

A Cumulative Error Suppression Method for UAV Visual Positioning System based on Historical Visiting Information

Ye Chen, RuHui Huang, and Yi Zhu

Abstract—The cumulative error is one of the main factors affecting the accuracy of visual positioning system. How to effectively suppress it has become a key issue for indoor navigation of unmanned aerial vehicle(UAV). In this paper, a new cumulative error suppression method based on historical visiting information (CES-HVI for short) is proposed, which uses the historical visiting points in flight as virtual landmarks to correct the positioning result of UAV. Experiment results show that the CES-HVI method can suppress the cumulative error in effect, improve the accuracy of visual positioning without pre-deploying landmarks in advance.

Index Terms—Accumulative error, historical visiting point, Visual positioning, UAV.

I. INTRODUCTION

TODAY, visual positioning has attracted more and more attention in indoor positioning area. As a special way, visual positioning determines the location of object only using captured environment images by camera. Compared with GPS way, it can avoid the problem of indoor wireless signal shielding. Compared with inertial navigation system, it has higher positioning precision. Therefore, now visual positioning is widely used in the indoor applications of unmanned aerial vehicle(UAV)[1], [2], [3]. As early as 2008, Iowa state university has already realized the autonomous flight of UAV in a regular corridor using Inertial Measurement Unit (IMU) and camera[4]. In 2010, With the support of German ministry of defense, Markus Kleinert et al. have successfully implemented the single-ocular visual positioning based on the Simultaneous Localization and Mapping algorithm and IMU data[5]. In the same year, Jeffrey Byrne, a researcher of Scientific Systems Company, proposed a way of estimating the barrier distance by the comprehensive analysis of image feature points and IMU data, and then successfully applied in the indoor flight experiment without GPS[6], [7].

However, today visual positioning technology still faces some challenges. One of the key problems affecting the positioning accuracy is the accumulative error[8], [9]. For visual positioning system, the current positioning result depends on the estimation of its previous location, so the accumulative error will be persistently increased as flight time goes on.

Manuscript received March 22, 2017; revised August 4, 2017. This work was supported by the Innovation Program for Graduate Student of Jiangsu Province (Grant No.SJZZ15_0133).

Ye Chen is with the School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China (phone: 86-182-5258-6463; e-mail: 185306947@qq.com).

RuHui Huang is with School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China (e-mail: 326304162@qq.com).

Yi Zhu is with the School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China (e-mail: zhuyi@ujs.edu.cn).

When the accumulative error becomes too large, it will lead to a significant shift of reconstruction path, especially for long distance flight of UAV. Due to the accumulative error affects the positioning accuracy greatly, so how to suppress it becomes the key to guarantee the running of visual positioning system. A traditional way of solving this problem is to set global landmarks with known coordinate in the flight line. Zhao Qi proposed a method for error mitigation based on adjustments of navigation parameters, which can reject the concussion error related to the rotational angular velocity[10]. Glenn P. Tournier et al. corrected the visual positioning result of Quadrotor UAV using landmark of Moire Patterns. This type landmark is composed of two pieces of parallel grating, it can enlarge UAV motion captured by camera and then improve the estimation precision[11]. Eduardo Rondon realized a UAV indoor navigation system by pre-deploying landmarks along the flight line[12]. Liu zhenyu designed a special kind of landmark with extensible color coding, the experiment results proved the effectiveness of this method[13]. Chen mingya improved the precision of monocular visual positioning system using natural landmarks under the situation of weak GPS signal[14]. Anna Gorbenko considered the landmarks selection problem of ρ -minimum overlapping region decomposition and then designed greedy algorithm to solve it[15].

Although these above works can relieve accumulative error problem for long distance flight of UAV, the pre-deployment of landmarks is impossible for most UAV application situations. Focusing on this limitation, how to design more feasible suppression method has become a hot topic.

In this paper, we consider the characteristic of flight path in UAV indoor cruise scenario and then propose a new method, its name is a cumulative error suppression method based on historical visiting information (abbreviated as CES-HVI). CES-HVI uses historical visiting points as virtual landmarks to correct accumulative error. Due to no physical landmarks needed to be pre-deployed, it has excellent adaptiveness for unknown environment than traditional way. The implementation of CES-HVI can be divided into two stages. First stage, a trusted historical visiting set is defined, its members (historical visiting points) must be satisfied two conditions, one is within a neighbor area of current location, other is with high image feature matching rate of current image. Second stage, CES-HVI traverses the trusted historical visiting set and estimates current coordinate using current image and each historical image one by one, then output the weighted sum of all estimated coordinates as corrected location result, where the weight is the matching rate of current image and historical image. Benefited from the extra information

of these virtual landmarks (historical visiting points), the accumulative error can be effectively suppressed by CES-HVI. This conclusion is proved by our flight experiments.

The remainder of this paper is organized as follows. Section II introduces the classic visual positioning algorithm. Section III details the proposed CES-HVI method. Section IV evaluates the performance of CES-HVI using experiments. Section V summarizes this paper and points out future work.

II. STEREO VISUAL POSITIONING OF UAV

A. System architecture

Due to the background of this paper is UAV application, first we should establish the basic understanding of UAV system with stereo visual odometry. The structure of UAV visual positioning system is shown in Fig.1, it is consisted of two cameras, ARM board, flight controller (typical type is Pixhawk). All above components are installed in an aircraft, e.g. quadrotor UAV.

During flight, the environment images are periodically captured by camera and then send to ARM board. Using the visual processing algorithms running in ROS (Robotic Operating System), the relative rotation and displacement during a short time interval can be calculated by analyzing two images captured in current and previous instants. Based on the results of rotation and displacement, the real-time position is estimated and fed to the flight controller. Then the flight controller uses it to adjust the flight attitude and realizes indoor visual navigation.

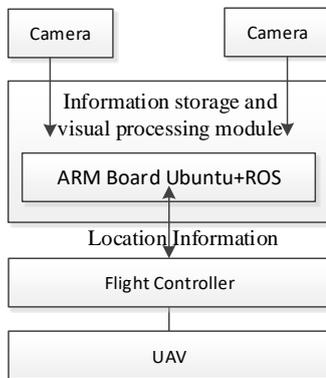


Fig.1. System architecture of UAV with stereo visual odometry

B. Stereo visual positioning method

Usually, the visual positioning processing of UAV can be divided into four steps, image acquisition and processing, feature point extraction, feature point matching and motion estimation. In this paper, we adopt SIFT (Scale Invariant Feature Transformation) algorithm to extract feature point of image[16], [17]. Fig.2 shows the feature matching procedure based on SIFT, where two key point descriptors are derived from the images of left and right camera respectively, and then the matching relationship is obtained by calculating the Euclidean distance of two key point descriptors.

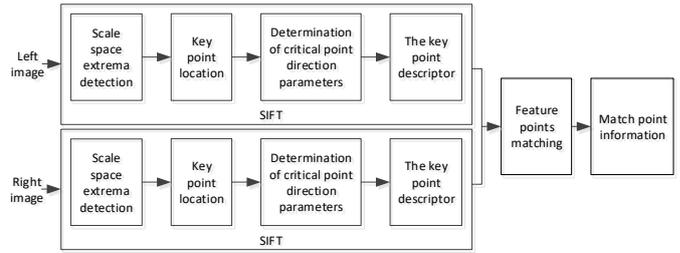


Fig.2. The features matching procedure based on SIFT

Now we assume the position of UAV at instant t_i is P_{t_i} , and its coordinate in GCS (Ground Coordinate System) is $S_{t_i} = (X_{t_i}, Y_{t_i}, Z_{t_i})$. Then, $I_{t_i,L} = f(u_{t_i,L}, v_{t_i,L}, t_i)$ and $I_{t_i,R} = f(u_{t_i,R}, v_{t_i,R}, t_i)$. indicate the image captured from the left camera and right camera at t_i respectively, where u and v is the pixel coordinate in ICS (Image Coordinate System), I is the gray value matrix of image. The following steps are the detailed operations of standard stereo visual positioning.

1) At t_i , extract the feature points from images $I_{t_i,L}$ and $I_{t_i,R}$ of position P_{t_i} using SIFT. Depending on the intrinsic parameters of camera, the coordinate matrix at t_i can be calculated and expressed as $L_{t_i} = \begin{pmatrix} \vec{X}'_{t_i} \\ \vec{Y}'_{t_i} \\ \vec{Z}'_{t_i} \end{pmatrix}$, where L_{t_i} is a matrix with $a_{t_i} * 3$ if the number of matched feature points is a_{t_i} . The elements of L_{t_i} are the coordinate vectors of feature points in ICS, so $X'_{t_i} = (X'_0, X'_1, \dots, X'_{a_{t_i}})^T$, $Y'_{t_i} = (Y'_0, Y'_1, \dots, Y'_{a_{t_i}})^T$, and $Z'_{t_i} = (Z'_0, Z'_1, \dots, Z'_{a_{t_i}})^T$;

2) At t_{i+1} , capture the images $I_{t_{i+1},L}$ and $I_{t_{i+1},R}$ of position $P_{t_{i+1}}$, and then repeat step 1 to calculate the corresponding coordinate matrix $L_{t_{i+1}} = \begin{pmatrix} \vec{X}'_{t_{i+1}} \\ \vec{Y}'_{t_{i+1}} \\ \vec{Z}'_{t_{i+1}} \end{pmatrix}$;

3) Use the feature points at t_i to track feature points at t_{i+1} , then two new matrixes L_{t_i}' and $L_{t_{i+1}}'$ of tracked feature points with size $b_{t_{i+1}} * 3$ are obtained. Substituting these two matrixes into equation $L_{t_{i+1}}' = R_{t_i} L_{t_i}' + T_{t_i}$, the relative rotation matrix R_{t_i} and displacement matrix T_{t_i} of time interval $[t_i, t_{i+1}]$ is calculated;

4) Next, the coordinate of UAV in GCS at t_{i+1} can be expressed as $S_{t_{i+1}} = R_{t_i} S_{t_i} + T_{t_i}$, that means we can estimate the location of UAV at t_{i+1} according to its previous coordinate S_{t_i} and the images captured at t_i and t_{i+1} .

Suppose $P_{t_0}, P_{t_1}, P_{t_2}, \dots, P_{t_n}$ are the successive positions in the flight line of UAV. If its initial coordinate at t_0 is S_{t_0} , UAV must repeat the above steps in each position and estimate its coordinate. Finally, the coordinate at t_n can be calculated as equation (1).

$$\begin{cases} S_{t_1} = R_{t_0} S_{t_0} + T_{t_0} \\ S_{t_2} = R_{t_1} S_{t_1} + T_{t_1} \\ \vdots \\ S_{t_n} = R_{t_{n-1}} S_{t_{n-1}} + T_{t_{n-1}} \end{cases} \quad (1)$$

III. CUMULATIVE ERROR SUPPRESSION USING CES-HVI

A. Design Idea

Although visual positioning is an effective alternative to traditional positioning method, the accumulative error prob-

lem seriously restricts its reliability. As shown in equation (1), the coordinate of UAV at t_i depends on its previous coordinate at t_{i-1} . During the processing of successive positioning of flight line, the positioning error will be gradually accumulated and then make flight navigation fail.

At present, the typical method against this problem is to pre-deploy landmarks with known coordinate in UAV flight environment. By identifying these landmarks, UAV can effectively reduce the cumulative error. But this way isn't convenient for UAV to fly in an unknown environment. Considering that the regional cruise is the commonest scenario for UAV indoor application, e.g. regional search and rescue inside a building. When UAV executes the mission of regional cruise, it needs to regularly traverse this area and repeatedly passes the same flight line. In Fig.3 and Fig.4, the typical flight lines of regional cruise are shown. Fig.3 is the spiral trajectory for scanning an area, Fig.4 shows the grid trajectory for round trip fly. From the trajectories shown in Fig.3 and Fig.4, we can find that the UAV will repeatedly pass the same position when it is cruising in a designated area. This flight characteristic provides a new idea, that is, the feature points of historical images and the historical positioning results in neighbor area of current position, can obviously be used to eliminate the current cumulative error. Based on this consideration, we propose CES-HVI method to suppress the cumulative error by using the information of historical visiting points.

For the example of Fig.3, when we calculate the coordinate of point $S_{t_{27}}$, its historical neighbor points ($S_{t_5}, S_{t_{17}}, S_{t_{18}}, S_{t_{26}}$) can provide useful extra information for correcting current positioning result. Also as shown in Fig.4, when current location is $S_{t_{22}}$, the image features and positioning results of historical neighbor points ($S_{t_7}, S_{t_8}, S_{t_{16}}, S_{t_{17}}, S_{t_{18}}, S_{t_{21}}$) can be used for correcting current result.

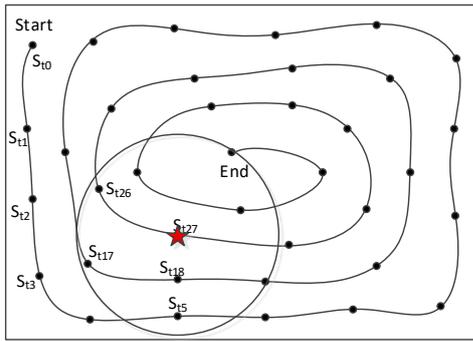


Fig.3. Spiral trajectory of regional cruise

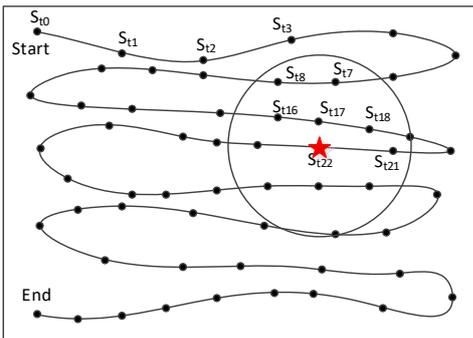


Fig.4. Grid trajectory of regional cruise

B. Algorithm Description

Definition 1 (Matching rate): Assume n_{t_i} is the number of extracted feature points from image captured at time t_i . Tracking these feature points in the extracted feature points from image captured at time t_j , if n_{t_j} is the successful tracked number, we can define $r_{t_i, t_j} = n_{t_j} / n_{t_i}$ ($i < j$) as the matching rate of images between n_{t_i} and n_{t_j} .

Definition 2 (Trusted neighbor point): If the stored historical image at t_j satisfies the conditions as shown in equation (2), then we define S_{t_j} as a trusted neighbor point of location S_{t_i} , where λ denotes the radius of neighbor area of S_{t_i} , δ is a threshold of matching rate to measure the similarity of two images at t_i and t_j . If these two conditions are satisfied simultaneously, we can consider the S_{t_i} is close to S_{t_j} and the images of two location are also similar. For these trusted neighbor points, we also call as trusted historical visiting points.

$$\begin{cases} |S_{t_i} - S_{t_j}| \leq \lambda \\ r_{t_i, t_j} \geq \delta \end{cases}, (i < j) \quad (2)$$

Based on above definitions, we propose CES-HVI method for suppressing the cumulative error by considering the extra information of trusted neighbor points. The processing of CES-HVI method is to traverse all trusted neighbor points of current position, then estimate current coordinate using the information of each trusted neighbor point one by one, further output a weighted coordinate as final positioning result, where the weight is the image matching rate of current location and trusted neighbor point. Next, we will show the steps of CES-HVI in detail.

Step1: Estimate the coordinate $S_{t_m} = (X_{t_m}, Y_{t_m}, Z_{t_m})$ at t_m using the historical positioning result $S_{t_{m-1}}$ and the images captured at instant t_m and t_{m-1} .

Step2: Define a circle as neighbor area of S_{t_m} , its center point is S_{t_m} and its radius is λ . Search the trusted neighbor points at this area then put their corresponding time t_j ($0 \leq j \leq m$) of trusted neighbor points to set $HP = HP \cup t_j$, where the neighbor point at t_j should satisfy the conditions of Definition 2. Now we define the size of HP is d , namely $\text{card}(HP) = d$.

Step3: Re-estimate the coordinate of t_m using the stored historical images of t_j and captured images of t_m , we use $S_{t_m}^k$ ($1 \leq k \leq d$) to denote the estimated result, where $S_{t_m}^k$ is calculated by $S_{t_m}^k = R_{HP(k), t_m} S_{HP(k)} + T_{HP(k), t_m}$, $R_{HP(k), t_m}$ is the relative rotation matrix of time interval $[t_m, HP(k)]$ and $T_{HP(k), t_m}$ is the relative replacement matrix.

Step4: Repeat step 3, traverse all trusted neighbor points in HP , and then obtain d estimated coordinates of t_m .

Step5: The corrected coordinate is derived by equation (3). This final output is the weighted sum of all estimated coordinates obtained in the previous two steps.

$$S_{t_m}^* = \sum_{k=1}^d r_{HP(k), t_m} S_{HP(k), t_m} \quad (3)$$

Where, $r_{t_m, HP(k)}$ indicates the matching rate of images between two instants t_m and $HP(k)$, $S_{HP(k), t_m}$ is the estimated coordinate using the images of instants t_m and $HP(k)$, $S_{t_m}^*$ is the final corrected positioning result at t_m .

In order to reduce the computational overhead of CES-HVI, our program running on ARM board will record the feature points information calculated in each historical visiting point and then use them for the subsequent estimation work. After each flight cycle is completed, the ARM board will update the stored data in next flight cycle.

IV. EXPERIMENTAL RESULTS

To evaluate the performance of CES-HVI method, a Quadrotor UAV based on the flight controller of Pixhawk is built by us for experiments. This UAV adopts Pixhawk autopilot as flight controller, adopts Odroid XU4 ARM board as image processing module, adopts two special cameras (model number is mvBlueFOX MLC200wC) with the frame rate up to 120fps and wide-angle lens for capturing image. During flight, the captured images will be dealt with by the visual processing software developed by ourselves, which is run in the ROS of Odroid XU4. The positioning result will send to Pixhawk for navigation through MAVROS module. Fig.5 shows the photo of this UAV. In our experiments, we will assign some flight missions to UAV, and collect the flight data recorded in ARM board after mission complete. These recorded data will be further analyzed using Matlab.



Fig.5. The UAV for our experiments

A. Programming in ROS

ROS is an excellent open source middle-ware for robot development. In our experiments, we write flight controlling program in ROS. For each program running under ROS, it can be regarded as a ROS node and the communication between nodes is based on subscribe/publish mode. For example, if node A wants to get the callback message from ARM board, it only needs to subscribe the topic of "/mavlink/from". When a new message of this topic arrives, node A will receive the notification. The following Fig.6 shows the subscribe/publish relationships of all ROS nodes in our program.

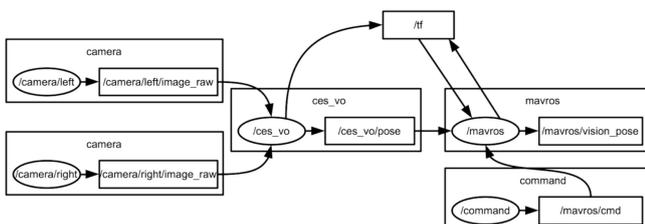


Fig.6. The relationship of ROS nodes in experiments

Where, the outside rectangles denote the domain name, the inside rectangles denote the topics, the ellipses denote ROS nodes. In Fig.6, the "/ces_vo" node is the core, it subscribes the messages of "/camera/left/image_raw" and "/camera_right/image_raw" published by left and right cameras, and then publishes the positioning results to topic of "/ces_vo/pose" after executing CES-HVI algorithm. The "/mavros" node subscribes the topic from "/ces_vo" node, it also subscribes the "/mavros/cmd" topic to receive destinations information. The output messages from "/mavros" node are sent to Pixhawk for adjusting the flight attitude of UAV, including estimated positioning information and orientation information, Fig.7 shows one output instance during our experiments.

```
header:
seq: 1071
stamp:
secs: 1357012634
nsecs: 495277648
frame_id: /cam
pose:
pose:
position:
x: -0.14582229433
y: -0.117983919863
z: -0.0015229562133
orientation:
x: -0.00651617942939
y: 0.00825718582049
z: -0.00910194241316
w: 0.999945427306
covariance: [1.5690719662979858e-06, -3.7875973276981045e-07, 1.8259838270866216e-07, 3.389285395789587e-07, 1.3049539
49432299e-06, 2.279596824134233e-07, -4.2359895417623047e-07, 4.567943728354522e-06, -1.2480697230551413e-07, -4.655532
49395714e-06, -3.983880191742535e-07, 9.656175764138476e-08, 3.288992762477526e-07, 3.6959284847442424e-07, 3.33189898
97958484e-07, -4.8893185048678936e-07, 2.7662918853888393e-07, 5.251752307747314e-08, 3.832788277708478e-07, -4.66207579
92737985e-06, 1.709271000224184e-08, 4.39111845400057e-06, 2.9399340371898565e-07, -1.611579922566079e-07, 1.334678751
7543845e-06, -2.220919815204712e-07, 2.0077071973559248e-07, 2.67184568478139e-07, 1.154289486883862e-06, 1.5477455943
384542e-07, 4.3801133246258286e-07, -6.265792280815585e-07, 5.941602757195755e-08, 6.155563768386355e-07, 3.59539340525
1294e-07, 3.273485613835577e-07]
```

Fig.7. One output result of experiments

B. Observation of Accumulative Error

Based on the programming in ROS, we first design an indoor autonomous flight experiment for disclosing the accumulative error problem of visual positioning in long distance flight. Through this experiment, we can observe the impact of accumulative error for UAV indoor navigation.

The experiment is deployed inside the building of school of computer science of Jiangsu University, as shown in Fig.8. In order to guarantee the reliability of visual positioning, we lay some cardboards with images of different colors and textures on the ground. These colorized cardboards are used to improve the effect of extracting feature points from captured image. From Fig.8, we can see the area of laid cardboards is an exact rectangle, and it is the flight area for UAV. The flight mission of this experiment is to circularly fly along a rectangle path of 400*200cm. In Fig.8, (a) shows the startup position of UAV, (b)~(d) show three angular points of flight path respectively. The default coordinates of these four positions are set as (0,0), (400,0), (400,200), (0,200) in experiment. With flying along the path circularly, UAV will record the real positioning results when it passes these four angular points. Table.I gives the recorded data of first four cycles in one experiment.



(a) Experiment image of location (0,0)



(b) Experiment image of location (400,0)



(c) Experiment image of location (400,200)



(d) Experiment image of location (0,200)

Fig.8. Flight Experiment under Mission of Rectangle Path

 TABLE I
 POSITIONING RESULTS OF FLIGHT EXPERIMENT

	Position 1	Position 2	Position 3	Position 4
Cycle 1	(0, 0)	(409, 1)	(409, 205)	(16, 208)
Cycle 2	(16, 5)	(428, 6)	(427,215)	(32,215)
Cycle 3	(32, 9)	(443, 9)	(444,224)	(49,225)
Cycle 4	(50, 17)	(461,18)	(461,231)	(68,230)

From Table I, we can clearly see that the accumulative error gradually becomes large with the cycle of flight increases. As the statistics of our experiments, when the number of flight cycle is more than 7 or 8 cycles, the indoor navigation mission will fail due to the UAV will completely lost its location. This experiment discloses the damage of accumulative error problem. Next, we will realize CES-HVI method and evaluate its performance.

C. Recommended Setting of Selecting Trusted Neighbor Points

According to the steps of CES-HVI described in III.B, the values of neighbor area radius λ and matching rate threshold δ are the key to this method. For exploring the available range of λ and δ , we design another experiment and analyze the experiment data by Matlab. In this experiment, we first capture images in two known positions respectively, then estimate the coordinate of one position using the two images and the known coordinate of another position. By enlarging the distance between two positions, we can see the change of image matching rate and positioning error. The results

are given in Table II and Fig.9. When the distance become large, the matching rate of two images become low, and then lead to the degradation of both absolute positioning error and relative positioning error. If the distance is 200cm, the result is well. But if the distance expands to 240cm, the result become unacceptable. So, we recommend the λ is set as 200cm and δ is set as 0.25 according to our experiment data. We ensure that the CES-HVI method can run smoothly by our suggestion.

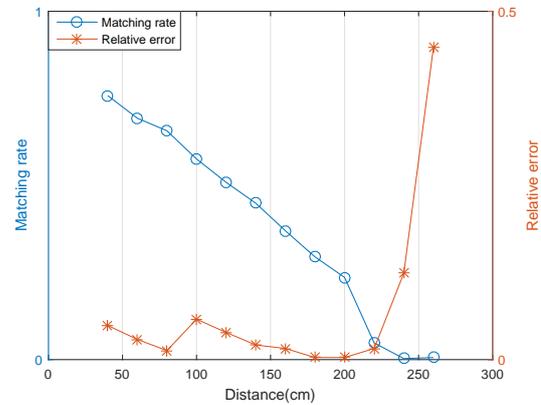


Fig.9. Matching rate and relative error vs. Distance

 TABLE II
 RELATIONSHIP BETWEEN DISTANCE, MATCHING RATE AND POSITIONING ERROR

Distance (cm)	Results of Visual Positioning Algorithm	Matching Rate	Absolute Error	Relative Error
-3,-4	-3,-4	-3,-4	0/0	0/0
836,6	866,9	854,6	30/18	3/0
945,100	1008,109	972,113	63/27	9/13
9,101	977,226	27,107	98/18	10/6
846,209	202,322	886,209	131/40	17/0
938,308	1105,328	1004,325	167/66	20/17
-1,299	202,322	59,428	203/60	23/23
848,402	1087,430	918,525	239/70	28/26
938,507	1211,540	1036,532	273/98	33/18
-1,495	308,530	91,632	309/92	35/37
835,606	1178,646	950,738	343/115	40/26
936,709	1311,752	1068,733	375/132	43/29
4,698	411,744	123,733	407/119	46/35

D. Positioning Results of CES-HVI

This section evaluates the effect of CES-HVI method. In the experiment of this section, the flight mission is to fly along the grid line in a laboratory with area of 120 m^2 . For making the authors to understand the visual positioning more clear, we use Fig.10 ~ Fig.13 to show some key steps. Fig.10 are the images captured by left and right camera at the start instant 0s. Fig.11 is the result of extracted feature points of left camera image at 0s. Fig.12 are the images captured by left and right camera at 0.025s during flight. Fig.13 gives the feature points tracing result of two instants 0s and 0.025s, the green points are the successful tracked feature points in the image of 0.025s, their tracking objectives are the feature points in the image of 0s.



(a) Image captured by left camera



(b) Image captured by right camera

Fig.10. Image captured at instant 0s



Fig.11. Extracted feature points of left camera at 0s



(a) Image captured by left camera



(b) Image captured by right camera

Fig.12. Image captured at instant 0.025s



Fig.13. Tracked feature points between 0s and 0.025s

In this experiment, we select the standard visual positioning method described in section II.B as comparison algorithm. The flight results of standard method and CES-HVI method are shown in Fig.14. In this Figure, the line with circle sign is the flight mission assigned to UAV, the line with square sign is the flight trajectory of standard method, and the line with star sign is the flight trajectory of CES-HVI. Comparing the line of circle with line of square, we can see the trajectory of standard method obviously deviates the mission path in the later part of flight. But the line of star derived by CES-HVI is close to mission path. Now we have known the deviation phenomenon is caused by the cumulative error. Due to there is no any suppression design for cumulative error, the standard method of visual positioning become unreliable with the increasing of flight time. So, it is not suitable for long distance flight. But with the extra information of historical visiting points, CES-HVI can effectively mitigate the cumulative error. Table III gives the detailed data of positioning results and relative errors of two methods.

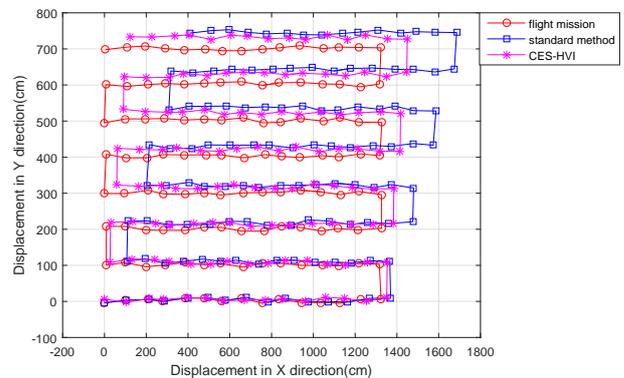


Fig.14. Comparison of flight mission, positioning results of standard method and CES-HVI

In order to clear disclose the effect of CES-HVI, we repeat the above flight experiment around 50 times and then obtain the statistics of cumulative error of two methods. The statistics is given in Fig.15. From this figure, we can find the cumulative error of standard method almost increases linearly with the increasing of cumulative flight distance. Although the cumulative error of CES-HVI also increases with the accumulating of flight distance, it shows marked effect for suppressing cumulative error. When the cumulative flight distance is 4000cm, the cumulative error of standard method is 135cm and the cumulative error of CES-HVI is 55.8cm. CES-HVI improves 58.7% comparing with standard method. When the cumulative flight distance is 11400cm, the cumulative error of two methods is 391cm and 121.2cm respectively. CES-HVI obtains 69% promotion for suppressing the cumulative error. So, it is an effective improvement for standard visual positioning.

TABLE III
RELATIONSHIP BETWEEN DISTANCE, MATCHING RATE AND
POSITIONING ERROR

Coordinate (cm)	Positioning Result of the Standard Method	Positioning Results of the CES-HVI	Relative Error of X Direction (Standard Method/ CES-HVI)	Relative Error of Y Direction (Standard Method/ CES-HVI)
-3,-4	-3,-4	-3,-4	0.000/0.000	0.000/0.000
836,6	866,9	854,6	0.036/0.022	0.500/0.000
945,100	1008,109	972,113	0.067/0.029	0.090/0.130
9,101	977,226	27,107	10.889/2.000	0.099/0.059
846,209	202,322	886,209	0.155/0.047	0.081/0.000
938,308	1105,328	1004,325	0.178/0.071	0.065/0.055
-1,299	202,322	59,428	203.000/60.000	0.077/0.077
848,402	1087,430	918,525	0.282/0.083	0.070/0.065
938,507	1211,540	1036,532	0.291/0.105	0.065/0.036
-1,495	308,530	91,632	309.000/92.000	0.071/0.075
835,606	1178,646	950,738	0.411/0.138	0.066/0.043
936,709	1311,752	1068,733	0.401/0.142	0.061/0.041
4,698	411,744	123,733	101.750/29.750	0.066/0.050

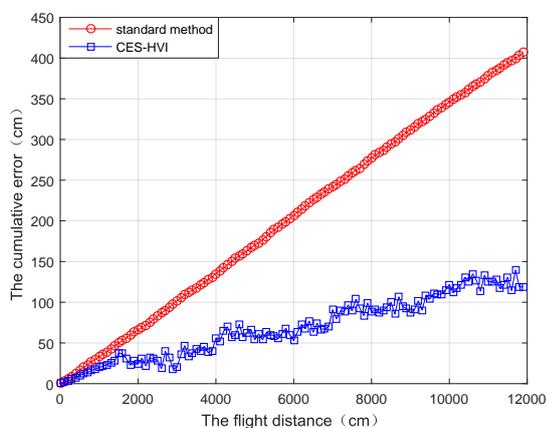


Fig.15. Comparison of cumulative error between standard method and CES-HVI

V. CONCLUSION

Today, the application of UAV in indoor environment become more and more widely, but the cumulative error problem of visual positioning technology significantly restricts its practicability. How to improve the accuracy of visual positioning by suppressing the cumulative error becomes one of hotspots of current research. Compared with the traditional method by pre-deploying the landmarks, this paper uses historical visiting points as virtual landmarks for position estimation, which can be better applied to unknown environment. The experimental results in real environment show the CES-HVI method can effectively reduce the cumulative error of UAV in long distance flight.

In the future study, we will emphasize on reducing the complexity of CES-HVI method, and further optimize the algorithm running time.

ACKNOWLEDGMENT

This research was financially supported by the Innovation Program for Graduate Student of Jiangsu Province (Grant No. SJZZ15_0133).

REFERENCES

- [1] Arras, K. O., and N. Tomatis, "Improving robustness and precision in mobile robot localization by using laser range finding and monocular vision," *Advanced Mobile Robots, 1999. (Eurobot '99) 1999 Third European Workshop on IEEE*, pp. 177-185, 1999.
- [2] Xu, Zezhong, et al, "Map building for indoor environment with laser range scanner," *International Conference on Intelligent Transportation Systems, 2002. Proceedings IEEE*, pp. 136-140, 2002.
- [3] Li Q Q., "Research on Laser Range Scanning and Its Application," *Journal of Wuhan Technical University of Surveying and Mapping*, vol. 25, no. 5, pp. 387-392, 2000. published.
- [4] Celik, K, S. J. Chung, and A. Somani, "Mono-vision corner S-LAM for indoor navigation," *IEEE International Conference on Electro/information Technology IEEE*, pp. 343-348, 2008.
- [5] Kleinert, Markus, and S. Schleith, "Inertial aided monocular SLAM for GPS-denied navigation," in *Multisensor Fusion and Integration for Intelligent Systems IEEE*, pp. 20-25, 2010.
- [6] Cohen, Benjamin, and J. Byrne, "Inertial aided SIFT for time to collision estimation," *IEEE International Conference on Robotics and Automation IEEE Press*, pp. 4465-4466, 2009.
- [7] Byrne, Jeffrey, and C. J. Taylor, "Expansion segmentation for visual collision detection and estimation," *IEEE International Conference on Robotics and Automation IEEE*, pp. 875-882, 2009.
- [8] Yan, Qin Lao, et al, "Double three-point fix and objective diagram fitting controlling of farming robot," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 40, no. 9, pp. 171-174, 2009.
- [9] Zou, Xiangjun, H. Zou, and J. Lu, "Virtual manipulator-based binocular stereo vision positioning system and errors modelling," *Machine Vision and Applications*, vol. 23, no. 1, pp. 43-63, 2012.
- [10] Qi, Z., and Q. Wang, "Error analysis and the development of an error mitigation approach for use in the rotation fiber optic gyro inertial navigation system," *Engineering Letters*, vol. 21, no. 4, pp. 356-364, 2013.
- [11] Tournier G, Valenti M, How J, et al, "Estimation and control of a quadrotor vehicle using monocular vision and moire patterns," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, pp. 6711, 2006.
- [12] Rondon, E, L. R. Garcia-Carrillo, and I. Fantoni, "Vision-based altitude, position and speed regulation of a quadrotor rotorcraft," *Ieee/rsj International Conference on Intelligent Robots and Systems IEEE*, pp. 628-633, 2010.
- [13] Liu, Zhen Yu, and N. Jiang, "Self-localization of mobile robots based on artificial landmarks and stereo vision," *Computer Engineering & Applications*, vol. 46, no. 9, pp. 190-192, 2010.
- [14] Chen, Ming Ya, Z. Y. Xiang, and J. L. Liu, "Assistance localization method for mobile robot based on monocular natural visual landmarks," *Journal of Zhejiang University*, vol. 48, no. 2, pp. 285-291, 2014.
- [15] Gorbenko, A., and V. Popov, "Visual landmark selection for mobile robot navigation," *Iaeng International Journal of Computer Science*, vol. 40, no. 3, pp. 134-142, 2013.
- [16] Lowe, D. G., "Object recognition from local scale-invariant features," *International Conference on Computer Vision IEEE Computer Society*, pp. 1150, 1999.
- [17] Lowe, David G, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.