The Spherical Simplex Unscented Transformation

Simon J. Julier
IDAK Industries
901 Missouri Blvd., #179
Jefferson City, MO 65109 USA
sjulier@idak.com

***Abstract*— This paper describes a new and better-behaved sigma point selection strategy for the Unscented transformation (UT). The UT approximates the result of applying a specified nonlinear transformation to a given mean and covariance estimate. The UT works by constructing a set of points, referred to as *sigma points*, which have the same known statistics as the given estimate. This paper describes a sigma point selection strategy that requires, for n dimensions, $n+2$ sigma points. $n+1$ of these points lie on a hypersphere whose radius is proportional to $\sqrt{n}$. The weights on each point are proportional to $1/n$. We illustrate the algorithm through an example which uses simultaneous localisation and map building.**

I. Introduction

Many types of practical systems, such as head trackers, missile tracking systems and mobile robotics systems are challenging for trackers for three main reasons. First, the underlying process and observation models are nonlinear. Second, the algorithms must be run at extremely high update rates. Third, there is only extremely limited information available about the environment. Probably the best-known and most widely implemented nonlinear estimator is the extended Kalman filter (EKF) or one of its many variants [1]. The EKF applies the Kalman filter to nonlinear systems by simply linearising all the nonlinear models so that the traditional linear Kalman filter equations can be applied. In so doing, the EKF inherits many of the advantages of the Kalman filter including its limited computational costs ($O(n^3)$ for an n-dimensional system), and the fact that only the first two moments of an estimate need to be specified[1]. Unfortunately, the EKF suffers two well-known problems: First, the required Jacobian matrices, i.e., the matrix/linear approximations of nonlinear functions, can be extremely difficult and error-prone to derive. Second, the EKF linearized approximations can be extremely inaccurate and lead to filter instability [2].

Many approaches have been developed to address these problems. These include: calculating higher order terms in the Taylor Series expansions [3] (which compounds the difficulty in calculating the matrices of partial derivatives), developing exact solutions [4] (which can only be applied to systems with very special structures), and using particle filters [5]. The latter approximate distributions using a set of points. Although the method is potentially extremely powerful, it requires a large number of sample points for reliable estimates (Thrun [6] reports several thousand for a 3 dimensional navigation system).

The Unscented Transform (UT) forms a middle ground between these two methods [7, 8]. It works by constructing a set of points, referred to as *sigma points*, that are deterministically chosen to have the same known statistics, e.g., first and second moments, as a given measurement or state estimate. A specified nonlinear transformation can be applied to each sigma point, and the unscented estimate can be obtained by computing the statistics of the transformed set. For example, the mean and covariance of the transformed set approximates the nonlinear transformation of the original mean and covariance estimate. The deterministic component of the UT avoids the random sampling errors introduced by Monte Carlo and other sampling methods and therefore dramatically reduces the number of points required to achieve the same transformation accuracy.

Because the computational costs are proportional to the number of sigma points used, there is a strong incentive to minimise the number of points required. In [9] we proved that, for an n-dimensional state, $n+1$ points are required to represent the mean and covariance fully. We derived the *minimal skew* set of simplex points that minimise the magnitude of the third order moments. However, these points have the problem that the radius which bounds the sphere of the points is $2^{n/2}$. Therefore, at even relatively low dimensions there are potential problems with numerical stability.

In this paper we develop a new sigma point selection strategy which defines a simplex of points that lie on a hypersphere. For an n-dimensional space, only $n+2$ points are required, the radius which bounds the points is proportional $\sqrt{n}$ and the weight applied to each point is proportional to $1/n$.

The structure of this paper is as follows. Section II reviews the unscented transform and describes the minimal skew point selection method. Section III derives the *spherical set*. The performance of this point selection algorithm is studied in Section IV. Conclusions are drawn in Section V.

II. Problem Statement

Let $\mathbf{x}$ be an n-dimensional random variable with (a not necessarily Gaussian) probability density function $p_x(\mathbf{x})$, mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{xx}$. A second random variable $\mathbf{y}$ is related to $\mathbf{x}$ through the nonlinear transformation

$$\mathbf{y} = \mathbf{f}[\mathbf{x}] \tag{1}$$

The objective is to calculate the mean $\bar{\mathbf{y}}$ and covariance $\mathbf{P}_{yy}$ of $\mathbf{y}$.

The Unscented Transform (UT) builds on the principle that it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function. A set of $p+1$ weighted points $\mathcal{S} = \{W_i, \boldsymbol{\mathcal{X}}_i\}$ are deterministically chosen so that they obey a condition of the form

$$\mathbf{g}[\mathcal{S}, p_x(\mathbf{x})] = \mathbf{0}.$$

where $\mathbf{g}[\cdot,\cdot]$ determines *what* information should be captured about $\mathbf{x}$. It is possible to meet this condition and still have some degree of freedom in the choice of the points. This ambiguity can be resolved by assigning a penalty function $c[\mathcal{S}, p(\mathbf{x})]$ to the different solutions. The purpose of this function is to incorporate features which are desirable, but do not necessarily have to be met. As the value of the penalty function increases, the solution becomes less desirable.

[1]It is frequently stated that a necessary condition for the Kalman filter is that all of the process and observation noises are Gaussian. However, this is not true.

The sigma point set which is used is that which is most desirable and confirms to the necessary conditions. In other words, the sigma points are given by the solution to the equation

$$\min_{\mathcal{S}} c\left[\mathcal{S}, p_x(\mathbf{x})\right] \text{ subject to } \mathbf{g}\left[\mathcal{S}, p_x(\mathbf{x})\right] = \mathbf{0}.$$

Given the set of points, each point is instantiated through the nonlinear function, $\mathcal{Y}_i = \mathbf{f}[\mathcal{X}_i]$. The appropriate statistics for $\mathbf{y}$ are approximated from the set $\{W_i, \mathcal{Y}_i\}$.

In this paper, we are concerned with the propagation of the first two moments through the nonlinear transformation. Therefore, the constraint equation is

$$\mathbf{g}\left[\mathcal{S}, p_x(\mathbf{x})\right] = \begin{bmatrix} \sum_{i=0}^{p} W_i - 1 \\ \sum_{i=0}^{p} W_i \mathcal{X}_i - \bar{\mathbf{x}} \\ \sum_{i=0}^{p} W_i \left\{\mathcal{X}_i - \bar{\mathbf{x}}\right\}\left\{\mathcal{X}_i - \bar{\mathbf{x}}\right\}^T - \mathbf{P}_{xx} \end{bmatrix}$$

and the mean and covariance of the transformed set are

$$\bar{\mathbf{y}} = \sum_{i=0}^{p} W_i \mathcal{Y}_i \tag{2}$$

$$\mathbf{P}_{yy} = \sum_{i=0}^{p} W_i \left\{\mathcal{Y}_i - \bar{\mathbf{y}}\right\}\left\{\mathcal{Y}_i - \bar{\mathbf{y}}\right\}^T. \tag{3}$$

In [7] we examined the following symmetrically-distributed set of points which match the mean and covariance:

$$\begin{aligned} \mathcal{X}_0(k \mid k) &= \hat{\mathbf{x}}(k \mid k) \\ \mathcal{X}_i(k \mid k) &= \hat{\mathbf{x}}(k \mid k) + \left(\sqrt{(n+\kappa)\mathbf{P}(k \mid k)}\right)_i \\ \mathcal{X}_{i+n}(k \mid k) &= \hat{\mathbf{x}}(k \mid k) - \left(\sqrt{(n+\kappa)\mathbf{P}(k \mid k)}\right)_i \\ W_0 &= \kappa/(n+\kappa) \\ W_i &= 1/\{2(n+\kappa)\} \\ W_{i+n} &= 1/\{2(n+\kappa)\} \end{aligned} \tag{4}$$

where $\kappa \in \Re$, $\left(\sqrt{(n+\kappa)\mathbf{P}(k \mid k)}\right)_i$ is the ith row or column of the matrix square root of $(n+\kappa)\mathbf{P}(k \mid k)$ and W_i is the weight that is associated with the ith point. κ scales the third and higher order terms of this set and, if $(n+\kappa) = 3$, it is possible to match some of the fourth order terms when $\mathbf{x}$ is Gaussian [8].

The computational costs of the UT are directly proportional to the number of sigma points which are used. Therefore, minimising the number of sigma points minimizes the computational costs. Such considerations are crucial if the process model is expensive to calculate or if real-time performance is required. In [9] we proved that, for an n-dimensional space, only $n+1$ points are required. We also developed the *minimal skew* sigma points. These points have the property that they minimise the third order moments. The algorithm for choosing these points is given in Figure 1.

Although these points have the same accuracy as the symmetric set, they have half the computational costs. However, there is a problem with this set: for a state of n dimensions, the weights on each point vary by a factor of 2^n. Furthermore, the values on the coordinates for each point vary by a factor of $2^{n/2}$. Therefore, for even moderately sized systems, this set of sigma points can lead to significant numerical problems.

To develop a set of points which have better behaviour at higher dimensions, we explored an alternative hypothesis: why not place all the sigma points (apart from the 0th point) on a hypersphere centered about the origin?

1) Choose $0 \leq W_0 \leq 1$.
2) Choose weight sequence:

$$W_i = \begin{cases} \frac{1-W_0}{2^n} & \text{for } i = 1 \\ W_1 & \text{for } i = 2 \\ 2^{i-2} W_1 & \text{for } i = 3, \ldots, n+1 \end{cases}$$

3) Initialize vector sequence as:

$$\mathcal{X}_0^1 = [0], \; \mathcal{X}_1^1 = \left[-\frac{1}{\sqrt{2W_1}}\right] \text{ and } \mathcal{X}_2^1 = \left[\frac{1}{\sqrt{2W_1}}\right]$$

4) Expand vector sequence for $j = 2, \ldots, n$ according to

$$\mathcal{X}_i^{j+1} = \begin{cases} \begin{bmatrix} \mathcal{X}_0^j \\ 0 \end{bmatrix} & \text{for } i = 0 \\ \begin{bmatrix} \mathcal{X}_i^j \\ -\frac{1}{\sqrt{2W_{j+1}}} \end{bmatrix} & \text{for } i = 1, \ldots, j \\ \begin{bmatrix} \mathbf{0}_j \\ \frac{1}{\sqrt{2W_{j+1}}} \end{bmatrix} & \text{for } i = j+1 \end{cases}$$

Fig. 1. The Point Selection Algorithm for the Minimal Skew Simplex Unscented Transform for an n dimensional system.

III. The Spherical Simplex Points

The spherical simplex sigma set of points consists of points which lie on the origin or on a hypersphere centered at the origin. We derive this set of sigma points using a similar procedure to that described in [10].

Let $\mathcal{X}_i^j$ be the ith sigma point in the set for the jth dimensional space. It is assumed, without loss of generality, that $\bar{\mathbf{x}} = \mathbf{0}$ and $\mathbf{P}_{xx} = \mathbf{I}$ (the $n \times n$ identity matrix)[2].

First consider the problem of choosing a set of points which capture mean and covariance in a single dimension, $\mathbf{e}_1$. Three points are used: $\mathcal{X}_0^1 = [0]$, $\mathcal{X}_1^1 = [-x_1]$ and $\mathcal{X}_2^1 = [x_2]$. The weights for these points are W_0, W_1 and W_2. From the condition that the means and covariances must be $\bar{\mathbf{x}}$ and $\mathbf{P}_{xx}$,

$$W_0 + W_1 + W_2 = 1 \tag{5}$$

$$-W_1 x_1 + W_2 x_2 = 0 \tag{6}$$

$$W_1 x_1^2 + W_2 x_2^2 = 1 \tag{7}$$

There are four unknowns but only three constraints and so these conditions are not sufficient. From the assumption that the simplex points lie at the same distance from the origin, we add the constraint that $x_1 = x_2$. From Equations 5 and 6, $W_1 = W_2 = (1 - W_0)/2$. Substituting into Equation 7, $x_1 = 1/\sqrt{2W_1}$. W_0 is a free parameter whose value affects the fourth and higher moments of the sigma point set. For space reasons it is not possible to analyse the full impact of the term, and we refer the reader to [9] and [8] for a fuller analysis.

To extend the set to two dimensions, points 1 and 2 are translated in the $\mathbf{e}_2$–direction by $-x_3$ and a new point — labelled 3 — is added at $(0, s_3 x_3)$ with weight W_3 and s_3 is a scaling factor. The constraints above ensure that the mean and covariance constraints

[2] The set can be transformed to arbitrary mean and covariance through the appropriate affine transformation. If a random variable $\mathbf{z}$ has mean $\bar{\mathbf{z}}$ and covariance $\mathbf{P}_{zz}$, the ith sigma point is

$$\mathcal{Z}_i^j = \bar{\mathbf{z}} + \sqrt{\mathbf{P}_{zz}}\mathcal{X}_i^j$$

where $\sqrt{\mathbf{P}_{zz}}$ is a matrix square root of $\mathbf{P}_{zz}$.

Proceedings of the American Control Conference
Denver, Colorado June 4-6, 2003

1) Choose $0 \leq W_0 \leq 1$.
2) Choose weight sequence:

$$W_i = (1 - W_0)/(n+1).$$

3) Initialize vector sequence as:

$$\mathcal{X}_0^1 = [0], \ \mathcal{X}_1^1 = \left[-\frac{1}{\sqrt{2W_1}}\right] \text{ and } \mathcal{X}_2^1 = \left[\frac{1}{\sqrt{2W_1}}\right]$$

4) Expand vector sequence for $j = 2, \ldots, n$ according to

$$\mathcal{X}_i^j = \begin{cases} \begin{bmatrix} \mathcal{X}_0^{j-1} \\ 0 \end{bmatrix} & \text{for } i = 0 \\ \begin{bmatrix} \mathcal{X}_i^{j-1} \\ -\frac{1}{\sqrt{j(j+1)W_1}} \end{bmatrix} & \text{for } i = 1, \ldots, j \\ \begin{bmatrix} \mathbf{0}_{j-1} \\ \frac{j}{\sqrt{j(j+1)W_1}} \end{bmatrix} & \text{for } i = j+1 \end{cases}$$

Fig. 2. The Point Selection Algorithm for the Spherical Simplex Unscented Transform for an n dimensional system.

are maintained in the $\mathbf{e}_1$ direction. The normalisation, mean and covariance constraints in the $\mathbf{e}_2$ direction are:

$$W_0 + 2W_1 + W_3 = 1, \quad (8)$$
$$-2W_1 x_3 + W_3 s_3 x_3 = 0, \quad (9)$$
$$2W_1 x_3^2 + W_3 s_3^2 x_3^2 = 1. \quad (10)$$

Because each point (apart from the 0th point) lies at the same distance from the origin, we assume that the weight on each point is the same and so $W_1 = W_3$. From Equation 9, $s_3 = 2$. Substituting into Equation 10 and rearranging,

$$6W_1 x_3^2 = 1. \quad (11)$$

The equi-distance constraint is satisfied if the distance of point 3 and point 1 from the origin is the same. This constraint is given by,

$$s_3^2 x_3^2 = x_1^2 + x_3^2. \quad (12)$$

Substituting $s_3 = 2$, $x_1^2 = 1/2W_1$ and $x_3^2 = 1/6W_1$, this equation is satisfied. In other words, assuming equal weights is sufficient to guarantee that all points lie at the same distance from the origin.

The general algorithm is summarised in Figure 2. This algorithm demonstrates two issues. First, the weight on each sigma point (apart from the 0th point) is the same and is proportional to $(1 - W_0)/(n+1)$. Second, all of the sigma points (apart from the 0th point) lie on the hypersphere of radius $\sqrt{n}/(1 - W_0)$.

IV. Example

We illustrate the use of the spherical simplex sigma point algorithm in an example using Simultaneous Localisation and Map Building (SLAM). A mobile robot is transported into an environment which is populated by a set of navigable features, canonically referred to as beacons. The robot uses onboard sensors to detect those beacons relative to the robot. These observations are used to initialise new beacons in the map, or update existing beacon and robot locations.

This is a benchmark problem from the field of mobile robotics. It is of great practical importance and is an example of a high order, nonlinear system with which most of the states are parameters whose states do not vary[3].

The structure of a joint vehicle-beacon system is as follows. The state of the vehicle at timestep k is $\mathbf{x}_v(k)$ and the state of the ith beacon is $\mathbf{p}_i(k)$. The complete state space for a system which comprises of n beacons is [11]:

$$\mathbf{x}_n(k) = \left[\mathbf{x}_v^T(k) \ \mathbf{p}_1^T(k) \ \ldots \ \mathbf{p}_n^T(k)\right]^T. \quad (13)$$

The mean and covariance of this estimate are

$$\hat{\mathbf{x}}_n^T(k \mid k) = \left[\hat{\mathbf{x}}_v^T(k \mid k) \ldots \mathbf{p}_T^n(k) k\right]^T \quad (14)$$

$$\mathbf{P}_n(k \mid k) = \quad (15)$$
$$\begin{pmatrix} \mathbf{P}_{vv}(k \mid k) & \mathbf{P}_{v1}(k \mid k) & \ldots & \mathbf{P}_{vn}(k \mid k) \\ \mathbf{P}_{1v}(k \mid k) & \mathbf{P}_{11}(k \mid k) & \ldots & \mathbf{P}_{1n}(k \mid k) \\ \mathbf{P}_{2v}(k \mid k) & \mathbf{P}_{21}(k \mid k) & \ldots & \mathbf{P}_{2n}(k \mid k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{nv}(k \mid k) & \mathbf{P}_{n1}(k \mid k) & \ldots & \mathbf{P}_{nn}(k \mid k) \end{pmatrix}$$

where $\mathbf{P}_{vv}(k \mid k)$ is the covariance of the robot's position estimate, $\mathbf{P}_{ii}(k \mid k)$ is the covariance of the position estimate of the ith beacon and $\mathbf{P}_{ij}(k \mid k)$ is the cross-correlation between the estimate of the ith and jth beacons.

By assumption, all of the beacons are stationary and no process noise acts upon them.

The vehicle model is the standard equation for a steered bicycle [12]:

$$\mathbf{x}_v(k+1) = \begin{pmatrix} x_v(k) + V(k)\Delta T \cos(\delta(k) + \theta_v(k)) \\ y_v(k) + V(k)\Delta T \sin(\delta(k) + \theta_v(k)) \\ \theta_v(k) + \frac{V(k)\Delta T \sin(\delta(k))}{B} \end{pmatrix},$$

where the position and orientation of the vehicle are (x_v, y_v, θ_v), the timestep is ΔT, the control inputs are the wheel speed $V(k)$ and steer angle $\delta(k)$ and the vehicle wheel base (distance between front and rear wheels) is B. The process noises are additive disturbances which act on $V(k)$ and $\delta(k)$. These are modelled as Gaussian, zero-mean disturbances with standard deviations 0.1ms^{-1} and $0.5°$ respectively. The value $B = 1m$ was adopted throughout this simulation.

The robot observes the range (r) and bearing (α) of beacons with respect to a platform-mounted sensor. The observation model is

$$\mathbf{h}_i[\mathbf{x}_v(k), \mathbf{p}_i(k), \mathbf{w}(k)] = \begin{bmatrix} \sqrt{(x_i - x_v)^2 + (y_i - y_v)^2} \\ \tan^{-1}\left(\frac{y_i - y_v}{x_i - x_v}\right) - \theta_v \end{bmatrix} \quad (16)$$

The specifications of the observation noises were taken from the specifications of the SICK™laser. They are modelled as zero-mean additive Gaussian terms with standard deviations 0.04m and 0.5°.

This equation is used in a conventional Kalman filter whenever a beacon is observed. A new beacon is initialised using the inverse of Equation 16:

$$\mathbf{g}_i[\mathbf{x}_v(k), \mathbf{w}(k)] = \begin{bmatrix} x_v(k) + r(k)\cos[\alpha(k)] \\ y_v(k) + r(k)\sin[\alpha(k)] \end{bmatrix}. \quad (17)$$

The vehicle also possesses an idealised compass which is capable

[3] It should be noted that the robotics community is aggressively tackling the issues of the computational and storage costs for this problem. It is not in the scope of this paper to discuss these issues and so we only consider the simplest and most computationally intensive implementation of the filter.

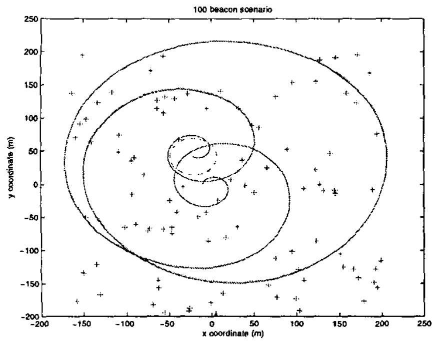


Fig. 3. An example scenario. The scenario consists of 100 beacons whose positions are denoted by crosses. The vehicle moves on the spiral path. The dashed circle shows the maximum sensor detection range. This is automatically calculated so that on average 1 beacon is visible per time step.

of measuring absolute orientation[4]. The standard deviation of this sensor is 2°.

Figure 3 shows the path the vehicle takes through the beacons. It starts at the origin and slowly spirals outwards. It then spirals inwards again. The scenarios vary according to the number of beacons (and hence the value of n).

A. 10 Beacon Case

In this situation $n = 23$ and the sensor detection range is 87.4m. Figure 4 (see next page) shows the results for the symmetric, skew and simplex forms. Each figure shows the 2σ bounds and estimation errors for x_v. As can be seen, the errors lie within the bounds. Furthermore, all three algorithms perform very similarly. The substantial variation in the position variance is a reflection of the fact that, as the vehicle moves, the constellation of visible beacons changes.

B. 11 Beacon Case

In this situation $n = 25$ and the sensor detection range is 83.3m. Figure 5 shows the results for the symmetric, skew and simplex forms. As before, each figure shows the 2σ bounds and estimation errors for x_v. Despite the fact that the number of states has increased by two, the results are dramatic: the skew filter fails catastrophically. Its estimates diverge. The reason is that the scaling on the skew point is 2^{25} or 33,554,432. This introduces rounding errors which break the subtle correlation structure needed to support SLAM.

C. 100 Beacon Case

The final example shows the behaviour when there are 100 beacons. In this situation, $n = 203$ and the sensor detection range is 27.6m. Figure 6 shows the results for the symmetric, and spherical forms. As can be seen, both algorithms continue to perform well despite the high dimensionality of the state space. The large step-change in the covariance at about time step 1750 occurs when the vehicle, returning on its circular path, encounters a set of beacons it had initialised as it was first departing the origin. The covariance on those beacons is much smaller than that on the vehicle, leading to the marked reduction in the vehicle position covariance.

[4]The reason why this sensor is needed is a direct result of linearisation errors. In [10] we algebraically proved that, for the system described here, the EKF is guaranteed to yield inconsistent estimates. An empirical study indicated that an absolute orientation sensor was required. Analysis with the UKF (not presented here) suggests that the second order accuracy provided by the UKF is still not sufficient to guarantee a consistent estimate and so it is used here

V. Conclusions

This paper has derived a new sigma point selection strategy, the spherical simplex set of points. These points have the property that, given a state of n dimensions, $n+1$ points are sufficient to capture the mean and covariance. Utilising earlier results, these points give second order accuracy in both the mean and the covariance. They also have the advantage that they are numerically better behaved than other sigma points.

References

[1] H. W. Sorenson, Ed., *Kalman filtering: theory and application.* Piscataway NJ, USA: IEEE Press, 1985.

[2] D. Lerro and Y. K. Bar-Shalom, "Tracking with debiased consistent converted measurements vs. ekf," *IEEE Transactions on Aerospace and Electronics Systems*, vol. AES-29, no. 3, pp. 1015–1022, July 1993.

[3] M. Athans, R. P. Wishner and A. Bertolini, "Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements," *IEEE Transactions on Automatic Control*, vol. TAC-13, no. 6, pp. 504–518, October 1968.

[4] F. E. Daum, "New Exact Nonlinear Filters," in *Bayesian Analysis of Time Series and Dynamic Models*, J. C. Spall, Ed. Marcel Drekker, Inc., 1988, ch. 8, pp. 199–226.

[5] N. J. Gordon, D. J. Salmond and A. F. M. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEE Proceedings-F*, vol. 140, no. 2, pp. 107–113, April 1993.

[6] S. Thrun D. Fox, W. Burgard and F. Dellaert, "Robust monte carlo localization for mobile robots," *Artificial Intelligence*, no. 101, pp. 99–141, 2000.

[7] S. J. Julier and H. F. Durrant-Whyte, "Navigation and Parameter Estimation of High Speed Road Vehicles," in *Robotics and Automation Conference, Nagoya, Japan.*, 1995, pp. 101–105.

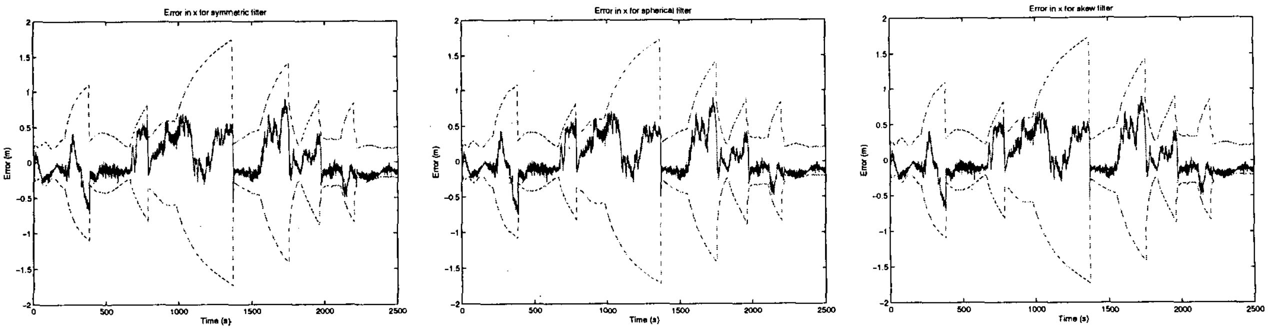
[8] S. J. Julier, J. K. Uhlmann and H. F. Durrant-Whyte, "A new approach for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, March 2000.

[9] S. J. Julier, and J. K. Uhlmann, "Reduced Sigma Point Filters for the Propagation of Means and Covariances through Nonlinear Transformations," in *Proceedings of the IEEE American Control Conference.* Anchorage AK, USA: IEEE, 8–10 May 2002, pp. 887–892.

[10] S. J. Julier and J. K. Uhlmann, "A Counter Example to the Theory of Simultaneous Localization and Map Building," in *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, Seoul, Korea, 21–26 May 2001.

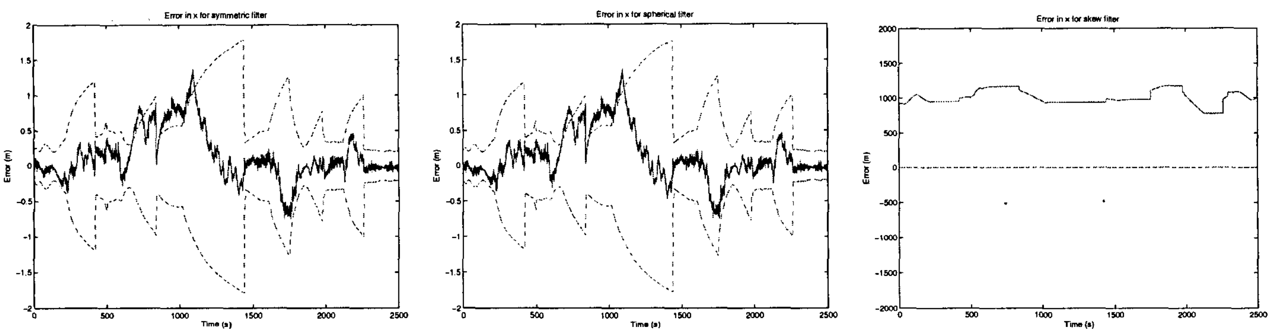
[11] R. Smith, M. Self and P. Cheeseman, *Estimating Uncertain Spatial Relationships in Robotics, Autonomous Robot Vehicles.* Springer-Verlag, 1989.

[12] H. F. Durrant-Whyte, "An autonomous guided vehicle for cargo handling applications," *International Journal of Robotics Research*, vol. 15, no. 5, pp. 407–440, October 1996.



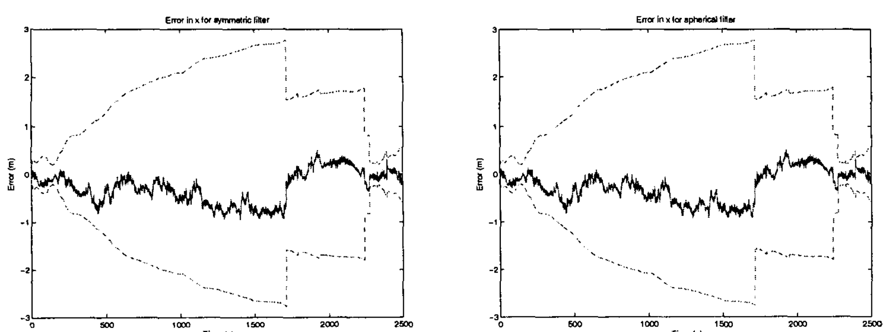
(a) Symmetric Sigma Points (b) Spherical Sigma Points (c) Skew Sigma Points

Fig. 4. The error in the estimate of x_v (solid) and the $\pm 2\sigma$ bounds (dashed)for the 10 beacon case.



(a) Symmetric Sigma Points (b) Spherical Sigma Points (c) Skew Sigma Points

Fig. 5. The error in estimate of x_v (solid) and the $\pm 2\sigma$ bounds (dashed) for the 11 beacon case. Note that the vertical scale for the skew filter is much larger (-2000 to +2000) than that for the other two filters (-2 to 2).



(a) Symmetric Sigma Points (b) Spherical Sigma Points

Fig. 6. The error in the estimate of x_v (solid) and the $\pm 2\sigma$ bounds (dashed) for the 100 beacon case. The skew sigma point filter is divergent and its results are not shown.