

Unscented Kalman Filtering for Visual-Inertial
Navigation with Applications to Small Unmanned
Aircraft Navigation

Virginia Tech

Joshua Galen Eddy

April 1, 2017

Chapter 1

Introduction

Throughout this document, there are several sections whose titles have been placed in parentheses. These sections have been included to give added context to the work, but are not strictly necessary to the reader's understanding.

1.1 (Personal Motivation)

I first took an interest in unmanned aircraft in the fall of 2012, my sophomore year of college. In search of an exciting engineering challenge, several of my friends and I founded the Cooperative Autonomous Robotics Design (CARD) team at Virginia Tech. Our core team consisted of a dozen students devoted to designing and competing with drones and other robotic vehicles. Our team, guided by my future graduate adviser Kevin Kochersberger, entered a number of design competitions and brought home several awards for the university. My early experiences with the team brought me into contact with microcontroller programming, Proportional-Integral-Derivative (PID) controller design, mechatronics, and computer-aided design (CAD) modeling.

After two years of involvement with the CARD team, I applied for an internship at the National Institute of Aerospace¹ (NIA). In the summer of 2014, I was part of a team of NIA

¹<http://www.nianet.org>

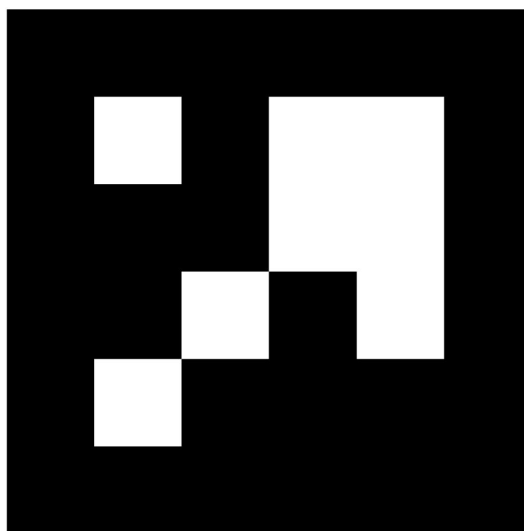


Figure 1.1: Example of an April tag from the `april_tags_ros` ROS package.

researchers working on the Flying Donkey Challenge², an international engineering competition centered around the idea of “flying donkeys,” full-sized autonomous airplanes capable of quickly carrying cargo between small airports in rural Africa. This competition, unfortunately now defunct, was divided into a number of sub-challenges focusing on different technical objectives such as precision landing and collision avoidance. Our team’s goal was to design an inexpensive system for GPS-denied navigation that could reliably guide unmanned aircraft during a GPS blackout. This project introduced me to many of the technologies and techniques that would later become my major research interests, particularly the Robot Operating System³ (ROS), Kalman filtering, and sensor fusion.

My internship at the NIA brought me into contact with Dr. Danette Allen, head of the NASA Langley Autonomy Incubator. During the 2014–15 academic year, Dr. Allen sponsored the CARD team to design and build two autonomous multirotor delivery drones. These aircraft were capable of delivering 5-lb packages to distances of up to 2.5 miles (or 5 miles, round trip). In addition, these vehicles were able to land precisely on 1

²<http://www.flyingdonkey.org>

³<http://wiki.ros.org>

m² April tags such as that found in Figure 1.1⁴. Following the completion of this project, I worked as a summer intern at the Autonomy Incubator.

During the summer of 2015, I began the research that eventually evolved into my thesis project, studying Visual-Inertial Navigation (VIN) and the unscented Kalman filter. In reading up on the unscented Kalman filter (UKF), I took a serious interest in the design of the algorithm. Unlike many other formulations of the Kalman filter, the UKF has a notably limited dependence on information about the system whose state is being estimated (this *system agnosticism* is discussed in detail later on in Chapter 3). In learning about the UKF, I became excited by the idea of taking advantage of this trait to build a minimalistic software interface by which a wide variety of disparate systems could one day be tracked and studied in a ROS framework. I envisioned a kind of “one-stop shopping” experience for massively reusable and customizable filtering profiles that could fulfill the needs of a great many researchers and roboticists with little overhead in the way of state estimation knowledge. This vision eventually drove my development of the `kalman_sense` ROS package and cemented my interest in state estimation and controls.

1.2 Project Overview

1.3 Organization of this Document

Prior Work

In Prior Work, we explore recent contributions to loosely coupled filter-based navigation and estimation. We focus primarily on a number of impactful publications coming from ETH Zurich’s Autonomous Systems Lab (ASL) and the University of Pennsylvania’s GRASP Lab. We define the current state of the art in filter-based navigation and establish the research context in which this work exists.

⁴http://wiki.ros.org/apriltags_ros

Algorithm Design and Implementation

Because of the naturally computational nature of this work, we explore in detail the design and implementation of the `kalman_sense` ROS package. We discuss plant model abstraction as well as code organization and data flow. We summarize the process by which one could extend `kalman_sense`'s functionality and the advantages of system-agnostic algorithm design.

Experimental Design

In this section, we first establish the goals of the testing regimen and then discuss the real-world execution of these goals. We discuss important statistical methods for characterizing the system's effectiveness as well as data collection procedures and post-processing. Physical testing infrastructure is explored in detail.

Experimental Results

In Experimental Results, we evaluate the system's performance during testing and seek out any limiting factors that influence estimation accuracy. We probe for possible improvements to the algorithm and provide a notional understanding of the system's theoretical effectiveness in real-world scenarios.

Conclusions

We briefly summarize the contributions made by this work, the effectiveness of the `kalman_sense` package, and any miscellaneous insights acquired during programming and testing.

Future Work

In Future Work, we expand upon the possible improvements proposed in Experimental Results and also offer a number of applications for this work. Specific examples of heterogeneous fleet management and unmanned traffic management (UTM) are explored.

Chapter 2

Prior Work

In [1], Simon Julier and Jeffrey Uhlmann discuss the development and demonstration of a new nonlinear estimator for the Kalman Filter. Recognizing that most applications for autonomous navigation are fundamentally nonlinear in both their dynamics and their observation models, Julier and Uhlmann proposed the use of a set of discretely sampled “sigma points” to determine the mean and covariance of a probability distribution. By recasting the prediction and correction steps of the Kalman filter in the form of unscented transforms (UTs), this new filter eliminates the need to calculate Jacobian matrices. Julier and Uhlmann argued that for this reason their formulation was easier to implement than the Extended Kalman Filter (EKF) and went on to suggest that its use could supplant the EKF in virtually all applications, linear or nonlinear.

In [2] (@year?), Julier acknowledges that the Kalman Filter (KF) has been successfully used in many nonlinear scenarios, but he also points out that the KF’s use of only the first two moments of the state estimate sigma points results in neglect of all higher order information (i.e., third order moments, or “skew”), a potentially rich source of new and useful information relating to symmetry of the state estimate. By extending the sigma point selection scheme of the conventional unscented transformation (UT), the author presents a new and tractable but computationally complex extension of the KF algorithm that can predict not only the first two moments of a sigma point distri-

bution (mean and covariance), but also the third moment (the so-called “skew”). Though formulated initially for unimodal distributions, the approach can, with additional mathematical considerations, be generalized for use with multimodal distributions. The author’s contention was that use of higher order information could promote better performance levels in autonomous vehicle navigation. The utility of maintaining and utilizing higher order information through the use of skewed filtering was assessed by the authors in a realistic tracking scenario. However, the results turned out to be somewhat disappointing as the change in performance turned out to be minimal, presumably due to the linearity of the KF’s update rule. Accordingly, research in this area continues, including examination of the use of nonlinear update rules in the filtering process.

In [Julier 2002] Simon Julier describes a novel approach to modifying the “unscented transformation” (UT) state estimation method, a subject of one of his previous studies in collaboration with Jeffrey Uhlmann. In the current paper, however, Julier takes the additional step of introducing a new framework for scaling sigma points as part of the state estimation process. The general framework of the new methodology allows preservation of the first two moments of any set of sigma points, thus providing a construct for limiting values to either the conventional unscented transform or the modified (scaled) transform. Providing detailed mathematical validations, the author shows that the new scaling algorithm is computationally manageable in that it is, in essence, little more than the conventional unscented transformation algorithm with the addition of a simple post-processing step, the only difference being the inclusion of an extra correction term. Thus, the new algorithm’s computational and storage costs are similar to that of the non-scaled transformation. The performance level of the scaled UT is thus demonstrably superior when the conventional non-scaled UT is used for propagating the two lower-order moments of a sigma point distribution. [Julier 2004] Julier and Uhlmann discuss the application of the extended Kalman filter (EKF) as an estimation algorithm and the associated difficulties. The EKF contains inherent problems and works best with linear systems. These limitations led to the development of the unscented transformation

(UT) for nonlinear applications. In this paper, Julier and Uhlmann describe the unscented transformation (UT) and its benefits including easier implementation and improved accuracy. The UT offers greater accuracy and reliability by applying higher order information using sigma points to the traditional mean and covariance information associated with linear applications. The authors provide examples, which may be tailored, that show how UT overcomes the limitations of the EKF.

[Mourikis 2007] Mourikis et al. expand investigation of the familiar topic of Vision-aided Inertial Navigation, with particular emphasis on small, low-cost, low-weight systems, including unmanned aerial vehicles (UAVs) using a monocular camera. Because technical advances have been made in the manufacture of inertial sensors, these robotic systems are capable of producing high precision state estimations even in uncontrolled urban environments, without utilizing a 3D feature position in the filter state vector of the Extended Kalman Filter (EKF). The exceptional accuracy produced in this real-time, real-world pose estimation model is highly desirable because it produces images that are rich in high-dimensional measurements. This robustness, however, invariably comes at a cost of computational complexity in the EKF-based algorithm, a trade-off that is unavoidable in the present state-of-the-art for vision-aided inertial navigation systems.

[Safonov 2007] Safonov revisits the decade of the 1970s, a period of transition in the theoretical framework underlying effective control of dynamic motion, marking a paradigm shift from a long-held theoretical construct centered on the "optimality" of Linear-Quadratic-Gaussian (LQG) feedback design theory to one centered on "robustness". The author observes that the shift stems from the poor performance of the LQG controllers involving early tests in real-world scenarios, a problem later recognized as an inadequate level of attention to the real-world need for high-level performance with respect to accuracy, flexibility, and overall robustness in motion control. A more pragmatic approach to motion control theory was thus born, with significant milestones being achieved throughout the 1970s, including advancements in the areas of diagonally structured uncertainty, LQ multivariable gain margins, and LQ phase-margins, to name

just a few. Further advancements came in the 1980s, culminating in the publication of the "Robust Control Toolbox" in 1988, thus making robust control design techniques widely available to students and engineers alike.

[Klein, 2007] In most Augmented Reality (AR) camera pose estimate scenarios, there exists at least some knowledge of the user environment. Complications arise when this is not the case. In this paper, Klein and Murray propose a method in which a user can track a calibrated hand-held camera in an unknown environment, while still building a map of the environment. Primarily useful only for small AR workspace areas, this design splits the tracking and mapping functions into two separate tasks. These two tasks are processed on a dual-core computer utilizing "parallel threads", with one thread directly tracking erratic motion of the hand-held camera while the other thread simultaneously produces a 3D map of the environment. In this approach, the authors estimate a dominant plane from mapped points and then populate it with virtual characters. To simplify the computationally complex tasks of simultaneously tracking and mapping, the nature of the scene to be tracked and mapped was constrained: it should be small and, for the most part, static. Experiments demonstrated that the system could produce highly accurate maps, rich in detailed landmark information trackable at frame-rate, all without prior knowledge of the scene and with minimal initialization procedures.

In [Weiss2010], Weiss et al. discuss how autonomous micro aerial vehicles (MAVs) provide access to environments that are otherwise difficult to traverse. Further, they mitigate the risk to people and the environment. For many applications, the MAV must navigate in a GPS-denied environment. This paper presents an approach to navigation that relies on a micro helicopter, a single camera, and onboard inertial sensors. A monocular simultaneous localization and mapping (SLAM) framework stabilizes the vehicle, which is used to overcome the issues with drift. This research is important because urban canyons limit GPS availability. The authors, in this paper, show that autonomous navigation in a GPS-denied environment is achievable.

[Shen 2011] Shen et al. extend the work of other authors on the topic of autonomous

MAV navigation, particularly as it relates to stable indoor flight and GPS-denied localization in constrained multi-floor environments. The research distinguishes itself by emphasizing the use of onboard sensors only, as well as fully autonomous, real-time internal computational capabilities, with no hands-on user interaction beyond basic high-level commands. The research extends to multi-floor MAV operation with loop closure. It also addresses specially designed controllers to help compensate for sudden changes in wind velocity and air flow as the MAV traverses constrained low-clearance areas with potentially strong aerodynamic disturbances.

[Weiss2011] Weiss and Roland Siegwart developed an algorithm that provides a metric scale to estimate that estimates monocular visual odometry or monoSLAM approaches. The authors accomplished the development of the metric scale by the addition of an inertial sensor with a three-axis accelerometer and gyroscope. Stephan Weiss and Roland Siegwart created a modular solution that is based on an Extended Kalman Filter (EKF) and provides both simulated results and data-based results. In this paper, the authors discuss their unique approach, its applications, versatility, and reliability of their estimating algorithm for visual odometry, such as visual SLAM, in real-time.

[Weiss 2012] Weiss et al present a versatile framework to enable autonomous flight of a Micro Aerial Vehicle (MAV). The MAV has only slow, noisy, delayed and possibly arbitrarily scaled measurements available. The use of these measurements directly for position control is not practical since MAVs exhibit great agility in motion. In addition, these measurements often come from a selection of different onboard sensors, hence accurate calibration is crucial to the robustness of the estimation processes. In this article, Weiss, et al address the problems using an EKF formulation which fuses these measurements with inertial sensors. The authors not only estimate pose and velocity of the MAV, but also estimate sensor biases, scale of the position measurement and self (inter-sensor) calibration in real-time. The authors show that it is possible to obtain a yaw estimate from position measurements only. The authors demonstrate that the proposed framework is capable of running entirely onboard a MAV performing state prediction at

the rate of 1kHz. Their results illustrate that this approach is able to handle measurement delays (up to 500ms), noise (std. deviation up to 20 cm) and slow update rates (as low as 1 Hz) while dynamic maneuvers are still possible. Stephan Weiss et al present a detailed quantitative performance evaluation of the real system under the influence of different disturbance parameters and different sensor setups to highlight the versatility of our approach.

[Weiss, Achtelik 2012] Weiss et al. explore the advantages of utilizing a high-performance navigation algorithm on a low-cost, low-weight micro aerial vehicle (MAV) equipped with a single camera and an inertial measurement unit (IMU) capable of both onboard processing and real-time operations, with focus on a speed estimation module to help control the speed of the MAV, all within an Extended Kalman Filter framework. The system was shown to be useful for real-time self-calibration of the sensor suite – critical to ensuring the robustness and flexibility of any state estimation process – and as a potential solution to some of the tracking failures common to keyframe-based modules.

[Huang 2013] Huang et al. explore solutions to two Unscented Kalman Filter (UKF) limitations that exist in current state-of-the-art Simultaneous Localization and Mapping (SLAM) systems. Specifically, the authors address the problems of cubic complexity in the number of state pose estimates, and the inconsistencies in those estimates caused by a mismatch between the observability properties of statistically-linearized UKF systems and the observability properties of nonlinear systems. To address the problem of cubic complexity, the authors introduce a novel sampling strategy which produces a constant computational cost which, while linear in the propagation phase, is quadratic in the update phase. Although this new sampling strategy was primarily proposed for resolving the above-referenced SLAM problem, it also has potential usefulness in other nonlinear estimation applications. To address the problem of inconsistency in state estimations, the authors propose a new UKF algorithm which, due to the imposition of observability constraints, ensures that the linear regression computations of the modified UKF system produce results similar to those of nonlinear SLAM systems and, in the process, provide

improved accuracy and consistency in state estimations. Importantly, these results have been validated with both real-world and simulation experiments. While the paper focused on 2D SLAM, the authors contend that their proposed methodology is also useful for robot localization in 3D, using inertial sensors.

[Lynen 2013] Lynen et al. report on the development of a generic framework to overcome known limitations in fusing information transmitted from multiple sensors in the navigation of robots, focusing mainly on the navigational needs of rotor-based micro aerial vehicles (MAV) which can more easily traverse from indoor to outdoor domains. The authors demonstrate that their Multi-Sensor-Fusion Extended Kalman Filter framework is capable of processing various measurements from an unlimited number of sensors, as well as sensor types, while simultaneously performing online self-calibrations of the overall sensor suite. Designed to be modular, the framework allows seamless handling of sensor signals during operation while performing other complex, iterative calculations to achieve near optimal linearization points for state updates.

[Engel 2013] Engel et al. propose a novel direct monocular Simultaneous Localization and Mapping (SLAM) algorithm unlike that of existing direct approaches which embrace pure visual odometry. The novelty of the authors's approach is that it permits the building of consistent, accurate, large-scale 3D maps of the environment while simultaneously tracking camera motion, incorporating any scale-drift in the environment and allowing for the detection and correction of any accumulated drift. The system is capable of running real-time on a central processing unit and as visual odometry on a modern smartphone.

[Rogers III 2014] Rogers et al. presented a methodology for overcoming some of the constraining conditions encountered in a GPS-guided autonomous robotic system, such as occlusion (blocking of GPS signals) and multipath (reception of indirect signals due to environmental reflections) and potentially to ameliorate the effects of jamming or spoofing resulting from adversarial activities. Specifically, the methodology incorporated GPS measurements into a feature-based mapping system, thus providing geo-referenced

coordinates allowing for better execution of high-level missions and providing the ability to correct accumulated mapping errors over the course of long-term operations in both indoor and outdoor environments.

[Hesch 2014] Hesch et al. investigated the relationship between system observability properties and estimator inconsistency for a Vision-aided Inertial Navigation System (VINS). By factorizing the observability matrix according to the observable and unobservable modes, they produced a new methodology for determining the unobservable directions of nonlinear systems. They then applied this new method to the VINS nonlinear model to determine analytically any of its unobservable directions. Subsequently, they leveraged their analysis to improve both the accuracy and consistency of linearized estimators as they applied to VINS. Through the use of extensive simulations and validation testing, the key findings of the analysis were evaluated for the purpose of demonstrating the superior accuracy and consistency of the proposed VINS framework.

[Faessler 2015] Faessler et al. report on the development and demonstration of a low-cost, low-weight, vision-based quadrotor micro aerial vehicle (MAV) with onboard sensing, computation, and control capabilities. These onboard capabilities eliminated reliance on external positioning systems such as GPS or motion capture systems. This development moves the MAV from its current line-of-sight control state to wireless communications with the ability to execute intricate processes autonomously and to transmit live feedback to a user. Reporting on both indoor and outdoor experiments, the authors believe that such a vehicle potentially would be a great enhancement in search-and-rescue missions, disaster response, and remote inspection of terrain.

Chapter 3

Algorithm Design and Implementation

We begin by defining the state vector $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{x} \triangleq \{x, y, z, q_x, q_y, q_z, q_w, \dot{x}, \dot{y}, \dot{z}, p, q, r, \ddot{x}, \ddot{y}, \ddot{z}\}^T. \quad (3.1)$$

The covariance associated with each state \mathbf{x} is a matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$. A set of $2n + 1$ sigma points is then derived from \mathbf{x} and \mathbf{P} :

$$\begin{aligned} \chi^0 &= \mathbf{x} \\ \chi^i &= \mathbf{x} + \left(\sqrt{(n + \lambda) \mathbf{P}} \right)_i, \quad i = 1, \dots, n \\ \chi^i &= \mathbf{x} - \left(\sqrt{(n + \lambda) \mathbf{P}} \right)_{i-n}, \quad i = n + 1, \dots, 2n \end{aligned} \quad (3.2)$$

where λ

Bibliography

- [1] S. J. Julier and J. K. Uhlmann, “A New Extension of the Kalman Filter to Nonlinear Systems,” in *Proc. SPIE*, vol. 3068, pp. 182–193, 1997.
- [2] S. J. Julier, “A Skewed Approach to Filtering,” in *Proc. SPIE*, vol. 3373, pp. 271–282, 1998.