

Article

Brain-Like Navigation Scheme based on MEMS-INS and Place Recognition

Chong Shen ¹, Xiaochen Liu ¹, Huiliang Cao ¹, Yuchen Zhou ^{1,2,*}, Jun Liu ^{1,*}, Jun Tang ¹, Xiaoting Guo ¹, Haoqian Huang ³ and Xuemei Chen ⁴

¹ Key Laboratory of Instrumentation Science & Dynamic Measurement, Ministry of Education, School of Instruments and Electronics, North University of China, Taiyuan 030051, China; shenchong@nuc.edu.cn (C.S.); s1706093@st.nuc.edu.cn (X.L.); caohuiliang@nuc.edu.cn (H.C.); tangjun@nuc.edu.cn (J.T.); guoxiaoting@nuc.edu.cn (X.G.)

² Research Institute of Tsinghua University in Shenzhen, Shenzhen 518057, China

³ College of Energy & Electric Engineering, Hohai University, Nanjing 210098, China; hqhuang@hhu.edu.cn

⁴ School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China; chenxue781@126.com

* Correspondence: 1506014440@st.nuc.edu.cn (Y.Z.); liuj@nuc.edu.cn (J.L.)

Received: 22 February 2019; Accepted: 22 April 2019; Published: 25 April 2019



Abstract: Animals have certain cognitive competence about the environment so they can correct their navigation errors. Inspired by the excellent navigational behavior of animals, this paper proposes a brain-like navigation scheme to improve the accuracy and intelligence of Micro-Electro-Mechanical System based Inertial Navigation Systems (MEMS-INS). The proposed scheme employs vision to acquire external perception information as an absolute reference to correct the position errors of INS, which is established by analyzing the navigation and error correction mechanism of rat brains. In addition, to improve the place matching speed and precision of the system for visual scene recognition, this paper presents a novel place recognition algorithm that combines image scanline intensity (SI) and grid-based motion statistics (GMS) together which is named the SI-GMS algorithm. The proposed SI-GMS algorithm can effectively reduce the influence of uncertain environment factors on the recognition results, such as pedestrians and vehicles. It solves the problem that the matching result will occasionally go wrong when simply using the scanline intensity (SI) algorithm, or the slow matching speed when simply using grid-based motion statistics (GMS) algorithm. Finally, an outdoor Unmanned Aerial Vehicle (UAV) flight test is carried out. Based on the reference information from the high-precision GPS device, the results illustrate the effectiveness of the scheme in error correction of INS and the algorithm in place recognition.

Keywords: inertial navigation system; visual navigation; place recognition; image matching

1. Introduction

Navigation technology plays a significant role in our life. Even in an unfamiliar environment, we can easily get the current location and reach our destination with the help of GPS. Nevertheless, GPS may be blocked in some environments, such as the places between tall buildings in cities, deep-sea area, deep mountains and forests, etc. Micro-Electro-Mechanical System Inertial navigation system (MEMS-INS) demonstrates its unique superiority, which can calculate the next-point location based on the continuously measured self-motion velocity and direction information, rather than external information, so it has a vast application scope. Whereas, the navigation information of INS is generated by the integration of velocity and direction information measured by sensors, so the error will increase over time, thereby resulting in poor positioning accuracy over a long time [1–3]. Additionally, inertial

sensors suffer from large measurement uncertainty at slow motion [4]. It has undoubtedly become one of the research hotspots to reasonably and effectively reduce these accumulated errors and improve the intelligence of the INS in recent years.

Studies have shown that the navigation strategies which integrate one or more navigation means, such as GPS [5–7], Spectral Redshift [8], Polarized Light [9] etc., with INS can obtain complementary advantages and achieve better navigation results. The Unmanned Aerial Vehicle (UAV) has played an increasingly important role in various fields in recent years [10], to achieve better navigation results, GPS/INS is the most widely used integrated navigation system especially for UAVs [11–13]. However, in GPS blocked environments, the GPS/INS system would be shut down and go back to INS standalone mode and the navigation error would be accumulated [14]. As a common integrated navigation method for UAVs, visual navigation also has many drawbacks [15]. On the one hand, the performance of visual navigation is inversely proportional to the velocity. Another limitation of visual sensors is due to their inability to distinguish rotational from translational motion, since it requires the summation of six motion components. Moreover, the mentioned integrated navigation systems are still simply a combination of different devices without any intelligent ability. To endow INS with intelligent ability, the attention of researchers is turned to the biological world.

Many animals have outstanding navigation ability, such as pigeon post, old horse knowing the way, etc. Whether across a myriad of rivers, thousands of hills, in a mighty storm or some other severe weather event, these amazing animals always know the way back home. Many people are good at recognizing the way too, and they seem to have a high-resolution map embedded in their mind so that they will not get lost somehow. It shows that a living organism can obtain the perception of external information just by its own organs, which can be transformed into accurate navigation information through a certain biological mechanism. The Nobel Prize in Physiology or Medicine in 2014 revealed this mystery for us. As early as in 1971, John O'Keefe discovered a type of neuron in an encephalic region called hippocampus in the rat brain, which was responsible for remembering location features [16]. More than 30 years later, May-Britt Moser and Edward Moser demonstrated that there was a mechanism for establishing spatial coordinate system in an animal's brain, which could generate path integration by obtaining motion information from speed cells [17] and head-direction (HD) cells [18,19]. In imitation of this mechanism, Michael E. Milford conducted the mathematical model of a rat brain, and established the SeqSLAM bionic navigation algorithm by using the vision-driven navigation system and achieved good results in outdoor navigation experiments [20]. SeqSLAM is a pure visual navigation algorithm, of which the core sensing information only come from the visual odometer. In fact, when animals or humans reach a familiar environment, the path integrator will be reset to adapt to the external environmental information sensed by the eyes. It means that the integrated position error can be corrected when animals or humans come into a familiar environment, where it can be recognized as an absolute position reference, therefore, the navigation accuracy can be remarkably enhanced by intermittently eliminating accumulated errors. The crux of stimulating such error correction mechanism of organisms lies in whether it is capable of rapidly and accurately matching current visual scenes with the feature scenario in "memory." SeqSLAM uses an algorithm based on image grayscale information to conduct closed-loop detection of scenes. Its advantages include the small computing amount and excellent real-time, but it has poor robustness and is vulnerable to environmental factors. By contrast, other visual matching algorithms such as Vector Field Consensus algorithm (VFC) [21], Guided Locality Preserving Matching algorithm (GLPM) [22] and Grid-based Motion Statistics algorithm (GMS) [23] are based on feature points. Most of these algorithms usually perform feature extraction first [24,25], and then match the extracted feature points in some way [26]. Therefore, they usually have highly accurate matching results, but the computation cost still needs to be improved for real-time application. To effectively eliminate the accumulative position errors of INS and enhance its intelligence, this paper first investigates the navigation mechanism of rats. The investigation finds that the reason why rats can eliminate the accumulative errors caused by path integration is that they have certain cognitive competence about the environment. Inspired by this, this paper

combines vision and INS together, putting forward a brain-like navigation scheme. This system can recognize “spots in memory” during movement to perform error correction for INS at those spots. Additionally, to better solve the place recognition problem of the scheme, this paper compares and analyzes traditional grayscale based matching methods with feature point based matching methods, and at last proposes a place recognition algorithm which is named SI-GMS. Experimental results suggest the scheme mentioned above can considerably improve the positioning accuracy of INS. The main contributions of this paper are summarized as follows:

The paper presents a “brain-like navigation” scheme based on the principle of rat brain navigation cells. The proposed brain-like navigation scheme corrects position errors caused by path integration through visual information, which improves the position measurement precision and intelligence level of INS.

The paper introduces a place recognition algorithm that integrates gray information and feature point information of images specific to the proposed brain-like navigation scheme. This algorithm has numerous advantages including real-time and accurate matching results. It can effectively overcome the shortcomings of the traditional place recognition algorithms, such as missed judgement, misjudgment and poor environmental suitability.

The paper is organized as follows: Section 2 is the proposed brain-like navigation model. Section 3 is the description of proposed place matching algorithm. Experiment and comparison results are presented in Section 4. The paper ends with a conclusion in Section 5.

2. Proposed Solution

Some animals have excellent navigation skills. For example, ants can directly go into their nests after foraging for food and migratory birds can fly over thousands of miles a year without getting lost. Humans also can remember different scenes for navigation. After years of research, the winner of the 2014 Nobel Prize in Physiology or Medicine discovered the brain localization system cells for animal navigation mechanisms [17]. To date, the primary neuronal cells related to animal’s environmental cognition that have been found mainly include place, head-direction (HD) and speed cells.

2.1. Place Cells

According to the experiment carried out by O’Keefe and Conway [12], when a rat runs to a particular area, certain cells fire consistently. If the rat runs to other areas, several other cells fire. These cells are called place cells. The figure below shows the trajectory of the rat in the open room. The orange point indicates the position of the rat when certain cells are active at a particular position, and the light orange round spot indicates the theoretical field of the cells in these positions (Figure 1).



Figure 1. Certain cells are active at a particular position.

The orange dots indicate that the place cells at this position are firing, and the black line represents the path of rat movement. When the rat reaches a specific site, cells are fired at a specific location, which indicates the position of the rat in the field.

2.2. HD Cell

In 1984, physiologists Ranck [17,18] identified HD cells at the presubiculum of rats, thereby elucidating the mechanism by which the animal encodes the direction of its own motion. Experiments have shown that when a rat's head is oriented in a certain direction, certain HD cells are firing. When the head deviates from this direction, this firing gradually decreases. As shown in Figure 2, the preferred direction of HD cells is at an angle of 150° . When the rat's head is oriented in this direction, the strength of the electrical signal could peak.

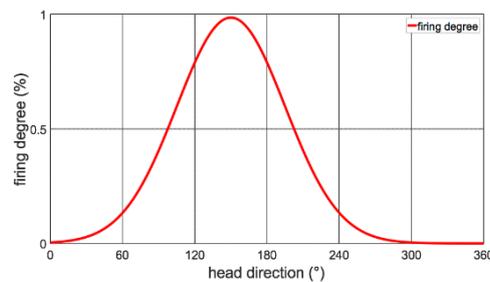


Figure 2. Response curve of firing of a head-direction (HD) cell.

2.3. Speed Cell

In 2015, after winning the Nobel Prize in Physiology or Medicine, May-Britt Moser and Edvard Moser [17] discovered that some nerve cells can increase their firing rate proportional to the increase of moving speed. By investigating the firing rate of such cells, it was possible to determine the speed at which an animal moves at a given point of time. The researchers named these cells as speed cells.

In their experiment, the researchers placed a rat in a top-opened box and induced it to run freely by randomly throwing food. The experiment was conducted in a dark environment to avoid the impact of visual information. Simultaneously, during data analysis, the effect of rat behavior on speed cells was avoided by ignoring the changes in cell viability of all moving speeds <2 cm/s. The activity intensity was embodied by a simple linear decoder [17], consisting of a distribution field and linear filter (Figure 3). The activity state information was transmitted to the pose cell where HD and place cells were fused, and then affected the linear track.

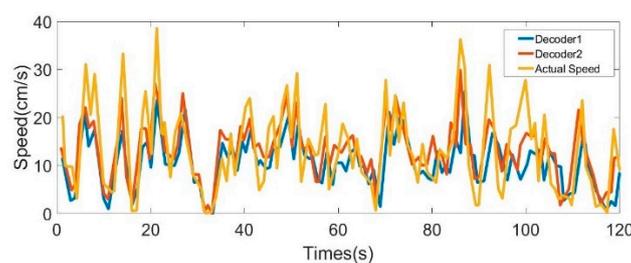


Figure 3. Actual speed and the results of linear decoders.

2.4. Proposed Brain-Like Navigation Model Based on Vision and INS

As indicated by research on physiology, the phenomenon that the animal's place cells are activated in the environment is determined by path integration, which is a consequence of integrating internal cues include directional heading and distance computations. Naturally, errors in path integration accumulate over time. When rats are placed in a familiar environment, the path integrator will be reset to be adapted to the external environment information perceived by the eyes. It has been demonstrated by these studies that rats can integrate the internal and external information for accurate navigation in various environments. In a familiar environment, the positioning error can be corrected with the external perceptive information as an absolute reference, and the same is true for human beings.

By simulating the navigation mechanism of these species, we put forward an intelligent brain-like navigation model based on vision and INS.

The attractor network model of place cells constructs a measurement model corresponding to the relative position in the actual external environment. The place cells model adopts continuous attractor networks, where the place cells on the fringe are connected to the ones on the other fringe and form an annular shape. The continuous attractor networks model creates a stable activity packet via the wraparound excitatory connections on the same neural plate. This attractor network is driven by the path integration system and reset by the graphic information derived from the current position. By using two-dimensional Gaussian distribution to construct $\chi_{x,y}$ which is the weight connection matrix of excitability, $\chi_{x,y}$ is expressed by:

$$\chi_{x,y} = e^{-(x^2+y^2)/E}, \tag{1}$$

where x and y are the distances between units in X and Y coordinate respectively, and E is the width constant for position distribution.

The place cell activity matrix P is used to describe the activity in the place cells. The variable quantity of place cells' activities induced by local excitation is given by:

$$\Delta P(X, Y) = \sum_{i=0}^{S_X-1} \sum_{j=0}^{S_Y-1} P(i, j)\chi_{x,y}, \tag{2}$$

where S_X, S_Y are the dimensions of the two-dimensional matrix of the place cells in the space and represent the attractor sub-model's range of activity on the neural plate. The precondition for the iteration of place cells and the matching of the visual template is to pinpoint the relative position of place cells attractors in the neural plate. The relative position coordinates are expressed by the subscript of the weight matrix, which is calculated by the equations:

$$x = (X - i)(\text{mod}S_X), \tag{3}$$

$$y = (Y - j)(\text{mod}S_Y). \tag{4}$$

Each place cell receives the global inhibitory signal in the same manner as an entire network. The symmetry of the excitatory and inhibitory connection matrix is a guarantee for proper neural network dynamics, which keeps the attractors in the space from unrestricted excitability. The variable quantity of place cells' activities induced by inhibitory connection weight is given by the equation:

$$\Delta P(X, Y) = \sum_{i=0}^{S_X} \sum_{j=0}^{S_Y} P(i, j)\Psi_{x,y} - \xi, \tag{5}$$

where $\Psi_{x,y}$ is the inhibitory connection weight, ξ controls the level of global inhibition. The activities of all place cells are nonzero and have undergone the normalization procedure. $P^{t+1}(X, Y)$, the firing rate of place cells with path integration is given by:

$$P^{t+1}(X, Y) = \sum_{i=\Delta X}^{\Delta X+1} \sum_{j=\Delta Y}^{\Delta Y+1} \alpha_{i,j}P^t(i + X, j + Y), \tag{6}$$

where $\alpha_{i,j}$ is the residual. ΔX and ΔY are the offset values via rounding down, which is determined by speed and direction, as shown in Equation (7):

$$\begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix} = \begin{bmatrix} C_i \vec{e}_\theta v \cos \theta \\ C_j \vec{e}_\theta v \sin \theta \end{bmatrix}, \tag{7}$$

where $\lfloor \cdot \rfloor$ refers to rounding down, C_i and C_j are the constants for path integration; v is the current velocity which is achieved from speed cells; θ is the current head direction which is achieved from HD cells, and \vec{e}_θ is the unit vector pointing at θ .

Similar to the mechanism of biological autonomous navigation, the INS is a type of autonomous system that neither relies on external information nor radiates energy outwards. By measuring the angular rate and acceleration of the carrier in the inertial frame of reference, the navigation information such as speed, yaw angle and position within the navigation coordinate system can be obtained after calculation. Nonetheless, error exists in the system output, and the error will be accumulatively increasing due to the fact that the position of INS is obtained via integral.

Regarding living rats, the place cells' relative position of firing is also obtained by path integration. However, when a rat comes into a familiar environment, rats will reset the firing for all the spatial cells involved in the path integration. The renewal of coordinates is in effect the renewal process of spatial cells' firing, whereas the renewal process at the close-loop point is the reset process of spatial cells. Based on this mechanism, the brain-like navigation model based on vision and INS is proposed. This system is capable of conducting real-time detection on whether the current visual information is matched with the pre-stored visual template. If it is successfully matched, it means that a "familiar place" is reached. Subsequently, the spatial cells within the entire path integration network will undergo a firing reset procedure, so as to regain the firing state of the previous close-loop point. Via this method, the accumulated errors can be effectively eliminated, and the navigation accuracy of INS will be increased.

The model overview is as follows:

Model: Brain-like navigation model based on vision and INS

1. The camera captures a frame of RGB image
 2. Collect the motion state information of the object, and renew speed cells and HD cells
 3. Execute path integration of spatial cell
 4. Obtain spatial geometric coordinates via geometric transformation
 5. Execute the image matching algorithm and obtain the return value U of the matching result using (19)
 6. **If** $U == True$
 7. Read the location coordinates of the template image
 8. Reset the firing for place cells
 9. Correct the position errors
 10. **End if**
-

The schematic diagram of the proposed brain-like navigation scheme is shown in Figure 4. The real trajectory of the UAV is a straight blue line that starts from point A, followed by point B, point C and point D. However, due to the accumulated errors, the trajectory of INS will gradually deviate from the real track. The red dots are visually corrected points, at which the place cells are reset. The brown dotted line represents the reckoned trajectory of the UAV and the purple line represents the trajectory of pure INS. It can be found that the trajectory of INS after visual correction is closer to the real trajectory of the UAV, and the navigation accuracy has also been significantly improved.

The brain-like navigation scheme studied in this paper is mainly applied to UAV. However, the landscape features near the ground are complicated due to various interference factors. The key to realize the proposed scheme is a fast and accurate place recognition algorithm. Therefore, this paper proposes a fast and accurate place recognition algorithm for the proposed brain-like navigation system in the next section.

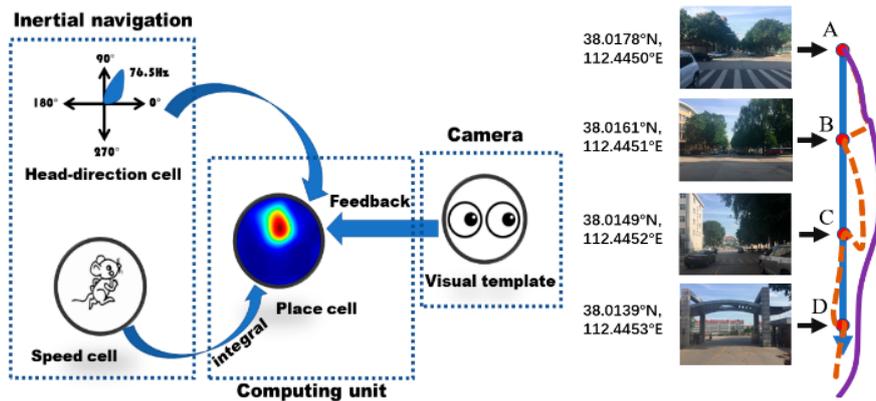


Figure 4. System composition and simulation trajectory map.

3. Algorithm

To achieve the effect of the proposed brain-like navigation scheme, the UAV needs to have the functions of memory and recognition to the scene. The memory process can be simulated by storing the coordinates and images of key locations in advance, including intersections, buildings, etc. Recognition refers to the UAV recognizing the template image corresponding to the scene and returning its real coordinate value when passing close to the location of the previous memory, to make the cell reset firing and eliminate the accumulated error. Traditional image matching algorithms generally consist of three categories, which are respectively based on grayscale [27], feature [28] and relation [29]. It is necessary for image matching algorithms to take both efficiency and precision into consideration. Michael Milford proposed place recognition algorithms RatSLAM and SeqSLAM based on continuous multi-frame image information in 2004 and 2012 respectively [17,22], and the algorithm of comparing the similarity of gray information between two images in a certain area was employed to determine whether the place cell node was encountered. This algorithm does not require large calculation and has the ideal real-time performance, but the algorithm tends to be affected by external environmental factors including weather and traffic conditions. This algorithm calculates the similarity value between the two graphs and comparing it with the threshold, thus knowing that whether the location of the image is in one place. If the threshold value is set too low, it may have omissions in judgment, but if too large, it may generate errors in judgment. Grid-based motion statistics (GMS) [20–22] is proposed for fast and ultra-robust feature correspondence. Compared with the traditional SIFT feature matching algorithm, GMS eliminates the incorreced matching points more effectively with an execution speed of 30 frames per second on the PC, which meets the real-time requirements, but the speed is getting slower on the lower computer and not suitable for real-time application. To solve the above problems, a novel place recognition algorithm for place cell correction is proposed by integrating these two algorithms with complementary advantages.

3.1. Scanlines Intensity Algorithm

The scanlines intensity (SI) algorithm is for identifying the place by comparing the similarity of the intensity profile of two image scanlines. In detail, the algorithm is to make the template image shift above the reference image, and calculate the difference between the normalized values of the two intensities, then sum them up. The smaller the sum of the differences, the more similar the two images are.

For template image T and reference image B , $T(x,y)$ and $B(x,y)$ are the intensity values (grayscale value) of corresponding pixel points on the image respectively. Taking the reference image B as an example, the sum of the intensity values of the j -th column is calculated as S_j

$$S_j = \sum_{i=0}^m B(i, j) \tag{8}$$

where m represents the number of rows.

The intensity of each column is normalized to obtain the normalized set I of the reference image.

$$I_j = \frac{S_j}{\sum_{j=1}^n S_j}, \tag{9}$$

$$\mathbf{I} = \{I_1, I_2, I_3, \dots, I_n\}, \tag{10}$$

where I_j represents the normalized result of the j -th columns, n represents the number of columns.

The normalized set I' ($I' = \{I'_1, I'_2, I'_3, \dots, I'_n\}$) of template image T can be obtained similarly.

The similarity of two images is represented by β .

$$\beta = \sum_{i=1}^n \text{abs}(I_i - I'_i). \tag{11}$$

The smaller the β , the more similar the two images are.

When the camera is applied to a UAV, if the relative height and inclination of the camera remain the same as that of the ground, two different images taken at the same location, but different times can be considered as having horizontal deviation without vertical deviation. Therefore, the region of interest (ROI) of the template image can shift above the reference image, the similarity of the corresponding region of the two images after each shift can thus be calculated, and the value of β' , which is the minimum value of all similarity values calculated after each shift, is taken as the final similarity of the two images.

$$\beta' = \min \left\{ \min_{0 \leq s \leq \text{effect}} \left[\frac{\sum_{i=1+s}^W \text{abs}(I_i - I'_{i-s})}{W-s} \right], \min_{0 \leq t \leq \text{effect}} \left[\frac{\sum_{j=1+t}^W \text{abs}(I'_{j-t} - I_j)}{W-t} \right] \right\}, \tag{12}$$

where β' is the final similarity of the two images, *effect* is the maximum shift, and W is the width of ROI.

The threshold is set as h , if $\beta' \leq h$, then the two images are considered to be matched successfully.

As shown in Figure 5, the picture on the left and the picture in the middle are two images taken at different times in the same place, whereas the picture on the right is taken in another different place. Taking the image in the middle as the template, it can be seen that the scanline profile of the template image is similar to the image on the left, but different to the image on the right.

The advantages of this algorithm are that little computation is required and the real-time processing requirements can be met. The disadvantage is that the accuracy is not high enough, and environmental variation factors including light, pedestrians and other vehicles may lead to misjudgment and low reliability.

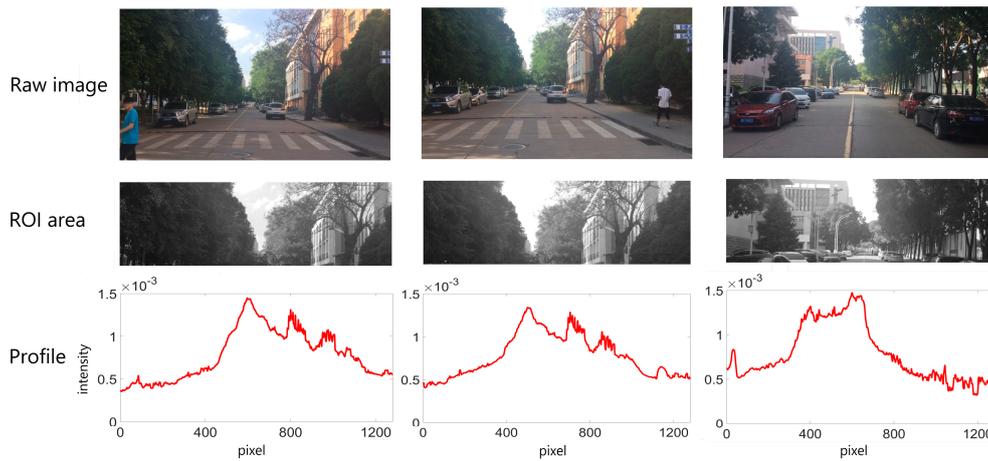


Figure 5. Raw image, region of interest (ROI) area and the scanline profile of the image.

3.2. Grid-Based Motion Statistics Algorithm

Grid-based motion statistics algorithm (GMS) is a simple statistics-based solution to the feature matching problem. The general process of algorithm execution is: firstly, perform any kind of detection and the description calculation of feature points (ORB features are used in our experiment). Then, perform a BF (Brute Force) algorithm and finally perform a GMS to eliminate the mismatch.

The principle of the GMS algorithm is as follows.

For a certain matching point (x_i, x'_i) , N_i is the score of the match which can be thought as a measure of neighborhood support. It can be deduced that N_i obeys a binomial distribution:

$$N_i \sim \begin{cases} B(M, p_r), & \text{if } x_i \text{ is the right match point} \\ B(M, p_w), & \text{if } x_i \text{ is the wrong match point} \end{cases} \quad (13)$$

where M is the number of supporting features of x_i in its neighborhood. p_r is the probability that, when (x_i, x'_i) is the correct match, the matching points corresponding to other feature points in the neighborhood of x_i are in the neighborhood of x'_i . p_w is the probability that, when (x_i, x'_i) is the wrong match, the matching points corresponding to other feature points in the neighborhood of x_i are in the neighborhood of x'_i .

The mean value and standard deviation are:

$$\begin{cases} \mu_r = Mp_r, & s_r = \sqrt{Mp_r(1-p_r)} & \text{if } x_i \text{ is the right match point} \\ \mu_w = Mp_w, & s_w = \sqrt{Mp_w(1-p_w)} & \text{if } x_i \text{ is the wrong match point} \end{cases} \quad (14)$$

The target function is:

$$\max G = \frac{\mu_r - \mu_w}{s_r + s_w}. \quad (15)$$

In order to turn the above theoretical analysis into an algorithm that can be used in practice, four problems need to be considered:

- (a) Efficient calculation of the score through grid-cells;
- (b) Grouping match neighborhoods for robustness;
- (c) Number of grid-cells to use;
- (d) Calculation of the effective threshold S ;

The solutions are summarized as follows:

(a) Divide the image into $G = 20 \times 20$ grids, iterate through each potential cell-pair once to compute the scores, accept all true matches between cell-pairs.

(b) For robustness purposes, group the cell-pairs, i.e., neighborhoods that match. A $3 \times 3 = 9$ grid around the network is calculated, as shown in the red region in Figure 6.

For grid m and grid n , the similarity score is calculated by the following formula:

$$S_{mn} = \sum_{i=1}^9 |\Omega_{m^i n^i}|, \tag{16}$$

where $|\Omega_{m^i n^i}|$ is the number of matching points between grid m^i and grid n^i .

- (c) The number of grids is usually the empirical value $G = 20 \times 20$ cells for 10,000 features.
- (d) Categorize the cell-pairs into the true set and the false set by thresholding S_{mn} .

$$C(m, n) = \begin{cases} 1, & \text{if } S_{mn} > \alpha \sqrt{t_i} \\ 0, & \text{otherwise} \end{cases}, \tag{17}$$

where $\alpha = 6$ is determined from experiment and t_i is the total number of features present in a single cell of the nine cells in Figure 6. The value of $C(m, n)$ indicates whether the grid area where m and n are a pair of corrected matching points.



Figure 6. Nine regions around cells (m, n) used in score evaluation.

The advantages of GMS are high matching accuracy, strong adaptability to the environment and almost zero misjudgment rate of the scene. Meanwhile, compared with the traditional feature matching algorithm, the matching time cost of GMS is significantly reduced, which basically meets the requirements of real-time processing (it takes about 31 ms to process a pair of images under the configuration Intel i7 CPU+GTX980 GPU). The disadvantage is that when the algorithm is applied to the lower computer, the computing speed is becoming slower than that on the desktop, and the real-time performance still needs to be improved. Moreover, this algorithm is applicable on images of different scales, but such application may produce matching errors when applied to actual scene matching. For example, when the UAV is flying over from far to near, a large number of successful matching points appear for images taken within a certain distance from the place of template images (Figure 7).

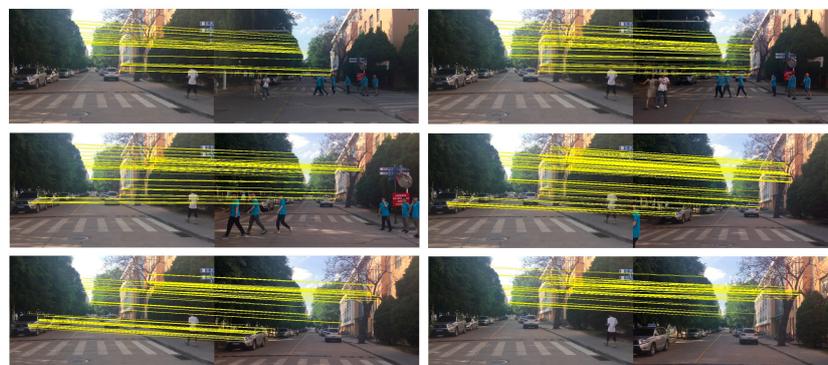


Figure 7. For a particular template (the image on the left part), images taken within a certain distance (the image on the right part) from it will produce a great number of points that match the template image successfully, which reduces the accuracy of positioning.

3.3. The Proposed SI-GMS algorithm

The above two algorithms have some disadvantages when applied to the actual scene matching independently. Therefore, a combination image matching algorithm based on SI and GMS is proposed, the advantages of SI on fast matching and GMS on accurate matching are all possessed.

Step 1: SI algorithm is used to calculate the similarity β' of the reference image R and template image T_i , where $T_i \in \mathbf{T}$, \mathbf{T} is the set of all template images, and T_i is the i -th template. The threshold value h can be slightly adjusted to avoid omissions in judgment. If $\beta' < h$, the two images are considered relatively similar, then go to the next step; otherwise, read the next frame of the reference image.

Step 2: the longitude and latitude coordinates calculated by INS are read to calculate the distance to the location of the template image successfully matched:

$$D = \sqrt{\left[\frac{T_i(x) - I_t(x)}{180} \times \pi R\right]^2 + \left[\frac{T_i(y) - I_t(y)}{180} \times \pi R\right]^2}, \quad (18)$$

where $\{T_i(x), T_i(y)\}$ represents the latitude and longitude coordinates of template T_i , $\{I_t(x), I_t(y)\}$ represents the coordinate value returned by INS at time t , R is the radius of the earth.

When $D \leq \sigma$, the data is considered valid and the next step is executed, where σ is the maximum reasonable distance between the position obtained by INS and place matching.

Step 3: the matching results that satisfy the distance constraint are checked by GMS algorithm. If the number of successful matching points N between the two images is greater than the threshold value Ω , it is considered as a correct set of matching points; otherwise, it is omitted.

$$U = \begin{cases} True & N \geq \Omega \\ False & N < \Omega' \end{cases} \quad (19)$$

where U represents the matching result.

The completion of the above three steps helps the system meet the real-time matching requirements and effectively eliminates some matching points of misjudgment. The proposed SI-GMS algorithm is shown as follows:

Algorithm: SI-GMS algorithm

- 1: Calculate the scanline intensity of two images.
 - 2: Find the minimum difference β' of the intensity distribution using Equation (12)
 - 3: **If** $\beta' < h$ **then**
 - 4: Read the latitude and longitude information returned by INS
 - 5: Calculate the distance D between the coordinates of the template and the coordinates returned from INS using Equation (18)
 - 6: **If** $D < \sigma$ **then**
 - 7: Detect the feature points of the two images
 - 8: Match the feature points
 - 9: Use GMS to eliminate erroneous matching points
 - 10: Counting the number of remaining points N
 - 11: **If** $N > \Omega$
 - 12: Output the latitude and longitude coordinates of the current face template image
 - 13: **Return** *True*
 - 14: **End if**
 - 15: **End if**
 - 16: **End if**
 - 17: **Return** *False*
-

We used some groups of images from KITTI, which is a publicly available standard dataset, to test the proposed SI-GMS algorithm. As shown in Figure 8:



Figure 8. Two groups of images from KITTI.

The similarity values β' of these two groups of images are calculated as 0.051 and 0.039 respectively. In addition, the number of the successfully matching points after using the GSM algorithm are 114 and 178 respectively. As shown in Figure 9:

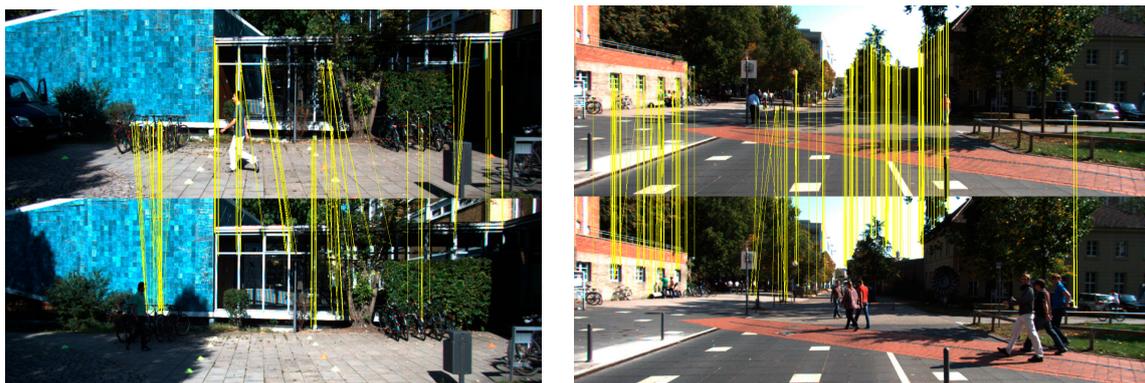


Figure 9. The successfully matching points after using the grid-based motion statistics (GMS) algorithm.

4. Experimental Test and Comparison

4.1. Experiment

To verify the proposed brain-like navigation scheme and the SI-GMS algorithm, experiments were carried out. The experimental results fully demonstrate the reliability of the proposed scheme and the superiority of the proposed algorithm. The experimental platform used in the experiment is a six-rotor UAV equipped with a self-made brain-like navigation system, which consists of a LattePanda (a kind of card type computer), a MEMS INS (consisted by STIM 202 and 1521 L), a camera and a high-precision GPS device (NovAtel ProPak6) used as reference (Figure 10). We used a laser ranging module (TW10S) to make the height of the UVA relative to the ground unchanged. Detailed equipment parameters are given in Tables 1–3.

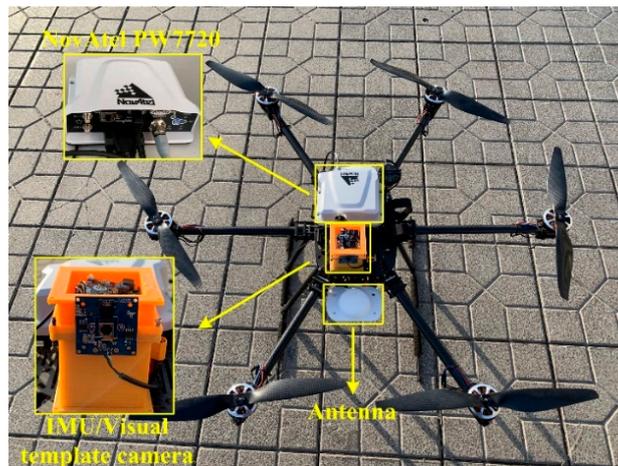


Figure 10. Experimental facilities.

Table 1. Characteristics of the Micro-Electro-Mechanical System based Inertial Navigation Systems (MEMS-INS)/GPS.

Gyroscope (STIM202)	Bias	0.5° √h
	Scale Factor	200 ppm
	Random Walk	0.2°/√h
Accelerometer(1521L)	Bias	0.5%–1%
	Scale factor	0.5%–1%
GPS (NovAtel ProPak6)	Position accuracy	1 cm + 1 ppm
	Velocity accuracy	0.03 m/s
	Time accuracy	20 ns

Table 2. The performance parameters of the camera.

Performance	Parameter
Power	5 V/240 mA
Working temperature	−10~70 °C
Sensor	Sony IMX179
Lens Size	1/3.2 inch
Pixel size	1.4 μm
Signal-to-noise Ratio	40 dB/69 dB
Pixel size	1280 × 720
Frame rate	12 fps
Low illumination	0.5 lux
Sensitivity	0.65 V/lux-sec@550 nm

Table 3. The performance parameter of LattePanda.

Performance	Parameter
Processor	Intel Cherry Trail Z8350 Quad Core 1.8 GHz
Operation System	Windows 10
Ram	4 GB DDR3L
Storage Capability	64 GB
GPU	Intel HD Graphics 200-500 MHZ
Power	5 V/2 A
Dimension of board	88 × 70 mm/ 3.46 × 2.76 inches

LattePanda is a single board computer with Windows 10 operating system and rich I/O interfaces, mainly undertaking the task of implementing the image matching algorithm in the experiment. The INS

mainly acquire the position, velocity and direction information. The INS communicate with LattePanda through the serial port protocol, so that the precise position information fed back by LattePanda after place matching can be received in real time and its position can be corrected when the place cell is activated. The power of the entire device is supplied by a USB and placed at the front of the six-rotor UAV. A camera is connected to the LattePanda to capture RGB images at 12 fps, which is then used for matching with the visual template when the UAV is flying. High-precision GPS is applied to provide the reference coordinate while acquiring the template images.

We conducted two sets of experiments. The experimental sites are located on a direct flight area about 2 m from the ground and 500 m in length for experiment 1 and an approximately 280-meter-long folding airspace for experiment 2 (shown as the red tracks in Figure 11), and the UAV flight at a speed of around 5 m/s. The position of the nodes of place cells for the visual matching is as shown in Figure 11. The place cells are activated for three times throughout the entire process for each experiment (at the green points in the Figure 11).



Figure 11. The driven routes and the places of the nodes for visual matching.

4.2. Comparison of Algorithms

Before the experiment is carried out, a ROI (region of interest) should be delimited first, therefore the computation cost can be further reduced, and the matching result would be more accurate and reliable. The area should include as many scenario features as possible on both sides of the road and try to avoid uncertainties such as pedestrians and vehicles on the road (the area inside the blue rectangle in Figure 12).

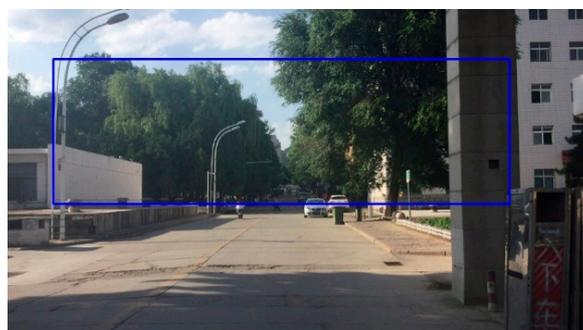


Figure 12. The ROI in this experiment.

In this experiment, the value of the threshold h of the SI algorithm is set as 0.075 by experiences, and the scanline profiles at the place where the nodes of place cells are located are shown as the following:

The three similarity values β' of the three nodes shown in Figure 13 are 0.069, 0.066 and 0.0704, respectively.

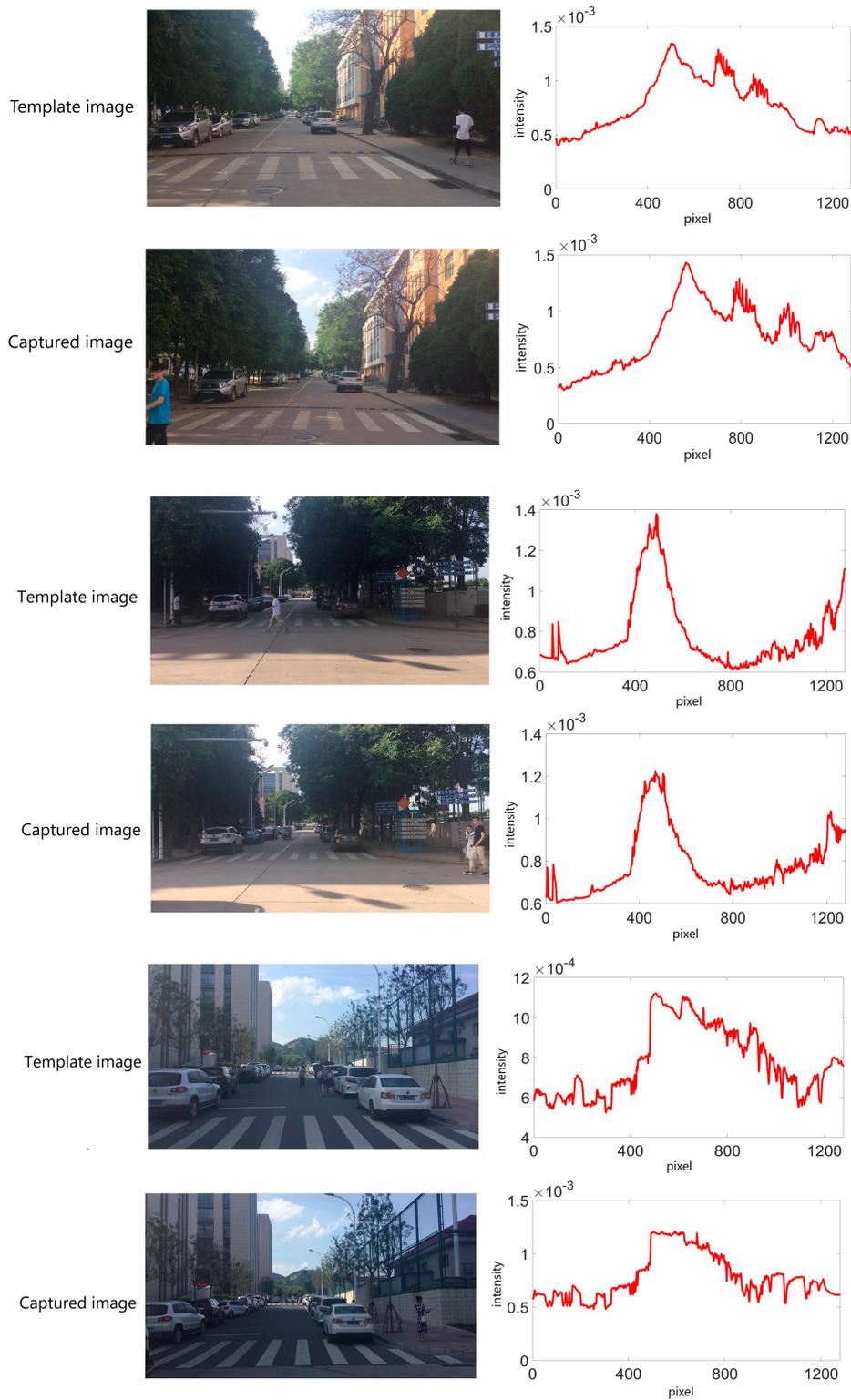


Figure 13. The captured images and their scanline profiles at the place where the three templates are captured.

However, the misjudgment of place matching happens sometimes when the SI algorithm is used independently. For example, in experiment 1, a misjudgment happens, as shown in Figure 14. The calculated β' value there is 0.0743 (Figure 14), which is smaller than the threshold $h = 0.075$.

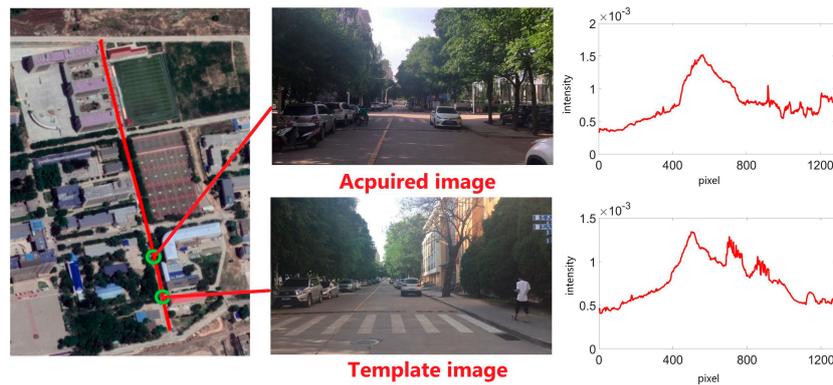


Figure 14. The specific locations on the map, raw images and their scanline profiles where the misjudgment happened. (The value of β' of these two images is 0.0743.).

Therefore, the SI algorithm, which is only based on the scanline intensity profile, is considered not reliable, and would cause serious consequences in the actual application if misjudgment occurs.

The scanline intensity of the acquired images is calculated and compared to the scanline intensity of the template images, and the acquired image would be retained, and its location would be thought of as a potential place cell node if the calculated similarity is less than 0.075. Then the position constraint is applied, which means that the potential node is matched properly only if the distance between the coordinate value given by the template image and INS is less than 100 m. After that, the template image and the acquired image are matched by using the GMS algorithm. The matching result is considered reliable if the number of matching points (N) that are successfully matched meets the constraint of $N \geq \Omega$. The place cell is then firing to reset. The constraint of $\Omega = 75$ is set for this experiment, and a total of no more than 500 characteristic points are detected for a single image. The experimental results indicate that the proposed SI-GMS algorithm can effectively improve the recognition accuracy of place cells, as shown in Figure 15. From Figure 15 we can see that: at the nodes of the place cells, the numbers of feature points where the real-time acquired images successfully match the visual template are 115, 248 and 100 for experiment 1, and 97, 183 and 204 for experiment 2, respectively. Therefore, the three nodes could be considered as trustworthy nodes of place cells.

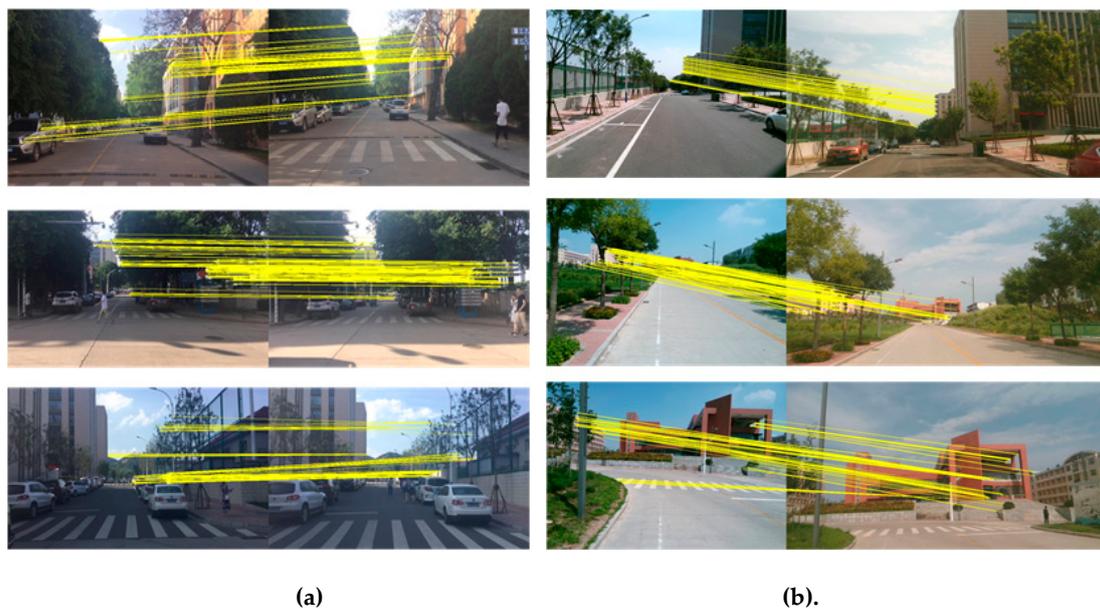


Figure 15. The result of the successfully matched feature points between real-time acquired images and template images by using the SI-GMS algorithm. (a) Experiment 1; (b) Experiment 2.

However, as shown in Figure 16, at the misjudged location by the SI algorithm, only three characteristic points are successfully matched by using the GMS algorithm, which means that the misjudgments could be effectively avoided.



Figure 16. The number of the successfully matched feature points of these two images is only three, whereas the value of β' calculated by the scanline intensity algorithm is 0.0743.

By using LattePanda, the time of matching a pair of images only takes 53 ms by GMS. However, in practical applications, there would be many template images. This means that each acquired image is matched to multiple template images. Take the number of templates as 10 as an example, the matching time cost approximately 624 ms when using GMS algorithm, which is impossible to meet real-time requirements. Whereas, the SI-GMS algorithm proposed in this paper only performs the GMS algorithm on images of which the SI algorithm matches successfully. It takes no more than 27 ms to match per frame and 76 ms for 10 template images on average, then the real-time application requirements can be reached. The algorithm execution time mentioned above is the time when the algorithm runs on the CPU and ignore the initialization time of the algorithm. Table 4 shows the execution time comparison of several algorithms.

Table 4. Comparison of the execution time.

	SI	GMS	SI-GMS (Average)
Single template	<10 ms	53 ms	27 ms
Ten templates	58 ms	624 ms	76 ms
Twenty templates	110 ms	1335 ms	148 ms

4.3. Comprehensive Comparison Results

The final experimental results are shown in Figure 17:

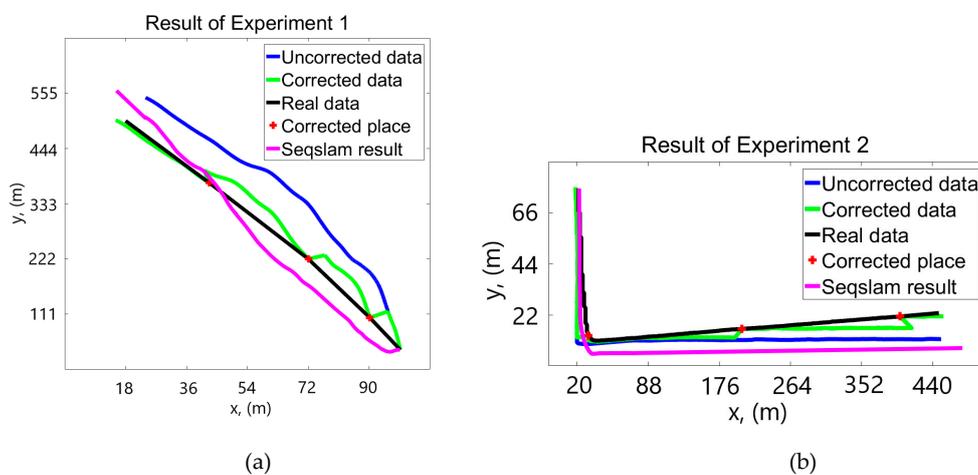


Figure 17. Experimental results and the comparison of different navigation schemes. (a) Comparisons of algorithmic results in experiment 1. (b) Comparisons of algorithmic results in experiment 2.

As shown in Figure 17, the black curve is the true UAV trajectory, the blue curve is the trajectory measured by pure INS, the purple curve is the trajectory of the results after using the Seqslam algorithm, and the green curve is the trajectory measured by the proposed brain-like navigation scheme. We can see that the positioning error of INS is accumulated when it works. However, after place cell recognition and position correction, the accumulation of INS errors could be effectively reduced, thus the navigation accuracy is improved. In experiment 1, according to the high-precision GPS reference information, the actual coordinates at the terminus are 38.017588°N, 112.44490°E, whereas the coordinates calculated by pure INS are 38.017921°N, 112.44493°E, and the error distance is about 37.2 m. Meanwhile, by using the proposed brain-like navigation scheme, the coordinates at the terminus calculated are 38.017513°N, 112.44489°E, and the error distance is only 8.4 m. It turns out that the navigation accuracy of INS has been improved by about 28.8 m per 500 m. As for experiment 2, the value of this data is about 16.6 m per 500 m. Tables 5 and 6 gives the latitude and longitude information of two groups of experiment terminus, and Table 7 shows the error comparison of several methods.

Table 5. Coordinates at terminus of experiment 1.

	High Precision GPS	Pure INS	Brain-Like Navigation Scheme	SeqSLAM
North latitude (°)	38.017588	38.017921	38.017513	38.018032
East longitude (°)	112.44490	112.44493	112.44489	112.44489

Table 6. Coordinates at terminus of experiment 2.

	High Precision GPS	Pure INS	Brain-Like Navigation Scheme	SeqSLAM
North latitude (°)	38.016407	38.016305	38.016394	38.016269
East longitude (°)	112.44755	112.44756	112.44758	112.44771

Table 7. Distance error.

	Pure INS	SeqSLAM	Brain-Like Navigation Scheme
Error of expt.1(m/500 m)	37.2	53.1	8.4
Error of expt.2(m/500 m)	22.8	56.4	6.2

It is obvious that the positioning accuracy can be improved by using the proposed brain-like navigation scheme compared to the pure vision-based algorithm and pure INS.

5. Conclusions

This paper has proposed a brain-like navigation system scheme to improve the intelligence level of the INS. Compared with pure INS, by incorporating visual information, the scheme can effectively eliminate the accumulated error of INS. Aiming at the place recognition problem in the scheme, this paper also proposes a place recognition algorithm that takes both image gray information and image feature points information into consideration. Experimental results demonstrate that this algorithm is superior to the traditional image matching algorithm in terms of calculation speed and accuracy, and this scheme is also practical and reliable.

As for future work, more sensors can be introduced into the system for further promotion, like polarized light and magnetic sensors. In order to save time when creating the template image database, we are trying to replace the existing database with image data from Google Earth. We also consider using the position error information to refine the equivalent biases of the gyroscopes and accelerometers, which is a major source of errors in the unconstrained IMU integration. Besides, we are going to perform more longer experiments that include indoor complex areas to test the algorithm further.

Author Contributions: C.S. and H.C. contributed equally to this paper. Conceptualization, C.S., J.L.; methodology, X.L., H.H.; software, X. G., X.C., H.C.; writing—original draft preparation, Y.Z., J.T.

Funding: This work was supported in part by the National Natural Science Foundation of China (61603353, 51705477, 51705021, U1764261, 91420203, 61703098), and the Pre-research Field Foundation (6140518010201) and the Scientific and Technology Innovation Programs of Higher Education Institutions in Shanxi (201802084), and the Shanxi Province Science Foundation the Pre-research Field Foundation (6140518010201) and the Scientific and Technology Innovation Programs of Higher Education Institutions in Shanxi (201802084), the Shanxi Province Science Foundation for Youths (201601D021067), the Program for the Top Young Academic Leaders of Higher Learning Institutions of Shanxi, the Young Academic Leaders Foundation in North University of China, Science Foundation of North University of China (XJJ201822), the Fund for Shanxi “1331 Project” Key Subjects Construction, Natural Science Foundation of Jiangsu Province under Grant BK20160699.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhao, L.; Wu, M.; Ding, J.; Kang, Y. A Joint Dual-Frequency GNSS/SINS Deep-Coupled Navigation System for Polar Navigation. *Appl. Sci.* **2018**, *8*, 2322. [[CrossRef](#)]
2. Cao, H.; Zhang, Y.; Shen, C.; Liu, Y.; Wang, X. Temperature energy influence compensation for MEMS vibration gyroscope based on RBF NN-GA-KF method. *Shock and Vibration.* **2018**, *2018*, 2830686. [[CrossRef](#)]
3. Cao, H.; Zhang, Y.; Han, Z.; Shao, X.; Gao, J.; Huang, K.; Shi, Y.; Tang, J.; Shen, C.; Liu, J. Pole-Zero-Temperature Compensation circuit design and experiment for dual-mass MEMS gyroscope bandwidth expansion. *IEEE/ASME Trans. Mechatron.* **2019**, *24*, 677–688. [[CrossRef](#)]
4. Shen, C.; Yang, J.; Tang, J.; Liu, J.; Cao, H. Note: Parallel processing algorithm of temperature and noise error for micro-electro-mechanical system gyroscope based on variational mode decomposition and augmented nonlinear differentiator. *Rev. Sci. Instrum.* **2018**, *89*, 1–3. [[CrossRef](#)] [[PubMed](#)]
5. Li, Z.; Wang, R.; Gao, J.; Wang, J. An Approach to Improve the Positioning Performance of GPS/INS/UWB Integrated System with Two-Step Filter. *Remote Sens.* **2018**, *10*, 19. [[CrossRef](#)]
6. Guo, H.; Zhu, Q.; Tang, J. A temperature and humidity synchronization detection method based on microwave coupled-resonator. *Sens. Actuators B* **2018**, *261*, 434–440. [[CrossRef](#)]
7. Valiente, D.; Gil, A.; Paya, L.; Sebastian, J.M.; Reinoso, O. Robust Visual Localization with Dynamic Uncertainty Management in Omnidirectional SLAM. *Appl. Sci.* **2017**, *7*, 1294. [[CrossRef](#)]
8. Wei, W.; Gao, Z.; Gao, S.; Jia, K. A SINS/SRS/GNS Autonomous Integrated Navigation System Based on Spectral Redshift Velocity Measurements. *Sensors* **2018**, *18*, 1145. [[CrossRef](#)]
9. Zhao, H.; Xu, W.; Zhang, Y.; Li, X.; Zhang, H.; Xuan, J.; Jia, B. Polarization patterns under different sky conditions and a navigation method based on the symmetry of the AOP map of skylight. *Opt. Express* **2018**, *26*, 28589–28603. [[CrossRef](#)]
10. Wang, Z.; He, G.; Du, W.; Zhou, J.; Han, X.; Wang, J.; He, H.; Guo, X.; Wang, J.; Kou, Y. Application of Parameter Optimized Variational Mode Decomposition Method in Fault Diagnosis of Gearbox. *IEEE Access* **2019**, *7*, 44871–44882. [[CrossRef](#)]
11. Nemra, A.; Aouf, N. Robust INS/GPS Sensor Fusion for UAV Localization Using SDRE Nonlinear Filtering. *IEEE Sens. J.* **2010**, *10*, 789–798. [[CrossRef](#)]
12. Guo, H.; Chen, Y.; Wu, D.; Zhao, R.; Tang, J.; Ma, Z.; Xue, C.; Zhang, W.; Liu, J. Plasmon-enhanced sensitivity of spin-based sensors based on a diamond ensemble of nitrogen vacancy color centers. *Opt. Lett.* **2018**, *42*, 403–406. [[CrossRef](#)]
13. Wang, Z.; Zhou, J.; Wang, J.; Du, W.; Wang, D.; Wang, J.; Han, X.; He, G. A novel Fault Diagnosis Method of Gearbox Based on Maximum Kurtosis Spectral Entropy Deconvolution. *IEEE Access.* **2019**, *7*, 29520–29532. [[CrossRef](#)]
14. Shen, C.; Song, R.; Li, J.; Zhang, X.; Tang, J.; Shi, Y.; Liu, J.; Cao, H. Temperature drift modeling of MEMS gyroscope based on genetic-Elman neural network. *Mech. Syst. Sig. Process.* **2016**, *72–73*, 897–905.
15. Song, C.; Wang, X.; Cui, N. Mixed-Degree Cubature H-infinity Information Filter-Based Visual-Inertial Odometry. *Appl. Sci.* **2019**, *9*.
16. Diba, K.; Buzsáki, G. Forward and reverse hippocampal place-cell sequences during ripples. *Nat. Neurosci.* **2007**, *10*, 1241–1242. [[CrossRef](#)]

17. Kropff, E.; Carmichael, J.E.; Moser, M.B. Speed cells in the medial entorhinal cortex. *Nature* **2015**, *523*, 419. [[CrossRef](#)]
18. Burgess, N.; Barry, C.; O'Keefe, J. An oscillatory interference model of grid cell firing. *Hippocampus* **2007**, *17*, 801–812. [[CrossRef](#)]
19. Zhang, K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *J. Neurosci.* **2010**, *16*, 2112–2126. [[CrossRef](#)]
20. Milford, M. Vision-based place recognition: how low can you go? *Int. J. Rob. Res.* **2013**, *32*, 766–789. [[CrossRef](#)]
21. Ma, J.; Zhao, J.; Tian, J.; Yuille, A.L.; Tu, Z. Robust Point Matching via Vector Field Consensus. *IEEE Trans. Image Process.* **2014**, *23*, 1706–1721.
22. Ma, J.; Jiang, J.; Zhou, H.; Zhao, J.; Guo, X. Guided locality Preserving Feature Matching for Remote Sensing Image Registration. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4435–4447. [[CrossRef](#)]
23. Yan, K.; Han, M. Aerial image stitching algorithm based on improved GMS. In Proceedings of the 2018 Eighth International Conference on Information Science and Technology (ICIST), Cordoba, Spain, 30 June–6 July 2018; pp. 1–6.
24. Zhang, W.; Shui, P. Contour-based corner detection via angle difference of principal directions of anisotropic Gaussian directional derivatives. *Pattern Recognit.* **2015**, *48*, 2785–2797. [[CrossRef](#)]
25. Kahaki, S.M.M.; Nordin, M.J.; Ashtari, A.H. Contour-Based Corner Detection and Classification by Using Mean Projection Transform. *Sensors* **2014**, *14*, 4126–4143. [[CrossRef](#)]
26. Kahaki, S.M.M.; Nordin, M.J.; Ashtari, A.H.; Zahra, S.J. Deformation invariant image matching based on dissimilarity of spatial features. *Neurocomputing* **2016**, *175*, 1009–1018. [[CrossRef](#)]
27. Saitoh, F. Image matching based on relation between pixels located on the maximum and minimum gray-levels in local area. *IEEJ Trans. Electr. Electr. Eng.* **2007**, *2*, 169–178. [[CrossRef](#)]
28. Kahaki, S.M.M.; Nordin, M.J.; Ashtari, A.H.; Zahra, S.J. Invariant Feature Matching for Image Registration Application Based on New Dissimilarity of Spatial Features. *PLoS ONE* **2016**, *11*.
29. Zhang, L.L.; Koch, R. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency. *J. Vis. Commun. Image Represent.* **2013**, *24*, 794–805. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).