2017 IEEE International Conference on Robotics and Automation (ICRA)
Singapore, May 29 - June 3, 2017

Robust Visual Localization in Changing Lighting Conditions

Pyojin Kim[1], Brian Coltin[2], Oleg Alexandrov[2], H. Jin Kim[1]

***Abstract*— We present an illumination-robust visual localization algorithm for Astrobee, a free-flying robot designed to autonomously navigate on the International Space Station (ISS). Astrobee localizes with a monocular camera and a pre-built sparse map composed of natural visual features. Astrobee must perform tasks not only during the day, but also at night when the ISS lights are dimmed. However, the localization performance degrades when the observed lighting conditions differ from the conditions when the sparse map was built. We investigate and quantify the effect of lighting variations on visual feature-based localization systems, and discover that maps built in darker conditions can also be effective in bright conditions, but the reverse is not true. We extend Astrobee's localization algorithm to make it more robust to changing-light environments on the ISS by automatically recognizing the current illumination level, and selecting an appropriate map and camera exposure time. We extensively evaluate the proposed algorithm through experiments on Astrobee.**

I. Introduction

Astrobee, a free-flying robot, is being built to autonomously navigate on the International Space Station (ISS), where it will assist astronauts, ground controllers, and researchers. It will replace the SPHERES robots, which currently operate on the ISS, but are limited to a two meter cube where they localize based on fixed ultrasonic beacons [1]. Astrobee will localize anywhere on the station through natural visual features (BRISK) stored in a pre-built sparse map [2]. See [3] for details about Astrobee's hardware. The lighting conditions on the ISS are controllable by the astronauts and change frequently. In particular, the lights are dimmed at night when the astronauts sleep. However, Astrobee's sparse maps are only constructed under a single lighting condition, and the effectiveness of this map under other lighting conditions is unknown. From the existing literature, it is unclear how the performance of visual localization systems are affected by illumination changes.

We analyze the effects of changing lighting conditions on our current visual localization algorithm, and then improve the algorithm to effectively localize regardless of lighting condition. First, we perform extensive experiments to observe how Astrobee's localization algorithm performs under various lighting conditions. To the best of our knowledge, this is the first work to empirically investigate and analyze the performance of a visual localization system under changing lighting conditions. We learn which conditions are most effective for visual localization and map building.

[1]The authors are with Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, South Korea. {rlavywls,hjinkim}@snu.ac.kr
[2]The authors are with SGT, Inc., NASA Ames Research Center, Moffett Field, CA, 94035, USA. {brian.j.coltin,oleg.alexandrov}@nasa.gov

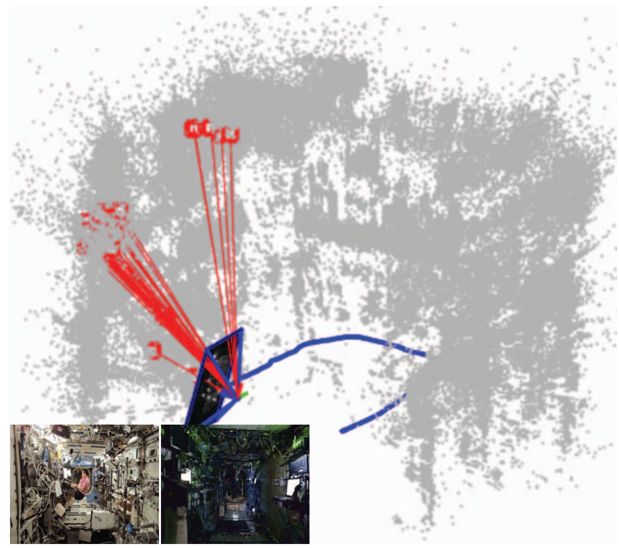

Fig. 1. Astrobee, a free-flying robot designed to autonomously navigate on the International Space Station (ISS), can localize robustly under changing lighting conditions within multiple maps reconstructed offline using structure from motion (SfM). Two pictures show light conditions on the ISS during day (left) and night (right).

Second, we present an algorithm for improved localization under changing lighting conditions. Instead of using a single pre-built map, we build multiple maps for various lighting conditions. We estimate the current brightness by comparing the image intensity distribution to the most similar images in each of the pre-built maps, and localize in the map with the closest brightness level. Furthermore, if the estimated brightness level is too dark or bright, we modify the exposure time of the camera to improve feature matching. Our work focuses on Astrobee, but the challenge of illumination-robust visual localization and our proposed solution apply equally to other robots, such as self-driving cars and UAVs.

II. Related Work

In the past decade, image-based localization has been an active research area in robotics and computer vision. From the vast literature in visual localization, we review work related to illumination changes in image sequences.

Appearance-only SLAM localizes not in the metric space but in the appearance (image) space to find the most similar images in a map. It has proven successful at loop-closure and kidnapped robot problems. FAB-MAP [4], one of the most famous image retrieval approaches, shows impressive performance through a probabilistic model of natural features in a bag-of-words [5]. However, lighting variations drastically degrade FAB-MAP's place detection performance [6].

978-1-5090-4633-1/17/$31.00 ©2017 IEEE

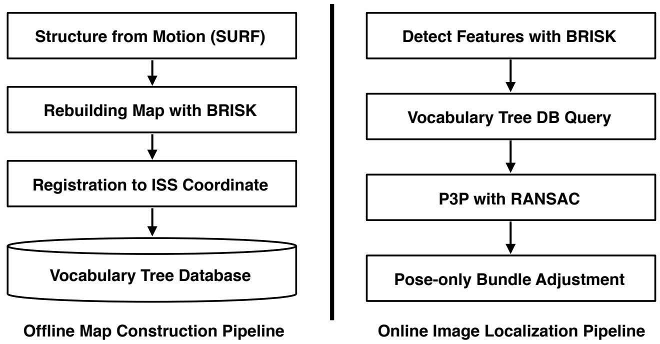

Fig. 2. Astrobee's mapping and localization algorithms.

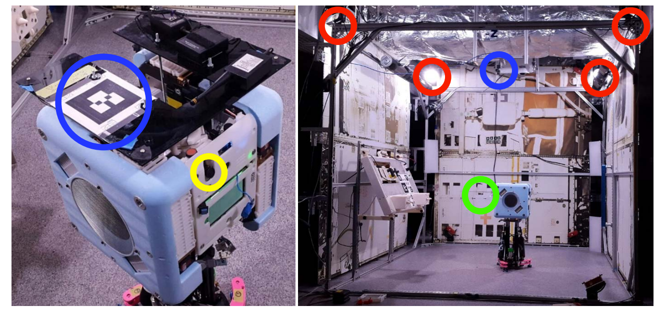

Fig. 3. Astrobee (left) and experimental environment in the granite lab (right) simulating the interior of ISS. Red circles on the right denote the adjustable lights. The green circle is the place where the intensity of light is measured with a digital light meter. The blue circles indicate the AR tag and overhead camera for providing the ground-truth position. The yellow circle indicates Astrobee's navigation camera.

To deal with the weakness of image feature algorithms such as SIFT and SURF under lighting changes [7], the fine vocabulary method [8] was introduced to learn lighting invariant feature descriptors [9]. Illumination changes have also been addressed with a high dynamic range (HDR) camera, which is used to construct a set of SIFT descriptors captured at different exposure times [10]. Another approach, SeqSLAM [6], localizes under severe lighting changes by comparing sequences of normalized images rather than comparing images pairwise. However, SeqSLAM is extremely sensitive to the changes in the camera field of view, and its computational complexity ($O(nm)$, where n is the map size and m is the sequence size) is much higher than the vocabulary tree Astrobee uses ($O(\log n)$) [11].

Numerous metric visual SLAM algorithms are based on either features [12], [13], [14] or direct (dense) image comparisons [15], [16], [17]. Some have been successfully implemented on micro aerial vehicles (MAVs) [18], [16] and smartphones [19], and show promising results. However, most SLAM and visual odometry do not explicity address and have not been tested under changing lighting conditions. An affine model for lighting correction is taken into account in [20] and [21] to avoid performance degradation from unexpected illumination changes for direct visual odometry and SLAM approaches. In [22] and [23], an illumination invariant transform for images [24] is adopted for robust visual localization of autonomous vehicles, and one-dimensional greyscale images where intensity values primarily depend on the material property have been used successfully in a metric monocular localization application. However, it is not easy to apply many existing image feature algorithms like SIFT, SURF to the illumination-invariant color space directly, and a strong assumption is required to make use of the illumination invariant image transformation.

III. Astrobee's Current Localization System

We briefly explain Astrobee's localization and mapping system. For full details, see [2].

A. Offline Map Construction

We build an offline map because the operating region is fixed, and offline maps provide higher stability and accuracy than existing SLAM systems. An overview of the offline mapping algorithm is shown in Figure 2. First, a sequence of images for offline map construction is collected by Astrobee. We match features on the images with approximate nearest neighbors on SURF features [25]. From the features, we construct tracks of features seen across multiple images [26]. An initial map is formed from the estimated essential matrices between consecutive images. Incremental and then global bundle adjustment processes optimize the camera poses and the 3D positions of landmarks by minimizing reprojection error. The map with SURF features is then rebuilt with BRISK binary features [28], which are less accurate and robust than SURF features (critical for accurate offline map building) but also much faster (critical for online localization). The map is registered to the pre-defined ISS coordinate system. Finally, a hierarchical vocabulary tree database [29] is constructed for fast retrieval of similar images for localization. The final constructed map, composed of fast BRISK features and a vocabulary database, enables quick metric localization.

B. Online Image Localization

The algorithm for online image localization, which computes the 6 DoF camera pose, is shown in Figure 2. First, we detect BRISK features in the query image. We search for these features in the bag of words vocabulary tree database to find the most similar images in the map. After finding the candidate matching features in the map, the robot pose is estimated with RANSAC and the P3P algorithm [30].

IV. Effect of Changing Lighting Conditions

We investigate the effect of changing lighting conditions by recording images under ten different conditions simulating day, night and intermediate lighting levels on the ISS. For each lighting level (ranging from 5-135 lux), nearly 2500 images were recorded, and a digital light meter on a fixed point on the wall measured the luminance in lux. The images are captured on a granite table which simulates the indoor environment of the ISS (see Figure 3). The robot slides freely on the surface of the table, constrained to motion on the two dimensional plane, but the localization algorithm computes the full 6 DoF camera pose. An overhead camera and an AR

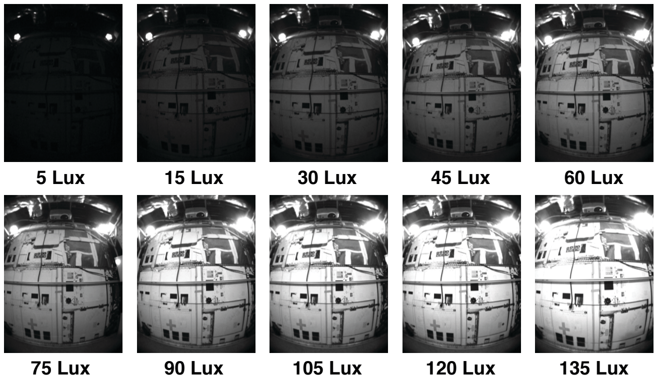

Fig. 4. Images taken under different lighting conditions in the granite lab. The wall panels imitate the interior of the ISS.

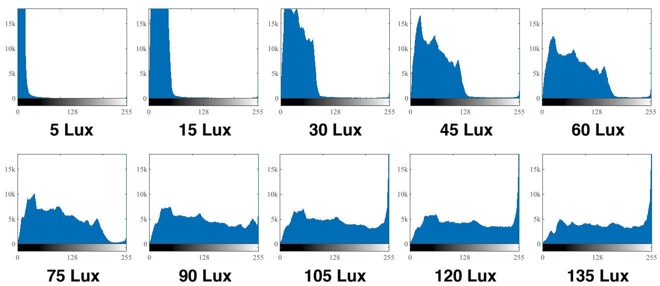

Fig. 5. Brightness distributions of the images in Figure 4. Each pixel is represented as 8-bit grayscale from 0 (black) to 255 (white).

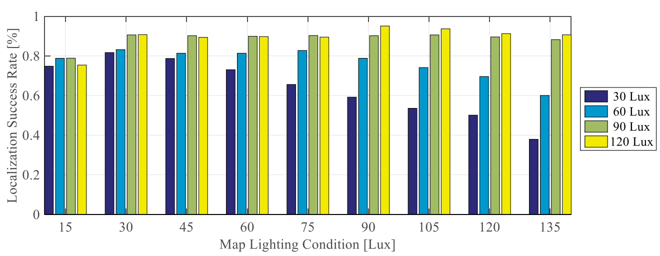

Fig. 6. The success rate for various lighting / map combinations.

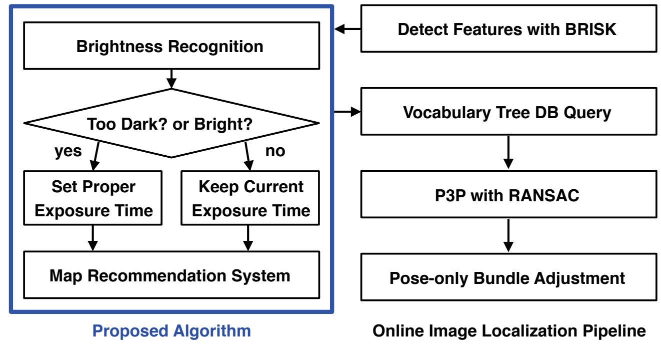

Fig. 7. The illumination-robust visual localization pipeline. The proposed algorithm (blue box) is inserted into the original pipeline (Figure 2).

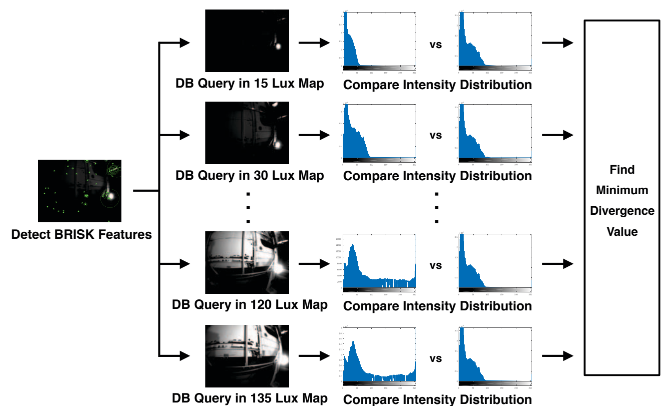

Fig. 8. Illustration of the proposed brightness recognition algorithm.

tag on Astrobee measure the ground truth pose. Figure 4 shows example images recorded for each lighting level, and Figure 5 shows the corresponding brightness distributions.

Nine maps for different lighting conditions were produced with the offline map construction algorithm. The 5 lux lighting condition is too dark to detect features to build a map, and is excluded from the remainder of the analysis.

Figure 6 shows the success rate for a subset of lighting condition and map combinations. An image is labelled as a failure if the estimated pose is outside the plane of the granite table (a 1.5x1.5 m area) or if localization fails. The number of features detected is roughly constant across all lighting levels, and the number of inlier features varies in a manner mirroring the success rate. Given that localization succeeds, the translational error varies little with map and lighting conditions, averaging around 5 cm. Dark maps below 45 lux show over an 80% success rate regardless of lighting conditions. Bright maps over 60 lux work well only with bright images. The reason for this behavior is likely that the feature descriptors cannot describe features in bright lighting conditions well because many image intensity values are saturated. In dark conditions, saturation rarely occurs. Therefore, we expect that Astrobee will fail to localize at night if the pre-built map is constructed with images captured in the day. However, a map constructed from images collected at night will work well in all lighting conditions.

Still, as one would expect, the most effective combination for stable localization is to use a map constructed in the same lighting conditions as the test images. Therefore, to make Astrobee's localization succeed regardless of the current lighting conditions, we propose to automatically recognize the environmental brightness, and choose the best of multiple maps for different lighting conditions.

V. Illumination-Robust Visual Localization

We present an illumination-robust visual localization algorithm which estimates the current lighting condition and selects an appropriate sparse map. As shown in Figure 7, the proposed algorithm is easily inserted into the existing localization algorithm. The proposed approach can be employed not only for Astrobee, but for any system that localizes in environments with changing lighting conditions.

A. Brightness Recognition

We design a brightness recognition algorithm on the premise that images taken at similar places under similar lighting will have similar intensity distributions. Figure 8 outlines the brightness recognition algorithm. We first detect BRISK features in the query image. Next, we query the bag of words vocabulary database for the images in each constructed map most similar to the detected BRISK descriptors. Then, we compute intensity distributions for each similar image, and compare them to the query image. We choose the most similar image with the symmetric KL-divergence:

$$D_{symKL}(P,Q) = \sum_{i=1}^{n} P(i) \log \frac{P(i)}{Q(i)} + \sum_{i=1}^{n} Q(i) \log \frac{Q(i)}{P(i)}$$

where P is the intensity distribution of the query image and Q is the intensity distribution of the retrieved images from each map. The smaller the KL-divergence, the more similar the two images are, hence we pick the map with the nearest brightness. Additionally, the exposure time of the camera can be controlled to extract more valuable features based on the estimated brightness. For example, the exposure time can be increased if the estimated lighting condition is too dark to detect enough features, or can be decreased to remove intensity saturation in the image if the lighting is too bright.

Note that we develop the proposed algorithm for a camera with adjustable exposure time, not for a camera with automatic exposure control. Unpredictable light variations caused by automatic exposure control make many existing visual odometry and localization algorithms unstable and unreliable. However, by changing the exposure time of the camera in a predictable manner, visual localization can account for the changes and perform stably.

B. Map Recommendation System

To achieve the best localization, the constructed map closest to the estimated current lighting condition is selected. For Astrobee's case, using the lighting conditions and maps from Section IV, we match each estimated lighting condition with a map constructed under that same condition and the same exposure time, except for in the 15 lux case, where the 30 lux map is used with double the exposure time. This selection is based on the earlier analysis, which showed that it is most effective to use a map constructed under the same lighting conditions as the test images, and from the observation that there is a linear relationship between the estimated brightness and exposure time. Note that the number of maps used can easily be adjusted according to the robot's hardware, the environment, and user preference. This algorithm requires the construction of multiple sparse maps, but enables localization under changing lighting conditions.

VI. Evaluation

We evaluate the effectiveness of the illumination-robust visual localization approach with experiments performed on the granite table (see Section IV and Figure 3).

TABLE I
EVALUATION RESULTS OF ILLUMINATION-ROBUST VISUAL LOCALIZATION ON CONSTANT EXPOSURE TIME

Experiment	Algorithm	Success Rate	Mean Inliers / Matches	RMSE (cm)
Circle	Proposed	0.99	83 / 132	5.92
	Current	0.81	73 / 130	5.92
Sideways	Proposed	0.66	55 / 115	6.28
	Current	0.51	58 / 122	5.46
Stationary	Proposed	0.89	77 / 127	7.05
	Current	0.67	82 / 131	7.11

A. Constant Exposure Time

First, we captured images from Astrobee's navigation camera at 15 Hz with a constant exposure time to confirm the effectiveness of the brightness recognition algorithm. With a changing exposure time, we are not able to directly compare to the original algorithm on the same image sequence. We recorded test runs with three movements patterns: a circle facing outwards, side to side, and stationary. As the robot moved, we brightened and dimmed the lights repeatedly, simulating the range of expected ISS lighting conditions.

For every test run, we tested the localization system with and without the proposed brightness recognition and map recommendation algorithm. Figure 9 shows sample images, the estimated brightness, the number of inlier feature matches, and the success or failure of localization for the stationary experiment with a map built under bright conditions. The estimated brightness levels closely match the actual lighting conditions. The proposed algorithm increases the localization success rate, while the original algorithm without brightness recognition fails when it is dark.

We plotted estimated moving trajectories of Astrobee for the circle and sideways experiments with 3D landmarks and ground truth trajectories in Figure 10. The estimated trajectories with the proposed method are qualitatively similar to the true moving trajectories.

We also analyzed localization accuracy quantitatively by comparing the estimated pose with the ground-truth position provided by the overhead camera. The translational error with and without our proposed improvements is shown in Figure 11. As we showed in Section IV, given that visual localization succeeds, the translational accuracy depends little on the lighting conditions of either the current environment or the map. However, we can still observe a small improvement in positional accuracy from choosing a map through brightness recognition. All of the experimental results with the constant exposure time are summarized in Table I.

B. Dynamic Exposure Time

Additional experiments were conducted on Astrobee when the exposure time is changed based on the estimated lighting condition. The proposed illumination-robust algorithm is implemented on the robot (see [3] for platform details).

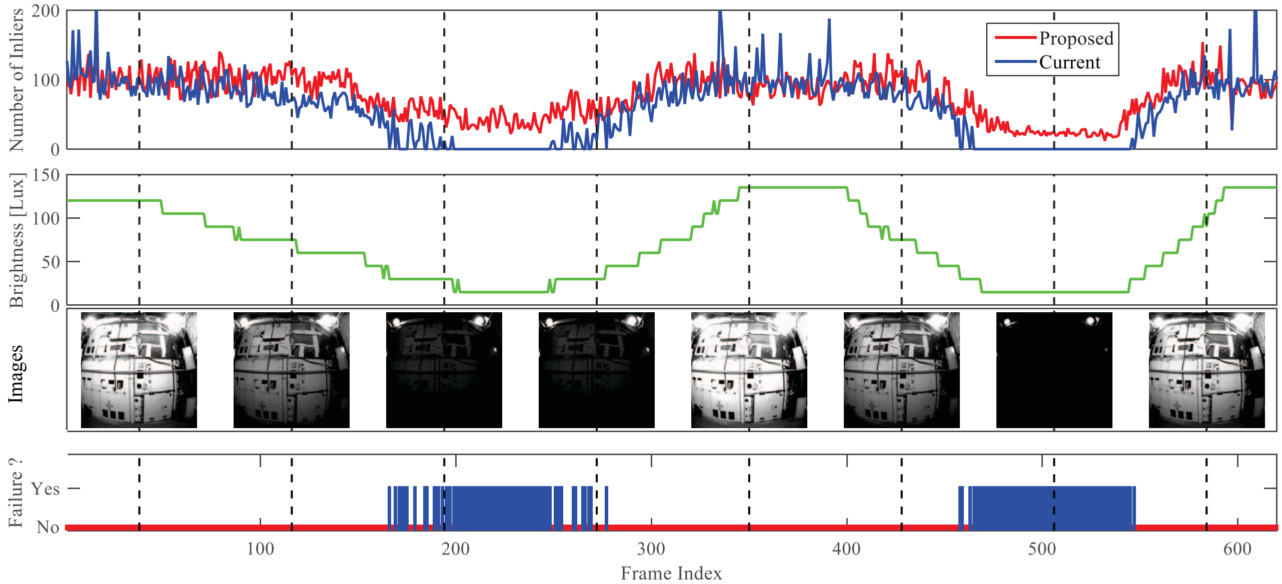

Fig. 9. Evaluation results comparing the proposed and current localization methods with constant exposure time, a bright map, and no motion. The dotted vertical lines represent the time instants at which each snapshot is taken. The current estimated lighting condition (green line) shows similar behavior to the brightness level of the actual images in the third row. Although the lights dimmed from frames 200 to 250, the proposed algorithm (red line) shows no failure and maintains the proper number of inliers whereas the current method (blue line) cannot localize.

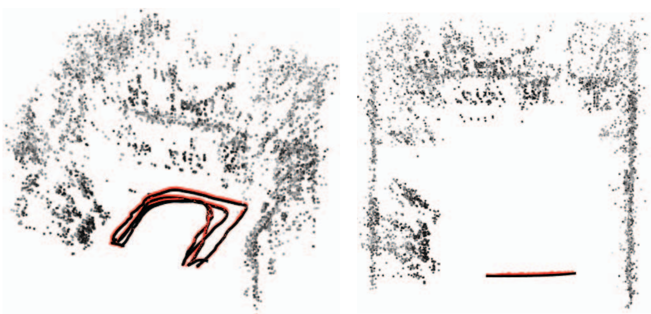

Fig. 10. 'Circle' (left) and 'Sideways' (right) trajectories estimated by our method (in red) are very similar to the ground truth trajectories (in black).

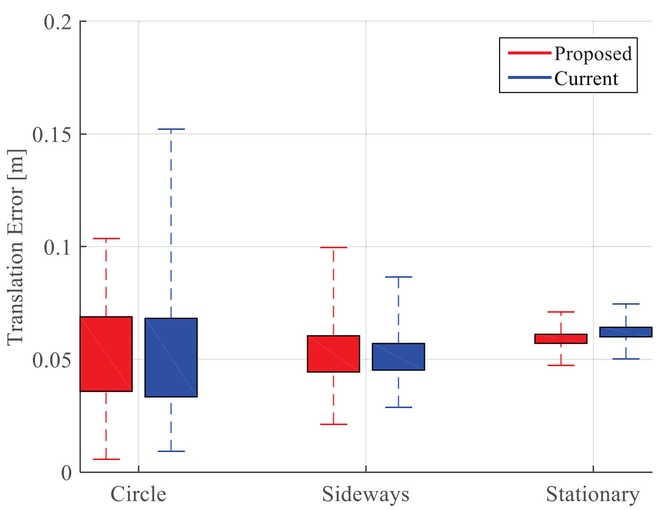

Fig. 11. Translational error of each method in the experiments.

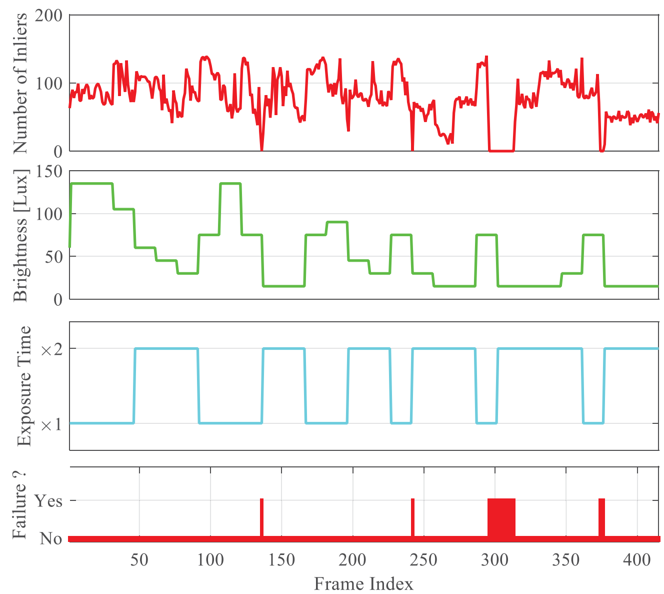

Fig. 12. Results on Astrobee in the stationary case with dynamic exposure time. The exposure time setting (cyan line) changes if the estimated lighting condition (green line) is too dark or too bright. Failures occasionally occur when the lighting condition is too dark to detect features (almost black).

Figure 12 shows that the number of correct matches, an indicator of whether localization will succeed or fail, continues to remain high by adapting the exposure time if the current estimated lighting condition is too dark (below 30 lux) or too bright (over 135 lux). Unlike existing algorithms which are very sensitive to the light variations caused by unexpected automatic exposure control, we can also achieve high stability and robustness by adjusting the exposure time in a pre-defined predictable manner. The high number of

TABLE II

ILLUMINATION-ROBUST LOCALIZATION WITH DYNAMIC EXPOSURE

Experiment	Algorithm	Success Rate	Mean Inliers / Matches
Circle	Proposed	0.92	72 / 129
	Current	0.81	73 / 130
Sideways	Proposed	0.74	61 / 112
	Current	0.51	58 / 122
Stationary	Proposed	0.94	83 / 133
	Current	0.67	82 / 131

correct matches and good success rate compared to the original algorithm are shown in Table II.

Please refer to the video clips submitted with this paper showing more details about the experiments.[1]

VII. CONCLUSION

We have investigated the performance of Astrobee's visual localization algorithm under changing lighting conditions, and presented an illumination-robust visual localization algorithm that automatically recognizes the brightness level to select an appropriate camera exposure time and map. This approach enables Astrobee to localize robustly under changing lighting conditions at the cost of building multiple lighting-specific maps. Our approach assumes uniform lighting changes; future work should consider the effects of irregular lighting changes, such as bright spots and shadows caused by sunlight coming through the windows.

ACKNOWLEDGEMENTS

We would like to thank the Astrobee engineering team and the NASA Human Exploration Telerobotics 2 project for supporting this work. The NASA Game Changing Development Program (Space Technology Mission Directorate) and ISS SPHERES Facility (Human Exploration and Operations Mission Directorate) provided funding for this work. This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science, ICT & Future Planning (MSIP) (No. 2014034854).

REFERENCES

[1] S. Nolet, "The SPHERES navigation system: from early development to on-orbit testing," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2007, p. 6354.

[2] B. Coltin, J. Fusco, Z. Moratto, O. Alexandrov, and R. Nakamura, "Localization from visual landmarks on a free-flying robot," in *IROS*. IEEE, 2016, pp. 4377–4382.

[3] J. Barlow, E. Smith, T. Smith, M. Bualat, T. Fong, C. Provencher, and H. Sanchez, "Astrobee: A new platform for free-flying robotics on the international space station," in *i-SAIRAS*, 2016.

[4] M. Cummins and P. Newman, "Fab-map: Probabilistic localization and mapping in the space of appearance," *IJRR*, vol. 27, no. 6, pp. 647–665, 2008.

[5] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," in *ICCV*. IEEE, 2003, pp. 1470–1477.

[6] M. J. Milford and G. F. Wyeth, "Seqslam: Visual route-based navigation for sunny summer days and stormy winter nights," in *ICRA*. IEEE, 2012, pp. 1643–1649.

[7] C. Valgren and A. J. Lilienthal, "Sift, surf and seasons: Long-term outdoor localization using local features." in *EMCR*, 2007.

[8] A. Mikulík, M. Perdoch, O. Chum, and J. Matas, "Learning a fine vocabulary," in *ECCV*. Springer, 2010, pp. 1–14.

[9] A. Ranganathan, S. Matsumoto, and D. Ilstrup, "Towards illumination invariance for visual localization," in *ICRA*. IEEE, 2013, pp. 3791–3798.

[10] K. Irie, T. Yoshida, and M. Tomono, "A high dynamic range vision approach to outdoor localization," in *ICRA*. IEEE, 2011, pp. 5179–5184.

[11] N. Sünderhauf, P. Neubert, and P. Protzel, "Are we there yet? challenging seqslam on a 3000 km journey across all four seasons," in *ICRA*, 2013, p. 2013.

[12] J. Civera, A. J. Davison, and J. M. Montiel, "Inverse depth parametrization for monocular slam," *TRO*, vol. 24, no. 5, pp. 932–945, 2008.

[13] G. Klein and D. Murray, "Parallel tracking and mapping on a camera phone," in *ISMAR*. IEEE, 2009, pp. 83–86.

[14] R. Mur-Artal, J. Montiel, and J. D. Tardós, "Orb-slam: a versatile and accurate monocular slam system," *TRO*, vol. 31, no. 5, pp. 1147–1163, 2015.

[15] M. Bloesch, S. Omari, M. Hutter, and R. Siegwart, "Robust visual inertial odometry using a direct ekf-based approach," in *IROS*. IEEE, 2015, pp. 298–304.

[16] C. Forster, M. Pizzoli, and D. Scaramuzza, "Svo: Fast semi-direct monocular visual odometry," in *ICRA*. IEEE, 2014, pp. 15–22.

[17] J. Engel, T. Schöps, and D. Cremers, "Lsd-slam: Large-scale direct monocular slam," in *ECCV*. Springer, 2014, pp. 834–849.

[18] S. Weiss, M. W. Achtelik, S. Lynen, M. C. Achtelik, L. Kneip, M. Chli, and R. Siegwart, "Monocular vision for long-term micro aerial vehicle state estimation: A compendium," *JFR*, vol. 30, no. 5, pp. 803–831, 2013.

[19] T. Schops, J. Engel, and D. Cremers, "Semi-dense visual odometry for ar on a smartphone," in *ISMAR*. IEEE, 2014, pp. 145–150.

[20] P. Kim, H. Lim, and H. J. Kim, "Robust visual odometry to irregular illumination changes with rgb-d camera," in *IROS*. IEEE, 2015, pp. 3688–3694.

[21] J. Engel, J. Stuckler, and D. Cremers, "Large-scale direct slam with stereo cameras," in *IROS*. IEEE, 2015, pp. 1935–1942.

[22] P. Corke, R. Paul, W. Churchill, and P. Newman, "Dealing with shadows: Capturing intrinsic scene appearance for image-based outdoor localisation," in *IROS*. IEEE, 2013, pp. 2085–2092.

[23] W. Maddern, A. Stewart, C. McManus, B. Upcroft, W. Churchill, and P. Newman, "Illumination invariant imaging: Applications in robust vision-based localisation, mapping and classification for autonomous vehicles," in *ICRA*, vol. 2, 2014, p. 3.

[24] G. D. Finlayson, S. D. Hordley, C. Lu, and M. S. Drew, "On the removal of shadows from images," *TPAMI*, vol. 28, no. 1, pp. 59–68, 2006.

[25] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *CVIU*, vol. 110, no. 3, pp. 346–359, 2008.

[26] P. Moulon and P. Monasse, "Unordered feature tracking made fast and easy," in *CVMP*, 2012, p. 1.

[27] S. Agarwal, K. Mierle, and Others, "Ceres solver," http://ceres-solver.org.

[28] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *ICCV*. IEEE, 2011, pp. 2548–2555.

[29] D. Gálvez-López and J. D. Tardos, "Bags of binary words for fast place recognition in image sequences," *TRO*, vol. 28, no. 5, pp. 1188–1197, 2012.

[30] X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng, "Complete solution classification for the perspective-three-point problem," *TPAMI*, vol. 25, no. 8, pp. 930–943, 2003.

[1]Video available at `https://youtu.be/Nuyq74wbz7I`