20pt 20pt

Evaluation of Computer Vision Algorithms for Autonomous Navigation in Polar Terrains

N. Vandapel R. Chatila S. J. Moorehead D. Apostolopoulos
S. Lacroix W. Whittaker

LAAS-CNRS
7 av. du Colonel Roche
31077 Toulouse, France

Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213,USA

Abstract. In this paper we address the issue of Computer Vision in Antarctica for robot navigation by analysing images collected at Patriot Hills, Antarctica in the Fall of 1998. Conditions produced by polar weather and terrain are unique and challenging for perception equipment and computer vision algorithms. The later aspect will be studied here through the evaluation of a colour segmentation algorithm, a stereovision algorithm and a feature points detector and tracker algorithm.

1 Introduction

In the Fall of 1998, Carnegie Mellon University deployed the mobile robot *Nomad* to Patriot Hills, Antarctica, as part of the Robotic Antarctic Meteorite Search program [1]. Two main demonstrations were conducted: autonomous classification of rocks [10] and autonomous navigation [7]. During the same expedition, LAAS-CNRS tested a set of sensors - cameras and laser - in order to evaluate their usefulness and capabilities to be used for robot navigation, scene interpretation and environment modelling. Applications envisioned were going from logistics for supporting human activities to an autonomous scientific rover. A preliminary investigation is reported in [12].

Three kinds of environments were encountered during the 35 day expedition over an area of 10 by 2 km around our base camp: snow fields, blue ice and moraine. Snow fields are made of hard packed snow more or less eroded by the wind. Erosion produces snow dunes, called sastruggi. Sastruggi can range in height from centimetres to metre. Blue ice, snow transformed into ice by a geological process, is covered by snow patches and by suncups, 10 cm scale undulations in the ice surface. A moraine is an area of blue ice where rocks have been deposited by ice movement. Rock size varies from centimetre to metre heights. Weather conditions experienced can be divided into four classes: sunny, sunny with blowing snow, overcast and white-out.

In such an environment, robot safety is challenged by slippery slopes, negative obstacles between sastruggi where robot wheels can get stuck, terrain discontinuity, hidden crevasses under snow bridges on blue ice field and rocks on moraines. Furthermore, useful visual information may not be accessible or even exist, for example meteorites hidden under snow cannot be seen, or an absence of terrain features for localization and mapping.

Polar terrain and weather produce challenging conditions for computer vision algorithms. Images seem, at least *a priori*, featureless, colourless, uniform, homogeneous,

textureless, highly reflective and dynamic (blowing snow). In this paper we propose to evaluate three computer vision algorithms related to the above mentioned tasks. In the second part we evaluate a colour segmentation algorithm for sastruggi detection, snow patches segmentation, and skyline detection. In the third part we evaluate a stereovision algorithm for the purpose of obstacle detection. In the fourth and last part we evaluate a feature point detector and tracker for landmark based navigation or egomotion. Algorithms used were developed or implemented in our lab. But first we will state where our paper fits in the context of computer vision algorithm evaluation.

2 Algorithm evaluation

In [3] Christensen summarized commonly given arguments against "Performance characteristics of vision algorithms". There is no room in this paper to start a discussion on that broad and passionate topic. Instead we choose the 10 reported objections to detail our approach, goals and motivation. We tested algorithms in a *task dependant manner*. Functionalities tested were envisioned as *a module of a complex system*, namely a robot, see [7]. *Vision is complex* and we attempted to present some of the issues related to computer vision in polar terrains by using representative data, but not exhaustive, and well known CV methods. Images were collected using Nomad's sensors - two sets of black and white stereo cameras and a colour camera - and by using 3-CCD stereo colour cameras mounted on a tripod [12]. Due to the experimental set-up and harsh environment it was not possible to acquire the same scene twice under different weather conditions. So representative images of the classes of environment and weather mentioned in the introduction will be compared. *Models used were wrong* but approximation was enough to complete our tasks. *The quality measure* was used to decide when and where such algorithms can be used and not to compare several algorithms for a given task. No *theoretical validation* was performed. *Tuning parameters* were well identified and tested. They were selected by hand after several set and trial cycles for each test in order to give the best results. *Ground truth* was not available so results were evaluated visually by the authors.

3 Colour segmentation

The colour images checked exhibited interesting features. For example the green channel tends to saturate whereas the red and blue channels do not. Also shadows and snow under direct sunlight exhibit different colour compositions, see the left part of figure 2. This gave us the idea to use colour information to segment snow patches on blue ice, sastruggi and skyline in colour images by using the following algorithm in the framework described in section 1.

We used a colour segmentation algorithm available in our lab developed by R. Murrieta-Cid during his PhD thesis [9]. The image space is divided into square cells - the level to start division is the first of the two parameters of the method. A colour space division is performed by analysing the colour histogram. Each image cell is associated with an histogram class and by using an adjacency graph, cells from the same class are merged. Small areas - under a threshold, the second parameter - are merged into the closest adjacent region. This algorithm was successfully used to segment outdoor images of natural scenes into large areas. Its drawback is that it tends to oversegment images.

Criterion for evaluation is the ability to find edges between snow and ice, between sky and ground and between different sastruggi. Four qualitative levels - unable, poor, medium and good - were defined. Table 1 reports the results for a set of 23 images and figure 1 and 2 show three segmented images.

	White-out	Overcast	Sunny
Sastruggi	unable	poor	poor
Snow patches	medium	good	poor with specular reflection, good otherwise
Skyline	unable	poor	medium

Table 1: Qualitative evaluation of colour segmentation algorithm.

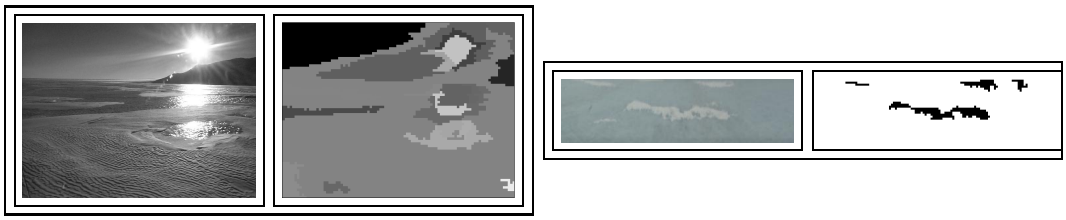

Figure 1: Colour segmentation examples of snow patch on blue ice. Left : during a sunny day with specular reflection. Right during an overcast day.

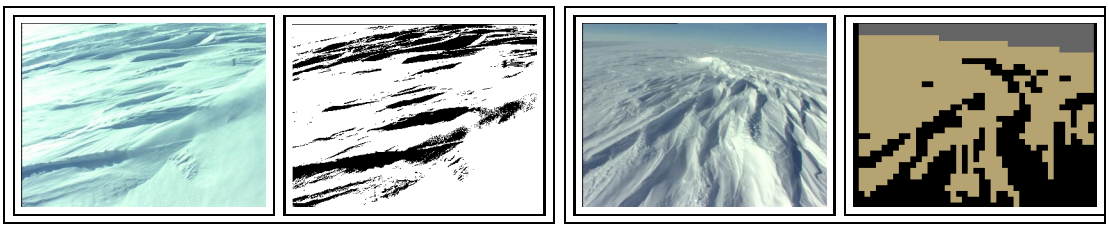

Figure 2: Colour segmentation examples. Left: example of shadow detection. Original colour image of dm scale sastruggi taken at 5 m and segmented image, in black areas the blue channel dominates. Segmented image is filtered with an erosion and dilatation algorithm. Right sastruggi field during a sunny day. Original colour image dm scale sastruggi and segmented images. In grey skyline, in brown snow background and in black extracted area.

In figure 2 the algorithm correctly detected the overall sastruggi field but not individual sastruggi even if it can detect shadows. The best results were gathered for snow patch detection on blue ice. An example is given in figure 1-right. But specularities, as shown in figure 1-left, degraded results. Skyline detection was correctly performed during a sunny day over the Patriot Hills. Otherwise the algorithm is unable to distinguish the sky from the ground for white-out and overcast conditions. Even for sunny conditions, the haze on the horizon line mislead the algorithm.

4 Evaluation of stereo-vision

Pixel-based stereovision is now widely used in several natural terrain robotics projects and developments [8, 2, 5]. Indeed, most natural environments (and especially planetary-like environments) are so textured that pixel correlation based algorithms are very efficient, and produce dense and precise 3D images. As compared to other range sensors

(namely Ladars and Radars), stereovision requires much less energy. Moreover, the direct association of intensity, colour or texture information to the 3D points produced by stereo facilitates the implementation of various functionalities required by autonomous navigation (*e.g.* motion estimation [6], terrain modelling ...): it is therefore worth investigating the possibility of using stereovision in polar terrains.

However, polar environments are challenging for stereovision: due to the low texture of the scenes and the lighting conditions, and the usual pixel correlation-based algorithms do not give satisfactory results. As a comparison, figure 3 shows the result of the same "classical" algorithm (ZNCC correlation [5]), with images provided by the same cameras.

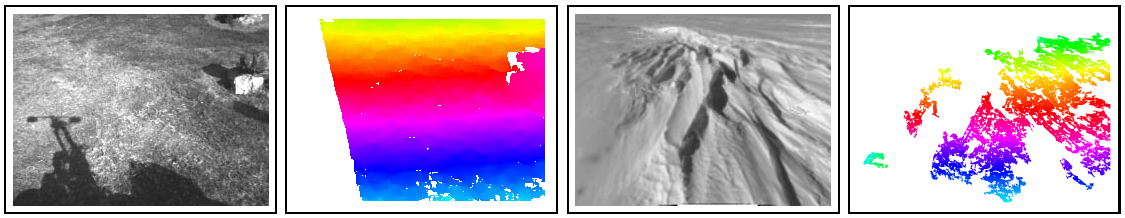

Figure 3: The same stereovision algorithm applied on images acquired by the same cameras. Top: an "almost perfect" disparity map obtained in the natural terrain test-field at LAAS. Bottom: much less disparity information is recovered on Antarctic images (the colour encodes the disparity values - the inverse of the depth, and the image size is 256×192 pixels).

It is essentially the lack of texture that deters the ZNCC algorithm from finding reliable pixel matches. Figure 4 shows the correlation curves for two pixels† : the one on the bottom left corresponds to the blue pixel in the top left image, which has been successfully matched. Indeed, the peak of this correlation curve (which indicates the most probable match) has a quite large value, and is quite different from the second local maximum. As a comparison, the bottom right curve corresponds to the red pixel, for which no reliable match has been found: since this pixel is located in a low texture zone of the image, the possible candidates for a match are very ambiguous. A necessary threshold on the value of the highest peak and on the difference between this peak and the second peak lead the algorithm to a "no reliable match found" answer for this pixel.

We studied various possible improvements to the algorithm in order to obtain more dense disparity images: we checked the utility of applying filters on the images, we tried other matching scores (especially the one defined by a Hamming distance on census images [14], an interesting score which is at least as efficient as the ZNCC score), we tried to find matches in parallel in the three colour planes of the images, and spent some time tuning the various parameters. All these improvements gave slightly better results than the usual algorithm, but still provided very poor disparity maps.

However, we can obtain some significantly better results by *focusing the disparity search zone*. Indeed, the classical algorithm can give dense qualitative disparity maps on images subsampled by a large factor††. The idea is to use these first disparity estimates to focus the search at the next resolution, up to the desired level. Using this procedure, we obtain much denser disparity maps, as shown in figure 5.

5 Selection and tracking of landmarks

Mobile robots need to extract feature points [6] in their surrounding environment to navigate reliably. In this part we evaluate the ability of an algorithm to detect and

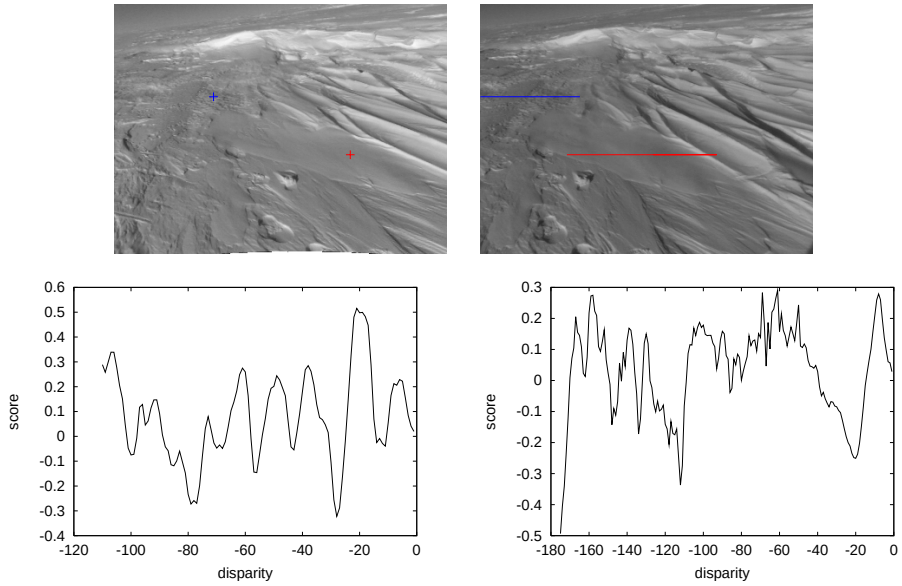

Figure 4: Example of two pixel correlation curves. Top: original stereo pair: the two selected pixels are shown on the left image, the search zone are shown on the right image - horizontal epipolar lines after rectification of the images. Bottom: the correlation curve for the two selected pixels (see text for explanations)

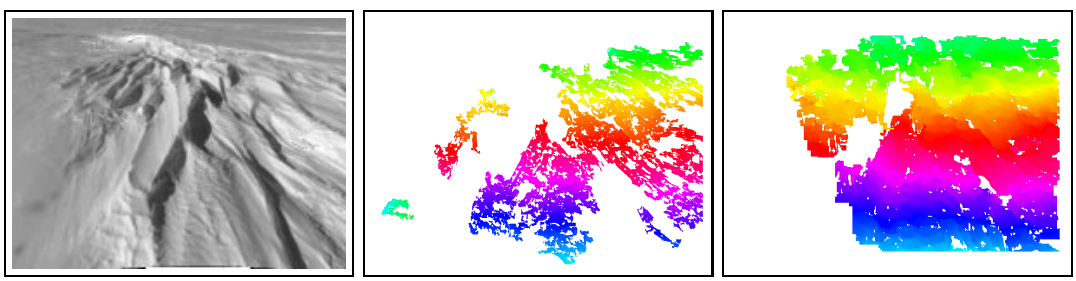

Figure 5: Improvements due to focusing the disparity range. Left image of the stereo pairs. Middle: disparity maps using the classical stereovision algorithm. Right: disparity maps produced after focusing the disparity search.

reliably track feature points in polar terrain image sequences.

We used the algorithm developed by Kanade, Lucas and Tomasi [11]. The algorithm assumes an affine motion between two consecutive images. So tracking a template from image I under such assumption in image J means finding the 6 coefficients of the affine transformation defined by $J(Ax+d) = I(x)$, where a point x in image I moves to $Ax+d$ in image J, d is the translation and $A = Id+D$, D the deformation matrix. Selection of feature points is done by using the pure translation model of the affine transformation which leads to a 2×2 linear system. Feature points are defined as an area for which the matrix of that system is well conditioned and for which the eigenvalues are above a threshold. Monitoring of feature points is done by using the full affine model. Kept feature points are areas for which the dissimilarity, defined as the difference between the template and the template transformed by the affine parameters estimated by a minimization procedure, is under a threshold.

The criteria used for evaluation are the number of frames in which a set of points

	Sunny	Overcast	White-out
Sastruggi	good, several	poor,2	null
Blue ice & snow patches	poor,2	poor,2	null
Moraine	poor,2-3	poor,2-3	no data

Table 2: Qualitative evaluation of the feature point detector and tracker. Reliability and maximum number of consecutive frames with successfully tracked points

can be tracked and the number of false matches over the feature points detected. Algorithm parameters are: the size of the support window around feature points, the eigenvalue threshold to select feature points and the dissimilarity threshold for feature point monitoring. With small displacements between images and without image depth discontinuities, like on a flat blue ice or sastruggi field, this assumption is valid. Tests on blue ice and moraine were performed with a set of 10 sequences of *Nomad* images, each sequence having up to 99 images. Tests on sastruggi fields were done on a more restricted set of images from LAAS-CNRS, 2 sets of 20 images.

Figure 6 shows results for a sastruggi field during a sunny day. Forty features were detected in the original image. Thirty eight of these feature points were tracked, 9 of which were false matches. Later in that sequence, 60 points were extracted, 26 tracked and 17 correctly matched. Due to the repetitive pattern of suncups on the blue ice, the number of false matches increases dramatically - 5 false matches over 5 tracked points on a 4 frames sequence, see figure 7. A qualitative interpretation of the results obtained are provided in table 5. Even with snow patches the quality of tracking (reliability, length of sequence) is too poor for use in robot navigation. The major reason for the poor tracking performance is the lack of texture in the images.

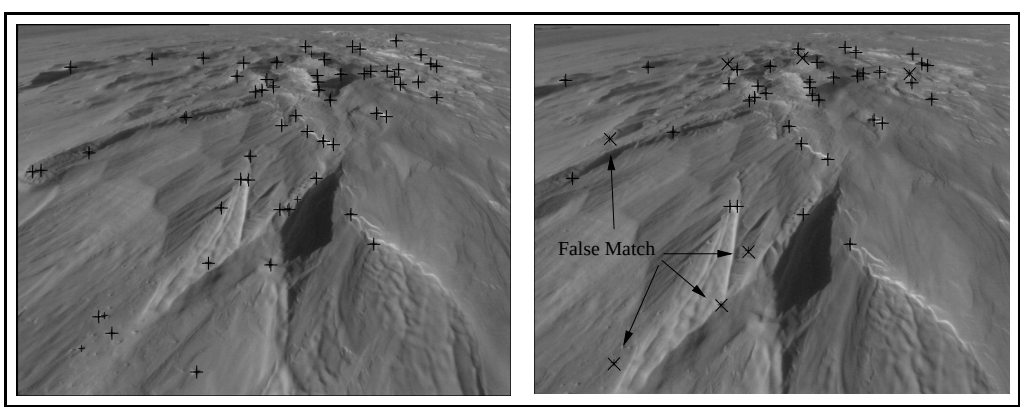

Figure 6: Feature point detection and tracking for sastruggi. Left: original image with feature points detected (shown as +). Right: next image with correctly identified feature points (+) and false matches (x)

6 Conclusion

In this paper we extented results from preliminary analyses of data collected during the RAMS'98 expedition and published in [12]. We evaluated, in a task dependent manner, three CV algorithms - color segmentation, stereo vision and feature points selection and tracking - on a representative set of images of environment and weather conditions encountered. For each test we selected tuning parameters optimizing the

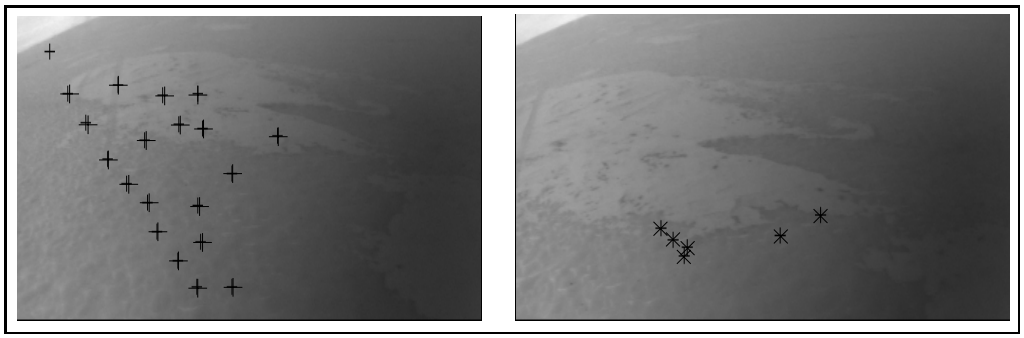

Figure 7: Feature point detection and tracking on blue ice. Left: original image with feature points detected (+). Right: next image in sequence with correctly identified feature points (+) and false matches (x)

results, - criteria used were detailed; evaluation was performed visually without ground truth.

Our main goal was to decide when and where cited algorithms could be used. Results were summarized qualitatively for color segmentation and feature points detection and tracking in table 1 and 5. The lack of texture prevent stereo-vision algorithms to provide good disparity maps, several matching score were tested. But by focusing the disparity search zone, in a hierarchical manner, denser maps were produced but still of medium quality. We tried also to understand specific issues related to CV in polar terrains and bad weather. The next step on which the work is currently on going is to propose algorithms adapted to the specific extreme situations of polar environments.

Notes

† These curves plots the ZNCC score as a function of disparity
†† With a sub-sampling factor of 16 (which produces 48×36 images), the disparity map is most of the times fully filled.

Acknowledgments

This expedition was funded by "Institut Français pour la Recherche et la Technologie Polaire" (IFRTP) and the "Région Midi-Pyrénées". Part of this work was partly funded by the grant NAG2-1233 "Accurate Localization from Visual Features" from NASA Ames Research Center. The Robotic Antarctic Meteorite Search program is funded by the Telerobotics Program NAG5-7707 of NASA's Office of Space Science Advanced Technology & Mission Studies Division. A special thanks to Martial Hebert, the RAMS expedition team, the Chilean Air Force, Rafael Murrieta-Cid and Michel Devy.

References

[1] D. Apostolopoulos, M. Wagner and W. Whittaker, "Technology and Field Demonstration Results in the Robotic Search for Antarctic Meteorites", *International Conference on Field and Service Robotics*, 1999

[2] L. Boissier and B. Hotz and C. Proy and O. Faugeras and P. Fua, "Autonmous planetary rover (V.A.P): on-board perception system concept and stereovision by correlation approach", *IEEE ICRA*, 1992

[3] H.I. Christensen and W.F. Forstner "Performances Characterics of Vision Algorithms" *Machine Vision and Applications*, vol 9, 1997

[4] S. Gautama, S. Lacroix and M. Devy, "On the Performance of Stereo Matching Algorithms", *LAAS Report 99416*, 1999

[5] E. Krotkov and M. Hebert and R. Simmons, "Stereo Perception and Dead Reckoning for a Prototype Lunar Rover", *Autonomous Robots*, 2-4, 1995

[6] A. Mallet and S. Lacroix and L. Gallo, "Position Estimation in Outdoor Environments using Pixel Tracking and Stereovision", *To be published to IEEE ICRA*, 1999

[7] S. Moorehead, R. Simmons, D. Apostolopoulos, and W.L. Whittaker, "Autonomous Navigation Field Results of a Planetary Analog Robot in Antarctica", *i-SAIRAS*, 1999

[8] L. Matthies, "Stereo Vision for Planetary Rovers: Stochastic Modeling to Near Real Time Implementation", *International Journal of Computer Vision*, 8-1, 1992

[9] R. Murrieta-Cid , M. Briot and V.Ayala-Ramirez, "Scene interpretation: segmentation and identification of regions by color and texture analysis", *LAAS Report N98183*, 1998

[10] L. Pedersen, D. Apostolopoulos, W.L. Whittaker, W. Cassidy, P. Lee, and T. Roush, "Robotic Rock Classification Using Visible Light Reflectance Spectroscopy: Preliminary results from the Robotic Antarctic Meteorite Search program", *Lunar and Planetary Science Conference*, 1999

[11] J. Shi and C. Tomasi, "Good features to track", *IEEE CVPR*,1994

[12] N. Vandapel, S. J. Moorehead, W. Whittaker, R. Chatila and R. Murrieta-Cid, "Characterization of Stereo and Laser Sensors in Antarctica" *International Symposium on Experimental Robotics*, 1999.

[13] N. Vandapel, M. Hebert and R. Chatila, "Active and attentive vision for automatic natural landmark selection" *LAAS Report N99434*, 1999

[14] R. Zabih and J. Woodfill, "A non-parametric approach to visual correspondence", *Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence*