

In the Dead of Winter: Challenging Vision-Based Path Following in Extreme Conditions

Michael Paton, François Pomerleau and Timothy D. Barfoot

Abstract In order for vision-based navigation algorithms to extend to long-term autonomy applications, they must have the ability to reliably associate images across time. This ability is challenged in unstructured and outdoor environments, where appearance is highly variable. This is especially true in temperate winter climates, where snowfall and low sun elevation rapidly change the appearance of the scene. While there have been proposed techniques to perform localization across extreme appearance changes, they are not suitable for many navigation algorithms such as autonomous path following, which requires constant, accurate, metric localization during the robot traverse. Furthermore, recent methods that mitigate the effects of lighting change for vision algorithms do not perform well in the contrast-limited environments associated with winter. In this paper, we highlight the successes and failures of two state-of-the-art path-following algorithms in this challenging environment. From harsh lighting conditions to deep snow, we show through a series of field trials that there remain serious issues with navigation in these environments, which must be addressed in order for long-term, vision-based navigation to succeed.

1 Introduction

Appearance-based localization and mapping algorithms have enabled mobile robots to navigate autonomously through their environments using inexpensive, commercial sensors. This is appealing in that it opens the door for many exciting applications such as autonomous motor vehicles, search-and-rescue robots, and hazardous exploration robots. However, in order for these applications to succeed, robots must have the

M. Paton (✉) · F. Pomerleau · T.D. Barfoot
University of Toronto Institute for Aerospace Studies, Toronto,
Ontario M3H 5T6, Canada
e-mail: mpaton@robotics.utias.utoronto.ca

F. Pomerleau
e-mail: francois.pomerleau@robotics.utias.utoronto.ca

T.D. Barfoot
e-mail: tim.barfoot@utoronto.ca

ability to navigate reliably through their environments over long periods of time. This poses a serious problem in outdoor environments where lighting, weather, and seasonal changes quickly alter the appearance of the scene.

This problem is exacerbated in winter and polar environments where the appearance of the scene has the potential to change on a daily basis. The low elevation of the sun and short time between sunrise and sunset cause drastic changes in lighting. Light snow forms small patches of texture that melt on sunny days, while heavy snow blankets the environment in a featureless landscape as well as causing issues for path-tracking controllers. Some of these difficulties were recently observed during a field trial in the Canadian High Arctic. In August 2014, our autonomous path-following code was deployed to Alert (Nunavut, Canada) in collaboration with Defence Research and Development Canada (DRDC) (Fig. 1). Results were unsatisfactory due in part to the difficult environment.

Environments with highly variable appearance are especially difficult for applications that require vision-in-the-loop navigation. This specific task requires the vision system to provide constant, accurate, metric localization to the control loop to keep the robot driving. An example of a such a system is Stereo Visual Teach & Repeat (VT&R) [4], an autonomous path-following algorithm that navigates using vision. Proposed solutions for localization across appearance change either provide only topological localization [10, 11], require offline collection of the scene in multiple appearances [2, 9], or have under-performed in winter environments [14]. In this paper, we classify some of the difficulties associated with autonomous path following in winter environments and test the limits of two of our VT&R algorithms [14, 15] in two challenging winter field trials. We also discuss issues that need to be overcome to provide reliable, long-term, outdoor navigation using vision.

The remainder of this paper is outlined as follows. Work related to vision in feature-limited environments and localization across appearance changes is presented in Sect. 2. Brief details of the two tested VT&R systems are discussed in Sect. 3. Field trials, environmental information, and evaluation metrics are described in Sect. 4. Results are presented in Sect. 5. Lessons learned and challenges related to winter field deployments are discussed in Sect. 6 before concluding the paper.

Fig. 1 Multi-Agent Tactical Sentry (MATS) vehicle performing autonomous path following in Alert (Nunavut, Canada). Polar environments cause issues for vision-based systems such as ice, snow, and 24-h sunlight with a peak elevation of 12° . This leads to unsatisfactory results for current vision-based systems



2 Related Work

This paper presents the performance of autonomous path-following techniques in winter environments that are especially difficult for vision algorithms. These environments are difficult for a variety of reasons: snow accumulates and melts at a rapid pace, visual feature detectors do not fire on contrast-free snow, and low sun-elevation accelerates the effects of lighting change, to name a few.

Motion estimation through Visual Odometry (VO) is typically not affected by appearance change, but can suffer in these feature-limited environments. Williams and Howard [17] apply Contrast Limited Adaptive Histogram Equalization (CLAHE) to increase feature matches in images with snowy foregrounds. They show an increase in feature match count by an order of magnitude. Operating in feature-limited, volcanic fields, Otsu et al. [13] extract and track different features depending on the terrain, they show an improvement in feature count and computation speed.

Lighting change is typically the first issue seen by vision-based localization system with regard to appearance change. Color-constant images, which are partially invariant to lighting conditions [16], have recently been used to great success in vision algorithms. Corke et al. [3] show an improvement in place recognition across lighting changes using these images. McManus et al. [7] localize by switching between greyscale and color-constant images. They show improved results on a challenging dataset with significantly different lighting conditions.

While these techniques help overcome issues with lighting, general appearance change over time remains an issue. Naseer et al. [11] align sequences of images through a probabilistic network flow problem. Churchill and Newman [2] treat localization failures as new experiences and build a system of parallel localizers. Neubert et al. [12] build a dictionary that encodes the transformation of a scene between winter and summer. McManus et al. [9] train custom features that describe a specific element of the scene. While these methods are capable of localizing across appearance changes, they are not suitable for applications that require vision-in-the-loop navigation, such as autonomous path following. Some methods only provide topological localization [11], while others require that examples of the scene in multiple appearances are manually collected prior to reliable operation [2, 9, 12].

The autonomous path-following algorithms presented in this paper are built upon the Stereo VT&R work presented by Furgale and Barfoot [4]. Because this system navigates by comparing visual features from greyscale images, it is highly susceptible to lighting change. This can be overcome by using an active sensor. McManus et al. [8] perform VT&R using keypoints formed from lidar-generated intensity images and range data. While it is invariant to lighting conditions, it suffers from motion distortion issues. Krüsi et al. [6] perform autonomous path following through dense, point-cloud registration at the cost of potential failure cases in open spaces that lack geometric information. Vision-based path-following algorithms do not share these limitations, but are less stable in terms of appearance change. This paper examines the performance of the legacy system [4] as well as two improvements to the VT&R

algorithm [14, 15] that attempt to mitigate the effects of appearance change. These are presented in further detail in the following section.

3 VT&R Solutions

As an application context for visual navigation, we selected three previously published variants of VT&R solutions labeled here: *Legacy* [4], *Color-Constant* [14], and *Multi-Stereo* [15]. The details of these solutions are fully described and evaluated in their respective publications. Therefore, we only introduce them at a high level and point out the main differences. Figure 2 overviews the processing pipeline for each solution. A key element to compare is the number of images required for each pipeline, which gives an idea of the computation power required to track the robot position. The color-constant solution is the most expensive with three inputs, but remains within the range of real-time computation [14].

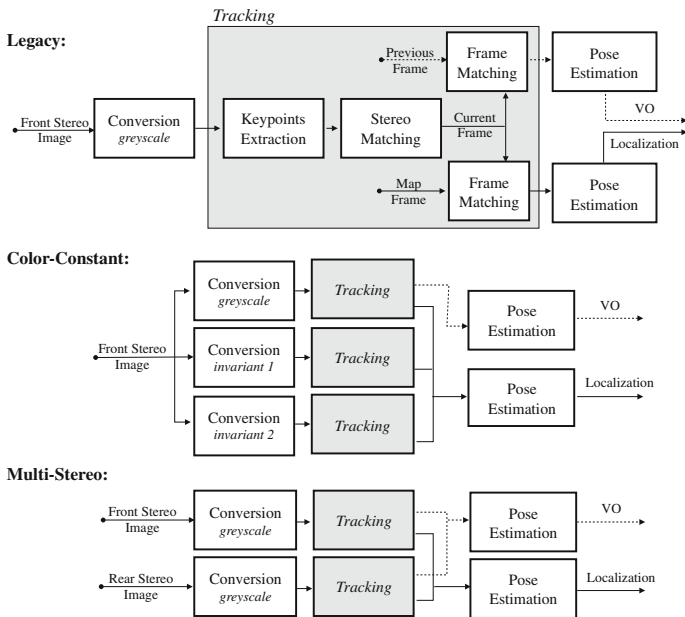


Fig. 2 Localization pipelines for the different stereo VT&R systems under investigation. The input to the system is a *left/right* RGB stereo image pair. The output is a pose estimate relative to a small subsection of the map (localization) and a pose estimate relative to the last image (VO). Incoming stereo images are first converted to different sources (i.e., greyscale, Invariant 1, and/or Invariant 2). Keypoints are extracted from each image source independently. Those keypoints are matched *left-to-right* for each respective image source to obtain depth for each feature. The 3D keypoints are then matched to a small subsection of the map to obtain feature correspondences between the live keyframe and the map keyframe. The *grey box* named *Tracking* is the same for all three solutions

(1) *Legacy VT&R*: This appearance-based path-following system is built upon the generation and tracking of keypoints, SURF [1] features with 3D position and uncertainty. Keypoints calculated from a stereo pair are organized into a keyframe. In the teaching phase, a robot is manually driven along a path while building a pose graph of keyframes connected by relative transformations. To repeat the path, the live keyframe, the collection of keypoints observed from the live stereo pair is matched to a map keyframe, a small subset of keyframes from the teach map relaxed into a single privileged coordinate frame. Data associations found between the live and map keyframe are used to obtain an estimate of the pose relative to the path, which is used to control the robot. The localization pipeline for this algorithm is illustrated in the upper section of Fig. 2.

(2) *Color-Constant VT&R*: Inspired by recent developments in the research area of color constancy, this stereo VT&R algorithm aims at increasing robustness against changes in lighting conditions. Color constancy is the ability to perceive the color of objects as constant under varying illuminations. Changes in the lighting of a scene is a major problem for appearance-based, localization algorithms that use passive sensors. This stereo VT&R pipeline is an autonomous path-following algorithm that is capable of handling significant lighting changes in a variety of outdoor environments. By expanding on the idea introduced by McManus et al. [7], the algorithm combines the accuracy of greyscale images with the robustness of color-constant images to achieve superior localization. This algorithm is identical to the *Legacy* system, with the exception of the generation of a set of two color-constant images that are partially invariant to lighting conditions. The localization pipeline is depicted in the middle section of Fig. 2. Note that tracked keypoints from each image source are fused to a single pose estimate.

(3) *Multi-Stereo VT&R*: Through multiple field deployments of Color-Constant VT&R, it was observed that failure situations were primarily due to a lack of successfully matched visual features in the environment. In the Alert field trial, we observed the camera pointing directly at the sun, causing glare. The probability of sun glares augments during the winter as the sun stays low on the horizon. The Multi-Stereo solution uses a second camera pointing behind the robot in order to augment the general number of matches and reduce the impact of glare. This pipeline is very similar to the Color-Constant solution, with the exception that image sources are coming from different cameras instead of multiple versions of the same image. Point clouds from all cameras are transformed into one common coordinate frame to obtain a single pose estimate. The localization pipeline is presented in the lower section of Fig. 2.

4 Methodology

As the goal of this paper is to quantify difficulties in harsh conditions, and not to introduce new algorithms, we describe here the datasets and evaluation metrics we explored to quantify the impact of extreme environments on visual navigation.

Throughout all the experiments, two components were kept stable: (1) the hardware and (2) the sky condition. The robot used is the Grizzly RUV from Clearpath Robotics, displayed in different environments in Fig. 3. The Grizzly is equipped with a payload that includes a suite of interoceptive and exteroceptive sensors. For the purpose of this evaluation, only the stereo cameras were used. More precisely, localization and mapping relied solely on forward and/or rear facing PGR Bumblebee XB3 stereo cameras. All experiments were executed outdoors under clear sky conditions (i.e., few or no clouds with the sun casting hash shadows on the ground).

4.1 Datasets

Three datasets demonstrate the impact of winter on visual navigation systems. As a nominal scenario, we included a summer experiment recorded at the *Canadian Space Agency (CSA)* on the Mars Emulation Terrain. We also conducted a set of trials in a *Meadow* and a field covered by *Snow* surrounding the campus of the University of Toronto Institute for Aerospace Studies (UTIAS) with the purpose of testing the limits of vision-based navigation algorithms in challenging winter environments. Displayed in Fig. 3, the two winter environments consisted of open fields with trees and buildings on the horizon, with and without the presence of heavy snowfall.

(1) *Canadian Space Agency (CSA)*: This kilometer-long dataset was recorded during the summer of 2014 in the CSA Mars Emulation Terrain and its surrounding forest. Key components of the environment include a balance of desert, marsh and forest. A continuous trajectory was recorded through those different biomes and autonomously repeated 26 times over the period of four days between sunrise and sunset in late May. We consider this dataset as our nominal scenario in terms of environment complexity and use it for comparison against winter scenarios. More details about this dataset can be found in the work of Paton et al. [14].



Fig. 3 Examples of winter environments that are challenging for vision-based navigation systems. *Left* Winter meadow consisting of dead vegetation, sparse snow patches, and trees at the horizon. *Right* Open field with deep snow cover

(2) *Winter Meadow*: This dataset was designed to test our system's robustness against lighting change and sun-stare in a challenging environment. The recording occurred in the early winter, before large snow storms covered the entire landscape. Displayed in Fig. 3-Left, this environment consists of a large field containing dead vegetation and sparse snow patches surrounded by trees and buildings in the background. This environment is difficult for vision systems for a number of reasons: (i) the dead vegetation is uniform in color and often matted to the ground, producing little contrast, (ii) tall grass moves with the wind, resulting in feature matches that are inconsistent to the movement of the robot, (iii) small patches of snow shrink and change shape as they melt, (iv) the low elevation of the sun accelerates lighting change between traverses and is often directly in the camera's field of view, which significantly changes the exposure of the image. This field trial proceeded by teaching an approximately 100m loop through this environment. The path was taught when the sun was at its highest elevation point. The robot autonomously repeated the path six times between 15:20 and 16:50 when the sun was setting (i.e., sunset happens much earlier during winter).

(3) *Snowy Landscape*: This dataset was designed to test our system's robustness against autonomous navigation through snowy environments. Snow is an especially difficult environment for vision-based systems as it is practically contrast free, causing a lack of visual features in most of the scene. Snow cover changes shape quickly as well. It accumulates, melts, turns to ice and can be blown by the wind changing the shape of the ground within minutes. Snow is also highly reflective; on sunny days this can lead a camera's autoexposure to generate images that are overexposed. An example of this environment can be seen in Fig. 3-Right, where the Grizzly is traversing through a snow covered field. A 250m path was manually driven through a large field with fresh snow cover as a teaching pass. During the teach, the sun was at its highest point in the sky, causing significant overexposure of the camera. The path was autonomously repeated approximately 3h later, when the elevation of the sun was significantly different. The complexity of the deployment and hardware limitations during this cold and windy day lead to a smaller number of repeats compared to the other dataset. Nonetheless, it is enough to draw a comparison with other environments and initial conclusions about winter deployments.

4.2 Evaluation Metrics

To evaluate the impact of extreme conditions on visual navigation, we selected three quantitative metrics: *Feature Quantity*, *Feature Uncertainty*, and *Feature Sparsity*. In this section, we describe these metrics and analyze examples from a nominal scenario (i.e., CSA dataset), which will be used as foundations for the discussion of results in Sect. 5.

(1) *Feature Quantity*: This is a notion of the amount of total inlier matches observed at any point in time between the live keyframe and the map keyframe during an autonomous traverse. Over the course of a day, this number is guaranteed to decrease

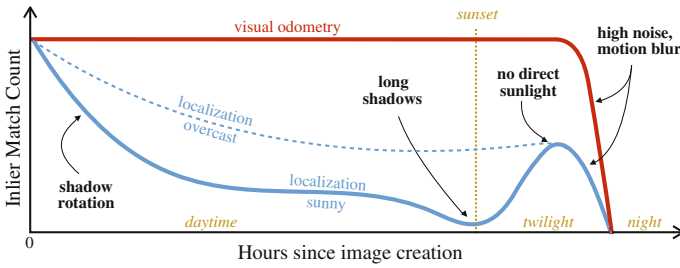


Fig. 4 Illustration of the evolution of the number of inlier feature matches through a nominal day. Time zero corresponds to when the reference images are collected (teaching phase) and the *blue line* represent the typical slow degradation of the number of matches when matching current images to the teaching phase. The difference between a sunny day (*solid line*) and an overcast day (*dashed line*) is also included. The *red line* represents the number of features used during VO, which stays constant up to the limit of the sensor. *Yellow annotations* refer to time events and *black annotations* refer to the main causes of inlier decreases or increases

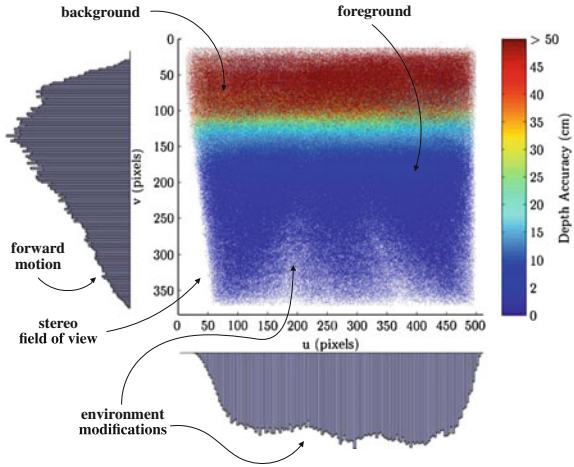
with time. If the number of inlier matches drops too low, the system will be forced to rely on VO, and eventually fail at following the taught trajectory. Figure 4 shows an illustration of the trend associated with the number of inlier matches typically observed over the course of a day. This figure sums up the experience collected during prior field tests, as reported in [14]. On overcast days, there is a gradual decline in feature matches, because the appearance of the scene is generally constant. This is not true on sunny days, where an early drop is caused by the sun changing position and creating sharp, moving shadows on the ground. Feature quantity begins to rise again at the beginning of twilight, when the light from the sun is not directly observable, generating a shadowless environment similar to an overcast day.

(2) *Feature Uncertainty*: Only considering the number of features is insufficient to ensure precise trajectory following. 3D landmarks measured with a stereo camera have an uncertainty in their depth associated with the disparity between the left and right feature matches. As this disparity decreases, the uncertainty associated with the depth reconstruction increases. High uncertainty is correlated to features observed far from the camera (i.e., in the background of the image). A reliance on background features leads to a pose estimation that is inaccurate in translation.

An example of the typical distribution of inlier matches observed between the live keyframe and map keyframe during an autonomous traverse with respect to depth uncertainty and measurement location is displayed in Fig. 5. This feature distribution is typical for a forward-looking camera on a moving robot. When moving forward, features close to the lower image border are typically not in the field of view of both the live keyframe and the map keyframe, leading to a skewed distribution of points on the vertical axis. In addition, the platform moves through the environment, generating changes in the re-observed images. On soft ground, a heavy vehicle will generate ruts that modify the deployment area over time.

(3) *Feature Sparsity*: Lastly, features can be distributed unevenly through a given trajectory. The previously mentioned metrics (i.e., feature quantity and feature

Fig. 5 Matched features with respect to the pixel coordinates aggregated through a full trajectory during the CSA field deployment. Side histograms represent the distributions of matches projected on the vertical (v-axis) and horizontal (u-axis) fields of views. All matches are colored by their depth accuracy with *dark red* being poor (>50cm) and *dark blue* being optimal. Key elements are labeled in *black*



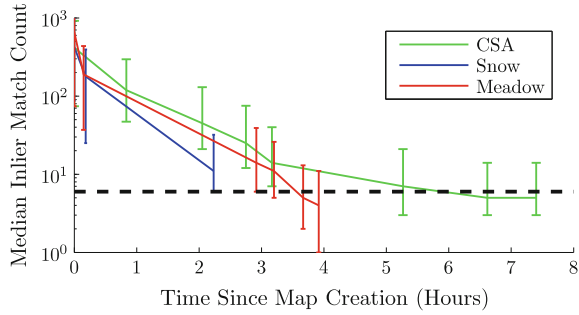
uncertainty) aggregate the data through a full repeat trajectory, limiting the analysis on consecutive successful localizations. We can indirectly observe this metric using the distance the robot relied on VO before being able to localize within its taught images. A short distance relying on VO is sign of a robust solution against the environment traversed. A system relying entirely on VO for a long period of time will increase its position uncertainty and will drift away from its reference trajectory leading to a mission failure.

5 Results

This section provides an overview of the results of our field trials with respect to the metrics defined in Sect. 4. We first perform a dataset comparison, where we look at the quantity and quality of inlier visual feature matches observed during autonomous traversals of each dataset. Results from the CSA dataset were obtained with the Color-Constant VT&R algorithm, and results from the winter trials were obtained with the Multi-Stereo VT&R algorithm. We note that color-constant images are underperforming in winter environments, and multi-stereo produces better results. We then analyze the performance of our VT&R algorithms with respect to the sparsity of successful localization matches to the map.

Figure 6 shows the rate of feature loss with respect to time since map creation. For each data set, we show the rate of feature degradation from map creation to sunset. The black horizontal line denotes our threshold match count where we can safely localize. It can be seen that when compared to the baseline dataset, the winter datasets have an accelerated decay rate. This can be primarily contributed to lighting having a much higher effect on localization, due to the low elevation of the sun and the poor performance of the color-constant images in these environments. Further reasons

Fig. 6 Feature degradation over time. In outdoor environments, the quantity of visual feature matches between the live view and map begins declining immediately after map creation. It can be seen that the rate of decline varies between data sets. Note log scale on the y-axis



for the accelerated decay rate include featureless snowy foregrounds, overexposed images, melting snow, and dead matted vegetation.

Related to feature loss, we also observe an accelerated migration of the distribution of observed feature matches towards the horizon as time passes in the winter environments. This is displayed in Fig. 7, where the distribution of inlier matches with respect to their vertical pixel coordinates over three repeats is shown for all three datasets. The green line shows this distribution when the map is compared to images collected during map creation. This is the upper limit on feature quantity as well as quality. For each data set, this distribution is nearly uniform. The blue line shows the distribution when the map is compared to the autonomous traversal taken as soon to map creation as possible, and the red line shows when the map is compared to an autonomous traversal several hours after map creation.

The distribution of our baseline comparison, the CSA data set, shows a slight migration towards the horizon after 5.2 h, yet retains a fair amount of foreground matches. In contrast, the winter data sets both show a fast shift to horizon matches only. Looking at the red lines of Fig. 7b,c, there is a significant positive skew of the distribution of matches. This means that after only a few hours in this environment, the majority of matches were obtained from the background of the image. The ramification of this is an increase in uncertainty in our localization estimate.

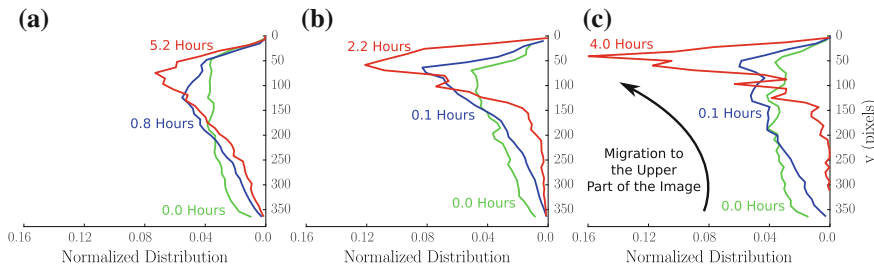
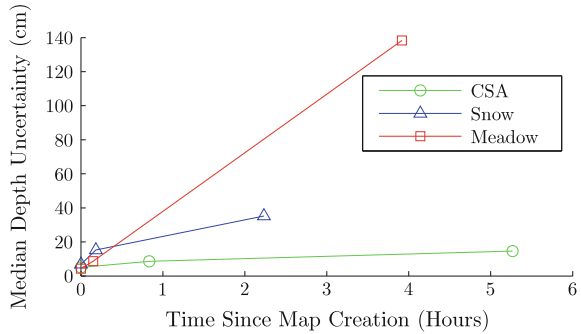


Fig. 7 Vertical distribution of the matched inlier features in the image coordinate frame. On the v-axis, 0 corresponds to a feature at *top* of the image and 360 at the *bottom* of the image. All distributions are normalized and represented over a time period of several hours for different datasets. **a** CSA, **b** Snow, **c** Meadow

Fig. 8 Median uncertainty of inlier matches over a period of several hours for all data sets. This demonstrates that the migration of inlier matches to the *upper* part of the image leads to an increased uncertainty in the estimation of the feature’s depth, which can lead to an inaccurate state estimation



This is confirmed in Fig. 8, where we plot the median uncertainty in our depth estimates for all of the inlier features observed during autonomous traversals of each data set. As expected, the CSA data set maintains a low uncertainty, while the uncertainty seen in matches during the winter data sets quickly rise. The CSA data set maintains a median uncertainty less than 20cm after 5 h, while in a fraction of the time, the Snow and Meadow data sets reach a median uncertainty level of 40 cm and 1.4 m, respectively.

If the count of inlier feature matches at a specific time step is below our threshold of six features, we discard the localization results and rely on VO for navigation. If navigation relies on dead reckoning for too long, the drift in error will cause the robot to stray from its path. We analyze the distance the robot would have driven on VO using the various VT&R methods detailed in Sect. 3. For results on the baseline CSA dataset, we refer the reader to [14]. Results with respect to sparsity are displayed in Fig. 9. These figures show the Cumulative Distribution Function (CDF) of the distance the robot would have driven on VO during the most difficult traverse of each trial. For the Snowy Landscape, this was the repeat that occurred 2.2h after map creation, for the Winter Meadow the repeat at 4.0h was chosen. The figure reads as: “for Y% of the traverse, the robot drove less than X m on VO”. The black dashed vertical line denotes the mission failure point of 20 m.

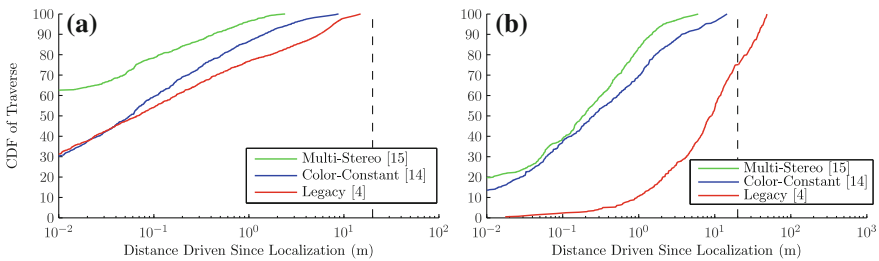


Fig. 9 Cumulative distribution of the distance the robot would have driven on VO for various algorithms on both winter datasets. *Left* Results from the second repeat of the Snow dataset, which occurred 2.2h after map creation. *Right* Results from the fourth repeat of the Meadow dataset, which occurred 4.0h after map creation. Note log scale on the x-axis. **a** Snow, **b** Meadow

For both environments, we see the trend for multi-stereo to outperform color-constant, and color-constant to outperform the legacy system. This comes as no surprise, as color-constant images were shown to underperform in these environments. This is possibly due to a lack in color information in the snow and dead vegetation. The multi-stereo system is based on greyscale images only, but has a wider field of view, having the ability to acquire more stable visual features.

6 Challenges/Lessons Learned

Snow: During the teaching phase of the Snowy Landscape data set, it was bright and sunny. Due to the high reflectivity of the snow, this caused unforeseen issues for our stereo cameras. The brightness of the scene brought the factory settings of the autoexposure algorithm of the Point Grey Research (PGR) Bumblebee XB3 to the limit. The result was saturated images, which reduced details in the foreground.

The Snowy Landscape data set was collected when there was light snow cover. We also attempted to perform autonomous path following in deep snow conditions with unsatisfactory results. In light snow, small vegetation is often visible in the foreground, providing visual features with high contrast. In deep snow, these features are gone and what remains in the foreground is nearly featureless. The only usable matched features were on the horizon not only for localization, but also for VO. This caused frequent inaccurate pose estimates, which caused issues for the path tracker. Figure 10 shows the vertical distribution of features only 0.1 h between the teach and the repeat phase, for deep snow, light snow, and meadow. The majority of matched features in the Deep Snow trial are concentrated on the upper part of the images explaining the poor performance.

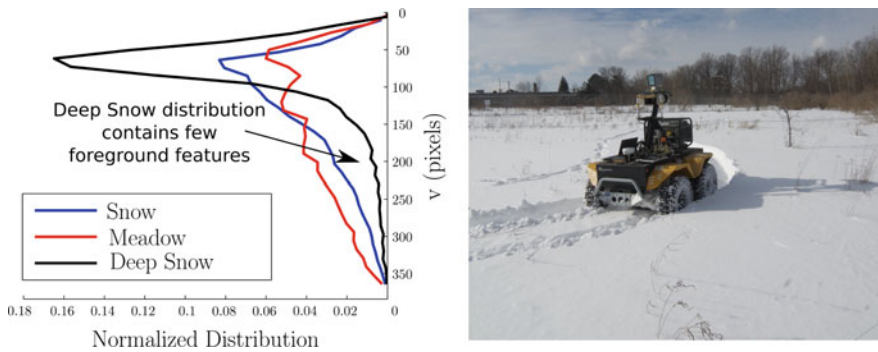


Fig. 10 Figures from the Deep Snow attempt. A lack of visual features in the foreground resulted in poor localization and VO estimates. *Left* Distribution of inlier feature matches with respect to vertical pixel location. The distribution is seen after 0.1 h for all data sets. *Right* Grizzly robot autonomously traversing in the deep snow before the failing point

Glare: An initial hypothesis motivating those field deployments was the assumption that the low elevation of the sun would cause glare in the camera, making localization impossible. Due in part to the attitude of the stereo cameras, glare was never an issue. With the cameras tilted to the ground by 20° , the sun was in the worst case only at the top of the image. We even observed cases where sun glare increased the contrast of horizon features, providing a significant boost in feature count. However, glare would be an issue if the cameras were pointed at the horizon.

Color-Constancy: The color-constant images are designed to remove the effects of lighting from the scene. These images were used to great success in the CSA field trials presented in [14]. In these trials, the robot repeated a 1 km route 26 times with an autonomy rate of 99.9 % of distance travelled in nearly every daylight condition. With this prior knowledge, the color transformations were expected to boost performance in the winter field trials presented here, but this was not the case. A hypothesis is that the color-constant images were tuned to perform in green vegetation and red-rocks-and-sand. It is possible that the dead vegetation and snowy landscapes lack the color information to remove the effects of lighting from the images.

Feature-Migration: As explained in Sect. 5, we found that the distribution of features with respect to vertical pixel location migrates to the horizon as time passes. We found that this process is accelerated in winter environments encountered in these trials. This migration results in an increase in the uncertainty of the robot's pose estimate during autonomous navigation. As the depth of observed features increase, the scale estimate becomes only loosely observable, degenerating the problem to localization based on a mono-camera. Further investigation will be required to account for this unforeseen consequence.

7 Conclusion/Future-Work

This paper presented the results of conducting a series of field trials that tested autonomous path-following algorithms in challenging winter environments. When compared to a summer dataset, we show a significant decrease in the quantity and quality of visual features matched over time. Furthermore, color-constant images that increase robustness to changes in lighting conditions have shown to be ineffective in these environments. In order for vision-based navigation to reliably navigate in these environments, we must address some of these difficult issues.

Future avenues of research may involve further classification of appearance-based matching performance in varying environments, variations in camera configurations to mitigate the issue of pose uncertainty as features migrate to the horizon, and the use of image pre-processing [17] and intelligent exposure techniques [5] to increase foreground matching in snowy environments.

Acknowledgments This work was supported by the Natural Sciences and Engineering Research Council (NSERC) through the NSERC Canadian Field Robotics Network (NCFRN). We would also like to thank Jack Collier and the DRDC for conducting a field trial on our behalf in the Canadian High Arctic and Chris Ostafew for assisting in the field trials leading up to this paper.

References

1. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). *Comput. Vis. Image Underst.* **110**(3), 346–359 (2008)
2. Churchill, W.S., Newman, P.: Experience-based navigation for long-term localisation. *Int. J. Robot. Res.* **32**(14), 1645–1661 (2013)
3. Corke, P., Paul, R., Churchill, W., Newman, P.: Dealing with shadows: capturing intrinsic scene appearance for image-based outdoor localisation. In: *Proceedings of the International Conference on Intelligent Robots and Systems (IROS)*, Nov 2013
4. Furgale, P., Barfoot, T.D.: Visual teach and repeat for long-range rover autonomy. *J. Field Robot.* **27**(5), 534–560 (2010)
5. Hrabar, S., Corke, P., Bosse, M.: High dynamic range stereo vision for outdoor mobile robotics. In: *Robotics and Automation (ICRA)* (2009)
6. Krüsi, P., Bücheler, B., Pomerleau, F., Schwesinger, U., Siegwart, R., Furgale, P.: Lighting-invariant adaptive route following using ICP. *J. Field Robot.* (2014)
7. McManus, C., Churchill, W., Maddern, W., Stewart, A., Newman, P.: Shady dealings: robust, long-term visual localisation using illumination invariance. In: *Robotics and Automation (ICRA)* (2014)
8. McManus, C., Furgale, P., Stenning, B., Barfoot, T.D.: Visual teach and repeat using appearance-based lidar. In: *Robotics and Automation (ICRA)* (2012)
9. McManus, C., Upcroft, B., Newman, P.: Scene signatures: localised and point-less features for localisation. In: *Robotics Science and Systems (RSS)* (2014)
10. Milford, M.J., Wyeth, G.F.: SeqSLAM: visual route-based navigation for sunny summer days and stormy winter nights. In: *Robotics and Automation (ICRA)* (2012)
11. Naseer, T., Spinello, L., Burgard, W., Stachniss, C.: Robust visual robot localization across seasons using network flows. In: *AAAI* (2014)
12. Neubert, P., Sunderhauf, N., Protzel, P.: Appearance change prediction for long-term navigation across seasons. In: *Mobile Robots (ECMR)* (2013)
13. Otsu, K., Otsuki, M., Kubota, T.: Experiments on stereo visual odometry in feature-less volcanic fields. In: *Field and Service Robotics. Springer Tracts in Advanced Robotics*, vol. 105, pp. 365–378 (2015)
14. Paton, M., McTavish, K., Ostafew, C., Barfoot, T.D.: It's not easy seeing green: lighting-resistant visual teach & repeat using color-constant images. In: *Robotics and Automation (ICRA)*, May 2015a
15. Paton, M., Pomerleau, F., Barfoot, T.D.: Eyes in the back of your head: robust visual teach & repeat using multiple stereo cameras. In: *Computer and Robot Vision (CRV)*, June 2015b. To Appear
16. Ratnasingam, S., Collins, S.: Study of the photodetector characteristics of a camera for color constancy in natural scenes. *J. Opt. Soc. Am. A* **27**(2), 286–294 (2010)
17. Williams, S., Howard, A.M.: Developing monocular visual pose estimation for arctic environments. *J. Field Robot.* **27**(2), 145–157 (2010)