A monocular-based navigation approach for unmanned aerial vehicle safe and autonomous transmission-line inspection

Xiaolong Hui, Jiang Bian, Xiaoguang Zhao and Min Tan

Abstract
This article presents a monocular-based navigation approach for unmanned aerial vehicle safe and continuous inspection along one side of transmission lines. To this end, a navigation model based on the transmission tower and the transmission-line vanishing point was proposed, and the following three key issues were addressed. First, a deep-learning-based object detection and a fast and smooth tracking algorithm based on the kernelized correlation filter were combined to locate transmission tower timely and reliably. Second, the vanishing point of transmission lines was computed and optimized to provide unmanned aerial vehicle with a robust and precise flight direction. Third, to keep a stable safe distance from transmission lines, the transmission lines were first rectified by optimizing a homography matrix to eliminate the parallel distortion, and then their interval variation was estimated for reflecting the spatial distance variation. Finally, the real distance from transmission tower was measured by the triangulation across multiple views. The proposed navigation approach and the designed UAV platform were tested in a field environment, which achieved an encouraging result. To the best of authors’ knowledge, this article marks the first time that a safe and continuous navigation approach along one side of transmission lines is put forward and implemented.

Keywords
UAV, safety navigation approach, intelligent transmission-line inspection, monocular vision

Date received: 24 August 2018; accepted: 12 January 2019

Introduction
Transmission lines and associated infrastructures are vital to economic construction. They are often exposed to the severe weather condition and lack the regular maintenance. Over the past few decades, autonomous transmission-line inspection has been a hot issue. Many research are typically conducted based on the following two platforms: multi-rotor unmanned aerial vehicles (UAVs) and climbing robots.\(^1\)\(^-\)\(^3\) By contrast, the UAV inspection is becoming more and more popular owing to its nice maneuverability.

For UAVs, safe and continuous autonomous inspection along overhead transmission lines has been a problem. A direct but effective way is to navigate UAVs by global positioning system (GPS). For GPS-based navigation systems, reliable recognition of transmission lines and real-time and precise GPS information of transmission towers are very crucial. Therefore, a navigation model based on the tower and the vanishing point of transmission lines was proposed, and the following three key issues were addressed. First, a deep-learning-based object detection and a fast and smooth tracking algorithm based on the kernelized correlation filter were combined to locate the tower timely and reliably. Second, the vanishing point of transmission lines was computed and optimized to provide the vehicle with a robust and precise flight direction. Third, to keep a stable safe distance from the transmission lines, the lines were first rectified by optimizing a homography matrix to eliminate the parallel distortion, and then their interval variation was estimated for reflecting the spatial distance variation. Finally, the real distance from the tower was measured by triangulation across multiple views. The proposed navigation approach and the designed UAV platform were tested in a field environment, which achieved an encouraging result. To the best of authors’ knowledge, this article marks the first time that a safe and continuous navigation approach along one side of transmission lines is put forward and implemented.
and UAV are indispensable. Lu et al.\textsuperscript{6} addressed the line-tracking issue utilizing GPS data of transmission towers during autonomous inspection. Luque et al.\textsuperscript{7} developed a quadrotor-based inspection system and achieved an autonomous flight along predefined GPS waypoints. However, such a system cannot work under the condition of unstable GPS and lacks the consideration of surroundings.

Despite the good performance of GPS-based navigation, the transmission-line-based navigation has also attracted a wide attention of researchers. Ceron et al.\textsuperscript{8} navigated UAV in a virtual environment by the direction of transmission lines and the intervals between transmission lines. Chen et al.\textsuperscript{9} made some improvements in the conventional radon transform to enable the sub-pixel extraction of transmission lines from the high-resolution satellite images. Cern et al.\textsuperscript{10} proposed a transmission line detection method based on the geometric relationships, which takes less time than Hough transform\textsuperscript{11} and has an advantage on long line detection. Zhou et al.\textsuperscript{12} proposed an edge detection method to overcome the problem of threshold selection when the background changed and achieved a fully automatic power line tracking. Tian et al.\textsuperscript{13} adopted a type of double-side filters to enhance power lines followed by recognition and tracking in Hough space by utilizing the parallel constraints. Throughout all the similar works, researchers typically focus only on the robust extraction of transmission lines and rely on the parallel hypothesis between transmission lines. Not to be affected by perspective distortion, UAVs have to fly above the transmission lines, which will threat the power supply security for the potential crash accident. In addition, navigation performance certainly will be degraded by the unreliable extraction of transmission lines.

Along the electricity transmission corridor, transmission tower is another important visual reference for the aid in the autonomous navigation. Sampredo et al.\textsuperscript{14} employed the features of histograms of oriented gradients (HOG)\textsuperscript{15} to train two multi-layer perceptrons (MLP)\textsuperscript{16} separately for the classification between tower and background and between four different types of towers and sought the object position by a sliding-window approach. Ceron et al.\textsuperscript{17} developed a descriptor for classifying regions of interest (ROIs) that are generated around key points. All ROIs, belonging to the same transmission tower, will jointly determine the final position of bounding box. Their algorithm took less than 50 ms, which was faster than the previous ones. Nevertheless, these methods above have to be faced with a complicated process of feature design, and their accuracy and robustness cannot yet meet the requirements of autonomous navigation. Since 2012, deep-learning-based methods have achieved breakthrough constantly in the field of object detection. Some typical works, such as Faster R-CNN,\textsuperscript{18} SSD,\textsuperscript{19} and YOLO\textsuperscript{20} outperform traditional methods. These deep-learning-based methods can be conveniently trained end-to-end but highly consume the computational resources of Graphics Processing Unit (GPU) and suffer from the poor real-time performance. Currently, these excellent methods have not yet been tested in a field inspection environment, and there has not been a navigation method based on transmission tower so far.

In terms of inspection safety, laser radar or stereo camera is typically used to enable UAV to percept surroundings. Hrabar et al.\textsuperscript{21} equipped the UAV with both stereo vision system and laser scanner to detect large static objects ahead and realized automatic obstacle avoidance during inspection. Deng et al.\textsuperscript{22} used laser scanner to follow transmission lines under the circumstance of a near distance. Zhang et al.\textsuperscript{23} employed the stereo camera to locate the wires in the study of autonomous transmission-line landing. These methods based on measurement of distance suffer from observation distance, thus it is difficult to meet safety distance requirement of high-voltage inspection.

These vision-based navigation methods above deliberately avoid some adverse impacts caused by perspective distortion, which in return limits the development of intelligent inspection. Following the projection theory, the projection of one spatial direction is defined as the vanishing point (VP). Actually, VP has been widely applied to many fields such as robotic navigation and road segmentation.\textsuperscript{28–32} There are many methods for VP detection, which are typically categorized into the edge-based ones and the texture-based ones. Taking road segmentation as an example, edge-based algorithms rely on high-contrast edges, which are mainly from clear road boundaries and lane marks. By contrast, texture-based methods show superiority in unstructured off-road environments, but have relative high computational complexity. In unstructured road, the computational VP generated by voting typically reflects a dominant direction of road. Likewise, in transmission-line inspection, due to sag effect, the computational VP can only reflect the spatial transmission-line direction approximately. Thus, how to compute a VP that meets practical navigation requirements including precision, speed, and robustness will be an important issue. In addition, there has not been a precedent that navigating the UAV by means of the VP of transmission lines.

The main objective of this article is to develop a monocular-based autonomous navigation system to achieve a safe and continuous inspection along one side of transmission lines. To this end, we combine the deep-learning-based detection with the KCF\textsuperscript{33} (kernelized correlation filter) tracking to achieve a real-time and reliable localization of transmission tower. To acquire a precise and reliable flight direction, we compute and optimize the VP of transmission lines by Levenberg–Marquardt (LM) and Cauchy loss. To keep the safety distance from transmission lines, we propose a two-stage method comprising distance perception and distance measurement. At the first stage, we optimize a homography matrix by a proposed objective function to restore the parallel attribute between transmission lines, which enables the measurement of their intervals. The spatial distance variation can be indirectly
reflected by variation of intervals between transmission lines. At the second stage, we make a triangulation to measure the real distance from transmission tower. UAV can be adjusted to the desired safety distance before the next inspection between adjacent towers. The whole scheme was evaluated in a practical inspection environment and achieved satisfactory results. The contributions of this article are as follows:

1. This article proposes a navigation approach based on transmission tower and transmission-line VP for inspection along one side of transmission lines. The approach depends only on a monocular pan-tilt-zoom (PTZ) camera.
2. This article proposes an approach for computing transmission-line VP, which provides the UAV with a reliable heading during autonomous flight.
3. This article proposes a two-stage approach for maintenance of the safety distance during inspection.

The remainder of this article is organized as follows. The second section presents the experimental platform. The third section details the navigation approach and addresses three related issues. The experiments and analysis are given in the fourth section. In the fifth section, the conclusions are summarized.

**Experiment platform**

The flight platform adopted is a refitted DJI Matrice 100 (M100) quadrotor (SZ DJI Technology Co., Ltd.), as shown in Figure 1. The flight control system comprises global positioning system (GPS), inertial measurement unit (IMU), barometer, and downward-looking stereo camera (DJI Guidance). To achieve visual navigation along transmission lines, a PTZ camera (DJI Zenmuse X3) is mounted below the M100. In addition, an onboard computational platform is built, which consists of two advanced embedded processors DJI Manifold, NVIDIA TX2, and a router used to connect these two processors. TX2 is mainly responsible for algorithms related to visual navigation. Manifold communicates with flight controller and PTZ camera, serving as a bridge to connect M100 and TX2. Finally, a robot operating system (ROS) network is built to facilitate information exchange and to share computational resource.

**Proposed navigation approach**

To have a better understanding of proposed navigation approach, we first introduce a navigation model. Then, we discuss three related issues, such as transmission tower localization, computation of transmission-line VP, and maintenance of safety distance. Finally, we details the whole navigation process.

**Navigation model.** The navigation model, as shown in Figure 2, describes some important projection relations between the first perspective and the third perspective during navigation along one side of transmission lines. The first perspective is the PTZ camera perspective, corresponding to the upper left image plane. The third perspective is the god perspective, corresponding to the whole image. To keep a simultaneous observation of transmission tower and lines, the PTZ camera needs to point to transmission tower all the time. At this time, transmission tower is located in the center of image plane, and most of transmission lines are concentrated in the upper right region that is above the horizon line. To plan a spatial flight direction, we first define three coordinate frames: inertia reference frame $C_g$, body coordinate frame $C_b$, and camera coordinate frame $C_c$. To simplify the problem, we assume that two coordinate origins $o_b$ and $o_c$ coincide. Next, let $^gR_b$ denote the attitude matrix of the camera frame with respect to the inertial frame, $^gR_b$ denote the attitude matrix of the body frame with respect to the inertial frame. The relative attitude between the body frame and the camera frame is computed as follows

$$^cR_b = (^gR_b)^{-1}gR_b \tag{1}$$

With respect to the camera frame $C_c$, we define two unit 3-D vectors $\beta_1$, $\beta_2$ separately corresponding to the computational VP and current flight direction, which are deduced from the following equations

$$\beta_1 = K^{-1}\hat{v}/\|K^{-1}\hat{v}\|_2 \tag{2}$$

$$\beta_2 = ^cR_bx_b \tag{3}$$

where $\hat{v}$ is the homogeneous coordinate of VP, $x_b$ is a unit 3-D vector parallel to the $x$-axis in the body frame, $K$ is the intrinsic matrix, and $\| \cdot \|_2$ denotes the L2-norm. To quantize the relationship between $\beta_1$ and $\beta_2$, rotation axis $l = (l_x,l_y,l_z)$ and rotation angle $\gamma$ are introduced, which follow the right-hand rule. They are computed directly by $l = \beta_1 \times \beta_2$ and $\gamma = \arcsin\|l\|$. The rotation matrix $^\gamma R_{\beta_2}$ between two vectors is deduced from the Rodriguez formula.
\[ b_1 R_{\beta_2} = \cos \gamma E + (1 - \cos \gamma) I^T + \sin \gamma I^A \]  

where \( E \) is a \((3 \times 3)\) identity matrix and \( I^A \) is a \((3 \times 3)\) skew-symmetric matrix generated by \( l \). The horizontal component \( \phi \) between the current flight direction and the transmission-line direction can be obtained by decomposing \( b_1 R_{\beta_2} \).

Ideally, when transmission lines are parallel to the ground, by adjusting \( \phi \) around axis \( z_b \), UAV flight direction \( b_2 \) can be consistent with transmission-line direction \( b_1 \). Considering sag effect of transmission lines, after rotating \( /C_30 \), UAV can only reach an approximate direction parallel to transmission lines. Typically, since the distance between adjacent towers is relatively far, the sag phenomenon looks not serious. In other words, the computational VP and \( \phi \) value can meet the requirement of practical navigation. The model above is constructed in 3-D space, in which UAV adjusts flight direction according to the computational \( \phi \). An alternative solution is to adjust UAV by horizontal pixel difference between projection \( p_A \) of flight direction and the computational VP of transmission lines. As shown in Figure 2, \( p_A \) is a virtual projection that can be obtained by the following homogeneous equation

\[ \tilde{p}_A \simeq K^* R_{b} x_b \]  

where the notation \( \simeq \) denotes the homogeneous equivalence and \( \tilde{p}_A \) is the homogeneous coordinate corresponding to \( p_A \). This alternative solution has more advantages in practical application. The effectiveness of the VP can be judged according to the detected transmission tower
bounding box. If VP is located within the detected bounding box, the result is not effective. When the VP is not effective, UAV should adjust direction to make its projection $p_A$ located out of bounding box to prevent UAV from flying toward transmission tower and eventually colliding with it. Judgement of VP effectiveness is performed only once, just at the beginning of inspection between adjacent transmission towers, which will be further introduced in section “Navigation process.” In addition, during inspection between adjacent towers, it is quite convenient to assess flight direction safety based on whether its projection $p_A$ is located within the detected bounding box.

**Transmission tower localization strategy.** As mentioned above, the transmission tower detection is crucial for the judgement of VP validity and flight direction safety. In addition, transmission towers are used to achieve continuity of inspection along power corridor, which will be detailed in section “Navigation process.” Therefore, it is necessary to seek a method of locating transmission tower, which meets the requirements of real-time and robust performance of navigation. Nowadays, in the field of object detection, deep-learning-based methods have the best precision and robustness. We trained and tested several state-of-the-art detection frameworks and finally adopted Faster R-CNN for transmission tower detection due to its lowest false detection rate, which will be further discussed in the section of “Transmission tower detection experiment.” Considering the limited computational capability of onboard embedded processor, we discarded the deep VGGNet$^{34}$ and ResNet$^{35}$ but selected a shallow ZFNet$^{36}$ as the core of network.

To improve real-time performance of Faster R-CNN, we combine it with real-time tracking algorithm KCF. KCF generates training samples by circularly shifting the matrix of images and avoids time-consuming matrix inverse operation via frequency domain calculation. These properties make KCF have a high tracking accuracy and low computational cost.

In transmission tower localization strategy, detection is used to initialize tracking and to judge the validity of tracking. Considering the runtime difference between two algorithms, we design a queue to store tracking results temporarily, as shown in Figure 3. Then, we wait for the late detection result to judge the effectiveness of tracking. Specifically, when a new image arrives, it is assigned a unique time stamp followed by a simultaneous execution of tracking and detection. Upon completion of the tracking, its tracking result and corresponding time stamp are together pushed into the queue. After the queue is updated, a new round of image acquisition and tracking begins. When the detection completes, its corresponding tracking result can be found in the queue according to the unique time stamp. Finally, if the central distance between two bounding boxes is beyond a predefined threshold, we think the tracking has failed and reinitialize it with the latest detection result.

**Transmission-line VP detection.** As mentioned in section “Navigation model,” the VP of transmission lines can provide UAV navigation with the information of direction. Benefiting from the proposed navigation approach, a large number of transmission lines are located in the region where the sky is background, which greatly reduces the risk of false detection. The proposed transmission-line VP detection algorithm consists of three parts: extraction of transmission lines, VP initialization, and VP optimization.

At the stage of line extraction, we perform first the Line Segment Detector (LSD)$^{37}$ to extract line segment regions.
and their binary contours, which helps to enhance edges of transmission lines while reducing the noise interference. Then, Progressive Probabilistic Hough Transformation (PPHT)\(^\text{38}\) algorithm is applied to detect straight line segments in the binary edge image. Figure 4 shows the practical processing effects.

At the stage of VP initialization, we adopt linear least-squares to compute the initial VP \(\bar{v} = (\bar{v}_x, \bar{v}_y)\). The linear set of equations are defined as follows

\[
\begin{bmatrix}
  a_1 & b_1 & c_1 \\
  \vdots & \vdots & \vdots \\
  a_n & b_n & c_n
\end{bmatrix}
\begin{bmatrix}
  v_x \\
  v_y
\end{bmatrix} =
\begin{bmatrix}
  -c_1 \\
  \vdots \\
  -c_n
\end{bmatrix}
\]  

where \((a_i, b_i, c_i)\) is a parameter vector of the \(i\)th straight line segment, \((v_x, v_y)\) denotes the intersection of these straight lines, and \(n\) represents the number of straight lines. Define the notation \(\Delta\) as the coefficient matrix of linear equations. Typically, the set of equations \(\Delta v = -c\) is over-determined, whose least-squares solution is equivalent to minimizing the algebraic distance \(\|\Delta v + c\|\). Thus, \(\bar{v} = -\Delta^+ c\), where \(\Delta^+\) is the pseudo-inverse of \(\Delta\). To obtain a robust initial VP, RANdom SAmple Consensus (RANSAC)\(^\text{39}\) strategy is added. For each sampling, randomly select two straight line segments to compute the VP, then divide the remaining line segments into inliers and outliers according to a predefined distance threshold. Finally, according to the equation (6), leverage the largest number of inliers, namely maximum consensus set, to reestimate the VP.

At the stage of VP fine-tuning, we adopt the optimization approach based on the geometric orientation consistency. As shown in Figure 5, \(L_k\) denotes the \(k\)th detected straight line segment. \(e_{k1}\) and \(e_{k2}\) are two endpoints of \(L_k\), and \(e_{km}\) is the midpoint. \(\hat{L}_k\) denotes the ideal straight line corresponding to \(L_k\), which passes through the VP \(\bar{v}\) and the midpoint \(e_{km}\). The measurement error \(e_k\) is defined as the geometric distance from the endpoint \(e_{k1}\) to straight line \(\hat{L}_k\), whose form is as follows

\[
e_k = \text{dis}(e_{k1}, \hat{L}_k) = \text{dis}(\hat{e}_{k1}, \hat{e}_{km} \times \hat{v})
\]  

where \(\text{dis}(\cdot)\) is an operation that computes the distance from point to straight line. \(\times\) denotes the cross-product operation. \(\hat{e}_{k1}\), \(\bar{v}\), and \(\hat{e}_{km}\) represent the homogenous forms of \(e_{k1}\), \(\bar{v}\), and \(e_{km}\), respectively. The objective function minimizes the sum of all square error items, whose specific form is as follows

\[
\bar{v}^* = \arg \min_{\bar{v}} \sum_k \rho(e_k^2)
\]  

\(\rho(\cdot)\) is the Cauchy loss function, whose definition is as follows

\[
\rho(x^2) = \log(1 + x^2)
\]

Cauchy improves the robustness of optimization by limiting the abnormal amplitude of gradient. The optimization with Cauchy loss makes VP gradually move to the position that meets the overall orientation consistency, which improves the accuracy of estimate for transmission-line direction. In addition, RANSAC may fail when outliers are dominant. To solve the problem, we use the VP of previous frame to filter wrong line segments in current frame in practice and achieve quite stable detection results.

**Distance perception from UAV to transmission lines**

**Image rectification by homography optimization.** To eliminate the parallel distortion of transmission lines in image, as illustrated in Figure 6(a), we rotate the current camera...
Algorithm 1. Adaptive line segment clustering algorithm.

Input: a sorted collection \( \{y_i\} \) of y-axis coordinates of line segments \( \{d_i\} \) in rectified image, a line-interval set \( \{d_{i+1}\} \) based on \( \{y_i\} \), balance factor \( \lambda = 0.5 \).

Output: clustering threshold \( \tau \), average line interval \( d_{m} \) inside a bundle, average line interval \( d_{out} \) between bundles, line bundles \( \{S_k\} \), number \( M \) of \( \{S_k\} \), a collection \( Q \) of y-axis coordinates of line-bundle centers.

1: initialize \( d_{in}^{pre} = d_{in}^{cur} = d_{m} \) and \( d_{out}^{pre} = d_{out}^{cur} = d_{m} \), in which \( d_{m} \) and \( d_{m} \) are randomly selected from \( \{d_{i+1}\} \). Beside \( \Phi_{in} = \Phi_{out} = \text{Null} \).
2: repeat
3: \( d_{in}^{pre} = d_{in}^{cur}, d_{out}^{pre} = d_{out}^{cur} \).
4: for \( i = 1 \) : size(\( \{d_{i+1}\} \)) do
5: if \( |d_{i+1} - d_{in}^{cur}| < |d_{i+1} - d_{out}^{cur}| \) then
6: add \( d_{i+1} \) to \( \Phi_{in} \).
7: else
8: add \( d_{i+1} \) to \( \Phi_{out} \).
9: end if
10: end for
11: update: \( d_{in}^{cur} = \text{average}(\sum \Phi_{in}(i)), d_{out}^{cur} = \text{average}(\sum \Phi_{out}(i)) \).
12: until \( |d_{in}^{pre} - d_{in}^{cur}| \geq 1 \) or \( |d_{out}^{pre} - d_{out}^{cur}| \geq 1 \).
13: end for
14: create a new line bundle \( S_k \).
15: add \( y_i \) to \( S_k \).
16: end if
17: end for
18: \( M = k \).
19: for \( i = 1 \) : \( M \) do
20: \( m_i = \text{average}(\sum y_j \in S_i) \).
21: add \( m_i \) to \( Q \).
22: end for

where \( \mathbf{c} \) is the homogeneous coordinate corresponding to \( \mathbf{p} \) and \( cR_c \) defined by equation (12) represents the rotation matrix from \( C_c \) to \( C_c \) with the y-axis as a rotation axis and \( 90^\circ - \theta \) as a rotation angle. The \( \theta \) is shown in Figure 6(a) and is the same as the one defined in Figure 2, since the x-axis in the body frame is expected to be consistent with the VP direction during navigation.

\[
\begin{align*}
\mathbf{cR_c} &= \begin{pmatrix}
\cos(90^\circ - \theta) & \sin(90^\circ - \theta) \\
-\sin(90^\circ - \theta) & \cos(90^\circ - \theta)
\end{pmatrix} \\
&= (\mathbf{cR_c})^T.
\end{align*}
\] (12)

Equation (11) is abbreviated to \( \mathbf{p'} \simeq H\mathbf{p} \), where \( H \) is the homography matrix with one rotation degree of freedom.

Based on the detected transmission line segments, we propose an objective function to optimize the \( H \) for the best recovery of parallel property.

\[
\begin{align*}
\theta^* &= \arg \min_{\theta} \sum_k \left\| 1 - \left( \mathbf{n}_k(\theta), (1, 0)^T \right) \right\|^2 \\
&\quad \times \left( (\mathbf{q}_{k1}, \mathbf{q}_{k2}, \mathbf{q}_{k3}) \simeq H(\theta) \cdot \mathbf{c}_k \times H(\theta) \cdot \mathbf{c}_{k2} \right) \quad (13)
\end{align*}
\]

In (13), \( \langle \cdot \rangle \) denotes an inner product operation, and \((1, 0)^T\) represents a horizontal unit vector that is parallel to the x-axis of image plane. In equation (14), \( \mathbf{c}_k \) and \( \mathbf{c}_{k2} \) are defined in section “Transmission-line VP detection” and \( \mathbf{q}_{k1}, \mathbf{q}_{k2}, \mathbf{q}_{k3} \) are parameters of the rectified line segment \( L_k \). Thereafter, the univariate objective function can be solved by LM iterations.

Distance perception algorithm. Affected by detection noise, the rectified straight line segments cannot reflect accurately the number and the positions of transmission lines. Thus, to solve the problem, we first group these line segments followed by measuring the intervals between groups instead of measuring the intervals between line segments directly, which achieves more reliable estimates. To make these line segments be grouped automatically, an adaptive K-means-based grouping algorithm is proposed, whose details are described in Algorithm 1. Benefitting from the previous homography rectification, the positions of line segments can be simplified into 1-D coordinates along y-axis of image. As illustrated in Figure 6(b), after grouping line segments, we extract the central positions of these line bundles and convert the problem of estimating spatial distance variation into a 1-D point-set-registration.

Further, we define two notations \( s \) and \( t \) to separately represent the scale and translation of the similarity transformation. The \( s \) and \( t \) can be computed by minimizing the point set registration error, as shown in the following equation.

\[
\begin{align*}
(s^*, t^*) &= \arg \min_{s, t} \sum_i \left\| n_i - (sm_i + t) \right\|^2 + \lambda \|t\|^2 \\
&= (s^*, t^*) \\
&\quad \max \left( |s^*| \right) \end{align*}
\] (15)
where $m_i$ and $n_i$ are a pair of normalized and matched bundle-center coordinates, and $\lambda$ is the regularization parameter, which can reduce the risk of mismatch by limiting the magnitude of $t$.

**Distance measurement from UAV to transmission tower.** UAV not only should be able to perceive the distance variation relative to transmission lines but also measure the real distance. According to the section “Transmission tower localization strategy,” when UAV is close to a transmission tower, the PTZ will rotate fast so as to track the tower. Benefiting from the fact that the camera pose varies significantly in a short term, the triangulation based on multiple views can be adopted to measure the real distance from UAV to the transmission tower at the end of inspection. As illustrated in Figure 7, the whole process begins at position $D$ with $\theta = 25^\circ$ and ends at position $F$ with $\theta = 90^\circ$. Taking $5^\circ$ as an interval, the UAV records its positions and orientations at the same height. Based on the recorded data, 14 dotted lines passing from the optical center of PTZ camera to the center of the detected bounding box can be plotted in a 2-D plane. Finally, at position $F$, the intersection of these dotted lines solved by linear least-squares can represent the tower center. According to the real distance, UAV is able to adjust its position to an expected one before the next inspection.

**Navigation process.** A long distance inspection task can be decomposed into several short subtasks between adjacent transmission towers. As shown in Figure 8, the subtask consists of initialization stage and inspection stage. At the stage of initialization, the UAV first detects transmission tower and initializes tracking by the detected bounding box. Next, the PTZ constantly adjusts pitch and yaw angle until the tracking bounding box is located in the center of image. Then, the UAV computes the VP of transmission lines and judges its validity according to the principle introduced in “Navigation model” section. If the VP is effective, the UAV will adjust heading to align its heading projection with VP. Otherwise, the UAV heading will be aligned with the empirical direction. At the stage of inspection, the UAV flies along the established direction while performing distance perception algorithm to ensure safety distance. At the end of inspection, it performs the triangulation algorithm.

**Experiment and analysis**

**Transmission tower detection experiment**

**Experiment setup.** For this experiment, 1300 sheets of transmission tower pictures with different resolutions and backgrounds were collected from aerial videos and annotated manually. The comparison were made among three state-of-the-art deep-learning-based detection frameworks: Faster R-CNN, SSD, and YOLOv2. We adopted 10-fold cross-validation to find the best models. Following this scheme, the data set is randomly partitioned into 10 subsets with equal size, then the training and validation are conducted for 10 times. Each time, a different subset is taken out for validation while the remaining union of ninefolds
are used for training. We used the Caffe framework\textsuperscript{41} to implement the training process on a GTX TitanX GPU and the validation process on TX2.

### Quantitative evaluation methodology

For quantitative evaluation of the detection task, we followed the evaluation standard of the PASCAL Visual Object Classes challenge.\textsuperscript{42} A detection result is considered correct when the bounding box overlap ratio $r$ between the ground truth $B_{gt}$ and the predicted bounding box $B_p$ exceeds 50%. The notation $r$ is defined as follows:

$$r = \frac{\text{area}(B_{gt} \cap B_p)}{\text{area}(B_{gt} \cup B_p)} \quad (16)$$

where $\text{area}(B_{gt} \cup B_p)$ represents the union of the ground truth bounding box and the predicted bounding box and $\text{area}(B_{gt} \cap B_p)$ denotes their intersection. According to $r$, detections can be divided into three types: true positive (TP, the tower is correctly detected), false positive (FP, the background is mistaken as the tower), and false negative (FN, the tower is not detected). The three different cases are illustrated in Figure 9. Based on the notations above, the precision and recall are defined as follows:

$$\text{precision} = \frac{TP}{TP + FP} \quad (17)$$
$$\text{recall} = \frac{TP}{TP + FN} \quad (18)$$

The average precision (AP) is also adopted to evaluate the comprehensive performance of detection, whose value is approximately equal to the area under the precision–recall curve.

### Experimental results

The comparison was made from the following three aspects: runtime, AP, and the false detection rate (precision–recall curve). As shown in Table 1, SSD300 has the fastest runtime and a relatively high AP, but its low input resolution (300 × 300) may cause frequent FPs and FNs. YOLOv2 has a satisfactory speed of 5.6 frames per second (FPS), but its AP is relatively low. As illustrated in Figure 10, Faster R-CNN (VGG16) and (ZF), denoted by the black and red solid lines, respectively, maintain a 100% precision over a fairly wide range of recall, which clearly outperform SSD and YOLOv2. At this point, both SSD and YOLOv2 encounter different degrees of false detection, even at a low level of recall. Since the false detection (FP) can bring a significant threat to the navigation safety, we finally choose Faster R-CNN (ZF) that achieves a lower false detection rate in the tower data set and is faster than VGG16 to enable the reliable and fast localization of transmission tower for inspection task.

Figure 11, extracted from a recorded inspection video, shows the fusion process of Faster R-CNN (ZF) and KCF. As shown in Figure 11(a), when the FN happened, the tracking result that corresponded to the blue bounding box was not affected, which achieved a continuous and smooth localization of transmission tower. With the yaw angle of PTZ changing, as shown in Figure 11(b) to (f), the sun moved gradually to the left side of the tower, affected by which the blue tracking bounding box also slowly drifted away from the target. When the pixel disparity between the two bounding box centers was beyond the prespecified threshold, UAV began to slow down, and finally re-initialized the tracker after the stable hover, whose process is illustrated in Figure 11(g) and (h). By virtue of complementary advantages of detection and tracking, locating the transmission tower becomes robust and fast.

---

Figure 8. Flowchart of the proposed approach. Initialization stage and inspection stage are distinguished by green and purple dotted bounding boxes. The pictures on the left show the real processing results corresponding to the annotations on the right. The red bounding box represents the result of detection. The blue bounding box denotes the result of tracking. At the stage of heading adjustment, the red circle denotes the computational VP, the yellow circle denotes the projection of the expected flight direction, the green lines are the optimized lines passing through VP. $\theta$ represents the yaw angle between the expected flight direction and the PTZ camera optical axis. VP: vanishing point; PTZ: pan-tile-zoom.
VP detection experiment

Experiment objectives. In this experiment, we will verify the feasibility of navigating UAV by VP of transmission lines followed by demonstrating that Cauchy loss and motion-continuity constraint play an important role in improving the precision and robustness of VP detection.

Experimental effect analysis. To verify the feasibility of navigation approach based on VP of transmission lines, we conducted four flight experiments between adjacent transmission towers, which considered the influence of light, clouds, and variation of both view angle and distance on the detection of VP. We separately extract eight detection results of different view angles from each experiment and list them in one subplot in the Figure 12. It can be seen that in the image plane, the transmission lines close to the camera appear approximately straight in spite of their sag problem. During the process of navigation between adjacent towers, the PTZ camera continuously rotates right, accordingly, the estimated VP gradually moves left. The estimated VPs can well reflect the direction of transmission lines.

Further, we select three typical detection results to make an effect display about the optimization, which are as shown in Figure 13. It can be seen that line segment detection results usually contain a small number of false edges that do not match the VP consistency, which may be caused by the propeller, the sun, and the clouds. However, due to the addition of Cauchy loss function, the optimization is not

Table 1. Speed and average precision.

| Method           | AP1 (%) | AP2 (%) | AP3 (%) | AP4 (%) | AP5 (%) | AP6 (%) | AP7 (%) | AP8 (%) | AP9 (%) | AP10 (%) | FPS  |
|------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|-----------|------|
| Faster R-CNN (ZF)| 89.5    | 89.8    | 89.9    | 89.4    | 89.3    | 89.4    | 89.5    | 89.5    | 89.5    | 89.2      | 2    |
| Faster R-CNN (VGG16)| 89.6 | 89.8 | 89.9 | 90.1 | 89.3 | 89.4 | 89.5 | 89.7 | 89.7 | 89.8 | 0.8 |
| SSD300           | 88.9    | 88.6    | 88.5    | 88.4    | 88.7    | 88.7    | 88.3    | 88.2    | 88.3    | 88.5      | 7    |
| SSD512           | 89.2    | 89.1    | 88.9    | 89.1    | 89.3    | 88.8    | 88.9    | 88.8    | 89.1    | 89.1      | 2    |
| YOLOv2           | 86.8    | 86.5    | 86.9    | 86.3    | 86.2    | 86.7    | 86.7    | 85.9    | 86.7    | 86.8      | 5.6  |

AP: average precision.

*For Faster R-CNN, ZF and VGG16 represent two different networks of feature extraction. SSD300 and SSD512 separately correspond to 300 × 300 and 512 × 512 input resolutions. YOLOv2 corresponds to the 544 × 544 input resolution. The different 10 APs correspond to the results of 10-fold cross-validation, in which the bold font indicates the best of each framework. FPS, namely frames per second, is measured on NVIDIA TX2.
Figure 11. Fusion of deep-learning-based detection and smooth tracking. The red bounding box denotes the result of Faster R-CNN (ZF). The blue bounding box represents the KCF result. The red character string beginning with character # and displayed in upper right region indicates the frame number. The cyan straight line segment shown in the upper left region denotes the horizontal pixel error between the centers of detection result and tracking result. KCF: kernelized correlation filter.

Figure 12. VP detection results. The red line segment denotes the detected transmission lines. The red circle denotes the initial VP computed by the linear least-squares. The blue circle represents the optimized VP by LM and Cauchy loss. The green line segment, connecting the optimized VP and the midpoint of red line segment, denotes the optimized straight line $L_k$ introduced in section “Transmission-line VP detection.” The red character string beginning with character # indicates the frame number. LM: Levenberg–Marquardt; VP: vanishing point.

Figure 13. Effect display of VP optimization. The yellow and purple line segments represent a small number of false edges that do not match the VP consistency. The other color marks are the same as the ones in Figure 12. VP: vanishing point.
affected by these false edges and gives the estimate approximate to the truth value.

When line segment detection results contain a large number of false edges, the initialization based on RANSAC will fail. Since the PTZ rotates slowly, the VP position will not change a lot across adjacent frames. Thus, we first leverage this prior knowledge to filter the detected straight line segments to ensure a right initialization. At this point, we make a comparison between detection results with and without motion-continuity constraint, as shown in Figure 14. In the Figure 14(a), a large number of cloud edges make the initial position of VP appear a serious offset, which results in the optimization eventually failing, while applying motion continuity constraint can help remove these edges, thus ensuring the correctness of the optimization. The correct optimization corresponds to the Figure 14(b).

Quantitative evaluation. For the quantitative evaluation, we extracted 60 images with 1920 × 1080 resolution at equal intervals from each video and annotated them. The ground truth is the mean of five manual annotations. We performed statistics on pixel error between the ground truth and the estimate and then used accumulative-pixel-error to assess the detection accuracy.

Figure 15(a) shows the four accumulative-pixel-error curves between the optimized VP and the ground truth, which correspond to the four experiments shown in Figure 12, respectively. In experiment (a), since the background is relatively clean, the algorithm achieves the best detection performance, whose final cumulative error is less than 200 pixels. In experiment (b), the first three-quarters of the curve is in a fast upward trend, which is in agreement with the time of illumination interference caused by the sunrise. Thereafter, as the camera rotates to the right, the sun finally disappears from the field of view, accordingly the curve slope also tends to be mild. The entire process of experiment (d) is affected by clouds, leading to a 300 pixel error in total. Because of the harsh environment in experiment (c), VP detection has the largest uncertainty with a cumulative error close to 600 pixels and about 10 pixel error each image in average. In terms of the 1920 × 1080 resolution, the 10-pixel swing around the ground truth is acceptable in practical inspection. Figure 15(b) displays
the accumulative-pixel-error curves of initial VP, which retains the same color configuration as the ones in Figure 15(a). By contrast, the errors between the initial VP and the ground truth increase significantly.

**Distance perception and measurement experiments**

**Experiment for safety distance from transmission lines.** This experiment shows the performance of distance perception algorithm, in which the UAV is deliberately operated to fly along transmission lines and slowly close to them. The results of homography optimization, line clustering and scale estimation were recorded throughout the flight. As shown in Figure 16(a), four representative frames were extracted in order from the video. The different colors assigned to the detected transmission lines represent the different clusters that are generated by the adaptive line segment clustering algorithm. As the UAV slowly approaches the transmission lines, by virtue of applying homography optimization and line clustering to the raw images, it can be intuitively observed that the intervals of parallel bundles in rectified images are gradually increasing and their variation has been well reflected by the gradually increasing scale factor. To further demonstrate the accuracy of the solved scale $s$ and translation $t$, we used the $s$ and $t$ to recover original positions of current bundles with respect to the reference frame (the first frame). The result is shown in Figure 16(b), where the estimations during the

**Figure 16.** Performance of distance perception from transmission lines. The first row of (a) shows the snapshots of PTZ camera and the second row displays the corresponding rectification results. The red character string displayed in left upper region indicates the frame number and scale estimation. The circles in (b) represent the observed bundle center positions. By applying scales and translations to current observations, the plus signs in (b) mean the estimated bundle center positions with respect to the reference frame. The black circles and arrows indicate this estimation process. PTZ: pan-tilt-zoom.
whole flight, denoted by plus signs, are in good agreement with their corresponding original positions.

After the individual performance analysis of each algorithm component, a comprehensive experiment was designed to demonstrate the effectiveness of the entire navigation approach in maintaining the safety distance from transmission lines. In this experiment, a segmented controller with dead-band, taking scale $s$ as input, was used to control the safety distance. The test distance covered two adjacent intervals, approximate 450 m. The transmission towers have the typical 220-kV double-circuit lattice steel structure, which has a height of 40 m and a base-width of 6 m. The tower can be approximately enveloped by a $8 \times 8 \times 40$ m$^3$ (length, width, and height) cuboid. The GPS data of transmission tower were provided by power companies and they were based on WGS84 World Geodetic Coordinate System 1984 (WGS84). The locations of transmission lines were determined by the adjacent towers they connect. To ensure the UAV safety, we limited the wind force within three and set the flight speed to 1 m/s. As shown in Figure 17, nine experimental trajectories with different takeoff positions (a range of 5 to 15 m for the safety distance and a range of 22 to 31 m for the flight altitude) were plotted. They were recorded based on the fusion result of visual localization system (DJI Guidance) and GPS system of the UAV. The centers of the two pentagrams denote the tower centers relative to North East Down (NED) coordinate system. Figure 17(a) is a planar view of trajectories. It can be seen that all trajectories achieved a relatively stable safe distance as expected. The UAV could adjust themselves when the safety distance from the transmission lines were getting closer or farther. Furthermore, Figure 17(b) displays the altitude component of the corresponding flight trajectory. It can be seen that the proposed navigation approach can adapt to a certain degree of variation of height, which endows the transmission-line inspection with much flexibility.

**Experiment for safety distance from transmission tower.** To verify the effectiveness of distance measurement, we hope that the UAV can adjust itself to an expected safety distance by applying the triangulation algorithm at the end of inspection between adjacent towers. The experiment environment is the same as the one mentioned in section “Experiment for safety distance from transmission lines.” As shown in Figure 18, to have a clear view, we selected two typical trajectories to demonstrate the effectiveness of triangulation. The two different tests started with different initial positions, whose corresponding trajectories were separately denoted by the red and blue color. At the position G, UAV began to collect the data, which comprised the current position of the UAV and the corresponding yaw
angle of the PTZ camera. At the position H, UAV hovered on the side of the transmission tower followed by estimating the distance from the transmission tower leveraging the collected data. Since the blue trajectory was farther than the expected safety distance (12 m), UAV made an adjustment before the next flight as expected. The red trajectory corresponded to the opposite circumstance. It can be seen that the triangulation is effect to improve the quality and safety of the autonomous flight.

Conclusion

In this article, we proposed an autonomous navigation approach based on transmission lines and transmission towers to enable the UAV continuous navigation along one side of transmission lines. To achieve the navigation approach, three following issues were addressed. First, to locate the UAV, the transmission tower was viewed as the landmark, which was positioned robustly and timely by combining Faster R-CNN detector with KCF tracker. Second, to obtain a robust and precise flight direction, VP of transmission lines was calculated followed by an optimization. The motion-continuity constraint and Cauchy loss function were added to improve the robustness of initialization and optimization to adapt to the harsh inspection environments. Third, to ensure the safety of inspection, the perception and measurement of safety distance from UAV to transmission lines were proposed. The perception algorithm was designed to sense the variation of spatial distance according to the variation of intervals between transmission lines, which was used at the early stage of navigation between adjacent towers. The measurement algorithm was designed to estimate the real spatial distance from transmission lines by the proposed triangulation across multiple views, which was used at the later stage of navigation between adjacent towers. To verify these methods, a UAV flight platform carrying advanced embedded processors was developed. Finally, the designed flight platform and the whole navigation approach were tested in a real-world field environment, which achieved an encouraging result.

In the future, we will integrate an online fault detection into the navigation system and improve the cruise duration for a long distance inspection.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (61673378,61421004).

ORCID iD

Xiaolong Hui https://orcid.org/0000-0003-0737-952X
Jiang Bian https://orcid.org/0000-0002-5125-4882

References

1. Pagnano A, Höpf M, and Teti R. A roadmap for automated power line inspection, maintenance and repair. Procedia CIRP 2013; 12: 234–239.
2. Chen C, Wu G, Wang Q, et al. Design of the gripper for power lines inspection robot. In: 2014 11th world congress on intelligent control and automation (WCICA), Shenyang, China, 29 June–4 July 2014, pp. 3340–3344. Piscataway: IEEE.
3. Zhao D, Yang G, Li E, et al. Design and its visual servoing control of an inspection robot for power transmission lines. In: 2013 IEEE international conference on robotics and biomimetics (ROBio), Shenzhen, China, 12–14 December 2013, pp. 546–551. IEEE.
4. Pouliot N, Richard PL, and Montambault S. Linescout technology opens the way to robotic inspection and maintenance of high-voltage power lines. IEEE Pow Ener Technol Sys J 2015; 2(1): 1–11.
5. Tavares L and Sequeira JS. RIOL-robotic inspection over power lines. IFAC Proc Vol 2007; 40(15): 108–113.
6. Lu M, Sheng G, Liu Y, et al. Research on auto-tracking algorithm for power line inspection based on unmanned aerial vehicle. In: 2012 Asia-Pacific power and energy engineering conference (APPEEC), Shanghai, China, 27–29 March 2012, pp. 1–5. Piscataway: IEEE.
7. Luque-Vega LF, Castillo-Toledo B, Loukianov A, et al. Power line inspection via an unmanned aerial system based on the quadrotor helicopter. In: 2014 17th IEEE Mediterranean electrotechnical conference (MELECON), Beirut, Lebanon, 13–16 April 2014, pp. 393–397. Piscataway: IEEE.
8. Cerón A, Prieto F, and Mondragón BIF. Towards visual based navigation with power line detection. In: International symposium on visual computing, Las Vegas, NV, USA, 8–10 December 2014, pp. 827–836. Berlin: Springer.
9. Chen Y, Li Y, Zhang H, et al. Automatic power line extraction from high resolution remote sensing imagery based on an improved radon transform. Pattern Recogn 2016; 49: 174–186.
10. Ceron A, Prieto F, and Mondragón BIF. Power line detection using a circle based search with UAV images. In: 2014 international conference on unmanned aircraft systems (ICUAS), Orlando, FL, USA, 27–30 May 2014, pp. 632–639. IEEE.
11. Illingworth J and Kittler J. A survey of the Hough transform. Computer Vision, Graphics, and Image Processing 1988; 44(1): 87–116.
12. Zhou G, Yuan J, Yen II., et al. Robust real-time UAV based power line detection and tracking. In: 2016 IEEE international conference on image processing (ICIP), Phoenix, AZ, USA, 23–28 September 2016, pp. 744–748. IEEE.
13. Tian F, Wang Y, and Zhu L. Power line recognition and tracking method for UAVs inspection. In: 2015 IEEE
16. Riedmiller M. Advanced supervised learning in multilayer perceptrons from backpropagation to adaptive learning algorithms. *Comput Stand Inter* 1994; 16(3): 265–278.

17. Ceron A, Mondragon I, and Prieto F. Real-time transmission tower detection from video based on a feature descriptor. *IET Comput Vis* 2016; 11(1): 33–42.

18. Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. In: *Advances in Neural Information Processing Systems (NIPS)*, Montreal, Canada, 7–12 December 2015.

19. Liu W, Anguelov D, Erhan D, et al. SSD: single shot multi-box detector. In: *European conference on computer vision*, Amsterdam, Netherlands, 11–14 October 2016, pp. 21–37. Springer.

20. Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA, 27–30 June 2016, pp. 779–788.

21. Hrabar S, Merz T, and Frousheger D. Development of an autonomous helicopter for aerial powerline inspections. In: *2010 1st International conference on applied robotics for the power industry (CARPI)*, Montreal, QC, Canada, 5–7 October 2010, pp. 1–6. IEEE.

22. Deng C, Liu JY, Liu YB, et al. Real time autonomous transmission line following system for quadrotor helicopters. In: *International conference on smart grid and clean energy technologies (ICSGCE)*, Chengdu, China, 19–22 October 2016, pp. 61–64. IEEE.

23. Zhang X, Li X, Yin WR, et al. A binocular vision based auxiliary navigation system of the unmanned aerial vehicle for power line inspection. In: *2016 3rd International conference on systems and informatics (ICSAI)*, Shanghai, China, 19–21 November 2016, pp. 291–296. IEEE.

24. Wang J, Zhao H, Bi Y, et al. Quad-rotor helicopter autonomous navigation based on vanishing point algorithm. *J Appl Math* 2014; 2014.

25. Elloumi W, Guissous K, Chetouani A, et al. Indoor navigation assistance with a smartphone camera based on vanishing points. In: *International conference on indoor positioning and indoor navigation (IPIN)*, Montpellier-Belfort, France, 28–31 October 2013, pp. 1–9. IEEE.

26. Lu Y and Song D. Visual navigation using heterogeneous landmarks and unsupervised geometric constraints. *IEEE Trans Robot* 2015; 31(3): 736–749.

27. Ji Y, Yamashita A, and Asama H. RGB-D slam using vanishing point and door plate information in corridor environment. *Intell Serv Robot* 2015; 8(2): 105–114.

28. Yoo JH, Lee SW, Park SK, et al. A robust lane detection method based on vanishing point estimation using the relevance of line segments. *IEEE Trans Intell Transp* 2017; 18(12): 3254–3266.

29. Lee S, Kim J, Yoon JS, et al. Vpgnet: vanishing point guided network for lane and road marking detection and recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Venice, Italy, 22–29 October 2017, pp. 1947–1955. IEEE.

30. Kong H, Sarma SE, and Tang F. Generalizing Laplacian of Gaussian filters for vanishing-point detection. *IEEE Trans Intell Transp* 2013; 14(1): 408–418.

31. Shi J, Wang J, and Fu F. Fast and robust vanishing point detection for unstructured road following. *IEEE Trans Intell Transp* 2016; 17(4): 970–979.

32. Fan X and Shin H. Road vanishing point detection using weber adaptive local filter and salient-block-wise weighted soft voting. *IET Comput Vis* 2016; 10(6): 503–512.

33. Henriques JF, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters. *IEEE Trans Pattern Anal* 2015; 37(3): 583–596.

34. Simonyan K and Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:14091556* 2014.

35. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA, 27–30 June 2016, pp. 770–778. IEEE.

36. Zeiler MD and Fergus R. Visualizing and understanding convolutional networks. In: *European conference on computer vision*, Zurich, Switzerland, 6–12 September 2014, pp. 818–833. Springer.

37. Gjio RGV, Jakubawicz J, Morel JM, et al. LSD: a fat line segment detector with a false detection control. *IEEE Trans Pattern Anal* 2010; 32(4): 722–732.

38. Matas J, Galambos C, and Kittler J. Robust detection of lines using the progressive probabilistic Hough transform. *Comput Vis Image Und* 2000; 78(1): 119–137.

39. Fischler MA and Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 1981; 24(6): 381–395.

40. Refaelzadeh P, Tang L, and Liu H. Cross-Validation. In: Liu L and Özsu M. (eds) *Encyclopedia of Database Systems*. New York, NY: Springer, 2016.

41. Jia Y, Shelhamer E, Donahue J, et al. Caffe: convolutional architecture for fast feature embedding. In: *Proceedings of the 22nd ACM international conference on multimedia*, Orlando, Florida, USA, 3–7 November 2014, pp. 675–678. ACM.

42. Everingham M, Van Gool L, Williams CK, et al. The Pascal visual object classes (VOC) challenge. *Int J Comput Vision* 2010; 88(2): 303–338.