Geo-localization using Ridgeline Features Extracted from 360-degree Images of Sand Dunes

Shogo Fukuda[1], Shintaro Nakatani[2], Masashi Nishiyama[2] and Yoshio Iwai[2]
[1]*Graduate School of Sustainability Science, Tottori University, Tottori, Japan*
[2]*Graduate School of Engineering, Tottori University, Tottori, Japan*
nishiyama@tottori-u.ac.jp

Keywords: Geo-localization, ridgelines, 360-degree images, sand dune

Abstract: We propose a method to extract the features of sand-dune ridgelines using a 360-degree camera to improve the accuracy of estimating geo-locations. It is difficult to estimate geo-locations in an outdoor environment with almost no texture such as in sand dunes. We focus on the feature of the ridgeline, which is the boundary between the ground region and the sky region. A 360-degree camera can quickly detect the ridgeline signal in all directions in a sand dune. Our method determines the current location by searching for the nearest ridgeline signal from target signals and pairing with their geo-locations. We evaluated the accuracy of this geo-localization method using synthesized images generated from a digital elevation model. We also evaluated it using real 360-degree images collected in sand dunes. We confirmed that our method significantly outperformed the existing geo-localization method on both synthesized and real images.

1 INTRODUCTION

Geo-localization systems have attracted attention to increase the use of tourism resources. A geo-localization system can guide users to specific locations and provide digital content suitable for each location. To guide users accurately, there is a need for a technique to estimate the geo-location of a user. In this paper, we discuss how to estimate geo-locations in an outdoor environment that is nearly textureless. We consider sand dunes, which are a noted tourist resource of Tottori prefecture in Japan. In sand dunes, simple and similar textures are widely spread. Thus, we have to carefully consider how to accurately estimate the geo-locations.

Many existing methods for estimating geo-locations have been proposed. Here, we classify the existing methods into three types. In the first type (Li et al., 2015; Hofmann-Wellenhof et al., 2001; Groves, 2013), the location is acquired using a GNSS (Global Navigation Satellite System) such as GPS or GLONASS. The existing methods have a problem in that the accuracy of the geo-location estimation depends on the arrangement and number of satellites that are observed from the location of the user. In the second type (Kim et al., 2017; Brejcha and Čadík, 2017; Lowry et al., 2016), the geo-location is estimated using texture patterns in the images acquired from a camera. We must consider that outliers frequently occur when simple and similar textures repeat in a sand dune environment. In the third type (Piasco et al., 2018; Chen et al., 2018; Dusha and Mejias, 2012), the accuracy of estimating the geo-location is improved using both GNSS and the camera. These existing methods set the initial location using GNSS and determine the current location using the image acquired from the camera. These methods should improve the accuracy compared with the methods that use only GNSS or a camera. However, the difficult problem of little texture still remains in sand dune environments.

Our aim is to improve the accuracy of geo-location estimation in a nearly textureless outdoor environment. We investigate robust features that can be detected from the images of sand dunes. The images include two regions, the ground and the sky. The ground region is covered with sand, so it is unsuitable for feature extraction because there is little texture. The sky region contains many objects, such as clouds and the sun, that disturb the feature extraction because these always move over time. Hence, it is difficult to use either region to improve the accuracy of estimating geo-location. In this paper, we focus on the feature of the ridgeline, which is the boundary between the ground region and the sky region. Existing methods (Saurer et al., 2016; Baatz et al., 2012; Nicolle et al., 2017;

Porzi et al., 2016) have introduced ridgelines that are robust to changes in vegetation, illumination, and season in the outdoor environment. We assume that the ridgelines work effectively even in a nearly textureless environment such as sand dunes. Note that the existing methods are designed using a general camera with a limited angle of view to extract the features. We need to consider that ridgelines spread to all 360-degree directions in sand dunes. It cannot be said that the existing methods are sufficient to represent the features of the ridgelines.

In this paper, we propose a method to extract the features of the ridgelines using a 360-degree camera. Our aim is to improve the accuracy of geo-localization in sand dunes. The 360-degree camera can quickly acquire an image containing all angles of the view and can easily detect the ridgelines of all directions. We regard the ridgelines as continuous signals and extract their amplitude components as features. Our method sets the search range of the geo-location using the initial location obtained by GNSS and determines the current location by searching for the nearest ridgeline feature. To confirm the effectiveness of our method, we evaluated the accuracy of the geo-location estimation using synthesized images generated from a digital elevation model. Furthermore, we collected the 360-degree images of a sand dune environment and confirmed that our method improves accuracy.

This paper is organized as follows. Section 2 describes the details of our method, Section 3 presents the results of a preliminary experiment using GNSS, Section 4 describes the results of geo-localization using the synthesized images, and Section 5 describes the result of the evaluation using real 360-degree images. Our concluding remarks are given in Section 6.

2 Estimation of geo-location using ridgeline signals

2.1 Ridgelines in a sand dune environment

A ridgeline generally refers to a narrow area of high land along the top of a line of hills. In this paper, the ridgeline refers to the boundary between the ground region and the sky region in the image acquired from a 360-degree camera. Figure 1(a) shows an example of a ridgeline. We represent a ridgeline as a one-dimensional waveform signal. When the user moves, parallax occurs between the ridgeline signals observed before and after movement. Therefore, we

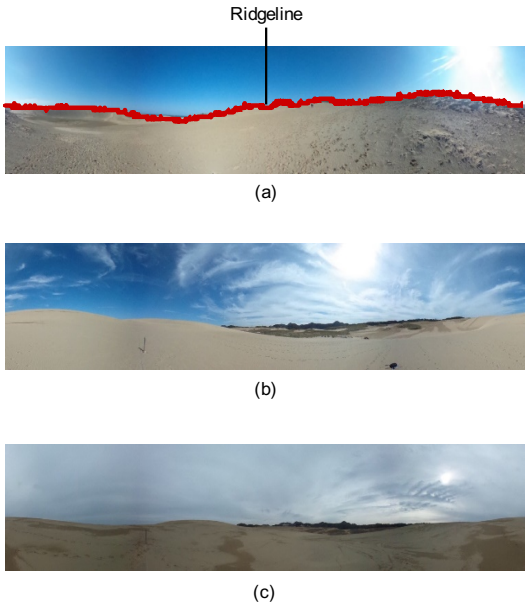

Figure 1: Example of a ridgeline detected from a 360-degree image of a sand dune environment (a). Two 360-degree images acquired under different weather conditions (b) and (c).

can use the ridgeline signals as features when estimating geo-location. We discuss the characteristics of the ridgelines below. A sand dune does not change its shape rapidly over time because of coastal conservation activities. There are no obstacles such as buildings at most locations on the sand dunes. Furthermore, it is difficult to see the buildings that stand several kilometers away because of the height difference in the sand dunes. We thus can stably observe the parallax included in the ridgeline signals in all directions. There is also an advantage in that the ridgelines can be easily detected from 360-degree images because the textures of the ground and sky regions are very different. Figures 1(b) and (c) show examples of 360-degree images acquired from the sand dunes. Given the above characteristics, we believe that geo-localization using ridgeline signals works effectively in sand dune environments.

2.2 Overview of our method

Figure 2 illustrates an overview of our method. In step S1, we generate a target database to store the pairs of ridgeline signals and their geo-locations. The system collects a 360-degree image at each location in the sand dune area and detects the ridgeline signal in advance. Steps S2 to S4 are used to estimate the geo-location of the user. In step S2, the user acquires a query 360-degree image using the camera

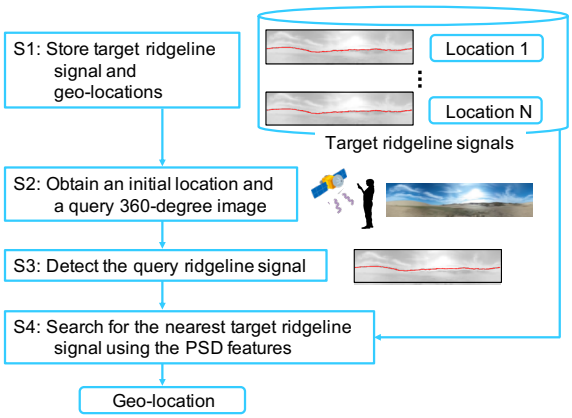

Figure 2: Overview of our method. We set the search range of the geo-location using the initial location obtained by GNSS and detect the current position by searching for the nearest ridgeline signal.

and his or her initial location using GNSS. In step S3, our method detects the ridgeline signal from the 360-degree image. We divide the 360-degree image into the sky and ground regions using color distributions, and determine the region boundary to detect the ridgeline signal. In step S4, we determine the geo-location by searching for the target ridgeline signal that is most similar to the query ridgeline signal. It is not practical to search through all ridgeline signals contained in the target database. Instead, our method searches only the neighborhood of the initial location acquired from GNSS by limiting the search range in the database. Our aim is to prevent outliers and reduce processing time. Our method determines the search range using the error of the initial location (see Section 3). The search for ridgeline signals is described in detail below.

2.3 Searching for the nearest target ridgeline signals using Power Spectrum Density (PSD) features

Our method searches for the target ridgeline signal that has the waveform that is most similar to the given query ridgeline signal. To accurately search for the ridgeline signal, it is necessary to consider the variation of camera poses when acquiring the 360-degree images. Recently available 360-degree cameras have various tilt correction functions, which suppresses the variation in attitude. To more stably extract features from ridgeline signals, our method uses the PSD (Power Spectrum Density). The PSD is a frequency-based analysis and is represented by the power spectrum components for each frequency band. The frequency analysis assumes that the query signal is a

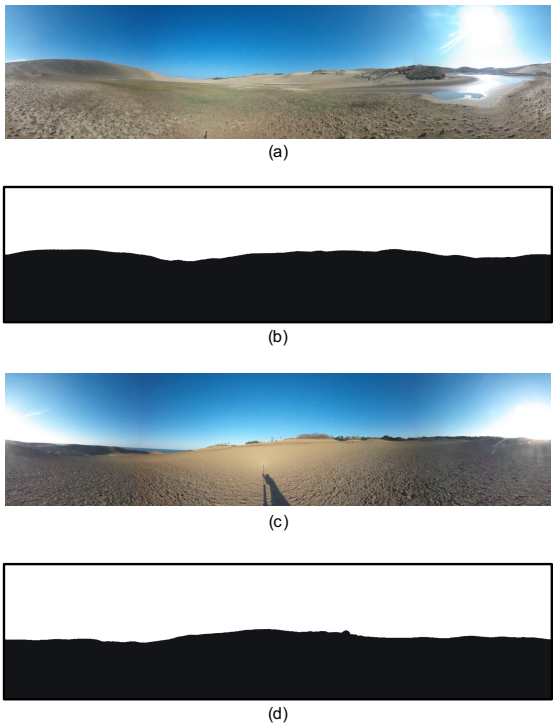

Figure 3: Examples of 360-degree images acquired from a real camera (a) and (c). Synthesized images generated from the digital elevation model (b) and (d).

continuous waveform, so it has a high affinity with the ridgeline signal acquired from a 360-degree image. The PSD also has the advantage of not being affected by phase shift. Note that we remove the components of the frequency band above $f_{\max}$ because high frequency components sometimes contain noise when detecting ridgeline signals. Our method determines the target ridgeline signal corresponding to the query ridgeline signal when the L1 norm between the two PSD features is the smallest. Our method outputs the final geo-location linked to the target ridgeline signal.

3 Preliminary experiment: the error of the initial location obtained by GNSS

Our method uses GNSS to set the search range for the target ridgeline signals from the initial location. To determine this range, we investigated how much error occurs in commonly available GNSS terminals. We used a mobile terminal (GARMIN: GPSMAP64SJ) that records the current position using GPS, GLONASS, and QZSS. We measured known location points managed by the Geospatial Informa-

tion Authority of Japan, which are generally used for map making and various topographical surveys. We acquired the current values of the latitude and longitude using the mobile terminal at each known point and compared them with the correct latitude and longitude values of the known point. We computed the error in distance between the current estimated point and the known point using these values. Three known points (L010000001070, TR35334212403, and TR45334214401) were used. We performed one-minute measurements three times on sunny days. The results of the evaluation show that the estimation error of the initial location using GNSS is 2.5 ± 1.3 m. In the following experiments, we set the search range of the target ridgeline signals to a radius of 5 m from the initial location, which is a margin of 2σ of the estimation error for the search range.

In recent years, RTK-GPS (Langley, 1998; Odijk et al., 2017) has attracted attention because it is a highly accurate GNSS and a low-cost device. The RTK-GPS system uses a reference station on the ground and corrects the estimated locations using positioning satellites. We set the reference station less than 5 km of the known points. Here, we evaluated the estimation error of the current locations using the single-band RTK-GPS (NEO-M8T, u-blox). We confirmed that the estimation error was within several tens of centimeters when the RTK-GPS system obtained fixed solutions. However, it was necessary to wait for several tens of seconds while standing the same location to obtain the fixed solution. Furthermore, there were some cases where the fixed solution could not be obtained even after waiting a long time. We believe that the use of RTK-GPS is still premature. We thus used the estimation error computed from the commonly available GNSS terminal to determine the search range of the target ridgeline signal.

4 Basic performance of geo-localization evaluated using the synthesized images

4.1 Dataset generated from the digital elevation model

We generated synthesized images of the ridgeline signals using the digital elevation model (hyo, 2 12) provided by the Geospatial Information Authority of Japan. The model contains three-dimensional meshes at intervals of 5 m. We generated the synthesized images of the ridgelines using graphics rendering engine (Unity). Figures 3(a) and (c) show real 360-degree images, and Figs. 3(b) and (d) show synthesized images assuming the same location as the real images. When generating the synthesized images, we rendered the ground regions as black and the sky regions as white. We generated a 360-degree image by connecting the center pixel rows of candidate images. We made a candidate image of each orientation by rotating a camera with a field angle of θ by A. The parameter settings $\theta = 100$ degrees and $A = 0.5$ degrees were adjusted by comparing the synthesized image with the waveform of the actual 360-degree image. The resolution of the synthesized image was set to 720×180 pixels.

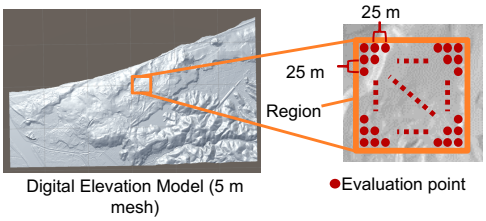

Figure 4: Evaluation points on the digital elevation model of sand dunes.

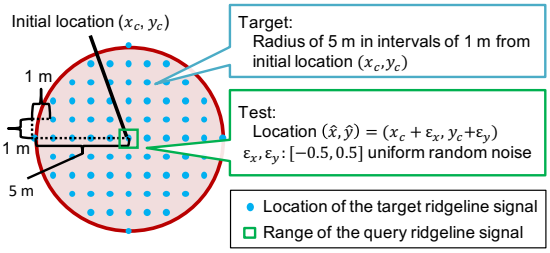

Figure 5: Locations used to generate the synthesized images at certain evaluation points.

We used an area of a 500×500 m square to generate the synthesized images. Figure 4 shows the evaluation points on the digital elevation model. We set $20 \times 20 = 400$ evaluation points on a 25-m grid. At each evaluation point, synthesized images for the target and query were generated. Figure 5 shows the locations used to generate the synthesized images at a certain evaluation point. In the 5 m radius from an initial position (x_c, y_c), which represents the center of the search range described in Section 2.3, 73 synthesized target images were generated at locations in intervals of 1 m. The 10 synthesized query images were generated at an initial location $(\hat{x}, \hat{y})$ with uniform random noise $(\varepsilon_x, \varepsilon_y)$ in the range $[-0.5, 0.5]$. The height of the camera from the ground surface was fixed at a height of 100 cm.

Table 1: Average errors of the geo-localization evaluated using the synthesized images of the digital elevation model.

	Average error [m]
Existing method	0.79 ± 0.95
Baseline	0.60 ± 0.75
Our method	$\mathbf{0.42 \pm 0.60}$

4.2 Results for synthesized images

We compared the accuracies of several geo-localization methods to confirm the effectiveness of our method by evaluating them using the synthesized images. We extracted features using the following methods:

- **Existing method**: We extracted the histogram features of the ridgelines for geo-localization using the existing method (Saurer et al., 2016). The ridgeline signals were divided into small segments. The distributions of gradients of contours in the segments were computed to generate the contour words, which are similar in spirit to visual words. We set the number of the contour words to 32.
- **Baseline**: We used the fast Fourier transform (FFT) to extract features from the ridgeline signals. The FFT is a representative approach of frequency analysis. We used the amplitude components as the feature, and set the maximum frequency band of 15 Hz to increase the accuracy.
- **Our method**: We used the PSD features of the ridgeline signals. The maximum frequency band was $f_{\max} = 15$.

We used the 360-degree images of the ridgelines synthesized in Section 4.1. All experimental conditions except for the features were the same for all three methods. We computed the Euclidean distance from the ground-truth location to the estimated location at each evaluation point. The errors of the geo-locations were calculated at all evaluation points and their averages are reported.

Table 1 shows the average errors of the geo-localization evaluated using the synthesized images of the digital elevation model. We confirmed that our method obtained higher accuracy than the existing method and the baseline. We visualized the average errors of the evaluation points with a heatmap. Figure 6(a) shows the area including all evaluation points on the digital elevation model. In Figs. 6(b), (c), and (d), the heatmaps of the average errors for the existing method, the baseline, and our method are represented, respectively. We see that the existing method (Fig. 6(b)) and the baseline (Fig. 6(c)) yielded large errors for many evaluation points. In contrast, we confirmed that our method (Fig. 6(d)) was able to estimate the geo-location at 1 m or less at most evaluation points, although the average errors were still large at some points. We discuss the reason for these large errors below.

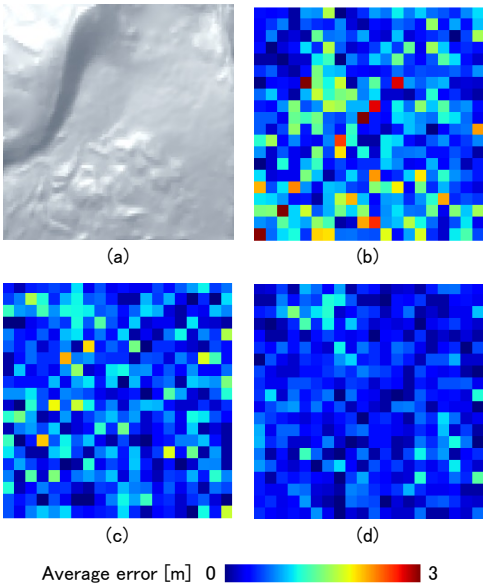

Figure 6: Visualization of the average errors at all evaluation points. (a) The three-dimensional shape of the all evaluation points when viewed from above. The average errors of the (b) existing method, (c) baseline, and (d) our method.

4.3 Discussion of the errors of geo-localization using our method

We investigated the locations at which the accuracy of the geo-localization decreases. We hypothesized that the errors are large when the ridgeline signals contain deep depth values and the parallax is very small. We generated histograms of the depth values of the rendered ridgeline signals for evaluation points with good accuracy and poor accuracy and compared them. Figure 7 shows the distribution of depth values on the top-10 and bottom-10 ridgelines with respect to evaluation point accuracy. We see that the distributions for low-accuracy points tend to have deeper depth values on the ridgeline signals than the distributions for high-accuracy points. In fact, we found that there are many low-accuracy evaluation points on the top of a sand dune. We believe that the top of a sand dune has a good view and distant mountains are observed, so the change in parallax of the ridgeline signals decreases.

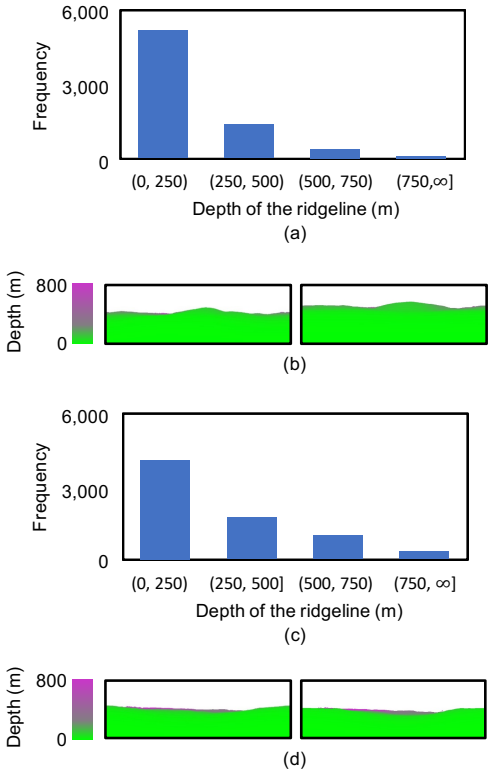

Figure 7: Distribution of depths on ridgelines and depth maps. (a) and (c) Depth distribution of the 10 most and least accurate evaluation points. (b) and (d) Depth maps of the most and least accurate evaluation points.

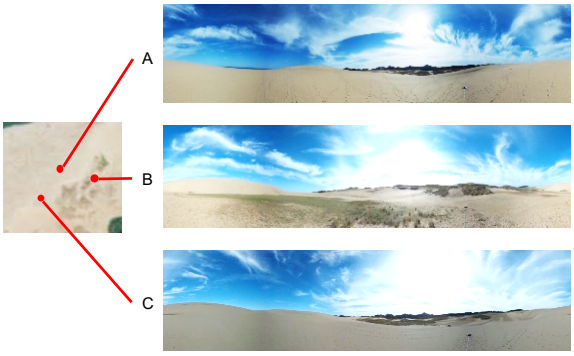

Figure 8: Examples of real 360-degree images. The target and query images were acquired from each area of the sand dunes.

5 Evaluation using real 360-degree images

We evaluated the accuracy of our method using actual 360-degree images collected in the sand dune environment. We used piles managed by Tottori prefecture to determine the locations where the 360-degree images were acquired. The piles were struck in a grid every 100 m. We defined three areas, each centered on a pile, and collected the 360-degree images in each area. Figure 8 shows examples of real 360-degree images, taken using a 360-degree camera (RICOH THETA m15) and a tripod. To generate a target database of ridgeline signals, we collected $19 \times 3 = 57$ target images per meter in the range of ± 9 m to the north starting from each pile. We collected $9 \times 3 = 27$ query images per meter in the range of ± 4 m to the north. The resolution of the 360-degree image was set to 720×180 pixels. We manually detected the ridgeline signals using a graphics editor (Adobe Photoshop). The search range of the target database was set to ± 5 m. The ridgeline signal feature extraction process was the same as that described in Section 4.2.

Table 2 shows the average errors of the geo-localization obtained using real 360-degree images.

Table 2: Average errors of geo-localization evaluated using real 360-degree images collected in a sand dune environment.

	Average error [m]
Existing method	3.62 ± 1.84
Baseline	2.42 ± 1.72
Our method	$\mathbf{1.81 \pm 1.44}$

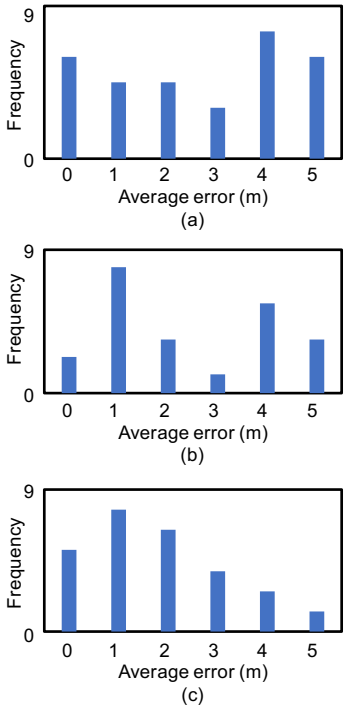

Figure 9: Distributions of the errors of the geo-localization evaluated using the real 360-degree images. Results for the (a) existing method, (b) baseline, and (c) our method.

The results show that our method obtains a smaller average error (less than 2 m) than the existing method and baseline.

We investigated how much the error was distributed for each feature. Figure 9 shows the distributions of the errors of the geo-localization using the real 360-degree images. In the results of the existing method (Fig. 9(a)) and baseline (Fig. 9(b)), the peaks of the distributions have an error of 4 m. In contrast, in our method (Fig. 9(c)), almost all cases had errors less than 1 m, though there were a few cases with errors of 5 m. The above results confirm that the proposed PSD features extracted from the ridgeline signals are effective for geo-localization.

6 Conclusions

We proposed a method for estimating geo-location using ridgeline features extracted from 360-degree images. We evaluated the accuracy of the proposed geo-localization method using synthesized images generated from a digital elevation model. We confirmed that our method substantially outperformed the existing method. Furthermore, we conducted an experiment to evaluate geo-location using real 360-degree images collected in the sand dunes. We confirmed that the average error of our method was less than 2 m.

In future work, we will further evaluate our method on various datasets of with low texture using 360-degree cameras. We will also explore the use of synthesized images to create a target database to reduce the cost of database generation.

REFERENCES

(Reference 2019-2-12). *Digital Map five meter Grid (Elevation), Geospatial Information Authority of Japan*.

Baatz, G., Saurer, O., K. Köser, K., and Pollefeys, M. (2012). Large scale visual geo-localization of images in mountainous terrain. In *Proceedings of the European Conference on Computer Vision*, pages 517–530.

Brejcha, J. and Čadík, M. (2017). State-of-the-art in visual geo-localization. *Pattern Analysis and Applications*, 20(3):613–637.

Chen, X., Hu, W., Zhang, L., Shi, Z., and Li, M. (2018). Integration of low-cost gnss and monocular cameras for simultaneous localization and mapping. *Sensors 18(7) 2193*.

Dusha, D. and Mejias, L. (2012). Error analysis and attitude observability of a monocular gps/visual odometry integrated navigation filter. *International Journal of Robotics Research*, pages 1–41.

Groves, P. D. (2013). Principles of gnss, inertial, and multisensor integrated navigation systems. *Artech House*.

Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J. (2001). Global positioning system theory and practice. *Springer-Verlag Wien*.

Kim, H. J., Dunn, E., and Frahm, J. (2017). Learned contextual feature reweighting for image geo-localization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3251–3260.

Langley, R. B. (1998). Rtk gps. In *GPS WORLD*, pages 70–76.

Li, X., Zhang, X., Ren, X., Fritsche, M., Wickert, J., and Schuh, H. (2015). Precise positioning with current multi-constellation global navigation satellite systems: Gps, glonass, galileo and beidou. *Scientific Reports*, 5(8328):1–14.

Lowry, S., Sunderhauf, N., Newman, P., Leonard, J. J., Cox, D., Corke, P., and Milford, M. J. (2016). Visual place recognition: A survey. *IEEE Transactions on Robotics*, 32(1):1–19.

Nicolle, L., Bonneton, J., Konik, H., Muselet, D., and Tougne, L. (2017). Towards an electronic orientation table: using features extracted from the image to register digital elevation model. In *Proceedings of the International Conference on Computer Vision Theory and Applications*, pages 28–38.

Odijk, D., Nadarajah, N., Zaminpardaz, S., and Teunissen, P. J. G. (2017). Gps, galileo, qzss and irnss differential isbs: estimation and application. *GPS Solutions*, 21(2):439–450.

Piasco, N., Sidibe, D., Demonceaux, C., and Gouet-Brunet, V. (2018). A survey on visual-based localization: On the benefit of heterogeneous data. *Pattern Recognition*, 74:90 – 109.

Porzi, L., Rota, B. S., and Ricci, E. (2016). A deeply-supervised deconvolutional network for horizon line detection. In *Proceedings of the 24th ACM International Conference on Multimedia*, pages 137–141.

Saurer, O., Baatz, G., Köser, K., Ladický, L., and Pollefeys, M. (2016). Image based geo-localization in the alps. *International Journal of Computer Vision*, 116(3):213–225.