Abstract—In the Unmanned Air Vehicle (UAV) navigation the main challenge is estimating and maintaining the accurate values of UAVs position and orientation. The onboard Inertial Measurement Unit (IMU) provide the measurements but it is mainly affected from the accumulated error due to drift in measurements. Traditionally the Global Position System (GPS) measurements of vehicles position data can be fused with IMU measurements to compensate the accumulated error. But the GPS signals is not available everywhere and it will be degraded or fully not available in hostile areas, building structures and water bodies. Researchers already evolved methods to handle the UAV navigation in GPS denied environment by using Vision based navigation like Visual Odometry (VO) and Simultaneous Localisation and Mapping (SLAM). In this survey paper we attempted to understand the existing research towards vision based navigation and finally proposed a Modular Multi-Sensor Data Fusion technique for UAV navigation in the GPS denied environment.

Keywords: Unmanned Air Vehicle (UAV), Inertial Measurement Unit (IMU), Visual Odometry (VO), Simultaneous Localisation and Mapping (SLAM)

I. INTRODUCTION

Unmanned Air Vehicle (UAV) is an upcoming technology with great value to modernize the military applications and facilitate innovative civilian applications. In the initial days the UAVs were mostly utilized only for the military applications like aerial reconnaissance. But recent days UAVs have seen a dramatic increase in civilian uses like surveying the required geographical area, collecting the images and required data from remote areas, forest fire monitoring and agriculture information gathering [1].

Early Drones were controlled by human operator from the ground station by radio unit. But the recent UAVs are more sophisticated systems can fly autonomously based on pre-programmed flight plans. Autonomous navigation of UAVs uses onboard IMU. The IMU has three accelerometer and three gyro to find the linear movement and angular movement from the starting point. But the IMU measurements will suffer with accumulated error over time. Even the drift error is very small it will accumulate to a large value over time. GPS location information was used to compensate the IMU accumulated error by fusing with IMU measurements.

In the real world scenario when the UAV will be flying close to vegetation, water bodies, hostile environment and inside structures the GPS signals will be normally degraded or fully absent. In the absence of GPS the navigation of Unmanned Air Vehicle is very challenging because the navigation is fully depend on inaccurate on-board IMU [2], [3].

To overcome the above mentioned difficulty of UAV navigation in GPS-denied environment, already different techniques have been evolved by the researchers by using vision based navigation [4], [5]. In this survey paper we attempted to analyse the already existing research on methods to get better pose estimation of the navigating vehicle in the GPS denied environment by fusing vision information with the IMU measurements.

II. TECHNIQUES

The basic characteristic of IMU and Camera is perfectly suitable to work with each other for providing accurate estimation. IMUs are ideally chosen for navigation purpose, because it can provide the measurement data at a higher sampling rate. But over time IMU measurements are drifted and generate the accumulated error. Vision sensor behaves in the other way, the measurements are at low sampling rate due to the long computation time for processing the captured image and the measurement will not drift over time. By fusing vision and inertial sensors provides a better pose estimation of the UAV.

In the vision based navigation approach the camera is mounted on the UAV and the camera moves through its surroundings and the motion of image features can be used to find the UAV trajectory. Most of the researchers follow the following two techniques for this vision based approach,

- Simultaneous Localisation and Mapping
- Visual Odometry

A. Simultaneous Localisation and Mapping

SLAM incrementally builds a consistent live map of the surrounding environment and estimates the pose of the vehicle by the doing matching between this live map of vehicle surrounding environment built with the currently captured image. The SLAM problem structure, convergence result and acronym SLAM was first presented by H. Durrant-Whyte [6] at the ISRR 1995. The advancement in SLAM helps the vehicle to construct a live map of the surrounding environment at the same time localize the vehicle by using the built map of its environment. This localization technology helps the vehicle to navigate autonomously in the remote environments. Vehicle
mounted with camera moving through the environment and capturing the unknown image features can be seen in Figure 1 as provided by Hugh Durrant Whyte and Tim Bailey [7].

\[ T_k = [R_{k,k-1} t_{k,k-1}; 0 1] \]

Where, \( R \) denotes rotation matrix and \( t \) represents the translation vector.

The set of camera poses \( C_{0:n} = [C_0, ..., C_n] \) contains vehicle camera transformations taken with respect to the frame taken at instant \( k=0 \). The current pose \( C_n \) can be estimated by integrating all the camera transformations \( T_k \), \( k= 1...n \),

\[ C_n = C_{n-1} T_n \]

III. LITERATURE SURVEY

In the proceeding work, the research papers were analysed which uses the SLAM or VO techniques for vision based navigation of the vehicle when the GPS signal is degraded or fully not available. Visual sensor like camera posses so many advantages, mainly the cameras are less in weight and require less power which makes it can be easily mounted on the UAV. The camera in UAV can simultaneously support two purposes. First, it can be used for visual surveillance of the required geographical areas and second, it can be used for estimating the pose of UAV in GPS denied environments.

Omead Amidi [9] designed Visual Odometry in the year 1996 is one among the first researchers who proposes using the vision data for estimating the unmanned vehicles position.
The VO was used for unmanned helicopter position estimation. The odometer locks on to ground objects viewed by stereo camera. By using high speed template matching of consecutive images it estimates the helicopter position. The VO unit is integrated with midsized Yamaha R50 helicopter for vision guided autonomous flight.

Bryce B Ready [2] presented a system for estimating the Micro Air Vehicle (MAV) position using the Omed Amidi [9] designed VO method for accurate position estimation. The author used an Iterative registration Method of VO for UAV pose estimation. The algorithm is verified by generated a synthetic video and corrupted the true pose by adding Gaussian noise. As well this was validated with MAV flight data. The result shows that GPS/INS measurement with VO has the better performance with maximal standard deviation of around 2.5 meters against 5.6 meters deviation without VO.

Evan D. Anderson [4] claimed that the iterative registration method proposed by Bryce B Ready [2] requires an extensive computation time for pose estimate and proposed a new method to do pose estimation with GPS/IMU data along with camera data without the 3D location calculation. The homography matrix is formed by combining the visual sensor measurements and it has been provided as the measurement variable input to the fusion. This is the first proposed method to use UKF to enable vision and IMU measurements data Fusion. The algorithm is evaluated by the simulated data.

Clark N Taylor [3] has utilized the epipolar constraint based UKF framework for fusing vision, IMU data air pressure sensors. Epipolar constraint is used over SLAM-based technique to achieve faster calculations by tracking features between two frames. The following three main drawbacks of epipolar constraint are addressed in this paper. 1. Bias introduced when visual sensor moved towards the image point centre. 2. Scale ambiguity in the camera measurement. 3. Visual sensor navigation state is always estimated with its prior navigation states.

Chaolei Wang [5] claims that Omed Amid [9] design of using stereo camera is successful in small UAVs, but the distance and calibrating the stereo camera each time before flight is the biggest limitation and it also incur more computational cost. Monocular camera provides good scalability and accuracy and less calculation than stereo camera. He suggested using other sensors like ultrasonic sensor or barometric altimeter for measuring image features depth data. Barometric altimeter measures the helicopter height and it is also used in the Visual Odometry. The algorithm is verified on the Test bed which is gas-powered radio-controlled model helicopter integrated with required sensors like IMU, altimeter, GPS and etc.

Chaolei Wang [10] also suggested to use monocular visual SLAM for aerial vehicles pose estimation in the GPS degraded environment. The estimate of the vehicle movement is done using homography matrix. The EKF takes the visual SLAM from monocular camera and IMU measurements and estimate the pose of the UAV and build the real-time map of the surrounding environment.

Dachuan Li [11] says that the when cameras and laser range finders are used for position estimation of UAVs it basically takes an assumptions about the structures and surrounding environment. These assumptions limit the effectiveness of the estimation in the 3D scenes. At the same time processing the 3D information required more computation power and time. To overcome these issues the author proposed fusing the IMU data with the RGBD sensor. The advantage of RGBD sensor is, in addition it gives the depth information of the environment. By EKF, the RGBD sensor (Microsoft Kinect is used) data are combined with GPS/INS measurement. The algorithm is verified in the indoor environment with the Kinect integrated quadrotor and observed maximum deviation of approximately 8 cm.

S Rady [12] presented that the SLAM algorithms perform very well when UAVs fly in the indoor environment but it is still a big challenge when the UAV fly in the outdoor environments. The author proposes utilizing the maps which are pre-existing for UAV pose estimation. The construction of the surrounding map is reduced with set of most informative landmarks by information-theoretic. From the map only the more informative ones are passed and less informative data is filtered out to provide more localization accuracy and less computational requirement.

Justin M. Barrett [13] presented a navigation system for unmanned vehicle which is hybrid in nature, it has on-board stereo camera and INS to estimate the vehicle position and orientation. The linear and angular movement of the autonomous vehicle is differentiated by the image processing algorithm which receives the inverse depth of the image features of the scene from the stereo camera. The accumulated drift in the IMU data need to be corrected to achieve accurate position estimation of the vehicle. The author proposes Bayesian Information Filter nothing but the EKF which fuses the camera data with IMU data to correct the vehicle position drift introduced by IMU. The algorithm is verified on the Husky A100 UGV.

Daniel Magree [17] proposed a navigation system for the unmanned vehicle which integrates both the Visual SLAM and Laser SLAM for achieving better pose estimation of the vehicle. The purpose of visual SLAM and laser SLAM integrated is even if one fails other will help in providing robust pose estimation. Generally the laser based SLAM can fail with ambiguous laser return profiles and the visual SLAM can fail with inadequate or ambiguous image features in the scene. These algorithms are verified individually in visual SLAM and laser SLAM in the Georgia Tech UAV Simulation Tool (GUST) and as well flight tested on quad rotor.

Annett Chilian [14] proposed the Information filter for fusing the IMU and Visual Odometry measurements to achieve Better Pose Estimation The information filter is numerically equivalent to the Kalman filter.

Andrew Chambers [15] suggested using screening gates before MSDF to create smooth estimate of the UAV position and orientation. For better estimation, the measurements provided to the MSDF have to be very accurate. The figure 4 shows the overall filter design to achieve better pose estimation of
In this, the measurement gating module will reject the erroneous measurement from sensors which are above the predefined threshold value before fusing.

![Image of Robust MSDF](image)

**Fig. 4. Robust MSDF.**

Davida Scaramuzza [16] provides the challenges faced while executing the autonomous navigation project called Swarm of Micro Flying Robots (SFLY). It is an European project implemented and verified for a period of around three years. These micro flying robots can autonomously navigate in GPS degraded environments. The major contributions in this SFLY project are new hexacopter developed with computer vision, SLAM based Navigation was developed. For global navigation, maps from different MAVs were collected and combined in the offline and position of the vehicle is calculated by using adaptive optimization algorithm.

Gianpaolo Conte [18] utilized the available satellite images in addition to Visual Odometry data to avoid the visual odometry drift. In the proposed method he used the existing satellite images as the reference images along with the real time image from the onboard camera for the image registration algorithm. With this vision aided method, visual odometry drift will be avoided.

Duo-Yu Gu [19] proposed to use Geographic Information System (GIS) for image aided UAV navigation. The limitations like more storage requirement of scene matching and high computation requirement of terrain matching are addressed in this GIS using method. From the satellite images the important feature points like bridges, road and rivers are extracted and GIS model is prepared. Then the location of the UAV is identified by the image registration method using the image from onboard camera and the reference image from GIS model.

Jay Farrell [20] provided a presentation for high degree of accuracy required for vehicle navigation can be achieved by fusing the information from IMU, GPS, camera and signals of opportunity. The Signals of opportunity is nothing but the RF signals from digital TV, digital radio and cellular. Since these are opportunistic signals they are not synchronized with receiver clock. But with time difference of arrival (TDOA) techniques localization of vehicle can be achieved with this asynchronous opportunistic signals [21].

**IV. PROPOSED ARCHITECTURE**

For the UAV navigation, vision aided approaches in the GPS signal degraded/absence environment is analysed. In the analysis of existing researches we noticed that most of the implementations were using either Visual SLAM or Visual Odometry, EKF with inertial data for better pose estimation of UAV. The Visual SLAM algorithm is slower and need more computation power and memory for building and maintaining the live map of the features (landmarks) compare to Visual Odometry.

In practical scenario with flying UAV the processing time, computation power and memory is the main constraint. In the proposed approach by keeping these constraints we are planning to use Visual Odometry along with GPS/IMU telemetry data. The Figure 5 provided after the TABLE 1 shows the proposed architecture for UAV navigation in the GPS degraded or denied environments. This is the modular MSDF approach where we can add any required sensor in future to get better pose estimation of the UAV without much change in the architecture and code. The aim of this modular MSDF approach is to accurately estimate UAV pose which may be represented in terms of three dimensional location (x,y and z) and orientation (pitch, yaw and roll).

In the proposed design the screening gate is used to validate(accept/reject) the actual GPS measurement based on the threshold. This screening gate will make sure only the undegraded GPS data will be passed on to fusion algorithm.

Before fusing the measurements from the IMU and VO, the first stage is to do the required preprocessing of the measurements. The IMU measurements have deterministic and dynamic errors like bias and scale factor errors. The error compensation block in the MSDF first remove the biases and correct the scale factors before fusion. As well the pose estimation from Visual Odometry needs to be translated to E-Frame before fusion. The E-frame is Earth-centred earth fixed frame which resolved earth rotation with respect to I-Frame. Generally if the measurements from the sensors are linear functions and affected by only Gaussian noise then we can use the Kalman filter approach. But in our UAV navigation application linearity assumption will not hold well, so we need to use Extended Kalman filtering to fuse the measurements. EKF is based on approximation of the Bayes rule using linearization.

For EKF, the measurements from the GPS and IMU modules are considered as the filter state observed value and these measurements are the pose vector current true value. The Visual Odometry will provide the pose estimate to MSDF which is non linear with respect to previous pose estimate. EKF calculates the Visual Odometry based present pose estimate with known previous pose estimate.

**V. COMPARISON STUDY**

In this survey we have done the study on the existing methods of Unmanned Vehicle navigation in the GPS denied Environments. The survey reveals that most of the methods using either monocular Camera or stereo camera along with IMU to do the pose estimation in the GPS degraded environments. With this extensive study on this topic we prepared the comparison summary as shown in TABLE 1.
<table>
<thead>
<tr>
<th>Sl.No</th>
<th>Type of Vehicle</th>
<th>Strategy</th>
<th>Sensors Used</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AscTec Pelican Quadrotor</td>
<td>Visual Odometry</td>
<td>Stereo Camera</td>
<td>2015</td>
<td>[26]</td>
</tr>
<tr>
<td>2</td>
<td>Quadrotor (GTQ)</td>
<td>Visual SLAM and Laser SLAM with EKF</td>
<td>IMU, Sonar, Scanning Laser and camera</td>
<td>2014</td>
<td>[17]</td>
</tr>
<tr>
<td>3</td>
<td>Hexacopter</td>
<td>Visual SLAM with EKF</td>
<td>IMU, Monocular Camera</td>
<td>2014</td>
<td>[16]</td>
</tr>
<tr>
<td>4</td>
<td>Mikrokopter</td>
<td>EKF</td>
<td>IMU, monocular Camera, GPS, Barometric altimeter</td>
<td>2014</td>
<td>[15]</td>
</tr>
<tr>
<td>5</td>
<td>six wheeled UGV</td>
<td>Bayesian Information filter (EKF)</td>
<td>IMU and Stereo Camera</td>
<td>2013</td>
<td>[13]</td>
</tr>
<tr>
<td>6</td>
<td>AscTec Pelican MAV</td>
<td>VO and SLAM</td>
<td>Stereo Camera</td>
<td>2013</td>
<td>[27]</td>
</tr>
<tr>
<td>7</td>
<td>Simulator with flight data</td>
<td>UKF</td>
<td>IMU, GPS and Camera</td>
<td>2013</td>
<td>[34]</td>
</tr>
<tr>
<td>8</td>
<td>Quadrotor</td>
<td>Invariant EKF</td>
<td>IMU and RGBD Odometry(Kinect)</td>
<td>2013</td>
<td>[11]</td>
</tr>
<tr>
<td>9</td>
<td>AscTec Pelican Quadrotor</td>
<td>UKF</td>
<td>IMU, Monocular Camera</td>
<td>2013</td>
<td>[24]</td>
</tr>
<tr>
<td>10</td>
<td>AscTec Firefly MAV</td>
<td>EKF</td>
<td>IMU, Pressure Sensor and Monocular camera</td>
<td>2013</td>
<td>[25]</td>
</tr>
<tr>
<td>11</td>
<td>Quadrotor</td>
<td>Visual SLAM with KF</td>
<td>IMU with monocular camera</td>
<td>2013</td>
<td>[22]</td>
</tr>
<tr>
<td>12</td>
<td>Hexacopter</td>
<td>Visual SLAM with EKF</td>
<td>IMU and WVGA monocular Camera</td>
<td>2012</td>
<td>[29]</td>
</tr>
<tr>
<td>13</td>
<td>Simulator with data</td>
<td>EKF</td>
<td>IMU, Monocular Camera</td>
<td>2012</td>
<td>[23]</td>
</tr>
<tr>
<td>14</td>
<td>Multi-Stereo Helmet tracking system</td>
<td>EKF</td>
<td>IMU and Monocular Camera</td>
<td>2012</td>
<td>[35]</td>
</tr>
<tr>
<td>15</td>
<td>Test bed which is gas powered radio controlled model helicopter</td>
<td>Visual SLAM with EKF</td>
<td>IMU, Monocular Camera</td>
<td>2012</td>
<td>[10]</td>
</tr>
<tr>
<td>16</td>
<td>Quadrocopter</td>
<td>Visual SLAM with EKF</td>
<td>IMU, Pressure Sensor, USB Firefly monocular camera</td>
<td>2011</td>
<td>[28]</td>
</tr>
<tr>
<td>18</td>
<td>Scout B1-100 Helicopter</td>
<td>Using Pre-Existing Maps</td>
<td>IMU and Monocular Camera</td>
<td>2011</td>
<td>[12]</td>
</tr>
<tr>
<td>19</td>
<td>Six Legged Crawler</td>
<td>Visual Odometry</td>
<td>IMU with Stereo Camera</td>
<td>2011</td>
<td>[14]</td>
</tr>
<tr>
<td>20</td>
<td>Simulator with Flight Data</td>
<td>Image Registration using GIS Data</td>
<td>IMU, GPS, Camera and GIS Data</td>
<td>2010</td>
<td>[19]</td>
</tr>
<tr>
<td>21</td>
<td>Simulator with Flight data</td>
<td>Visual SLAM</td>
<td>Camera</td>
<td>2010</td>
<td>[38]</td>
</tr>
<tr>
<td>22</td>
<td>Quadrotor</td>
<td>Visual SLAM with EKF</td>
<td>IMU, Stereo camera, Monocular color camera with Laser finder</td>
<td>2010</td>
<td>[32]</td>
</tr>
<tr>
<td>23</td>
<td>HMAV</td>
<td>EKF</td>
<td>IMU and Wi-Fi Camera</td>
<td>2009</td>
<td>[30]</td>
</tr>
<tr>
<td>24</td>
<td>Quadcopter</td>
<td>KF</td>
<td>IMU and VGA Camera</td>
<td>2009</td>
<td>[31]</td>
</tr>
<tr>
<td>25</td>
<td>Yamaha RMAX Helicopter</td>
<td>KF with Image Registration</td>
<td>IMU, GPS, Camera and Satellite Images</td>
<td>2008</td>
<td>[18]</td>
</tr>
<tr>
<td>26</td>
<td>Simulator with Vehicle data</td>
<td>Kalman Filter</td>
<td>IMU with Laser scanner</td>
<td>2008</td>
<td>[37]</td>
</tr>
<tr>
<td>27</td>
<td>Simulator with synthetic MAV flight data</td>
<td>UKF framework utilizing epiopolar constraint</td>
<td>IMU and Stereo Camera</td>
<td>2008</td>
<td>[3]</td>
</tr>
<tr>
<td>28</td>
<td>Simulator with MAV flight data</td>
<td>Iterative Registration method, UKF</td>
<td>IMU with monocular camera</td>
<td>2007</td>
<td>[4]</td>
</tr>
<tr>
<td>30</td>
<td>Acrobatic 23cc helicopter</td>
<td>Non Linear observer</td>
<td>IMU and Webcam</td>
<td>2007</td>
<td>[33]</td>
</tr>
<tr>
<td>31</td>
<td>Simulator with Vehicle data</td>
<td>EKF</td>
<td>IMU with Camera</td>
<td>2007</td>
<td>[36]</td>
</tr>
</tbody>
</table>
VI. CONCLUSION

In this survey paper we presented the different techniques used by the existing researchers for the UAV navigation in the GPS denied environment using vision based navigation. A comparison study chart has been prepared and listed in the Table 1. Based on our survey we found that the Visual Odometry based approach will be faster and required less memory and computational power compare to SLAM. In Section IV we proposed modular Multi Sensor Data Fusion approach for UAV navigation with better pose estimation in the GPS denied environment.

REFERENCES


