KLT Based Predictive Aircraft Tracking in Low Resolution Images

By

Khawar Ali

2011-NUST-MS-PhD- ComE-14

MS-69 (CE)

Submitted to the Department of Computer Engineering in fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

COMPUTER ENGINEERING

Thesis Supervisor

Dr. Shoab Ahmad Khan

College of Electrical & Mechanical Engineering

National University of Sciences & Technology

2013

DECLARATION

I hereby declare that this thesis has been built on my personal efforts under the genuine supervision of my supervisor Dr. Shoab A. Khan. No data of this thesis has been shared as part of any research work presented in any other institute for the fulfillment of degree requirement.

____

Student’s Signature

ACKNOWLEDGEMENT

First of all, I am grateful to Almighty Allah who gave me the ability and courage to complete this task. He is the only one I always looked to in the event of happiness and trouble and He always helped me in the time of need. I would not be what I am today if He did not want me to be. With His sympathy I have been able to complete my work.

I would like to express my deep-felt gratitude to my advisor, Dr. Shoab Ahmad Khan, Head of Computer Engineering Department for his valuable advice, great support, positive encouragement and enduring patience.

I gratefully acknowledge my committee members Dr. Aasia Khanum, Dr. Saad Rehman, Dr. Arsalan Shaukat, and Dr. Usman Akram for their support and guidelines for the completion of my thesis.

Abstract

Tracking of a particular object in an image using the feature tracking algorithms faces significant challenges under conditions of severe aircraft rotation, highly cluttered background presence, arrival and collision with other aircrafts and sun, excessive noise, and varying lightening conditions due to weather changes. In this research, a real-time and robust algorithm is presented for the tracking of an aircraft in video sequences of low resolution with the use of Kanade-Lucas-Tomasi (KLT) feature tracker, random sample consensus (RANSAC) algorithm, texture modeling and matching techniques. Features of an aircraft are tracked using KLT feature tracker while developing aircraft model at each frame and finding transition matrix with the use of each pair of consecutive frames. Features of aircrafts are excluded when not following the transition of the aircraft and new features are introduced if they are within the bounding range of the aircraft and match with the current aircraft model. The varying weather conditions and noise reduction are handled in the algorithm without the use of averaging filters to increase the overall efficiency of the framework.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

A comprehensive introduction of the thesis based on the common terms and basic concepts about the research work done is discussed in this chapter.

The aircraft tracking systems mounted on the guided missiles are of great importance in fighting planes. The utilization of infra-red sensors in the past has proved that there are certain limitations of their use which cannot be improved with the improvement in technology. This has led to the basis of the utilization of cameras and computer vision plus image processing techniques to tackle similar problems. An algorithm has been developed for this purpose which has been explained in detail throughout this thesis report.

1.1 INTRODUCTION TO THESIS

The thesis intends to deal with the problems associated with the use of camera and image processing techniques in guided missiles. Missile guidance basically refers to techniques, methods, algorithms, and processes associated with guiding a missile to a specific targeted object which in most cases is an opponent aircraft. The use of computer vision techniques for tracking and following a particular object in the flight path of a mounted camera based guided missile is the main topic of investigation in the paper. [6][7]

Many different types of sensing methodologies and equipments have been in utilization in the past for following tracking trajectory of an object, but all the techniques have displayed a large number of disadvantages and negative impacts associated with them. The failure to reach equilibrium between efficiency and accuracy has a convincing argument of replacing the current methods by attempting other techniques and technologies. This thesis focuses on the use of camera and utilizing image processing and computer vision techniques to deal with the tracking of a rotating and fast-paced moving aircraft in a quickly varying and cluttered background environment with high efficiency and accuracy [8][9].

The thesis focuses on the Kanade-Lucas-Tomasi algorithm for the major part of the tracking by making modifications and attaching algorithms and techniques to raise the accuracy while keeping a high efficiency percentage. Apart from detection and tracking, another focus is to predict the future location of the aircraft in a few frames time based on the transition matrix and angular velocity measurements to estimate the exact possible locations if the sight of the aircraft is lost.

An application of the Kanade-Lucas-Tomasi algorithm is presented with the handling of the challenges associated with the problem. The KLT and other optical flow algorithms like SIFT and SURF face significant problems in object tracking when the tracked object goes through excessive and quick rotations and scaling. In this thesis, a robust and real-time algorithm is provided with high accuracy is presented handling low resolution video sequences with the sequential use of a combination of methods including KLT feature tracker, random sample consensus, texture modeling and matching. Features of a aircraft are tracked using KLT in the first frame and the aircraft model is developed, the tracked feature in the next images and the arrival of the new features matching with the aircraft model present the transition matrix using RANSAC algorithm defining the angular velocity and the translation, rotation, and scaling of the jet by which a new model of the aircraft is developed on each presiding frame.

1.2 MOTIVATION

There are different types of sensors that are implemented in the missile guidance systems which have their own techniques and trajectories that they follow to track or target the object in attention. There are different types of guidance options including inertial guidance, preset guidance, celestial guidance, and terrestrial guidance. The most important technique used in all of these guided systems is the utilization of infra-red based sensors. Though, these sensors have been used in a large number of missile systems in the recent past to present, they have a set of disadvantage associated with their use.

Some of the major issues involved with IR sensors include its lesser accuracy, its problems associated with direct diversion towards sunlight, the possibility for the pilots to deceive the intelligence of these systems, and the possibility of a very narrow beam. The accuracy of infrared sensors decreases to a great extent when they are exposed to detection or tracking of dark colored objects. As discussed earlier, the IR sensors have a narrow beam due to which they work only till a distance of 5 feet. Another alternative is the use of an ultrasonic sensor which works to a distance of around 30 feet. [1] Cameras are able to detect and track objects at much higher distances which is a huge plus point regarding their use.

Along with the major problems associated with the use of infrared sensors, the use of camera for guidance systems also involves certain challenges which need to be managed and catered. The most common tracking method in computer vision currently is the use of optical flow algorithm, but quick rotations have an immense effect on the accuracy of detection and tracking while these rotations are inherent in any application associated with the detection and tracking of a high speed moving object like an aircraft.

The camera mounted on the head of the missile goes through excessively fast movement resulting in excessive noise and low resolution of the image. The low resolution of the image is a major factor as it results in a rapid loss of features while tracking. These low resolutions and noise need to be handled in the current applications using camera. The efficiency and accuracy equilibrium maintenance in computer vision tracking techniques requires the reduction of important techniques like morphological filter convolution etc which affect accuracy. Therefore, important techniques are also required to manage these problems.

The radar based systems that are currently in use for the guided missile based systems for tracking the target in real world situations have a set of disadvantages attached to their use. The major disadvantage of such systems is the fact that if an aircraft is being tracked by radar based system, the object under tracking conditions is informed of this by the system deployed on it. This situation provides a warning to the pilot to hover the aircraft in some other direction to avoid collision. The camera based system completely handles this issue due to the fact that the tracked object has no idea about its tracking by a camera mounted on some system. Therefore, even if some disadvantages and problems arrive in the camera based missile guidance systems, the main advantage and requirement is the ability of such tracking system to keep on tracking without being detected.

1.3 SCOPE OF THE RESEARCH

The latest researches in the field of object tracking in computer vision have significant effects in terms of efficiency and accuracy. Some of the major problems out of those issues are identified and catered in this research. The low resolution, excessive noise, fast movement and rotation of the aircraft, low efficiency of the overall algorithm implementation, presence of heavily cluttered and rapidly moving background environments, and the problems associated with the correct prediction of the future position of the aircraft on the basis of transition matrix, rotation, translation, scaling, and angular velocity are catered in this thesis and algorithm.

There are some further sub fields which require attention to perfectly execute this application which is not catered for in this algorithm. The video sequences utilized in the development and testing of the algorithm are made from stationary cameras with small translation motions. In actual camera mounted missile guidance systems, the camera is mounted on the head of the missile, and therefore it also goes through excessive transitions creating many further problems including the possibility of aircraft losing from the line of sight. In such situations, the probability of the location prediction of the aircraft is important to place the missile back in the correct zone and direction. These issues are presently not catered for in this thesis, and can be worked on for further research in the domain.

1.4 OBJECTIVES OF THE RESEARCH

The objectives of this research are to develop a system for the detection, tracking, and location prediction of an aircraft with the help of a high speed moving camera mounted on a guided missile system. The research tends to handle low resolution video sequence images with extensive noise, highly cluttered and varying backgrounds, and high speed rotations and translations of the aircraft. An enhancement of the KLT algorithm with certain other techniques is also presented to improve the overall efficiency of detection while sustaining high rate of accuracy of detection.

1.5 PROBLEM DEFINITION

The tracking of a fast moving object on a video sequence received from a fast moving camera is the main problem addressed. The fast movement of the camera due to the attachment with a high-speed moving missile results in a low resolution, low frequency, and high noise image which make the detection of features and the tracking of aircrafts a difficult task. The high speed movement of the aircraft adds to the recognition challenges which are further handled in the problem.

1.6 THESIS OUTLINE

The thesis is divided into six chapters. Chapter 2 is focused on the description of the problem, previous techniques of aircraft detection, and the problems and challenges associated with those techniques. Chapter 3 focuses on the previous work done in the sub domain of object tracking in computer vision with details of the different tracking algorithms and techniques that have been used in the tracking of object with major focus on the algorithms related to aircraft detection. Chapter 4 is the description of the technique proposed in this thesis for the detection, tracking, and location prediction of the aircraft. Chapter 5 contains the evaluation of the research work in the project, explanation of results, comparison with other algorithms, and explanation of dataset used for the implementation and testing of the algorithm. Chapter 6 contains the conclusion of the thesis and the recommendation for future work and further improvements possible to enhance the quality of the algorithm.

1.7 PROBLEM OVERVIEW

Infra-red guided aircraft tracking systems are widely used in guided missiles, but the limitations, problems, and shortcomings associated with these systems have given rise to the need of aircraft tracking systems using high speed cameras mounted on the guided missiles. The heat seeking missile systems are expensive and have further shortcomings which can be catered with the use of image guided aircraft tracking systems with the use of a camera mounted on the front of guided missile. IR tracking systems get severely blind in weather conditions with high humidity; they can be deceived by creating clutters in the background and have to compete with massive heat source of sun, and requires a huge computing speed and power in the system to deal with the problem of false alarming.

The use of computer vision techniques for a guided missile with a low resolution camera is investigated in this paper. Such a system will help the camera based guided missile systems to track the current position and predict the future possible position of the aircrafts replacing the IR based sensors and catering for the issues involved in their utilization. The main challenges involved in this problem include the low resolution of the camera images, the extensive rotation and movement of the aircraft under tracking, constant change in the overall texture of the aircraft due to rapid translations, rotations, and low resolution of the image, presence of the image noise, high speed moving camera, and the possibility of the presence of highly textured and cluttered backgrounds. The algorithm also needs to be highly efficient to be able to run in real-time on dedicated hardware.

Tracking problems under the conditions of fast translations, rotations, illumination, and occlusion require complex and extensive calculations which require a fully dedicated and computationally fast dedicated hardware otherwise the algorithms fails to run in real time. One example is the high and low pass filters and morphological operations which even though are useful for tracking objects even in non rigid motion difference and occlusion but they require a large number of sample for supporting reliability of the detection and tracking of the target object.

Certain algorithms have been put forth for the tracking of aircrafts including SIFT feature tracker, SURF feature tracker, and other optical flow algorithms. Stavros [6] used cascade training, non-maximal suppression, and temporal filtering for detecting aircrafts with high efficiency, but the algorithm required high computational power, and only works for stationary cameras while fails for the condition of high speed moving cameras . Jeffrey [8] uses a predictive optical flow algorithm with spatial consistency and predictive testing for tracking aircraft but the trade-off is that high frequency vibrations decreases the effectiveness of the spatio-temporal smoothing, hence the efficiency is low with the miss rate of around 18%.

In this thesis, a new framework is presented that includes processes for handling all the issues present in the above mentioned system with increased efficiency and accuracy including the removal of noise, handling of the camera rotation issue in the KLT algorithm (Translation-based KLT tracker under severe camera rotation using GPS/INS Data), change in the aircraft texture due to extensive rotations and low resolution of the image, and the presence of a highly cluttered and rapidly varying background.

CHAPTER 2: KANADE-LUCAS-TOMASI FEATURE TRACKER

2.1 INTRODUCTION

Kanade-Lucas-Tomasi is an efficient feature tracking technique used in the domain of computer vision for the detection and tracking of objects. It is basically a feature extraction technique, but certain comparisons to find the matching features in different frames help in estimating the transition of the whole image, as a result providing the tracking dimension. KLT tends to increase the accuracy of feature extraction from the previously available algorithms but its sole purpose of introduction and implementation was to deal with the problem that the tradition techniques used for image registrations were generally costly. Spatial intensity information of the image is used by the KLT feature tracker to search for the particular positions that yield best performance in terms of matching. If the number of features to be extracted and tracked, and the number of matching features to be found between the images are less in number then the KLT feature tracker is better in performance and faster in efficiency.

2.2 BASIC TECHNIQUE OF FEATURE EXTRACTION AND TRACKING

The implementation of the KLT feature tracker is based on the early work done by Lucas and Kanade [2], but the complete algorithm was developed by Tomasi and Kanade [3], and the clear description of the algorithm and its working was explained in the paper written by Shi and Tomasi [4]. Certain modifications in the KLT feature tracker algorithm were also proposed in the coming years to improve its accuracy, efficiency, performance, limitations, and challenges. A slight modification was proposed later on by Tomasi which actually tends to makes the computation symmetric with respect to the two images. The resulting equation as a result of this modification is explained by Stan [5]. The number of features to be tracked in the KLT feature tracker can be changed, and limiting the number can have positive and negative impacts. The features are extracted and arranged on the basis of their strengths in an ascending order. Good features in an image are located by the examination of the least eigen values of each gradient matrix with dimensions 2 by 2, and the features are tracked by the use of Newton-Raphson technique of minimizing the different between two windows. Sometimes the displacements between the two consecutive images are huge which are handled in an accurate manner in the competitive algorithms like SIFT and SURF. For this purpose, multi-resolution tracking is used in the KLT feature tracker. The affine computations for evaluating the features consistency between non-consecutive frames [3] was implemented by Thorsten Thormaehlen[10] [11].

2.3 THE REGISTRATION PROBLEM

A large number of image registration applications are available in the field of computer vision technology. The image registration techniques used in the past are costly; therefore KLT uses a different technique for image registration that is based on the spatial intensity gradient of the images for the purpose of finding a good match among features with the use of Newton-Raphson iteration [12]. This technique can be used to register and estimate the scaling, rotation, and translations of the features among consecutive and non-consecutive image frames. The techniques present before the Kanade-Lucas-Tomasi feature tracker suffer in performance in the conditions of high rotations of an object in the image. KLT provides a better image registration technique using spatial intensity gradient information for directing the search of a particular location or position that provides the best match.

The translational image registration problem faced in the previous algorithm can be categorized in the following manner: There are two different functions G(x) and F(x) which contain the pixel values at the location in two images where x is basically a vector. The intention is to calculate the disparity vector h which minimizes a particular measure of the different between G(x) and F(x+h), for vector x in a particular region of interest R.

Difference measures between G(x) and F(x+h) are:

$$\text{L1 norm} = \sum_{XER} \text{abs}(F(x+h)-G(x))$$

$$\text{L2 norm} = \left(\sum_{XER} [F(x+h)-G(x)]^2\right)^{1/2}$$

The negative of normalized correlation is:

$$= -\sum_{XER} F(x+h)G(x) \,/\, \left(\sum_{XER} F(x+h)^2\right)^{1/2} \left(\sum_{XER} G(x)^2\right)^{1/2}$$

The registration problem identified and the algorithm provided for handling this issue in the KLT feature tracker gives rise to certain important application implementations. The generalized problem of registration implemented in this algorithm can be applied to extract the information of depth from the stereo images. The problem of depth extraction information with the help of a stereo pair basically has four different principle components: finding different objects in pictures, matching the same objects in the two different views, determining the parameters of the camera, and determining the camera to object distance. KLT uses the approach of matching the object with solving for the distance of the object and for the camera parameters by the use of a form for the registration technique of the algorithm in a rapid manner.

The techniques for locating the object include an interest operator [13], the zero crossings in the band-pass filtered images [14], and the linear features [15].

2.4 PYRAMID IMPLEMENTATION

The standard KLT algorithm deals with the small displacements between the images. The solution to this is the use of a pyramidal implementation. Let $I^0=I$ be the image's zero level then the representation of the pyramid is built respectively in the following order:

$I^0 \text{->} I^1 \text{->} I^2 \text{->} I^3 \text{->} I^{Lm}$

where the superscript of I coordinate with the level of the image

At any given point u in the image I, it has to find the corresponding position by the help of the formula $v=u+d$

The simple overall algorithm for the tracking of the pyramid works as follows:

d^{Lm} is computed at the L_m level of the pyramid and then the d^{Lm-1} is computed with the help of making an initial guess of the d^{Lm} at the previous level. This process is repeated up till it reaches the level zero. After all the calculations are performed for the tracking of the features in image sequences, the final equation of the optical flow solution d that is retrieved is as follows:

$d = \sum 2^L d^L$

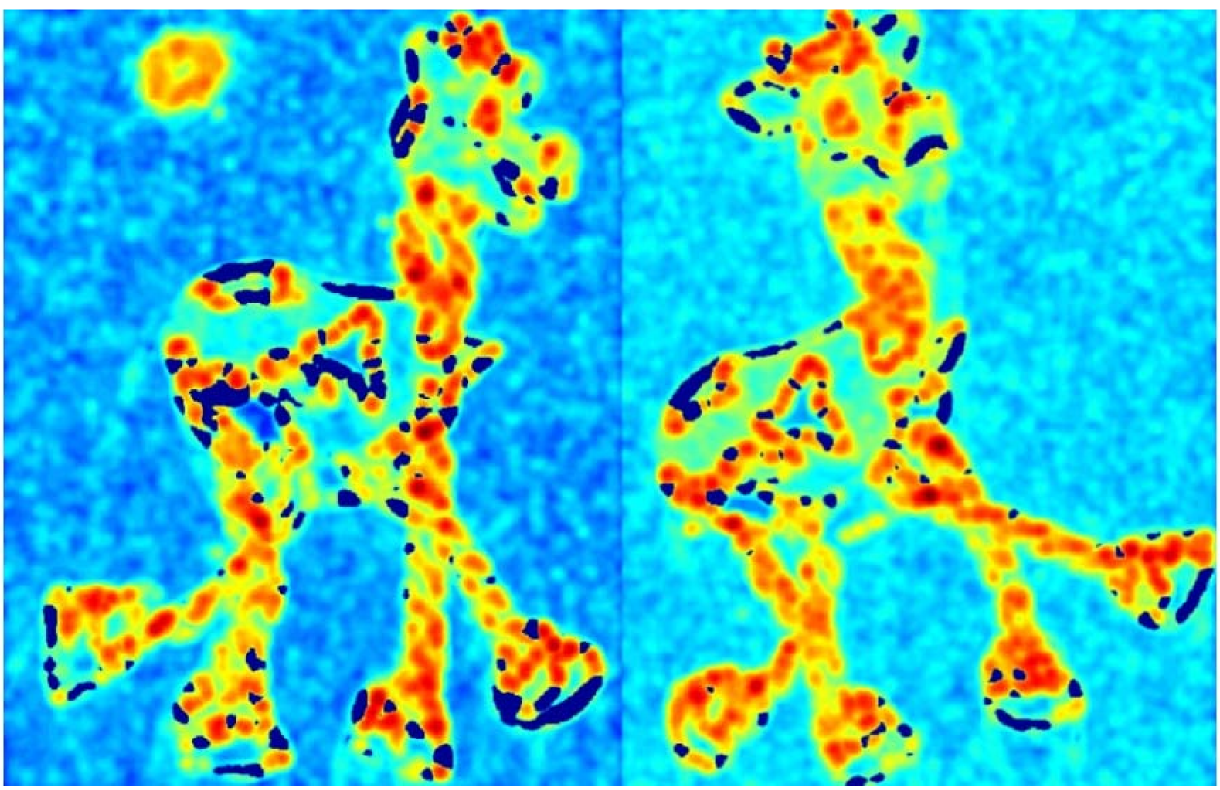
Figure 2.1: Different levels of the images for pyramid implementation of KLT

2.5 FEATURE SELECTION IN ITERATIVE KLT ALGORITHM

The KLT feature tracker revolves around the extraction and tracking of features in consecutive frame in the tracking of objects in video sequences. The algorithm for the selection of features in the KLT algorithm is based on the following steps:

1. Compute the matrix G and the minimum Eigen value at every pixel in the image.
2. Pick up the maximum value among all the Eigen values over the complete image.
3. Retain the pixels of the image that have Eigen Values larger than a predefined mentioned value of the threshold.
4. Retain the maximum value of pixels within their own localities. A particular pixel is only kept of it has the maximum Eigen value of all the pixels that exist in its 3*3 neighborhood.
5. Now keep a subset of all of these pixels so that any pair of pixels has a minimum distance between them larger than a predefined threshold distance value.

The following images below define the process of feature selection with the use of KLT feature tracker.



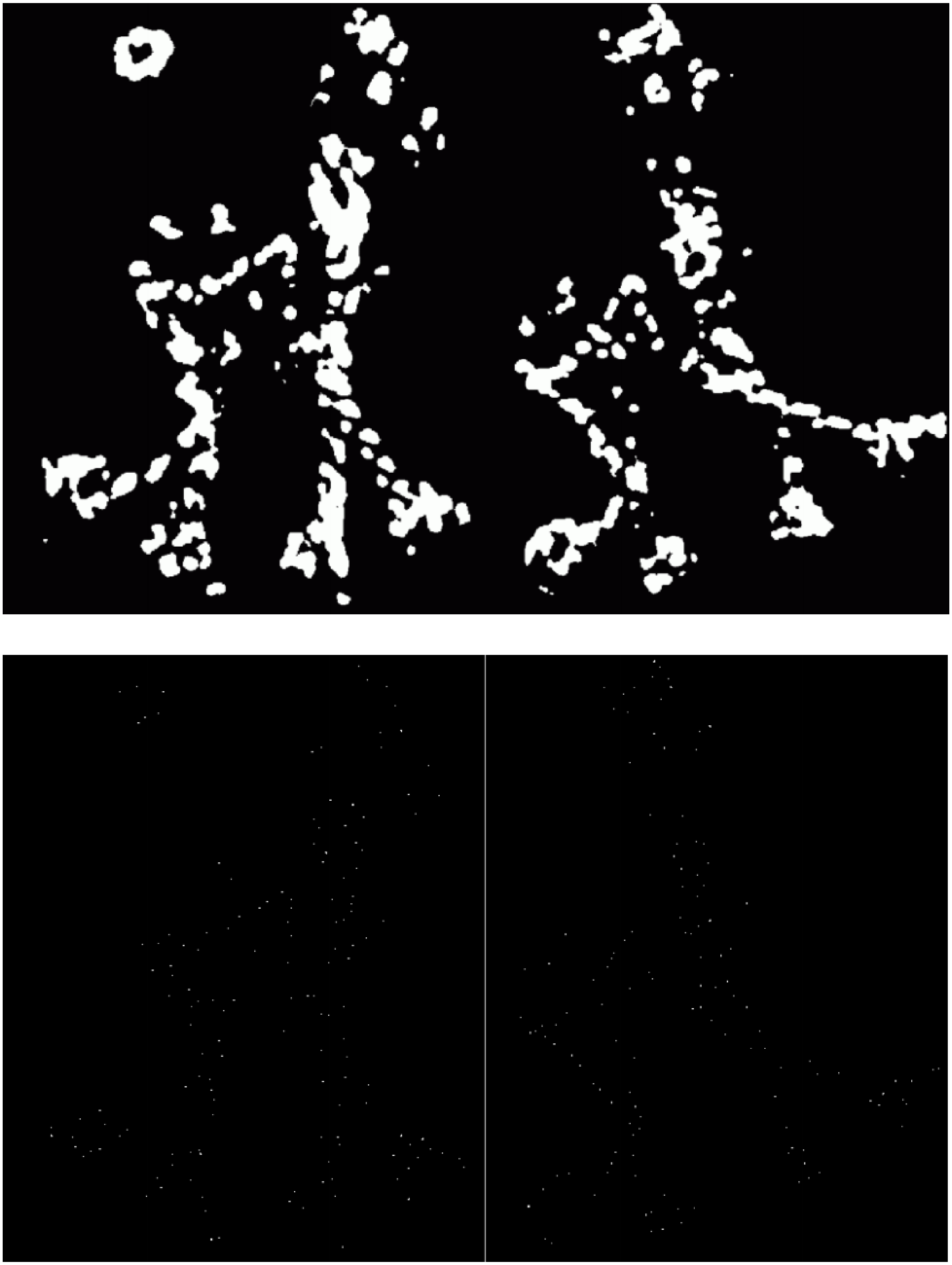


Figure 2.2: Resulting images of KLT feature tracker using pyramid technique [16]

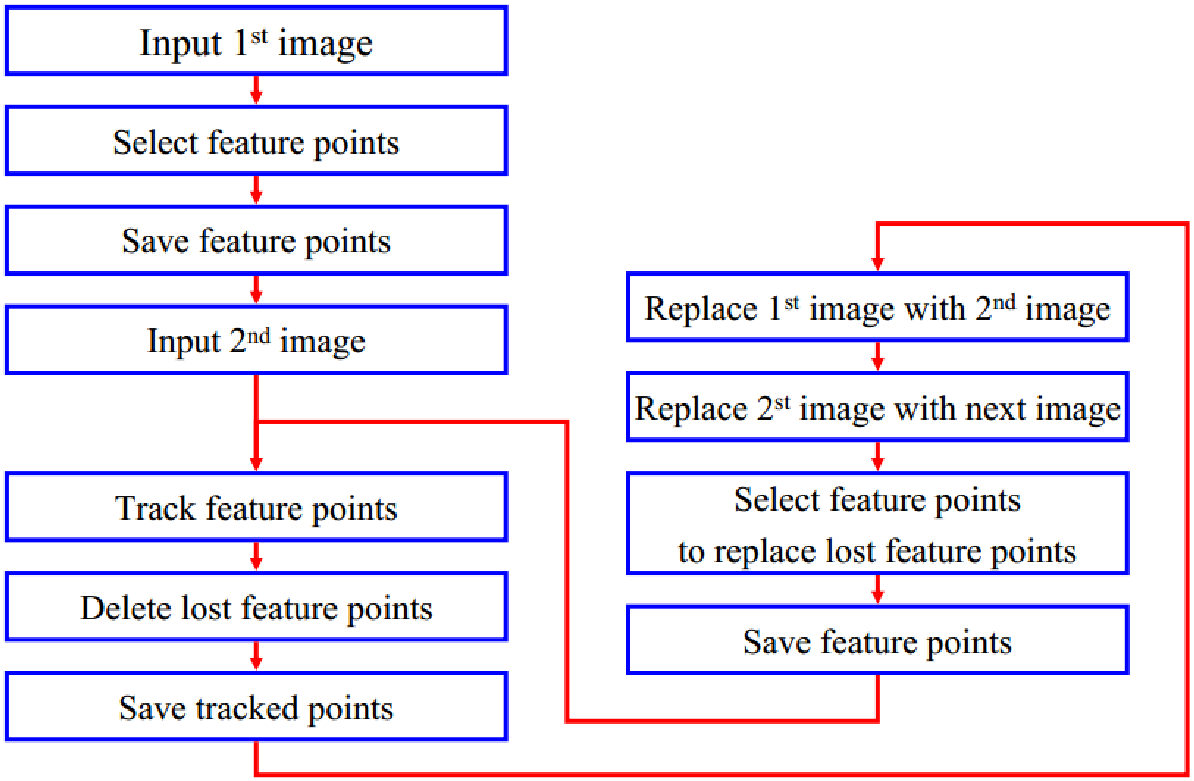


Figure 2.3: Implementation algorithm of pyramid KLT feature tracker [16]

CHAPTER 3: LITERATURE REVIEW

This chapter focuses on discussing the previous work done in the field of object detection and tracking with particular focus on the techniques and algorithms used for the tracking of aircrafts. Different feature tracking and matching algorithms in the category of optical flow techniques are discussed with their pros and cons. The need and importance of the aircraft tracking, the limitations of such an application, the associated challenged, and the need for high efficiency are discussed in detail. This chapter basically focuses on providing details about the previous work done in the field which will help in comparison with the algorithm designed in this thesis and explained in the next chapters.

3.1 DETECTION OF AIRCRAFT WITH CASCADE TRAINING

3.1.1 INTRODUCTION

Stavros et al [6] has implemented an application of aircraft tracking system with the use of Viola and Jones object detector. The overall approach tends on the machine learning rather than the typical techniques that were previously used that included filtering and morphological techniques. The problem of aircraft detection with the use of computer vision techniques and methodologies is a very challenging problem since the movement of the aircraft is very quick and changes the subpixels wise location rapidly, plus the background can be cluttered heavily. This type of system can be used as a part of a collision avoidance system for the sake of warning the pilots about the potential possibilities of collision. This paper presents the overall result of detection to be around 80% with a false positive rate of detection that is comparable with the previous approaches that are used which involve the morphological operations and filtering techniques. The system in this application was evaluated on more than 15000 frames which were extracted from the real sequences of video that were recorded by NASA.

3.1.2 ALGORITHM

The algorithm in this particular algorithm works as follows: Images are retrieved from the camera and are converted at run time into frames. After the frames of the camera images are extracted, the next step is the scanning of those images[17]. After scanning, the frames are passed through the detection procedure which has been developed on the basis of cascade training. After the passing through the detecting point of the algorithm, the non-max suppression is done. The image is moved through the temporal filtering and the targets are detected [18]. The collision risk estimation is done once the temporal filtering is done and the targets are detected [19]. These estimations are not important within the domain of out aircraft tracking application's domain.

3.1.3 STAVROS ET AL RESULTS

The results of the algorithm have been tested on a number of frames, and the correct detections as well as the false positives are estimated. Results have been estimated on 4 different sequences, and the detection rate has been estimated. The detection rate has been calculated on the basis of correctly detected aircrafts out of the total number of present aircrafts in the frame. The detection rate in sequence 1 is 81.68 %; in sequence 2 is 84.3%, in sequence 3 is 96.38 %, in sequence 4 is 89.14 %, providing an averaging detection rate of 87.88 % [20].

Another testing is done to estimate the number of correct detection from a large number of frames. There are 2 different sequences that have been tested, each contain the number of frames equal to 1181, and 1301 respectively. The number of false positives in these frames is 264 and 115 which provides an average percentage of the false positives equal to 15.26 %.

3.2 DETECTION USING PREDICTIVE OPTICAL FLOW ALGORITHM

The work presented by Mccandless JW in 1999 [8] focuses on the detection of aircraft in the video sequences based on the optical flow algorithm. The optical flow technique is one of the most commonly used techniques in tracking especially when the movement of objects is at a rapid pace. The work presented in the paper belongs to the category of computer vision based systems that segregates spurious optical flow artifacts for the detection of a moving object. The algorithm is comprised of six steps. First step is to compensate all of the pixels of an image for the rotation of camera. The second step is the smoothing of the images with the use of a spatio-temporal Gaussian filter. The Gaussian filters are low pass filters that are used for the smoothing of images and for the reduction of edges and noise. In the third step of the algorithm, the optical flow calculations are done with the use of a gradient based technique. The fourth step in the process is that the optical flow vectors are discarded which have small magnitudes. Then the vectors with similar magnitudes, directions, and locations are combined together by the use of a spatial consistency test. In the sixth step of the process, the similar optical flow vectors are temporally extended for the purpose of making predictions about the future location, directions, and magnitudes in the coming frames of the video sequence. The optical flow vectors that come out to be consistent with the predictions are actually the ones that are associated with a moving object.

3.2.1 RESULTS

The algorithm devised in the framework of this algorithm was tested on a large number of images retrieved from video sequences that were captured by the use of a camera mounted below the noise of a Boeing 737. Two different sequences of the same time were recorded by the cameras using another camera that was mounted on another Boeing 737 flying aircraft. The aircraft was detected in almost 82 % of the frames by the algorithm in the first sequences and in 78 % of the frames in the second sequence. The false alarm rate in all the sequences was exactly zero. The results obtained from this algorithm show the effectiveness of a comprehensive technique based on prediction in the process of the detection of moving objects.

Table 3.1: Stimulus-response matrix for providing the classification of the performance of the overall algorithm. For the sake of providing an example, probability for P=12 and for P`=3 are shown in each of the box [8].

Hit (82 %)	Miss (18 %)
False alarm (0 %)	Correct rejection (100 %)

Table 3.2: ROC results ([hits]/[false alarms]) for the first scenario. These results demonstrate the effect of different number of frames used in the predictive technique. The term P refers to the number of frames used in the prediction process. The term P' refers to the number of those frames which satisfy the temporal consistency test. For example, if P'/P = 75% and P=8 frames, then 6 of those 8 frames must be satisfied by the temporal consistency test. If the prediction technique is not used (P=0) false alarms occur in every frame. [8]

(P`/P) (%)	0 (off)	4	8	12
0	97 / 100	- -	- -	- -
25	- -	86 / 34	85 / 1	82 / 0
50	- -	76 / 1	73 / 0	69 / 0
75	- -	54 / 0	33 / 0	18 / 0
100	- -	23 / 0	4 / 0	1 / 0

To further test the accuracy and effectiveness of the algorithm, the results were tested and examined for another sequence i.e. sequence number 2 which contained a total of one hundred and sixty (160) frames containing an aircraft in the moving state. The second sequence of images has a certain similarity to the first sequence of images except for the fact that the background of the aircraft under target is the sky and not the earth in this case. The typical ratio of signal to noise in these images is 25 dB which is lesser than the signal to noise ratio of the first used sequence. In this sequence, the hit rate as well as the false alarm decrease when the overlap (P` / P) is increased and the duration P is increased. Once again, an overlap of 25 % with a total duration of P = 12 frames provides the maximum hit rate of all testing which turns out to be at 78% for the zero false alarms. The results for the both of the predictive parameters P` and P is shown in the table below:

Table 3.3: ROC results ([hits]/[false alarms]) for the second scenario. The parameters here are identical to those in Table 3.2.

(P`/P) (%)	0 (off)	4	8	12
0	96 / 100	- -	- -	- -
25	- -	96 / 14	93 / 1	78 / 0
50	- -	77 / 1	35 / 0	11 / 0
75	- -	28 / 0	3 / 0	0 / 0
100	- -	3 / 0	0 / 0	0 / 0

3.2.2 EFFECTIVENESS OF THE OPTICAL FLOW AND PREDICTIVE ALGORITHM

The predictive technique used in the tracking and detection of the aircrafts in the video sequences help in providing an effective mean for the reduction of false alarms in the optical flow computations. The hit rate as well as the false alarm rate decrease when the overlap increases in the predictive technique. The main reason for this decrease is the addition of a new stringency level beyond the test of the spatial consistency for the calculation and computation of the translational optical flow vectors. The effectiveness of the optical flow computations is limited by the quantization and temporal filtering, and as result inaccurate optical flow vectors are created in the target region under consideration [21].

The main benefit attached with the use of this technique is its provision of a further improved and more accurate assessment of the optical flow vectors by discarding the inconsistent vectors. The highest hit rate among both of the sequences is attached with a zero false alarm in the case when at least three of the last 12 frames are in consistency and correspondence with the current frame under consideration. These parameters help in providing a much higher hit rate while eliminating the false alarms in the aircraft detection.

Another mechanism that is possible for the reduction of the false alarms is by the use of the pre-processing for ignoring the regions that have much lower contrast. However, a region with low contrast can actually still correspond to the target due to the hazy conditions of weather. In addition, the projected shape, the surface reflection, and the size properties of the target aircraft are completely unknown, as a result making the contrast assumption quite inappropriate. As a result of this the computations of the optical flow are not limited solely to the image regions with high contrast.

The frameworks in the domain of the optical flow predictive models are designed for the production of accurate results without any consideration for the time required for the detection of the target object. At present, the algorithm I refined in the real-time environment with a low possible decrement in the performance for instane, the total weights associated with the equation of gradient-constraint may be changed to the byte wise format i.e. either 1 / 8, 4 / 8, or 5 / 8, thereby minimizing the total bits number required for the implementation on the hardware.

A CCD mounted camera with the availability of a total of 256 gray levels and a 30 Hz frequency provide an acceptable figure for the camera to be used as a sensor for this particular application. In the example explained above with the use of the camera, the target subtended approximately around 27 pixels with the formation of 9 * 3. For greater distance target, the field of view of the camera can be reduced such that the total number of pixels driven by the target remains same and consistent. However, the tradeoff is basically that the vibrations at the high frequency can cause an excessive motion on the surface of the image, as a result making the spatio-temporal switching less effective. As a consequence, the aircraft vibration accuracy measurement and motion may be considered when the field of view of the camera is decreased for providing further pixels of the target.

The predictive technique defined in the above mentioned optical flow algorithm is similar to the predictive technique [22]. The predictive technique that has been proposed in this framework is much similar to the theory of temporal coherence that is proposed for the explanation of the human motion perception. This theory basically accounted for the psychophysical experiments in which the basic subject that attempted for the detection of a single dot moving with a consistent motion while surrounding the dots moving with random motion.
This theory that has been accounted for the experiments in the domain of psychophysics in which the subjects are attempted towards the detection of a single for that is moving with a consistent track of motion. The information of the speed was discarded in the implementation of the framework. The prediction based technique provided is different in a way that the information of the speed is actually utilized for the validation of the optical flow vector computations. Another different exists in the terms of the signals of velocity themselves. With

the theory of biological temporal coherence, the smoothing was applied for the sake of the biological plausibility to the outputs associated with the directionally tuned cells which results in the addition of new noise to the overall system. With the use of the predictive technique, the signals of the velocity are directly used for the sake of estimating the coherence. A similar and related approached to the temporal coherence was taken in consideration by Kruger et al for the detection of the automobiles on the road [23]. The key assumption in this framework was the fact that the rotation of the camera was zero and the ground surface was much more planer. Kalman filtering was also utilized for the same of estimating the depth and tracking the optical flow vectors. Tracking of these potential obstacles was terminated in the base if there were no appropriate measurements of the optical flow because these measurements are inherently noisy. The predictive technique utilized in the framework described here does not really need an accurate measurement of the optical flow in each of the frame. As a consequence, this particular technique gives rise to the provision of a much more robust means for the detection of the moving objects in the frames retrieved from the video sequences.

3.2.3 COMPARISON OF FEATURE TRACKING ALGORITHMS

The experimentation done on a large number of frames for different object tracking algorithms including the optical flow KLT algorithm and the two-frames matching algorithms like SURF and SIFT. The KLT algorithm has shown almost equally good results in terms of efficiency with much better results in terms of accuracy. SURF proves to be the most inefficient of all algorithms but the performance of the algorithm is comparatively fine. Some of the charts below try and demonstrate, in the form of charts, some of the comparisons between the different algorithm using images that are mentioned in the table headings.

Figure 3.1: Average tracking error comparison of feature tracking algorithms

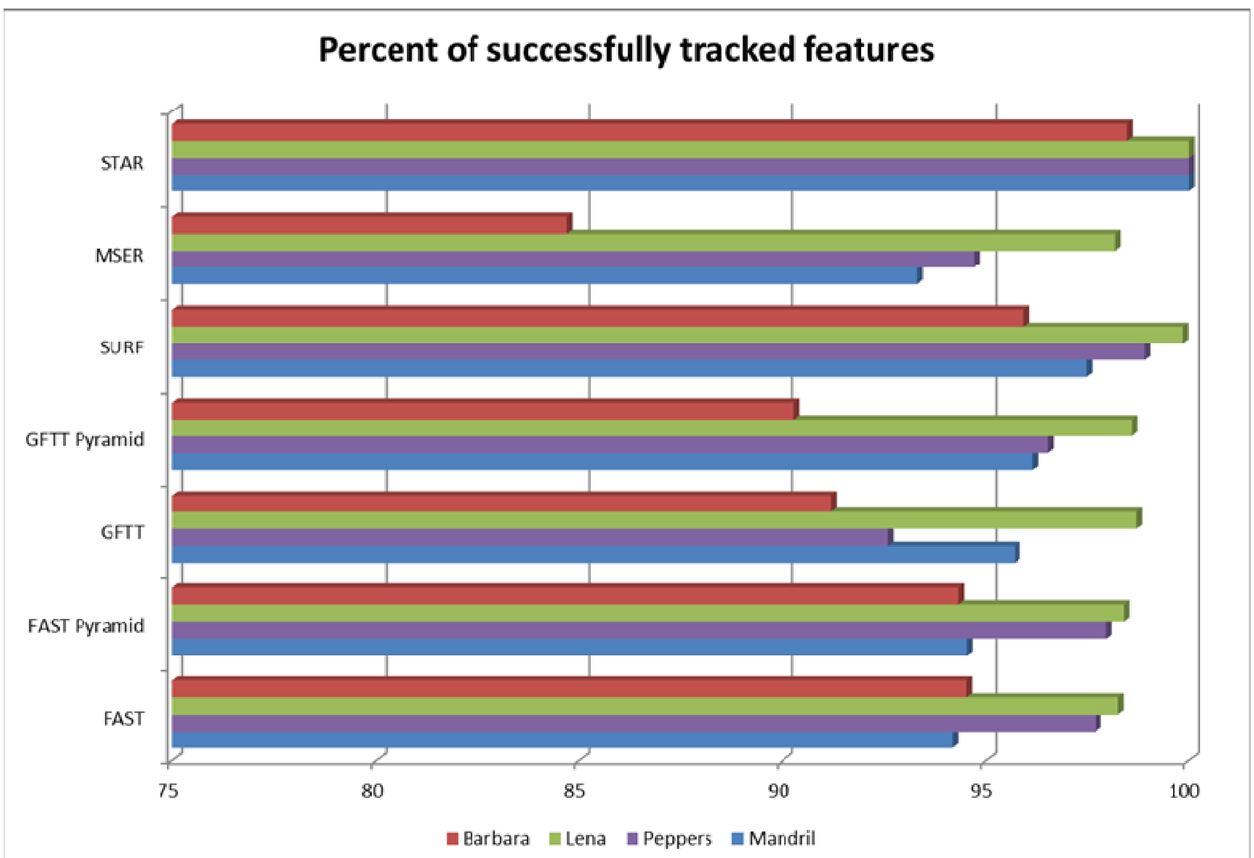


Figure 3.2: Percent of successfully tracked features comparison of feature tracking algorithms [24]

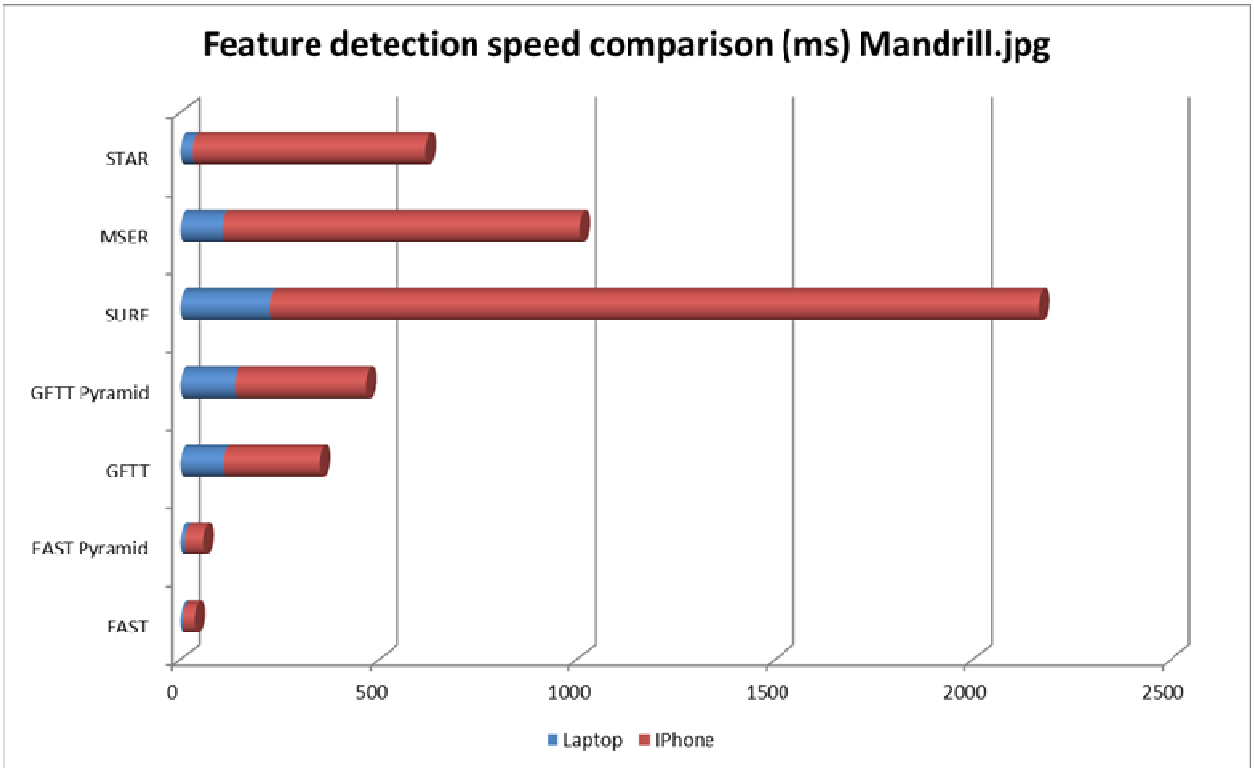


Figure 3.3: Feature detection speed comparison of feature tracking algorithms [24]

CHAPTER 4: PROPOSED TECHNIQUE

In this chapter, the complete framework and algorithm developed in the thesis is explained in detail with the help of flowchart and steps following for the detection, tracking, and probabilistic location prediction of the aircraft in low resolution images retrieved from video sequences in real time systems.

4.1 OVERVIEW

The framework and algorithm proposed in this paper is based on Kanade-Lucas-Tomasi (KLT) features [2][3][4][5]. KLT feature tracker is an efficient feature tracking technique as explained in the discussion above. KLT feature tracker tracks the features from one frame to the next. The features are scattered throughout the image depending upon the overall texture of the background, and the number of objects in the image. A bounding box of the aircraft is given in the first frame for locking the target object for tracking. In this frame, all the features within the bounding box are kept in consideration while the rest of the features are temporarily thrown out of the interest. The aircrafts are detected by correlating the matching features in consecutive frames, but KLT alone provides a low accuracy and prediction performances with majors errors in aircraft detection and a huge miss ratio in the presence of highly textured background with a large number of objects. For handling this complete problem, a framework and algorithm has been established, which is presented in this paper. This chapter provides details of the algorithm including the preprocessing techniques applied on image frame including filters and morphological operations, the improvements and enhancements done in the KLT feature tracker, random sample consensus algorithm, and aircraft modeling and comparison in the consecutive frames. Two different algorithms have been suggested in this report with one without the image pre-processing i.e. implementation of filters and morphological operations while the other involving the pre-processing. The former focuses on the increased efficiency while keeping accuracy at a high while the later focuses on further improving the accuracy of the algorithm by affecting efficiency to some extent.

The overview of the algorithm can be explained as follows: KLT feature tracker tracks the features from one frame to the next. The features are scattered throughout the image depending upon the overall texture of the background, and the number of objects in the image. The aircrafts are detected by correlating the matching features in consecutive frames, but KLT alone provides a low accuracy and prediction performances with majors errors in aircraft detection and a huge miss ratio in the presence of highly textured background with a large number of objects. For handling this complete problem, a framework and algorithm has been established, which is presented in this chapter.

The algorithm works as follows:

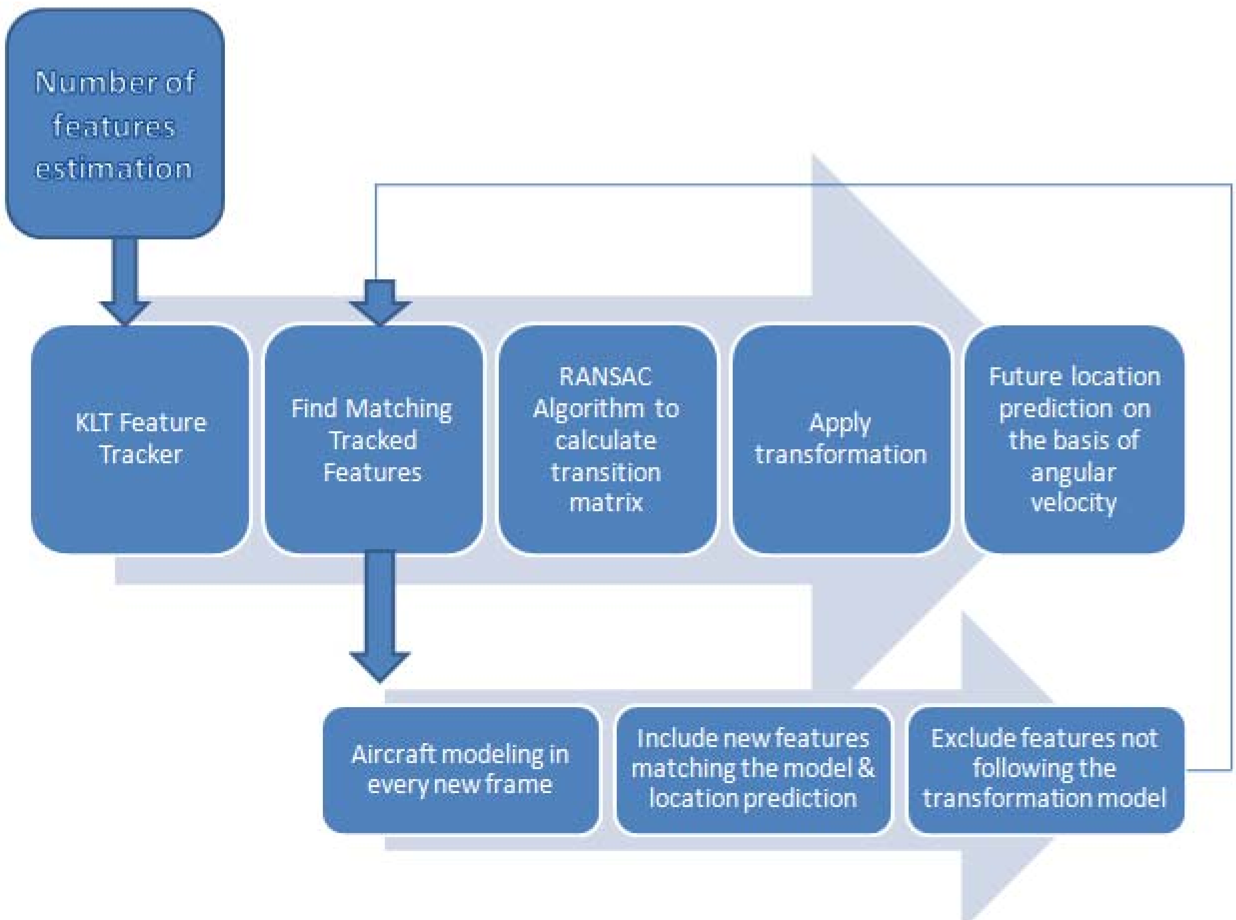


Figure 4.1: Flowchart of overall algorithm

4.2 PREPROCESSING OF IMAGE

An image of a fast moving aircraft retrieved from a low resolution camera connected to a fast moving guided missile system is bound to inherit excessive noise and blurring. The low resolution will also cause the image to be dull and unclear. These images are although clear enough to be viewed properly by the human eye but when they are given as an input to a computerized system for applying image processing and computer visions techniques, there is a high possibility and risk of faulty detections and estimations as a result of applied operations. For this basic purpose, it is necessary to apply certain operations on the images i.e. to perform some preprocessing on the image before applying algorithms and techniques to ensure high performance of the system. Certain techniques are applied to perform preprocessing techniques on image frames to remove noise, enhance the image, and adjust the contrast, sharpness and brightness of the image to make sure the image is ready to be provided as an input to computerized software.

4.2.1 FRAME RETRIEVAL FROM VIDEO

Image processing operations are all performed in frames i.e. images instead of videos. Therefore, the first step in the process of performing operations in videos or live camera feed is to extract frame sequences from the video. This process in the algorithm is actually slightly different from the process that needs to be implemented in the real-time systems. In this algorithm, a set of videos have been used, and the videos are converted into frame sequences before initialization any operations on them, which means that all the image frames were available at the time of starting the operation. On the contrary, in the actual real-time system, the camera feed will be used, which will provide one frame at a time. In such systems, it will be necessary to estimate how many frames to retrieve, and there will also be a need to find out the efficiency of the algorithm on live camera feed based systems.

4.2.2 NOISE REMOVAL AND IMAGE SHARPENING

Low resolution image frames captured from moving camera, of an environment that contains a highly textured background, has the inherent input of excessive black and pepper noise as well as blurring. It has been tested in the development of the algorithm that if an averaging filter for noise removal and high pass filter for edge enhancement for sharpening the image are used before feeding the images to the KLT feature tracker, the results have much higher accuracy, but at the same time, the efficiency is affected, as a result resulting in a low frame processing rate per second. For the noise removal, a weighted average filter and a median filter is used to reduce salt and pepper noise whereas high pass ‘Sobel’ filter is used to enhance the edges of the image.

Sharpening of edges helps the feature tracker to find out better features easily and efficiently in the frame.

4.2.3 CONVERSION TO PORTABLE GRAY MAP

KLT feature tracker works with the images in the format of portable gray maps. The retrieved image sequence for the video is in ‘jpeg’ format. A code is written to convert the image to the required portable gray map before sending over to the KLT feature tracker. All the preprocessing is done on the normal jpeg images, and all the algorithm after the KLT feature tracker is applied is also done on the normal jpeg images. The portable gray maps of the images are used just because of the requirements to be input to a KLT feature tracker. Once the features are extracted from the image, the normal images are utilized for any further operations on those features or the image itself.

4.3 KLT FEATURE TRACKER

A wide range of feature extraction and comparison algorithms are available but KLT is utilized in the algorithm for two basic reasons: One being the better performance in terms of the efficiency when the number of features required to be found out are small and secondly due to its much better performance in the tracking of features. Other algorithm are available that have better performance in future extraction, but KLT is better in terms of the tracking of features because that is the major aim of the KLT feature tracker algorithm.

The images are fed into the KLT feature tracker to detect the features in a frame and try to track those features in the next frame. If KLT fails to track a particular feature during the course of running on different frames, it informs that the particular frame is lost, and finds a new one as an alternative for making sure that the total number of features remains the same throughout. The features are scattered throughout the image depending upon the overall texture of the background, and the number of objects in the image, and these features need to be handled properly and in detail to execute the tracking system with high accuracy which is discussed later in the algorithm in detail.

4.3.1 ESTIMATING THE NUMBER OF FEATURES FOR KLT

In contrast to most of the other feature extraction algorithms, KLT takes a number as an input to decide the number of feature to be extracted and tracked. In most of the other algorithms, this number cannot be controlled; the number of total features and the matching features can actually be controlled in the KLT feature tracker.

There are certain reasons why this number needs to be controlled. The higher the number of features needed to be extracted and tracked by KLT, the higher is the time required to compute and track them, and hence resulting in lower efficiency. Apart from efficiency, the accuracy is also affected: both when the number of features is too small or too large. When the number is too small, many important features are lost, and when the number is too large, many unimportant features are extracted which are either noise or belonging to unimportant objects in the frame which require much handling in manipulation of those features and as a result affect the performance of the algorithm. If the number of extracted features is too small then many important features are not extracted at all. Since, an aircraft goes through excessive rotations, many features are lost in every consecutive frame, and so dealing with a small number of features largely reduce the accuracy of tracking. Due to these problems, the number of features has to found out such that equilibrium between efficiency and accuracy is reached.

KLT feature tracker finds out ‘n’ number of features in the first image, and tracks these features in the upcoming images. If a feature is lost, an alternative feature is found which is tracked from their onwards. The number ‘n’ can be provided to KLT feature tracker, and the manual attempt at trying different values varies the effects on the efficiency and accuracy of the tracking and detection of objects. If the number ‘n’ is significantly high, the KLT algorithm consumes a lot of time and the efficiency of the algorithm is affected. Many weak features and image noise is also tracked if ‘n’ is very large. On the contrary, a small value of ‘n’ leaves behind many important and strong features affecting the accuracy of the detection and tracking to a large extent. Though this value varies largely depending upon the application, the accuracy and efficiency of the algorithm has been tested by selecting different values of ‘n’ on 8 different videos to estimate the correct number of features to be tracked for best performance and results. All of the videos on which the code is tested are low resolution videos with resolution of 640*360. The value present in the table below is given by the formula.

Efficiency E= (Time Consumed in feature tracking T/ Maximum Time Consumed M for varying ‘n’) + (Time consumed in algorithm T2/ Maximum time Consumed M2 for varying ‘n’)

E percentage=100-((T/M)+(T2/M2)*100)

Accuracy percentage is calculated by the number of correctly tracked features.

Accuracy % = 100- (Miss rate/ Correctly tracked features) *100

V= (efficiency percentage+ accuracy percentage)/2

Table 4.1: Testing results for efficiency and accuracy of KLT for varying number of features

n	V1	V2	V3	V4	V5	V6	V7	V8
50	85	81	80	86	88	90	80	88
100	90	88	82	90	90	94	86	90
150	90	89	93	90	91	95	90	88
200	88	86	86	88	90	92	88	86
250	82	83	87	87	87	89	83	84
300	79	84	75	84	85	87	82	83
350	75	74	74	82	83	86	81	81
400	71	69	71	79	82	85	77	80
450	68	69	69	80	80	83	76	79
500	65	68	68	80	77	80	73	74

The average calculated values provided an estimate to select the value of ‘n’ between 100 and 150 features for best performance in terms of the efficiency and accuracy of code. The large number of features not only has a negative impact on the efficiency but also affect the accuracy by capturing a large number of features containing salt and pepper noise. The number of features ‘n’ should ideally be kept between 100 and 150 in the case of low resolution images with moderately textured aircrafts and backgrounds. These tests have been conducted on eight different videos of the same resolution, and the consistency in the trend is visible.

As evident from the above table, the value increases and reaches the maximum at 100 and 150, and then start to decrease again. The graph of this table is displayed below to give a better idea. Though the average of efficiency and accuracy follow this trend, they separately follow a comparatively different trend than that demonstrated above. Efficiency continue to decrease somehow from less value of 'n' to higher value of 'n' but the accuracy reaches its peak at around 150 which makes the average equilibrium to come out at the mentioned point.

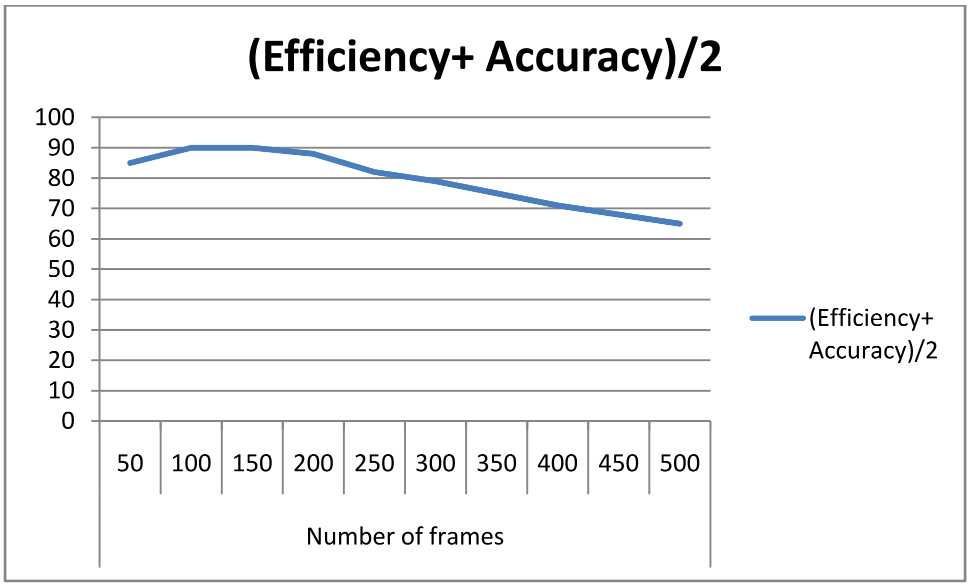


Figure 4.2: Chart displaying efficiency and accuracy average against number of features

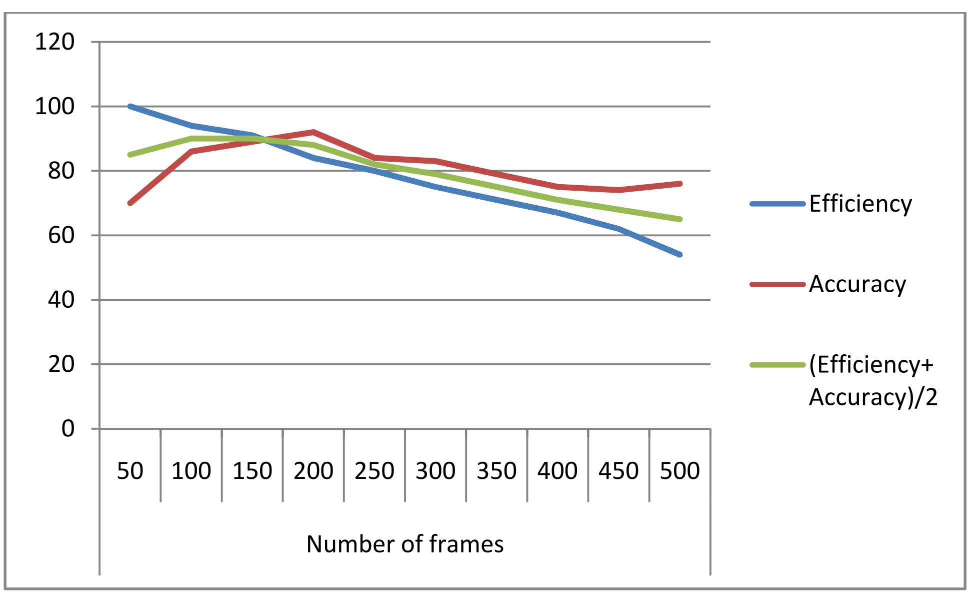


Figure 4.3: Chart displaying efficiency and accuracy separately

It is also required to be noticed that all of this testing to find out the estimated number of frames for the best possible results are done on low resolution and blurred images of a particular dimension. These results will be affected if frames from high quality cameras in stable environments are taken.

4.3.2 APPLYING KLT ALGORITHM

After the frames are extracted from the video, preprocessing and conversion to portable gray maps is done, and the number of features to be extracted to track the object in the image is decided, the next step in the algorithm is to apply the KLT feature tracker on the image sequences. The features are provided by KLT in a feature table from which these features are extracted by the help of 3 different matrices. The information of the features is saved in the x, y, and values matrices which each matrix having an n*f dimension. The 'n' number of rows represents the number of features to be tracked in an image whereas the 'f' number of columns is associated with the total number of frames in the video sequences. If a particular value of the location x(i,j) is taken, it gives the position in the x-coordinate of the i'th features in the j'th frame. Similarly, y also works in the same manner all together, but it gives the y-coordinate position of the feature.

Therefore, if the position of a particular feature 'i' has to be checked in any frame 'j', it can be recovered by getting the value of x(i,j) and y(i,j).

Position of feature 'i' in frame 'j'= [x(i,j) y(i,j)]

The third matrix value is the most important matrix in the KLT feature tracker, and the provision of this feature helps in manipulating the features with much more control than any other feature extraction and tracking algorithm available. The value matrix informs when a particular feature is lost in the process of tracking. Let's suppose, the below matrix represents the first features in 10 frames, the 'value' matrix appears something like this:

[0 0 0 0 0 0 6 0 0 0]

The value 6 in the 7^{th} frame represents that the first feature is lost in the 7^{th} frame. Once a feature is lost, an alternative feature is found out and then it's tracking start. The algorithm in this thesis handles the lost frames but eradicating them from the interest, and stop tracking them once they are lost. Handling of this issue is important, because a large number of features are lost in such applications in which the objects or camera or both are moving and rotating at a very face speed as they are in this particular application.

The above discussion of the KLT algorithm can be summarized in the following manner:

KLT provides features in the form of 3 matrices: x, y, and values with each having dimensions of n*f with 'n' being the number of features tracked and 'f' being the number of frames in the video. Each (i, j) of x and y matrices contains the coordinates or position of the feature in the image in a particular frame. The value matrix prompts about the features which KLT fails to tracks in the next frame. When the aircraft is locked in the first frame of the image, the algorithm decides to keep only those features in the calculations which exist within the range of the locked aircraft. The 'values' matrix provides information to decide which of the features are tracked in the next frame and which of the features are lost in between.

4.3.3 LOCKING AND TRACKING BOUNDING BOX

The first step of a guided missile system is the locking of target. The detection of the aircrafts in heavily cluttered background environments in low resolution images result in the lack of accuracy, plus the application of the guided missile system does not require the recognition rather requires the tracking of an aircraft. Multiple aircrafts can exist in the frame at a time, but only one of those targets can be locked. This basic problem forces the locking of the target aircraft in the first frame of the image.

After the KLT algorithm is applied on the image sequence, and the value of all 3 required matrices is received, the target aircraft is locked in the first frame by giving a bounding rectangle. The purpose from here on is to keep on tracking the features that exist in the box, but since the aircraft moves around, the box has to change its position i.e. it has to track the aircraft throughout the course of its flight. Certain features within the box will lose with time resulting in the arrival of new frames, handling of this transition of the box, inclusion and exclusion of frames under conditions requires further handling which is explained in detail in the next section.

4.4 RANSAC (RANDOM SAMPLE CONSENSUS) ALGORITHM

As discussed earlier, the value matrix prompts about the features which KLT fails and which it succeeds in tracking in the successive frames. When the aircraft is locked in the first frame of the image, the algorithm decides to keep only those features in the calculations which exist within the range of the locked aircraft. The ‘values’ matrix provides information to decide which of the features are tracked in the next frame and which of the features are lost in between.

The extensive rotation, small size of the aircraft in the image, and low resolution of the overall image results in an excessive feature loss. The lost features are removed from the consideration and only the matching features are considered.

When the aircraft is locked in the first frame of the image, the features that exist within the bounding box of the locking rectangle are considered as the features of the aircraft whereas the rest of the features are not considered. The value matrix gives information about which of these features are tracked in the next frame, and the x and y values corresponding to those coordinates provide the location of those features in both of the frames. There are certain features that are lost; as a result certain new features are included. Some features are lost because KLT fails to track them while some are intentionally given up if they fail to follow the trajectory of the rest of the features of the aircraft which is discussed in detail in the next sections. The new features that arrive in consideration have to move through certain modeling matching and other requirements that are also discussed further.

RANSAC algorithm is basically used to find the transition matrix providing scaling, rotation, and translation of the whole image in the two frames, but it is intended in this application to only find out the transition of the aircraft in the two frames, that is the transition of only a particular object in an overall relatively consistent and static image. This requires proper handling in a way that only the features that belong to the aircraft need to be provided to the RANSAC algorithm, and no other features are provided to it.

RANSAC algorithm only deals with the features that are matched in the two consecutive frames which are extracted with the help of the value matrix. RANSAC algorithm takes the coordinates of the matching features in the two frames, and estimates the transition matrix. The transition is initially found from the second frame to the first frame, therefore the inverse of the transition matrix is taken to estimate the exact transition of the aircraft from the first frame to the second frame.

It provides the translation, rotation, and scaling of the aircraft within the image in the form of a 3*3 transition matrix. This transition matrix can be applied on the initial aircraft locking rectangular area given in the first frame in every next frame to calculate the current position of the aircraft. The transition matrix provided has the format:

[sc –ss 0; ss sc 0; tx ty 1]

where sc=s*Cos(theta) and ss=s*Sin(theta)

The translation of the aircraft is found with the help of tx and ty. Scaling and rotation are found by the following formulae:

Scale= sqrt (ss*ss + sc*sc)

Rotation in degrees= atan2 (ss, sc)*180/pi

These equations provide the rotation, scaling, and translation of the aircraft which is then applied on the bounding rectangle to move, scale, and rotate it along with the aircraft. All of these steps in the RANSAC algorithm and their application in the algorithm framework in this work are mentioned in detail in this section.

4.4.1 FIND MATCHING TRACKED FEATURES

The matching tracked features between the frames are found out in the next frame by considering the value matrix. Those features that show a 0 value in the value matrix are tracked and those that have any other value are lost. It is checked that which of the features in the previous frame are tracked in the next frame, and their positions are given to the RANSAC algorithm. These feature points are then converted to SURFPoints to be taken as an input in the RANSAC algorithm.

4.4.2 CALCULATION OF TRANSITION MATRIX AND INVERSE

Once the positions of the matching features in consecutive frames are provided to the random sample consensus algorithm, a transition matrix is calculated. The transition matrix is calculated in the reverse order that is from the second to the first frame which means that the transition matrix needs to be inversed to calculate the exact transition on the aircraft from the first frame to the second. The transition matrix has the following format:

[sc –ss 0; ss sc 0; tx ty 1]

where sc=s*Cos(theta) and ss=s*Sin(theta)

4.4.3 FIND SCALING, ROTATION, AND TRANSLATION

After the above mentioned transition matrix is calculated, the next step in the process is to estimate the scaling, rotation, and translation on the basis of those features. The rotation gives the estimated of how much the aircraft has rotated between the last two frames, the scaling gives an estimate of how much distance the aircraft has travelled between the last two frames in terms of the z-axis, and the translation provides an estimate of the distance travelled by the aircraft in the x and y direction.

sc and ss provide the scaling of the aircraft along the x and y axis, whereas the overall scaling of the aircraft is found by the following formula:

scale= sqrt (ss*ss + sc*sc)

The rotation angle of the aircraft from the first frame to the second is calculated with the help of the following formula:

Rotation in degrees= atan2 (ss, sc)*180/pi

The translation of the aircraft can be simply found by the values of tx and ty where tx gives the translation of the aircraft along the x direction and ty gives the translation of the aircraft along the y direction.

4.4.4 TRANSFORMATION ON THE BOUNDING RECTANGLE

This transition matrix is then applied to the bounding rectangle tracking the aircraft. This basically keeps the bounding rectangle to keep the aircrafts within its bounds at all time, since it is following the same path and operations as that of the aircraft.

4.4.5 CALCULATION OF ANGULAR VELOCITY AND FUTURE LOCATION PREDICTION

This transition matrix not only helps in the tracking of the aircraft in the current frame but it also helps in providing an estimate of the angular velocity of the aircraft at the current frame. This angular velocity is calculated with the help of the transition matrix found out by the random sample consensus algorithm. It helps to predict the position of the aircraft in the frames to arrive. The results of these predictions are compared with the actual results in the result section to estimate the percentage correctness of the prediction. The location (x,y) of the aircraft in the future time is estimated by applying angular velocity on the current position of the frame. The prediction is made on the delay 5, 10, 15, 20, 30, and 60 frames to estimate correctness of the estimation at each interval to reach a threshold point between time delay and accuracy. The results are shown in the results section.

The angular velocity of the aircraft at any given point is found out with the help of the following formula:

Angular velocity in x= $(C_l + C_{lp}^B w_x d/2v)\ qSd$

Angular velocity in y= $(C_m + C_{mp}^B w_y d/2v)\ qSd$

Angular velocity in z= $(C_n + C_{nr}^B w_z d/2v)\ qSd$

where

C_l is the aerodynamic coefficient as a result of the roll moment

C_m is the aerodynamic coefficient as a result of the pitch moment

C_n is the aerodynamic coefficient as a result of the yaw moment

C_{lp} is the aerodynamic coefficient as a result of the roll damping

C_{mq} is the aerodynamic coefficient as a result of the pitch damping

C_{nr} is the aerodynamic coefficient as a result of the yaw damping

v is the total mean square magnitude of the airframe velocity d is the overall aerodynamic length [25].

4.4.6 CHECKING RESULT ACCURACY

After this angular velocity is estimated, and since the current position of the aircraft is already known along with the frame rate of 30 frames per second, the position estimation is done. The estimates are made at the current point to a delay of 5, 10,15,20,30, and 60 frames i.e. from 166milliseconds to 2 seconds, and the accuracy of the result is estimated at each delay. The accuracy of estimations definitely decreases with the number of frames but the estimation for a very small value of n is quite unimportant in terms of the application under consideration. Due to this reason, equilibrium has been found between the number of frames delay and the value of the accuracy of the estimation. The details of these results are displayed and discussed in detail in the results section.

4.5 DYNAMIC AIRCRAFT MODELING

The modeling of an object in image processing is associated with the concept of developing an estimate of the color combinations of the object to match it with the other images of the same object. Such modeling helps in the matching of the object with another counterpart of the same object.

Another major disadvantage of using low resolution video sequences is the lack of consistency in the gray scale model of the tracked object. The aircraft continuously changes its overall model of histogram during the video sequence, sometimes due to the excessive noise while sometimes due to the excessive rotation and scaling of the aircraft. The histogram and other image estimating measures of the aircraft change their course throughout the video. The figure below gives an idea of this change between two frames of the same video:

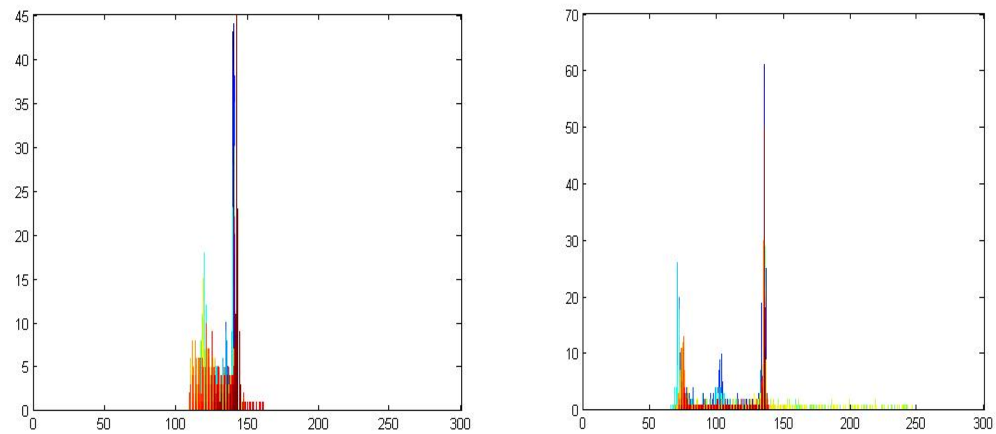


Figure 4.4: Different images of the same video sequence and their histograms

The two displayed frames are of the same video sequences and the histograms in the below images show the histograms of only the aircraft features. As evident that both the histograms belong to the same aircraft but have very different and largely spread histograms. The mean, median, histogram, standard deviation, and other image properties vary largely within the same video sequence. This is basically a consequence of low resolution frames, and morphological filters fail to handle the problem effectively.

Due to this reason, no static model of the aircraft can be kept throughout; there is a need for developing the aircraft model dynamically as it changes during the course. Keeping this problem in mind, a model of the aircraft is made at each frame. At any frame, the features that belong from within the rectangular area tracking the aircraft are kept, and the model including histogram, mean, median, minimum and maximum value is made. At every new frame, some of the features are lost and some new features arrive, while some features get tracked. The tracked features in the next frame help in forming the model of the aircraft in each successive frame, and the process if repeated throughout.

To summarize, the need for modeling, as the aircraft revolves and the noise in the image changes randomly, the model or the pixel range of the aircraft varies to a large extent. This causes a constant change in the model of the aircraft and cause a decrease in accuracy if a static model development in the first frame is followed. Therefore, there is a need to develop a system of consistently dynamic configuration of the aircraft model in every new frame.

This dynamic modeling is done in the following manner: At every new frame, the matched features are found. The new positions of those matched features can be used to extract the pixel configurations i.e. the intensity values of the aircraft. On their basis, the model of the aircraft is changed at every successive frame. The four different parameters that are used for the purpose of modeling the aircraft are the calculations of the mean, median, standard deviation, and histogram of the aircraft features. This model is used later on in matching which is explained in this section below.

4.5.1 MODEL MATCHING

Matching of image parameters based models is done by finding the mean square distance with a threshold value, and check if the mean square distance is lesser than or greater than the threshold value. Of the distance is lesser than the threshold value then it means that the model matches and if it is greater than the threshold value then it means that the model does not match. The parameters to develop and match the model are found out from the features of the aircraft. These features are found out by checking which of the features are present within the bounds of the bounding rectangle tracking the aircraft. All the features inside this bounding box are assumed to be belonging to the aircraft. The color intensity values at these features are used to develop the model at each frame, and whenever a model is developed, it is utilized in the exactly next frame with the purpose of comparison. At each frame, the model from the features in the current frame is compared with the model developed from the features in the previous frame to check all that if a new feature has been found out by the KLT algorithm within the range of the bounding box then weather it belongs to the aircraft or not. This model matching is basically utilized for two purposes: one is to check if a particular feature belongs to the aircraft and second is to check if a set of features belong to a completely different object that currently resides within the range of the bounding rectangle tracking the aircraft.

The matching criteria that is selected for the feature comparison is the range estimation i.e. comparison with the minimum and maximum value. If a new feature arrives near the outer ranges then continuous checking is done on it. The matching criteria for the features of a complete object including histogram matching, standard deviation matching, and median matching.

4.5.2 EXCLUDING FEATURES FROM CONSIDERATION

As discussed earlier, many features are lost in the process of excessive rotations of the aircraft in low resolution images with excessive noise and fast translations and rotations of the tracked object. But these features are all lost by the KLT feature tracker as they become less visible or completely hidden in the image frame. Apart from these features that are lost by the KLT feature tracker, there are some other features that are voluntarily left out and excluded by the algorithm framework developed in this thesis. There are two reasons for eradicating these features from consideration: Either these features are just noise or these features belong to some other object in the image other than the aircraft. The criteria and methodology behind the decision regarding this is discussed in the next sections in detail.

4.5.2.1 EXCLUDING FEATURES CONTAINING NOISE

Noise reduction in the images is done with the help of low pass filters. There are different filters that are available for the reduction of noise, each having a performance advantage over another on the basis of the type of noise that the application requires to reduce. The common low pass filters implemented are Gaussian filter, averaging filter, and median filter. These filters remove the strong edges in the image, and since noise is also strong edges i.e. high frequency components, therefore it is also removed with the help of averaging filters. There are certain problems associated with the use of averaging filters specific to the application of aircraft tracking in low resolution images. The average filters blur out the edges, and since low resolution images are used in this algorithm testing, further blurring of edges further increases the difficulty of detecting features in the aircraft. Therefore, it is disadvantageous to use low pass filters in low resolution blur images as they affect the tracking functionality of the KLT algorithm.

Although the reduction of high frequency components i.e. edges by the use of low pass filters can be covered by the use of edge strengthening methods, another major problem associated with the use of filtering techniques is the reduced efficiency caused by repetitive convolutions. Convolution is a process that takes a lot of time to compute, and especially in this application where efficiency is important and several filters instead of a single one is required to be applied, the effect of convolution on the overall efficiency is huge.

Due to this reason, the method implied for the reduction of noise in this algorithm has been decided after repetitive testing. The noise reduction has also been done by the use of the features that are detected by KLT feature tracker as explained in the start of the algorithm. Since noise components in an image are high frequency components, KLT detects them as features, but the features of noise only remain in an image sequence retrieved from a low resolution video for a certain period of time. After a time period of 20 to 30 frames, the noise features are lost by the tracker. Though this technique lacks the real time approach due to the need for considering the past 30 features, but if a particular feature is not tracked consecutively for the past 20 frame, the algorithm does not consider it considering it as a feature belonging to noise. This simple calculation of measuring if a feature has stayed for the past 20 frames or not does not take away from the efficiency of the algorithm at all, and the performance has been tested on a large number of videos and it provides high rate of accuracy.

4.5.2.2 EXCLUDING FEATURES WITHOUT MATCHING TRAJECTORIES

The noise portions in the image are tracked as features but are eradicated from the consideration as explained in the previous section, but there are also features in the images that are tracked by KLT but they belong to some object other than the aircraft. There are certain criteria for the inclusion of new features into the consideration due to which such objects do not get to consideration but if they do then the algorithm uses the path matching technique to remove them from consideration. This can be explained with the help of an example, suppose there are two objects in the image; one is the aircraft and the second being the sun. The aircraft features tracked by the KLT feature tracker follow a particular transition between the two frames. The translation, scaling, and rotation factors in those features in the consecutive frames are comparable to each other. If any features belong to an object other than the aircraft, its transition is not the same. This gives rise to the reason for eradicating this feature from consideration.

A typically difficult condition to handle in such situations is when an aircraft is moving from top of a sun. The features of the sun come within the bounding box of the tracking rectangle and some of those sun features arrive in the consideration of the algorithm due to matching histogram and mean values with the aircraft model. Such features need to be eradicated before they mislead the tracking, and the testing of the matching trajectories and transition following of the features help in the removal of such features from the calculation consideration. If the aircraft is moving from top of the sun to the right then all the aircraft features would move to the right in the next frame whereas all the sun features will move to the left, clearly displaying a mismatch of the trajectory following. This way it is implemented in the algorithm to handle with collisions between the tracked aircraft and another object.

4.5.3 INCLUDING NEW FEATURES INTO CONSIDERATION

As a large number of features are lost in the process of tracking and many others are voluntarily left out, the total number of features continues to decrease which causes problems in terms of accuracy of the code. Almost every feature in an image sequence is lost at least once when the aircraft rotations are high and backgrounds are highly cluttered. This problem indicates the importance of including new features into the consideration at the run time. The criteria and methodology for the dynamic exclusion and inclusion of features into consideration is explained in much detail in the next sections.

4.5.3.1 INCLUDING FEATURES WITHIN BOUNDING RANGE

As explained earlier, there is a tracking bounding box that keeps the aircraft within its range at all times. The transitions of the aircraft are calculated and applied on the bounding box to make it move along the aircraft in different frame sequences. As some of the features are lost in the aircraft tracking due to low image resolution, high noise ratio, high speed rotations and translations of the aircraft, and highly cluttered background environments, therefore new features need to be included in the consideration for further calculations. It is required to make sure that the new features that are put under consideration belong to the aircraft and there is certain criterion that needs to be satisfied before it is decided. The first criterion is explained in this part. A new feature is only put to the next text for selection if it belongs from within the tracking rectangle bounding box. If it is outside of this range, it is not considered for the test explained in the next section.

4.5.3.2 MODEL MATCHING

A model of aircraft with the help of all its detected features in a particular frame is developed which includes the histograms, minimum and maximum gray scale ranges, mean, median, and standard deviation values of the aircraft pixels. If a new feature has to be included in the consideration for further calculations and it has passed the criteria of presence within the range of the tracking rectangle bounding box then the next test it has to clear is the matching of the pixel values with the aircraft model developed with the help of matching features between the previous and the present frame. If the pixel value of the feature in consideration is within the minimum and maximum ranges of the pixel values of the aircraft features then it is considered as a feature belonging to the aircraft, but there are certain other issues that are handled in this particular condition.

A particular set of features that exist within the bounding box of the tracking rectangle is also used to develop a histogram as well as mean and median calculation of all of such features, and this set of features is compared with the model of the aircraft. If the models match than the new complete set of features belong to the aircraft and if it does not match then it is an indication that the new set of features belong to another object and hence that object should be dealt with since it is in collision with the aircraft at that particular time frame. Any features belonging to the model of that other object and following its trajectory are neglected from the consideration for tracking in the future permanently to make sure that the collision of the aircraft with sun, clouds, another aircraft, or any other object can be handled properly.

CHAPTER 5: EXPERIMENTAL RESULTS

The accuracy of results in the guided missile based system is of upmost importance. Maintaining high rate of accuracy is a very critical issue especially when using low resolution images and attempting to keep high rate of efficiency. An automated system based on cameras instead of radars and sensors is crucial for the tracking and detection of aircrafts with high rate of accuracy. The intention of the framework devised for this particular thesis is to maintain a high rate of accuracy while keeping a considerably high rate of efficiency. This chapter contains the material for the evaluation of the results of the proposed algorithm. Comparison of the algorithm with the state of the art techniques and the accuracy in terms of special challenges faced in the application is provided in this chapter in detail. The accuracy, efficiency, and the visual inspection of the results of the proposed algorithm are presented in this chapter in detail.

This chapter focuses on providing the results of different measures considered in the testing of this application. The aircraft detection accuracy, the aircraft detection accuracy in terms of features with respect to hit ratio and miss ratio, the position prediction accuracy of the aircraft depending upon the delayed number of frames, the tracking accuracy in terms of features when collision occurs, and the effectiveness of the noise reduction algorithm. All of these results are focused on and discussed in detail in this chapter.

5.1 DATASET

The testing of an algorithm is done on a particular set of data that helps in the analysis of the results of the overall framework. The application under consideration in this algorithm needs a set of database containing a number of videos of long length so that the number of frames on which the testing is done is considerably enough to devise conclusions. There are certain particular extreme cases like excessive rotations of the aircraft, fast movement of the aircraft, excessive noise in the image frames, collisions of the aircraft, and the movement of the target aircraft from over sun or other objects that must be present in the dataset selected for the testing for the purpose of fully finalizing and deriving the results of the framework. A database having dataset of varying cases having all possibilities that input data can be is useful. In guided missile based systems, an automated system is used, where there is possibility of having different types of input data, such a database is useful for testing and training of the system that contains a complete variety of all possible cases. The database used for the assessment should contain data having good quality.

NASA provides a database for the detection of multiple aircrafts in the images but the intention of this application is much different than the one presented by NASA. The papers published in the domain of aircraft tracking have used normal video frames from the internet due to the lack of availability of a dataset for the testing of such an application. For this purpose, videos are captured from the internet to test the algorithm on that input for devising the calculated results to an accurate manner. A large data set has been selected for the testing of the application to improve the accuracy of the results. The algorithm has been tested on a total of 20 different videos of approximately 10 minutes each. The resolution of all of the videos is 640*360, and the frame rate is 30 frames per seconds. The overall data input that has been provided to the algorithm is approximately around 18 thousand frames.

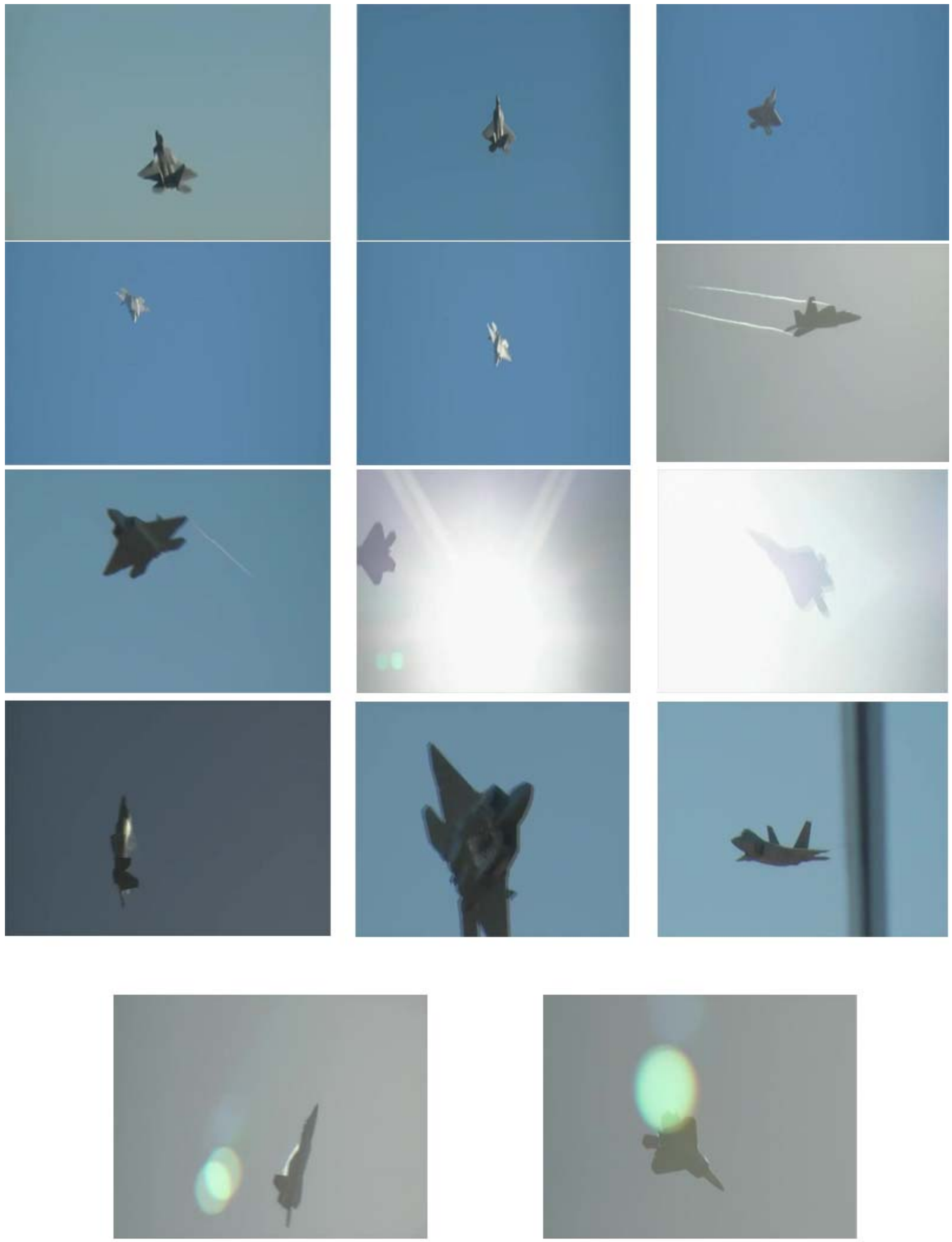


Figure 5.1: Dataset frames from different conditions

5.2 RESULTS

The results of the aircraft tracking accuracy, prediction accuracy, tracking accuracy in conditions of collision, noise reduction techniques, and the snapshots of results with the comparisons with other algorithms are provided in this section of the report in detail.

5.2.1 ESTIMATING NUMBER OF FEATURES TO BE TRACKED

KLT feature tracker finds out 'n' number of features in the first image, and tracks these features in the upcoming images. If a feature is lost, an alternative feature is found which is tracked from their onwards. The number 'n' can be provided to KLT feature tracker, and the manual attempt at trying different values varies the effects on the efficiency and accuracy of the tracking and detection of objects. If the number 'n' is significantly high, the KLT algorithm consumes a lot of time and the efficiency of the algorithm is affected. Many weak features and image noise is also tracked if 'n' is very large. On the contrary, a small value of 'n' leaves behind many important and strong features affecting the accuracy of the detection and tracking to a large extent. Though this value varies largely depending upon the application, the accuracy and efficiency of the algorithm has been tested by selecting different values of 'n' on 8 different videos to estimate the correct number of features to be tracked for best performance and results. All of the videos on which the code is tested are low resolution videos with resolution of 640*360. The value present in the table below is given by the formula.

Efficiency E= (Time Consumed in feature tracking T/ Maximum Time Consumed M for varying 'n') + (Time consumed in algorithm T2/ Maximum time Consumed M2 for varying 'n')

E percentage=100-((T/M)+(T2/M2)*100)

Accuracy percentage is calculated by the number of correctly tracked features.

Accuracy % = 100- (Miss rate/ Correctly tracked features) *100

V= (efficiency percentage+ accuracy percentage)/2

Table 5.1: Testing results for efficiency and accuracy of KLT for varying number of features

n	V1	V2	V3	V4	V5	V6	V7	V8
50	85	81	80	86	88	90	80	88
100	90	88	82	90	90	94	86	90
150	90	89	93	90	91	95	90	88
200	88	86	86	88	90	92	88	86
250	82	83	87	87	87	89	83	84
300	79	84	75	84	85	87	82	83
350	75	74	74	82	83	86	81	81
400	71	69	71	79	82	85	77	80
450	68	69	69	80	80	83	76	79
500	65	68	68	80	77	80	73	74

The average calculated values provided an estimate to select the value of 'n' between 100 and 150 features for best performance in terms of the efficiency and accuracy of code. The large number of features not only has a negative impact on the efficiency but also affect the accuracy by capturing a large number of features containing salt and pepper noise. The number of features 'n' should ideally be kept between 100 and 150 in the case of low resolution images with moderately textured aircrafts and backgrounds. These tests have been conducted on eight different videos of the same resolution, and the consistency in the trend is visible.

5.2.2 AIRCRAFT TRACKING ACCURACY

The tracking of the aircraft has been tested on video sequences of 18000 frames with perfect results but some of the features are lost and some unnecessary features arrive in the image. The accuracy of the tracking of aircraft has been calculated on the basis of loss of features. As explained in the algorithm, features that do not belong to the aircraft are not included even if they fall within the bounding rectangle tracking the aircraft, and if they are included by mistake then they are excluded if they do not follow the same trajectory as the aircraft. The aircraft tracking accuracy has been improved for handling all conditions and the overall tracking done by a number of features is exactly accurate but some of the features are included and excluded which should not be. This has been tested by examining the features that are previously included but are excluded later. The overall result of the feature based aircraft tracking is described in the table below.

Table 5.2: Aircraft tracking accuracy results

Hit Ratio 92%	Miss Ratio 8%
False Alarm (included features that do not belong to aircraft) 9.6%	Correct Rejection (Excluded features that belong to the aircraft) 1.1 %

5.2.3 PREDICTION ACCURACY

The position of the aircraft in the future time is computed with the help of its current angular velocity as explained in the algorithm. The angular velocity helps in the estimates of the predictions of the aircraft at delays of 10, 20, 30, and 60 frames which accommodate for the estimated location after 333ms, 667ms, 1 second, and 2 seconds. The accuracy of the prediction is checked with the actual results received after a few frames to check the prediction accuracy. The frames with extensive camera movements resulting in low accuracy in certain frames (not due to the algorithm) are rejected in the calculation. The below table estimates the prediction accuracy of the algorithm.

Table 5.3: Aircraft prediction accuracy results

10 frames	20 frames	30 frames	60 frames
98.2%	96.4%	92.2%	85.1%

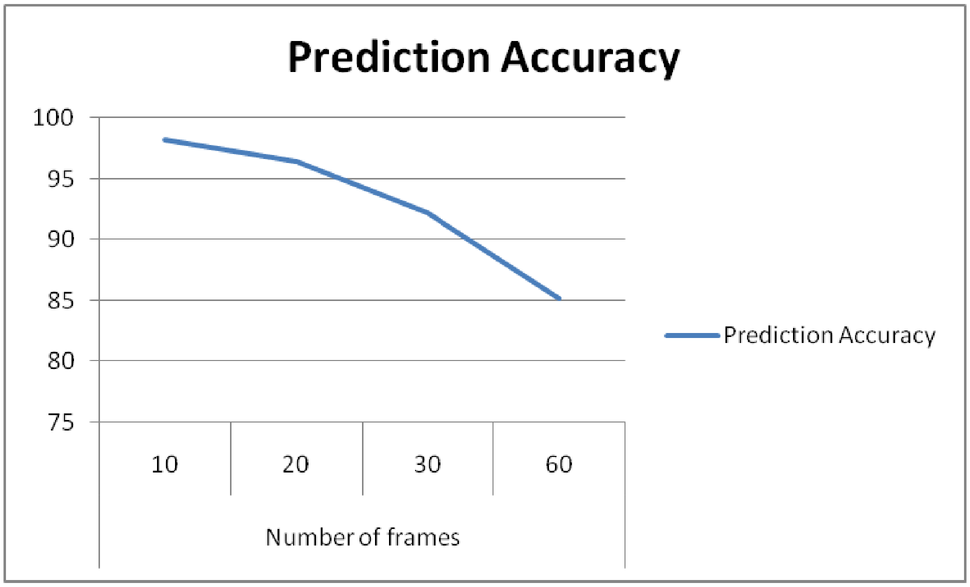


Figure 5.2: Aircraft prediction accuracy graph against number of frames

The prediction accuracy of the algorithm is calculated by finding the mean square distance from the initial position to the final position, and then calculating the distance from the position 'f' frames later with the estimated final position. These two distances are used to find out the mean square distance percentage of the algorithm which is shown in the table and graph above.

5.2.4 TRACKING ACCURACY IN COLLISIONS

The accuracy of tracking features decreases with the increased background cluttering and collisions of the aircraft with other objects as well as with the increasing noise. The estimates of accuracy provided below only deal with the testing on the frames that have collisions and other background cluttering and excessive noise issues. These tests have been conducted on a smaller number of frames and on different video sequences. The overall result has been estimated to be around 70%.

5.2.5 RESULTS OF NOISE REDUCTION TECHNIQUE

The noise reduction technique employed in this framework is unconventional and does not deal with the traditional techniques of reducing the noise in the images. The noise reduction technique has been tested and the results have been calculated in the terms of number of frames that are removed as noise, and the number of features that are missed out in terms of noise and are not removed. Hit ratio is estimated as a measure of the correctly reduced noise features whereas the miss ratio is estimated as the number of noise features that are missed.

Hit ratio = correctly identified noise features / Total noise features = 740016/889068= 83.23%

Miss ratio = 1 – hit ratio = 16.76%

The figures below show the results of the noise reduction technique. The first image shows the image without the reduction of noise whereas the second image shows the result of the image after applying noise reduction technique.

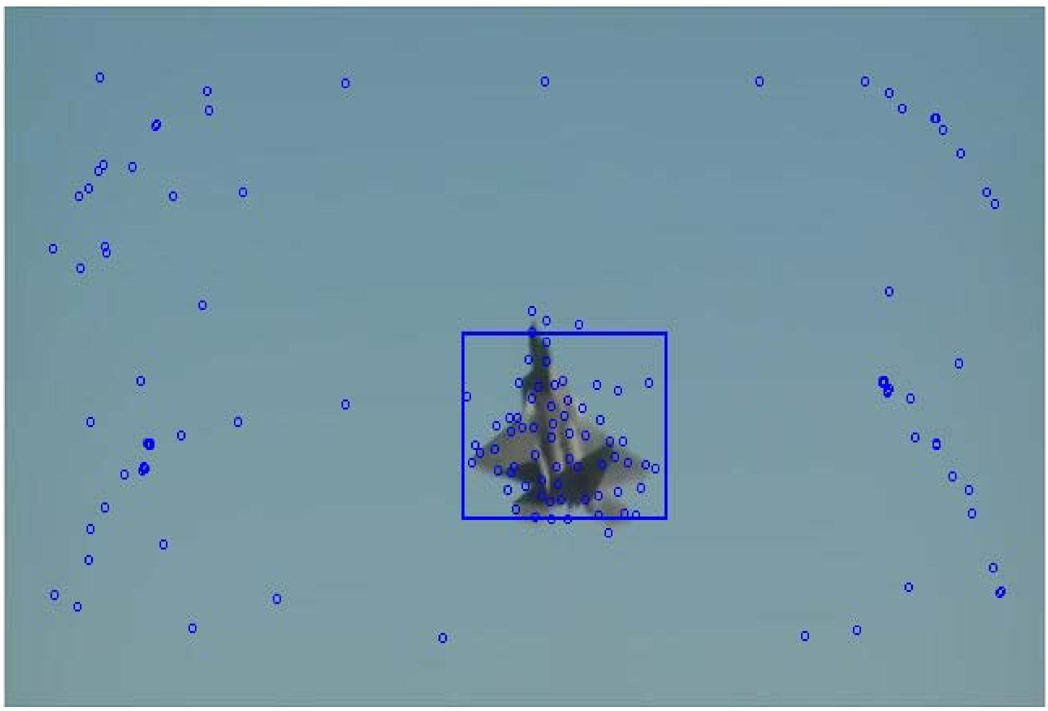


Figure 5.3: Image features without noise reduction

Figure 5.4: Image features after noise reduction

5.3 RESULT SNAPSHOTS

This section displays the output image frames of the results of the algorithm. The description of the images is written in the figure title details. The position prediction estimator box is also show in the last figure for displaying the output.

Figure 5.5: Result with bounding tracking box

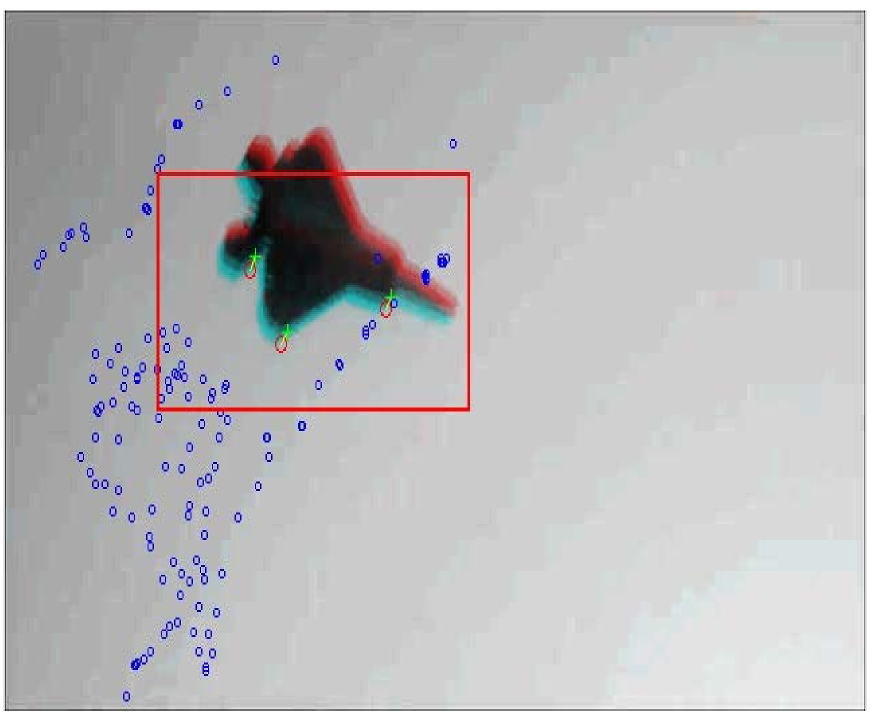


Figure 5.6: Result with another condition of aircraft

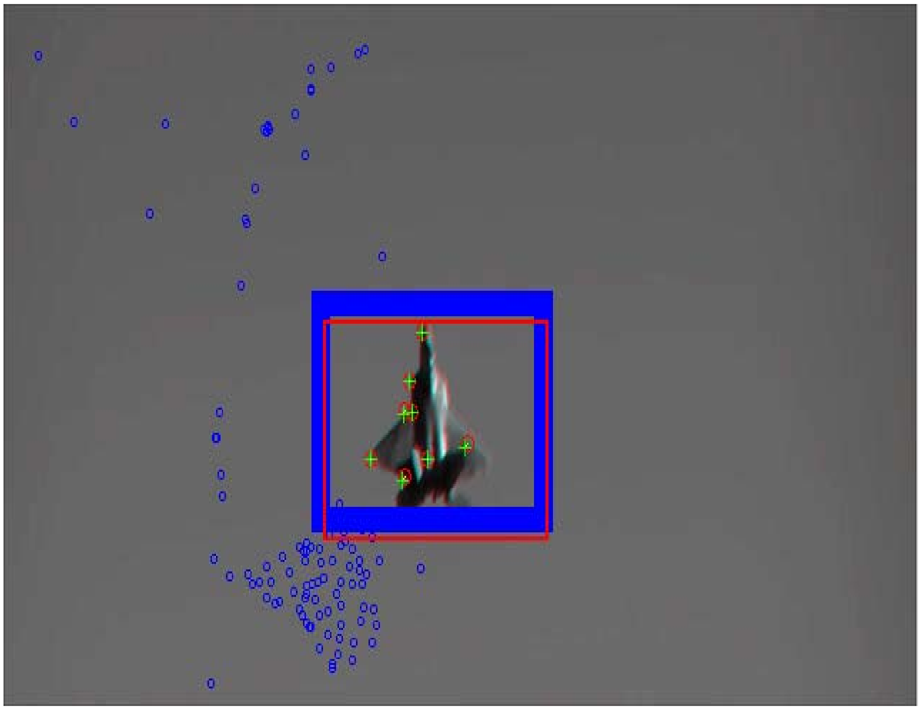


Figure 5.7: Result with red bounding box and blue prediction box

5.4 COMPARISON WITH OTHER ALGORITHMS

The comparison of the framework implemented in this research is done with the two other techniques that are discussed in the literature review section of the thesis. The below figure shows the accuracy estimation of the results of this algorithm with both of the algorithm discussed in the literature review: One by Jeffrey et al. and other by Stavros et al.

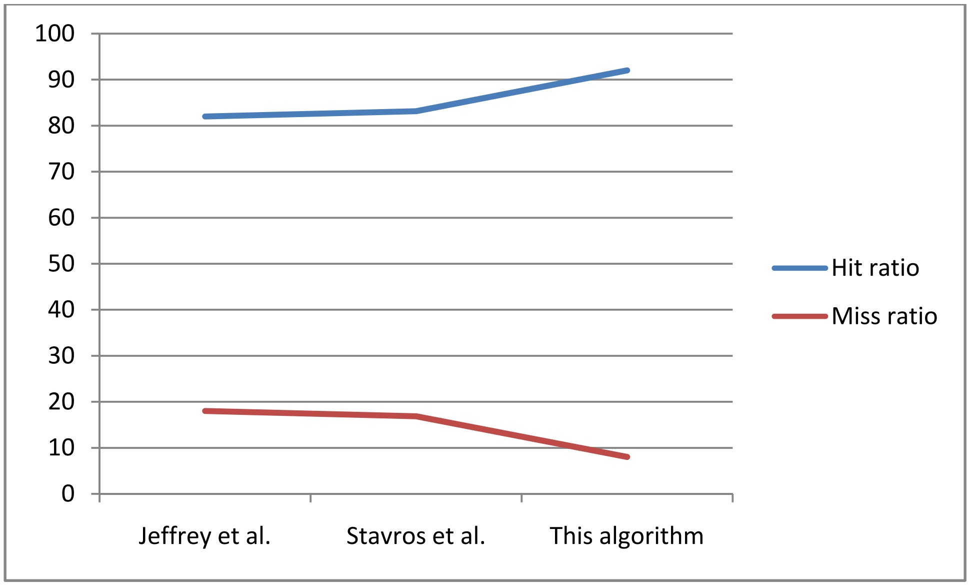


Figure 5.8: Comparison with other algorithms

CHAPTER 6: CONCLUSION AND FUTURE WORK

This chapter focuses on the explanation of the overall conclusion of the framework developed for this thesis, and the directions for the future for further betterment and extension of the system.

6.1 CONCLUSION

Object detection and tracking has been a field of interest for research for a long time now, but the detection, tracking, and prediction of high speed moving objects with mobile camera and low resolution images is relatively new with lesser work being done. The camera based guided missile systems have recently became an important topic of research due to the problems associated with the IR based guided missile systems and the use of radar. The algorithm that has been developed in this thesis used Kanada-Lucas-Tomasi feature tracker, random sample consensus algorithm, feature manipulation, noise reduction techniques, local histogram matching, and movement prediction using angular velocity calculations.

KLT feature tracker tracks the features from one frame to the next. The features are scattered throughout the image depending upon the overall texture of the background, and the number of objects in the image. The aircrafts are detected by correlating the matching features in consecutive frames, but KLT alone provides a low accuracy and prediction performances with majors errors in aircraft detection and a huge miss ratio in the presence of highly textured background with a large number of objects. For handling this complete problem, a framework and algorithm has been established. First of all, the number of features that are required to be tracked by the KLT feature tracker are estimated, and that number is fed into the feature tracker to make sure the accuracy and efficiency maintain an equilibrium. The matching features between the two frames are used to compute the transition matrix with the help of random sample consensus algorithm that helps in providing an estimate of the scaling, rotation, and translation of the object. The transition matrix provides an estimate of the new location and the calculation of the angular velocity of the aircraft at any particular time. This helps in predicting the future location of the aircraft within the 2D path under consideration. Certain features are removed and certain new futures are arrived under consideration during the course of calculations on different frames. The features are lost when they cannot be tracked i.e. they are lost due to aircraft rotations, the features belongs to noise in the image or they don't follow the same trajectory as that of the aircraft features i.e. they belong to some other object. The new features are arrived if they match with the image properties of the aircraft features. This matching is done with the dynamic model of the aircraft that is changed on every new frame. The new features are only arrived within the consideration if and only if they belong to the tracking rectangle that bounds that aircraft within itself at all times.

The research contributes the following things towards the improvement in the algorithm for the tracking of aircrafts.

- Presents calculations for predefining the number of features to be tracked for best detection results
- Presents an efficient technique for noise removal without convolution techniques with low pass filters
- Presents method for calculations excessive rotations and scaling of aircraft
- Provides algorithm for the position prediction of an aircraft based on angular velocity after time 't'

6.2 FUTURE WORK

Improvement in angular velocity calculations

The angular velocity is currently estimated in the 2 dimensions framework based on the transition matrix. The 2 dimensional calculations of the angular velocity measures cannot prove high accuracy measures as compared to the 3 dimensional calculation of the angular velocity. Another problem associated with the angular velocity calculation at present is the movement of camera to keep the aircraft within its range at all times. This movement of camera results in low values of accuracy so the testing results can never be accurate.

Calculations of yaw, pitch, and roll angles

The calculation of yaw angle, pitch angle, and roll angle can help estimate the accurate measure of the angular velocity making accuracy of the application increase.

Use of morphological operations and filters

The main focus of this framework has been on the high efficiency of the framework due to which all image operations have been replaced with the feature manipulations. The noise reduction and other operations are performed by manipulating the features of the image detected by the KLT feature tracker. The use of morphological operations will further improve the accuracy of the framework and will make it less application dependent.

Testing on wide data set

The data set could not be found for this application testing due to unavailability across the internet and computer vision datasets. As a result of this unavailability, the testing has been done on a number of videos downloaded from the internet, but the testing on the actual dataset can prove better in estimating the correct measures of the accuracy and the improvements required in the application.

Use in multiple applications

The framework has been developed not only for an aircraft detection system but for any applications that tends to track high speed moving objects with the help of a mobile camera with low resolution images. The focus of low resolution images is maintained due to highly efficient performance of the KLT feature tracker against its counterparts like SURF and SIFT when the number of features to be tracked is less.

Deployment in real-time systems

The efficiency has been improved in the framework developed for this application for the aircraft tracking to be deployed in the real-time systems for testing. The implementation on the real time systems has not been done up till now; the actual implementation on those dedicated hardware based system will actually provide a better estimate regarding the improvements required in the application efficiency.

REFERENCES

[1] Pro’s and con’s of range sensors. http://www.rssc.org/content/pros-and-cons-range-sensors (visited: 11 May 2013)

[2] Bruce D. Lucas and Takeo Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. International Joint Conference on Artificial Intelligence, pages 674-679, 1981.

[3] Carlo Tomasi and Takeo Kanade. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.

[4] Jianbo Shi and Carlo Tomasi. Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition, pages 593-600, 1994.

[5] Stan Birchfield. Derivation of Kanade-Lucas-Tomasi Tracking Equation. Unpublished, January 1997.

[6] Stavros Petridis, Christopher Geyer, Sanjiv Singh. Learning to detect aircraft at low resolution. Robotics Institute, Carnegie Mellon University.

[7] Carnie, R., Walker, R., Corke, P.: Image processing algorithms for UAV Sense and Avoid. In: Proc. ICRA, pp. 2848–2853 (2006)

[8] Jeffrey McCandless. Detection of aircraft in video sequences using predictive optical flow algorithm. Optical Engineering, 3:523-530 (1999)

[9] Y. Tang and R. Kasturi, Tracking moving objects during low altitude flight, *Machine Vision and App.* 9, 20-31 (1996).

[10] The Newton-Raphson Method: http://www.math.ubc.ca/~anstee/math104/104newtonmethod.pdf (visited: 11 May 2013)

[11] Newton-Raphson Method: http://www.macalester.edu/aratra/edition2/chapter3/chapt3a.pdf (visited: 11 May 2013)

[12] Hieu Lu Nguyen. Newton-Raphson method in complex form [power system load flow analysis], IEEE power and energy society, 1997

[13] Marr, D. and Pogglo, T. "A Computational Theory of Human Stereo Vision." *Proceedings of the Royal Society of London* 8-*204* (1979), 301-328.

[14] Moravec, Hans P. Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover. Tech. Rept. CMURI-TR-3, Carnegie-Mellon University Robotics Institute, September, 1980.

[15] Baker, H. Harlyn. Edge Based Stereo Correlation. Proceedings: Image Understanding Workshop, DARPA, April, 1980, pp. 168-175.

[16] Kanade Lucas Tomasi (KLT) Feature Tracker: http://web.yonsei.ac.kr/jksuhr/articles/Kanade-Lucas-Tomasi%20Tracker.pdf (visited: 19 May 2013)

[17] Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. CVPR 2001, pp. 511–518 (2001)

[18] Xiao, R., Zhu, L., Zhang, H.J.: Boosting chain learning for object detection. In: Proc. ICCV, pp. 709–715 (2003)

[19] Bourdev, L., Brandt, J.: Robust object detection via soft cascade. In: Proc. CVPR 2005, pp. 236–243 (2005)

[20] Sung, K., Poggio, T.: Example-based learning for view-based human face detection. IEEE Trans. On PAMI 20(1), 39–51 (1998)

[21] J.L. Barron, D.J. Fleet, and S.S. Beauchemin, "Performance of optical flow techniques," *Inter. J. Comp. Vis.* 12, 43-77 (1994).

[22] N.M Grzywacz, S.N.J. Watamaniuk, S.P. McKee, "Temporal coherence theory for the detection and measurement of visual motion," *Vis. Res*. 15, 3181-3203 (1995).

[23] W. Kruger, W. Enkelmann, S. Rossle, "Real–time estimation and tracking of

optical flow vectors for obstacle detection," *Proc. Intelligent Vehicles '95 Symposium*, 304-309 (1995)

[24] Comparison of the OpenCV's feature tracking algorithms: http://computer-vision-talks.com/2011/01/comparison-of-the-opencvs-feature-detection-algorithms-2/ (visited: 19 May 2013)

[25] B. Etkin, Dynamics of Flight – Stability and Control, John Wiley and Sons, Inc., New York, 1982.)