Cooperative Machine-Vision-Based Tracking using Multiple Unmanned Aerial Vehicles

Ali Pirshayan¹, Hadi Seyedarabi², Siamak Haghipour³

¹ Department of Mechatronics Engineering, Tabriz Branch, Islamic Azad University
Tabriz, IRAN
a.pirshayan@yahoo.com

² Faculty of Electrical and Computer Engineering, University of Tabriz
Tabriz, IRAN
seyedarabi@tabrizu.ac.ir

³ Department of Biomedical Engineering, Tabriz Branch, Islamic Azad University
Tabriz, IRAN
haghipour@iaut.ac.ir

Abstract
This paper presents an automatic cooperative tracking of the interested target using two quad-rotors UAVs equipped with stereo vision systems. The system includes vision-based algorithms for searching and detecting of target on the video stream. The data processing algorithm used to estimate the position of the target, in addition a method offered based on Kalman filter. This method applied to fusion of each UAV’s estimation together. This combination’s result increased accurate estimation of target location comparing with single ways.

Keywords: Unmanned Aerial Vehicles, Machine vision, Kalman Filter, Data Fusion, Vision-Based Tracking, Cooperative.

1. Introduction
Unmanned aerial vehicles (UAV) have attracted considerable attention in recent years. UAVs are applied in pilotless systems. These systems are well suited to use in a wide range of missions and operations that dangerous or difficult for human, such as search and rescue operations after natural disasters, as well as inspection and surveillance in military or urban areas. In these operations, the use of video cameras and machine vision techniques is special interest due to the wealth of information they can provide. Researchers have used on-board camera for remote sensing of forest fires [1-4]. Colfman and McCord groups developed a machine vision-based roadway traffic monitoring system by using fixed-wing unmanned aerial vehicle [5]. In addition a number of researchers have used vision-based techniques for target detection and tracking. The use of vision for estimation and tracking of ground based targets has also been studied by Bethky and Yorks groups [6], [7]. Using of a quad-rotor UAV with an on-board camera for autonomously identifying and tracking of marine animals has been developed in another paper [8]. Application of video cameras on unmanned aerial systems for tracking targets, poses several challenges. In first step, an algorithm must be proposed to identify targets of interest, in second step, if the target is found, its location must be estimated. Finally, the UAV may use this information to track the target.

When the target moves or the environment has some obstacles, the ability to estimate the exact location of a target with single UAV could be limited. Multiple systems for tracking targets can be useful and provide a more accurate estimation of the target position.

This paper, proposes an automatic vision-based estimation and tracking algorithm that takes advantage of the cooperation between two quad-rotors UAV that equipped with a stereo-vision system. This system contains two critical parts 1) computer vision technics for stereo vision based depth estimating and 2) a computer vision algorithm for detecting and identify the target. Kalman filters are often employed to estimate object state [9]. Designed system has a Kalman filter based Algorithm to estimate the target location in the environment. To provide a more accurate result, cooperative system proposed. Combining estimation of UAVs will achieve the desired objective of accuracy.

The paper is organized as follows. Section 2 presents a method for identifying and estimating target position using stereo cameras. Section 3 explains how to locate the target in environment. Section 4 describes a method for combination of UAVs estimated data. Section 5 explains the tracking system. Section 6 presents the experiments and the results obtained. The conclusions and future trends in section 7 complete the paper.
2. Machine Vision

The goal of the proposed machine vision system is searching and identifying the target in the video stream and also estimating target position relative to the camera. At first, images are garbed from a stereo camera with BGR color system (OpenCV library standard color system). In order to implement a stereo depth estimation with a reasonable level of accuracy, it is important to calibrate the camera system. It is a process of finding the intrinsic and extrinsic parameters of the camera. Because of different light conditions prevailing in various parts of the test environment, images are converted to the HSV color system. Thresholds are defined in each image channel for improving robustness of the target recognition system in various light conditions. After thresholding process, there is still some noise on the binary image. The median filter is used for removing noise from the image and smoothing. After multiple steps of pre-processing on the image, it is necessary to find the external contours of interested target to compute the center of the target in the images. When the center of the target in both images is obtained, the position of the target relative to the camera can be estimated using the Eq. (1), (2) and (3) [10].

Fig. (1) shows the diagram of proposed machine vision system.

\[ Y_{\text{cam}} = y_t \cdot \frac{Z}{f} \]  

Equation (3)

\[ X_{\text{cam}} = x_t \cdot \frac{Z}{f} \]  

Equation (2)

\[ Y_{\text{cam}}, X_{\text{cam}}, Z_{\text{cam}} \]  

are coordinates of the target relative to the camera, B is the stereo camera based line and f is the focal length of camera and \( x_t, y_t \) Represent coordinate of target on the left image plane.

3. Target Localization

This section explains about generating an estimation of the target position in the testing environment. To get the coordinates of target in the test field, we need to have coordination of UAV in the environment. This information can be obtained via the localization sensors which are mounted on the UAV. Also a simulated inertial measurement unit (IMU) mounted on each drone to obtain the rotational angles.

The target location can be expressed using Eq. (4)

\[
\begin{pmatrix}
X_{\text{target}} \\
Y_{\text{target}} \\
Z_{\text{target}}
\end{pmatrix} = \begin{pmatrix}
x_{\text{uav}} \\
y_{\text{uav}} \\
z_{\text{uav}}
\end{pmatrix} + \begin{pmatrix}
x_{\text{cam}} \\
y_{\text{cam}} \\
z_{\text{cam}}
\end{pmatrix}
\]  

Equation (4)

\( x_{\text{target}}, y_{\text{target}} \) and \( z_{\text{target}} \) represent the position of the target in the environment, \( x_{\text{uav}}, y_{\text{uav}} \) and \( z_{\text{uav}} \) are the position of UAV and \( x_{\text{cam}}, y_{\text{cam}} \) and \( z_{\text{cam}} \) are the position of the target in the camera image, respectively. The rotation around Z axis is expressed using Eq. (5).

\[
\begin{align*}
x' &= x \cos \psi - y \sin \psi \\
y' &= x \sin \psi + y \cos \psi \\
z' &= z
\end{align*}
\]  

Equation (5)

\( x, y, \) and \( z \) are the coordinates of the target on the image plane, and \( x', y' \) and \( z' \) are target coordinates by applying UAV rotation around the z axis, which is characterized by \( \psi \).

It assumes that the position of the UAV and target position measured relative to the camera, are not perfectly accurate due to the sensors noise or other artifacts. It leads to uncertainties in both knowledge of UAV and target location. In this work Kalman filter was applied to reduce this uncertainty.

This paper uses system model with state vector according to Eq. (6):

\[ X = [x, y, z, \dot{x}, \dot{y}, \dot{z}] \]  

Equation (6)

Process model and measurement model are then given by Eq. (7) and (8).
\[ X_{k+1} = AX_k + w_k \]  
\[ Y_k = HX_k + v_k \]

\( w_k \) and \( v_k \) are process noise and measurement error of the system model. \( A \) denotes the state transition matrix and \( H \) is measurement matrix. \( A \) and \( H \) given by Eq. (9), (10):

\[ A = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \]  
\[ H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \]  
\[ (9) \]

\[ (10) \]

4. Data Fusion

In this section, a data fusion algorithm for creating a cooperative estimated of the interested target position is developed. The basic idea of the proposed fusion algorithm is to combine local estimate of multiple UAV’s. The aim of the algorithm is to provide an accurate target state estimate and to allow better tracking of the target. Ideal tracking is not needed for a single vehicle to perform maneuvers to achieve a better vantage point.

In this system, observation of each UAV provides an estimation of target position by the local Kalman filter and this estimation is input of a global Kalman filter for fusion of other observations. Fig. (2) displays the schematic of the data fusion system.

5. Tracking System

In order to track a target, the target location is transmitted to the decision-making unit. According to the location and position of the UAVs and target direction, the control commands are generated and sent to UAVs.

The commands production is considered in two cases. First, the arrangement of the drone must have some conditions. They should not collide with each other. A default arrangement for this purpose is considered. The second objective is that the target must always be in the camera field of view. In this case, UAV visually locks on target. Fig. (3) shows scheme of the target tracking system using machine vision.

6. Results

In this section the results of tracking algorithm are presented. In order to increase the similarity of an actual camera image in simulating environment, Gaussian noise with mean \( \mu = 0.1 \) and variance \( \sigma^2 = 0.01 \) is added. Also the Gaussian white noise with signal-to-noise ratio (SNR = 25 dB) is combined by UAV location coordinates data in the simulation environment. The webots software package used to verify the performance of stereo vision estimation and tracking system. For simulating system, two quad-rotor UAV equipped with a stereo vision system was used as the test platform. The UAVs are simulated version of Airrobot drone. A small UGV robot was used as a target on the ground. In the first experiment, the

---

Fig. 2 The diagram of data fusion system

Fig. 3 Scheme of the target tracking system using machine vision
objective of drones was two goals. Hovering with the target in the field of view and estimating the target position in test field. This experiment performed with a single drone and also this test repeated cooperatively with two UAV.

Fig. (4) indicates the configuration of the experiment using a screenshot from testing environment.

The result of the first experiment with single and two UAV are shown in Fig. (5) and Fig. (6). The plot shows the x-y position of the target and black rectangle illustrates the target actual position. Table 1 and Table 2 are demonstrating Standard deviation and Mean absolute error of the estimation.

The estimated location of the target for X and Y axis remains within 0.016 m and 0.01 m of its true location and its great result considering that the target itself is over 0.2m length.

The second experimental result shows the benefits of cooperation rather than the single UAV solution for tracking scenario. In this experiment, the drones hover on the moving target. The drones are trying to track and keep the target in the camera field of view. It should be noted, the ground vehicle moved to outside of drons field of view, once the UAVs not moved on with it.

This experiment was accomplished with single UAV and cooperatively by two UAVs. Fig. (7) and Fig. (8) demonstrate the position of the target in traveled path and black rectangle described the actual path, respectively. Table 3 and 4 are displaying Standard deviation and Mean absolute error of the estimation.

The results demonstrate that the UAVs were able to estimate the position of ground vehicle with reasonable precision (within 0.02 m of its true location) even in case they were moving together in order to keep the vehicle in the field of view of two UAVs.

Table 1: Standard deviation (first experiment) via single and cooperative methods

<table>
<thead>
<tr>
<th>Method</th>
<th>X (m)</th>
<th>Y (m)</th>
<th>Z (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>0.031911</td>
<td>0.034547</td>
<td>0.032965</td>
</tr>
<tr>
<td>Cooperative</td>
<td>0.016195</td>
<td>0.010007</td>
<td>0.020809</td>
</tr>
</tbody>
</table>

Table 2: Mean absolute error (first experiment) via single and cooperative methods

<table>
<thead>
<tr>
<th>Method</th>
<th>X (m)</th>
<th>Y (m)</th>
<th>Z (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>0.026801</td>
<td>0.046052</td>
<td>0.064371</td>
</tr>
<tr>
<td>Cooperative</td>
<td>0.013394</td>
<td>0.007586</td>
<td>0.050459</td>
</tr>
</tbody>
</table>
The third experiment shows the tracking of target in bumpy path. Note that the estimator also provides the Z position of the target. In this experiment target moves up the ramps with different height. Fig. (9) and Fig. (10) shows the target position in traveled path and black vector is actual direction. This experiment was accomplished with single UAV and cooperatively by two UAVs. Table 3 illustrates Standard deviation of estimation and Table 4 is demonstrating Mean absolute error of the estimate. According to the results it can be found that the system was able to estimate the position of ground vehicle with good precision and location of the target for Z axis remains within 0.028 m of its true location. The experiment results show that cooperative system remains very accurate in all three dimensions.

Table 3: Standard deviation (second experiment) via single and cooperative methods

<table>
<thead>
<tr>
<th>Method</th>
<th>X (m)</th>
<th>Y (m)</th>
<th>Z (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>0.038997</td>
<td>0.044139</td>
<td>0.105924</td>
</tr>
<tr>
<td>Cooperative</td>
<td>0.021164</td>
<td>0.026268</td>
<td>0.057198</td>
</tr>
</tbody>
</table>

Table 4: Mean absolute error (second experiment) via single and cooperative methods

<table>
<thead>
<tr>
<th>Method</th>
<th>X (m)</th>
<th>Y (m)</th>
<th>Z (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>0.031820</td>
<td>0.051173</td>
<td>0.089985</td>
</tr>
<tr>
<td>Cooperative</td>
<td>0.016778</td>
<td>0.023611</td>
<td>0.044774</td>
</tr>
</tbody>
</table>
Table 5: Standard deviation (Third experiment) via single and cooperative methods

<table>
<thead>
<tr>
<th>Method</th>
<th>X (m)</th>
<th>Y (m)</th>
<th>Z (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>0.038295</td>
<td>0.031898</td>
<td>0.052920</td>
</tr>
<tr>
<td>Cooperative</td>
<td>0.017074</td>
<td>0.017838</td>
<td>0.028494</td>
</tr>
</tbody>
</table>

Table 6: Mean absolute error (Third experiment) via single and cooperative methods

<table>
<thead>
<tr>
<th>Method</th>
<th>X (m)</th>
<th>Y (m)</th>
<th>Z (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>0.031035</td>
<td>0.038741</td>
<td>0.041519</td>
</tr>
<tr>
<td>Cooperative</td>
<td>0.013627</td>
<td>0.013378</td>
<td>0.022943</td>
</tr>
</tbody>
</table>

7. Conclusions

This work proposes a vision-based system for cooperative tracking and following of a dynamic target. The system uses stereo cameras for identifying and estimating target position relative to the camera. This work uses a Kalman filter for estimating the target position and also for fusion of each UAV’s estimated data. Simulation results verify the performance of the system in cooperatively tracking of target and offer more accurate result against single UAV and robust in face of estimation errors cause location sensors and other factors. As future work we would like to test the system in hardware in loop system and on a real UAV.

References