International Journal of Electrical and Computer Engineering (IJECE)
Vol. 9, No. 5, October 2019, pp. 3504~3511
ISSN: 2088-8708, DOI: 10.11591/ijece.v9i5.pp3504-3511

Obstacle avoidance and distance measurement for unmanned aerial vehicles using monocular vision

Aswini N., Uma S. V.
Department of Electronics and Communication, RNS Institute of Technology, India

Article Info

Article history:

Received Jan 3, 2019
Revised Mar 22, 2019
Accepted Apr 9, 2019

Keywords:

Distance measurement
Obstacle avoidance
Scale invariant feature transform (SIFT)
Speeded up robust features (SURF)
Unmanned aerial vehicle

ABSTRACT

Unmanned Aerial Vehicles or commonly known as drones are better suited for "dull, dirty, or dangerous" missions than manned aircraft. The drone can be either remotely controlled or it can travel as per predefined path using complex automation algorithm. To make it completely autonomous, the most challenging problem faced by UAVs is obstacle avoidance. In this paper, frontal obstacles are detected using monocular vision by extracting features using Computer Vision algorithms like Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Feature (SURF). Distance of the obstacle from camera is calculated by measuring the pixel variation in consecutive video frames. To meet the defined objectives, designed system is tested with self-developed videos which are captured by DJI Phantom 4 pro.

Corresponding Author:

Aswini N,
Department of Electronics and Communication Engineering,
RNS Institute of Technology,
Dr. Vishnuvardhana Road, R R Nagar Post, Channasandra, Bangalore, 560098-India.
Email: shijiaswini@gmail.com

1. INTRODUCTION

Unmanned Aerial Vehicles, commonly known as drones are now very much popular in both military and civil applications. UAVs have grown significantly covering various applications ranging from surveillance in military to commercial applications like, product delivery, firefighting, precision farming etc. They have unlimited potentials. The growth of UAVs has been remarkable and in the coming years, they are going to be a big success. The various applications of drones and their growing demand in future is shown in Figure 1.

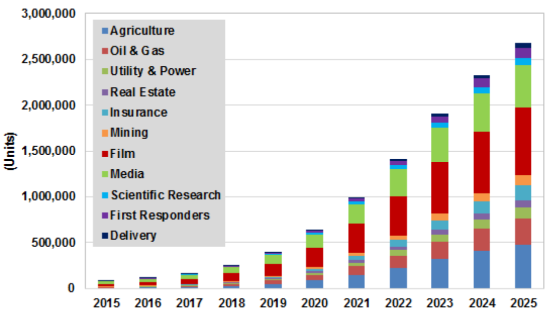

Figure 1. Growing demand of drones in the future

As in [1], there are ten levels of autonomous mission control ranging from automatic mission control (remotely guided aircraft) to partial autonomous mission control (UAV equipped with failure adaptation algorithms) to fully autonomous systems (UAV free from human operator). The commercial applications require the small UAVs to fly at lower altitude or operating inside buildings, where they are exposed to many hazards and obstacles. Current UAV technology in automatically sensing, detecting, and avoiding fixed obstacles such as power line, building, tower, tree, and moving obstacles such as birds, and other aircraft is still immature compared to manned aerial vehicle. So, there is a great scope of research in embedding Sense-Avoid Detect algorithms on board UAV. By acquiring the sensing mechanism, optimum routing of UAV as in [2] can be very effective.

The limitations of carrying heavy weight, costly equipments like Radar, Lidar etc in commercial drones can be overcome by using cameras for obstacle sensing. To address the problems of visual tracking, object recognition and avoidance, the concept of extracting key points is applied in this paper. In section 2, related works on vision-based obstacle avoidance are discussed. The use of computer vision algorithms is a promising solution for easier obstacle detection in real time video processing. Accuracy performance in vision-based navigation depends on object tracking and performance. Feature descriptor algorithms include Speeded Up Robust Features (SURF) [3] and Scale Invariant Feature Transform (SIFT) [4]. Although SIFT has proven to be very efficient in object recognition applications, it requires a large computational complexity which is a major drawback especially for real-time applications. Speed up Robust Feature (SURF) technique, which approximates SIFT, performs faster than SIFT without reducing the quality of the detected points. These two robust feature descriptors are invariant to scale changes, blur, rotation, illumination changes and affine transformation. In the existing works, various techniques for feature extraction and hence detection of obstacles are explained. In our work which is given in section 3, after extracting features, we are measuring the approximate distance between the camera and obstacle by calculating the pixel variation from consecutive video frames. The results given in section 4 shows that the proposed method is an efficient way to avoid obstacles in the path of UAV.

2. RELATED WORKS

Abdulla Al-Kaff et.al [5] proposed a real-time collision avoidance and object detection algorithm for UAV. A singular monocular camera is mounted on vehicle to capture the image. SIFT and Brute Force algorithm is used for generating and matching keypoints. Processing time required is 52.4ms with 62-degree Field of View (FOV). In this work they can detect objects within 90 to 120 cm range.

Levente Kovacs et.al [6] proposed a deconvolution technique to discover the object region to take out features of that object and to create feature map which is usually called as D-map. Monocular camera is employed to capture obstacle with low collision ratio and frequently used in various environments. In Future, in they need to fuse the feature (Map) with other features of image. Methodology used in this work is helpful in navigation system, surveillance, military, mapping and odometry.

Omid Esrafilian et.al [7] proposed a collision avoidance scheme for Aerial Quadrotor (Drone). Video streams recorded using frontal camera and the navigation data measured by Aerial Quadrotor is transmitted to ground station through wireless network connection. Simultaneously Localization and Mapping (SLAM) is helpful in navigation and mapping. The navigation data received is processed by Oriented Fast and Rotated Brief (ORB) and SLAM to compute 3D maps and three-dimensional position of robot. The scaling parameter of monocular SLAM is figured out using linear filtering. Kalman Filter (KF) is used for fusing sensor in monocular camera of Aerial Quadrotor. For controlling three-dimensional position of obstacle, Proportional Integral Derivative (PID) controller is designed. The deisgn of a PID controller for AR. Drone is given in [8].

Jakob Engel et.al [9] explained an UAV navigation method in Global Positioning System (GPS) enabled surroundings. The Quadrocopter based system includes SLAM Algorithm, Extended Kalman Filter (EKF) for fusing the sensor. PID offer steering command, to control and direct accuratly. EKF computes scale map up to ± 1.7% if value is true. This system is applicable for outdoor surroundings having average location accuracy 18.0 cm, indoor environment with average position accuracy of 4.9cm, having an acceptable delay of 400ms. This system offers robustness against visual tracking loss, it eliminates odometry flow due to SLAM, provides accurate navigation estimation.

Yuki Sakai et.al [10] has proposed object detection and tracking systembusing SIFT and SURF methods. In this method the accuracy found for matched key points using SURF algorithm is higher when compared to SIFT. In the future, this technology of detecting a specific object in video or images are expected to further expand to a wide range of applications, like car detection functions for ITS and other systems. Symmetrical SURF detector is introduced by Jun Wei Hsieh et.al [11] to determine the object of interest for vehicles on road. In this method, the SURF algorithm is used to extract vehicle features. For real-time applications, SURF is efficient and along with it eliminates background subtraction. However, the ambiguity issues resulting from vehicles having similar shapes posed a major challenge. To address the ambiguity issues given image is sectored into several grids. Histogram of Gradient (HOG) and SURF is applied on the sectored grid to extract numerical features. Finally, Support Vector Machine (SVM) classifier is used to classify vehicle category.

Ahlem Walha et.al [12] has proposed an object detection and video stabilization system. The data provided by aerial surveillance system suffers variations due to the motion of the camera. To stabilize the aerial surveillance video and to detect the moving object, Kalman filtering and SIFT are used. A matching algorithm is used for feature matching purpose. Random Sample Consensus (RANSAC) estimate dominant object motion and provide outlier to match key points. For object detection, an adaptive clustering algorithm is used. Once moving object is detected, to track the object and to smoothen the movement, Kalman filtering is used. Desired motion is retained using median filtering.

Hailing Zhou et.al [13] presents a vehicle tracking system in UAV videos. A graph cut method is used to extract the specified road in ROI (Region of Interest). Fast feature technique is used to collect the numerical features and Kanade-Lucas-Tomasi (KLT) feature tracker is applied to estimate the motion. RANSAC estimator is preferred to outline the valid key point's features. The preferred system performance is analyzed with drift error and zigzag contour problems. Experimental results show that, this technique provides an effective solution to two problems in UAVs i.e. when the video is captured at low altitudes and with high speeds.

Pouria Sadeghi-Tehran et.al [14] has designed an autonomous template matching based object tracking model. This approach was tested using pre-recorded videos which are taken by AR Drone. Key feature points are detected using FAST detector. Here template matching is performed by matching features of the reference frame with search frame. Brute force algorithm is used for initial feature matching of key points. RANSAC estimator compute fundamental H matrix. RANSAC find inliers and the outliers get eliminated during H matrix computation.

A context-aware motion descriptor (CMD) is designed by Tao Chen et al [15] for detecting object-lvel motion using moving cameras. They calculate the contextual information like optical flow of the image background surrounding the objectof interest. The inconsistency of the histograms between the object and the surroundings is measured.

Wilbert G. Aguilar et.al [16] has proposed real-time micro aerial-based object detection and collision avoidance system. In this method, SURF descriptor extracts the obstacle feature points. These extracted features get compared between the images from the database without incrementing the computational cost. To avoid an obstacle, a control law is implemented. This method is tested in real time on low-cost UAV and the result shows that it effectively detects and avoids the collision.

Trung Nguyen et.al [17] has proposed three-dimensional visual navigation techniques on funnel lane theory for quad rotor vehicle, to overcome drawbacks of KLT feature. To develop funnel lane theory navigation on the quad rotor and to improve the system robustness, SURF feature is used. It is more robust and less computational than KLT feature. Features are tracked using feature matching. Robot Operating System and Gazebo simulator are employed for simulation. In future, this method addresses the problem of visual obstacle avoidance during path following and improve the self-localization problem.

Jagdeep Kaur et.al [18] has proposed video stabilization and moving objects detection module. SIFT and SURF are used as descriptors. While estimating parameters of camera, this method tends to find moving object using Kalman filtering. Referred module recognizes the object and rearranges motion of moving the object into a stabilized position. Here SIFT is used for stabilizing video and to detect moving object. Feature extraction and descriptor matching algorithm are used for camera motion estimation. In future work, this method tries to implement video stabilization techniques which will use more frames or massive video size and extremely quick processing speed.

3. RESEARCH METHOD

The database is created by "*DJI Phantom 4 pro*" [19] drone with 1920x1080 frame resolution. The drone is made to fly 6 feet above from the earth with 1 Meter/sec speed. Considering the different state of the object multiple videos are captured. The data base has been generated by using static and moving car in the field of view of quadcopter. For each database, the angle of view is 70^0 and covered distance is approximately 30 Meter. The minimum difference maintained between drone and the static/moving object is $\cong$ 3.5 to 4.5 Meter. The resolution of the Videos captured are very high for individual frame analysis, hence it is necessary to reduce or convert the frame rate and picture resolution to speed up the analysis rate. An "Apowersoft Video Converter Studio tool" [20] is used to convert high resolution and high frame rate video data into acceptable video sequences (640x480). The given RGB24 video frame is converted into a 2D grayscale level, which makes the frame analysis simple and effective. The level of pre-processing is extended to one more step to smooth object edges in the frame using filtering concept. The median filter is applied to each histogram equalized video sequences to smooth its edge content. The signal to noise ratio variation for various types of noise was analysed and it is seen that median filter is a better choice as shown in Figure 2. The further approach used to detect the obstacle and calculate distance to the camera is as follows:

Algorithm 1: Obstacle Detection and Measurement

Input: Input Video.
Output: Obstacle is detected.

Step1. *Generate video frames.*
Step2. *Apply Histogram equalization and median filtering to remove unwanted noise.*
Step3. *Extract key points of each frame using SIFT and SURF descriptor. (Separately done to compare performance)*
Step4. *Match key points of each frame using Feature Matching Metric Algorithm.*
Step5. *The convex hull is applied around matched key points to create a region of interest.*
Step6. *if current frame convex hull size is greater than the previous frame then obstacle is detected*
Step7. *Estimate change in position of previous with respect to the current frame in pixel unit using Euclidean distance.*
Step8. *Calibrate distance between object and camera.*

End of Algorithm

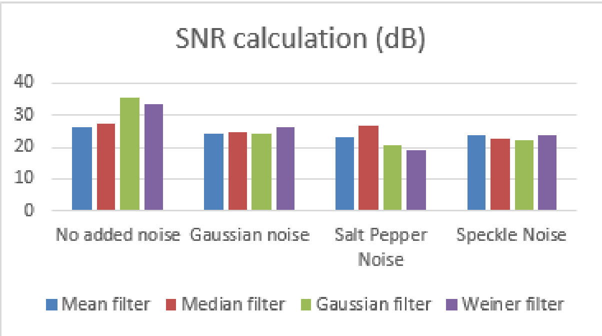

Figure 2. Signal to Noise Ratio calculated with and without added noise

For feature extraction from consecutive frames, we have tried with both SIFT and SURF algorithms. The various steps followed in SIFT and SURF descriptor algorithm for the detection of key points is given in Figure 3 and Figure 4.

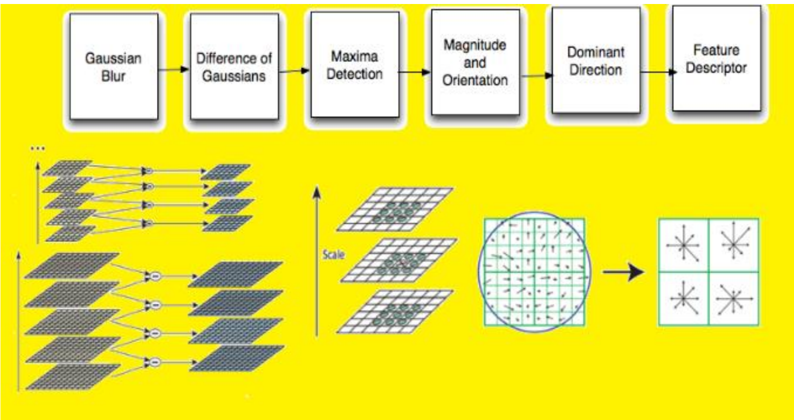

Figure 3. SIFT Algorithm steps

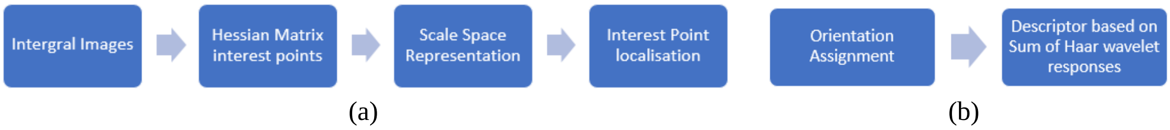

Figure 4. (a) SURF interest point detection, (b) SURF description of interest points

Once the key points are detected, next step is to calculate the distance of object frame. The x, y coordinates of both current and previous frames matched key points are extracted for pixel co ordinate difference calculation.

$$Pixel\ Co - ordinate\ Diff. = \sqrt{(Xc - Xp)^2 + (Yc - Yp)^2} \quad (1)$$

where (Xc, Yc) are co ordinates of current frame and (Xp, Yp) of previous frame. This calculation is repeated for the entire matched key point's coordinates (i.e. both current and previous frame). The coordinates difference vector size equals the number of matched key points in both frames. The mathematical representation is given in Equation (2).

$$Pixel\ Co - ordinate\ Diff.Vector\ size = 1x\ No.Matched\ key\ Points \quad (2)$$

Once the difference vector is formed, a mean of the difference vector is estimated approximately the pixel difference between successive frames. This is measured in terms of pixel difference and this difference is directly proportional to change in object state. Mathematical equation involved in approximating the pixel difference (Diff.) between matched key points is shown in Equation (3).

$$Pixel\ Diff. = mean(Pixel\ Co - ordinate\ Diff.) \quad (3)$$

This pixel difference between the matched key points of both current to the previous frame is used to measure the distance between object and UAV.

The basic principle involved in distance measurement between the moving object and camera is shown in Figure 5. Consider the moving object '*N*' with height '*H*'. Assume that the initial distance between the moving object and camera is '*P*'. At distance '*P*' the image height formed is '*A*', similarly, after covering a distance the image height will be '*B*'. From Figure 4, change in image height will be directly proportional to the distance covered by a physical object.

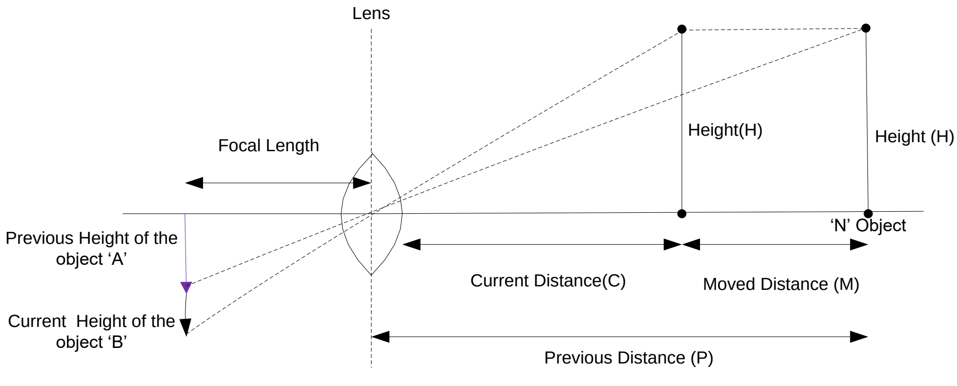

Figure 5. Distance measurement

This principle is applied to the proposed system to measure the distance between and moving object. From the camera calibration, we calculated that approximately 3.31985 Pixel difference of current to previous frames denotes the 1 Meter change in obstacle state. The analysis is performed using key points extracted by SIFT and SURF algorithms. So the pixel variation is used to calculate the change in obstacle state in meters.

4. RESULTS AND ANALYSIS

The time taken for feature extraction in the case of SURF and SIFT are given in Table 1. Figure 6 shows the extracted keypoints for consecutive frames using SURF descriptor. From the Table , SURF took bit more time to extract feature compared to SIFT but the ratio of matched key points in i^{th} and $i^{th}+15$ frames are maximum with respect to SIFT. Due to the extraction of the strongest key points from both the frame, system accuracy is increased in finding the pixel difference between current and previous video frames.

Tabel 1. SIFT and SURF performance comparison

Algorithm	Time Taken for Feature Extraction in i^{th} Frame	Time Taken for Feature Extraction $i^{th}+15$ Frame	Extracted Key Points for i^{th} Frame	Extracted Key Points for $i^{th}+15$ Frame	Matched Key points between i^{th} and $i^{th}+5$ frame
SIFT	0.014834 Sec	0.013015 Sec	1792	1664	5
SURF	0.010106Sec	0.013177 Sec	40	37	14

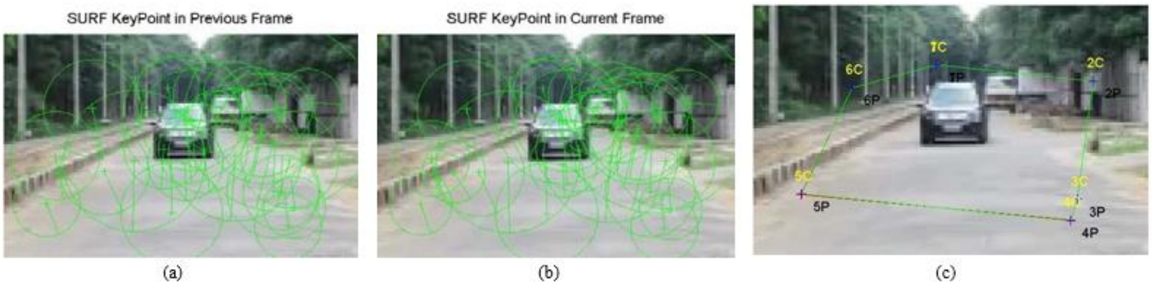

Figure 6. (a) SURF key points at previous frame; (b) SURF key points at current frame; (c) Matched key points and convex hull in both previous and current frame

Figure 7 and Figure 8 gives the various stages of distance measurement for static and dynamic objects. This method of detection and calculation of obstacle distance using monocular vision can be incorporated on board the UAV. It is a better option in terms of size, weight and cost. A high-speed Graphic Processor Unit (GPU) can do the operation at a tremendous speed for the safe navigation of UAVs

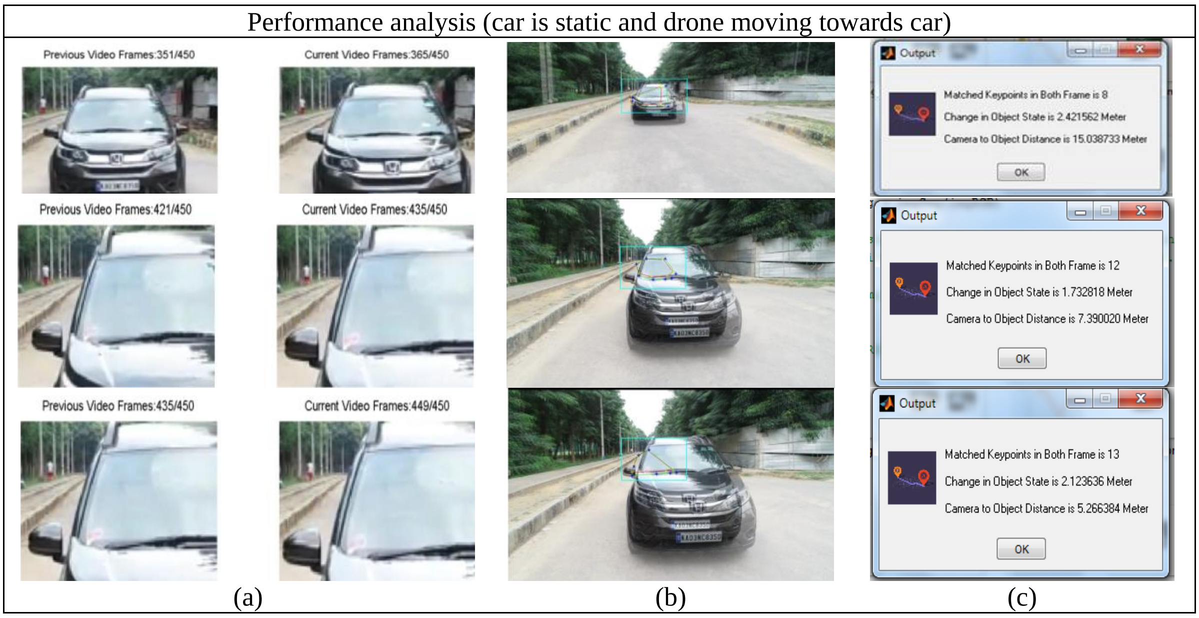

Figure 7. (a) Previous and current frame; (b) Matched key points; (c) Measured distance between static object and UAV

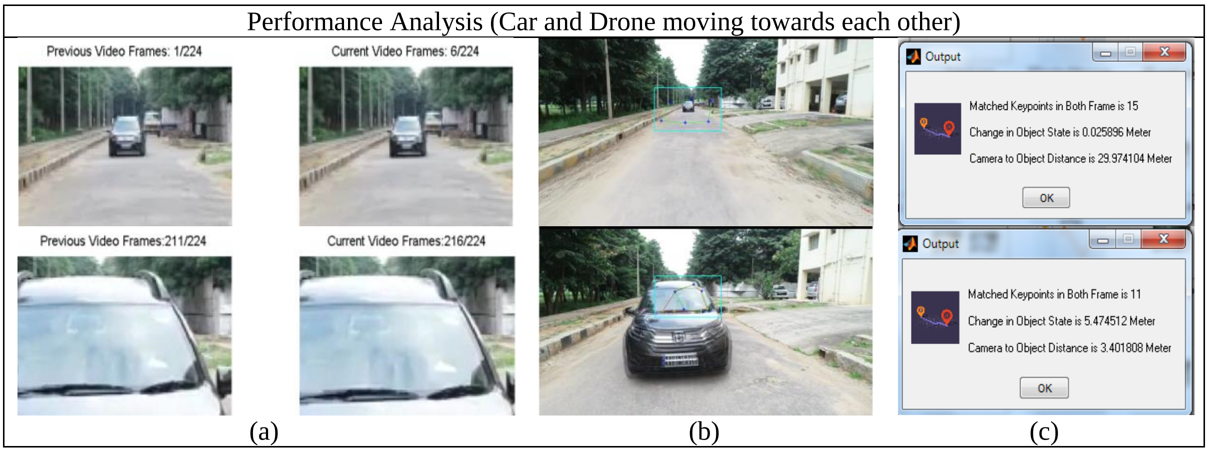

Figure 8. (a) Previous and current frame; (b) Matched key points;
(c) Measured distance between dynamic object and UAV

5. CONCLUSION

In the proposed work, SURF descriptor algorithm effectively calculates the key points of interest at a faster rate than SIFT. Using camera calibration techniques and pixel difference calculation, the distance between camera and obstacle is measured. The whole process takes less than one second, which is very efficient in real time obstacle avoidance point of view. The further work is to incorporate Convolutional Neural Networks (CNN) based detection of obstacles which are approaching from sideways and maneuver the drone accordingly.

REFERENCES

[1] M. Suresh and D. Ghose, "Role of information and communication in redefining UAV autonomous control levels," *Proc. IMechE, Part- G: Journal of Aerospace Engineering*, vol. 224, pp. 171-197, 2010.
[2] A. Fadavi, et al., "Offine/Online Optimum Routing of a UAV using Auxiliary Points," *International Journal of Electrical and Computer Engineering (IJECE)*, vol/issue: 7(1), pp. 392-401, Feb 2017.
[3] H. Bay, et al., "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol/issue: 110(3), pp. 346-359, 2008.
[4] D. G. Lowe, "Distinctive Image features from scale invariant KeyPoints," *International Journal of Computer Vision*, vol/issue: 60(2), pp. 91-110, 2004.
[5] A. Al-Kaff, et al., "Monocular Vision-Based Obstacle Detection/Avoidance for Unmanned Aerial Vehicles," *Intelligent Vehicles Symposium (IV), IEEE*, 2016.
[6] L. Kovács, "Visual Monocular Obstacle Avoidance for Small Unmanned Vehicles," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pp. 59-66, 2016.
[7] O. Esrafilian and H. D. Taghirad, "Autonomous Flight and Obstacle Avoidance of A Quadrotor by Monocular SLAM," *in Robotics and Mechatronics (ICROM), 4th International Conference on IEEE*, pp. 240-245, 2016.
[8] A. Prayitno, et al., "Fuzzy Gain Scheduling PID Control for Position of the AR. Drone," International *Journal of Electrical and Computer Engineering (IJECE)*, vol/issue: 8(4), pp. 1939-1946, Aug 2018.
[9] J. Engel, et al., "Camera-Based Navigation of A Low-Cost Quadrocopter," *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on IEEE*, 2012.
[10] Sakai and Yuki, "An Object Tracking System Based on SIFT and SURF Feature Extraction Methods," *in Network-Based Information Systems (NBIS), 2015 18th International Conference on*, pp. 561-565, 2015.
[11] J. W. Hsieh, et al., "Symmetrical SURF and Its Applications to Vehicle Detection and Vehicle Make and Model Recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol/issue: 15(1), pp. 6-20, 2014.
[12] A. Walha, et al., "Video Stabilization with Moving Object Detecting and Tracking for Aerial Video Surveillance," *Multimedia Tools and Applications, Springer*, vol/issue: 74(17), pp. 6745-6767, 2015.
[13] H. Zhou, et al., "Efficient Road Detection and Tracking for Unmanned Aerial Vehicle," *IEEE Transactions on Intelligent Transportation Systems*, vol/issue: 16(1), pp. 297-309, 2015.
[14] P. S. Tehran and P. Angelov, "ATDT: Autonomous Template-Based Detection and Tracking of Objects from Airborne Camera," *in Intelligent Systems, Springer*, pp. 555-565, 2014.
[15] T. Chen and S. Lu, "Object level motion detection from moving cameras," *IEEE Transactions on Circuits and Systems for Video Technology*, vol/issue: 27(11), pp. 2333-2343, Nov 2017.
[16] W. G. Aguilar, et al., "Obstacle Avoidance for Low-Cost UAVs," *in Semantic Computing (ICSC), 2017 IEEE 11th International Conference on*, pp. 503-508, 2017.

[17] T. Nguyen, et al., “Vision-Based Qualitative Path-Following Control of Quadrotor Aerial Vehicle with Speeded-Up Robust Features,” *in Computer and Robot Vision (CRV), Canadian Conference on IEEE*, pp. 321-327, 2014.
[18] J. Kaur and A. K. Bathla, “Video Stabilization for an Aerial Surveillance System using SIFTS and SURF,” *in Next Generation Computing Technologies (NGCT), IEEE 2nd International Conference*, pp. 742-747, 2016.
[19] https://www.dji.com/phantom-4-pro/info
[20] https://www.apowersoft.com/faq/video-converter-studio-guide.html.

BIOGRAPHIES OF AUTHORS

N Aswini is a research scholar, in Department of Electronics and Communication, RNS Institute of Technology, Bangalore, India. She has done her Masters in VLSI Design and Embedded systems and has more than 10 years of experience in teaching. She has guided various post graduate and under graduate student projects. Her area of research is Obstacle sensing, detection and avoidance for Unmanned Aerial Vehicles under Visvesvaraya Technological University, Karnataka, India.

S V Uma is presently working as Associate professor in Department of Electronics and Communication, RNS Institute of technology, Bangalore India. She has done her Doctorate in the area ‘Congestion control and Improved QoS in multimedia networks from Bangalore University. She has nearly 20 years of teaching experience and her interests include Communication, Network Security, Signal, and Image Processing.