3, the system must be able to detect when the probe has touched the workpiece, as well as the probe's position relative to the selected coordinate-frame when it does. The touch detection is relatively simple. For testing purposes, a simple circuit was used with an LED light that activates when the probe touches a metal object (see Figure 1). The more complicated input is the position of the probe (relative to the end-effector). This can be obtained by manually measuring the position of the probe relative to the camera, or (in the case of this project) by using a Kalman Filter to pre-calibrate the position of the probe. Given these inputs, over time the system will estimate the parameters of the prior model, as well as its pose relative to the chosen coordinate-frame (position and rotation).

3.1 Problem Formulation

The underlying system that will be used to solve the problem must track a workpiece of approximately-known pose and shape. The problem will be defined using the state-space convention, since that is the most commonly-used when discussing recursive filtering algorithms. The filtering algorithm will take the following input as measurement:

$$z = \begin{bmatrix} x & y & z \end{bmatrix}^T \tag{1}$$

which represents the position of the touch probe relative to the chosen coordinate-frame. The state-vector is defined as follows:

$$x = \begin{bmatrix} p & r & c_0 & \dots & c_n \end{bmatrix}^T \tag{2}$$

where p is the center-point of the workpiece relative to the chosen coordinate-frame, r is the rotation of the workpiece about p and $c_0 \dots c_n$ is the set of important parameters that determine the shape of the workpiece (for instance: a cuboid is $n = 3$ for x, y and z-side lengths)

These parameters should be enough to fully represent the state of the system, no matter how complex the workpiece is. After each observation, the chosen state-estimation technique should use the observation to generate a new estimate of these parameters.

4 Comparison of State-Estimation Techniques

This section will compare different recursive Bayesian estimation techniques from a theoretical perspective and evaluate their suitability for performing sensor fusion in this problem. Each technique was compared by following a specific set of criteria. This involved the performance criteria detailed in Section 1.1, such as the accuracy, error-tolerance and practicality (on-line and convergence speed) of the filter. It also involved considering the nature of the problem, and researching the types of problem each technique is capable of solving. This particular problem is multimodal due to potential ambiguity in the tactile sensor data. In addition, measurement errors are non-Gaussian, and the measurement transformation between the prior and posterior distributions is nonlinear and potentially discontinuous. It is worth noting that for this particular set of requirements, a mathematically optimal algorithm has not yet been found. [Ristic *et al.*, 2004]

4.1 Grid-based Filters

Grid-based filters involve partitioning the state-space into a grid of N segments and performing filtering on each individual segment. Each segment is weighted and used to approximate the posterior distribution. Some variants of this technique, such as the Beneš and Daum filters, are capable of optimally estimating the posterior distribution for linear problems and a subset of nonlinear problems. [Beneš, 1981; Daum, 1986] Unfortunately, this particular problem falls outside of that subset because the transition function from state-space to observation-space is nonlinear (only the process function can be nonlinear). Alternatively, the Gaussian Sum filter is a suboptimal estimator, but can be used to estimate all nonlinear systems. It uses a Gaussian mixture-model to approximate the posterior distribution at each time step. [Sorenson and Alspach, 1971] There are two common subclasses of this filter, the Static Multiple Model Estimator [Alspach and Sorenson, 1972] and the Dynamic Multiple Model Estimator [Bar-Shalom *et al.*, 2002]. This paper will only consider the static variant, as the dynamic estimator is more suited to target tracking. The key advantage of the Gaussian Sum approach is that it is suitable for estimating multimodal posterior densities, even up to higher-order moments. This means that theoretically, the Gaussian Sum filter is capable of modelling the problem. However, the disadvantage of this approach is that computational complexity grows exponentially as the dimension of the state vector increases. This could potentially affect the viability of an on-line implementation for the problem because complex shapes may require many parameters in the state vector.

4.2 Kalman Filters

Kalman Filters are one of the original and most widely-used methods for performing state-estimation. The original Kalman Filter is an optimal estimator, but is only capable of estimating linear systems. [Kalman, 1960] For this reason, it is not appropriate for this problem. However, many suboptimal Kalman Filter variants exist that are designed to handle nonlinear problems. One of the most commonly-used variants is the Extended Kalman Filter (EKF). The primary advantage of this filter is that it can perform both linear and nonlinear estimation. [Jazwinski, 1970] Like the Kalman Filter, it is far more computationally-efficient than most Grid-Based and Monte Carlo approaches, making it suitable for online use. Unfortunately, there are also several disadvantages to the EKF. The first of these is that it requires the explicit calculation of Jacobians, making it not only difficult to implement, but also incapable of handling discontinuous measurement functions. [Ristic *et al.*, 2004] Therefore, using an EKF would require complex shapes to be represented as continuous functions, which could potentially lead to a loss of accuracy for shapes with hard edges. Additionally, estimates only model first-order moments in the standard EKF. A second-order EKF exists, but is rarely used due to its difficulty to implement. From a theoretical perspective, the EKF could be used, but it would be unlikely to reach the desired performance requirements of this problem.

A more-recent alternative to the EKF is the Unscented Kalman Filter (UKF). [Wan and Van Der Merwe, 2000] It involves propagating the mean and a small set of sigma points through each nonlinear transition function. The advantage of this is that unlike the EKF, Jacobians are calculated implicitly by the filter, thereby allowing discontinuous transition functions. The posterior estimates are captured up to at least second-order moments (potentially third-order if parameters are correctly set), which usually allows for higher accuracy than the EKF. The computational-complexity of this filter is slightly higher than the Kalman Filter and EKF due to the requirement for multiple sigma points. However, this difference is negligible for most applications except ones with large state-vectors, making the UKF still suitable for on-line estimation. The main disadvantage of the standard UKF is that it is not suitable for multimodal systems such as the problem specified in this paper. However, it is possible to use multiple UKFs simultaneously to achieve multimodal approximation. There have also been several improvements made to the UKF, such as the Square-Root UKF (SR-UKF) for better numerical stability and the Scaled Unscented Transform for more control over sigma point distribution spread. [Van Der Merwe and Wan, 2001; Julier, 2002] The UKF is a good theoretical fit for this problem, provided that it is possible to use as a multimodal estimator.

4.3 Monte Carlo Simulation

This family of filters make use of a large number of particles to approximate the posterior distribution. The most common type is the Particle Filter, which is a form of Sequential Monte Carlo (SMC) estimation. [Arulampalam *et al.*, 2002] The original Particle Filters used Sequential Importance Sampling (SIS), which have been replaced by Sequential Importance Resampling (SIR). [Doucet *et al.*, 2000; Gordon *et al.*, 1993] Hereafter, the term "Particle Filter" will be used to refer to SIR Particle Filters. The advantages of the Particle Filter are as follows:

- It does not distinguish between linear and nonlinear systems (unlike Kalman Filters)
- It is insensitive to the distribution of the state (unlike the Kalman Filter, which performs best with Gaussian distributions)
- It is capable of estimating multi-modal problems
- It can often be more accurate than a grid-based system, as the representation of the state has not been discretised

- It is possible to use negative observations when the touch probe is not touching the object to eliminate particles

However, despite its numerous advantages, the Particle Filter also has several disadvantages. One of the largest problems is the high-computational complexity required. This complexity often increases exponentially as the dimension of the state vector increases. It is common for Particle Filters to require thousands of particles for accurately estimating the posterior distribution. It can also be difficult to choose a mathematically-rigorous weighting function to determine the importance of a particular particle. If the weighting function is chosen incorrectly, the filter can become unstable and the weights of the particles may collapse. A specific issue with using the Particle Filter with this problem is that it will be used to estimate static parameters. Particle Filters were designed to track moving targets, not the parameters of a stationary object. This could potentially lead to a lack of diversity in the particles, which often results in particle impoverishment and irrecoverable collapse. However, the likelihood of this occurring, and the degree to which it could affect the estimate in this problem will need to be determined using experimentation.

Recently, modifications have been made to prevent degeneration in static parameter Particle Filters. One such modification is the Artificial Dynamics approach which involves creating an artificial motion-model to maintain diversity in the particles. [Liu and West, 2001] However, it is difficult to formulate the motion-model such that biases are not introduced into the estimation. Other online techniques are compared in a paper by Kantas et al. [2009] such as Resample-Move, Expectation Maximisation (EM) and on-line Gradient approach. The paper finds that these techniques are either computationally-expensive, do not entirely solve the degeneracy problem, or suffer from the local-maximum problem. Another modification, The Extended Parameter Filter, has been shown to reduce the degeneracy problem at a lower computational-cost. [Erol *et al.*, 2013] Despite the disadvantages of the Particle Filter for this problem, it fits the multimodal, nonlinear, non-Gaussian nature of the problem reasonably well.

5 Implementation

This section describes the details of the UKF implementation, as well as providing information on tuning the filter, initialisation, and modifications. A simple example of estimating a cuboid will be used to provide additional clarity. This paper will not discuss the details of the original UKF implementation. Instead, it will provide details on the implementation that is specific to this particular problem. For details on the original UKF, see [Wan and Van Der Merwe, 2000].

5.1 Setting the Parameters

This implementation of the UKF uses the scaled unscented transform, proposed by Julier [2002], for a higher degree of control over the distribution of the sigma points. It has three tunable parameters (α, β, κ) that can be adjusted to yield better performance for certain problems. α and κ control the spread of the sigma points while β relates to the expected distribution ($\beta = 2$ is optimal for Gaussian distributions). A common heuristic for setting these parameters is $\alpha = 1e-3$, $\beta = 2$ and $\kappa = 0$ if the state distribution is Gaussian. [Wan and Van Der Merwe, 2000] However, as discussed by Turner [2011], setting $\beta = 2$ can result in exact inference for non-Gaussian state distributions. After testing the UKF with these parameters, it was found that they indeed caused instability in the filter, often leading to divergence from the solution or undefined values in the state vector. Experiments showed that the parameters suggested by Thrun et al. [2005], $\alpha = 1$, $\beta = 0$ and $\kappa = 3$ provided the most stable performance.

5.2 Initialising the Filter

The filter has several parameters that need to be initialised when it is started, before any observations have been made. These parameters are:

L: the dimension of the state vector (x)
m: the dimension of the measurement vector (z)
x: a $L \times 1$ matrix representing the state vector
P: an $L \times L$ matrix representing the covariance between state variables
Q: an $L \times L$ matrix representing the process noise
R: an $m \times m$ matrix representing the observation noise

L should be initialised to the number of parameters required for modelling the shape of the workpiece, as well as three additional parameters for the center point and three for the rotation (or four if using a quaternion). The m variable should always be initialised to 3 as the measurement vector needs to store the Cartesian position of the touch probe. The center point position and rotation in the x matrix can be initialised to any value, but it should be within the range of expected values if possible. For this implementation, Cartesian coordinates are used for the center point and a unit quaternion is used to represent rotation. Quaternions were chosen instead of Euler angles due to their lack of singularities. The shape parameters in the x matrix should be initialised with the values given by the prior model. For instance, in the example cuboid implementation, if the cuboid has sides of length 10, 20 and 30, the initial x matrix could be:

$$x = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 10.1 & 19.8 & 30.05 \end{bmatrix}^T \quad (3)$$

where the values are center point x, y, z, quaternion x, y, z, w, and the shape side lengths respectively. The slight difference from the real side lengths is to simulate inaccuracies in the prior. The P matrix should be initialised to a diagonal matrix. The values on the diagonal corresponding to the center point and rotation should be initialised to 1, as these values are completely unknown. The values corresponding to the shape parameters should be set to a very low value to indicate that some information is already known about these values (from the prior). This value should be chosen depending on the accuracy of this prior, a value of 0.01 was used during testing. For instance, in the example, the P matrix could be set to:

$$P = \begin{bmatrix} 100 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.01 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.01 \end{bmatrix} \tag{4}$$

The high values corresponding to the variance of the center point indicate that the initial estimate of the cuboid's position is highly uncertain. The values corresponding to the rotation indicate a high level of uncertainty because, relative to the range of possible unit quaternion values, 1 is a high variance. The low value for the variance of the shape parameters indicates that the initial estimate should be relatively accurate (due to the prior). These initial values were found to significantly impact the performance of the filter.

5.3 Definition of Process and Measurement Functions

The process function f determines the expected movement of the variables in the state-vector over time. In the scope defined by this paper, this function is unnecessary as the given coordinate-system is assumed to be fixed (or tracked elsewhere). However, this function could be used if the coordinate system is moving, such as if it is defined by a fiducial marker fixed to a moving workpiece. This will be explored in future research. The measurement function (h) is where the state-vector is transformed into observation space to compare it with the current observation. Regardless of the shape being estimated, the general approach for performing this transformation is shown in Listing 1.

Note that the original UKF paper defines the h function as only taking the state vector as an input. [Wan and Van Der Merwe, 2000] This approach requires that the observation z also be passed to the h function.

```
Input: (x, z) where x = (p, r, c0 ... cn)
    and p, z are Cartesian points and r is a
     unit quaternion

for each set of parametric equations:
    construct surface at origin
    rotate surface by r
    translate surface by p
    np = nearest point on surface to z
    d = dist(np, z)
    if d < dist(closest , z)
        closest = np
end

Output: closest
```

Listing 1: Pseudocode showing proposed formulation of h function

5.4 Modifications

Two significant modifications to the standard UKF were necessary to improve the performance of the filter. The first modification was required to address issues with the unit quaternion in the state vector. The problem was that the normalisation constraint imposed on the quaternion could often be violated by the linear measurement updates of the filter. To overcome this problem, we used a technique proposed by Crassidis and Markley [2003] to generate quaternion sigma points from a Modified Rodrigues Parameters (MRP) three-component attitude-error vector. This served to maintain the normalisation constraint during both the generation of sigma-points and the measurement update. The second modification was the use of a global Gaussian Sum filter which combined the posterior estimates from a bank of UKFs into a Gaussian mixture. This technique is based on a paper by Fang and Wu [2006], and was implemented to improve performance on multimodal state spaces. At the time of writing, this modification had not been optimised, and did not appear to improve performance. However, experiments are planned in the future to determine its effect on performance.

6 Preliminary Findings

Some early experiments were performed to benchmark the performance of the current implementation. These experiments were performed on a single UKF. The global Gaussian Sum filter was not used in the testing due to it not being properly optimised. It should also be noted that these results are not conclusive, they are intended to serve only as an approximate indicator of performance. More rigorous testing is planned in the future (see Section 8).

The touch-probe and workpiece were both simulated in order to simplify the experiments and provide more control. Measurement noise was not simulated. Each

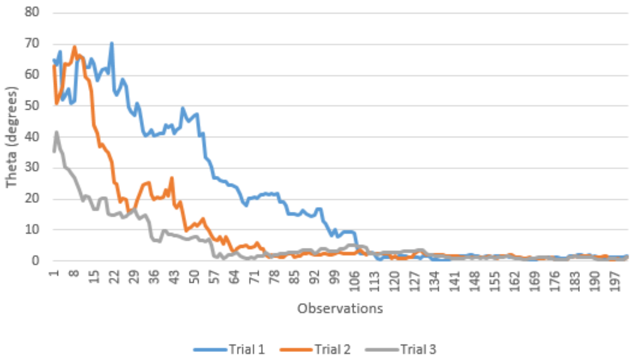

Figure 4: Graph showing the difference in rotation between the estimate and the real object in three trials.

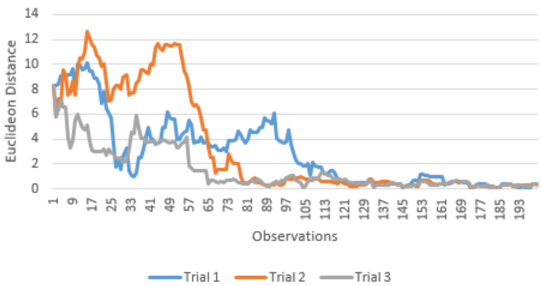

Figure 5: Graph showing the Euclidean distance between the center point of the estimate and the real object in three trials.

touch position was generated using a deterministic random number generator, and each trial used the same seed for consistency. The workpiece was defined as a cuboid with varying x, y and z-side lengths. In each of the trials, the workpiece was moved to a different position and orientation (see Figure 6). In each case, the filter was initialised to within ±10mm of the center point and ±1mm of the actual side lengths (as these would normally be given as priors). As can be seen in Figures 4 and 5, the error in the center point position and orientation of the workpiece converges toward 0 after approximately 120 observations in all trials.

Trial	Center Point Error	Orientation Error
1	0.2996	1.2019
2	0.2931	0.4243
3	0.2162	0.5780

Table 1: Center Point error and Orientation error (degrees) after 200 observations

Table 1 shows that the error in both the center point and orientation of the estimate is relatively insignificant after 200 observations. While 200 individual observations may seem impractical, it is a favourable result considering that the filter is unoptimised and the observation points are random. Identifying observation points that yield the most information about the workpiece is part of ongoing research, and is expected to further improve convergence time in the future. In addition, a sliding, rather than probing, touch sensor or laser may be considered to acquire the observations more rapidly.

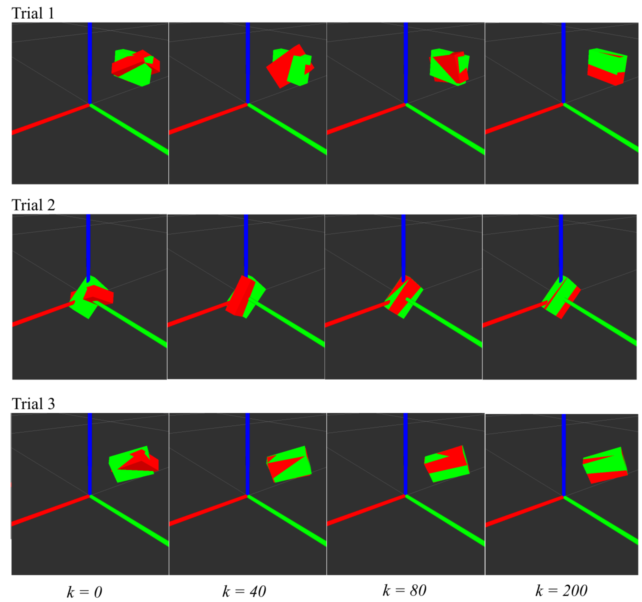

Figure 6: Visualisation of estimate (red) and simulated workpiece (green) at various times k during trials.

7 Conclusion

Humans are often required to perform object manipulation tasks that demand a high degree of accuracy, precision and control. In these scenarios, vision can provide a holistic approximation of the location and shape of an object, but is not accurate enough to be useful on its own. By contrast, touching the object can be significantly more accurate, especially if performed with the probe tool (to avoid accumulated errors). However, this requires some prior knowledge of the object's location. By combining the information from both vision and touch, humans are capable of performing these demanding tasks. The techniques presented in this paper aim to artificially mimic this behaviour, and are a step towards developing robots to perform these tasks autonomously. Through comparison of several modern state-estimation techniques, it was found that the theoretical best fit for this problem was a Gaussian Sum filter using a bank of UKFs. The implementation of this filter was successful and this paper proposed several techniques and modifications for maximising performance. Preliminary experimental findings (Section 6) were supportive of the research conducted in this paper. It showed that the techniques presented appear to be suitable for solving the problem, as each trial resulted in convergence to an acceptable degree of accuracy.

8 Future Work

As shown in the Preliminary Findings (Section 6), the filter converges to the correct solution. However, it is not optimised or fitted to a specific real-world problem. Further experiments are required to evaluate the performance in real-world situations. These scenarios would include more realistic measurements and noise, and the filter parameters would be tuned to provide the best performance for the problem. The first planned real-world problem is the estimation of the pose and shape of a real cuboid. To simplify the problem, a tactile sensor will initially be guided by a human. At a later stage, the tactile exploration will be performed autonomously based on information from other sensors. A tactile probe attached to the Baxter Research Robot would be used to provide the observations. These experiments will be repeated using a KUKA LBR iiwa series robot.

A limitation of the current implementation is that it requires manual parameterisation of the workpiece. This is acceptable for simple primitives such as planes, spheres and cubes. However, it becomes difficult or impossible with complex objects such as meshes. In future, we plan to extend the filter to automatically parameterise meshes of complex objects based on other low resolution sensors such as vision, 3D sensors and other prior knowledge. In addition, several recent modifications have been proposed for the UKF such as the SR-UKF and the Adaptive Unscented Kalman Filter (AUKF) which could potentially provide better stability, accuracy or computational complexity than the existing implementation. [Van Der Merwe and Wan, 2001; Jiang *et al.*, 2007] These modifications will be considered for future iterations of the filter.

References

[Alspach and Sorenson, 1972] D.L. Alspach and H.W. Sorenson. Nonlinear bayesian estimation using gaussian sum approximations. *Automatic Control, IEEE Transactions on*, 17(4):439–448, Aug 1972.

[Ansari *et al.*, 2006] A.-N. Ansari, M. Abdel-Mottaleb, and M.H. Mahoor. A multimodal approach for 3d face modeling and recognition using deformable mesh model. In *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the*, pages 1–6, Sept 2006.

[Arulampalam *et al.*, 2002] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2):174–188, Feb 2002.

[Bar-Shalom *et al.*, 2002] Yaakov Bar-Shalom, Thiagalingam Kirubarajan, and X.-Rong Li. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, NY, USA, 2002.

[Beneš, 1981] V. E. Beneš. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, 5(1-2):65–92, 1981.

[Crassidis and Markley, 2003] John L Crassidis and F Landis Markley. Unscented filtering for spacecraft attitude estimation. *Journal of guidance, control, and dynamics*, 26(4):536–542, 2003.

[Daum, 1986] Frederick E. Daum. Exact finite-dimensional nonlinear filters. *Automatic Control, IEEE Transactions on*, 31(7):616–622, Jul 1986.

[Doucet *et al.*, 2000] Arnaud Doucet, Simon Godsill, and Christophe Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

[Dragiev *et al.*, 2011] S. Dragiev, M. Toussaint, and M. Gienger. Gaussian process implicit surfaces for shape estimation and grasping. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 2845–2850, May 2011.

[Erol *et al.*, 2013] Yusuf Erol, Lei Li, Bharath Ramsundar, and Stuart J Russell. The extended parameter filter. *arXiv preprint arXiv:1305.1704*, 2013.

[Fang and Wu, 2006] Bingyi Fang and Siliang Wu. Angle-parameterizations range-only target tracking for scalar miss distance measurement system. In *Signal Processing, 2006 8th International Conference on*, volume 1, pages –, 2006.

[Fox *et al.*, 2012] C. Fox, M. Evans, M. Pearson, and T. Prescott. Tactile slam with a biomimetic whiskered robot. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 4925–4930, May 2012.

[Gong *et al.*, 2004] Lixin Gong, S.D. Pathak, D.R. Haynor, P.S. Cho, and Yongmin Kim. Parametric shape modeling using deformable superellipses for prostate segmentation. *Medical Imaging, IEEE Transactions on*, 23(3):340–349, March 2004.

[Gordon *et al.*, 1993] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *Radar and Signal Processing, IEE Proceedings F*, 140(2):107–113, Apr 1993.

[Jazwinski, 1970] Andrew H. Jazwinski. *Stochastic processes and filtering theory.* Number 64 in Mathematics in science and engineering. Acad. Press, New York, NY [u.a.], 1970.

[Jiang *et al.*, 2007] Zhe Jiang, Qi Song, Yuqing He, and Jianda Han. A novel adaptive unscented kalman filter for nonlinear estimation. In *Decision and Control, 2007 46th IEEE Conference on*, pages 4293–4298, Dec 2007.

[Julier, 2002] S.J. Julier. The scaled unscented transformation. In *American Control Conference, 2002. Proceedings of the 2002*, volume 6, pages 4555–4559 vol.6, 2002.

[Kalman, 1960] R. E. Kalman. A new approach to linear filtering and prediction problems. *ASME Journal of Basic Engineering*, 1960.

[Kantas *et al.*, 2009] Nicholas Kantas, Arnaud Doucet, Sumeetpal Sindhu Singh, and Jan Marian Maciejowski. An overview of sequential monte carlo methods for parameter estimation in general state-space models. In *15th IFAC Symposium on System Identification (SYSID), Saint-Malo, France.(invited paper)*, volume 102, page 117, 2009.

[Liu and West, 2001] Jane Liu and Mike West. Combined parameter and state estimation in simulation-based filtering. In De Freitas and Nj Gordon, editors, *Sequential Monte Carlo Methods in Practice.* Springer-Verlag, New York, 2001.

[Liu, 1997] Yan Liu. Comparison of tactile probe cmm and laser scanning techniques for reverse engineering. Master's thesis, University of Windsor, 1997.

[Meier *et al.*, 2011] M. Meier, M. Schopfer, R. Haschke, and H. Ritter. A probabilistic approach to tactile shape reconstruction. *Robotics, IEEE Transactions on*, 27(3):630–635, June 2011.

[Ristic *et al.*, 2004] Branko. Ristic, Sanjeev. Arulampalm, and Neil. Gordon. *Beyond the Kalman filter : particle filters for tracking applications / Branko Ristic, Sanjeev Arulampalm, Neil Gordon.* Artech House Boston, Ma. ; London, 2004.

[Sorenson and Alspach, 1971] H.W. Sorenson and D.L. Alspach. Recursive bayesian estimation using gaussian sums. *Automatica*, 7(4):465 – 479, 1971.

[Strub *et al.*, 2014] C. Strub, F. Worgotter, H. Ritter, and Y. Sandamirskaya. Correcting pose estimates during tactile exploration of object shape: a neuro-robotic study. In *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2014 Joint IEEE International Conferences on*, pages 26–33, Oct 2014.

[Thrun *et al.*, 2005] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents).* The MIT Press, 2005.

[Turner, 2011] Ryan Darby Turner. *Gaussian Processes for State Space Models and Change Point Detection.* PhD thesis, University of Cambridge, Cambridge, UK, July 2011.

[Van Der Merwe and Wan, 2001] R. Van Der Merwe and E.A. Wan. The square-root unscented kalman filter for state and parameter-estimation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, volume 6, pages 3461–3464 vol.6, 2001.

[Wan and Van Der Merwe, 2000] E.A. Wan and R. Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158, 2000.