Acronyms

AR Augmented Reality

EKF Extended Kalman Filter

GPS Global Positioning System

KF Kalman Filter

IMU Inertial Measurement Unit

MSF-EKF Multi-Sensor Fusion Extended Kalman Filter

PTAM Parallel Tracking and Mapping

SLAM Simultaneous Localization and Mapping

ROS Robot Operating System

UAS Unmanned Aircraft System

UAV Unmanned Aerial Vehicle

UKF Unscented Kalman Filter

UT Unscented Transform

UTM Unmanned Traffic Management

VIN Visual-Inertial Navigation

Chapter 1

Prior Work

1.1   Development of the Unscented Kalman Filter

tk: OVERVIEW

In [1], Simon Julier and Jeffrey Uhlmann presented a nonlinear estimation approach 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 EKF and went on to suggest that its use could supplant the EKF in virtually all applications, linear or nonlinear.

In [2], Julier acknowledges that the (linear) Kalman Filter has been used successfully in many nonlinear scenarios, but notes that the use of only the first two moments of the state estimate sigma points results in the neglect of all higher order information (that is, 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 transform, Julier was able to present a tractable but computationally complex extension of the Kalman Filter that could predict not only the first two moments of a sigma point distribution but also the skew. Though formulated initially for unimodal distributions, Julier stated that the approach could, with additional mathematical considerations, be generalized for use with multimodal distributions. Julier's contention was that the 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 filter's update rule. Accordingly, research in this area continues, including examination of the use of nonlinear update rules in the filtering process.

In [3], Julier describes a novel approach to modifying the unscented transformation state estimation method. In the new approach, Julier takes the additional step of introducing a 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 to the unscaled UT for propagating the two lower-order moments of a sigma point distribution.

In [4] Julier and Uhlmann discuss the application of the EKF as an estimation algorithm and the associated difficulties in doing so. Because the EKF is fundamentally a

linearizing approach to estimation, its effectiveness is thus tied to the veracity of the local linearity assumption for the system under scrutiny. These limitations led to the development of the UT for nonlinear applications. In this paper, Julier and Uhlmann describe the 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 to various process and observation models, that show how the UT overcomes the limitations of the EKF.

1.2 Visual-Inertial Navigation (VIN)

tk: OVERVIEW

In [5], Georg Klein and David Murray proposed a method for tracking a handheld camera in unknown environments for use in small augmented reality (AR) workspaces. In contrast to many previous Simultaneous Localization and Mapping (SLAM) approaches to camera tracking, Klein and Murray split the tracking and mapping functions into two separate computational tasks. They performed these tasks on a dual-core computer utilizing parallel threads, with one thread directly tracking erratic motion of the hand-held camera and the other thread constructing a 3D map of the environment. Through the use of this Parallel Tracking and Mapping (PTAM) algorithm, Klein and Murray were able to take advantage of computationally expensive batch-optimization techniques for map reconstruction which were rarely ever used in real-time applications previously. This, in turn, allowed Klein and Murray to forego the common approach of creating a sparse map of high quality features in favor of a much denser map whose features could vary widely in quality. The resulting system could produce detailed maps tracking thousands of features at frame-rate and could recover gracefully from a variety of intermittent tracking failures. That being said, the researchers made certain relaxing assumptions regarding the scenes which would be tracked. PTAM, by nature of its orientation toward AR applications, operates best in small, static, planar environments (such as on the surface of a desk or

the floor of an office). PTAM's value to the robotics community quickly became obvious due to its independence of *a priori* knowledge of the scene and its minimal initialization procedure (explored in Chapter tk).

Four years later, Stephan Weiss et al. presented a VIN system for autonomous UAV navigation which employed PTAM [?]. The researchers presented the results of several experiments in which a UAV equipped with only a monocular camera and inertial sensors navigated through unknown environments without the aid of GPS satellites or other external sensing infrastructure. All calculations were performed online in real time using an EKF framework, proving that this minimalist combination of sensors could be employed in real-world GPS-compromised flight scenarios to great effect. At approximately the same time, Shen et al. [6] conducted experiments aimed at stable indoor flight and GPS-denied localization in constrained multi-floor environments with a similarly limited suite of purely onboard sensors. 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 UAV operation with loop closure. It also addresses specially designed controllers to help compensate for sudden changes in wind velocity and air flow as the UAV traverses constrained low-clearance areas with potentially strong aerodynamic disturbances.

In [?] Weiss and Siegwart went on to tackle the problem of metric scale in monocular VIN systems. The researchers developed a general algorithm that provides metric scale to monocular visual odometry and monocular SLAM systems using inertial measurement unit (IMU) data. The authors accomplished the development of the metric scale by the addition of an inertial sensor with a three-axis accelerometer and gyroscope. Weiss and Siegwart created a modular solution that is based on an 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.

In 2012, Weiss et al. built upon this metric scale algorithm to present a versatile sensor fusion framework for autonomous flight [?]. Due to latency, noise, and arbitrary scaling within the output of a UAV's sensors, it is both impractical and ill-advised to incorporate this sensor output for position control without calibration or post-processing. They addressed these problems using an EKF-SLAM formulation which fuses pose measurements with inertial sensors. The researchers not only estimate pose and velocity of the UAV, but also estimate sensor biases, scale position measurements, and perform inter-sensor self-calibration in real time. Their research demonstrates that the proposed framework is capable of running entirely onboard a UAV, performing state prediction at a rate of 1 kHz. Their results illustrate that this approach is able to handle measurement delays (up to 500 ms), sensor noise (with positional standard deviation up to 20 cm), and slow update rates (as low as 1 Hz) while dynamic maneuvers are still possible. Weiss et al. present a detailed quantitative performance evaluation of the system under the influence of different disturbance parameters and different sensor setups to highlight the versatility of their approach. That same year, Weiss et al. further developed their system, adding a speed-estimation module to the framework which fuses IMU and vision data to turn the monocular camera into a metric body-speed sensor [?]. They then demonstrated how this module could be used for self-calibration of the UAV's onboard sensor suite in real time.

Shortly thereafter, Huang et al. [?] presented solutions to two UKF limitations that exist in current state-of-the-art SLAM systems. Specifically, the researchers addressed 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, they introduced a novel sampling strategy which produces a constant computational cost. This sampling method, while linear in the prediction phase, is quadratic in the update phase. Although this new sampling strategy was primarily proposed for resolving the aforementioned SLAM problem, the researchers stressed that it has potential usefulness in other nonlinear estimation applications. To

address the problem of inconsistency in state estimations, Huang et al. proposed 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, the researchers validated their results with both real-world and simulation experiments.

In 2013, Simon Lynen et al. [?] presented a generic framework based on the EKF-SLAM system developed in [?] to be more robust to sensor blackouts and to be self-correcting in scale. The researchers demonstrated that their Multi-Sensor-Fusion EKF (MSF-EKF) framework was capable of processing measurements from an unlimited number of sensors, as well as sensor types, while simultaneously performing online self-calibrations of the overall sensor suite. It was the design of this software framework, which the researchers open-sourced shortly after publication, that inspired many of the design decisions behind the `kalman_sense` package.

[Engel 2013] Engel et al. proposed a novel direct monocular SLAM algorithm unlike that of the existing direct approaches which embrace pure visual odometry. The novelty of the authors' 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.

[Faessler 2015] Faessler et al. reported on the development and demonstration of a low-cost, low-weight, vision-based quadrotor UAV 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 moved the UAV 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 researchers believed that such a vehicle potentially would be a great enhancement in search-and-rescue missions, disaster response, and remote inspection of terrain.

Chapter 2

Algorithm Design and

Implementation

2.1 UKF Formulation

We begin by defining the following vector quantities:

$$tk : p, q, v, \omega, a \tag{2.1}$$

tk: define n and m

We then define the state vector $\mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{x} \triangleq \left\{ \mathbf{p}, \mathbf{q}, \mathbf{v}, \omega, \mathbf{a}, \right\}^T. \tag{2.2}$$

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 $\mathbf{P}$:

$$
\begin{aligned}
\chi^0 &= \mathbf{x} \\
\chi^i &= \mathbf{x} + \left(\sqrt{(n + \lambda)\mathbf{P}} \right)_i, \quad i = 1, \dots, n \\
\chi^i &= \mathbf{x} - \left(\sqrt{(n + \lambda)\mathbf{P}} \right)_{i-n}, \quad i = n+1, \dots, 2n
\end{aligned}
\tag{2.3}
$$

where λ is defined in terms of the following constants:

$$tk : \alpha, \beta, \kappa, \lambda \tag{2.4}$$

tk: TO BE CONT'D.

2.2 Software Design Considerations

Much of the impetus for creating the `kalman_sense` package came from a desire to create a generic UKF framework for estimating the state of arbitrary systems using any number of relative and absolute sensors. To achieve this, the `kalman_sense` package is organized in an object-oriented manner around an overarching abstract class called `UnscentedKf`. This abstract class contains a number of methods performing the different mathematical operations defined in Section **??**. These methods have been written in a generic manner in order to enable easy extension of `UnscentedKf` by other subclasses containing concrete implementations of various systems. Currently, the package contains exactly one such subclass, known as `QuadUkf`. This subclass contains methods and data structures related directly to estimating the state of a quadcopter or other rotorcraft UAV.

This object-oriented architecture is meant to allow for a certain degree of system agnosticism. The abstract `UnscentedKf` class encapsulates the generic mathematics of the Unscented Kalman Filter without knowledge of particular system constraints. This class does little other than matrix mathematics and is designed to take as input the number of a system's states n and its number of sensors m. With this knowledge, `UnscentedKf` is able to populate a set of mean and covariance weights and intelligently perform all of the requisite linear algebra for the UKF formulation. All other knowledge of particular states, sensors, vehicle geometry, and other particulars is hidden within subclasses such as `QuadUkf`.

`UnscentedKf` behaves in a manner similar to a Java interface in that it requires the extending class to supply functions codifying a process model and an observation model for the system under scrutiny. These two functions, along with n and m, form the entirety of what `UnscentedKf` "knows about the vehicle." All other details, including the fact that the class is being used in a ROS environment, are hidden from `UnscentedKf`. It is worth

noting that UnscentedKf's only dependency is on the Eigen[1] C++ linear algebra library.

The subclass (QuadUkf for the remainder of this thesis) handles all of the ROS communications for the given system. Specifically, this class has callback functions for receiving sensor data and is responsible for publishing state and covariance estimates. The kalman_sense main method handles setup and teardown of the necessary ROS publisher and subscriber nodes.

---

[1] www.eigen.tuxfamily.org

Chapter 3

Experimental Design

3.1 Testing Considerations

Before venturing further, we should summarize the goals of the UKF framework described previously with particular attention to unmanned aircraft system (UAS) operations. This ROS package was designed with the express intent of producing estimates of the position vector **p** and orientation quaternion **q** of a rotorcraft UAV in real time. Thus, the experiments testing kalman_sense's efficacy compare the filter's estimates of position and orientation to the "ground truth" as measured by a Vicon[1] motion capture system.

3.2 Materials

3.2.1 Computation and Sensing

1. One (1) MatrixVision mvBlueFOX-MLC Camera[2]

2. One (1) 1044_0 PhidgetSpatial Precision 3/3/3 High Resolution IMU[3]

3. One (1) Hewlett-Packard Spectre x360 Convertible Laptop 13-ac076nr[4]

---

[1] https://www.vicon.com

[2] https://www.matrix-vision.com/USB2.0-single-board-camera-mvbluefox-mlc.html

[3] http://www.phidgets.com/products.php?product_id=1044

[4] http://store.hp.com/us/en/pdp/hp-spectre-x360---13-ac076nr

4. Two (2) male Mini USB 2.0 to male USB Type A cables

3.2.2 Mobile Test Stand

1. One (1) Oklahoma Sound PRC200 Premium Presentation Cart[5]

2. Two (2) 4-inch C-clamps

3. One (1) 1.2-meter 80/20 1515 rail[6]

4. One (1) 15 Series "L" Handle Linear Bearing Brake Kit[7]

5. One (1) tk: other rail screw

6. Two (2) tk: the part that goes inside the rail for both the brake and the screw

7. Three (3) 1-inch Vicon infrared retroreflector balls

8. Two (2) 0.5-inch Vicon infrared retroreflector balls

9. One (1) 0.7-meter length of $\frac{1}{8}$-inch-thick carbon fiber rod

3.3 The Experiments

The following three experiments were devised to characterize the filter's behavior:

1. An x-directional "long walk,"

2. A y-directional "long walk," and

3. A square pattern translation.

In each experiment, the

---

[5]http://www.oklahomasound.com/products/product-category/single/?prod=9
[6]https://8020.net/1515.html
[7]https://8020.net/6800.html

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.

[3] S. J. Julier, "The Scaled Unscented Transformation," in *Proceedings of the 2002 American Control Conference (IEEE Cat. No.CH37301)*, vol. 6, pp. 4555–4559 vol.6, May 2002.

[4] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," in *Invited Paper Proceedings of the IEEE*, vol. 92, pp. 401–422, March 2004.

[5] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," in *Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)*, (Nara, Japan), November 2007.

[6] S. Shen, N. Michael, and V. Kumar, "Autonomous Multi-Floor Navigation with a Computationally Constrained MAV," in *IEEE International Conference on Robotics and Automation*, pp. 20–25, 2011.