# Panoramic Horizon Recognition

**David Rawlinson**
**david.rawlinson@eng.monash.edu.au**
**Ray Jarvis**
**ray.jarvis@eng.monash.edu.au**
**ARC Centre for Perceptive & Intelligent Machines in Complex Environments (PIMCE),**
**IRRC, Monash University**

## Abstract

This paper describes a method in which SIFT features on and around the horizon are used for image classification (matching or recognition). Ordinarily, transformational geometry relating two images is recovered from a set of point correspondances – but correspondance formation may be improved and accelerated using geometric constraints. The proposed solution to this chicken/egg problem is to first search for an approximate geometric relationship between images using only the scale, orientation and position of SIFT features; if found, the recovered transformational geometry is used to filter feature matching. Generous feature position tolerance simplifies the search to 1 dimension in unwarped panoramic images.

## 1 Introduction

### 1.1 The Horizon in Nature and Robotics

The horizon is a strong and easily recovered feature of many outdoor vistas, being absent during daylight hours only when obscured by dense vegetation, artificial structures or particularly rugged terrain. Furthermore the Sun is usually a distinct feature of the horizon and together these features suffice to determine a ground plane and a canonical orientation within it (assuming for a moment that the Earth is flat and not rotating). Animals including *Catagylphis* ants and honeybees have been shown to use the Sun and other visual features navigationally [Lambrinos, M¨oller 1999, Weber, Venkatesh 1997].

The distant horizon and Sun are rarely used as a navigational aid in terrestrial robotics (one notable exception being Sahabot2 [Lambrinos, M¨oller 1999]), perhaps because researchers' laboratories are typically located in exactly the type of structured environments that obscure it. In the buregoning Unmanned Aerial Vehicle (UAV) domain the horizon is an essential feature for the maintenance of straight and level flight as evidenced by the presence of an "artificial horizon" on most such vehicles. Although the primary purpose of the artificial horizon is attitude determination some UAV researchers have deliberately detected topographic features on the horizon for mapping purposes [Kweon and Kanade 1991].

Although in many natural environments a terrestrial robot is prevented from viewing the horizon the behaviour of the robot might be modified to deliberately approach lookouts and other far-horizon viewpoints. Humn walking-paths tend to encompass such places.

### 1.2 Panoramic Imaging

The horizon is most useful when much of it is visible i.e. in panoramic views. Moreover, given a rotationally-invariant image representation the robot localisation problem is reduced from 3 dimensions (x,y,bearing) to 2 (x,y). Of course, it is not necessary to use a panoramic camera: Perspective images may be registered and stitched together [Brown and Lowe 2003]: This is computationally expensive, but probably anthropically justified to some extent [Cohn 1997].

In the experiments described in this paper a panoramic camera apparatus is employed. A common concern with panoramic cameras is the effective loss of resolution in areas close to the robot and at low elevation. The horizon is distant and generally at high relative elevation thereby mitigating this problem. Interestingly, the animals studied in [Lambrinos, M¨oller 1999, Weber, Venkatesh 1997] tend to have wide-angle or nearly panoramic de-foveated vision.

### 1.3 Local Feature Image Analysis

Whole-image analysis methods are usually classified as "appearance"-based (generally statistics computed on all pixels e.g. histograms of colour distribution [Prasser, Milford 2005, Carson, Belongie 1997, Fitzgibbons and Nebot 2002]) or "feature" -based. To classify an entire image using point (more honestly, local) features, the features must be modeled collectively or measured for consistency in some way. Local features used in image classification include edges or lines [Canny 1986], corners [Smith and Brady 1995, Nist´er, Naroditsky 2004], shape (usually shape skeleton [Rosenfeld and Pfaltz 1966] or peaks in contour curvature [Barenholtz, Cohen 2003]). Local features may also include or be derived from image "regions" segmented from the image on the basis of operators such as those above or statistically [Hernandez and Barner 2000].

One problem with linear as opposed to point features is the retention of dimension: Lines parallel with the plane of movement are less useful for measuring

position within that plane and vertical lines are equivalent to point features, in such a system.

### 1.4 Image Classification Using SIFT Features

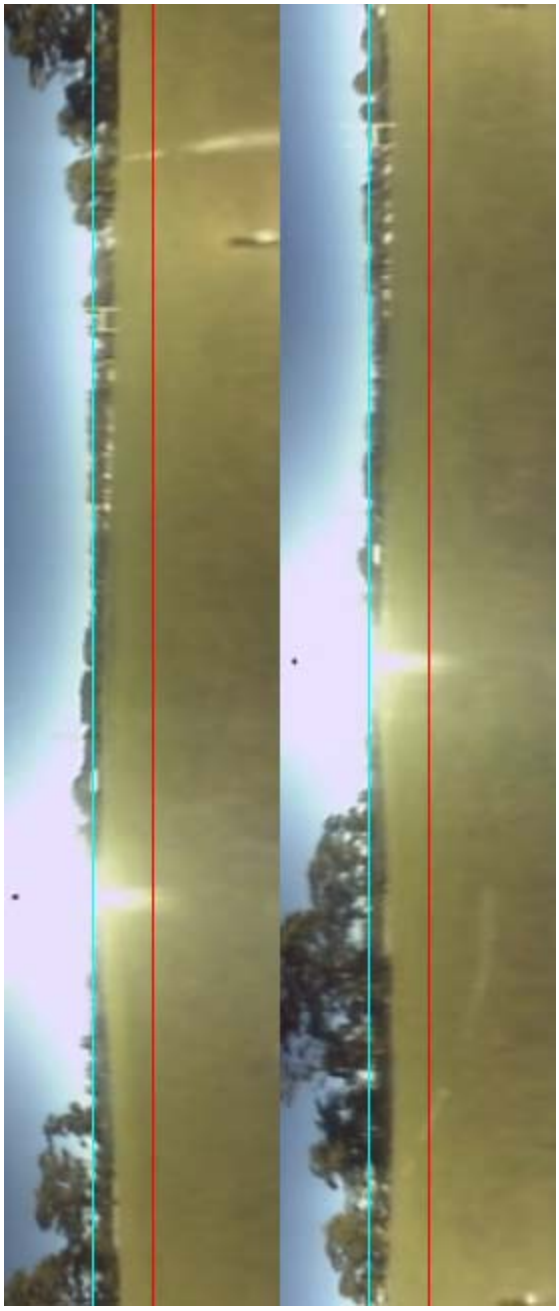

Figure 1: 2 panoramic images captured in an open area at Monash University; note the detected horizon and the Sun, a strong & distinctive landmark.

Recently the Scale-Invariant Feature Transform (SIFT) [Lowe 1999, Lowe 2004] has become a popular method of local feature detection because the resultant features are robust to many commonly experienced distortions: Scale changes, rotation, and other affine transformations.

Briefly, SIFT features are located at **positions** within the scale-space of an image where the Difference of Gaussian function is maximized; each feature has a **dominant orientation** (computed from peaks in a histogram of local greyscale gradient orientations), is detected at a known **scale**, and is augmented with a **descriptor** – essentially a 2-dimensional array of local greyscale gradient orientation histograms centred on the detected feature. The descriptor is constructed relative to the keypoint location and dominant orientation thereby providing rotation and translation invariance. The use of histograms and weighted interpolation of contributing pixels provides partial invariance to other affine changes. SIFT features have successfully been used for purposes such as visual SLAM [Se, Lowe 2002].

### 1.5 Correspondances and Transformational Geometry in Panoramic Images

An epipolar geometry for panoramic cameras has been derived by [Svoboda, Pajdla 1998]. Let $\boldsymbol{p} = [\, x_1, y_1, \ldots \,]^T$, the location of a pixel in a raster image $\boldsymbol{I}_1$. Similarly, $\boldsymbol{q} = [\, x_2, y_2, \ldots \,]^T$, the location of a pixel in $\boldsymbol{I}_2$. Given a set of correct and accurate point correspondances between $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ it is possible to solve for a transformational matrix $\boldsymbol{T}$ such that:

$$\boldsymbol{q} = \boldsymbol{T}\boldsymbol{p}$$

.. using this geometry. However, this system is dependent on confidence in no less than 8 correspondances between $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$. In a high-resolution or perhaps off-line application such as image mosaicing this is an attractive proposition, since it does not matter that the discovery & verification of the correspondances is an arduous task (see [Fischler and Bolles 1981, Brown and Lowe 2002] for a suitable method).

## 2 Proposed Approach

In this paper we adopt an alternative approach, relating pairs of unwarped greyscale panoramic images using only an affine transformation – specifically, a translation in the x-axis. Unusually, this approximate relationship is computed early – in fact, immediately post feature detection. The affine transformation may be easily implemented using a pair of matrices:

$$\boldsymbol{q} = \boldsymbol{A} \times \boldsymbol{p} + \boldsymbol{B}$$

.. where $\boldsymbol{A} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$ $\boldsymbol{B} = \begin{vmatrix} b_1 \\ b_2 \end{vmatrix}$

Given these coefficients $\boldsymbol{A}$ has no effect and may be eliminated; and $b_2$ is always 0 since we are only interested in translation in the x-axis. Since the x-axis represents orientation a circular number system should be used:

$b_1 = n \,\%\, w$ (w is the width of the images)

Thus the problem is reduced to 1 dimension.

Viewed objects move away from the median point of the y-axis as the robot approaches, and toward it as the robot recedes; given that the view is panoramic features 180 degrees disparate will be moving in opposite directions in the vertical axis. To accomodate this effect, - and error or variation in the localisation of corresponding SIFT features - the method permits keypoints to be matched if located within a Euclidean radial tolerance $r$ of the translated pixel position. Further, it is required that corresponding keypoints differ in scale by less than 1 Octave [Lowe 2004] and have dominant

orientations with angular distance < $\theta$. In our experiments, the unwarped image had dimensions 651x138 pixels and $r$ = 25; $s$ = 1 (octave) and $\theta$ = 2Π / 36 radians.

To quickly discover n an exhaustive search is initiated for all values of n between 0 and (w-1). The search is coarsely discretized into 36 uniformly distributed values from this range:

$$\max(f(n)),\ n = [0,35]$$

$a$ is a scalar being the number of keypoints detected in $\boldsymbol{I}_1$.
$b$ is a scalar being the number of keypoints detected in $\boldsymbol{I}_2$.
$v_j$ is a vector of keypoints in $\boldsymbol{I}_j$. $V_{1,m}$ is a keypoint from $\boldsymbol{I}_1$.

$$f(n) = \sum_{k=0}^{a-1} \delta(v_{1,k})$$

i.e. the sum over all keypoints from $\boldsymbol{I}_1$ of $\delta(v_{1,k})$:

$$\delta(v_{1,k}) = \begin{cases} 1, \exists(v_{2,i}), i = [0,b] \ni \begin{pmatrix} \alpha(v_{1,k}, v_{2,i}) \wedge \\ \beta(v_{1,k}, v_{2,i}) \wedge \\ \chi(v_{1,k}, v_{2,i}) \end{pmatrix} \\ 0 \end{cases}$$

In other words, $\delta(v_{1,k})$ is 1 iff there exists a keypoint from $\boldsymbol{I}_2$ (such that) conditions α,β, and χ are met.

$$\alpha(v_{1,k}, v_{2,i}) = Euc.Dist.(v_{1,k}, v_{2,i}) < r$$

$$\beta(v_{1,k}, v_{2,i}) = Angular.Dist.(v_{1,k}, v_{2,i}) < \theta$$

$$\chi(v_{1,k}, v_{2,i}) = Scale.Dist.(v_{1,k}, v_{2,i}) < s$$

## 2.1 Horizon Determination

The horizon is likely to contain strong features due to the high contrast between ground and sky; furthermore, features on the horizon are likely to be distant and therefore little affected by robot movement over small distances. Nearby ground, in contrast, is likely to contain strong texture and features across many scales that will quickly become obscured or distorted beyond recognition.

In our work the horizon is crudely determined by calculating the image row of the greyscale image at which the first derivative in the columnar dimension is maximised, and adding a fixed offset. The fixed offset is intended to permit the generation of gross-scale keypoints located below the horizon to which the horizon pixels contribute significantly, and to handle variations in horizon level within a single panoramic image caused by uneven terrain. SIFT features are computed everywhere in the image *above* the offset horizon, but not below it.

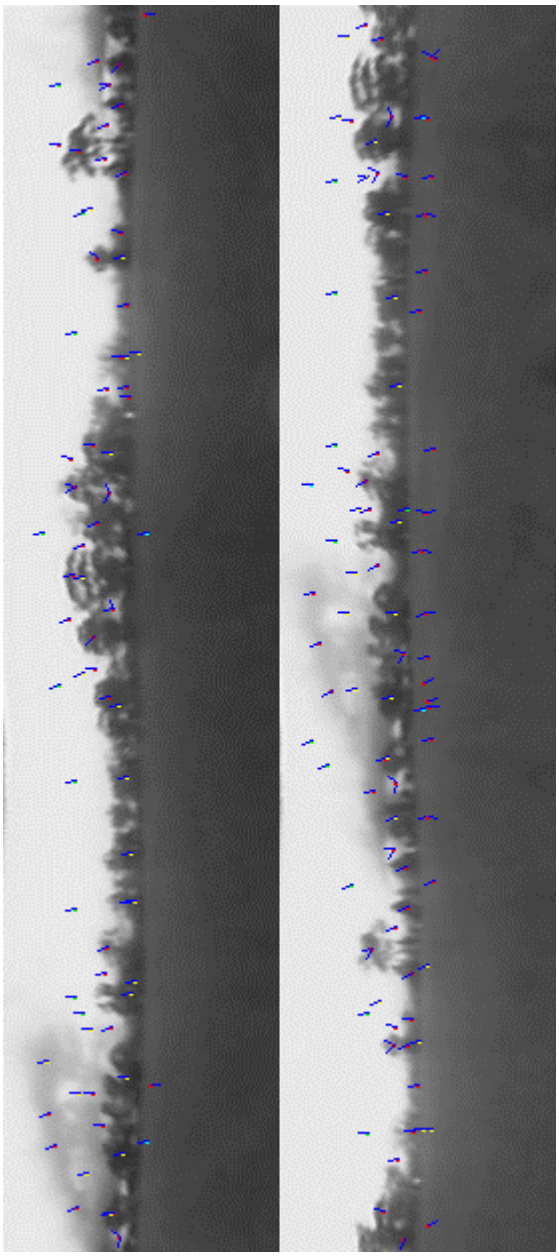

Figure 2: Keypoints detected on or around the horizon in 2 panoramic images. Tail indicates dominant orientation; colour indicates octave. Note that these images have been rotated 90 degrees anticlockwise to better fit the page.

## 2.2 Matching Feature Descriptors

Traditionally, SIFT features are matched by descriptor comparison. The proposed approach is no different, but the recovered translation $n$ is used to filter descriptor comparisons. For example, let $\boldsymbol{p'} = [\, x_1, y_1 \,]^T$, the location of a keypoint in $\boldsymbol{I}_1$. $\boldsymbol{q'} = \boldsymbol{p'} + \boldsymbol{B}$, the predicted location of the corresponding keypoint in $\boldsymbol{I}_2$. A circle of diameter $d$ is centred on $\boldsymbol{q'}$; any keypoints in $\boldsymbol{I}_2$ located within the circle

are candidate matches.

The purpose of the circular window is to permit correct feature matching despite changes in distance and elevation caused by camera egomotion. Simultaneously it caters for variations in keypoint location in the x-axis caused by parallax effects from egomotion and error in *n*. (These are mentioned only to enumerate expected sources of error).

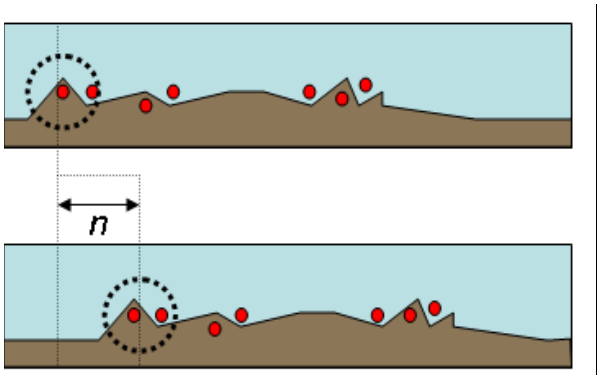


Figure 3: Two panoramic images, SIFT keypoints shown in red. A match is required for the circled keypoint in the top image. Matches are sought in the lower image but only in the circle now offset by *n* pixels. Two candidate matches are found; the correct match will be determined by descriptor comparison.

In our experiments, each keypoint descriptor has 128 dimensions (based on [Lowe 2004]). Generally, the most likely keypoint match is taken to be that which minimizes Euclidean distance between 2 descriptors over all 128 dimensions. There are a number of mechanisms by which descriptor comparison (effectively a "nearest neighbours" problem) may be accelerated, most notably using data structures optimized for fast descriptor comparison in all dimensions. [Beis and Lowe 1997] suggest using K-D Trees (a generalization of the binary tree for K dimensions); other authors have utilized the set intersection of lists of descriptors sorted in all dimensions independently (the "Projections" approach [Nene and Nayar 1997]). Regardless, filtering descriptor comparisons reduces matching expense. The proposed algorithm reduces the number of descriptor comparisons firstly by not generating SIFT features below the horizon and secondly by using SIFT feature location, scale and dominant orientation to predict corresponding feature location.

Imposition of suitable geometric constraints is also likely to improve correspondance consistency: It is possible that the closest feature in terms of descriptor similarity is not the "correct" correspondance. In these cases, it is probable that the correct match evinces only marginally greater descriptor dissimilarity. [Lowe 2004] describes precisely this situation and suggests discarding any keypoints within an image that posess similar descriptors. Alternatively, as in this paper, geometric constraints can be used to resolve conflicts of this sort.

## 3 Experiment

Experiments were performed to:

- Quantify reduction in keypoint comparisons due to use of approximate transform relating pairs of images
- Quantify effect of approximate transform filtering on accuracy of keypoint matching
- Quantify effect of approximate transform filtering on geometric consistency of keypoint matches
- Demonstrate the use of SIFT features located on or around the horizon, and show that these are effective for image registration in open spaces
- Quantify the stability of horizon features in open spaces – in particular, image registration reliability as inter-image distance increases

To address these aims 3 sets of 5 panoramic images were captured, each in a grossly different "open" location. Set 1 was captured in a small park. Set 2 was captured in a cricket oval at Monash University. Set 3 shows views from a terraced hilltop in a rural area.

Each set has at least 180 degrees' horizon (sky/ground border) at distances of more than 20 metres. Maximum feature distance is some 150-200m: Minimum 10-20m. Within each set, image capture locations are separated by between 10 and 100m horizontally. In set 3 there is a verticaly disparity of some 6m between some images due to the terraced terrain.

Two experimental conditions were implemented: In **condition A** (control) features were matched using only Euclidean descriptor distance, filtering possible matches using scale & dominant orientation. In **condition B** matching was identical but potential matches were filtered by precomputation of approximate transforms between image pairs in the manner previously desribed (see section 2). Within each set all possible pairwise comparisons were computed for both conditions i.e. 2 5x4=20 comparisons in condition A, 20 in condition B = 40 image comparisons per set.

Correspondance validity was assessed by manually estimating the change in camera orientation in the ground plane using a single feature for reference. In sets 2 this feature was the Sun. In sets 1 & 3 clouds obscured the sun and the relative bearing to the 1st column of the selected image was used as the reference point in these comparisons.

Just as the recovered transform was used to filter the search for correspondances in condition B, the manually estimated change in camera orientation was expressed as a translation in the x-axis $t_1$ and as a coefficient of a translation transform. This enabled us to define erroneous matches programmatically. Given images of width 651 pixels the maximum possible error in a circular system is 325.5 pixels. A small window of tolerance $\Omega$=32.5 pixels (10% of possible error) was permitted to account for errors in SIFT feature location and factors such as parallax. Note that this test is similar to that posed by Bearing-Only SLAM [Fitzgibbons and Nebot 2002].

The mean average translation in the x-axis of all correspondances ($t_2$) was compared to $t_1$ as a metric of global registration error. Error is shown in [Table 1,2] as a fraction of maximum error (10 $\Omega$). If error exceeded $\Omega$ the image was considered to be incorrectly matched.

For each correspondance generated between two images (valid or not) let $x_1$ represent the horizontal position of one keypoint of a correspondance pair in its image. Let $x_2$ represent the horizontal position of the corresponding keypoint in the other image. Let $\Delta x$ be the difference between $x_2$ and $x_1$. Geometric consistency was

evaluated by computing the standard deviation in $\Delta x$ for all correspondances relating a pair of images. (Since x-axis coordinates are in a circular number system the largest free span is moved to cover the 0/max point whilst calculating mean). The standard deviation in correspondances' keypoints' vertical positions was also computed.

Table 1: Results from 3 image sets in 2 experimental conditions; %F is percent of possible keypoint comparisons Filtered; Error is measured in terms of difference between average correspondance translation and manually determined ideal translation as a fraction of image width; if error exceeds 0.1 image is considered NOT correctly matched. Standard Deviation is measured in pixels within a set of correspondances:

| Image Set | Expt. Cond. | % F | Error | Correctly Matched | Std. Dev. x | Std. Dev. y |
|---|---|---|---|---|---|---|
| 1 | A | 42.8 | 0.12 | 10 / 20 | 66.1 | 7.2 |
| 1 | B | 97.9 | 0.01 | 20 / 20 | 10.8 | 4.2 |
| 2 | A | 47.5 | 0.39 | 7 / 20 | 89.6 | 10.3 |
| 2 | B | 98.7 | 0.02 | 19 / 20 | 11.6 | 5.6 |
| 3 | A | 49.9 | 0.37 | 4 / 20 | 95.5 | 12.0 |
| 3 | B | 98.6 | 0.04 | 18 / 20 | 14.1 | 5.2 |

Table 2: Results, aggregated:

| Image Set | Expt. Cond. | % F | Error | Correctly Matched | Std. Dev. x | Std. Dev. y |
|---|---|---|---|---|---|---|
| * | A | 46.7 | 0.29 | 21 / 60 | 83.8 | 9.8 |
| * | B | 98.4 | 0.02 | 57 / 60 | 12.1 | 5.0 |

## 4 Discussion

Overall results were unexpectedly good in all areas. Introducing an initial approximate transform relating pairs of images successfully filtered 98.4% of all keypoint comparisons, compared to only 46.7% filtered using only scale and dominant orientation constraints. Since computation of the approximate transform is comparatively cheap, the technique would yield significant acceleration to image registration & matching in suitable applications.

However, the approximate transform can only be found cheaply due to geometric constraints unique to the panoramic-camera-in-open-spaces scenario that effectively reduces search to 1 dimension. For example, when matching two perspective images from unknown environments an approximate translation transformation in x and y might be computed to filter keypoint matching. This would involve search in 2 dimensions, in other words from linear to polynomial in complexity.

It might be argued that it is common practice to apply geometric constraints after matching keypoints thereby eliminating inconsistent matches and improving matching accuracy. Certainly, matching performance in condition A would likely be greatly improved by such a measure, whereas condition B results would be largely unaffected since accuracy was already high. But the object of this experiment was to demonstrate that an approximate and cheaply computed transform could improve keypoint matching **prior** to geometric relation recovery, with the following benefits:

- subsequent fitting of an accurate transformation would be based on superior and more numerous correspondances.
- computation time would be significantly reduced by introduction of geometric constraints early in the matching process.

In both conditions, an accurate geometric transform could be subsequently applied to filter the selected correspondances.

It was expected that even a set of perfect correspondances would exhibit some variance due to apparent expansion, contraction & parallax of viewed objects from egomotion; the assumption remains that erroneous correspondances would increase this variance. The closer objects in the test images did change apparent shape and position considerably during the experiment. In theory, recovery of accurate optical flow during camera movement would appear to show local areas of expansion and contraction, in addition to translation caused by camera rotation.

The third image set was considered by the authors to be the hardest registration task due to the massive variation between images, much of which was caused by movement of the camera between terraces. However, much of this variation occurs in the foreground – the features & overall shape of the horizon are in fact relatively stable. It is encouraging that the approach works in both semi-natural and semi-structured outdoor environments.

Finally, as expected, the more accurate correspondances are strongly auto-correlated.

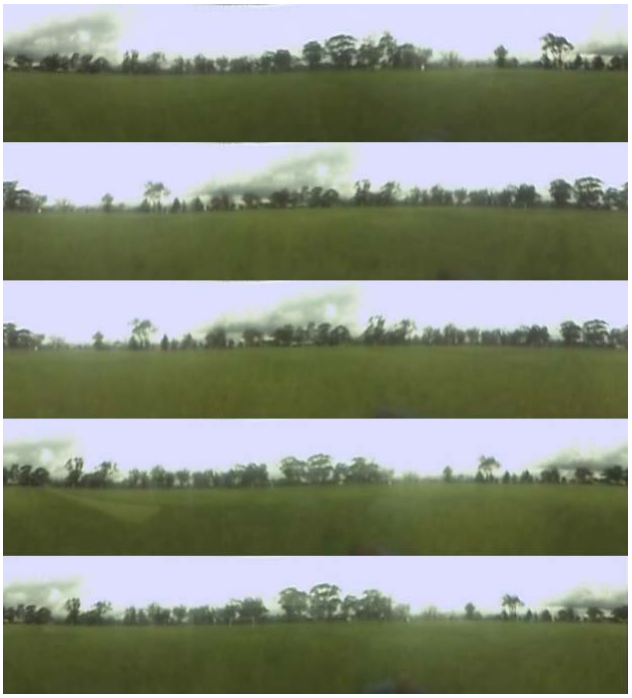

Figure 4: 5 Panoramic images ("Burnley", set 1). These images were taken at locations up to 30m apart; objects on the horizon are 20 to 70m from the camera. The features are predominantly natural (tree silhouettes). Notice that some features on the horizon vary in size by as much as 100% between images. The authors considered this set to be the least challenging of the three.

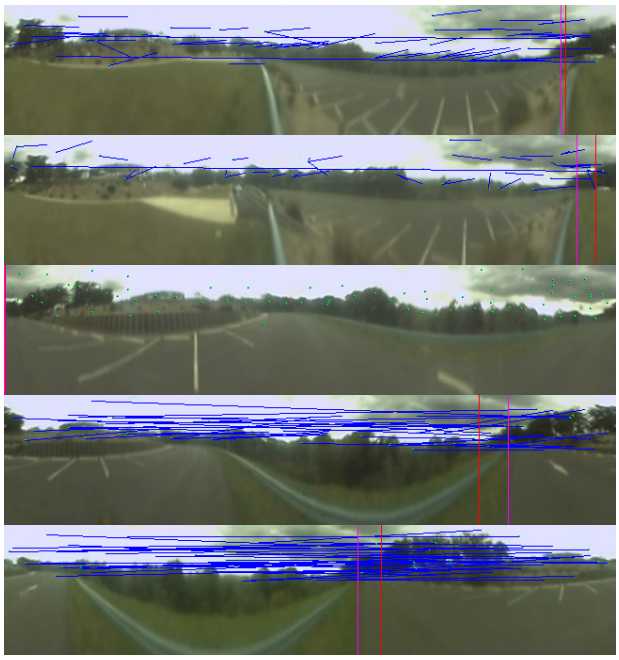

Figure 5: Successful matching of 5 images from the saddle at Mt. Dandenong (set 3). The land here is terraced; between images 2 & 3 the camera moves between terraces, a vertical transition of some 6m. Over the 5 shots the camera has moved approximately 60m horizontally. Objects on the horizon are between 5 and 100m distant. The most distinct features are again tree silhouettes, but part of the skyline is obscured by artificial structures (buildings). Notice that the foreground has changed significantly between shots, but the horizon is still recognisable.

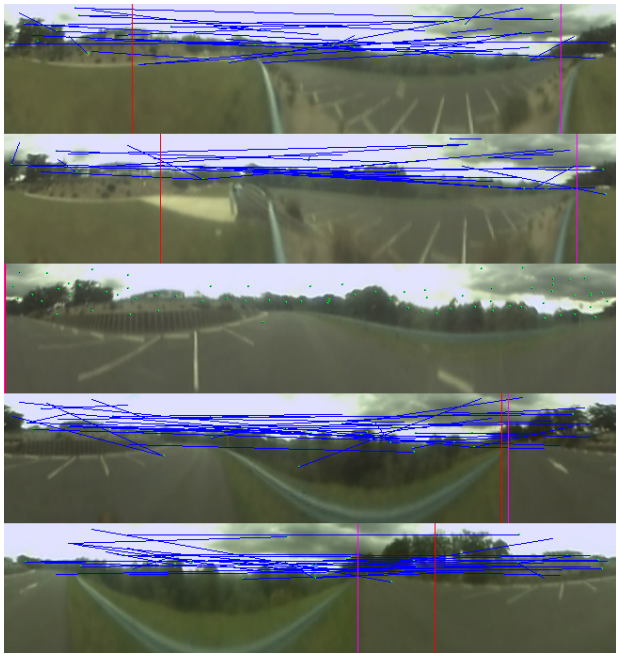

Figure 6: Matching using keypoint descriptors only. The central image is the baseline one, to which the others are matched. The first two images have not been matched; neither has the last. The 4th image is similar to the 3rd and the majority of the correspondances are correct.

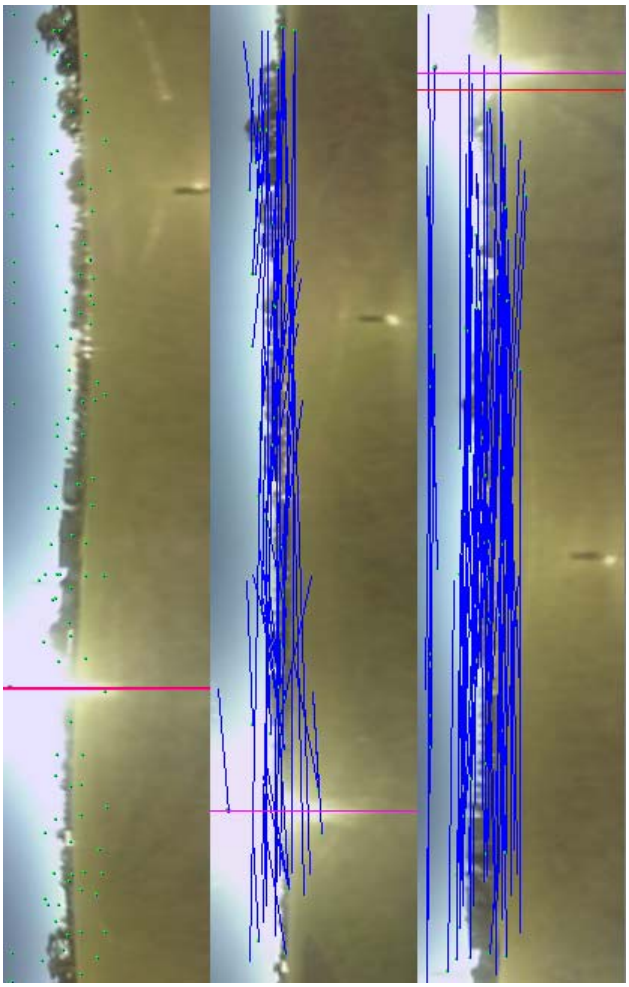

Figure 7: Images from set 2 matched using approximate translation filtering of keypoint descriptors. The rotation implied by averaging the translation of all correspondances is so accurate in the 2nd image that the recovered and correct lines overlap.

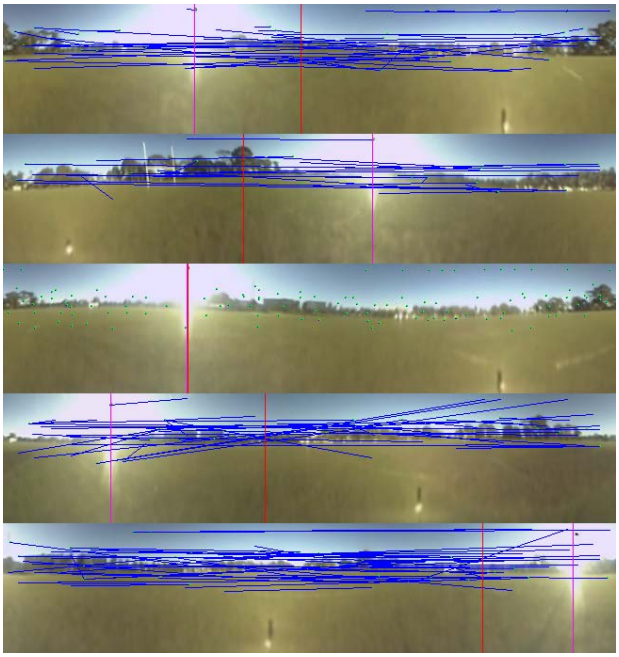

Figure 8: Images (from set 2) not matched (!) using only

keypoint descriptors (horizontal distance between red and pink lines greater than 10% of image width).

## 5 Conclusions

This paper shows that the horizon contains strong and robust SIFT features in open natural environments. In general horizon features are likely to remain visible and largely unaffected by small changes in viewing position, due to their distance. For these reasons horizon features are well-suited to topological visual localisation. Overall, the approach detailed in this paper could be considered a very coarse Bearing-Only SLAM [Fitzgibbons and Nebot 2002].

The Sun in particular was found to be consistently identified and matched in all images. In the set 1 clouds caused some matching errors by changing shape, position and intensity in the time that elapsed between shots.

Although the horizon has been used consciously in UAV navigation [Kweon and Kanade 1991] it has appeared incidentally in many other papers [Lambrinos, M¨oller 1999, Brown and Lowe 2003, Cohn 1997, Chahl and Srinivasan 2000, Yagi, Imai 2003].

It would be interesting to examine the effect of deliberate horizon viewpoint-seeking behaviours on terrestrial robot navigation in natural conditions but such experiments would probably be impractical without very effective local navigation and path-following algorithms to handle transitions between open spaces.

This paper has shown that the spatial distribution of local features around the horizon permits faster image comparison via. the recovery of approximate transformational geometry. This would not be possible (at least, less efficient) in totally unstructured images: It is dependent on the recovery of the horizon in panoramic images. The approach can also be applied (to a lesser extent) to perspective images of planar objects – perhaps initially searching for an approximate translation in 2-dimensions and then performing spatially-bounded keypoint matching; 4 correspondances are sufficient to recover a projective transformation.

Results reported in this paper also demonstrate that correspondance geometric consistency and accuracy are improved by prior determination of an approximate geometric relationship between pairs of images. Generation of more numerous and accurate initial correspondances accelerates & assists subsequent recovery of accurate geometric relationships between images (e.g. using RANSAC [Fischler and Bolles 1981] probability of success would be increased).

## References


[Lambrinos, M¨oller 1999]. Lambrinos, D., et al., *A Mobile Robot Employing Insect Strategies for Navigation.* Robotics and Autonomous Systems, special issue: Biomimetic Robots (to appear), 1999.

[Weber, Venkatesh 1997]. Weber, K., S. Venkatesh, and M.V. Srinivasan, *Insect Inspired Behaviours for the Autonomous Control of Mobile Robots.* From Living Eyes to Seeing Machines, 1997. **Oxford University Press**: p. 226-248.

[Kweon and Kanade 1991]. Kweon, I.S. and T. Kanade, *Extracting Topographic Features for Outdoor Mobile Robots.* International Conference on Autonomous Robots and Agents (ICRA), 1991.

[Brown and Lowe 2003]. Brown, M. and D.G. Lowe, *Recognising Panoramas.* International Conference on Computer Vision (ICCV), 2003.

[Cohn 1997]. Cohn, A.G., *Qualitative Spatial Representation and Reasoning Techniques.* Fundamenta Informaticae, 1997. **46**(1-2): p. 1-29.

[Prasser, Milford 2005]. Prasser, D., M. Milford, and G. Wyeth, *Outdoor Simultaneous Localisation and Mapping using RatSLAM.* International Conference on Field and Service Robots, 2005.

[Carson, Belongie 1997]. Carson, C., et al., *Region-Based Image Querying.* Computer Vision and Pattern Recognition (CVPR), 1997.

[Fitzgibbons and Nebot 2002]. Fitzgibbons, T. and E. Nebot, *Bearing-Only SLAM Using Colour-Based Feature Tracking.* Proc. 2002 Australian Conference on Robotics and Automation, 2002.

[Canny 1986]. Canny, J., *A Computational Approach to Edge Detection.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986. **8**: p. 679--698.

[Smith and Brady 1995]. Smith, S.M. and J.M. Brady, *SUSAN - A new approach to low level image processing.* ICPR96, 1995.

[Nist´er, Naroditsky 2004]. Nist´er, D., O. Naroditsky, and J. Bergen, *Visual Odometry.* IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), 2004.

[Rosenfeld and Pfaltz 1966]. Rosenfeld, A. and J.L. Pfaltz, *Sequential Operations In Digital Picture Processing.* Journal of the Association for Computing Machinery, 1966. **13**(4): p. 471-494.

[Barenholtz, Cohen 2003]. Barenholtz, E., et al., *Detection of change in shape: an advantage for concavities.* Cognition, 2003. **89**(1-9).

[Hernandez and Barner 2000]. Hernandez, S.E. and K.E. Barner, *Tactile Imaging Using Watershed-based Image Segmentation.* ASSETS'00, 2000.

[Lowe 1999]. Lowe, D.G., *Object Recognition from Local Scale-Invariant Features.* International Conference on Computer Vision (ICCV), 1999.

[Lowe 2004]. Lowe, D.G., *Distinctive Image Features from Scale-Invariant Keypoints.* International Journal of Computer Vision (IJCV), 2004.

[Se, Lowe 2002]. Se, S., D. Lowe, and J. Little, *Global Localization using Distinctive Visual Features.* Proceedings of the 2002 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, 2002.

[Svoboda, Pajdla 1998]. Svoboda, T., T. Pajdla, and V. Hlavac, *Epipolar Geometry for Panoramic Cameras.* European Conference on Computer Vision (ECCV), 1998.

[Fischler and Bolles 1981]. Fischler, M.A. and R.C. Bolles, *Random Sample Consensus: A Paradigm for Model Fitting with Apphcatlons to Image Analysis and Automated Cartography.* Graphics and Image Processing, 1981. **24**(6): p. 381.

[Brown and Lowe 2002]. Brown, M. and D. Lowe, *Invariant Features from Interest Point Groups.* British

Machine Vision Conference, 2002.
[Beis and Lowe 1997]. Beis, J.S. and D.G. Lowe, *Shape Indexing Using Approximate Nearest-Neighbour Search in High-Dimensional Spaces*. Computer Vision and Pattern Recognition (CVPR), 1997.
[Nene and Nayar 1997]. Nene, S.A. and S.K. Nayar, *A Simple Algorithm for Nearest Neighbour Search in High Dimensions*. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 1997. **19**(9): p. 989-1003.
[Chahl and Srinivasan 2000]. Chahl, J.S. and M.V. Srinivasan, *Panoramic vision system for imaging, ranging and navigation in three dimensions*. IEEE Workshop on Omnidirectional Vision (OMNIVIS'00), 2000: p. 104.
[Yagi, Imai 2003]. Yagi, Y., K. Imai, and M. Yachida, *Iconic Memory-Based Omnidirectional Route Panorama Navigation.* International Conference on Robotics and Automation (ICRA), 2003.