Analysis of the applicability of visual navigation methods in arctic conditions

A V Bakhshiev¹,², E U Smirnova¹, V N Ulanov², E A Abrosimov¹,²
¹Russian State Scientific Center for Robotics and Technical Cybernetics (RTC), St.Petersburg, Russia
²Peter the Great St.Petersburg Polytechnic University (SPbPU), St.Petersburg, Russia

alexab@rtc.ru; eus@rtc.ru

Abstract. The article discusses the features of the use of autonomous navigation algorithms for mobile robotics in the Arctic. The main difficulty of the work of visual navigation methods in such conditions is that large areas of frames are occupied by a uniform snow cover, which greatly complicates the search for characteristic points and their comparison, as a result, it becomes impossible to achieve the necessary localization accuracy and map construction. The paper describes ways to improve the quality of visual navigation in arctic conditions - determining the skyline, which allows a more detailed view of the area of movement of the robot, not taking into account variable cloud cover and finding characteristic points not only on large objects on the horizon, but also on snow; adaptive contrast equalization, which is used to display such imperceptible features of a snowy surface as ribbing left by the wind, which makes it possible to distinguish features even on a visually uniform plane; the use of scale-invariant features; a special approach to the selection of key frames.

1. Introduction
Robotics in the Arctic can be used to solve tasks such as search and rescue, inspection and control of supply systems and infrastructure, and environmental monitoring. In the harsh northern climate, autonomous navigation of robotic is preferable for all these activities, the implementation of which is especially complicated by the snowy terrain [1]. To determine the location is not enough to use only global positioning systems, and to detect obstacles and further movement with the required accuracy, alternative methods of localization are needed. Using cameras or laser rangefinders significantly improves the quality of localization and allows you to build a map of the area. This article will discuss the features of visual navigation in an arctic environment.

2. Analysis of visual navigation methods
Work in open snowy environment is difficult for visual navigation. The main difficulty is that large areas of frames have a relatively small number of features due to the lack of geometric structures and large sections of a uniform snow cover. This leads to difficulties in tracking features near the robot and clustering of points along the skyline used to generate the map, due to which tracking during movement is not of high quality, as well as increased sensitivity to map optimization errors if the signs are incorrectly matched. Due to the large scatter in the depth of the scene and the relative visual visibility of the skyline points, the clustered features receive depth estimates that do not correspond to reality, in
addition, when using classical metrics, key frame selection errors occur, which does not allow obtaining accurate results of group corrections.

Visual navigation algorithms track the position of the camera in three-dimensional space and build a map in the form of a cloud of features [2]. The mathematical foundations of visual navigation are independent of the robot or the environment, therefore, the general idea can be applied without changes in the Arctic terrain. But feature extraction algorithms are usually based on a search for local areas of gradient variation, which is the main problem when working with arctic images. If you focus only on areas of the image suitable for moving the robot, they will be almost entirely white sections of the snowy desert. As can be seen in Figure 1, in such areas of the image, slight changes in color due to texture or shadows are noticeable, but the large gradients necessary for the methods for selecting features in such areas are not represented.

![Figure 1. A typical example of a terrain with the absence of any characteristic features in a passable area](image)

Figure 2 shows an example of feature extraction using two different methods - using the Harris detector and the SIFT algorithm. It should be noted that both methods found many signs in the mountains against the background, but did not single out any in the snowy foreground plane.

![Figure 2. Feature extraction using Harris detector and SIFT algorithm](image)

One way to solve the problem of lack of features in the required area is to determine the skyline. Such segmentation allows you to process only the frame area that is significant for the robot and not consider moving clouds, which can adversely affect visual navigation. In addition, the horizon serves as a good guide for estimating the orientation of the robot. There are segmenting methods according to the color criterion, but the sky covered by clouds often has almost the same color as the snow surface, therefore, such methods are supplemented by various texture characteristics.

S. Williams and A. Howard in [3] proposed the following method for segmenting frames in Arctic terrain: first, a zone is selected in the form of a trapezoid (the area directly in front of the robot that most likely contains data on the movement surface), it computes a histogram of pixel intensity to obtain a central, median and quartile values. Then the Canny Border Detector is applied to the entire image. Each
boundary is simplified to linearly piecewise using the Ramer-Douglas-Pecker algorithm. The longer the segment obtained, the more likely it is to be part of a large structure - the horizon or the mountain slope, such segments receive a larger weight coefficient relatively short. If the segment is part of the horizon, then the area below this segment should be statistically like the trapezoidal zone. To verify the fulfillment of this condition, a histogram of the intensities of the region located under the candidate segment is constructed, quartile values are compared. Similarly, the area above this segment should be statistically different from the trapezoidal zone of the foreground, their quartile values are also compared. Line segments are often generated at the edges of snow-capped mountains or at the cloud-sky boundary. By sorting all the segments by the percentage of white (snow) pixels between the segment and the bottom of the image, you can get rid of extra borders. You can further estimate the distance from the obtained segments to the estimated horizon line, based on data on the position of the camera. Using the combination of all the described weights, the final horizon segment is selected, then additional lines are searched at the ends of the selected line. The relative value of this segment is calculated as the length of the extension segment from it to the edge of the image. The variable cost is also calculated for each candidate segment as the distance between the final segment and the candidate segment plus its extension to the edge of the image. This cost is then offset by the weighting factor of the segment. Thus, segments with poor visual visibility enhance the contour of the stronger, and noticeable segments can change the direction of the horizon. Then the segment with the least solution cost is added to the initial one, and the process is repeated until the edge of the image is reached.

Another way to identify the snowy area of the robot’s movement using skyline detection was proposed by A. Des and D. Kumer [4] - after the Canny boundary detector, apply the Hough transform to search for lines. The principle of detection is the “voting” procedure in favor of the possible presence of a direct one. The element with the highest value assigned - usually the most pronounced straight line in the original image - becomes the horizon. When solving this problem, only horizontal lines are detected.

The next step to improve the performance of feature extraction algorithms should be to increase the contrast of the foreground image. The basis of most methods for increasing contrast is the processing of histograms. Such methods are usually very quick and easy to use, while achieving good results. There are two types of histogram alignment methods: global and local. Global methods typically use pixel brightness compression to obtain more consistent exposure characteristics using statistics from the entire image. These methods do not consider local image statistics. Methods of local histogram alignment allow you to get better results, as they reveal more local image details and give a noticeable improvement.

To achieve the optimal result, the local adaptive histogram equalization (CLAHE) algorithm is used [5]. The image is first divided into a grid of rectangular fragments, for each fragment the histogram of intensity is calculated. In accordance with the set threshold for trimming the peaks of the histograms, the cropped pixels are evenly redistributed across all cells forming a new histogram. Then, based on the new histogram, the normalized integral distribution function is calculated. This distribution defines a new pixel value for each source pixel. For each image fragment, the resulting pixel values are calculated using bilinear interpolation. The computational complexity of the algorithm allows you to work in real time.

Even if the image is pre-processed, the scene still does not have enough features in the area of pristine snow. However, the wind leaves a slight ribbing on the snowy surface, alternating light and slightly darker stripes that can be used as features if extract its at a suitable scale. The SIFT (scale-invariant feature transform) descriptor [6] returns a 128-component feature vector that inclusive image statistics in small areas around each feature point. Using this vector when comparing features on frames allows us to find a match even for not pronounced snow features. About hundreds of matched features are found in Arctic images with increased contrast. Just five correct comparisons are enough to generate the correct camera movement. Thus, there is a margin in case of incorrect comparisons and for methods of increasing accuracy. Figure 3 shows an example of the detected and correlated SIFT features - now the features are distinguished not only on the background mountains, but also on the snow surface [7].
found correspondences of features on two consecutive frames are shown as connected points. There are not many correspondences, but most of them are true.

![Figure 3. Selection and comparison of features in processed images](image)

In real-time systems, algorithms for detecting FAST (features from accelerated segment test) signs are more widely used - the sequence of checks and their total number are selected and optimized in advance based on an extensive training set of images, as a result of which such features are detected faster. In the part of the frame below the selected skyline, smaller and more accurate features are distinguished; Above the skyline, large gross signs are considered.

Systems that use real-time visual navigation cannot use each frame for calculations when working with localization and map algorithms, so only certain frames, called key frames, are usually considered. For non-key frames, the position of the robot is unknown, so they do not contribute to the calculations. However, tracing features is a rather noisy process, and features are often lost on several frames and only then are detected again. Therefore, it is especially important to choose key frames correctly.

There are many approaches to the selection of key frames. Usually they are based on the triangulation of points, information on the “joint visibility” of features on the frames, or on the number of matches of the tracked points. In arctic conditions, researchers A. Des and D. Kumer [4] used a metric for reducing the entropy of features, resulting in a smaller map with the same localization accuracy.

### 3. Analysis of laser navigation methods

In addition to optical navigation methods in severe arctic conditions, methods for analyzing the environment using lidars are also applicable. As a device, the lidar is an active optical rangefinder that allows you to build two-dimensional or three-dimensional maps of the surrounding space.

In the conditions of uniformity of the relief, the most applicable are amplitude- and frequency-modulated continuous-wave lidars. Such lidars have a scanning range from several tens to hundreds of meters and make it possible to achieve submillimeter accuracy [8,9].

Scanning lidars have a low vertical angular resolution, which makes them suitable for constructing a map of the underlying surface only in a very limited area in front of the mobile robot. However, a large scanning range allows you to receive sparse maps of the surrounding space in the form of point clouds, if there are any obstacles in the scanning area.

Another way to use lasers when navigating in arctic conditions is to use structured laser light in the near zone in front of the robot, followed by analysis of the laser pattern.

The essence of the method is that a dense laser grid is projected in front of the robot during movement. The reflection of the laser grid from the surface is read by the camera. Since the distance to the backlight plane and the angle between the plane and the optical axis of the camera are specified constructively, the discrepancy between the received pixel coordinates of the grid and the reference values allows us to judge the nature of the surface immediately in front of the robot. Based on this
analysis, the so-called 2.5D maps are constructed in which each cell of a flat terrain map also reflects the height of the surface. This type of map allows the local navigation of a mobile robot.

An example of applying the laser grid projection method to determine the relief of the underlying surface is shown in Figure 4.

![Figure 4](image)

**Figure 4.** The use of structured laser illumination to determine the relief of the underlying surface

The advantages of using laser systems for navigation in arctic conditions include the fact that, unlike cameras, which are passive sources, laser systems can also function in dark conditions. In addition, scanning lidars produce a point cloud containing directly the distances to the obstacle. In this regard, there is no need for additional processing of the frame to identify the position of obstacles, as is the case with conventional cameras.

The disadvantages of lidars include, first, their high cost compared to conventional cameras in the optical range. Often, the price of such devices can reach several thousand dollars, which exceeds the cost of the mobile robot itself. Scanning lidars are quite sensitive to external temperature conditions in view of the complexity of the mechanical device and the presence of precisely calibrated optics. However, solid-state lidars are practically devoid of this drawback, since only the MEMS mirror rotates inside, which reflects the beam onto a stationary diffuser. Figure 5 shows a schematic diagram of the operation of the device and a real solid-state lidar.

![Figure 5](image)

**Figure 5.** Schematic diagram of the functioning of the solid-state lidar (a) and solid-state lidar Velodyne "Velarray" (b)

Often, the disadvantages of lidars include the impossibility of working in bad weather conditions, in particular with heavy snowfall. Undoubtedly, in the conditions of Arctic latitudes, snowfalls are an important factor that must be considered when designing a navigation system for a mobile robot. However, modern filtration methods can almost eliminate the influence of falling snow, while maintaining the key points of the environment, as shown in Figure 6 [10].
Figure 6. The cloud of points of the scanning laser range finder in the conditions of snowfall before (a) and after filtering (b)

Thus, considering all the advantages and disadvantages, the use of laser navigation systems in the Arctic is advisable, but only when combined with other types of sensors, such as cameras.

4. The discussion
All the above methods can be divided into groups according to the types of sensors used, the parameters determined and the output. The diagram of such a partition is graphically presented in Figure 7.

Figure 7. General scheme of visual navigation methods used in arctic conditions
Methods that use the image from the camera as input allow you to select skylines and characteristic features of the relief on the images. Determining the skyline allows you to process only a significant part of the frame. The determination of the relief features is based on the alternation of light and dark areas of ribbing on the surface of the snow created by the wind.

All these methods as output result in a segmented image that cannot be directly used as a navigation map. In addition, all camera-based methods are strongly influenced by weather conditions and changes in scene illumination, which makes them inapplicable in the dark and in severe snowstorms.

Methods based on laser technologies allow you to build detailed three-dimensional maps of the environment using scanning laser rangefinders. Such maps can be used by a mobile robot directly for navigation. However, scanning laser rangefinders are also highly influenced by weather conditions and have a high cost compared to cameras.

The distortion analysis method of structured laser light combines technology based on lasers and classic cameras. Because the laser is an active source of light, this method is applicable in conditions of poor visibility. Also, this method is slightly affected by bad weather conditions due to the high contrast of the laser and the environment. With its help it is impossible to get a full-fledged navigation map of the surrounding space, however it is quite suitable for local navigation.

A separate significant problem is climatic conditions. From this point of view, passive devices get advantages, because they are much easier to manufacture for operation in conditions of significant freezing temperatures. The problem of creating active navigation systems suitable for use in the climatic conditions of the Arctic is the subject of a separate study.

5. Conclusion

The use of additional add-ons to classical visual navigation algorithms, such as selection the region of interest by detecting the skyline, adaptive contrast alignment to features extraction on a uniform snow surface, as well as a special approach to the selection of key frames when creating a map, allows using such systems even in Arctic terrain difficult for computer vision.

In addition, a navigation system using laser scanning rangefinders or structured laser light can become an additional source of information for the vision system. Combining sensors of various types within a single vision system allows us to use the efficiency of each type of sensor.

Only a comprehensive approach to creating computer vision systems in the conditions of Arctic can lead to the creation of a system that works in conditions of uniformity and low geometric structure of the Arctic landscape.

Acknowledgments

This work was done as the part of the state task of the Ministry of Education and Science of Russia No. 75-01195-20-00.

References

[1] Williams S 2011 Visual Arctic navigation: Techniques for autonomous agents in glacial environments ProQuest Diss. Theses August 177
[2] Yousif K, Bab-Hadiashar A, Hoseinnezhad R 2015 An Overview to Visual Odometry and Visual SLAM: Applications to Mobile Robotics. Intell Ind Syst 1 289-311
[3] Williams S, Howard A M 2011 Horizon line estimation in glacial environments using multiple visual cues IEEE International Conference on Robotics and Automation 5887-5892
[4] Das A, Kumar D, Bably A 2016 St. Waslander. Taming the North: Multi-Camera Parallel Tracking and Mapping in Snow-Laden Environments Springer Tracts in Advanced Robotics 113 345-359
[5] Reza A M Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement Journal of VLSI Signal Processing 38(1) 35-44
[6] Patent US 6711293 B1. Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image. David Lowe's patent for the SIFT algorithm.
March 23 2004

[7] Williams S, Howard A M 2010 Developing monocular visual pose estimation for arctic environments *Journal of Field Robotics* **27**(2) 145-157

[8] Rosten E, Reitmayr G and Drummond T 2005 Real-time video annotations for augmented reality *In Advances in Visual Computing* 294-302

[9] Behroozpour B, Sandborn Ph A M, Ming C W, Boser B E 2017 Lidar system architectures and circuits *IEEE Communications Magazine* **55**(10)135-142

[10] Charron N, Phillips S and Waslander S L 2018 De-noising of Lidar Point Clouds Corrupted by Snowfall *15th Conference on Computer and Robot Vision (CRV)* 254-261