Real Time Video Mosaicing At Frame Level Fusion Using Different Image Fusion Techniques

S.Mukul Manohar[1], K.N.Muralidhara[2]
[1, 2] Dept of Electronics and Communication
[1]Vemana Institute of Technology
[2]PES College of Engineering, Mandya

Abstract- *In several search and rescue operations, humans cannot carry out the task in person due to physical limitation. In such cases, UAV (Unmanned Aerial vehicle) is used. The unforgiving terrestrial terrain, plants, bushes, trees, animals, etc, makes it difficult to perform search from the ground. It also becomes extremely exhausting and stressful. Thus, aerial bird's eye view is more efficient and simpler. To accomplish this, a camera can be mounted in the aircraft, which can be flown over the region of interest to conduct a quick survey of the neighborhood. On the downside, it is not easy to detect objects of interest in the aerial video surveillance. Tracking an object is even harder. As the aircraft will be moving, the frames or images in the video will be changing. Thus providing a continuity in the video with respect to an object is a challenging task. To perform such tedious task in real-time, video mosaicking is used, which stitches the frames in the video. This creates a continuous picture with fusion of previous frames which helps to track an object with a single image. Therefore, using real-time video mosaicking with UAV offers good efficiency in tracking an object with good resolution. This paper deals with the image fusion to perform video mosaicking for efficient object tracking by UAV. The UAV will be fitted with a transmitter to send the real-time video to the base station, which is majorly used in military, ground mapping, etc.*

Keywords- the UAV aerial images; image mosaic; image fusion; target tracking.

I. INTRODUCTION

In current situation, human life is more precious than anything else, and risking the human life for completing some task is not affordable. Encountering dangers head-on, is the last option. The modern technology offers great possibilities to deal with critical problems without direct human intervention. One such technology is the UAV, which helps in object detection and tracking. It can also be used in mapping different applications. However, getting a continuous video in form of a single large image is a challenging task. But it helps in tracking the target without losing it. This creates a continuous image that can be seen on a single planar image. In order to accomplish this, real time video mosaicing must be implemented.

Real Time Video Mosaicing in UAV is not an easy thing to do. There are several hardships in implementing video mosaicking, as it must be performed at each frame. Furthermore, an algorithm must be designed to carry out such a task. This is a critical part of the system design, as it will be used in military and tracking purposes. This research work focuses on the design, development and implementation of a Real time Video mosaicking and target tracking system for an UAV. The UAV will be transmitting the video wirelessly to the base station.

The contributions of this paper include

- Conduct a comprehensive literature review on various other mosaicing algorithm to help in designing a suitable algorithm for efficient video mosaicing.
- Design the construct the algorithm on a DSP for real time signal processing applications.
- Design and implement an effective tracking algorithm on the same DSP platform which is used for target tracking.
- Run experiments to check the functionality of both the algorithms under known environment conditions and targets.
- Design and equip the UAV with a wireless video transmitting system, to send the capture video to the base station.

II. LITERATURE REVEW

An algorithm was presented by Seyyed Ali Hoseini and Shahram Jafari [1] for video mosaicking of frame sequences using rotating camera. The geometric transformation was taken between the successive frames to construct the mosaiced image. Few additional features are extracted and matched using this algorithm. The parameters are later extracted using estimation of the mapping function.

Nuno Gracias and Jos´e Santos-Victor [2], presented a research on construction of mosaics of the sea bed with high resolution. They proposed an algorithm to assess the camera trajectory. The authors report that this technique should be completely automatic.

Lu Gregory D. Hager and Xiangtian Dai Le [3] introduced a video mosaicing algorithm using warping technique. However, this proposed algorithm suffered from a drawback, due to is slow computational process and latency. The system failed to perform under high speed target tracking.

Shmuel Peleg and Yair Poleg [4], Andrew J. Davison and Steven Lovegrove [5] describes that any type of miss-match or misalignment in the image stitching results in large-scale distortion in the image, rendering it useless. Therefore, the authors have addressed this issue by developing an algorithm which resolves the misalignment problems associated with mosaicing.

Aljoscha Smolic and Thomas Wiegand [6] contend that video processing will not be efficient if the video is not clear, or if it contains some form of distortion. Therefore, the authors have focused on high definition mosaicking for creating more crisp and high quality video mosaics. The use of high definition video is quite handy in object tracking.

Subhasis Chaudhuri, Udhav Bhosle and Sumantra Dutta Roy [7] explains that a video is composed of a series of continuous frames, in which each frame is regarded as an image. The authors have developed an algorithm which develops a single seamless image by aligning a series of spatially overlapped images. This technique resulted in an image with a larger field-of-view, compared to a single image.

III. UAV

UAVs or the Unmanned Aerial vehicle are air-borne devices which possess the ability to fly without an onboard pilot. The UAV are controlled remotely using autonomous or semi-autonomous capabilities. Albeit, UAVs have tremendous scope in modern day application. It is majorly used for image processing applications in surveillance and reconnaissance. UAV is extensively employed in the area of battlefield reconnaissance, natural disaster assessment, military monitoring and surveillance, environmental monitoring, etc. UAV are highly popular due to their low cost, high flexibility, high efficiency in taking low altitude video data with high resolution. For proper expansion of the field of vision, better interpretation, uniform processing and analysis, and deeper study of UAV image information is needed. Most frequently, video mosaicking of adjacent images has to be done to generate a panoramic image. Nevertheless, in comparison to high technological advancement of UAV, the UAV image processing technology is unable to keep up the pace. Hence, video mosaicking of UAV images create a hurdle for its effective utilization. Thus, video mosaicking gained an immense interest from the research community.

Before actually going into mosaicking and image fusion, understanding the nature of UAV image formation is quite important. The images generated by the UAV contains large amount of data, overlapping degree higher characteristic, several small phase amplitude, etc. For easier and better analytical understanding of the data collected by the UAV, a large panoramic image has to be generated. For effective image fusion, image mosaic technology is quite crucial. Image fusion of visual sensed data creates a whole new dimension of making the target tracking more effective. Tracking targets under fog, smoke, etc., are greatly enhanced.
Since image registration is the core technology of image mosaic, the effects and running speed of image mosaic is mainly decided by the precision and efficiency of registration. The focus is put on image registration in the image mosaic technology research at home and abroad. According to different images used in registration process, three categories can be classified: methods based on frequency domain, method based on feature and the method based on gray information.

UAV are embedded with imaging sensor system, which can be operated remotely without any onboard pilot. The image sensor is a small or high resolution camera capable of taking still images or capturing long videos. Some UAVs are also equipped with thermal or infra-red camera for capturing video during night times or during poor visibility. Thermal cameras provide more functionality than regular light-based camera. These onboard sensor are highly sophisticated, light-weight and easily portable.

IV. PROPOSED METHOD

The whole process can be carried out through the following steps:

Image acquisition: The UAV is remotely instructed to pan and rotate the onboard camera to point at the object of interest, and acquire continuous shot in the form of video. It helps in determining the nature of target, its size, position, speed (if moving) texture, color, etc. These information is then transmitted into base station. Based on the instruction given by the base station, the UAV continues to follow the target, and acquire more images.

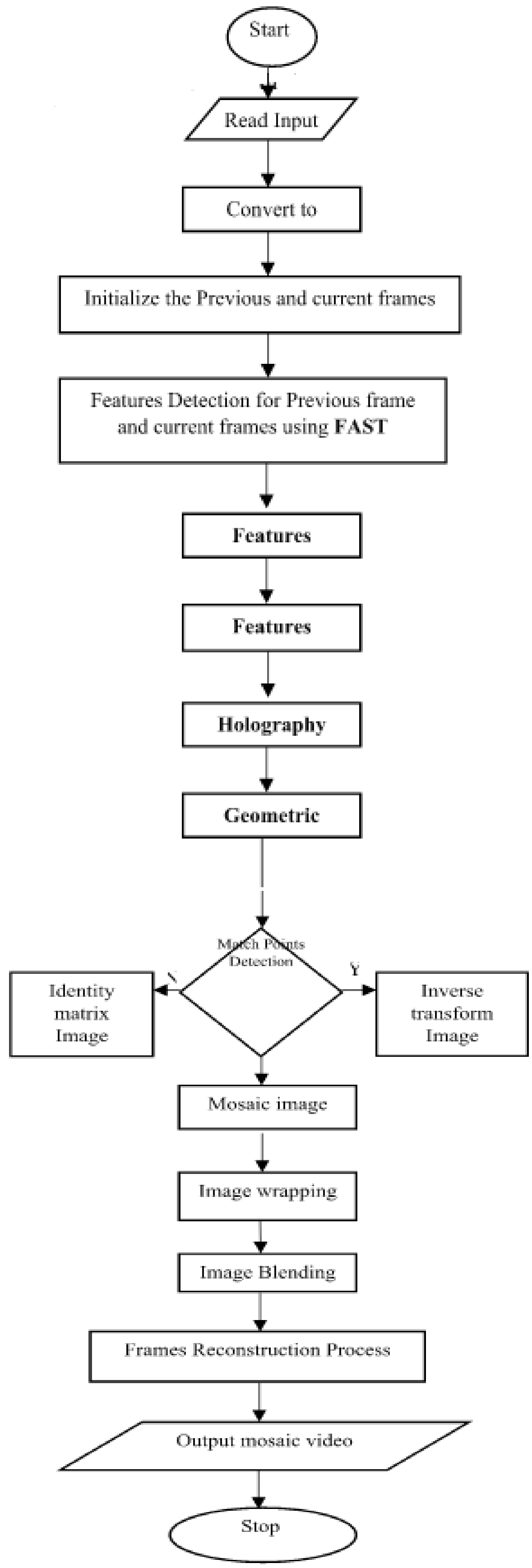
Video mosaicking: The UAV videos is generally composed of moving background due to the motion of the UAV. Hence, successive frames are aligned within a time window to generate a panoramic still image with static background, in order to localize the target more efficiently [8].

Target acquisition: It involves extraction of feature points of the images. These feature points are matched with those features present in the database. If there is a match, then the target acquisition is a success and the tracking process will be initiated. In any other case, the target will be re-localized and this step is repeated.

Target tracking: The consensus-based algorithm is implemented to perform target detection, localization, tracking, motion estimate, size, shape, etc.

UAV control: the altitude and position of the UAV is estimated with respect to the position of the target using the information form the onboard sensors. The movement and direction of the UAV can be controlled based on these parameters, while keeping the target as the fixed focus.

Ground base station: The video and other information or wirelessly transmitted to the base station located on the ground. The base station processes the video and information, controls the UAV and acts as a command center.



A. Image fusion process

Image fusion is an image processing technique, which involves combining two or more images to form a single larger image which incorporates only the important areas of the image, without losing any significant data. Generally, image fusion is needed for images captured from various imaging devices or techniques, which involves same object or region. Image fusion is predominantly used in microscopic imaging, medical imaging, computer vision, remote sensing, robotics, etc. Image fusion can be accomplished by using simpler algorithms like pixel averaging, to more sophisticated techniques like wavelet transform (WT), principal component analysis (PCA), etc. Some of the prominent applications of image fusion are, Medical Diagnosis, Concealed Weapon Detection, Digital Camera Applications, Defect Inspection, Night Time Surveillance, Remote Sensing, Automatic Landing System, etc.

As seen earlier, image fusion generally involves extraction of desired information from two or more images. This process creates a larger image, which conveys more information than the individual images. The limitations of single senor information can be addressed by using multi temporal, multi sensor and multi view information capturing techniques. Image fusion has several benefits. It offers extended spatial and temporal resolution, enhanced range of operation, higher accuracy, reduced uncertainty, good coverage, high reliability and compact representation.

Different imaging techniques such as visual, Infrared, Synthetic Aperture Radar (SAR), etc., uses different types of image fusion. However, these conventional image fusion algorithms suffer from few drawbacks. Direct fusion can result in blurring of images. The contrast of information might diminish by using image averaging method. The pixel based image fusion technique is quite complicated for high resolution images. Hence, a novel and efficient image fusion technique is required, as it plays a critical role in target acquisition and tracking.

Some of other significant applications of image fusion includes target detection, localization, image filtering, etc. Therefore, an efficient and reliable surveillance and reconnaissance system can be implemented with the help of image fusion technology. In some cases, the information in the image can be redundant and/or complementary. A general classification of image fusion approaches are given below. Wavelet Transform, Expectation Maximization, Generic Multiresolution Fusion Scheme, Linear Superposition, Optimization Approaches, Fuzzy Techniques, Image Pyramids, Nonlinear Methods, Artificial Neural Networks, to name a few.

B. Video mosaicking

The image mosaicking technique is composed of several steps. The user will be giving the input for overlapping images which are to be stitched. In this case, the target dictates the image mosaicking. The final image can be obtained after Image registration and merging. Image registration is composed of feature detection and feature matching, which is accomplished by generating a feature detector and feature descriptor [1].

Based on the content in the image, few regions in the image are more common. Hence number of features that can be detected using a given detector may tend to differ for different types of images. Subsequently, feature extraction eliminates redundancy and estimates feature keypoints [5]. The matching keypoints are used to compute feature descriptor. Local features are computed by detecting prominent locations such as, blob, angle and region detectors. Then, respective feature descriptors will be estimated.

Our proposed technique creates a panorama video mosaicking system. In the first step of image registration, the features of image are matched with the help of feature detection. Based on the matching, corresponding result is generated. The resultant image will possess greater quality. Feature matching results in keypoint detectors and keypoint descriptors. More matches are created if the feature matching is good. Depending on the number of matches, image transformation is performed, which also includes homography. Homography is a process of mapping between two spaces which are normally used to represent the correlation between two images of the same scene. In our research work, the homography technique provides a panoramic video with target at the center. Image warping **is** digital image processing which involves distorting any shape present in the image. It is generally used to correct any distortion in the image. Finally, image blending performs concealing the seams at the stitched regions of the image. It creates a seamless view of panoramic stitched images. It is implemented by using feathering, which uses weighted averaging color values to blend the overlapping pixels.

1. **Features Detection**
   Initially, the normalized region is obtained with the resolution 16×16. The 16×16 sub-region will then be subdivided into 16 blocks of a 4×4 resolution each. The 16 blocks are employed to encode the geometric

information. The smaller 4 × 4 subregion will be normalized to the localization error and noise.

2. **Features Extraction**
Based on the image composition, the number of features found with a particular detector may vary, SURF or FAST are the feature extraction technique used to eliminate redundancy. The Surf is based on scale space theory. It creates a stack without down sampling, in order to retain the same resolution. The estimation of local maxima is done with the help of Hessian matrix (H), as shown in equation (1).

3. **Features Matching**
The main objective of feature matching is to search for closest match of the test data in the feature space. The maximum and minimum distances are estimated between the keypoints. Only those good matches whose distance is less than 3*min_dist, are used. The keypoints from good matches are then extracted.

4. **Homograph Process**
Homography is a process of mapping between two spaces, which is generally used to represent the correlation between two images of the same scene. It is calculated by estimating the Probability of L consecutive failures using equation (2),
(2)
Where, is the Probability that a random data item fits the mode.
L is calculated by equation (3),
(3)

5. **Image Warping**
Warping manipulates the image in such a way that any shape in the image will be considerably distorted. The image warping is performed the images which will be mosaicked, using geometric transformation.

6. **Image Blending**
Blending is done by weighing each pixel with the nearest imae boundary, as given by equation ()

C. Tracking of objects in fused images

A visual video is initially acquired from the camera mounted on the UAV. The cameras must have good resolution and must be able to shoot smooth images even under motion. Hence, a light weight, high resolution, strong-featured, stabilized platform, compact and portable sensing equipment must be installed on the airframe. The images should be preprocessed to eliminate atmospheric disturbances, distortion and inherent noise. The visual images can then undergo image fusion. The captured video is then transmitted by the UAV to base station for further processing and analysis. Subsequently, the UAV can be controlled and further instruction can be given.

The captured video from the camera is extracted in .avi file format, which can be processed in MATLAB. MATLAB offers real time processing and efficient algorithm implementation. The idea is to detect the targets in the frames sequences. The individual frames are applied to an optical flow pattern, which creates vectors. This process helps in target detection and localization. The target detection is accomplished based on the feature matching and extraction from the commencing frame. Once the target is found, the corresponding region is effectively extracted. This algorithm optimizes the target detection process, which results in images with minimal noise, compared to other techniques. Once the target is identified, the region will be segmented, and marked. These individual frames are then combined in correct sequence to generate a video file.

Steps for object tracking:

- Obtain the aerial video from the UAV using a visual imaging sensor.
- Transmit the captured video wirelessly to base station on ground.
- The video is split into constituent frames.
- Consecutive images are to be considered with the same field of view (FOV).
- Target is identified and marked.
- Image fusion is initiated.
- Target tracking is initiated in MATLAB.
- Video mosaicking is performed to create a wider background.
- Fused mosaic images are combined to generate a continuous video.
- The resulting video provides real time target tracking.

The **Video Mosaicking Steps are given below;**

- Select and Reading the Input Video Process
- Initialize Video Corners Features Detection Process
- Initialize the Mosaic Parameters
- Video Mosaicking Process
    - Features Detection (FAST/SURF)
    - Features Extraction
    - Features Matching
    - Homograph Process
    - Image Warping
    - Image Blending

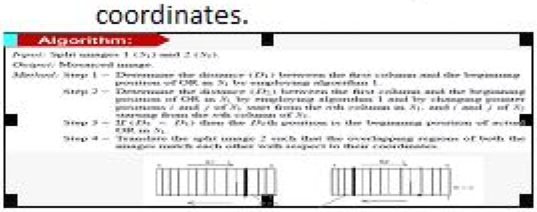
- Output Mosaicking Video

The pseudocode for the proposed mosaicking technique is presented below.

Algorithm: image mosaicing

Input: split the images into 1 (S1) and 2 (S2)
Output: Mosaiced image
Process:

- estimate the distance (D1) between the first column and the starting position of OR in S2 according to algorithm 1.
- estimate the distance (D2) between first column and the starting position of OR in S1 according to algorithm 1. The pointer positions i and j of S1 are changed from nth column in S1, i and j of S2 from nth column of S2.
- if (D2 < D1), then the starting position of actual OR in S1 is D2.
- translate the split image 2 such that the overlapping regions of both the images match each other with respect to their coordinates.



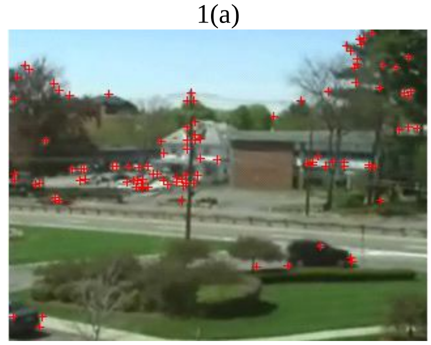
Algorithm 1: image mosaicking

- divide the set of input image into a disjoint sets based on OD location.
- Best quality image is selected in each subset as an anchor image using pair wise registration.
- register the remaining images in the subset to the anchor image with the help of pair wise registration.
- the remaining images are combined using blending technique in order to generate intermediate mosaic using local quality measure.
- repeat steps 3 and 4 to register and combine the intermediate mosaics to generate resultant mosaic.

V. RESULT

The visual images are considered for target tracking. The acquired video is split into its constituent frames. **Figure shows the** video acquired from the UAV, and target detected, respectively. **Based on figures a,b, a fused video mosaic is presented in figure c.**

1(a)



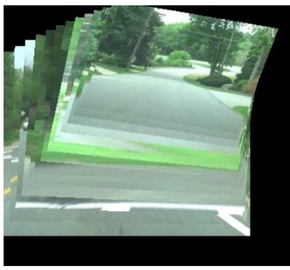
1(b)

1(c)

Figure 1 (a) Video captured by UAV ,1(b) Target detection 1(c) Image fusion and video mosiac

2(a)

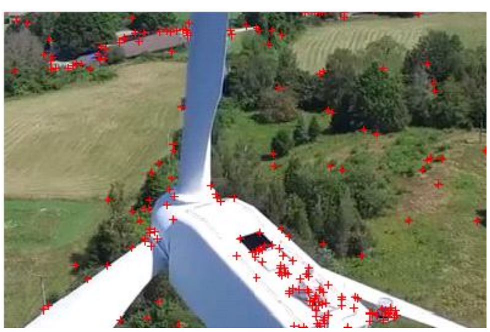
2(b)




2(c) 2(d)

Figure 2 (a) Video captured by UAV (b) Target detection (c) (d) Image fusion and video mosiac

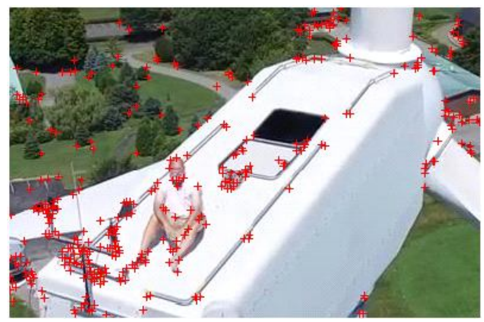
3(a)



3(b)

3(c)

3(d)



3(e)

(f)

Figure 3 (a) Video captured by UAV (b) (e)Target detection (c) (d) (f) Image fusion and video mosiac

VI. CONCLUSION

As per the proposed image fusion and video mosaicking system, experiments are carried out in MATLAB. The aerial video is captured and transmitted to base station. The video is then split into frames. The target is efficiently detected and identified form the images. Image fusion is applied to each frame. We have successfully designed and implemented the algorithm on a DSP for real time signal processing applications. The performance of the target tracking algorithm on the DSP is found to be functioning efficiently. The proposed technique offers great possibility for efficient and reliable target tracking in real-time applications.

REFERENCES

[1] Seyyed Ali Hoseini and Shahram Jafari " An automated method for mosaicing of video frames with projective constraint" International Journal of Science and Advanced Technology (ISSN 2221-8386) Volume 1 No 4 June 2011
[2] Nuno Gracias and Jos´e Santos-Victor " Underwater Mosaicing and Trajectory Reconstruction using Global Alignment" Instituto Superior T´ecnico & Instituto de Sistemas e Rob´otica Av. Rovisco Pais, 1049–001 Lisboa Codex, Portugal
[3] Xiangtian Dai Le and Lu Gregory D. Hager " Real-time Video Mosaicing with Adaptive ParameterizedWarping" CIRL Lab, Computer Science Department, Johns Hopkins University, MD 21218, USA
[4] Yair Poleg and Shmuel Peleg " Alignment and Mosaicing of Non-Overlapping Images" IEEE international conference
[5] Steven Lovegrove and Andrew J. Davison " Real-Time Spherical Mosaicing using Whole Image Alignment" Imperial College London, London
[6] Aljoscha Smolic and Thomas Wiegand " HIGH-RESOLUTION VIDEO MOSAICING" published in ICIP'01, Thessaloniki, Greece.
[7] Udhav Bhosle and Subhasis Chaudhuri and Sumantra Dutta Roy " A Fast Method for Image Mosaicing using Geometric Hashing"IEEE confrence at Indian Institute of Technology Bombay.
[8] Patil, Raju, Paul E. Rybski, Takeo Kanade, and Manuela M. Veloso."People detection and tracking in high resolution panoramic video mosaic." InIntelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on, vol. 2, pp.1323-1328. IEEE, 2004.