UAV Autonomous Motion Estimation Methodologies

Anand Abhishek and K. S. Venkatesh
Department of Electrical Engineering, IIT Kanpur, Kanpur, U.P, India

Keywords: Implicit Extended Kalman Filter, UAV, Autonomous Navigation, Epipolar Constraint, IMU, Planar Constraint.

Abstract: Unmanned aerial vehicle(UAV) are widely used for commercial and military purposes. Various computer vision based methodologies are used for aid in autonomous navigation. We have presented an implicit extended square root Kalman filter based approach to estimate the states of an UAV using only onboard camera which can be either used alone or assimilated with the IMU output to enable reliable, accurate and robust navigation. Onboard camera present information rich sensor alternative for obtaining useful information form the craft, with the added benefits of being light weight, small and no extra payload. The craft system model is based on differential epipolar constraint with planar constraint assuming the scene is slowly moving. The optimal state is then estimated using current measurement and defined the system model. Pitch and roll is also estimated from above formulations. The algorithms results are compared with real time data collected from the IMU.

1 INTRODUCTION

Autonomy is commonly defined as the ability to make decisions without human intervention. Currently a large amount of research is devoted for developing reliable and robust vision based formulations to aid UAVs in autonomous navigation. Pivotal for the autonomy of any UAV is the ability to obtain information of the surrounding environment to estimate the self-motion, computer vision algorithms provides a good alternative.

The IMU outputs of velocity, position, and attitude however, are not exact for a number of reasons (Siouris, 1993). The main reasons are sensor inaccuracies, gravity modeling and external disturbances. Because of their wide application a large amount of research has been carried out for integration computer vision with IMU sensors data. By state of UAV (χ) we mean

$$\chi = [V_x, V_y, V_z, \omega_x, \omega_y, \omega_z, \theta, \Psi, \Phi] \quad (1)$$

where V_x, V_y, V_z are linear velocities along X, Y, Z direction; ω_x, ω_y, ω_z are angular velocities about X, Y, Z axis; θ, Ψ, Φ are Pitch, Roll and Yaw as shown in Figure 1. Commonly aircraft motion estimation are based on feature points in successive frames. Optical flow with Horizon constraint gives the χ. (Adiv, 1985) has reconstructed the angular velocity and linear velocity of moving objects using optical flow. But their formulation has inherent ambiguity in the noisy environments and is based on linear optimization. (Jepson and Heegar, 1991) first calculated the translation parameter, then using this and other information estimated the rotation parameter. However there method is liable to be unreliable if the surface is not smooth. (Sazdovski et al., 2010) estimated the aircraft attitude using a single on board camera. The proposed approach assumes that position of aircraft and feature points in the camera images are known.

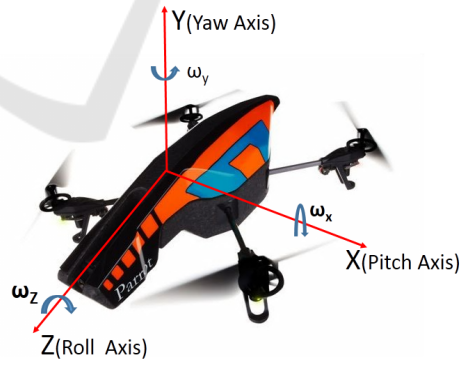


Figure 1: Axis Convention.

(Soatto et al., 1996) derived the implicit extended Kalman filter *(IEKF)* for estimating displacement and rotation that incorporates an implicit formulation into the framework of the IEKF on the random walk model. The IEKF implementation was applied on the non linear space to characterize the motion of a cloud of feature points about a fixed camera. (Grabe et al., 2012) had use *continuous homography con-*

Abhishek A. and Venkatesh K..
UAV Autonomous Motion Estimation Methodologies.
DOI: 10.5220/0005302805430550
In *Proceedings of the 10th International Conference on Computer Vision Theory and Applications* (VISAPP-2015), pages 543-550
ISBN: 978-989-758-091-8

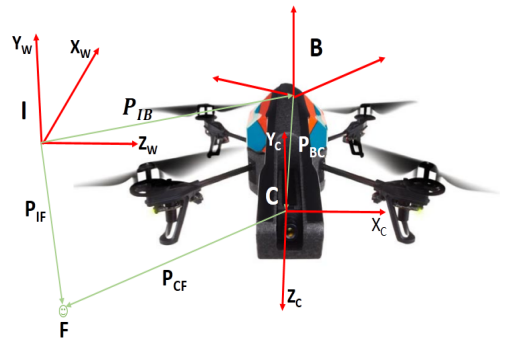


Figure 2: Relationship among Camera reference frame, Body reference frame and Inertial frame.

straint (Ma et al., 2004) for linear velocity estimation, integrated the result with IMU for partial state estimation using computer vision. (Dusha et al., 2007) proposed algorithm based on Kalman filter and optical for state estimation. Detected lines optical flow are tested against the motion filter. Then Kalman filter is used to remove the outlier. Their algorithm is not suitable for real world scenario. (Oreifej et al., 2011) have used horizon constraint to estimate the pitch and roll. First they had detected horizon from the image, slope and height of the detected horizon from reference point was used for pitch and roll estimation. In our formulation pitch and roll estimation does not require extra computations.

In Section 2 we discuss the theoretical background of this work. Section 3 describes how the continuous homography constraint is harnessed to find system model used in Kalman filter. Section 4 discusses the simulation results and practical considerations and lastly concludes with Conclusion.

2 THEORETICAL BACKGROUND

2.1 System Model

Position and velocity in $3D$ space has meaning only if described with respect to a frame of reference (coordinate system). Depending on the requirement different coordinate systems are used in different context. Here, we will be dealing with camera reference frame, body reference frame and inertial frame as shown in Figure 2. The orientation of a general aircraft, with a front camera is shown in this figure. The aircraft body frame $\boldsymbol{B}$, is fixed at the aircraft center of gravity and is located with respect to the any inertial frame $\boldsymbol{I}$, by the vector P_{IB}. The camera reference frame $\boldsymbol{C}$, is fixed to the aircraft and is located with respect to $\boldsymbol{B}$ by the vector P_{BC}. The tracked features point of the environment is denoted by $\boldsymbol{F}$, which is located relative to $\boldsymbol{I}$ by P_{IF}. The vector P_{CF} relates the relative position of $\boldsymbol{F}$ with respect to $\boldsymbol{C}$. These coordinate systems have different spatial configurations, so suitable transformation of coordinate frame is required for mapping a vector represented in one reference frame to another reference frame. $\Re_{BC}$ is the Euler angles transformation relating $\boldsymbol{B}$ and $\boldsymbol{C}$ and $\Re_{IB}$ is the Euler angles transformation relating $\boldsymbol{I}$ and $\boldsymbol{B}$. $\Re_{IB}$ and $\Re_{BC} \in SO(3)$. The mathematical relationship among them is given in Equation 2

$$P_{CF} = \Re_{BC} * \Re_{IB}(P_{IB} - P_{IF}) - \Re_{BC} * P_{BC} \quad (2)$$

We also assume that ground is planar. This assumption is perfectly valid when UAV is at certain height above the ground or ground surface has minor ups and downs. The Z_w and X_w axis of the world frame lie on the ground plane. Negative Y_w axis of world frame points downward to the centre of the earth and perpendicular to the surface of the earth. The Z_c axis of the camera is along the optical axis of the camera such that Z_c and Z_w lie in the same plane. The X_c and Y_c axis of the camera are parallel to the bottom edge and the left edge of the image plane respectively, with the origin at the optical centre as shown in Figure 2. Any point $P_w \in \Re^3$ in the world frame will have the coordinate $P_c \in \Re^3$ in the camera frame given by

$$P_c = R_w^c P_w \quad (3)$$

where R_w^c is the direction cosine matrix, representing the rotation from world frame to camera frame. Let the rigid object in the scene undergo translation $T = [T_x, T_y, T_z]$ and small rotation $\Omega = [\Omega_x, \Omega_y, \Omega_z]$, relative to the camera. Let the object position relative to the camera be $\mathcal{X}_{t1} = [X, Y, Z]$ at time $t1$ and $\mathcal{X}_{t2} = [X_2, Y_2, Z_2]$ at time $t2$, such that $t2 > t1$ then

$$\mathcal{X}_{t2} = R\mathcal{X}_{t1} + T \quad (4a)$$

$$R = \begin{pmatrix} 1 & -\Omega_z & \Omega_y \\ \Omega_z & 1 & -\Omega_x \\ -\Omega_y & \Omega_x & 1 \end{pmatrix} \quad (4b)$$

If $V = [V_x, V_y, V_z]$ is linear velocity and $\omega = [\omega_x, \omega_y, \omega_z]$ is angular velocity, then differentiating above equation gives the total velocity $\dot{\mathcal{X}}_{t1}$.

$$\dot{\mathcal{X}}_{t1} = \hat{\omega}\mathcal{X}_{t1} + V \quad (5)$$

where $\hat{\omega} \in so(3)$ is skew symmetric matrix of ω.

2.2 Continuous Homography Constraint

Epipolar constraint states that feature point $\boldsymbol{x_1} \in \Re^{3\times1}$ in a frame, its corresponding location $\boldsymbol{x_2}$ in the consecutive frame satisfies following constraint (Zisserman

and Hartley, 2004)

$$\boldsymbol{x_2}^T \mathbb{E} \boldsymbol{x_1} = 0 \tag{6}$$

where $\mathbb{E} = \hat{T}R \in \Re^{3\times3}$ is known as essential matrix, $R \in \Re^{3\times3}$ is rotation, $\hat{T} \in \Re^{3\times3}$ is skew symmetric matric of T. Assuming pinhole camera model, feature point $\boldsymbol{X} = [X,Y,Z]$ and its corresponding image coordinate $\boldsymbol{x} = [x, y, 1]$ satisfy

$$K\boldsymbol{X} = \lambda \boldsymbol{x}$$

where λ is unknown scale factor, K is intrinsic camera calibration matrix. Let $N \in \Re^{3\times1}$ be surface normal and d be the shortest distance of the plane from the camera reference frame then any point $\boldsymbol{X}$ lying on the plane satisfies

$$(N/d)^T \boldsymbol{X} = 1 \tag{7}$$

Using above equations we have (Ma et al., 2004)

$$\hat{\boldsymbol{x}}\boldsymbol{u} = \hat{\boldsymbol{x}} H \boldsymbol{x} \tag{8}$$

where $\boldsymbol{u} = [u,v,0]^T$ is optical flow, and $\mathbb{H}$ is the *Continuous homography matrix* given as

$$H = K\mathbb{H}K^{-1} = K(\hat{\omega} + V(N/d)^T)K^{-1}$$

$$H = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \tag{9}$$

3 STATE ESTIMATION

The Kalman filter is a set of mathematical equations that provides an efficient recursive means to estimate the state of a linear dynamic system discretized in the time domain on the basis a of noisy measurement. It models the state of a system at a time t_k, as a function of the prior state at time t_{k-1}, according to the equation

$$x_k = F_k x_{k-1} + B_k u_k + w_k \tag{10a}$$

$$z_k = H_k x_k + v_k \tag{10b}$$

where x_k is the state vector containing, w_k is state noise, F_k is the state transition matrix, z_k is the measurement vector, H_k is the transformation matrix, u_k is the control input, B_k is the control input matrix and v_k is measurement noise. We assume the following:

$$E\{w_k\} = E\{v_k\} = 0$$

$$E\{w_k w_k^T\} = R_w, E\{v_k v_k^T\} = R_v$$

Standard **Time Update** equations at time $t = t_k$ are given below (Kalman, 1960):

$$x_k^f = F_k x_{k-1}^a + B_k u_k \tag{11a}$$

$$P_k^f = cov(x_k - x_k^f) = F_k P_{k-1} F_k^T + R_w \tag{11b}$$

x_k^f is the predicted state at time t_k, P_k^f is the predicted state covariance at time t_k, x_{k-1}^a is the estimated value of x_{k-1} and R_w is process noise covariance. Also standard **Measurement Update** equations at time $t = t_k$ are

$$x_k^a = x_k^f + K_k(z_k - H_k x_k^f) \tag{12a}$$

$$P_k = cov(x_k - x_k^a) = P_k^f - K_k H_k P_k^f \tag{12b}$$

x_k^a is the estimated state of x_k and P_k is the Posteriori State Covariance, where innovation covariance Q and Kalman gain K_k are given by

$$Q = cov(z_k - H_k x_k^f) = H_k P_k^f H_k^T + R_v$$

$$K_k = P_k^f H_k^T Q^{-1}$$

We now derived a system model of UAV based on continuous homography constraint. Putting the values of all the parameters in Equation 8, we get

$$u = xh_1 + yh_2 + h_3 - x^2h_7 - xyh_8 - xh_9 \tag{13a}$$

$$v = xh_4 + yh_5 + h_6 - xyh_7 - y^2h_8 - yh_9 \tag{13b}$$

$$-yu + xv = -xyh_1 + x^2h_4 - y^2h_2 + xyh_5 - yh_3 + xh_6 \tag{13c}$$

Multiplying Equation 13a by $-y$ and Equation 13b by x on both sides yields Equation 13c. Therefore, it is a dependent set of equations, and only top two of them are used for estimating H. Above system model is linear w.r.t $[h_1,h_2,...h_9]$ but they are nonlinear on *Continuous homography space* $\mathcal{H}$. $\mathcal{H}$ belongs to a subset of $\Re^{3\times3}$ whose eigen values possess unique characteristics. Taken any random value of ω, N and V as given below

$$\omega = \begin{pmatrix} 41 & 13 & 47 \end{pmatrix}$$

$$N = \begin{pmatrix} 31 & 24 & 18 \end{pmatrix}$$

$$V = \begin{pmatrix} 18 & 10 & 13 \end{pmatrix}$$

Without the loss of generality assuming $d = 1$ then

$$\mathbb{H} = \begin{pmatrix} 558 & 385 & 337 \\ 357 & 240 & 139 \\ 390 & 353 & 234 \end{pmatrix}$$

The eigen values λ of $\mathbb{H} + \mathbb{H}^T$ are $(2082.5, 0, -18.51)$. $\lambda_1 > 0$, $\lambda_2 = 0$ and $\lambda_3 < 0$. This is true for any $\mathbb{H}$ which belongs to the continuous homography group $\mathcal{H}$. So after solving Equations 13a and 13b for $[h_1,h_2,...h_9]$, it will be projected to space $\mathcal{H}$. Let h_k be the state vector and z_k the measurement vector at time $t = t_k$

$$h_k = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 & h_5 & h_6 & h_7 & h_8 & h_9 \end{pmatrix}^T$$

$$z_k = \begin{pmatrix} x & y & u & v \end{pmatrix}^T$$

Rewriting Equations 13a and 13b in implicit form gives

$$\underbrace{\begin{pmatrix} x & y & 1 & 0 & 0 & 0 & -x^2 & -xy & -x \\ 0 & 0 & 0 & x & y & 1 & -xy & -y^2 & -y \end{pmatrix}}_{\zeta_k} h_k + \underbrace{\begin{pmatrix} -u \\ -v \end{pmatrix}}_{\mathcal{U}_k} = 0 \quad (14)$$

Denote above equation by $g(h_k, z_k)$. Assuming h_k and z_k are slowly varying, they can be modeled as

$$h_k = f(h_{k-1}) + w_k; w_k \in \mathcal{N}(0, R_w) \quad (15)$$

$$z_k = z_k^t + v_k; v_k \in \mathcal{N}(0, R_v) \quad (16)$$

where $f(\,)$ is nonlinear vector functions and z_k^t is true value of z_k. Linearising Equation 15 at time t_k using previous estimate h_{k-1}^a, we get

$$h_k \approx f(h_{k-1}^a) + J_f(h_{k-1}^a)(h_{k-1} - h_{k-1}^a) + w_k + \text{H.O.T}$$

where J_f is Jacobian of function f w.r.t h_k. We assume that all h_k is very close to h_{k-1}^a, so higher order terms can be ignored. Rewriting this in form as given in equation 10

$$h_k \approx F_k h_{k-1} + u_k + w_k \quad (17)$$

where

$$F_k = J_f(h_{k-1}^a)$$
$$u_k = f(h_{k-1}^a) - J_f(h_{k-1}^a) h_{k-1}^a$$

Therefore predicated value of h_k, h_k^f is given by

$$\begin{aligned} h_k^f &= F_k h_{k-1}^a + u_k \\ &= J_f(h_{k-1}^a) h_{k-1}^a + f(h_{k-1}^a) - J_f(h_{k-1}^a) h_{k-1}^a \\ &= f(h_{k-1}^a) \end{aligned} \quad (18)$$

Similarly *Predicted state covariance* P_k^f is

$$P_k^f = J_f(h_{k-1}^a) P_{k-1} J_f(h_{k-1}^a)^T + R_w \quad (19)$$

Linearising the $g(h_k, z_k)$ about (h_k^f, z_k^t) results,

$$g(h_k, z_k) \approx g(h_k^f, z_k^t) + J_h(h_k^f)(h_k - h_k^f) + J_z(z_k^t)(z_k - z_k^t)$$

where J_h is Jacobian of $g(h_k, z_k)$ w.r.t h_k and J_z is Jacobian of $g(h_k, z_k)$ w.r.t z_k. But, $g(h_k, z_k) = 0$ and $z_k - z_k^t = v_k$. Rearranging the equation in *Standard form*

$$-g(h_k^f, z_k^t) + J_h(h_k^f) h_k^f = J_h(h_k^f) h_k + J_z(z_k^t) v_k$$
$$\Rightarrow \tilde{g}(h_k, z_k) = H_k h_k + \tilde{v_k}$$

Comparing this with equation 10, measurement update equations are

$$\begin{aligned} h_k^a &= h_k^f + K_k(\tilde{g}(h_k^f, z_k) - H_k h_k^f) \\ &= h_k^f - K_k(\zeta_k h_k + \mathcal{U}_k) \end{aligned} \quad (20)$$

$$P_k = P_k^f - K_k \zeta_k P_k^f \quad (21)$$

Algorithm 1: State estimation using Kalman Filter.

Given set of n feature points and optical flow, find (ω, V), pitch and roll

1: **Initialisation of State Variable**
Form the equations as mentioned in the Equations 13a and 13b. Solve for h and stack it to form H. Normalize H to find $\mathbb{H} = H - \frac{e_2}{2} I$. Estimate the initial error covariance P_0 and h_0^f.

2: **Time update**
Predict the current state of system on the basis of previous estimate.
$h_k^f = h_{k-1}^a$
$P_k^f = P_{k-1} + R_w$

3: **Measurement Update**
Estimate the best current state using time update results.
$h_k^a = h_k^f$ - $K_k(\zeta h + \mathcal{U})$
$P_k = P_k^f - K_k J_h(h_k^f) P_k^f$
Above equation will be solved as follows

- Compute square root of P_k^f, $\acute{R}_v$ using LDL^T.
- Form the block matrix, O is zero matrix.

$$pre = \begin{pmatrix} \acute{R}_v^{T/2} & O \\ P_k^{fT/2} \zeta^T & P_k^{fT/2} \end{pmatrix}$$

- QR factorization of above will result in

$$post = \begin{pmatrix} Q^{T/2} & Q^{-1/2} \zeta P_k^f \\ O^T & P_k^{T/2} \end{pmatrix}$$

- $P_k = post(end - k : end, end - k : end)$ where end is number of rows and k is number of state vector.
- Similarly solve for other equations..

4: **Projection on Continuous homography space**
Stack h_k to get H and normalize it to obtain $\mathbb{H}$.

5: **Recover State**
Decompose $\mathbb{H}$ to estimate V, N and ω (Ma et al., 2004).
Using N estimate pitch and roll.

where

$$K_k = P_k^f \zeta_k^T Q^{-1}$$
$$Q = \zeta_k P_k^f \zeta_k^T + \acute{R}_v$$
$$\acute{R}_v = J_z(z_k^t) R_v J_z^T(z_k^t)$$

where T denotes the transpose of matrix. h_k found in above equations generally will not belong to $\mathcal{H}$. Hence h_k will be projected on to $\mathcal{H}$. For this stack the $[h_1, h_2, ...h_9]$ to form $H \in \Re^{3\times3}$. The continuous homography matrix $\mathbb{H}$ is given by

$$\mathbb{H} = H - \frac{e_2}{2} I \quad (22)$$

where I is identity matrix and e_2 is second largest eigenvalue of $H+H^T$. Once $\mathbb{H}$ is known, (ω, V) is estimated using steps mentioned in Algorithm 1. Now we will show procedure for pitch and roll estimation from N. Given axis convention in Figure 2, unit normal to the ground is $\acute{N} = [0,1,0]$. Due to different orientation of coordinate axes $\acute{N}$ will appear as N from the reference frame B, relation among them is given by Equation3, R_w^c is given by

$$R_w^c = \begin{pmatrix} \cos\Phi & 0 & \sin\Phi \\ 0 & 1 & 0 \\ -\sin\Phi & 0 & \cos\Phi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\Psi & -\sin\Psi & 0 \\ \sin\Psi & \cos\Psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

N is independent of yaw and translation, hence solving for N yields

$$N = [-\sin\Psi, \cos\Psi\cos\theta, \cos\Psi\sin\theta]^T \quad (23)$$

Pitch θ and roll Ψ are calculated from N using relations given below

$$\Psi = \frac{\arccos\sqrt{N_2^2+N_3^2} + \arcsin N_1}{2} \quad (24)$$

$$\theta = \frac{\arccos\frac{N_2}{\cos\Psi} + \arcsin\frac{N_3}{\cos\Psi}}{2} \quad (25)$$

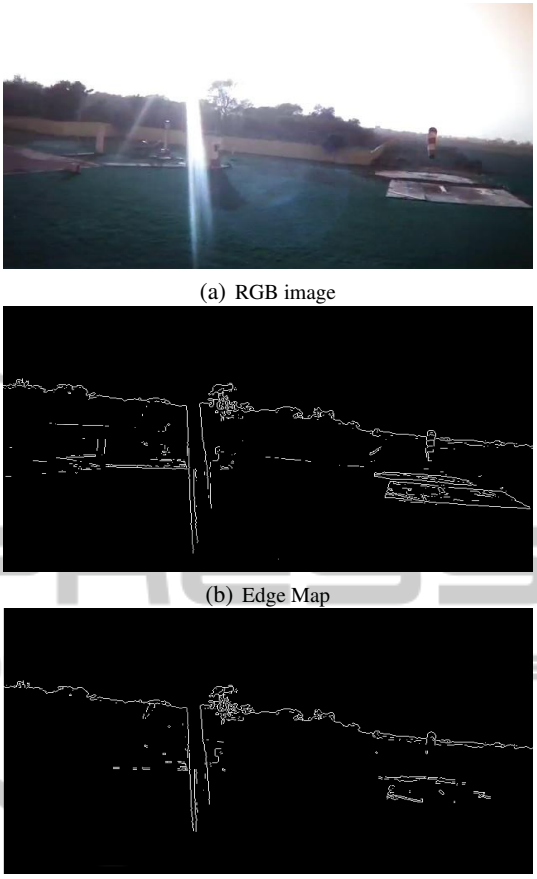
Above method for calculating pitch and roll depends only on feature points implying these can be estimated in wide range of environment opposed horizon constraint based methods.

Theocratically, P_k is symmetric positive definite, but numerical implementation sometimes led to P_k that became nonsymmetric and/or indefinite. The situation becomes worse if the matrix dimensions are large. Thus, to improve the condition number, and hence the numerical stability to the noise, square root filter is used. Square root filtering implementation provides twice the precision than standard kalman filter implementation. Cholesky factorization is used to find the square root of positive definite matrix. In practice, however, the estimation problems are severe as some outliers can seriously pollute the sample, thus making the matrix non positive definite. To avoid this issue, LDL^T factorization is used. The LDL^T factorization uniquely factors the symmetric matrix P as

$$P = LDL^T$$

where L is a lower triangular matrix with 1 as diagonal element, D is a diagonal matrix. It is more efficient than Cholesky factorization because it avoids computing the square roots of the diagonal elements and is always stable. Due to noise some insignificant eigenvalues are neglected. From LDL^T, the Cholesky factorization will be computed as follows

$$P = LDL^T = (LD^{1/2})(D^{1/2}L) \quad (26)$$



(a) RGB image

(b) Edge Map

(c) Edge map using Equation 27

Figure 3: Edge map comparison using different color to grayscale conversion relations.

4 SIMULATION RESULTS

Multiple experiments were conducted on outdoor environment using **Parrot AR Drone** shown in Figure 1, for the verification of algorithms proposed in this discussion. The algorithms were tested on both **OpenCv** and **Matlab** platforms. We have used **Parrot AR Drone** SDK for capturing video(30 frame per second) and IMU data.

For converting RGB to grayscale we have used Equation 27. We found that feature tracking gives better results when following relation is used for converting *RGB* image to grayscale image.

$$GS = min(min(255-R, 255-G), 255-B) \quad (27)$$

where R, G and B are red, green and blue color channel of an *RGB* image, and *GS* is grayscale image. This conversion removes the noise and makes feature detection more robust than other grayscale conversion methods. Figure 3(a) shows a frame captured by the

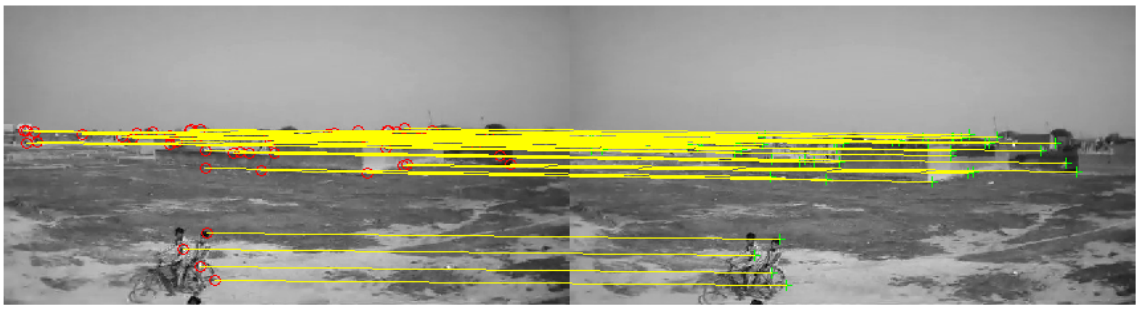
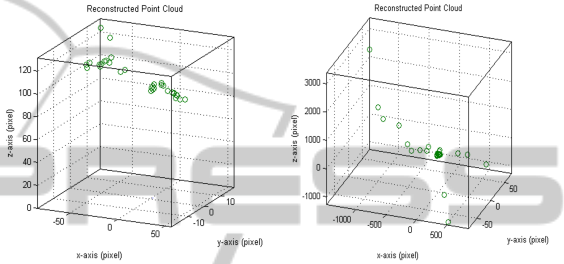


Figure 4: Feature points extraction and matching.

UAV. Due to the large height of UAV from the ground, the image transmitted over Wifi to the ground station, is not of a good quality and has poor contrast. Edge map of the corresponding image using Matlab inbuilt RGB to grayscale conversion is shown in Figure 3(b). The edge map has a large number of false and stray edges due to the poor quality of image and the presence of noise. The strong edges, caused due to sunlight are also undesired. In Figure 3(c), a large number of false edges are removed using relation mentioned in Equation 27. Feature points are the points of interest in an image. Most commonly used features are the edges and corners. We have used Harris corner descriptor for feature extraction and tracking. The Harris corner detector is a commonly used interest point detector, owing to its strong invariance to scale, rotation, image noise and illumination variation(Schmid et al., 2000). Correspondence based technique is used for feature tracking between frames. Correspondence based techniques extract a set of features from each frame, and then attempt to establish correspondences between consecutive sets of features. To deal with large motion hierarchical search strategy is used. The patches in consecutive frames are compared using a translational model, and then these location estimates are used to initialize an affine registration between the patch in the current frame and the base frame, where a feature was first detected(Shi and Tomasi, 1994). Also number of features point used for correspondence between consecutive images depends on type of environment.

4.1 ω and V Estimation

Various methodologies for state estimation were adopted in experiments, to find harmony between the theoretical concepts and the actual output of the sensor. States are estimated solving *Continuous homography constraint* for each frame and approach discussed above. Most of above mentioned algorithms give multiple solutions for linear and angular velocity. This inherent ambiguity is removed using positive



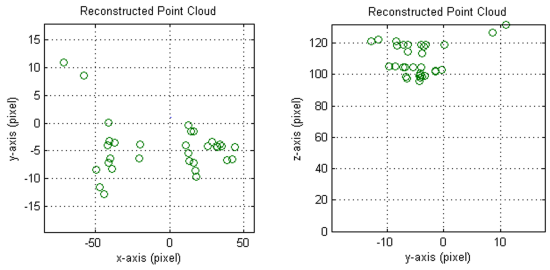
(a) Configuration 1 (b) Configuration 2

Figure 5: 3*D* point reconstruction.

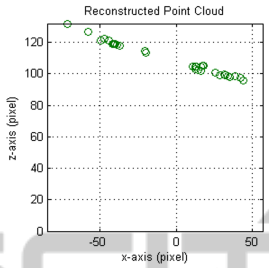
depth constraint. Here, we first show the results of 3*D* reconstruction, and then state estimation, followed by a comparison among them.

We have reconstructed the 3D point from the camera matrix obtained in our experiment. Figure 5, shows the two possible orientation of the 3*D* points with respect to camera, reconstructed from feature points extracted from frame taken by our UAV as shown in Figure 4. Other two possible 3*D* reconstructed points are mirror image of Figure 5(a) and 5(b) respectively. It is difficult to visualise the above point coordinates in 3*D*, thus the projection of Configuration 1 on X, Y, Z plane is shown in Figure 6. These points have been scaled with a scale factor of 100 (randomly selected), as the reconstructed points are obtained upto a scale factor. These reconstructed points have both positive and negative X and Y coordinates but always have positive depth i.e all have a positive Z coordinate, i.e the assumption we have taken about the body coordinate system(Refer to Figure 2). The other configuration is just a mirror image. Similarly, the projection of Configuration 2 on X, Y, Z plane is shown in Figure 7. Reconstructed points are spread along both positive and negative axis. The other configuration will the mirror image. Among these only Configuration 1 is valid justifying our assumption. This result also supports validity of computer vision approach for state estimation.

The angular velocity estimated from various algorithms is shown in Figure 8, where **IMU** stands

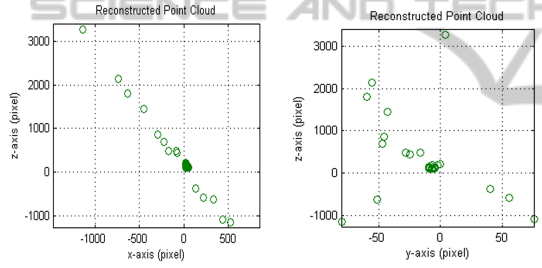


(a) Projection on $X-Y$ (b) Projection on $Y-Z$

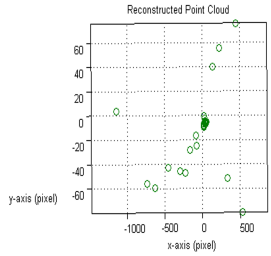


(c) Projection on $Z-X$

Figure 6: Configuration 1.



(a) Projection on $X-Z$ (b) Projection on $Y-Z$

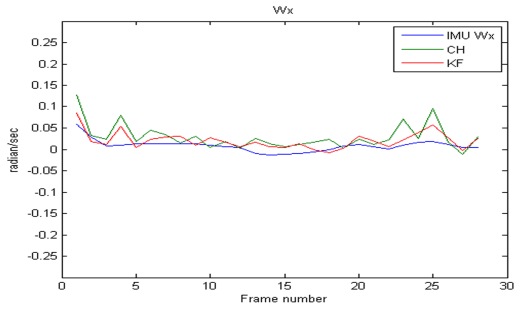


(c) Projection on $X-Y$

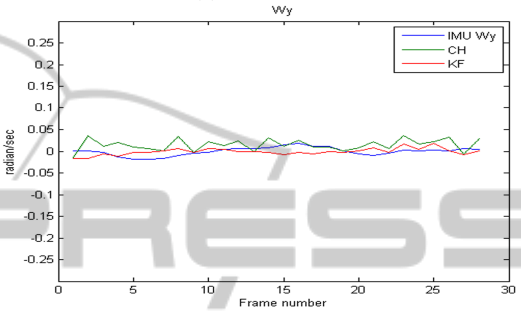
Figure 7: Configuration 2.

for measurement obtained form Inertial measurement unit, **CH** stands for *Continuous homography constraint* and **KF** stands for Implicit square root kalman filter. Angular velocity i.e ω_x, ω_y, ω_z estimated using these method approximately follows the IMU.

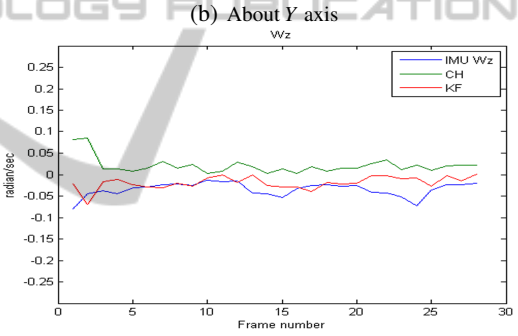
The linear velocity is also estimated using all the above mentioned methods. But, in the absence of any $3D$ information, it can be estimated only upto a scale factor. Implicit square root kalman filter based algorithm gives the better result than continuous homography based method.



(a) About X axis



(b) About Y axis

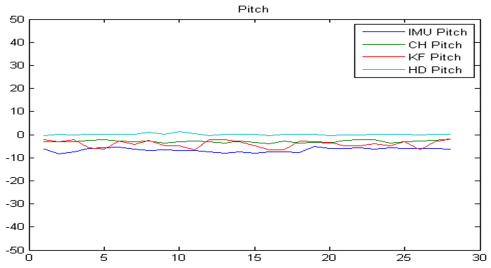


(c) About Z axis

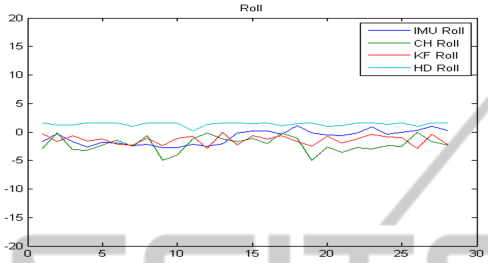
Figure 8: ω estimation.

4.2 Pitch and Roll Estimation

Pitch and roll estimated using horizon constraint(**HC**) is shown in the Figure 9. Pitch estimation is more prone to error than roll estimation using horizon constraint. Because pitch estimation depends on the height while roll depends on the slope of detected horizon line. Pitch estimation is only accurate if the constant height assumption is valid. If horizon is not present, then pitch and roll cannot be found. There are no such limitations in continuous homography constraint based algorithm (shown as **CH** in the plot). Figure 9 shows the comparison among them with ground truth. In our experiments Implicit square root kalman filter based algorithm proved to be more promising than horizon constraint based approach, because it solely depends upon the feature points. Proposed method can be used in any environ-



(a) Pitch



(b) Roll

Figure 9: Pitch and Roll Estimation.

ment provided sufficient number of features points are available.

5 CONCLUSIONS

Computer vision based approaches are attractive due to low weight, low cost and existing presence of a camera on the vehicle. This work gives an insight into the application of onboard cameras for state estimation, without using any additional sensors. Linear and angular velocity estimation algorithm has been developed using implicit extended Kalman filter. We have removed the horizon constraint in the estimation of pitch and roll. They can be estimated more accurately without any horizon constraint. Also feature points correspondence are required for calculating sparse optical flow, so pitch and roll estimation do not require extra processing as opposed to horizon constraint where horizon line segment is required for finding these parameters. The results clearly manifest the feasibility of algorithms for real time applications.

Although the results are inspiring, yet these approaches are limited to an environment where sufficient feature points are available. It cannot be used if a good number of feature points is not available. Also we have assumed planar surface, this assumption is not valid if UAV is closer to earth surface. This work can be extended to environment were less number of features points are available and earth surface has uneven ups and downs. We have assumed that a camera is fixed to the UAV, which can be relaxed with little complexity, so that the camera may always point towards the region where a sufficient number of feature points are available. Further computation complexity can be reduced. Finally, this work is based on RGB camera which is highly sensitive to lighting condition, and thus developing a system based on the IR camera would be beneficial because that will enable its use even in low lighting conditions.

REFERENCES

Adiv, G. (July 1985). Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Transactions on pattern analysis and machine intelligence*, 7(4):384 – 401.

Dusha, D., Boles, W., and Walker, R. (Dec. 2007). Attitude estimation for a fixed-wing aircraft using horizon detection and optical flow. In *Digital Image Computing Techniques and Applications, Australian Pattern Recognition Society on*, pages 485 – 492.

Grabe, V., Bulthoff, H., and Giordano, P. (2012). On-board velocity estimation and closed-loop control of a quadrotor uav based on optical flow. In *IEEE International Conference on*, pages 491–497.

Jepson, A. D. and Heegar, D. J. (1991). A fast subspace algorithm for recovering rigid motion. In *Visual Motion, Proceedings of the IEEE Workshop on*, pages 124 – 131.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. 82:35–45.

Ma, Y., Soatto, S., Kosecka, J., and Sastry, S. (2004). *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer.

Oreifej, O., Lobo, N., and Shah, M. (2011). Horizon constraint for unambiguous uav navigation in planar scenes. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 1159–1165.

Sazdovski, V., Silson, P., and Tsourdos, A. (2010). Attitude determination from single camera vector observations. In *Intelligent Systems (IS), 2010 5th IEEE International Conference*, pages 49–54.

Schmid, C., Mohr, R., and Bauckhage, C. (June 2000). Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172.

Shi, J. and Tomasi, C. (June 1994). Good features to track. In *Computer Vision and Pattern Recognition,IEEE Computer Society Conference on*, pages 593–600.

Siouris, G. M. (1993). *Aerospace Avionics Systems : A Modern Synthesis*. Academic Press Inc.

Soatto, S., Frezza, R., and Perona, P. (March 1996). Motion estimation via dynamic vision. *Automatic Control, IEEE Transactions on*, 4(3):393–413.

Zisserman, A. and Hartley, R. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press.