An Auto-Encoder enabled Fault Detection and Isolation Scheme for enabling a Multi-Sensorial Distributed Pose Estimation

Moumita Mukherjee, Avijit Banerjee, and George Nikolakopoulos

2022 IEEE 31st International Symposium on Industrial Electronics (ISIE) | 978-1-6654-8240-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIE51582.2022.9831458

***Abstract*—This article proposes a novel Auto-encoder based fault detection and isolation framework approach for supporting the operation of a novel multi-sensorial distributed pose estimations scheme. The proposed work detects weak and strong time-dependent anomalies in a decentralised fusion approach from the initial estimation layer. As it will be presented, at the end of the learning phase, the neural network-based auto-encoders provide synthetic actual position and orientation of the robotic system, based on the statistics of the learning data. As a result, the square error between the output and input signal of the auto-encoder can yield the actual outlier with reasonable success. On the other hand, an Extended Kalman Filter (EKF) based fault detection method has been introduced in this article, which consists of a set of judiciously designed EKF acts as filter assembly. Based on innovation obtained from each of the EKF an innovative detection logic is proposed to identify the outlier in sensor measurement autonomously at the appropriate time samples. Based on the degree of accuracy of detecting the anomaly, the estimated signal is accepted or rejected for each time sample in the second layer of the fusion architecture. Moreover, we will introduce two outlier detection methods for the demonstration purposes and outline a comparative study using experimental data from a micro aerial vehicle. An extensive analysis with supporting results demonstrate these two methods' effectiveness and accuracy.**

***Index Terms*—Auto encoder, Multi sensor fusion, Decentralize fusion, Filter bank, Maximum likelihood function, Optimal information filter.**

I. Introduction

Robotic applications involving operations in changing landscapes present a variety of challenges, where in many cases the systems may fail to provide accurate pose information depending on the sensors' accuracy, failures and noise in the operating environment. However, such corrupted sensor measurements, for a temporary operating period, are often encountered in reality as for example, in the absence of sufficient illumination conditions for some part of the surrounding environments, where the visual sensors fail to determine the robot's pose and thus further fusion with other sensorial onboard systems is needed as in the case of a laser/infrared technology-based sensors, or in general multiple redundant sensors for persistent fused pose estimation, as an effort to elevate the accuracy of the pose estimation and ensure resiliency simultaneously.

In most of the cases in the robotic community, the multi-sensor based fusion algorithms rely on a centralized architecture, where a central node is used to process the information by involving a Kalman filter from different sensor measurements [1], or distributed approaches as in [1], [2]. In centralized fusion architectures, a central node or a single node is utilized, where direct measurement data or raw data from multiple sensors are used to fuse by utilizing several type of Kalman filters, depending upon the system whether it is linear or nonlinear. In the centralised approaches, the overall performance of the fusion scheme degrades in the presence of sensor failure/inaccurate measurements, thus for increasing the overall resiliency, towards avoiding inaccurate sensor measurements, decentralized fusion architectures are more suitable as in [3].

However, even in the case of fusion approaches based on a decentralized architecture, it is not enough to overcome such circumstances without incorporating the full integration of a fault detection and isolation framework. Thus, approaches based on the Fault Resilient Optimal Information Filter (FROIF) are introduced as in [4] to address the gap where the failure detection is the most crucial task. The failure scenario in the measurement signals can be considered as an abnormal behaviour or deviation from the actual value and in general can be considered as an outlier and in most of the cases, the Kalman filter innovation can be utilized for that outlier detection. In this direction, the Innovation based Detection Filter Bank [5], the Robust Gaussian ESKF [6], and the Iterated Kalman filter detection architecture [7] have been proposed to overcome the anomalies from the estimated states to incorporate the concept.

Towards the anomaly detection or outlier detection, many approaches have been established e.g., signal processing [8], state estimation [9], machine learning [10], statistics [11], data mining [12]). In general, it has been found that anomaly detection is a challenging task, with different paradigms of detection schemes showing success on different types of data. Therefore, one of the challenging jobs is to select one method for addressing this kind of challenge faced in real-world applications. The quality of a neural network for classifying information has also been utilized to identify the anomaly in time-series signals. The most popular training method for neural networks is the back-propagation algorithm for feed-forward Neural networks. A feed-forward, fully connected neural network, consisting of three layers, is used for the outlier detection. This configuration is known as Auto-encoder that is able to provide synthetic reconstructed output and the square error between the network output and input values, which has been used to locate the deviation of the original signal as presented in [13].

Based on the presented background, this article is aiming in establishing a totally novel auto-encoder-based decentralized

This work has been partially funded by the European Unions Horizon 2020 Research and Innovation Programme under the Grant Agreement No. 869379 illuMINEation.

[1]Robotics and AI, Department of Computer, Electrical and Space Engineering, Luleå University of Technology, Luleå {moumuk, aviban, geonik}@ltu.se

fusion scheme for estimating the pose of a robot. Thus, the major contributions of this article stems from: a) the establishment of a unique detection logic to automatically identify the faulty sensors for each time instant by involving a collection of judiciously designed Extended Kalman filter (EKF), b) Introducing a generic AI based fault detection scheme by applying several neural networks based Auto-encoders, c) performance evaluation of the proposed Auto-encoder based and extended Kalman filter based fault detection over the FROIF framework with the help of actual sensor data, along with a comparative study between these two detection algorithms.

In the rest of the article, the platform based mathematical framework is established in Section II, while the Kalman filter based and Auto-encoder based fault resilient optimal information filter are explained elaborately in Section III. Section IV represents the experimental results that prove the proposed scheme's efficiency, while the conclusions are stated in Section V.

II. Platform Based Mathematical Framework Establishment

For the problem formulation we will consider the realistic use case of a Micro Aerial Vehicle (MAV) that carries a multi-sensorial suit containing an IMU, a 3D Velodyne Puck LITE lidar, a Real-sense camera T265 and an Ultra wide-band. In this sensorial setup, the IMU provides the raw acceleration, the angular velocity and the orientation. The Intel real-sense camera T265 combined with an IMU from Pixhawk 4 flight controller are used for providing a visual odometry [14], and the collected point-clouds from the 3D-lidar are also integrated with an IMU to provide a Lidar Inertial Odometry (LIO) [14]. The sensors mentioned above are mounted in the body-fixed frame $X_B - Y_B - Z_B$ of the MAV, as depicted in Fig. 1. The conventional East-North-Up (ENU) based world frame is considered in the present work, which is used to describe the position of the MAV. Ultimately, the goal is to build a unique and purely autonomous outlier detection system based on the fault-resilient optimal information fusion architecture [4], combining the information from multiple real-time sensors and calculating the most accurate MAV pose. The following nonlinear kinematic model for the MAVs is considered as in [15], where the kinematic equations are given as:

$$\begin{aligned} \dot{\mathbf{p}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= \mathbf{R}_{\mathbf{B}}^{\mathbf{W}}(\mathbf{q})\left(\mathbf{a}_m - \mathbf{a}_b - \mathbf{a}_n\right) + \mathbf{g} \\ \dot{\mathbf{q}} &= \frac{1}{2}\mathbf{q} \otimes \left(\boldsymbol{\omega}_m - \boldsymbol{\omega}_b - \boldsymbol{\omega}_n\right) \\ \dot{\mathbf{a}}_b &= \mathbf{a}_\omega \\ \dot{\boldsymbol{\omega}}_b &= \boldsymbol{\omega}_\omega \\ \dot{\mathbf{g}} &= \mathbf{0}_{3\times 1} \end{aligned} \tag{1}$$

where $\mathbf{p} \in \mathbb{R}^{3\times 1}$ indicates the position, $\mathbf{v} \in \mathbb{R}^{3\times 1}$ denotes the velocity, $\mathbf{q} \in \mathbb{R}^{4\times 1}$ stands for the orientation quaternion representation of the MAV. The acceleration and body rates are characterized as the input to the kinematic model, which

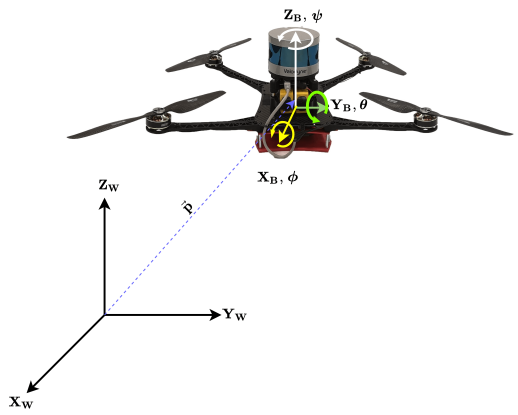

Figure 1: The utilized MAV for the problem formulation and the related utilized frame, where the subscript W indicates the global frame and subscript B indicates the body frame.

are typically measured using an IMU denoted as a_m and ω_m. In general, measurements from the sensors are noisy, which includes a sensor bias as well. $a_b \in \mathbb{R}^{3\times 1}$ denotes the accelerometer and $\omega_b \in \mathbb{R}^{3\times 1}$ represents the gyroscopic bias terms, a_n and ω_n signify the additive noise for the acceleration and the angular rate noise respectively and $g \in \mathbb{R}^{3\times 1}$ denotes the gravitational acceleration acting on the MAV. $a_\omega \in \mathbb{R}^{3\times 1}$ and $\omega_\omega \in \mathbb{R}^{3\times 1}$ are the accelerometer and the gyroscopic process noise respectively. The rotation matrix $R_B^W(q) \in SO(3)$ is introduced to perform the frame transformation from the body to the world frame. The equations of motion for the MAV is expressed in a compact mathematical notation with the following generalized form, provided as:

$$\dot{X} = f(X, u_m, w) \tag{2}$$

$$y_t = [\boldsymbol{I}_{7\times 7} \mid \mathbf{0}_{7\times 12}]X_t \tag{3}$$

where, the state vector is denoted as $X = [p, v, q, a_b, \omega_b, g]^T \in R^{19\times 1}$, the noisy measured input based on IMU reading is denoted as $u_m \in R^{6\times 1}$ and the random process noise $w \in R^{6\times 1}$.

III. Overview of a Fault Resilient Optimal Information Filter

As presented in [4], the decentralized multi-sensor fusion is a more suitable pose estimation method over the centralized fusion when all sensors are not working correctly or the measured raw data signals can be inaccurate at anytime sample. A brief overview of the Fault Resilient Optimal Information Filter (FROIF) is presented in sequel for the sake of completeness.

The construction of the FROIF inherits a distributed nodal architecture, which is introduced in the first layer of the decentralized filter. In the context of the multi-sensorial fusion based MAV pose estimation, position and orientation information, obtained from any two distinct sensors (arbitrarily selected) are utilized as the measurement information to the EKF for constructing a node. By exploring all of these possible

combinations of sensor measurements, collectively, a total number of seven nodes are built in the current setup and alphabetically denoted as node-l, and $l \in \{A,\ B, \ldots, G\}$. However, each node receives the IMU information as measured actuation/control input to the kinematic model associated with the Kalman filter. A more detail description of the nodes can be found in [4]

The kinematic model of the MAV in a continuous time form was presented in Eq. (2). In order to describe the decentralized nodes, associated with the first layer, in a compact mathematical form and in correlations to the EKF, an equivalent of Eq. (2) in discrete time representation, by using Euler [16] discretization, is given by:

$$x_k = f_{k-1}(x_{k-1}, u_{k-1}, \omega_{k-1}) \quad (4)$$
$$y_k = h_k(x_k, v_k) = [\boldsymbol{I}_{7\times 7} \mid \mathbf{0}_{7\times 12}]x_k + v_k \quad (5)$$

where, k denotes the discrete time instants. Note that, in order to account for the model inaccuracies we have considered also the process noise $\omega_k \in \mathbb{R}^{19}$. Moreover, an additional measurement noise vector $v_k \in \mathbb{R}^7$ is introduced in the Eq. (5) to represent a realistic output model, under the influence of noisy measurements, appearing from the utilization of real sensors. The process and measurement noises are assumed to follow the Gaussian distribution as:

$$\omega_k \sim \mathcal{N}(0,\, Q_k), , v_k \sim \mathcal{N}(0,\, R_k) \quad (6)$$

where Q_k and R_k represent the process noise, the co-variance matrix, and the measurement noise co-variance matrix respectively. Starting with an initial guess of a posterior estimate $\hat{x}^+_{l_0} = E(x_{l_0})$ and $P^+_{l_0} = E[(x_l - \hat{x}^+_{l_0})(x_l - \hat{x}^+_{l_0})^T]$, along with the assumption in Eq. (6), the l^{th} node is described as a local EKF with the following prediction-correction formalism:

Prediction Steps:

$$\hat{x}^+_{l_0} = E(x_{l_0}), P^+_{l_0} = E[(x_l - \hat{x}^+_{l_0})(x_l - \hat{x}^+_{l_0})^T] \quad (7a)$$
$$\hat{x}^-_{l_k} = f_{l_{k-1}}(\hat{x}^+_{l_{k-1}}, u_{k-1}, 0) \quad (7b)$$
$$K_{l_k} = P^-_{l_k} H^T_{l_k}(H_{l_k} P^-_{l_k} H^T_{l_k} + R^T_{kl})^{-1} \quad (7c)$$
$$P^-_{l_k} = F_{l_k} P^+_{l_k} F^T_{l_k} + L_{l_k} Q_{l_k} L_{l_k} \quad (7d)$$

where, the '$+$' symbol is used to denote an a priory estimate, the '$-$' symbol is designated for an a posterior estimate, the subscript l indicates the corresponding variable of the l^{th} node, where $l \in \{A,\ B, \ldots, G\}$. The mathematical operator E denotes the expectation and the superscript T indicates the transpose. The Jacobian matrices are defined as:

$$F_{l_k} = \frac{\partial f_{l_{k-1}}}{\partial x_{l_k}},\ L_{l_k} = \frac{\partial f_{l_k}}{\partial u_k},\ H_{l_k} = \frac{\partial h_{l_k}}{\partial x l_k} \quad (8)$$

Correction Steps:

$$\hat{x}^+_{l_k} = \hat{x}^-_{l_k} + K_{l_k}[y_{l_k} - h_{l_k}(x^-_{l_k}, 0)] \quad (9a)$$
$$P^+_{l_k} = (I - K_{l_k} H_{l_k}) P^-_{l_k} \quad (9b)$$

It is apparent that individually, each node is potentially capable of proving the information regarding the pose of the MAV. However, the accuracy of the pose information, obtained from each decentralized node, varies depending on the accuracy of the sensors that are involved to constitute the node. When the sensor measurements fail to provide an accurate information for a certain period, the FROIF [4] has the flexibility to momentarily isolate the corresponding node to proceed for the next layer fusion formalism, which combines the information from the first layer in an weighted combination as follows:

$$\sum_{l\in\{A,\ldots,G\}} \delta_l \bar{A}_{l_k} = \mathbf{I} \quad (10)$$

where $\bar{A}_{l_k}$ denotes the weighting parameter and $\delta_l \in \{0, 1\}$ is a scalar multiplying factor. FROIF imposes $\delta_l = 0$ as well the corresponding $\bar{A}_{l_k} = 0$ if the l^{th} node is found to be corrupted at k^{th} time instant, otherwise $\delta_l = 1$ and the corresponding $\bar{A}_{l_k}$ can be optimally constructed based on the following constrained co-variance minimization problem described as:

$$\min_{A_A,\ldots,A_G} J_k = \frac{1}{2} tr(P_k) = \frac{1}{2} tr(\mathbf{W}^T_k \boldsymbol{\Sigma}_k \mathbf{W}^T_k) \quad (11)$$
$$\textit{subjected to } \left(\mathbf{W}^T_k \boldsymbol{e}_{\boldsymbol{\delta_k}} - I\right) = \mathbf{0}$$

where,

$$\mathbf{W_k} = [\bar{A}_{A_k}, \bar{A}_{B_k}, \ldots \bar{A}_{G_k}]^T, \quad (12)$$
$$\boldsymbol{e}_{\boldsymbol{\delta_k}} = [\delta_{A_k} I,\ \delta_{B_k} I,\ \ldots,\ \delta_{G_k} I]^T \quad (13)$$
$$\tilde{\mathbf{x}}_{\mathbf{L}} = [\tilde{x}_{A_k}, \tilde{x}_{B_k}, \ldots, \tilde{x}_{G_k}] \quad (14)$$

Solution of Eq. 15 provides the optimally constructed weight matrices that are used for the inbuilt fault isolation.

$$\mathbf{W}_k = \Sigma^{-1}_k \boldsymbol{e}_{\boldsymbol{\delta_k}} (\boldsymbol{e}^T_{\boldsymbol{\delta_k}} \Sigma^{-1}_k \boldsymbol{e}_{\boldsymbol{\delta_k}})^{-1} \quad (15)$$

However, the formulation presented in [4], is not explicitly capable to autonomously identify the faulty measurements. Its fault detection mechanism solely relies on the assistance from a Vicon based accurate motion capture system. In typical field robotic application, where such accurate external support is unavailable, an autonomous fault recognition based on the onboard multi-sensorial unit is mandatory, which is the primary contribution of the present work ans will be described in the next Section.

IV. Novel Fault Detection Approaches

Identifying the incorrect measurement at the appropriate time instant is a crucial task for accurate localization. The presence of faulty measurements can be considered as an abnormal pattern in the entire motion or a deviation from the actual value and in general it is regarded as an outlier. However, accurately detecting the outlier is challenging in the multi-sensor pose estimation paradigm, while this article presents two types of novel outlier detection algorithms.

A. Kalman based Fault Detection Formulation

In order to locate the inaccurate measurements, from the first layer fused nodes, the novel automatic fault detection mechanism (as described in Fig. 2) is proposed. In general, when a fault occurs, the innovation vector $(y - \hat{y})$, associated

with the classical EKF, is unable to locate the fault, while it is only capable of detecting the distortion instantaneously [5] by identifying a small glitch, which is mostly insignificant in typical applications. Thus, a filter bank that consists of multiple modules to detect the faulty sensors appropriately is introduced in this Section.

Typically, in a multi-sensor framework, consisting of a total number of n sensors, n number of modules are required as depicted in Fig. 2. In the present context, the output from the i^{th} sensor is denoted as $y_i,\ \forall i = 1, 2, \ldots n$. Each module is constructed using a judiciously designed EKF, which consists of three sensors measurements as its input. The EKF of the i^{th} module considers the measurements y_i, except the ith measurement, where $i = 1, ...n$. One can consider the associated EKF in each module as an equivalent centralized filter and estimate the MAV pose individually. In the sequel, a series of residues ($R_j = x_1 - \hat{x}_j, \forall j = 2, 3, \ldots n$) are constructed by subtracting the estimated pose obtained from the first module ($\hat{x}_1$) to i^{th} module ($\hat{x}_i$). Afterwards, and by comparing all the residues with an arbitrarily chosen threshold (Δ) of the filter bank, the detection observation table, as depicted in Fig.2, is provided.

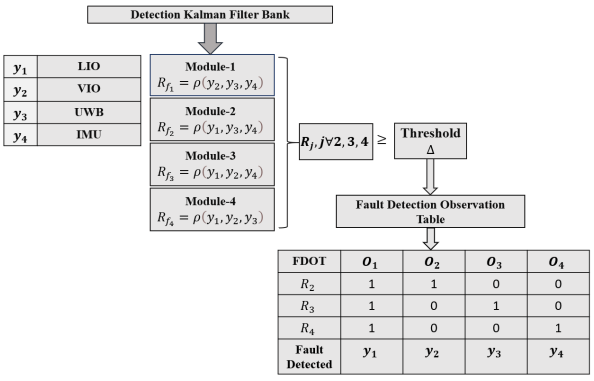

Figure 2: Kalman filter based detection architecture

In this case, the detection logic is presented as follows:

$$\begin{cases} \textit{if} & |R_j| \geqslant \Delta: \quad o_i = 1, \\ \textit{else} & o_i = 0, \forall i = 1, \ldots, n, j = 2, \ldots, n \end{cases} \tag{16}$$

$|.|$ is denoted for the absolute value operator and Δ is the threshold that can be chosen arbitrary. If an entire column of the observation table matches with the residue comparison, it will be convicted that the corresponding sensor is faulty. According to the aforementioned condition in Eq. (16), the observation table is constructed. Once the table is constructed, the pattern in the column are used for the detection. The observation table in Fig. 2 follows such a pattern of numbers and depends on that pattern, and thus it has the ability to detect which sensor is faulty at a specific time instant. For example, a pattern $(o_3 = (0, 1, 0))$ denotes the UWB is erroneous, while this detection method is only applicable if the fault occurs at a particular time instant. The limitation of this method is that it can't recognize a fault if two sensors are giving inaccurate measurements simultaneously. Once the fault detection unit identifies a sensor to be faulty, the scalar multiplying factor δ_{l_k}, for the corresponding nodes, are set as $\delta_{l_k} = 0$.

B. Auto-encoder based Fault Detection Formulation

Auto encoders are based on three layers of feed-forward neural networks, and the networks consists of an encoder, code and a decoder. The encoder compresses the N-dimensional input (a frame of sensor data) into an x-dimensional code (where $x < N$), which contains most of the information carried in the input but with less data. Hence, the encoder is similar to the principal component analysis framework, but auto-encoders can capture non-linear relationships. On the other hand, the decoder tries to regenerate the input from the lower-dimensional code or latent representation. Fig. 3 represents the fundamental single auto-encoder configuration. Each

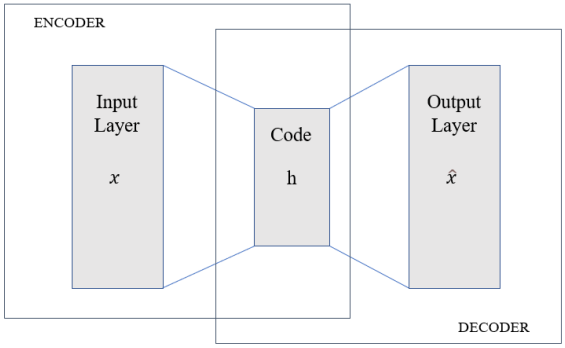

Figure 3: Auto-encoder based fault detection

auto-encoder is utilized in Eq. 17 and Eq. 18 to reconstruct the input signal(x). In Eq. 17 input(x) is multiplied by the optimal weight $z^{(1)}$ by using a back-propagation learning algorithm and added with the bias $b^{(1)}$. These two terms provide the function $h^{(1)}$ and the output variable $z^{(1)}$ is sent to the second layer of activation function, which is denoted as $h^{(2)}$ in Eq. 18. Hence, The reconstructed estimated signal $\hat{x}$ is produced with the help of the second layer of the activation function $h^{(2)}$. In general, there are many ways to perform an anomaly detection, while some methods regard the nature of the data used for the training of the algorithm. Three such possibilities are the supervised, weakly supervised and unsupervised methods.

$$z^{(1)} = h^{(1)}(w^{(1)}x + b^{(1)}) \tag{17}$$

$$\hat{x} = h^{(2)}(w^{(2)}z^{(1)} + b^{(2)}) \tag{18}$$

Unsupervised anomaly detection uses a data set without labels and statistical techniques to label samples anomalous. Under the assumption that the majority of the data set comes from the same distribution, the probability of a data sample belonging to the same distribution can be calculated using the mean and standard deviation of the data set. Supervised anomaly detection uses labelled data. This data is used to train a classifier to find a decision boundary, such that it separates anomalous and normal samples.

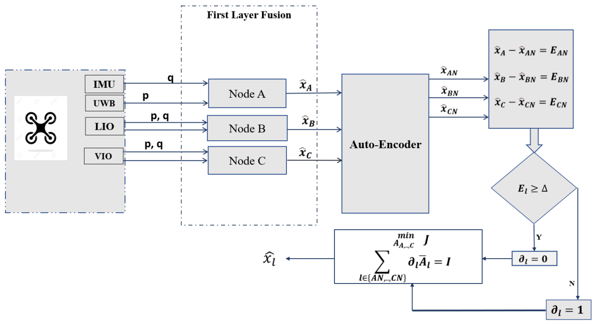

Figure 4: AE-based Decentralized Fusion

Weakly supervised anomaly detection uses a piece of information set for which imprecise labelling methods have been used. When using a weakly supervised annotation scheme, the information has been labelled using a set of rules or functions determined by the user. This means that there is no proof that the label of each sample is correct, but the labelling process can be done automatically. What data labelling scheme to use depends on the problem domain, while using a supervised labelling scheme may not be the best option for a task where the amount of labelled data is insufficient concerning the project's goals. A better option, under these circumstances, could be a weakly supervised or unsupervised scheme. In turn, in a situation where the margin for error is minimal, it would be inadvisable to use a weakly supervised labelling scheme. In this article, we demonstrate the supervised auto-encoder based detection, and isolation framework 4, where the poses coming from a Vicon motion camera system are considered as the ground truth, and based on this information the auto-encoders are trained once as depicted in Fig. 5, where the neural network of the auto-encoders has adopted the behaviours of the actual signal pattern. As a result, it will provide the exact error between the corrupted signal and the reconstructed output signal. If it is possible to find the precise deviation, then in the second layer of the decentralized fusion, corrupted estimated measurements can be easily eliminated by the logical concept.

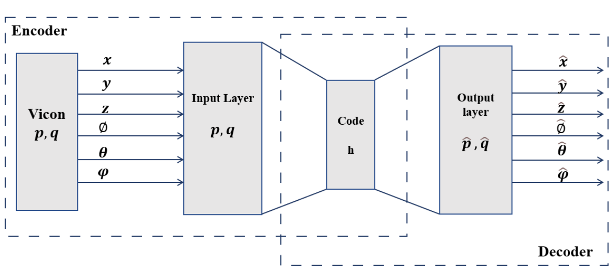

Figure 5: Train the autoencoder's three-layer structure with accurate position (p) and orientation (q)

V. Performance Evaluation

For the evaluation of the novel proposed scheme, a MAV is considered that it is manually flown in an approximate rectangular trajectory. The vehicle is equipped with a multiple pose sensor suite. The effectiveness of the Auto-encoder based detection and EKF based detection in the FROIF method are evaluated and compared using measured data collected from the sensors. To assess its effectiveness, the fused poses, from the two different proposed detection techniques, are compared with the pose from a highly accurate VICON based motion capture system, consisting of 19-cameras, which yields an accuracy in the sub-millimetre range, hence considered as the ground truth. A faulty measurement is introduced by artificially introducing a momentary fault in the measured data from Visual Inertial Odometry (VIO) in between (20-30)s. The measurements from other sensors remain unchanged. A comparative study with the EKF based multi-sensor fusion and the Auto-encoder based decentralized FROIF are presented in this article. Fig. 6 depicts the comparative clear analysis between the estimated trajectory, yields from the fusion methods as mentioned earlier.

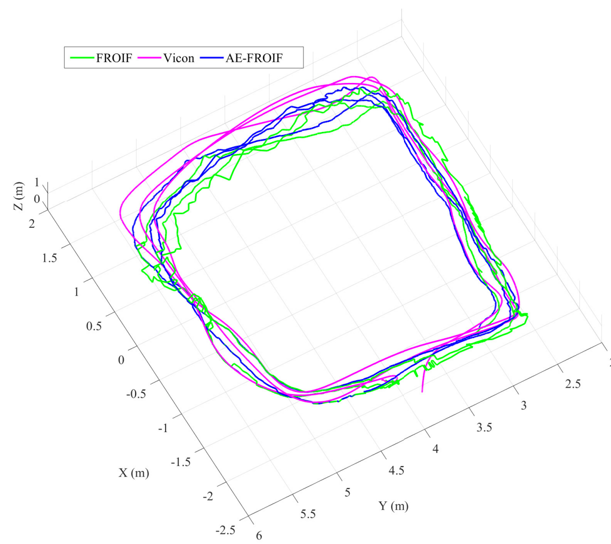

Figure 6: An estimation of the MAV trajectory based on FROIF and AE-FROIF and ground truth

Performance evaluation of Auto-encoder-based Fault Detection and Isolation

In this article, the Auto-encoder based Fault Detection and Isolation is performed with the help of on feed-forward neural network-based auto-encoder. In this experiment, each position and orientation measurement signal passes through the auto-encoder structure to reconstruct the signal through a modified back-propagation learning algorithm. The principal objective of reconstructing those signals is to recognize the outlier among samples of regular patterns. Initially, for the training of a three-layered neural network, we have used the ground truth poses. Once the training is completed, the auto-encoder circuit eventually copes with the nature of those signals that can provide the reconstructed signal almost similar to the ground truth signals. Effectively, the proposed scheme is able to identify the abnormal behaviour from the patterns. Consequently, the

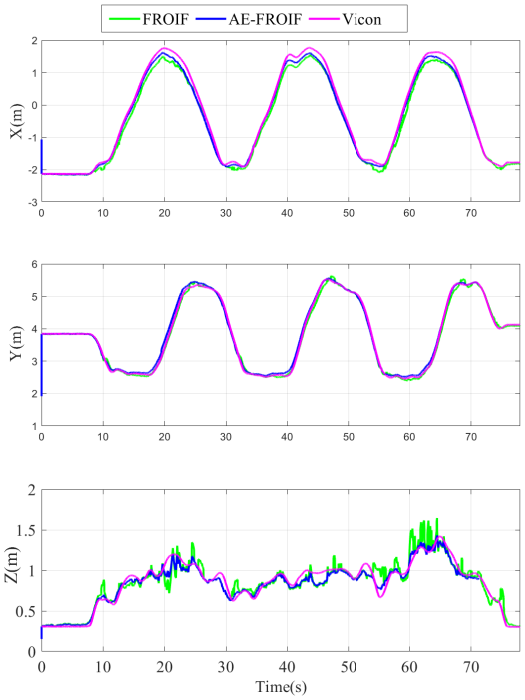

Figure 7: Variation of individual position components

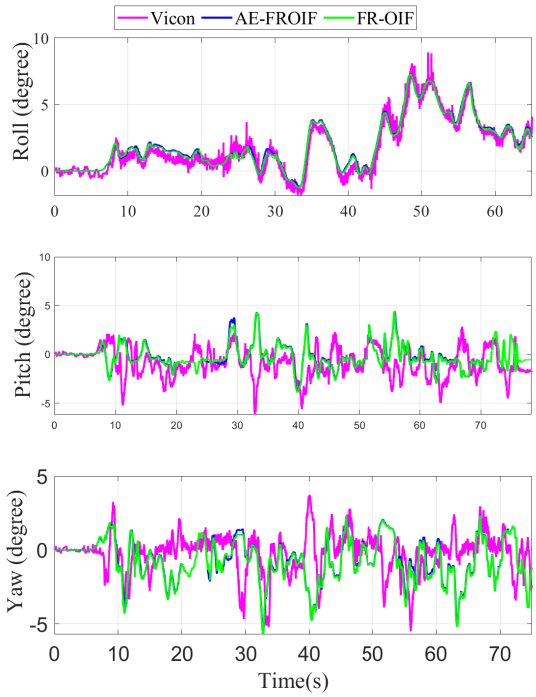

Figure 8: Variation of MAV's orientation with time

input layer signal is coded and decoded using the encoder and code layer of the neural network structure, and the decoder provides the reconstructed states $\hat{x}_N$. We have considered three nodes in the first layer of fusion, where node A and node B are the pure estimations of the information coming from LIO and VIO. Using an extended Kalman filter, the third or Node C fused the information (position from UWB and orientation from IMU). The estimated and fused states pass through the auto-encoder architecture in Fig. 4 to generate the reconstructed estimated states $\hat{x}_N$. Taking the difference between the reconstructed states coming from the auto-encoder and estimated/fused states from the first layer nodes provide the error E_l. Errors $E_l = [E_{AN}, E_{BN}, E_{BN}]^T$ are observed by the threshold Δ and passes through the logic circuit. The value of the threshold is taken as 0.4 in this experiment. Once the outlier is identified properly and observed through the logic circuit, the faulty part of the measurement is straight forward eliminated through the second layer fault resilient optimal information fusion structure. Fig. 7 and Fig. 8 show the estimated position and orientation comparison during the failure between $(20-30)$ seconds of operation period, respectively. The translation motion is more dominant than the rotational one in this experiment, due to which deviation is not evident in Fig. 10 and in comparison to Fig. 9.

Performance Evaluation of Kalman Filter based Fault Detection and Isolation

The FROIF fusion framework works in three different stages to find an exact pose of the MAV in the presence of inaccurate measurements. Firstly, multiple nodes denoted as $l = A, \ldots, G$ generate their equivalent estimate of fused states and the associated error co-variances for the whole time period. The process and measurement noise co-variance

Table I: RMSE comparison of estimated orientation in Eular Angles

Euler Angles	FROIF	AE-FROIF
Roll	0.4844	0.4998
Pitch	1.9820	2.0190
Yaw	2.0131	2.0007

Table II: RMSE for position estimation along X,Y,Z axis (in meter)

Position	FROIF	AE-FROIF
X	0.1278	0.2162
Y	0.0803	0.0695
Z	0.0620	0.1080

for the EKF associated with each node are given as $Q_{l_k} = 1000 \times I_{19\times19}$, $R_{l_k} = 10 \times I_{18\times18}$ respectively. $X_{l_0} = [0_{1\times6}, 1, 0_{1\times10}, 9.81]^T$, where as the initial error co-variance is set as $P_{l_0} = I_{19\times19} \forall l \in (A, \ldots, G)$ respectively. Secondly, the estimated state from the various nodes is processed by the auto-detection logic described in Section IV-A. The threshold associated with the detection logic is considered as $\Delta = 0.4$, which is used to construct the observation patterns from sequence (R_2, R_3, R_4). Based on the observation patterns, the faulty sensor is detected and the associated δ_{l_k} are nullified. Finally, the isolation algorithm (FROIF) eliminates the faulty

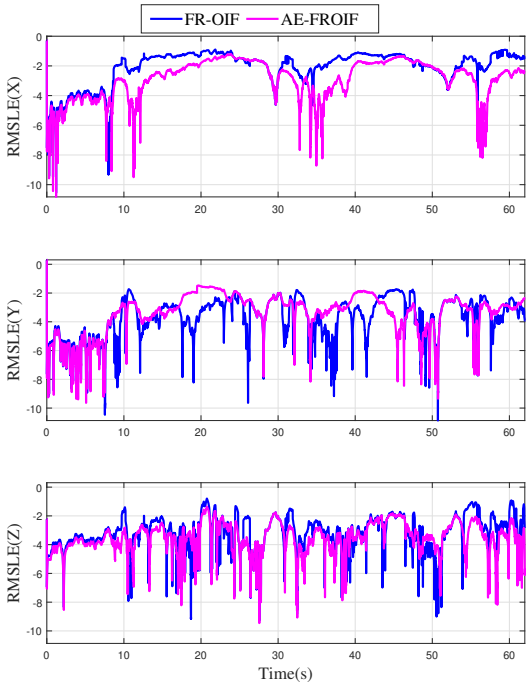

Figure 9: Variation of logarithmic sliding window RMSE for position

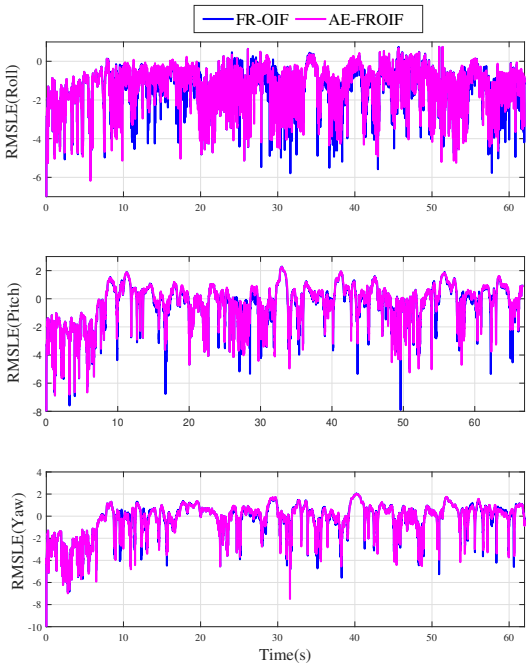

Figure 10: Variation of logarithmic sliding window RMSE for orientation

nodes for the time being and utilizes the other node's covariances for the second layer of fusion to yield a resilient pose estimation. The Fig. 9 and Fig. 10 provide the Root Mean Square Error comparison in logarithmic scale between the estimated positions and orientations, respectively. The obtained simulation results and the single value Root Mean Square Error (RMSE) for the position are depicted in Table II, while the orientation is presented in Table I that clearly show that the Kalman filter based FROIF has a reduced performance when compared to the Auto-encoder based FROIF, especially when the induced errors are of bigger time duration and more intense, while the AE-FROIF was giving better results in cases of errors during more intense rotations.

VI. Conclusion

A novel EKF and Auto-encoder based decentralized multi-sensor fault detection and isolation scheme for MAV pose estimation was presented in this article. The estimated and fused nodes are created using the extended Kalman filter to estimate the poses in the first layer. The EKF and Auto-encoder based fault detection methods are introduced between the first and the second layers to select the most accurate pose for each instant. The collected correct poses are merged into the second layer. Furthermore, both fault detection and isolation framework are performed very well with the second layer FROIF formulation that yields resiliency in the presence of inaccurate measurements.It was found that the classical OIF's performance closely resembles that of centralized EKF based multi-sensor fusion, without the use of external fault isolation mechanisms and the classical OIF's performance closely resembles that of centralized EKF based multi-sensor fusion without external fault isolation mechanisms. Therefore, these two detection methods are added significant value to accurately eliminating the erroneous measurement from the estimated pose of MAV. In conclusion, the proposed comparative study for both the FROIF framework is fairly generic. Based on simulation results, it can be observed that auto-encoder-based FROIF detection and Kalman filter-based FROIF detection are more effective in sensor failure scenarios.

References

[1] M. Liggins II, D. Hall, and J. Llinas, *Handbook of multisensor data fusion: theory and practice*. CRC press, 2017.

[2] D. Hall, C.-Y. Chong, J. Llinas, and M. Liggins II, *Distributed data fusion for network-centric operations*. Crc Press, 2017.

[3] H. F. Durrant-Whyte, B. Rao, and H. Hu, "Toward a fully decentralized architecture for multi-sensor data fusion," in *Proceedings., IEEE International Conference on Robotics and Automation*. IEEE, 1990, pp. 1331–1336.

[4] M. Mukherjee, A. Banerjee, A. Papadimitriou, S. S. Mansouri, and G. Nikolakopoulos, "A decentralized sensor fusion scheme for multi sensorial fault resilient pose estimation," *Sensors*, vol. 21, no. 24, 2021.

[5] H. Chafouk *et al.*, "Fault detection and isolation using kalman filter bank for a wind turbine generator," in *2011 19th Mediterranean Conference on Control & Automation (MED)*. IEEE, 2011, pp. 144–149.

[6] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended kalman filter for power system dynamic state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3205–3216, 2016.

[7] S. Piperakis, D. Kanoulas, N. G. Tsagarakis, and P. Trahanias, "Outlier-robust state estimation for humanoid robots," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 706–713.

[8] W. Lu and A. A. Ghorbani, "Network anomaly detection based on wavelet analysis," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1–16, 2008.
[9] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Model-based anomaly detection for power system state estimation," *Advances in Electric Power and Energy: Static State Estimation*, pp. 99–121, 2020.
[10] P. K. Chan, M. V. Mahoney, and M. H. Arshad, "A machine learning approach to anomaly detection," Tech. Rep., 2003.
[11] P. J. Rousseeuw and M. Hubert, "Anomaly detection by robust statistics," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 2, p. e1236, 2018.
[12] S. Agrawal and J. Agrawal, "Survey on anomaly detection using data mining techniques," *Procedia Computer Science*, vol. 60, pp. 708–713, 2015.
[13] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," *Special Lecture on IE*, vol. 2, no. 1, pp. 1–18, 2015.
[14] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and R. Daniela, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 5135–5142.
[15] J. Sola, "Quaternion kinematics for the error-state kalman filter," *arXiv preprint arXiv:1711.02508*, 2017.
[16] G. Yuksel and O. R. Isik, "Numerical analysis of backward-euler discretization for simplified magnetohydrodynamic flows," *Applied Mathematical Modelling*, vol. 39, no. 7, pp. 1889–1898, 2015.