Multi-resolution Visual Positioning and Navigation Technique for Unmanned Aerial System Landing Assistance

Chong Yu[1], Jiyuan Cai[2] and Qingyu Chen[3]

[1](*Asia-Pacific Research and Development Ltd, Intel Corporation, Shanghai, 200241, China*)
[2](*School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai, 200240, China*)
[3](*Robotics Institute, University of Michigan, Ann Arbor, MI, 48105, USA*)
(E-mail: dxxzdxxz@126.com)

To achieve more accurate navigation performance in the landing process, a multi-resolution visual positioning technique is proposed for landing assistance of an Unmanned Aerial System (UAS). This technique uses a captured image of an artificial landmark (e.g. barcode) to provide relative positioning information in the X, Y and Z axes, and yaw, roll and pitch orientations. A multi-resolution coding algorithm is designed to ensure the UAS will not lose the detection of the landing target due to limited visual angles or camera resolution. Simulation and real world experiments prove the performance of the proposed technique in positioning accuracy, detection accuracy, and navigation effect. Two types of UAS are used to verify the generalisation of the proposed technique. Comparison experiments to state-of-the-art techniques are also included with the results analysis.

1. INTRODUCTION.   Due to the decreasing cost and increasing capability, Unmanned Aerial Systems (UAS) are widely employed in military, commercial and personal scenarios, including reconnaissance, surveillance, photography, entertainment, etc. Due to low cruising capability and power storage, common UAS often need frequent take offs and landings in real applications. As a fundamental requirement, accurate landing is critical. However, this is a very tough task for many environments. At higher attitude, UAS hardly encounter obstacles, but when the UAS approach the ground, trees, buildings, high-voltage power cables, ground vehicles and so on will cause a lot of trouble. Accidents occur with high frequency during UAS landing processes. To improve this and reduce equipment damage,

it is essential to provide an accurate, reliable positioning and navigation technique during assisted landing.

In practice, Global Navigation Satellite System (GNSS)-based assisted positioning and traditional vision-based assisted positioning techniques are two common methods for navigation in UAS landing processes, however each of them has obvious defects.

1.1.   *GNSS-based assisted positioning technique.*   The most widely used GNSS in UAS outdoor positioning is the Global Positioning System (GPS). However, due to the following defects, GPS-assisted positioning and navigation techniques are not suitable for accurate landing processes in outdoor and indoor environments (Lee et al., 2008).

- For adopting single frequency GNSS (Global Navigation Satellite System) receivers, GPS positioning accuracy for civil usage is only 2–5 metres in commercial UASs (Groves, 2011).
- If adopting double frequency GNSS receivers and differential real time corrections (known as Real Time Differential GNSS Technique), civil users can achieve centimetre-level positioning accuracy. This technique is used in some high-end commercial UASs like the DJI Matrice 210 RTK. However, because of the system complexity and expense, this technique is not widely used in middle-end and low-end commercial UASs.
- Military GNSS systems are expensive and usually not available for commercial UASs.
- GPS signals cannot generally be used in an indoor environment.

1.2.   *Traditional vision-based assisted positioning technique.*   In traditional vision-based positioning techniques, artificial patterns are used to mark the landing target (Li et al., 2013). Common marker shapes include: 'H', 'H' surrounded by circles (Yang et al., 2014), concentric circles (Cocchioni et al., 2014), different shapes within a round marker, etc.

- Even though these common artificial patterns can mark the landing target for UASs, they cannot feed back numerical positioning information. Without positioning information, the system cannot achieve an accurate landing performance.
- Another limitation is that all these patterns are of single resolution, and if only a part of the marker pattern is captured by a visual sensor, the positioning and assisted landing capability could be lost or unreliable.
- The traditional system cannot differentiate multiple landing targets in one area.

In this paper, a Multi-Resolution Visual Positioning and Navigation (MVPN) technique is proposed. This technology can fix the issues that earlier systems cannot solve.

In the next section, we review and analyse the current state-of-the-art in UAS positioning and assisted landing techniques. In Section 3, we describe the main principle of our MVPN technique. In Section 4, we focus on design and implementation of the multi-resolution visual positioning algorithm. Section 5 focuses on the assisted landing control strategy. Section 6 shows experiment results obtained in simulation and real flight. Section 7 concludes this work and sets out the direction of future work.

2.   CURRENT STATE-OF-THE-ART TECHNIQUES.   Many researchers have focussed on vision-based positioning and assisted landing problems by capturing the images of artificial or natural markers. In Guan and Bai (2012), the proposed method is based on image matching between the current view from a monocular video camera and a previously known database of geo-referenced images. The key points of the image features are extracted using SURF (Speeded-Up Robust Features). The feature tracking algorithm only considers natural landmarks, without the help of visual beacons or artificial landmarks. Yang et al. (2014) presents a novel solution for micro aerial vehicles to autonomously search for and land on an arbitrary landing site using real-time monocular vision. The multi-scale ORB (ORiented Brief) features are extracted and integrated into the monocular visual SLAM (Simultaneous Localisation And Mapping) framework for landing site detection. In Cocchioni et al. (2014), an autonomous vision navigation and landing system is proposed with a predefined marker based on ellipse geometry. The developed vision algorithm provides the measurement of UAS position with respect to the landing platform using the predefined visual marker. At the same time, this system can automatically switch to an estimation of position which is independent from the visual marker. In Jung et al. (2014), the authors design a novel landing marker using concentric circles and the letter 'H' to overcome the problem that the entire shape of the landing marker cannot be captured as the UAS approaches it due to the limited field of view of the UAS's camera. By calculating the conic parameters from portions of ellipse curves, they can estimate accurate relative positioning for the precise autonomous landing system. Lee et al. (2014) presents a novel method of cooperation between two UASs at high altitude and low altitude for autonomous navigation and landing. With high flexibility and extensive vision, the high-altitude UAS estimates the position of the low-altitude UAS and controls it to finish the tracking marker and landing processes. In Roozing and Göktoğan (2013), the authors propose a low-cost vision-based positioning solution for UAS. An on board infrared tracking sensor with built-in vision processing is used to detect infrared markers and a point-based pose estimation algorithm is implemented to obtain six Degree Of Freedom (DOF) positioning estimation at high rates. In Lange et al. (2009), the authors describe their method for multirotor UAS autonomous landing and position control. They propose a landing pad and a vision-based detection algorithm that estimates the Three-Dimensional (3D) position of the UAS relative to the landing pad. A cascaded controller structure stabilises velocity and position in the absence of GPS signals by using a dedicated optical flow sensor. Li et al. (2015) discussed the solution for a small-scaled quad-rotor UAS to autonomously search for and land on a pre-designed marker placed on a rooftop. The UAS navigates to the landing site based on GPS and transitions to vision-guided positioning once the UAS identifies the targeted marker. Benini et al. (2016) proposes a real-time system for pose estimation of a UAS using parallel image processing of a known artificial marker. The system exploits the capabilities of a high-performance Central Processing Unit/Graphics Processing Unit (CPU/GPU) embedded system to provide autonomous take-off and landing.

3.   PRINCIPLE OF MVPN SYSTEM.   In this paper, we propose a Multi-Resolution Visual Positioning and Navigation (MVPN) technique for UAS landing assistance. Firstly, we design the specific multi-resolution visual barcode and coding algorithm. The multi-resolution coding algorithm ensures the UAS will not lose target barcode detection due to limited visual angle or camera resolution. It ensures the provision of consistent navigation

for the UAS during the whole landing process. Secondly, we design the visual positioning algorithm. This algorithm uses the captured image of a multi-resolution barcode to provide six DOF of relative positioning information in the X, Y and Z axes, and yaw, roll and pitch orientations to the flying UAS. With the six DOF positioning information feedback, the UAS can achieve a more accurate landing performance both in indoor and outdoor environments.

3.1. *Six DOF positioning feedback relies on MVPN barcode's special design.* To achieve a precise navigation performance in the landing process, the UAS should know the accurate ground truth of the relative position between landing target and itself. The MVPN system uses a Two-Dimensional (2D) barcode (Olson, 2011) as the landmark of the expected landing target. When the camera carried by the UAS captures the MVPN barcode, the barcode image can provide six DOF of relative positioning information feedback to the flying UAS, including relative displacements in X, Y and Z axes, and relative yaw, roll and pitch orientations.

Thus, if the physical size of target MVPN barcode is known, and the UAS camera is calibrated, then the relative position and orientation transform between the MVPN barcode and the UAS camera in a World Coordinate System (WCS) can be determined. Here, Cartesian coordinates of the target MVPN barcode in WGS84 (World Geodetic System 1984) are used to represent its absolute position. Because WGS84 is also the reference system of GPS, the absolute position of target MVPN barcode in WCS is determined with the provided GPS location. Then through coordinate system transformation (detailed in the next section), the UAS can determine its absolute position in X, Y and Z axes and its absolute yaw, roll and pitch orientations from a single image of MVPN barcode.

3.2. *Differentiating multiple landing targets relies on unique barcode coding number.* In real applications, the visual angle and resolution of the UAS's camera are limited. To provide a consistent assisted landing signal for positioning and navigating the UAS, we propose a multi-resolution coding method to combine multiple nested barcodes. Each small nested barcode is encoded with a unique Identification (ID) number. This ensures the UAS can detect at least one barcode at varying heights during the whole landing process, and can parse absolute positioning information by decoding the ID number.

Our solution can provide a low-cost, high-scalability and easy to use positioning and navigation technique for assisted landing for various standardised UASs on the market.

4. MULTI-RESOLUTION VISUAL POSITIONING AND NAVIGATION. The MVPN barcode detection process is comprised of five phases. The overall flow chart is illustrated in Figure 1.

4.1. *Pre-processing.* In this phase, we change the original captured MVPN image into greyscale and blur the frame to facilitate the calculation of gradient. In this phase, blur operation is necessary and useful, because it can greatly reduce the small noise in the original frame. Moreover, the pre-processing operations can also simplify the calculation of gradient.

4.2. *Line Segments Detection.* A graph-based line segmentation method (Felzenszwalb and Huttenlocher, 2004) is adopted to precisely estimate all lines in the captured MVPN image in this phase. This method starts by calculating gradient direction and magnitude at each pixel. Then adjacent pixels with a similar gradient are clustered into connected components based on graph theory. If two components x, y satisfy the following
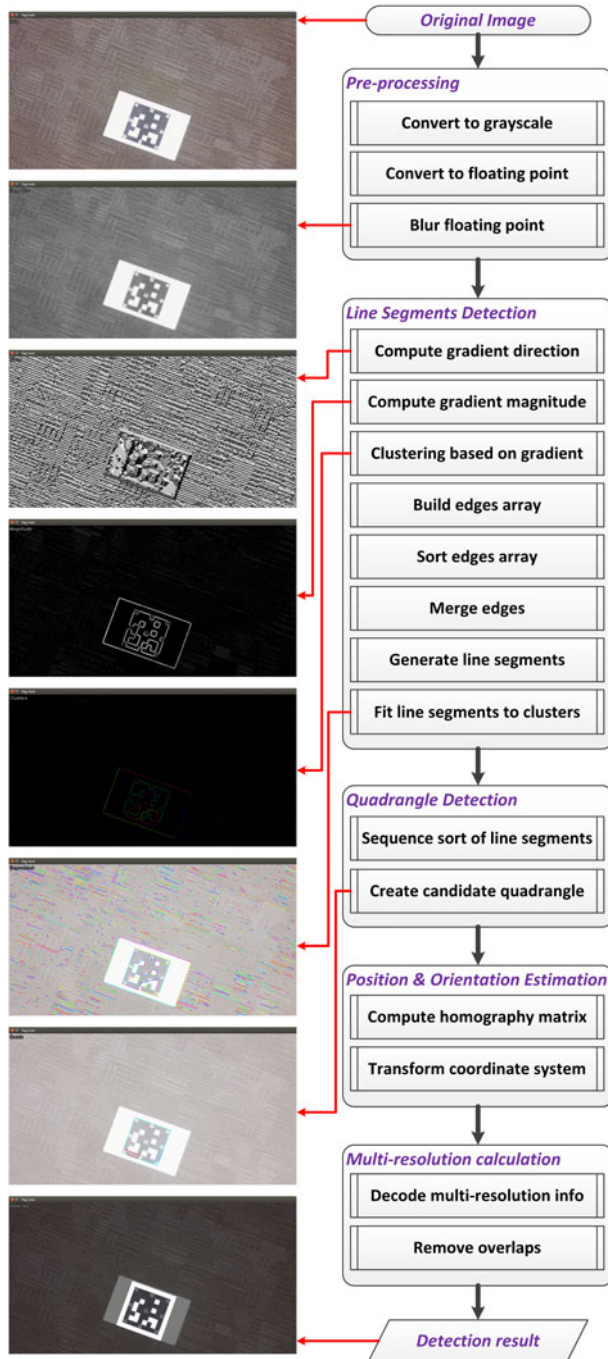
Figure 1. Flow chart of detection process for MVPN barcode.

two conditions, they are clustered together to generate a larger connected component.

$$D_g(x \cup y) \leq min\left(D_g(x), D_g(y)\right) + T_D/|x \cup y| \tag{1}$$

$$M_g(x \cup y) \leq min\left(M_g(x), M_g(y)\right) + T_M/|x \cup y| \tag{2}$$

Symbols $D_g()$ and $M_g()$ represent the difference between maximum and minimum values of gradient direction and magnitude respectively. $T_D/|x \cup y|$ and $T_M/|x \cup y|$ denote the intra-component variation threshold which shrinks as the connected component becomes larger.

4.3. *Quadrangle Detection.* The purpose of this phase is to find the quadrangle formed by a sequence of four line segments. The detection method is based on a recursive tree search with a depth of four. In the first depth, we choose one line segment as the tree root. From the second to the fourth depth, the line segment which is within the "gap threshold" to its previous line segment terminal point will be added as the child node. If four edges of this tree wind in same order (clockwise or counter-clockwise), the detection method regards it as a candidate quadrangle.

To improve the robustness of the quadrangle detection algorithm against uneven illumination, geometric transformation and even partial omissions, we set the "gap threshold" as a relatively large value. As a result, this setting leads to a very low false negative rate and a high false positive rate at the same time. The coding system in the following process is designed to reduce the false positive rate for the whole MVPN system.

4.4. *Position and Orientation Estimation.* This phase is to estimate the position transformation and orientation rotation from MVPN barcode in the 3D WCS to the 2D captured image coordinate system. The process needs three coordinate system transformations between the four coordinate systems.

The first transformation is from the 2D Pixel Coordinate System to the 2D Captured Image Coordinate System. Coordinate (u, v) represents the relative pixel position of target point (P_{Target}) in the 2D Pixel Coordinate System (u and v in pixels). Coordinate (x, y) represents the real physical position of P_{Target} in the 2D Captured Image Coordinate System (x and y in millimetres). Coordinate (u_0, v_0) represents the intersection of the optical axis and the 2D Captured Image Coordinate System. The real physical sizes of each pixel are depicted as dx, dy (millimetres/pixel). Then the transformation relation is expressed as follows.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{dx} & 0 & u_0 \\ 0 & \dfrac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3}$$

The second transformation is from the 2D Captured Image Coordinate System to the 3D Camera Coordinate System. Coordinate (X_c, Y_c, Z_c) represents the position of P_{Target} in the 3D Camera Coordinate System. f represents the camera focal length. According to the pinhole camera model, the transformation relation is expressed as follows.

$$Z_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{4}$$

From Equation (4) we can discover that Z_c represents the scale factor from the 2D Captured Image Coordinate System to the 3D Camera Coordinate System.

The third transformation is from the 3D Camera Coordinate System to the 3D WCS. Coordinate (X_T, Y_T, Z_T) represents the absolute position of P_{Target} in the 3D WCS. It is the Cartesian coordinates of P_{Target} in WGS84 (World Geodetic System 1984). WGS84 is an Earth-centred, Earth-fixed terrestrial reference system and geodetic datum, and it is the reference system for GPS. $\boldsymbol{R}$ represents the 3×3 orientation rotation matrix. $\boldsymbol{T}$ represents the 3×1 position transformation vector. According to the 3D object imaging theory, the transformation relation is expressed as follows.

$$X_c = R_{11}X_T + R_{12}Y_T + R_{13}Z_T + T_x$$

$$Y_c = R_{21}X_T + R_{22}Y_T + R_{23}Z_T + T_y$$

$$Z_c = R_{31}X_T + R_{32}Y_T + R_{33}Z_T + T_z$$

The elements $R_{n1}(n = 1, 2, 3)$ represent the rotation weights in the X axis. The elements $R_{n2}(n = 1, 2, 3)$ represent the rotation weights in the Y axis. The elements $R_{n3}(n = 1, 2, 3)$ represent the rotation weights in the Z axis. The elements T_x, T_y and T_z separately represent the position transformation weights in the X, Y and Z axes. The transformation relation can be summarised into matrix format:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \begin{bmatrix} X_T \\ Y_T \\ Z_T \\ 1 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_T \\ Y_T \\ Z_T \\ 1 \end{bmatrix} \tag{5}$$

Summarising the results of three times of the coordinate system transformation, Equation (6) can be derived from Equations (3), (4) and (5) as follows.

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{dx} & 0 & u_0 \\ 0 & \dfrac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \begin{bmatrix} X_T \\ Y_T \\ Z_T \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_T \\ Y_T \\ Z_T \\ 1 \end{bmatrix} \tag{6}$$

The homographic matrix can represent the projection relation from the MVPN barcode in the 3D WCS to the 2D Pixel Coordinate System. The 3×3 homographic matrix can be calculated with a Direct Linear Transform algorithm (Hartley and Zisserman, 2003). It can be expressed as follows.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s\boldsymbol{H} \begin{bmatrix} X_T \\ Y_T \\ Z_T \\ 1 \end{bmatrix} \tag{7}$$

Because the homographic matrix is the product of the 3×4 camera projection matrix and the 4×4 camera extrinsic matrix, from Equations (6) and (7), we can obtain Equation (8) as follows:

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = \frac{1}{sZ_C} \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

The parameters in the 3×4 camera projection matrix are determined by the camera's characteristics, which can be assumed as known in this paper. Nine elements of the 3×3 homography matrix can provide nine constraint conditions. Moreover, because the columns of the rotation matrix must be of unit magnitude and the columns of the rotation matrix must be orthonormal, there are four more constraint conditions. Then we can calculate the 13 unknown parameters including the scale factor: $1/sZ_C$, nine elements of the 3×3 orientation rotation matrix R and three elements of the 3×1 position transformation vector T. So eventually the position transformation and orientation rotation from the MVPN barcode in the 3D WCS to the 2D captured image coordinate system is well-determined.

4.5. *Multi-Resolution Calculation.*   In real applications, visual angles and resolution of UAS cameras are limited. To provide a consistent MVPN signal for navigating the UAS, we propose a multi-resolution coding algorithm to combine multiple nested MVPN barcodes. The more limited the visual angle and the lower the resolution, the more layers and nested barcodes are needed. Each small nested barcode is encoded with a unique ID number. This ensures the UAS can detect at least one barcode at varying heights during the whole landing process, and can parse absolute positioning information by decoding the ID number. One example of generated MVPN barcode with four layers and nine small barcodes nested in each layer is depicted in Figure 2.

When the UAS is at greater height, the nine smaller nested barcodes in the lower layer appear to be the black squares in the higher layer barcode, so the UAS's camera will obtain its positioning and navigation information from the higher layer barcode. When the UAS descends to a lower height, the camera can only capture a part of the higher layer barcode due to the limited visual angle. However, the lower layer nested barcodes can be captured with higher resolution. Then the UAS switches to obtain positioning and navigation information from the lower layer barcode. Since MVPN barcodes can be nested in a hierarchical structure, this can provide a robust and general multi-resolution coding solution to generate a MVPN barcode for the UAS to track during the whole landing process. Moreover, the coding system adds to the complexity of the barcode, as the square pattern consists of many rectangles and high complexity rarely occurs in a natural scene, so it helps to filter potential barcode candidates and lower the false positive ratio for the whole MVPN technique.

5.   ASSISTED LANDING CONTROL STRATEGY.   To set up the assisted landing control strategy, the UAS's flight state in the environment should first be defined. At each time moment t, UAS has a set of flight states $s_t \in S$, which includes its absolute positions in X, Y and Z axes, its absolute yaw, roll and pitch orientations, six speed components in X, Y
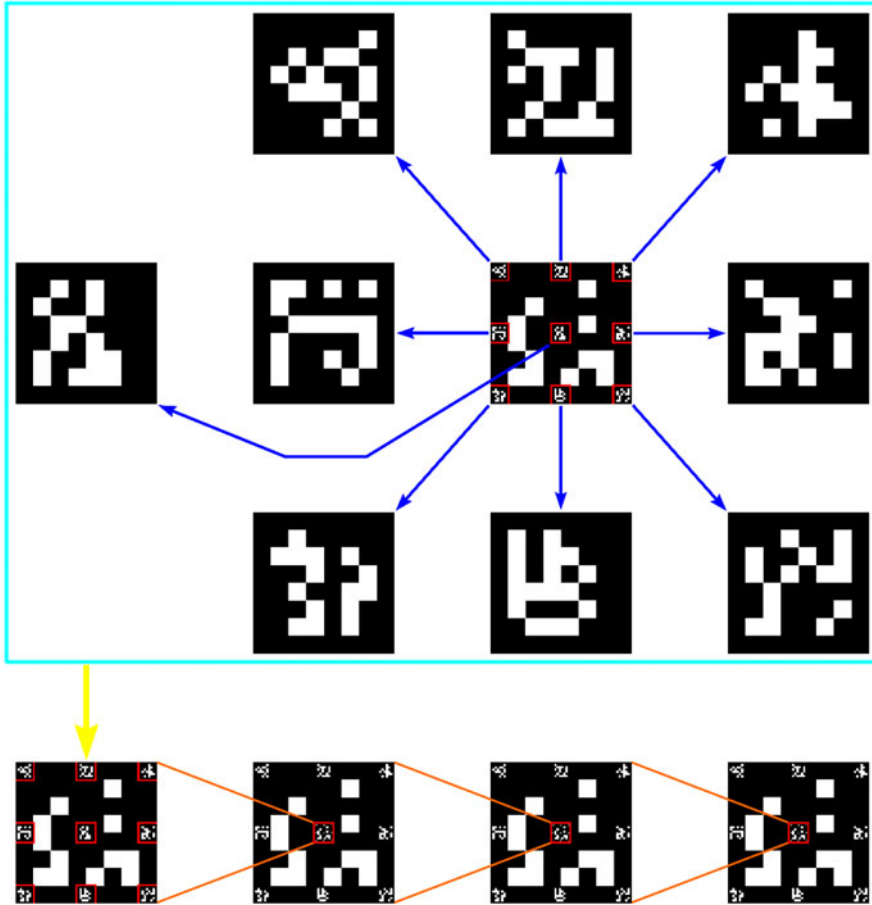
Figure 2. Four-layer nested MVPN barcode (Partial enlarged).

and Z axes and yaw, roll and pitch orientations, i.e.

$$S = \left\{ x, y, z, yaw, roll, pitch, v_x, v_y, v_z, v_{yaw}, v_{roll}, v_{pitch} \right\}$$

The final goal of landing assistance is navigating the UAS to land exactly on the target point $P_{Target} = (X_T, Y_T, Z_T)$ in 3D WCS, with no offset in yaw, roll and pitch orientations, and each speed component is at zero. So the final optimal landing state is: $S_T = \{X_T, Y_T, Z_T, 0, 0, 0, 0, 0, 0, 0, 0, 0\}$. The assisted landing control strategy is meant to:

- Minimise the total offsets in X, Y and Z axes.
- Minimise the total offsets in yaw, roll and pitch orientations.
- Reduce each speed component to zero.
- Minimise the total assisted landing time.

Figure 3.   UAS platform for experiments: Parrot AR. Drone 2.0 (left), DJI Mavic Pro (right).

Then the landing control strategy can be expressed as the following set of optimisation issues.

$$\begin{cases} \min_t \left\{ \omega_{XY} \left[(x - X_T)^2 + (y - Y_T)^2 \right] + \omega_Z \left(z - Z_T \right)^2 + \omega_{yrp} \left(yaw^2 + row^2 + pitch^2 \right) \right\} \\ \qquad \min_t \left\{ \omega_{v\_XYZ} \left(v_x^2 + v_y^2 + v_z^2 \right) + \omega_{v\_yrp} \left(v_{yaw}^2 + v_{row}^2 + v_{pitch}^2 \right) \right\} \\ \qquad\qquad \min \sum_{origination}^{Landing} t \end{cases}$$

Where ω_{XY} is the weight for position offsets in the X and Y axes, w_Z is the weight for position offset in the Z axis, ω_{yrp} is the weight for orientation offsets in yaw, roll and pitch, ω_{v_XYZ} is the weight for speed offset of speed components in X, Y and Z axes and w_{v_yrp} is the weight for speed offset of speed components in yaw, roll and pitch orientations. Different control strategies can be easily configured with the different combinations of these weight values. In the flight process, the values of yaw, roll and pitch orientations are derived from the Inertial Measurement Unit (IMU) installed on the UAS, while in the assisted landing process, the value of relative yaw, roll and pitch orientations are obtained from the MVPN barcode detection process.

6.   EXPERIMENTAL SETUP AND RESULTS.   In this paper, two types of UAS: the Parrot AR. Drone 2.0 and DJI Mavic Pro are used as the testing platforms to verify the feasibility and performance of the proposed MVPN technique, which are shown in Figure 3.

AR. Drone 2.0 is a quad-rotor UAS with a size of 55 cm from rotor-tip to rotor-tip and 380 to 420 grams in weight. It is equipped with a Micro-electromechanical System (MEMS)-based nine DOF miniaturised IMU, including an upgraded three-axis gyroscope, along with a three-axis accelerometer and magnetometer. The AR. Drone 2.0 uses a High Definition (HD) (720p, 30 frame per second (fps)) front-facing camera. This camera can be configured to stream either 360p (640×360) or 720p (1280×720) images. The bottom-facing camera is a QVGA (320×240, 60fps) camera. This camera pictures will be upscaled to 360p or 720p for video streaming.

DJI Mavic Pro is also a quad-rotor UAS with a diagonal size of 33·5 cm and is 734 to 743 grams in weight. It is equipped with an IMU, as well as GPS and GLONASS modules. DJI Mavic Pro uses a high-resolution front-facing camera. This camera can be configured to capture 4000×3000 resolution images, and to stream C4K (4096×2160, 24fps), HD (3840×2160, 30fps), 2·7K (2720×1530, 30fps), FHD (1920×1080, 96fps),
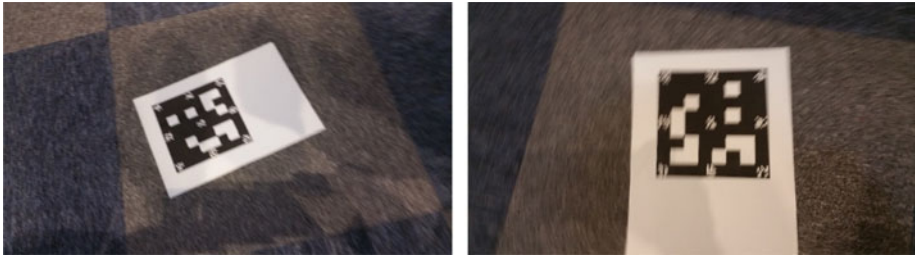
Figure 4.   MVPN barcode detection failure examples.

HD (1280 × 720, 120fps) videos. The gimbal can turn the front facing camera to allow working as a down-facing camera.

6.1.   *MVPN Barcode Detection Performance.*   To evaluate the MVPN barcode detection process, we captured 211 images with various UAS cameras, distances and angles. Then we achieved data augmentation by rotating each of the original images every one degree to generate a testing set of 77015 images. The detection accuracy of the MVPN barcode is 98.578%, while all the failure cases are due to images being excessively fuzzy. Some examples of detection failure images are shown in Figure 4.

The C++ implementation of the MVPN barcode detection system runs at 7.3 fps on 4000 × 3000 resolution images, 10 fps on 2592 × 1944 resolution images, and 30 fps on 320 × 240 resolution images with an Intel Core i7 CPU. From time consumption analysis, the major time cost comes from segmentation and clustering operations. In future research, we will accelerate the MVPN barcode detection algorithm with GPU as well as on-chip solutions.

6.2.   *Positioning and Navigation Performance in Simulation.*   To verify the positioning and navigation performance of the MVPN system, we built an experimental environment in Gazebo with Robot Operating System (ROS) libraries. We used the simulation environment because it can imitate various situations (different wind, height, velocity, payload weight, indoor, outdoor, etc.) for us to verify the performance of the MVPN technique in a safe manner. The experimental UAS type is AR. Drone 2.0.

The simulation landing experiments in indoor and outdoor environments are shown in Figures 5 and 6. From the main window, we can see the AR. Drone 2.0 hovering in the indoor or outdoor environment. The left upper window shows the image captured by the front-facing camera of the UAS. The left middle window shows the MVPN barcode image which is captured by the bottom-facing camera of the UAS. The left bottom window and the right upper window give the positioning and navigation feedback of positions in X, Y and Z axes and absolute yaw, roll and pitch orientations after processing the MVPN barcode image. The positioning and navigation error is only 0.7 cm in X, Y and Z axes, and 6.3° in yaw, roll and pitch orientations.

6.3.   *Positioning and Navigation Performance on Real UAS.*   To further check the positioning and navigation performance of the MVPN system, we made similar experiments on a real UAS. The experimental UAS type is DJI Mavic Pro. DJI mobile Software Development Kit (SDK) was utilised to send the control commands and obtain sensor feedback from the DJI Mavic Pro. Another UAS testing platform was adopted for the following reasons:
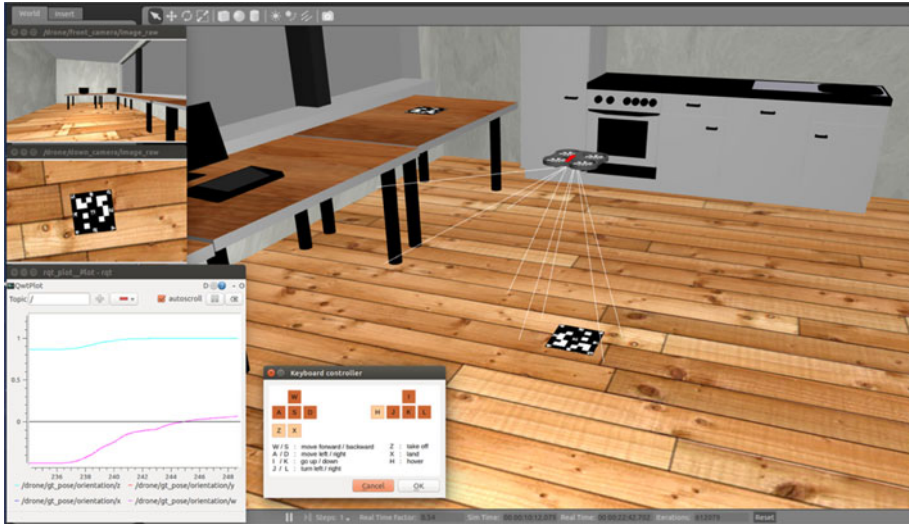
Figure 5.   Simulation UAS landing experiment with MVPN technique in indoor environment (AR. Drone 2.0).
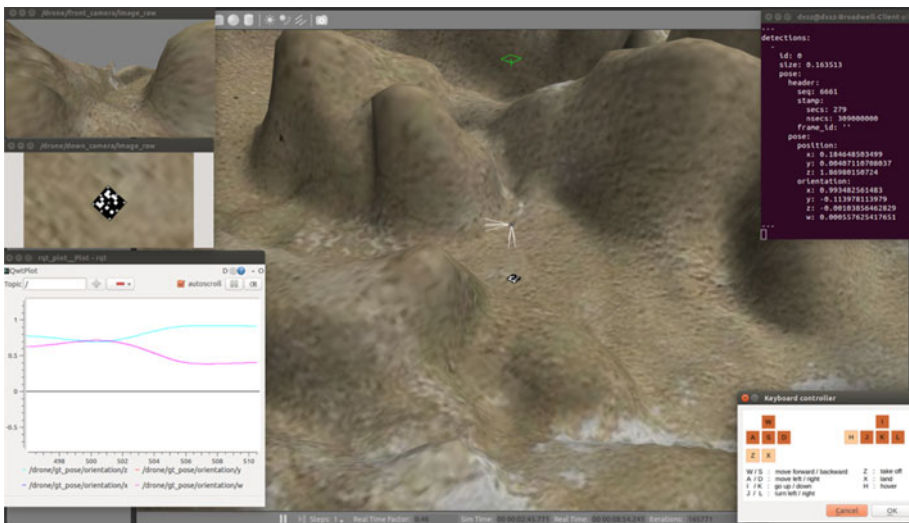

Figure 6.   Simulation UAS landing experiment with MVPN technique in outdoor environment
(AR. Drone 2.0).

- DJI Mavic Pro can turn its front facing camera to bottom facing, so it can provide higher resolution images when capturing the MVPN barcode.
- DJI Mavic Pro has a gimbal for camera stabilisation, which will improve the quality of captured MVPN images.
- AR. Drone 2.0 is very light in weight, so it can be affected by the surrounding airflow. DJI Mavic Pro is heavier in weight, so it is more suitable to be used as the landing performance testing platform, especially in the outdoor environment.
- By using two different UAS models, we can prove the generalisation capability of proposed the MVPN technique and assisted landing strategy.

Figure 7. Real UAS landing experiment with MVPN technique in indoor environment (DJI Mavic Pro).

The real landing experiment in an indoor environment is shown in Figure 7. We chose a lobby as the testing environment. The temperature of the indoor environment was $20°$ Celsius. There was nearly no wind and the landing altitude was 4.5 metres. The results of the real UAS landing experiment with the MVPN technique in an indoor environment are pretty close to ideal. The positioning and navigation error is only 0.3 cm in X, Y and Z axes, and $3.5°$ in yaw, roll and pitch orientations. In the simulation experiments, the resolution of the bottom-facing camera for AR. Drone 2.0 is only 320×240, while in the real indoor experiment, the resolution of the bottom-facing camera for DJI Mavic Pro is 1234×651. Thus, it is reasonable that the positioning and navigation performance is better in the real indoor experiment.

6.4. *Comparison Experiments.* Table 1 provides a comparison of the state-of-the-art research on vision-based positioning and assisted landing solutions. The positioning error is the square root of the sum of squares of positioning errors in X, Y, and Z axes.

$$E_{positioning} = \sqrt{E_X^2 + E_Y^2 + E_Z^2}$$

Table 1.   State-of-the-art vision-based UAS assisted navigation and landing systems.

References	Marker Shape	Heading Estimation	Asymmetry	Image Dimension	Positioning Error
Yang et al., 2014	'H' surrounded by a circle	Yes	No	640×480	1.5 cm
Cocchioni et al., 2014	Concentric circles + triangle for heading	Yes	Yes	640×480	2.0 cm
Guan and Bai, 2012	Natural landmarks	No	N/M[1]	320×240	N/M
Jung et al., 2014	'H' surrounded by 3 circles	No	No	N/M	N/M
Lee et al., 2014	'H'	No	No	176×144	$26 - 38$ cm
Roozing and Göktoğan, 2013	IR pattern	Yes	Yes	N/M	10.0 cm
Lange et al., 2009	Concentric circles	No	No	320×240	N/M
Li et al., 2015	Different shapes within the marker	Yes	Yes	N/M	10.0 cm
Benini et al., 2016	Concentric circles	Yes	Yes	640×480	2.0 cm
Our technique	MVPN barcode	Yes	Yes	Multi-resolution	0.3 cm/0.4 cm/ 0.7 cm[2]

[1]N/M is the abbreviation of Not Mentioned. It means the corresponding data is not provided by the reference papers.
[2]0.3 cm positioning error is achieved on 4000×3000 resolution image, 0.4 cm positioning error is achieved on 2592×1944 resolution image, while 0.7 cm positioning error is achieved on 320×240 resolution image.

From the comparison results, we can see our MVPN technique outperforms the state-of-the-art solutions on positioning accuracy and resolution flexibility.

7. CONCLUSION AND FUTURE DIRECTION.   In this paper, we propose a multi-resolution visual positioning and assisted navigation technique and prove its performance in simulation and real experiments. Comparing with existing solutions like GPS-based assisted positioning techniques and traditional vision-based assisted positioning techniques, our solution has the following advantages.

7.1. *VS. GPS-based assisted positioning technique.*   Centimetre-level positioning accuracy provided by the MVPN technique versus metre-level positioning accuracy of the GPS technique with single frequency GNSS receivers. GPS positioning accuracy for civil usage is only 2–5 metres with only single frequency GNSS receivers in commercial UASs. Our solution can achieve less than one centimetre positioning accuracy both in the indoor and outdoor environments. We have an enhanced design of a more condensed barcode grid on the ground (each barcode's size is 10 cm × 10 cm). Then we use the smaller barcodes to build up a giant barcode for far-sight recognition. Finally, we can provide centimetre-level accuracy positioning technique for commercial UASs from high-end to low-end types.

Because of its high cost, weight and volume, a real time differential GPS system with centimetre-level positioning accuracy cannot be widely used in all types of commercial UASs. By contrast, visual sensors are relatively low-cost, lightweight and small in volume. When close to the ground, visual sensors can typically provide more accurate positioning information than GPS, so our solution is a better choice to complete the positioning task in an assisted navigation and landing system, especially for middle-end and low-end commercial UASs.

A GNSS signal cannot be used in an indoor environment, while our solution is effective both in indoor and outdoor environments.

7.2. *VS. Traditional vision-based assisted positioning technique.* Traditional techniques use some common shapes like: 'H', 'H' surrounded by circles, and concentric circles to mark the landing place. However, these markers cannot provide relative positioning and numerical information to the UAS in the landing process. Without this positioning information feedback, the assisted landing effect is quite limited. In contrast, our visual positioning algorithm and multi-resolution visual barcode solution can provide six DOF of relative positioning information with regard to the landing place in X, Y and Z axes and yaw, roll and pitch orientations to the UAS. With this numerical positioning information feedback, our solution can achieve a more accurate landing effect both in indoor and outdoor environments.

In traditional techniques, another obvious limitation is that all the markers are of single resolution. If only a part of the marker pattern is captured by the UAS's visual sensor, the navigation and assisted landing capability would be lost. In our solution, we use a multi-resolution coding method to combine multiple nested barcodes. This ensures that the UAS can detect at least one barcode at varying heights during the whole landing process. Our solution can provide consistent navigation and assisted landing effect even if only a part of the whole MVPN barcode is captured by the UAS's visual sensor.

Our solution can embed much more landing navigation information than the traditional techniques. In our solution, we can fill the white areas with one set of barcodes (yellow + orange). The black areas can be filled with another set of barcodes (black + blue) as well. Then if the processing software uses a blue colour as the filter, the UAS's camera can still see the big black/white barcode. When the UAS comes closer, its camera can recognise the smaller barcodes with different colours.

Our solution is more extendable than the traditional techniques. In our solution, we can extend such MVPN barcodes to the vertical level. These can then provide precise height positioning and navigation for the best soft UAS landing.

7.3. *Future Direction.* In the future, we will go on with our research in the following directions.

In the current work, to achieve better computation performance, we use WiFi on the AR. Drone 2.0 to send the captured MVPN barcode to a Personal Computer for processing, then send the calculated navigation information back to the AR. Drone 2.0. This may lead to some latency in the real flight. We will try to optimise the MVPN detection algorithm in the embedded board to make the technique a full online solution.

The proposed MVPN barcode for UAS landing assistance can only be applied to the tasks performed on areas which are pre-prepared for landing the UAS. The application scope is restricted. There are also many scenarios in which a UAS needs to land in areas previously unreached or not equipped with the MVPN. One possible solution is to install the proposed MVPN barcode on the flat top of a ground vehicle. If the UAS needs to land in an unrecognised area, the ground vehicle can drive to that area as a moving airport. The UAS can safely land on the top of the ground vehicle through visual navigation provided by the MVPN barcode. The prototype we have made for this idea is shown in Figure 8.

In the 2016 DJI SDK Challenge, a similar challenge is mentioned, and requires participants to design an unmanned rescue aircraft which can take off from and land on a moving Ford pickup truck. We will further study the more general navigation techniques for UAS in the future.
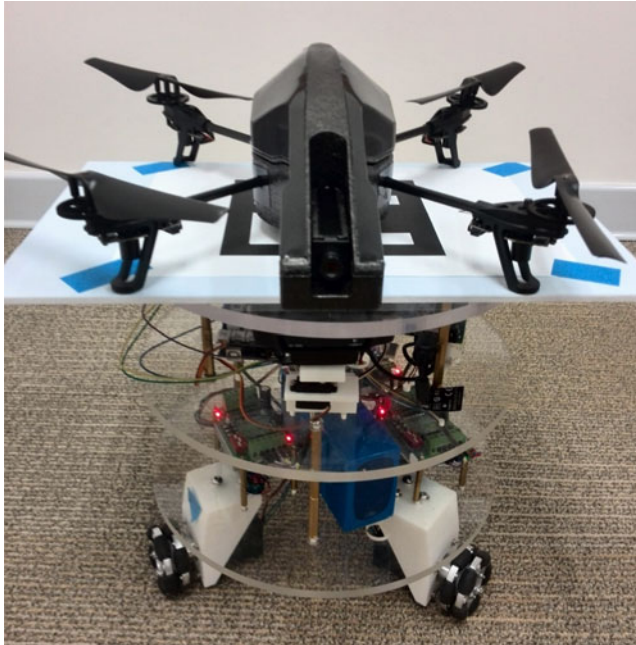
Figure 8.   Prototype of landing UAS on the top of a ground vehicle.

ACKNOWLEDGMENTS

REFERENCES

Benini, A., Rutherford, M.J. and Valavanis, K.P. (2016). Real-time, GPU-based pose estimation of a UAV for autonomous takeoff and landing. *IEEE International Conference on Robotics and Automation (ICRA)*, 3463–3470.

Cocchioni, F., Mancini, A. and Longhi, S. (2014). Autonomous navigation, landing and recharge of a quadrotor using artificial vision. *IEEE International Conference on Unmanned Aircraft Systems (ICUAS)*, 418–429.

Felzenszwalb, P.F. and Huttenlocher, D.P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, **59**(2), 167–181.

Groves, P.D. (2011). Shadow matching: A new GNSS positioning technique for urban canyons. *Journal of Navigation*, **64**(03), 417–430.

Guan, X. and Bai, H. (2012). A GPU accelerated real-time self-contained visual navigation system for UAVs. *IEEE International Conference on Information and Automation (ICIA)*, 578–581.

Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press.

Jung, Y., Lee, D. and Bang, H. (2014). Study on ellipse fitting problem for vision-based autonomous landing of an UAV. *IEEE 14th International Conference on Control, Automation and Systems (ICCAS)*, 1631–1634.

Lange, S., Sunderhauf, N. and Protzel, P. (2009). A vision based onboard approach for landing and position control of an autonomous multirotor UAV in GPS-denied environments. *IEEE International Conference on Advanced Robotics (ICAR)*, 1–6.

Lee, H.K., Soon, B., Barnes, J., Wang, J. and Rizos, C. (2008). Experimental analysis of GPS/Pseudolite/INS integration for aircraft precision approach and landing. *Journal of Navigation*, **61**(02), 257–270.

Lee, M.F.R., Su, S.F., Yeah, J.W.E., Huang, H.M. and Chen, J. (2014). Autonomous landing system for aerial mobile robot cooperation. *IEEE 15th International Symposium on Soft Computing and Intelligent Systems (SCIS), Joint 7th International Conference on and Advanced Intelligent Systems (ISIS)*, 1306–1311.

Li, K., Liu, P., Pang, T., Yang, Z. and Chen, B.M. (2015). Development of an unmanned aerial vehicle for rooftop landing and surveillance. *IEEE International Conference on Unmanned Aircraft Systems (ICUAS)*, 832–838.

Li, S., Lu, R., Zhang, L. and Peng, Y. (2013). Image processing algorithms for deep-space autonomous optical navigation. *Journal of Navigation*, **66**(04), 605–623.

Olson, E. (2011). AprilTag: A robust and flexible visual fiducial system. *IEEE International Conference on Robotics and Automation (ICRA)*, 3400–3407.

Roozing, W. and Göktoğan, A.H. (2013). Low-cost vision-based 6-DOF MAV localization using IR beacons. *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 1003–1009.

Yang, S., Scherer, S.A., Schauwecker, K. and Zell, A. (2014). Autonomous landing of mavs on an arbitrarily textured landing site using onboard monocular vision. *Journal of Intelligent & Robotic Systems*, **74**(1–2), 27–43.