Target Detection and Tracking for Unmanned Traffic Management using SWIR/MIR/EO Imagery Source– Final Report

Arik Intenam Mir[1], n9637567, Supervisor – Dr. Aaron McFadyen

Executive Summary: **This report intends to present the outcome for the detection and tracking aspect of the research project for managing unmanned aircraft traffic management using SWIR/MIR imagery. Relevant knowledge regarding the topic has been discussed and implementation techniques for target detection and tracking using temporal HMM filter are explained as well. Focus will be on the overview, implementation techniques and the results of this aspect of the project.**

I. Project Overview

Unlike the past, Unmanned Aircraft System (UAS) is a major field of interest in the aviation world now, as it carries a lot of potential for being widely used in both civil and military applications. Unmanned Aerial Vehicles (UAV) can provide an alternative, or a novel solution, to various existing aviation-based problems due to their lower cost and safety as no on-board human pilot is required for a UAV mission [1]. For an example, large scale land surveying is already being done using UAVs, where the task at hand can be done more proficiently and is observed to be financially more efficient as there are fewer issues regarding the safety and management of employees [2]. The concept of unmanned traffic management (UTM) system is new compared to that of the manned aircrafts [3]. For enhancing air traffic flow and maintaining safe operations, a well-structured traffic management is required for manned aircrafts [4]. If structured and managed properly, an unmanned traffic management system can provide strategic solutions for operating UAVs in the same airspace with manned aircrafts while minimizing the probability of collision risk [5]. With the focus on all the advancements and research done regarding unmanned traffic maangement, the final aim of this project will be to verify the usability of SWIR/MIR imagery in aiding unmanned traffic management(UTM).

The entire project can be broken down into three major parts- (a) Image pre-processing, (b) Detection & tracking and (c) Encounter generation. The completion of this project is subject to the successful implementation and integration among the parts.

The focus of this project is to process the raw imagery data obtained from the SWIR/MIR source to reduce the background clutter and improve detection and tracking accuracy for the detection-and-tracking stage of the project. This part is proposed to be done via a temporal filtering stage [7]. Firstly, the output from this stage will be to display the video output of the input data with a graphical overlay representing successful detection of targets and tracking. Secondly, the information obtained from the detection-and-tracking stage, will be utilized in the encounter generation stage, where hybrid encounter imagery will be generated for a range of different encounters. Which will lead to determining a set of avoidance waypoints in the 3D space. A path search, or traversal algorithm, will be utilized here for generating a new collision avoidance path. To better demonstrate how the project is supposed to be completed by implementing the different stages, a block diagram was created, which can be viewed in the Appendix Section -C.

When finished, the entire system will be evaluated for determining the compatibility of SWIR/MIR imagery in the context of UTM.

This report is structured as follows. Section II provides background on the relevant literature for this project. Section III outlines the specifications for HMM parameterization and filtering, section IV presents and discusses the results obtained. Finally, conclusions are presented

II. Literature Review

Detecting and tracking objects in the airspace is always extremely challenging, as the objects are not stationary, and the size appears relatively smaller [6]. Doing it successfully is one of the major obstacles to overcome in order to potentially integrate UAVs in the civil airspace, in future [11]. A two-stage processing paradigm has been explored extensively for detecting and tracking of sub-pixel sized targets from imagery datasets [11], namely - 1) Image pre-processing stage, and 2) Temporal filtering stage. In this paper, focus will only be in the temporal filtering stage, which exploits target dynamics across image frames to detect and track [8,9].

The temporal filtering stage is often based on a track-before-detect concept, where the main objective is to highlight and extract temporal features, which exhibits behaviour like that of a target across image sequences [9]. This stage can be approached in various techniques. This project takes the Viterbi-based approach because of its feasibility in the context of detecting and tracking targets, where under a number of predefined assumptions, the Viterbi algorithm can successfully determine the best track for the target given the dataset [8-11].

[1] Name is with Science and Engineering Faculty (SEF), Queensland University of Technology (QUT), Australia. This report is in partial fulfilment of EGH400-1 and/or EGH400-2 unit assessment requirements and **submitted on 18/06/2020**

arikintenam.mir@connect.edu.au

First proposed in 1967 as a method for solving estimation problems [12], the goal of the Viterbi algorithm is to generate the sequence of likely states that best explains a sequence of observations. It is an excellent dynamic programming algorithm which has been used for various digital estimation problems, as it is capable of providing an efficient method for finding the most likely state sequence for a finite-state, discrete-time Markov process and hidden Markov models [11, 14].For the tracking and detection aspect of the project, it was decided to take a Viterbi-based Hidden Markov Model filtering approach by implementing a single HMM-based filter as the temporal processing stage[8,9,15].This filtering approach was observed to offer excellent results in detection performance. [13, 15]. HMM parameterization and filtering regarding the target detection and tracking are discussed in the following section.

III. HMM FILTERING AND IMPLEMENTATION

A standard HMM filter requires two model parameters – *states(N)* and *observations(M)* and specifications of three probability measures – *transition matrix(A), observational probability matrix(B)* and *initial probabilities(π)* [10]. Which can be defined together as:

$$\lambda = (A, B, \pi) \tag{1}$$

Thus, the complete parameterization of an HMM is:

$$HMM: \{N, M, A, B, \pi\} \tag{2}$$

In the context of detection and tracking, a first-order discrete-time, discrete state Markov chain is utilized for modelling the projected motion of a target on the 2D image [8,9].

To operate with the HMM filter optimally, the pre-processed image frames are further processed to generate binary images using a threshold for pixel intensity. Setting up the threshold value to 70, pixels with intensity lower than the threshold are considered as zero elements and assigned a value of 0, while values larger than the threshold are considered as non-zero elements and a assigned value of 1. The pixels with non-zero elements are considered as the target pixels.

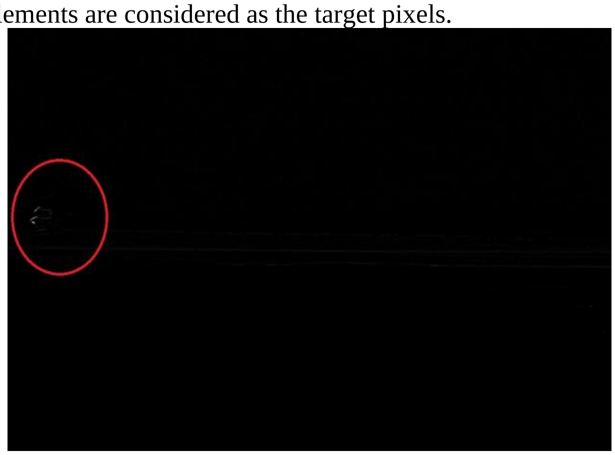


Figure. 1. Target (inside red circle) in the background after preprocessing stage

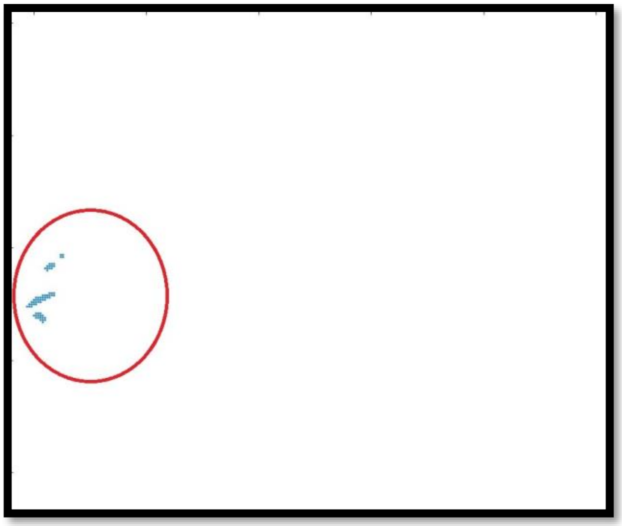


Figure. 2. After processing the image frame with the predefined threshold value of 70 to only extract the target pixels, which are above the threshold (inside red circle)

Figure 1 and 2 illustrates the presence of the target pixels, before and after further processing an image frame with the binary threshold respectively.

Assuming the Markov model describes a target located at pixel location (state) *'m,* which is relatively stationary and does not transition to distant states, it becomes redundant to consider all of the possibilities in the transition probability matrix and the observational probability matrix, that are far away from the current state *m.* Therefore, only a 9×9 matrix is taken around the current state for the HMM filter while the rest are negated. In the new 9×9 matrix, the current state *m* is placed at the center location of the matrix (5,5), and HMM parameterization and filtering are done on this 9×9 matrix.

A. States(N)

Under the assumption, the target is visible in the image plane, it can be represented by a fixed 2D plane. The 2D plane consists of discrete grid points and it can be represented by $\{(i,j)|1 \leq i \leq N_v, 1 \leq i \leq N_h\}$, where N_v and N_h are the vertical and horizontal resolution of the image plane. Thus, the total number of states/pixel locations can be denoted by $N = N_v \times N_h$[9,10]. Each pixel location {i, j} will be representing a unique state for the HMM model. A target indicator vector L_x is declared, by putting together every column of the 2D image plane to form a single column vector, which is ideally convenient for modelling the HMM temporal filter [9].

B. Observations(M)

Observations for an HMM generally means a sequence of obseravble symbols from each state[10]. For the temporal HMM filtering, the observations for the model are the pre-processed image sequences.

$$O = \{O_1, O_2, O_3, \ldots\ldots O_x\}, \quad (3)$$
for time steps till t = T_x

C. *Transition probability Matrix(A)*

In scope of detecting and tracking target, the HMM's *transition probability matrix A*, models the target dynamics by defining the likelihood of the target staying at the same pixel location or moving to different states. The *A* matrix for any given states *"m"* and *"n"* can be represented as following:

$$A^{mn} = P(L_x = m \,|L_{x-1} = n), \quad (4)$$
for $1 \le m \le N$ and $1 \le n \le N$

Each element A^{mn} of the matrix *A*, will define the probability of the target moving from the state *m* to the state *n*. Every element in this matrix is assigned a probability, where
For self-transition, the probability is 7/15 and for transitioning to any adjacent pixels, it is 1/15 [9]. Zero probabilities are assigned to all non-adjacent states.

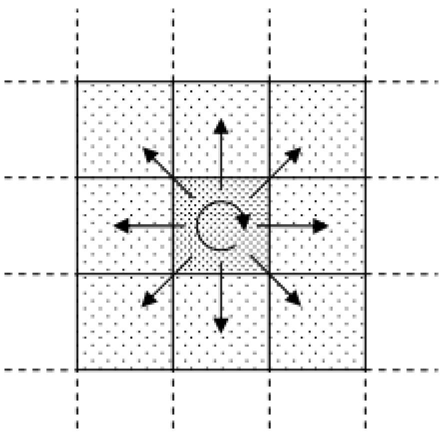


Figure. 3. State transitions possible for a target in an HMM filter [9].

Here, Figure 3 is representing the transitions possible for the target to the 8 adjacent states. As the *transition probability matrix A* is constructed to only allow a self-transition or a transition to one of the adjacent states from time step *x* to *x+1*, it is possible to create the matrix as a diagonal sparse matrix[8, 9]. Figure 4 visualizes the *transition matrix A* of the HMM filter.

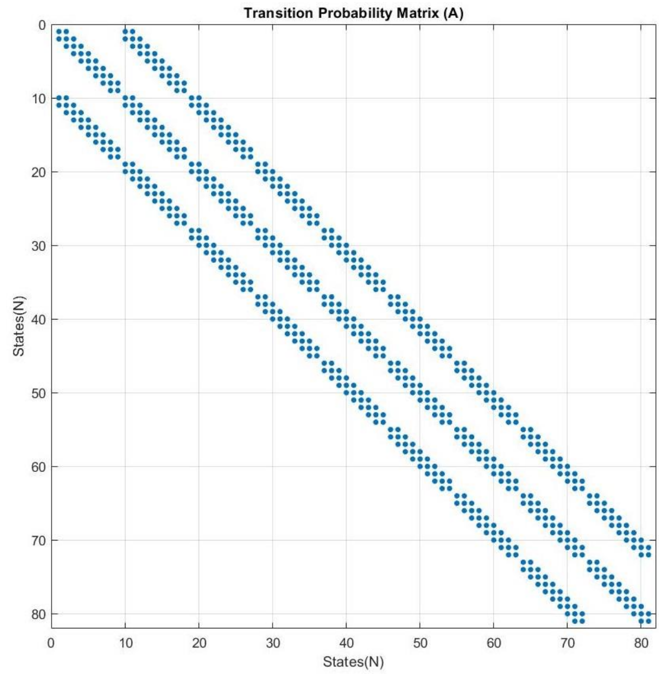


Figure. 4. Transition matrix A for the HMM filter. Dimension of the matrix is N× **N**, where N = 81.

D. *Observational Probability Matrix(B)*

According to standard HMM parameterization, *observational probability matrix (B)* defines the probabilities of obtaining a sequence of observations (*M*) given a set of states (*N*) [10]. In the detection-and-tracking application, for a target being present at a pixel location(state) *m* within the 2D image plane, its probabilistic relationship with the Observations O_x is denoted by the *observational probability matrix* $B^m(O_x)$, which is of the shape $N \times N$ and a sparse matrix like the *transition probability matrix A* [8].

$$B^m(O_x) = \frac{P(O_x^m|L_x=m)}{P(O_x^m|L_x \ne m)}, \quad (5)$$
for $1 \le m \le N$

Here in (5), for constructing $B^m(O_x)$, the elements of the matrix are calculated by taking the ratio of probability of the target being present and absent in state *m*[7-9]. The probability of the target being present $P(O_x^m|L_x = m)$, is calculated by taking division of the probability of the target residing at the state *m* and the total number of observations.

$$P(O_x^m|L_x = m) = \frac{P(L_x=m)}{\sum_{n=1}^{n=x} O_n} \quad (6)$$
for $1 \le m \le N$

while for $P(O_x^m|L_x \ne m)$, every non-location of the target, can be given as

$$P(O_x^m|L_x = m) = \frac{P(L_x=m)}{\sum_{n=1}^{n=x} O_n} \quad (7)$$
for $1 \le m \le N$

Illustration of the implemented *observational probability matrix B* is given below in Figure 5.

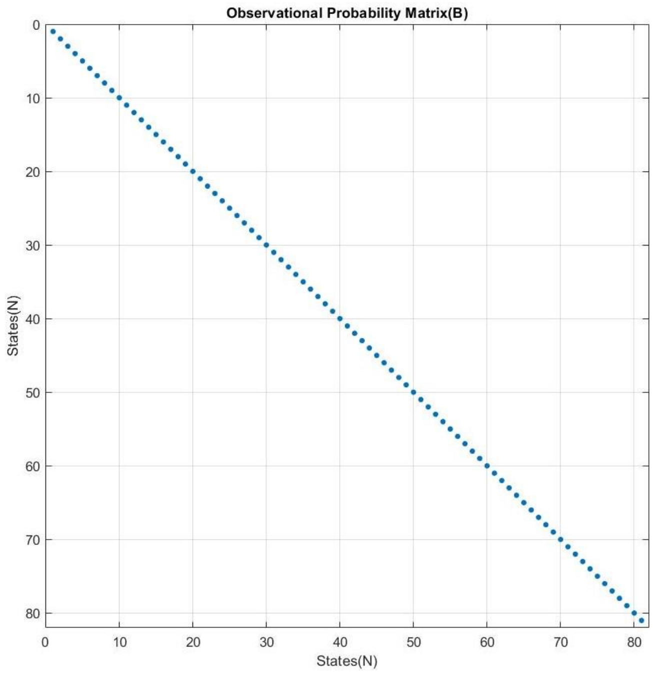


Figure. 5. Observational probability matrix B for the HMM filter. It is of size N× N, where N = 81.

E. Initial Probability matrix(π)

In the case of a given state, the vector *Initial probability matrix (π)* will define the initial probability of the target located at a certain pixel location(state). For a state *m*, the initial probability π^m can be denoted as:

$$\pi^{m} = P(l_1 = m), \quad \text{for } 1 \leq m \leq N \tag{8}$$

Like the *transition probability matrix A* and the *observational probability matrix B, matrix π,* is also constructed as a sparse matrix, containing the initial probabilities for every pixel location of the target in the image frame, where its shape is 1/N.

F. Detection Mechanism

When the parameterization of the temporal HMM filter is complete, the Viterbi Algorithm uses it to predict the likely states for a target over an input imagery data. Implementation of the temporal HMM filter and the operations of the Viterbi algorithm was completed using MATLAB. MATLAB's built-in function *hmmviterbi (sequences, transition probability matrix, observational probability matrix)* is used for generating the estimated states for the target across a given set of image sequences/observations from the implemented HMM filter [16]. The estimated states will be used for two purposes. Firstly, the pixel locations of the estimated states will be generated as a CSV file, so it can be used for the encounter generation stage and secondly, to display a graphical overlay on the actual image sequences to illustrate detection and tracking

IV. RESULTS

In this section, the outcome of the detection and tracking of target using a temporal HMM filter is discussed extensively.
For the evaluation of the performance of the HMM filter for detecting and tracking targets, two basic concepts of evaluation are discussed. Firstly, by comparing the trend of the estimated states and the actual states to determine the competency of HMM filter to detect and track targets from image sequences. Secondly, by making a visual comparison for a head-on collision encounter to explain if the estimated state appears to be at the same pixel location as the pixel location in the actual image sequences.

1. HMM Filter operations on Sample EO Image Sequences

The likely states obtained using the HMM filter and the Viterbi, for a target aircraft across a sample set of image sequences, obtained from an EO(Electro-optical) source are presented here.
The sample set consists of 9 image sequences/observations. Using the HMM parameterization and filtering technique discussed in section III, the image sequences were first processed and parameterized for the filter and then the likely states were predicted for the target using the Viterbi algorithm

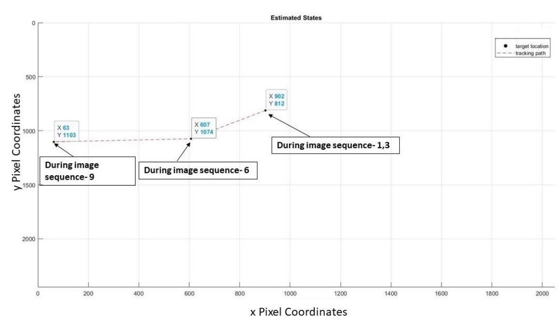
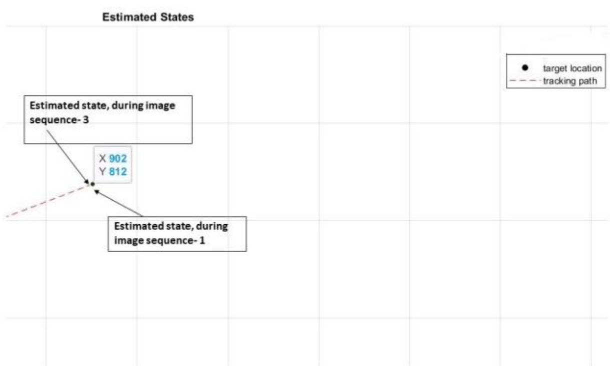


Figure. 6. Complete estimated track for the target across the set of sample image sequences. (---, tracking path), (●, target location).

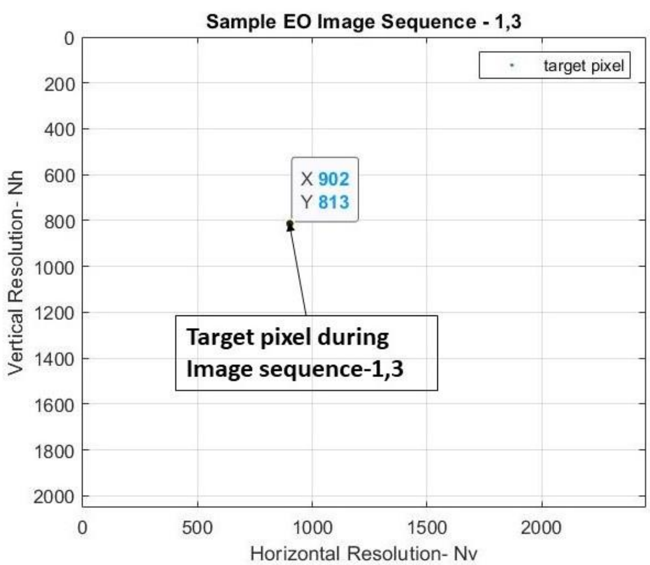
Please refer to the Appendix, Section -B to view a better and larger representation of Figure 6.
For demonstration purposes, the transition of the target across the set of image sequences were represented at an interval of every 3 image sequences. In figure 6, the target's position in the image plane were shown at image sequence 1, 3, 6 and 9 respectively.
Under the hypothesis, the estimated track of the target aircraft obtained through Viterbi algorithm should follow the trend of the target aircraft in the actual EO image sequences,

comparison has been done on the basis of pixel coordinates of the target across for EO image sequences and the estimated track.

To better explain and visualize the hypothesis, parts of figure 6 are zoomed in and compared with the EO image sequences below and the results of the comparison are presented in Table I.

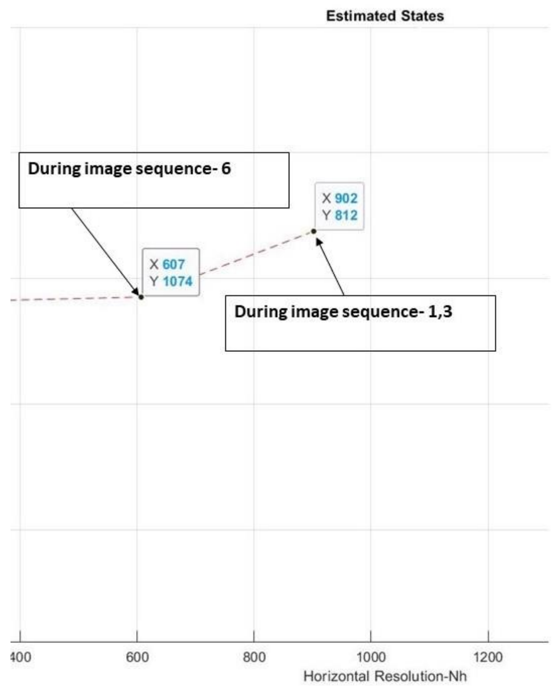


(a)

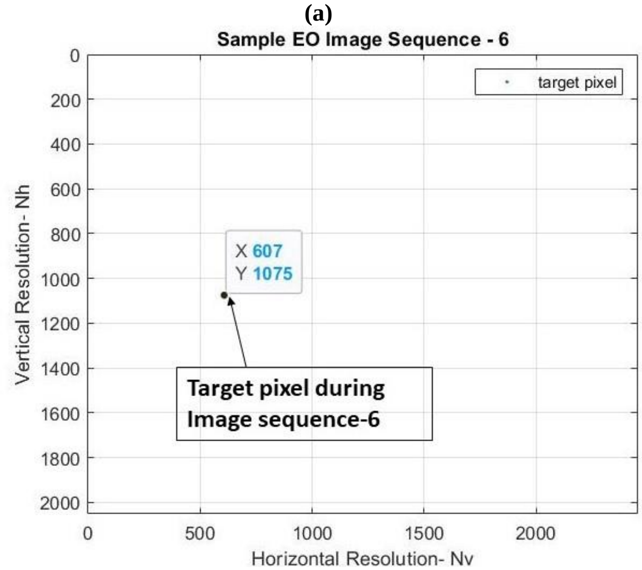


(b)

Figure. 7. (a)Shows the estimated pixel location(902,812) of the target across during its stay on image sequences 1 and 3. (b) Actual pixel location of the target (902,813) in the sample EO image sequences 1 and 3. As the target stays stationary, during transition between sequences 1 to 3, they are plotted in the same figure.

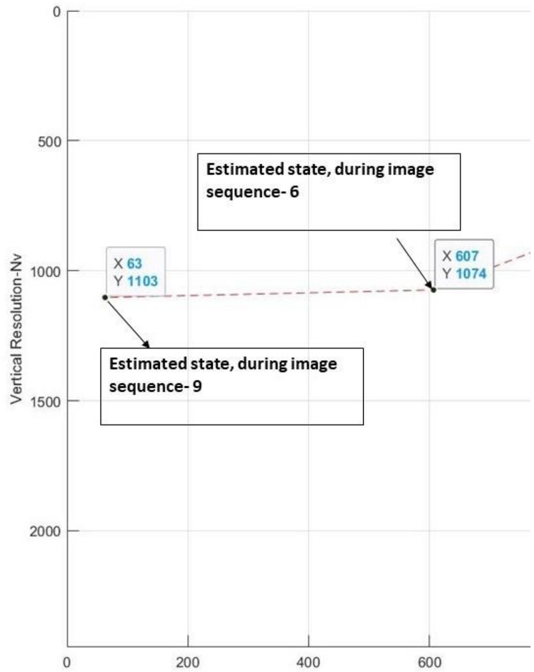


(a)

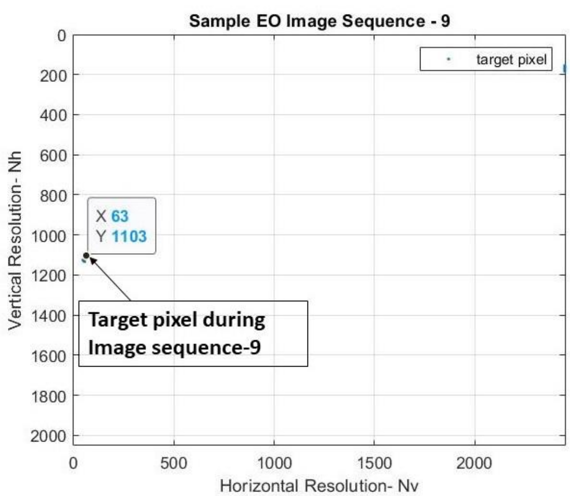


(b)

Figure. 8. (a)Shows the estimated pixel location (607,1074) of the target across during image sequence 6. (b) Actual pixel location of the target (607,1075) in the sample EO image sequence 6.



(a)



(b)

Figure. 9. (a)Shows the estimated pixel location (63,1103) of the target across during image sequence 9. (b) Actual pixel location of the target (63,1103) in the sample EO image sequence 6.

TABLE I
COMPARISON OF ESTIMATED TRACK AND ACTUAL EO OBSERVATIONS, BASED ON PIXEL COORDINATES

Image Sequence	**Pixel coordinates - Estimated Track (x,v)**	**Pixel coordinates- Trend across actual EO observations (x,y)**
Sequence- 1,3	902,812	902,813
Sequence - 6	607,1074	607,1075
Sequence -9	63,1103	63,1103

Looking at the pixel coordinates for each cases of image sequences across both the estimated track and the actual EO image sequences from Table I, it can be stated that the estimated track follows the trend of the aircraft in the sample EO image sequences. As there were only 9 sample image sequences available, prediction of the estimated states for the target had to be halted at the image sequence 9. But it can be assumed, if more image sequences were available, the Viterbi algorithm would have been able to estimate target states for more sequences [10].

Even though the implementation for the detection and tracking aspect of this project did not align with the timeline(Appendix, section-A), it was still possible to obtain a satisfactory outcome.

2. *HMM Filter Operations on Head-on Collision Encounter*

Output obtained from the Encounter Generation stage for collision encounter, were processed using the temporal HMM filter and the Viterbi algorithm to obtain the estimated states for a target. 450 image sequences were generated for the collision encounter. For ease of testing with the HMM filter, it was assumed that the target is located at the very center pixel location across the total number of image sequences.

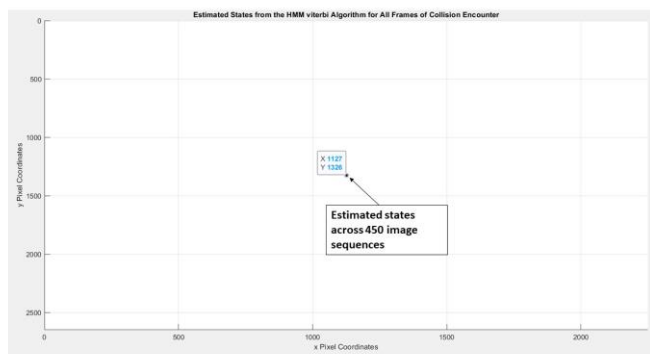


Figure. 10. Estimated state(1127,1326) obtained from the Viterbi operation on the HMM filter for head-on collision encounter

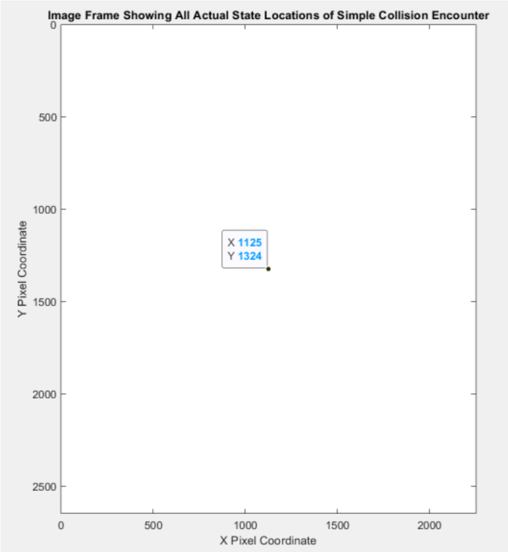


Figure. 11. Actual state(1125,1324) of the target in head-on collision encounter

After comparing the coordinates of the estimated and the actual state of the target from figure 10 and 11 respectively, it can be stated that the final estimated state is very close to the actual state. To better understand the performance of the HMM filter for this scenario, the actual states of the target across the image sequences are compared in figure 12 and a zoomed version of figure 12, is given in figure 13.

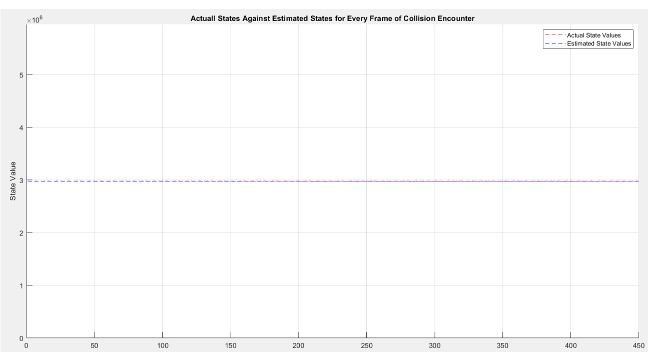


Figure. 12. Actual states of the target compared to the estimated states across every image sequence for collision encounter

It can be seen by observing the plot in figure 12, that the estimated states of the collision encounter follow the trend of the actual states, presenting a fairly correct prediction of states.

Please refer to the Appendix, Section -C to view a better and larger representation of Figure 12.

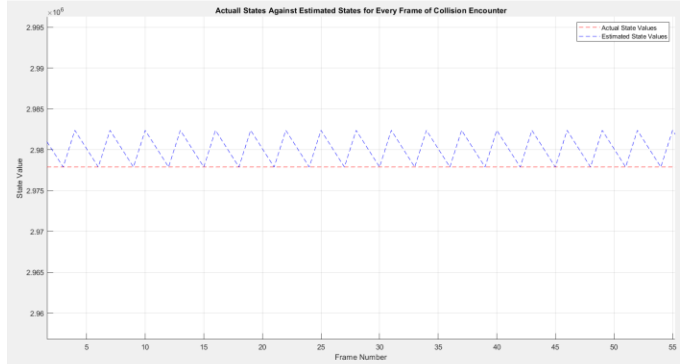


Figure. 13. Actual states of the target compared to the estimated states across every image sequence for collision encounter(zoomed)

From the zoomed plot of the comparison between the actual and the estimated states, it can be stated that the estimated states are deviating slightly from the actual states, in an even pattern. These repeated patterns are occurring because the estimated states are update every 3 image sequences.

Based on the limited output of the two operations of the HMM filter and the Viterbi algorithm discussed above, it can be stated, that detecting and tracking of target is possible, when the input imagery data is coming from a EO source and is viable in aiding UTM. Although a sophisticated evaluation metric is required to be implemented to ensure the best performance of the HMM filter for detecting and tracking targets in similar scenarios.

More tests and evaluations are required to be conducted on various encounter scenarios and situations to ensure the feasibility of using EO sources when it comes to aiding in UTM.

V. CONCLUSION

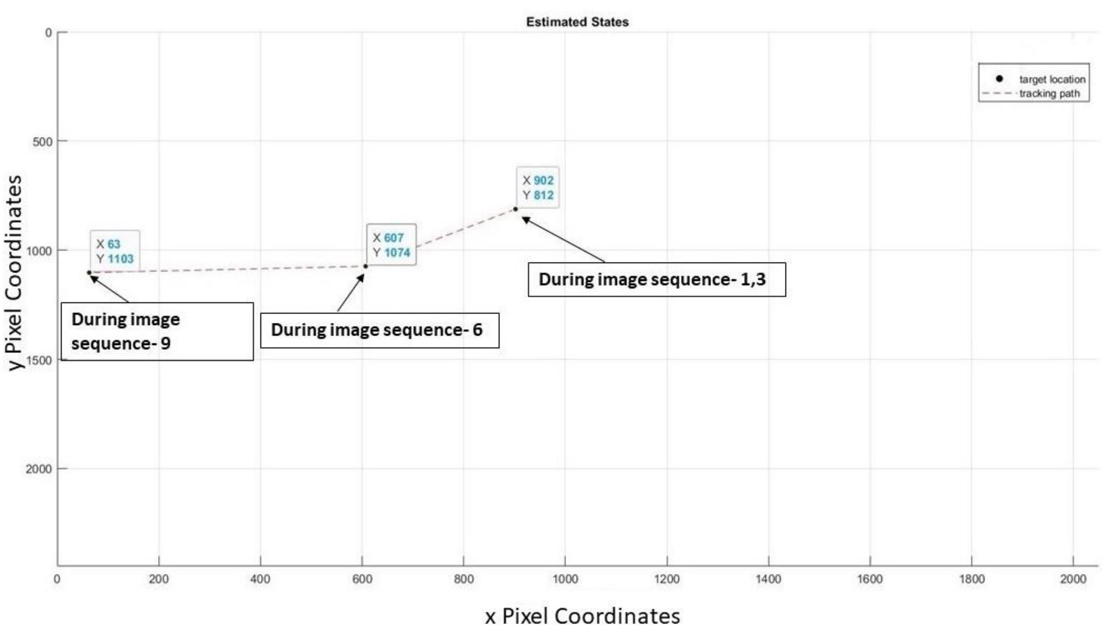
Due to the limited availability of genuine data to work with, it was not possible to fully test and evaluate the performance of the single temporal HMM filter for detecting and tracking targets, in a better manner. Because of the time constraints and other limitations, it was not possible to implement the HMM filter bank. Although, the performance of the implemented single HMM filter is sufficient enough for explaining the detection and tracking of targets in the airspace, an HMM filter bank would provide better results.

APPENDIX

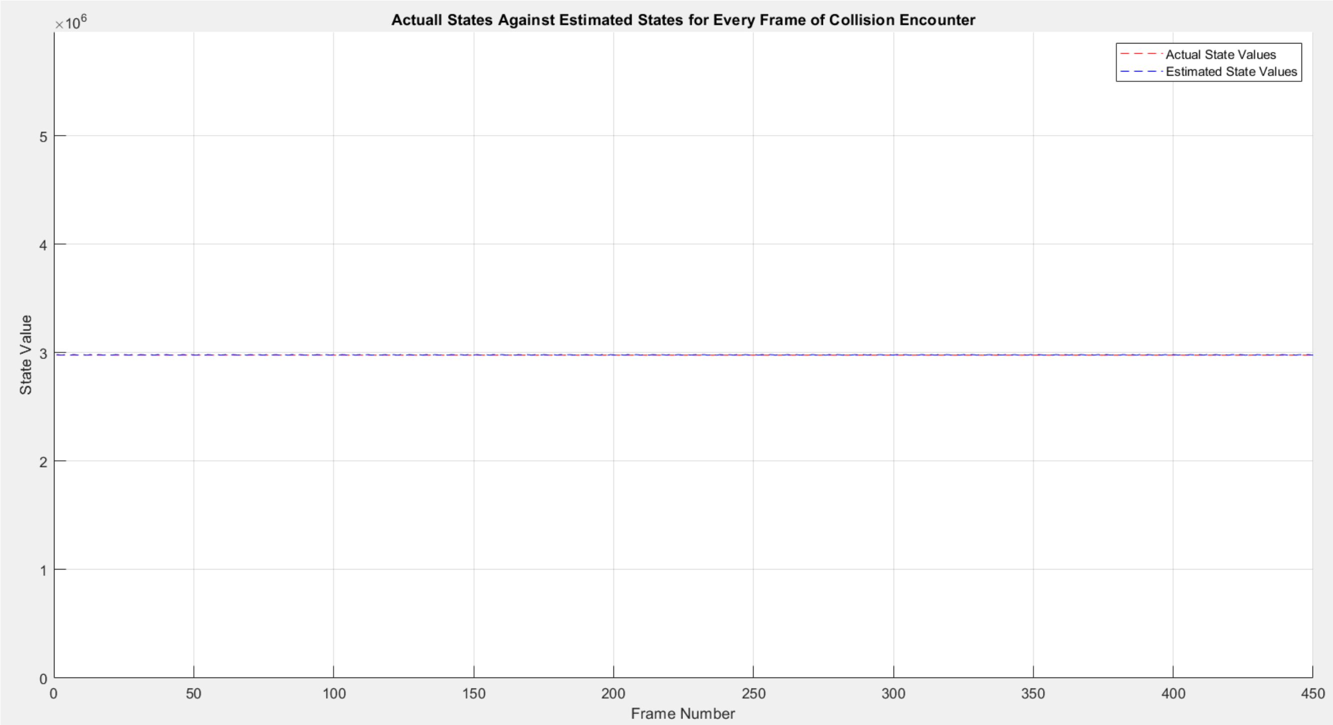
A. *Project Timeline Gannt Chart*

Task	Week	Date started	Date ended	01	02	03	04	05	06	07	08	09	10	11	12	13
SEMESTER- 01																
1A	Background Review of project	09/03/20	30/03/20			■	■	■								
1B	Research and Planning	02/03/20	07/04/20		■	■	■	■	■							
1C	Topic investigation report	09/03/20	23/03/20			■	■									
1E	Implementation of Detection & tracking Mechanism	01/04/20	05/05/20						■	■	■	■	■			
1F	Evaluation of Implementation	22/05/20	06/05/20									■	■			
SEMESTER- 02																
2A	Single HMM filter implementation	13/07/20	09/08/20	■	■	■	■									
2C	HMM filter bank implementation	02/08/20	06/09/20				■	■	■	■	■	■				
2D	Integration & testing	24/08/20	06/09/20									■				
2E	Evaluation & further testing	30/08/20	06/09/20										■			
2F	Final Integration & testing	09/10/20	23/10/20											■	■	■
2G	System Performance Evaluation	06/09/20	23/10/20												■	■

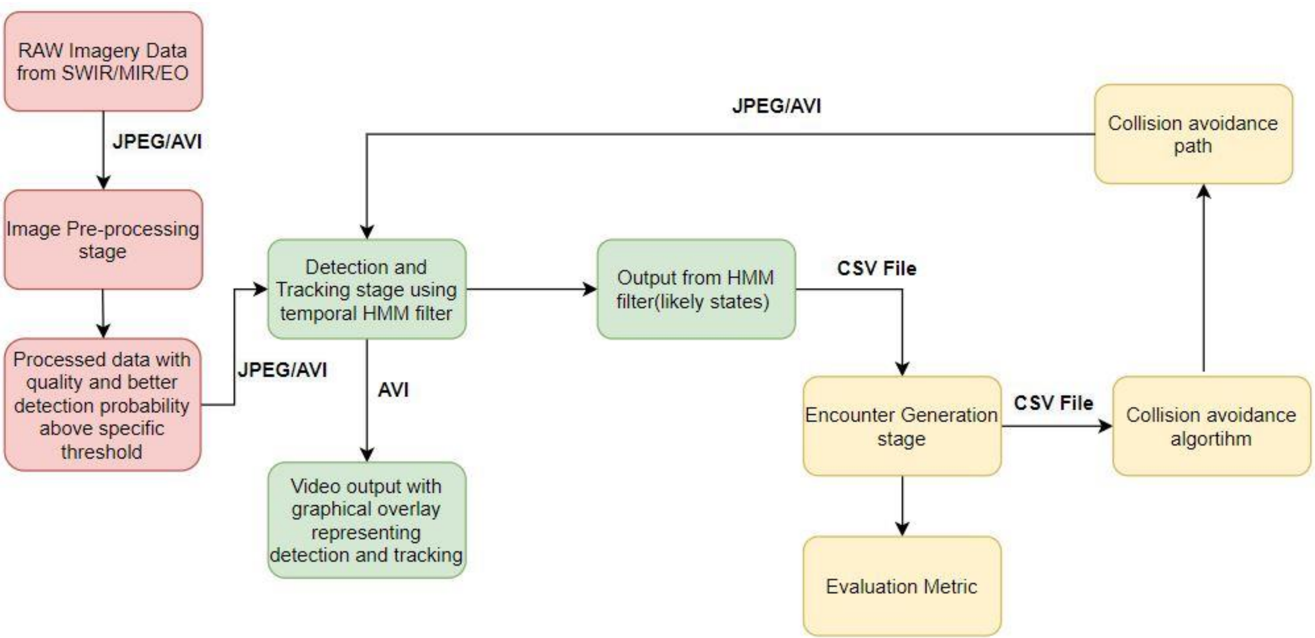
B. Estimated track of target across image sequences



C. Comparison of Actual States vs Estimated States for Collison Encounter



D. Block Diagram of the Project



VI. References

[1] H. Pham, S. A. Smolka, S. D. Stoller, D. Phan and J. Yang, "A survey on unmanned aerial vehicles collision avoidance systems," Dept. of Computer Science, Stony Brook University, NY, 2015.

[2] S. 3D, "Why UAVs Are Ideal for Large-Scale Surveys," 2016.

[3] P. Kopardekar, "“Safely Enabling Low-Altitude Airspace Operations: Unmanned Aerial System Traffic Management(UTM)," Technical Report, National Aeronautics and Space Administration (NASA), 2015.

[4] B. Alexander, "Aircraft density and midair collision," in Proceedings of the IEEE, 1970.

[5] Eurocontrol, “RPAS ATM Conops,” Technical Report, Feb. 2016*Computing: Techniques and Applications*, Noosa, QLD, 2011.

[6] O. Yu and G. Medioni, "Motion Pattern Interpretation and Detection for Tracking Moving Vehicles in Airborne Video," Institute for Robotoics and Intelligent Systems, University of Southern California, Los Angeles, CA, 2009.

[7] J. Lai, J. Ford, M. Luis and O. Peter, "Vision-Based Estimation of Airborne Target Pseudobearing Rate using Hidden Markov Model Filters," in *IEEE Transactions on Aerospace and Electronic Systems*, 2013.

[8] J. Lai, L. Mejias and W. R, "Detection versus False Alarm Characterisation of a Vision-Based Airborne Dim-Target Collision Detection System," in 2011 International Conference on Digital Image Computing: Techniques and Applications, Noosa, QLD, 2011.

[9] J. Lai, L. Mejias, P. O'Shea and W. R , "Hidden Markov Model Filter Banks for Dim Target Detection from Image Sequences," in 2008 Digital Image Computing: Techniques and Applications, Canberra,ACT, 2008.

[10] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," in IEEE, 1989.

[11] G. D. Forney, Jr, "The Viterbi Algorithm," IEEE, vol. 61, no. 3, pp. 268-278, 1973.

[12] M.T. DeGarmo. “Issues concerning integration of unmanned aerial vehicles in civil airspace,” The MITRE Corporation, Tech. Rep. MP 04W0000323, 2004.

[13] S. J. Davey, M.G. Rutten, and B. Cheung, “A comparison of detection performance for several track-before detect algorithms,” *EURASIP Journal on Advances in Signal Processing*, vol. 2008, pp. 1-10, 2008.

[14]] D. Neuhoff, "The Viterbi algorithm as an aid in text recognition (Corresp.)," in IEEE Transactions on Information Theory, vol. 21, no. 2, pp. 222-226, March 1975, doi: 10.1109/TIT.1975.1055355.

[15] S. Tuğaç and M. Efe, "Hidden Markov Model based target detection," 2010 13th International Conference on Information Fusion, Edinburgh, 2010, pp. 1-7, doi: 10.1109/ICIF.2010.5711878.

[16] Mathworks, "hmmviterbi," Mathworks, https://au.mathworks.com/help/stats/hmmviterbi.html (Accessed Oct. 30, 2020)