

Quantifying Feature Association Error in Camera-based Positioning

Chen Zhu*, Mathieu Joerger[†], Michael Meurer*

*Institute of Communications and Navigation, German Aerospace Center (DLR), Oberpfaffenhofen, Germany
Email: Chen.Zhu@dlr.de, Michael.Meurer@dlr.de

[†] Aerospace and Ocean Engineering Department, Virginia Tech, Blacksburg, VA, USA
Email: joerger@vt.edu

Abstract—Camera-based visual navigation techniques can provide six degrees-of-freedom estimates of position and orientation (or pose), and can be implemented at low cost in applications including autonomous driving, indoor positioning, and drone landing. However, feature matching errors may occur when associating measured features in camera images with mapped features in a landmark database, especially when repetitive patterns are in view. A typical example of repetitive patterns is that of regularly spaced windows on building walls. Quantifying the data association risk and its impact on navigation system integrity is essential in safety critical applications. But, literature on vision-based navigation integrity is sparse. This work aims at quantifying and bounding the integrity risk caused by incorrect associations in visual navigation using extended Kalman filters.

I. INTRODUCTION

Cameras have successfully been used in a variety of navigation applications including visual odometry for trajectory estimation of the Mars rover [1] [2] and vision-based positioning for unmanned aerial vehicle (UAV) landing [3] [4]. In safety critical applications such as autonomous driving, vision can potentially be used, not only for situational awareness, but also for landmark-based navigation [5]. Cameras can be incorporated in a navigation system as additional sensors to improve navigation performance in terms of accuracy, availability, continuity and integrity [6] [7] [8]. However, there is no widely implemented method to quantify integrity in vision-based navigation.

A crucial error source in visual positioning is feature matching error in data association. The probability of occurrence of incorrect associations can be high in urban environments, especially when multiple similar-looking objects are observed in a same image. Examples of difficult to distinguish features include series of windows on building walls, parked cars, traffic signs, etc. In this work, in order to evaluate the integrity risk in visual positioning, we propose an analytical framework to upper bound the probability of hazardous misleading information (HMI) while considering all possible incorrect feature associations. Based on prior work on laser-based navigation [9][10], we develop a method to quantify the probability of the correct association that uses the innovation vector from the extended Kalman filter (EKF).

The paper is organized as follows. First, we define the camera measurement equation and the EKF-based positioning algorithm. Second, we develop a separation metric

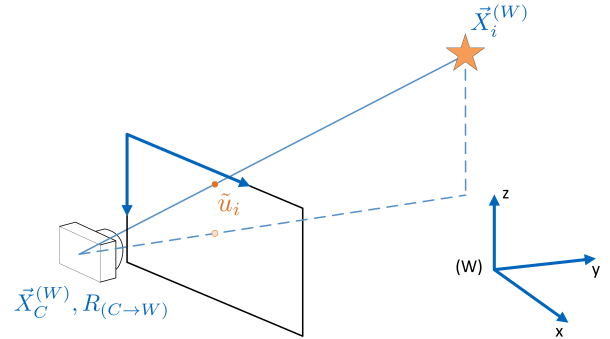


Fig. 1: System model of a pinhole camera

for extracted landmarks, and describe the data association criterion. Third, we derive an integrity risk bound as a function of the EKF innovation vector considering data association errors. Simulation and experimental results are then provided, analyzed and conclusions are drawn.

II. MEASUREMENT MODEL OF CAMERA-BASED POSITIONING

In the remainder of this paper, a superscript with parentheses $(\cdot)^{(W)}$ is used to denote the reference frame in which the vector is represented. Vectors and matrices are denoted with bold symbols while scalars are not. Vectors such as $\vec{\mathbf{X}}_j^{(W)} \in \mathbb{R}^3$ with geometric meanings are written with an arrow. The homogeneous coordinates in the extended Euclidean plane are written as $\tilde{\mathbf{u}} \in \mathbb{P}^2$. $[\mathbf{A}; \mathbf{B}]$ denotes vertical concatenation of two matrices.

In this work, a state-of-the-art pinhole model with corrected radial lens distortion is used as the sensor model of the camera. Fig. 1 illustrates the system model of a pinhole camera.

The projection of a point with coordinates $\vec{\mathbf{X}}_j^{(W)} \in \mathbb{R}^3$ to the camera is described as

$$\tilde{\mathbf{u}}_j = \mathbf{K} \mathbf{R}_{(C \rightarrow W)}^T (\vec{\mathbf{X}}_j^{(W)} - \vec{\mathbf{X}}_C^{(W)}), \quad (1)$$

where $\vec{\mathbf{X}}_C^{(W)}$ is the camera position in a global reference frame (W) , and $\mathbf{R}_{(C \rightarrow W)} \in \mathbf{SO}(3)$ is the rotation matrix between the camera body frame (C) and frame (W) . The matrix \mathbf{K} is called camera intrinsic matrix, the parameters of which are only dependent on the camera and the lens. $\tilde{\mathbf{u}}_j = d[u_x, u_y, 1]^T$ is the location of the point j on the image plane in homogeneous coordinates, where d is the depth of

the point, i.e., the distance of the point to the image plane. The depth can be expressed as

$$d = [0, 0, 1] \mathbf{R}_{(C \rightarrow W)}^T (\vec{\mathbf{X}}_j^{(W)} - \vec{\mathbf{X}}_C^{(W)}). \quad (2)$$

In Euclidean space, the 2D coordinates of a projected point in the image plane is related to the camera pose and the 3D location of the feature point as described in the following equation

$$\mathbf{u}_j = \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \frac{1}{d} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tilde{\mathbf{u}}_j. \quad (3)$$

We parameterize the camera attitude using Rodrigues parameters $\mathbf{r} = [r_1, r_2, r_3]^T = \boldsymbol{\theta} \mathbf{e}$, which can be related to the rotation matrix by Rodrigues rotation formula as [11]

$$\mathbf{R}_{(C \rightarrow W)} = \mathbf{I}_3 \cos(\boldsymbol{\theta}) + [\mathbf{e}]_{\times} \sin(\boldsymbol{\theta}) + (1 - \cos(\boldsymbol{\theta})) \mathbf{e} \mathbf{e}^T, \quad (4)$$

where \mathbf{I}_3 is a 3-dimensional identity matrix, and $[\mathbf{e}]_{\times}$ is the skew symmetric matrix using the normalized unit vector $\mathbf{e} = \mathbf{r} / \|\mathbf{r}\| \in \mathbb{R}^3$ as non-zero entries. The operator $[\cdot]_{\times}$ is defined as

$$[\mathbf{e}]_{\times} = \begin{bmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{bmatrix}. \quad (5)$$

It should be mentioned that the three Rodrigues parameters are corresponding to the coefficients of the Lie algebra of the $\mathbf{SO}(3)$ group [12]. The rotation can equivalently represented as

$$\mathbf{R}_{(C \rightarrow W)} = \exp([\mathbf{r}]_{\times}), \quad (6)$$

where $\exp(\cdot)$ is matrix exponential operator. Such minimal parameterization of attitude is convenient for operations on manifold. It is mathematically elegant, and is capable of avoiding the Gimbal lock problem of Euler angle parameterization [11].

Consequently, we can parameterize the camera pose with a 6 degrees-of-freedom (DoF) vector

$$\mathbf{x} = \begin{bmatrix} \vec{\mathbf{X}}_C^{(W)} \\ \mathbf{r} \end{bmatrix} \in \mathbb{R}^6. \quad (7)$$

In order to simplify the notation, a function $\pi_j(\cdot) : \mathbb{R}^6 \rightarrow \mathbb{R}^2$ is defined to project the 3D point $\vec{\mathbf{X}}_j^{(W)}$ to the 2D coordinates on the image plane of a camera with 6DoF pose parameterized as \mathbf{x} . The projection can be explicitly expressed as

$$\mathbf{u}_j = \pi_j(\mathbf{x}) = \frac{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{K} \mathbf{R}_{(C \rightarrow W)}^T (\vec{\mathbf{X}}_j^{(W)} - \vec{\mathbf{X}}_C^{(W)})}{\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{K} \mathbf{R}_{(C \rightarrow W)}^T (\vec{\mathbf{X}}_j^{(W)} - \vec{\mathbf{X}}_C^{(W)})}. \quad (8)$$

Consequently, for a feature point j extracted from the measurement image, the noisy measurement equation can be expressed as

$$\boldsymbol{\mu}_j = \mathbf{u}_j + \mathbf{n}_j = \pi_j(\mathbf{x}) + \mathbf{n}_j, \quad (9)$$

where \mathbf{n}_j is zero-mean Gaussian noise with covariance matrix $\boldsymbol{\Sigma}_{n_j}$, which can be conservatively modelled using the method in [13].

Given a known landmark M consisting of m_F feature points in the map (requiring $m_F \geq 4$ in general), the position and orientation of the camera can be estimated using corresponding 2D measurements $\{j = 1, \dots, m_F | \boldsymbol{\mu}_j\}$ by iteratively solving the following nonlinear optimization problem

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\boldsymbol{\mu} - \boldsymbol{\pi}(\mathbf{x})\|_{\boldsymbol{\Sigma}_n}^2, \quad (10)$$

where

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_{m_F} \end{bmatrix}, \boldsymbol{\pi}(\mathbf{x}) = \begin{bmatrix} \boldsymbol{\pi}_1(\mathbf{x}) \\ \vdots \\ \boldsymbol{\pi}_{m_F}(\mathbf{x}) \end{bmatrix}, \mathbf{n} = \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_{m_F} \end{bmatrix}. \quad (11)$$

$\boldsymbol{\Sigma}_n$ is the covariance matrix of noise vector \mathbf{n} . The initial estimation for the linearization point can be obtained using the direct PnP methods such as in [14].

III. SYSTEM MODEL

Let us assume that there are N_L repetitive landmarks with known poses in the map, and each landmark consists of m_F feature points with known geometric arrangements. Such situation is rather common in real-life scenarios, e.g., identical windows on the street. Fig. 2 demonstrates an instance of typical scenarios that may cause data association error in visual navigation. In this synthetic example, there are three identical chessboards landmarks lying next to each other, i.e., $N_L = 3$. A camera moves in a trajectory consisting

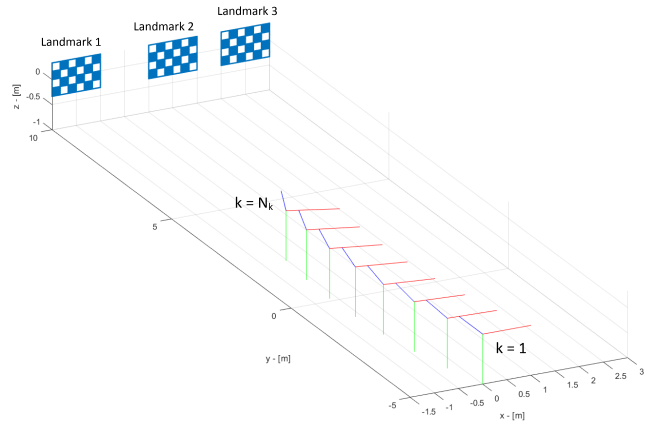


Fig. 2: A moving camera and visible identical landmarks in the map database.

of N_k keyframes as sampling epochs. The poses of the camera in the global reference frame (W) are denoted as $\{k = 1, \dots, N_k | \mathbf{x}_k\}$, and are measured at keyframes. The motion of the camera can be modelled as

$$\mathbf{x}_k = \boldsymbol{\Phi}_{k-1} \mathbf{x}_{k-1} + \mathbf{w}_{k-1}, \quad (12)$$

where $\boldsymbol{\Phi}_{k-1}$ is the linear transition matrix at time $k-1$, and \mathbf{w}_{k-1} is the process noise which follows a Gaussian distribution $\mathbf{w}_{k-1} \sim \mathcal{N}(0, \mathbf{W}_{k-1})$. We assume that the visible landmarks can be detected from the measurement images. The detection can be based on models for designed pattern, or based on advanced machine learning methods for specific class of landmarks [15]. Fig. 3 demonstrates an example

of a detected checkerboard landmark and the corresponding feature points extracted from the measurement image. Given a conservative feature error model such as CCGEM proposed in [13], the correctness of the landmark extraction can be tested by exploiting the feature arrangement of the known landmark. The error in landmark detection is not the main topic of this paper. Instead, in this work we focus on quantifying the feature association error for identical landmarks that have been successfully detected using the aforementioned methods. For landmark i consisting of m_F



Fig. 3: An example of correctly detected feature points on a landmark pad.

feature points with 3D location $\{j = 1, \dots, m_F \{\bar{\mathbf{x}}_j^{(W)}\}\}$, the corresponding projected 2D features in the image of camera at pose \mathbf{x} are written as

$$\mathbf{z}_i = \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{m_F} \end{bmatrix} = \begin{bmatrix} \pi_1(\mathbf{x}) \\ \vdots \\ \pi_{m_F}(\mathbf{x}) \end{bmatrix}. \quad (13)$$

Exploiting the feature points of the visible landmarks, the camera can estimate its 6DoF pose using an extended Kalman filter (EKF). The least squares optimization introduced in the last section can be applied to initialize the EKF. In this recursive process, given a posterior pose estimate $\hat{\mathbf{x}}_{k-1}$ with covariance $\hat{\mathbf{P}}_{k-1}$ at time instant $k-1$, the pose at time k as well as its covariance can be predicted using the motion model in Eqn. (12) as

$$\bar{\mathbf{x}}_k = \Phi_{k-1} \hat{\mathbf{x}}_{k-1}, \quad \bar{\mathbf{P}}_k = \Phi_{k-1} \hat{\mathbf{P}}_{k-1} \Phi_{k-1}^T + \mathbf{W}_{k-1}. \quad (14)$$

The measurements at time k including all the visible landmarks are written as

$$\hat{\mathbf{z}}_k = \begin{bmatrix} \hat{\mathbf{z}}_1 \\ \vdots \\ \hat{\mathbf{z}}_{N_L} \end{bmatrix} = \begin{bmatrix} h_1(\mathbf{x}_k) \\ \vdots \\ h_{N_L}(\mathbf{x}_k) \end{bmatrix} + \mathbf{v}_k = h(\mathbf{x}_k) + \mathbf{v}_k, \quad (15)$$

where $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{V}_k)$ is the measurement noise, and $h_i(\mathbf{x}_k)$ is the measurement equation of landmark i . The measurement function can be linearized at the predicted value $\bar{\mathbf{x}}_k$ as

$$h_i(\bar{\mathbf{x}}_k) = \begin{bmatrix} \pi_{i,1}(\bar{\mathbf{x}}_k) \\ \vdots \\ \pi_{i,m_F}(\bar{\mathbf{x}}_k) \end{bmatrix} \approx h_i(\mathbf{x}_k) + \mathbf{H}_{i,k}(\bar{\mathbf{x}}_k - \mathbf{x}_k), \quad (16)$$

where

$$\mathbf{H}_{i,k} = \left. \frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\bar{\mathbf{x}}_k}. \quad (17)$$

The innovation vector can be defined as the difference between the measurements and the prediction as

$$\boldsymbol{\gamma}_k = \hat{\mathbf{z}}_k - h(\bar{\mathbf{x}}_k). \quad (18)$$

The posterior estimate at time k can be updated using the following equations

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + \mathbf{G}_k \boldsymbol{\gamma}_k, \quad \hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{G}_k \mathbf{H}_k) \bar{\mathbf{P}}_k, \quad (19)$$

in which \mathbf{H}_k is the linearized matrix concatenated over all the landmarks $i = 1, \dots, N_L$, and \mathbf{G}_k is the Kalman gain calculated using the equation

$$\mathbf{G}_k = \bar{\mathbf{P}}_k \mathbf{H}_k^T (\mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T + \mathbf{V}_k)^{-1}. \quad (20)$$

IV. IDENTICAL LANDMARK ASSOCIATION CRITERION

Since the landmarks are identical, the association between different detected 2D features and the 3D point locations in the map is only dependent on the geometry between the landmark location and the camera pose. For N_L identical landmarks, there are $N_L!$ number of permutations of the features, one of which corresponds to the correct association. As a result, $N_L!$ permutation matrices $\{i = 0, 1, \dots, N_L! - 1 | \mathbf{A}_i\}$ can be created to mathematically express all the landmark association options. For example, in the scenario of Fig. 2, there are three landmarks, i.e., $N_L = 3$. The 2D features can be matched to the three landmarks in the map with 6 different orders. As an illustrative example, under the hypothesis of the association order $[\hat{\mathbf{z}}_2; \hat{\mathbf{z}}_3; \hat{\mathbf{z}}_1]$, the corresponding permutation matrix is constructed as

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{2m_F} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{2m_F} \\ \mathbf{I}_{2m_F} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (21)$$

As a result, the measurement equation under such hypothesis of association is

$$\hat{\mathbf{z}}_{k,A} = \mathbf{A} h(\mathbf{x}_k) + \mathbf{v}_{k,A}. \quad (22)$$

For landmark association hypothesis represented by \mathbf{A}_i , the innovation vector of the EKF at time k is [9]

$$\begin{aligned} \boldsymbol{\gamma}_{k,i} &= \hat{\mathbf{z}}_k - \mathbf{A}_i h(\bar{\mathbf{x}}_k) \\ &\approx h(\mathbf{x}_k) + \mathbf{v}_k - \mathbf{A}_i (h(\mathbf{x}_k) + \mathbf{H}_k(\bar{\mathbf{x}}_k - \mathbf{x}_k)) \\ &= (\mathbf{I} - \mathbf{A}_i) h(\mathbf{x}_k) - \mathbf{A}_i \mathbf{H}_k (\bar{\mathbf{x}}_k - \mathbf{x}_k) + \mathbf{v}_k \\ &= (\mathbf{I} - \mathbf{A}_i) h(\mathbf{x}_k) - \mathbf{A}_i \mathbf{H}_k \bar{\boldsymbol{\epsilon}}_k + \mathbf{v}_k. \end{aligned} \quad (23)$$

It can be seen from Eqn. (23) that both $\bar{\boldsymbol{\epsilon}}_k \in \mathbb{R}^6$ and $\mathbf{v}_k \in \mathbb{R}^{2m_F N_L}$ are zero-mean Gaussian variables. The innovation vector $\boldsymbol{\gamma}_{k,i}$ is unbiased if and only if $i = 0$ for $\mathbf{A}_0 = \mathbf{I}$. For other $i \neq 0$, there exists a bias term

$$\mathbf{b}_i = (\mathbf{I} - \mathbf{A}_i) h(\mathbf{x}_k). \quad (24)$$

Therefore, we can exploit the innovation vectors to find the most likely landmark matching from all possible association hypotheses.

For the correct association, $\boldsymbol{\gamma}_{k,0} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_0)$ is a zero-mean Gaussian random variable with

$$\boldsymbol{\Sigma}_0 = \mathbf{V}_k + \mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T. \quad (25)$$

However, for other hypotheses, the innovative vector $\boldsymbol{\gamma}_{k,i} \sim \mathcal{N}(\mathbf{b}_i, \boldsymbol{\Sigma}_i)$ with

$$\boldsymbol{\Sigma}_i = \mathbf{V}_k + \mathbf{A}_i \mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T \mathbf{A}_i^T. \quad (26)$$

In EKF based sequential estimation problems, we do not explicitly know which association is correct, so we calculate the squared L2-norm of the innovation vectors weighted by their covariance matrices for all hypotheses. The results are realizations of $N_L!$ different Chi-squared distributions with the same degrees of freedom, among which only one is central Chi-squared distribution. In our implementation, the variables are expressed as

$$\Gamma_i = \|\boldsymbol{\gamma}_{k,i}\|_{\boldsymbol{\Sigma}_i^{-1}}^2 \sim \chi^2(2m_F N_L + 6). \quad (27)$$

Therefore, we choose the permutation with minimum weighted norm of the innovation as the feature association criteria, i.e.,

$$\hat{i} = \arg \min_{i=0,1,\dots,N_L} \Gamma_i. \quad (28)$$

V. INTEGRITY RISK FROM INCORRECT LANDMARK ASSOCIATION

Following the outline of fault analysis for data association error in the work by Joerger and Pervan [9], we also express the hazardous misleading information (HMI) for visual navigation at time epoch N_k as

$$\begin{aligned} P(\text{HMI}_{N_k}) &= P(\text{HMI}_{N_k}, \text{CA}_{1\dots N_k}) + P(\text{HMI}_{N_k}, \text{IA}_{1\dots N_k}) \\ &= P(\text{HMI}_{N_k} | \text{CA}_{1\dots N_k}) P(\text{CA}_{1\dots N_k}) \\ &\quad + P(\text{HMI}_{N_k} | \text{IA}_{1\dots N_k}) (1 - P(\text{CA}_{1\dots N_k})), \end{aligned} \quad (29)$$

where $\text{CA}_{1\dots N_k}$ is the event that all the data associations are correct from time 1 till time N_k , while $\text{IA}_{1\dots N_k}$ is the complement event that there can exist an incorrect association at any time between time 1 and time k . We can overbound the integrity risk by conservatively assuming that incorrect association events always cause hazardous misleading information, i.e., $P(\text{HMI}_{N_k} | \text{IA}_{1\dots N_k}) = 1$. Consequently, Eqn. (29) becomes an inequality that can be written as

$$P(\text{HMI}_{N_k}) \leq 1 - (1 - P(\text{HMI}_{N_k} | \text{CA}_{1\dots N_k})) P(\text{CA}_{1\dots N_k}). \quad (30)$$

The probability $P(\text{HMI}_{N_k} | \text{CA}_{1\dots N_k})$ is obtained by considering other types of errors in visual positioning, e.g., the nominal case error by applying the methods in [13] and [16]. The most significant task of quantifying the data association error is to calculate the correct association probability $P(\text{CA}_{1\dots N_k})$. The joint distribution of correct association over all the time can be factorized as

$$P(\text{CA}_{1\dots N_k}) = P(\text{CA}_1) \prod_{k=2}^{N_k} P(\text{CA}_k | \text{CA}_{k-1}). \quad (31)$$

As a result, we can obtain an analytical upper bound for the integrity risk caused by data association if the correct association probability can be sequentially calculated.

Exploiting the criteria in Eqn. (28) to associate the detected landmarks to those in the map, the correct association probability can be calculated using the following expression

$$P(\text{CA}_k | \text{CA}_{k-1}) = P\left(\Gamma_0 < \min_{i=1,\dots,N_L-1} \{\Gamma_i\}\right). \quad (32)$$

The probability of correct association depends on the feature separation between the different ordering of landmarks. It has been proved in [9] that

$$P\left(\Gamma_0 < \min_{i=1,\dots,N_L-1} \{\Gamma_i\}\right) \geq P\left(\Gamma_0 < \min_{i=1,\dots,N_L-1} \left\{\frac{\mathbf{b}_i^T \boldsymbol{\Sigma}_i^{-1} \mathbf{b}_i}{4}\right\}\right), \quad (33)$$

where \mathbf{b}_i and $\boldsymbol{\Sigma}_i$ are defined in Eqn. (24) and Eqn. (26).

Consequently, we obtain an analytical lower bound for the probability of correct association. By substituting Eqn. (33) into Eqn. (31) and Eqn. (30), an upper bound for the risk of HMI can be obtained, which accounts for the landmark association error.

VI. SIMULATION RESULTS

Fig. 4 illustrates a test scenario that we used to analyze the impact of the identical landmarks on data association. A camera moving in a line while there are three identical landmarks distributed on one side of the route. The camera moves towards the features. This can be a realistic scenario for a camera mounted on a moving car, while identical windows are on the side of the street. In this simulation, the three landmarks are 7x5 chessboards, but only four outer corners are used as extracted feature points for each landmark, i.e., $N_L = 3, m_F = 4$ in this case. The choice of the 4 corners are as shown in Fig. 5. The standard deviation for measurement noise is set to 1 [pixel], and the process noise standard deviation is 0.5 [m] for translation and 0.1 [rad] for rotation.

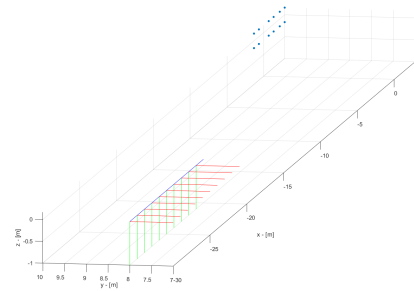


Fig. 4: Simulation scenario for repetitive patterns at one side.

Applying the method introduced in the earlier sections, the landmarks in the images can be associated with the ones in the map, and the probability of the correct association $P(\text{CA}_k | \text{CA}_{k-1})$ can be calculated. The result is shown in Fig. 6. It can be seen that as the camera moves closer to the landmarks, the separability of the features increases, so the probability of correct association also increases.

In order to investigate the impact of the complexity of the landmarks on the integrity risk due to association error, we simulate the same aforementioned scenario by using all

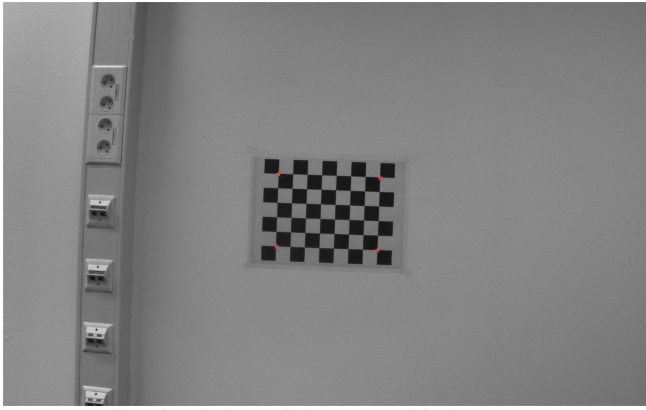


Fig. 5: Choice of the 4 specified corners.

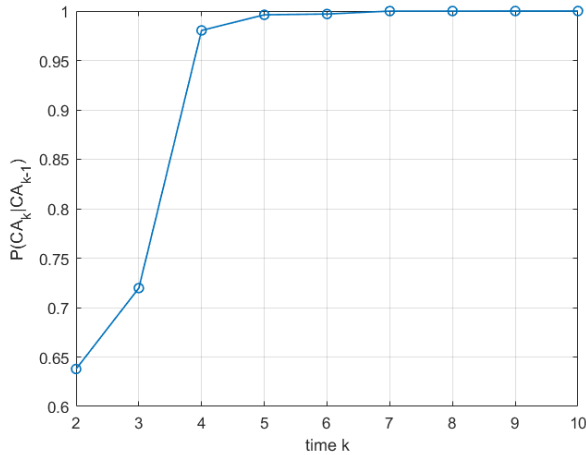


Fig. 6: Probability of correct association.

the corner points of the checkerboards. A demonstration is shown in Fig. 7. In such case, $m_F = 35$. The probability

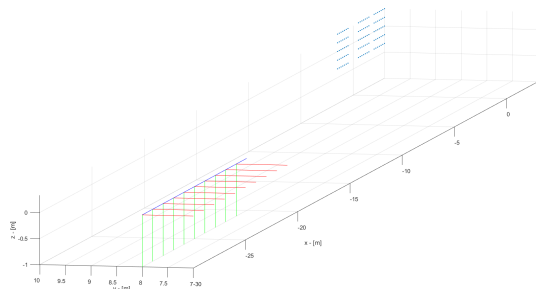


Fig. 7: Simulation scenario for repetitive patterns at one side.

of correct association by using all the 35 corners for each landmark is illustrated in Fig. 8, including the comparison with the 4 corner case. It can be concluded from the curves that the more known feature points each individual landmark contains, the smaller the probability of incorrect association between different landmarks. For better illustration, Fig. 9 shows the probability of incorrect association in logarithm scale. It can be seen that by using landmarks with only 4 features each, the probability of incorrect association is smaller than 10^{-7} if the camera moves as close to the features as $k = 10$, while the landmarks with 35 feature points

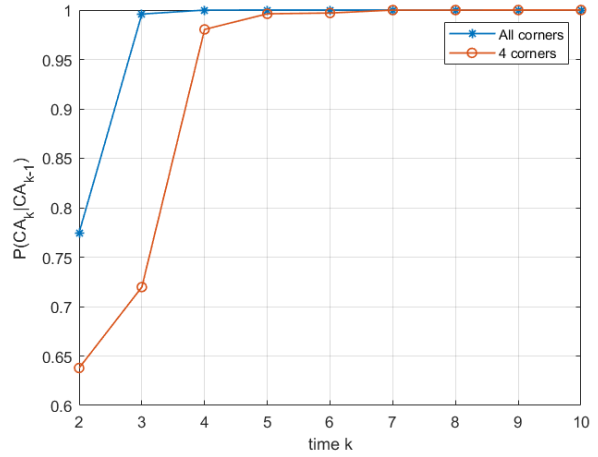


Fig. 8: Probability of correct association for different number of features per landmark.

has reached $P(IA_k | CA_{k-1}) < 10^{-7}$ from $k \geq 5$. Such data provide us information that from which distance on the visual navigation system starts to lose integrity due to incorrect association risks.

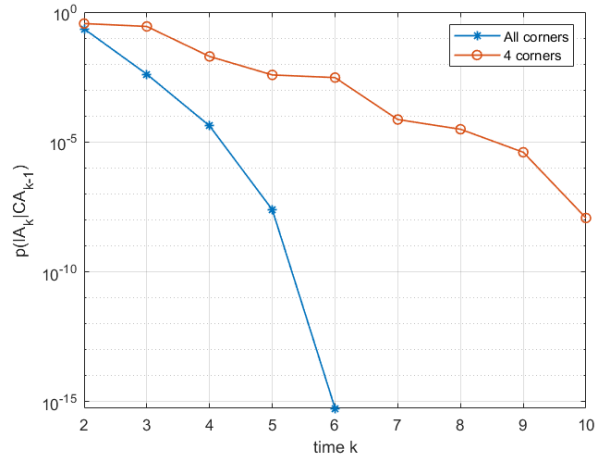
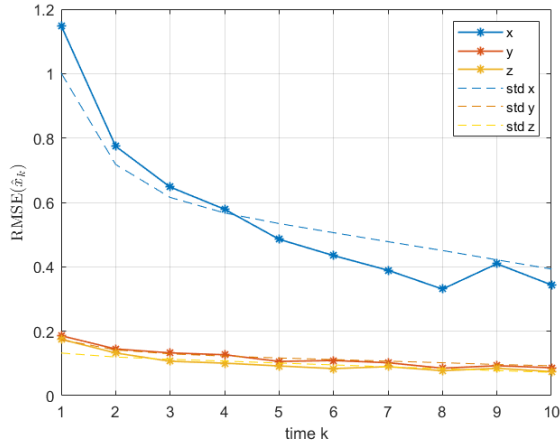
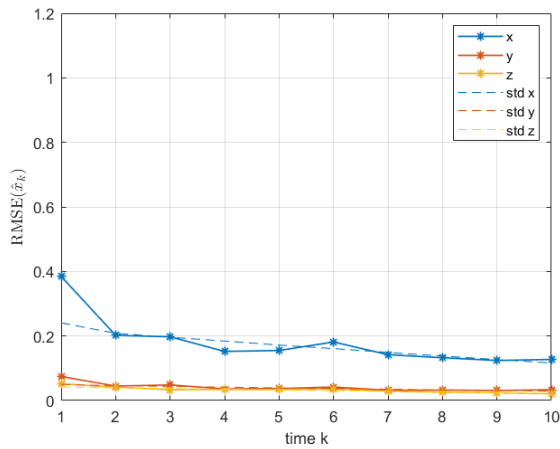


Fig. 9: Probability of incorrect association.

There is also accuracy gain by using more complex landmarks. Fig. 10 shows the position error of the EKF estimation. The error is demonstrated using root-mean-squared error (RMSE) from 100 runs of Monte-Carlo simulation. In our simulation, the standard deviation of measurement noise is set to 1 [pixel], and the process noise standard deviation is 0.5 [m] for translation. The upper figure shows the estimation error for $m_F = 4$, and the lower plot demonstrates the RMSE when using all the features, i.e., $m_F = 35$. The standard deviation of the estimated position is plotted using dashed line for the three dimensions respectively. It can be seen that the estimation exploiting all the features in the landmarks has better performance than utilizing a subset of them. However, the computational load also increases with the features per landmark.



(a) $m_F = 4$



(b) $m_F = 35$

Fig. 10: Performance of EKF-based visual navigation

VII. CONCLUSIONS

In this work we propose a method to evaluate an analytical bound for HMI probability in camera-based positioning while accounting for features association error. To the best of our knowledge, it is the first analytical description of the feature association error in visual navigation.

REFERENCES

- [1] M. Maimone, Y. Cheng, and L. Matthies, “Two years of visual odometry on the mars exploration rovers,” *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.
- [2] S. Sand, S. Zhang, M. Mühlegg, G. Falconi, C. Zhu, T. Krüger, and S. Nowak, “Swarm exploration and navigation on Mars,” in *Proceedings of International Conference on Localization and GNSS (ICL-GNSS) 2013*, Torino, Italy, June 2013.
- [3] C. S. Sharp, O. Shakernia, and S. S. Sastry, “A vision system for landing an unmanned aerial vehicle,” in *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*, vol. 2. Ieee, 2001, pp. 1720–1727.
- [4] A. Cesetti, E. Frontoni, A. Mancini, P. Zingaretti, and S. Longhi, “A vision-based guidance system for uav navigation and safe landing using natural landmarks,” *Journal of Intelligent and Robotic Systems*, vol. 57, no. 1, p. 233, Oct 2009. [Online]. Available: <https://doi.org/10.1007/s10846-009-9373-3>
- [5] M. Salarian, A. Manavella, and R. Ansari, “Accurate localization in dense urban area using google street view images,” in *2015 SAI Intelligent Systems Conference (IntelliSys)*, Nov 2015, pp. 485–490.
- [6] L. Narula, J. M. Wooten, M. J. Murrian, D. M. LaChapelle, and T. E. Humphreys, “Accurate collaborative globally-referenced digital mapping with standard gnss,” *Sensors*, vol. 18, no. 8, 2018. [Online]. Available: <https://www.mdpi.com/1424-8220/18/8/2452>
- [7] C. Zhu, G. Giorgi, Y. Lee, and C. Günther, “Enhancing accuracy in visual slam by tightly coupling sparse ranging measurements between two rovers,” in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, April 2018, pp. 440–446.
- [8] Y. Lee, C. Zhu, G. Giorgi, and C. Guenther, “Stereo vision-based simultaneous localization and mapping with ranging aid,” in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, April 2018, pp. 404–409.
- [9] M. Joerger and B. Pervan, “Quantifying safety of laser-based navigation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 1, pp. 273–288, Feb 2019.
- [10] A. Hassani, N. Morris, M. Spenko, and M. Joerger, “Experimental integrity evaluation of tightly-integrated imu/lidar including return-light intensity data,” in *Proceedings of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2019*, 1 2019, pp. 2637–2658.
- [11] M. D. Shuster, “A Survey of attitude representations,” vol. 1, no. 4, pp. 439–517, 1993.
- [12] R. M. Murray, *A mathematical introduction to robotic manipulation*. CRC press, 2017.
- [13] C. Zhu, C. Steinmetz, B. Belabbas, and M. Meurer, “Feature error model for integrity of pattern-based visual positioning,” in *Proceedings of ION GNSS+ 2019, Miami, Florida*, 2019.
- [14] V. Lepetit, F. Moreno-Noguer, and P. Fua, “EPnP: An accurate O(n) solution to the PnP problem,” *International Journal of Computer Vision*, vol. 81, no. 2, pp. 155–166, 2009.
- [15] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, “Scalable object detection using deep neural networks,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2014.
- [16] C. Zhu, C. Steinmetz, B. Belabbas, and M. Meurer, “Six degrees-of-freedom dilution of precision for integrity of camera-based localization,” in *Proceedings of ION GNSS+ 2019, Miami, Florida*, 2019.