Visual compass

Frédéric Labrosse

Department of Computer Science
University of Wales, Aberystwyth
Aberystwyth SY23 3DB, United Kingdom

e-mail: `ffl@aber.ac.uk`

Published in: *Proceedings of Towards Autonomous Robotic Systems*, pages 85–92 (2006), University of Essex, Colchester, United Kingdom, `http://www.taros.org.uk`.

Visual compass

Frédéric Labrosse
Department of Computer Science
University of Wales
Aberystwyth, Ceredigion
SY23 3DB
ffl@aber.ac.uk

Abstract

We present in this paper a method to incrementally estimate the heading of a robot using panoramic images grabbed while the robot moves in its environment. We present results that show a performance similar to a magnetic compass. In particular, the heading measured by the visual compass does not drift as normal odometry does. We discuss the performance of the system as well as its limitations. We also briefly give examples of applications where the visual compass could be used and constitute an improvement.

1 Introduction

One of the very important tasks of mobile robot navigation is to somehow learn where the robot is, i.e. what the current position and/or orientation of the robot is. Some methods only provide the position *or* the orientation, the other being inferred by integration.

One of the approaches uses a system, made of external devices and usually a device mounted on the robot, that allows the precise localisation of the robot. The Global Positioning System (GPS) is such a system. Sometimes, the system doesn't provide an absolute position/orientation but only one that is relative to the desired path of the robot for example. This is the case of visible (for example reflective), or otherwise (for example induction loops), guides. That approach suffers from the fact that the environment of the robot must be equipped with such a system, which is not always possible, is often expensive and sometimes constraining. For example, GPS needs a set of satellites orbiting above the robot in direct view of the robot, which thus cannot be indoors or in any closed environment. Guides need to be fitted in the environment and the robot is restricted to the area covered by the guides.

Another approach uses some map of the environment as well as a way of finding where on the map the robot currently is. This has the advantage of not needing any external apparatus, although the localisation often involves external beacons sometimes fitted to the environment. The approach however needs a precise and accurate knowledge of the environment, knowledge not available when exploring the environment. The localisation procedure often involves iterative mechanisms, especially when the robot's position is initially unknown.

A final approach only uses sensors on board the robot and their interaction with the environment of the robot. Some sensors provide an immediate position and/or orientation that is either absolute or relative to the environment. This is for example sensors such as altimeters, range finders (laser, sonar), and compasses. Other sensors provide a position and/or orientation that is relative to the previous position/orientation of the robot. The simplest sensors simply measure the rotation of the wheels of the robot (in the case of a wheeled robot) and work out the displacement of the robot by odometry. More sophisticated methods use accelerometers. These sensors all suffer from limitations. For example, a compass needs a well known and predictable magnetic field, which does not always exist (for example on Mars). Odometers are typically unreliable as soon as there is some slippage of the wheels, which almost always happens to some extend. Lasers need surfaces having "good" reflective properties and only work if the atmosphere surrounding the robot is clean enough.

In this paper, we propose to use images of the environment gathered by the robot while it moves to estimate the heading (orientation) of the robot relative to either an absolute reference (if available) or a starting orientation. In order to provide a full view of the robot's surroundings, we use panoramic images that are un-rotated to match as best as possible the previous image in the sequence. The amount of un-rotation needed corresponds to the rotation the robot underwent between the two images.

The work reported in this paper is one of the many projects undertaken by the author and colleagues on mobile robot navigation and localisation and map building using raw images only (Neal and Labrosse, 2004b, Mitchell and Labrosse, 2004), by opposition to extracting features from the images and using the features, as for example in (Winters and Santos-Victor, 2002, Regini et al., 2002).

Section 2 discusses panoramic images and their rotation invariance properties. Section 3 discusses the basis of our method: distance between images. Section 4 describes the presented work while Sections 5 and 6 present the experimental setup and the results of some experiments. Finally, a conclusion and some discussion of the proposed method is presented in Section 7.

2 Panoramic images and their rotation invariance

In this work we use panoramic images of the environment[1], Figure 1. Panoramic images are appealing

Figure 1: A typical panoramic image

as they capture in one image everything that is visible from the current view point. This makes them ideal for localisation and map building tasks. Moreover, there is some evidence that insects use wide-angle "images" of their environment to localise themselves with respect to a visual map of their environment. Furthermore, it seems that images are interpreted as a whole, without requiring any feature extraction (Collett, 1996, Judd and Collett, 1998).

Another advantage of panoramic images, often used in the literature (Pajdla and Hlaváč, 1999, Jogan and Leonardis, 2001, Neal and Labrosse, 2004b), is that they explicitly contain information about the orientation of the camera. All images taken from the same position but with different orientations can indeed be transformed one into another. This can be performed efficiently by just shifting columns of the images. This property has been used to extract rotation invariant representations of places, for example the *Zero Phase Representation* in (Pajdla and Hlaváč, 1999). However, this assumes that the robot can indeed rotate around the optical axis of the panoramic camera, which is often not the case, at least not in uncontrolled environments. The robot used in these experiments (Section 5) is not circular and turns using skid steering. Combined with the fact that we usually use it outdoors, there is no guarantee that the robot rotates around the optical axis of the camera or that it is even possible to do that given the uneven terrain.

[1]Section 5 describes the hardware and software setup to produce such images.

3 Distance in the image space

Another interesting property of panoramic images is that they smoothly change from places to nearby places. This property is fundamental to all the purely image-based methods.

Let us formalise this. Grey-level $n \times m$ pixels images can be viewed as points in an $n \times m$ dimensional space that thus can represent all possible images. This high-dimensional space is called the *image space*. With the case of colour images, the image space has a dimension of $n \times m \times 3$ for three-components colours (we used RGB images in this case)[2]. When a robot moves in its environment, all the images grabbed are organised on a 1D curve embedded in the image space. In more general cases, particularly if the robot has a trajectory that cannot be assimilated to a 1D trajectory, the curve becomes a manifold of dimensionality dependent on the actual trajectory of the robot and possibly factors such as lighting conditions.

Geometrical properties of these manifolds have been studied in (Lu, 1998, Lu et al., 1998). In particular, it has been shown, and nicely exemplified in (Bichsel and Pentland, 1994), that the manifolds are highly curved. The consequence is that the Euclidean distance is not necessarily a good measure of the similarity between images (the image midway, using Euclidean distance, between two images is *not* the image midway between the two corresponding positions of the robot). Figure 2 shows the distance between

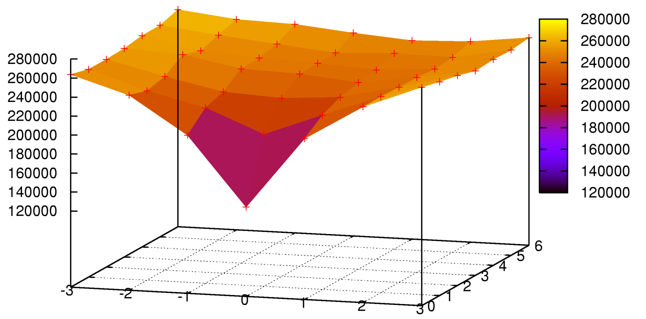


Figure 2: Distance between several images and a target image

an image taken in our lab close to one side of a pen and images taken from different places in the pen. Figure 7 shows such images. More can be found in (Mitchell and Labrosse, 2004). As can be seen, there is a clear relationship between Euclidean distance between points in the image space (in other words between images) and distance in Cartesian space. This has also been

[2]Note that the RGB colour space is not the best colour space but is enough for this work. A better colour space would be the CIE L*a*b* colour space as it is perceptually linear, a property that is very important when it comes to measuring distances between colours (Fairchild, 1998, Labrosse and Willis, 2001).

used in (Neal and Labrosse, 2004a) to build topological networks of places.

4 The visual compass

Having shown that panoramic images contain information about the orientation of the robot and that it makes sense to measure the Euclidean distance between images (points in the image space), we show how we can extract the heading of the robot.

The robot regularly grabs panoramic images while it moves. Each pair of successive images is used as follows. The current image is compared to the previous one and un-rotated so that it matches as best as possible the previous image (lowest distance in the image space). The un-rotation is performed by shifting columns of the image and corresponds to rotating to robot in the opposite direction of its real rotation. The column shift giving the lowest distance between the two images is thus related to the rotation of the robot. Because we use images having one column per degree of rotation, the shift is actually equal to the angle of rotation. By keeping track of the relative shifts (rotations of the robot) between consecutive images, the current heading of the robot can be integrated[3].

Figure 3 shows the pseudo-code of the algorithm.

```
main
  currentHeading = 0    // Or any reference.
  grab(previousImage)
  grab(currentImage)
  while not finished
    shift = bestMatch(previousImage,
                      currentImage)
    currentHeading += shift
    previousImage = currentImage
    grab(currentImage)
  endwhile
endmain

function bestMatch(image1, image2)
  bestShift = 0
  bestDist = infinity
  forall shift in 0..359
    shiftImage(image2, shift)
    dist = distance(image1, image2)
    if dist < bestDist
      bestDist = dist
      bestShift = shift
    endif
  endfor
  return bestShift
endfunction
```

Figure 3: Pseudo-code of the algorithm for the visual compass

5 Experimental setup

We used an all-terrain, four-wheel drive, skid steering, Pioneer2 AT robot. It is equipped with an omni-directional camera made as follows. A "normal" camera pointed upwards looks at a hyperbolic mirror[4] which gives circular images, Figure 4.

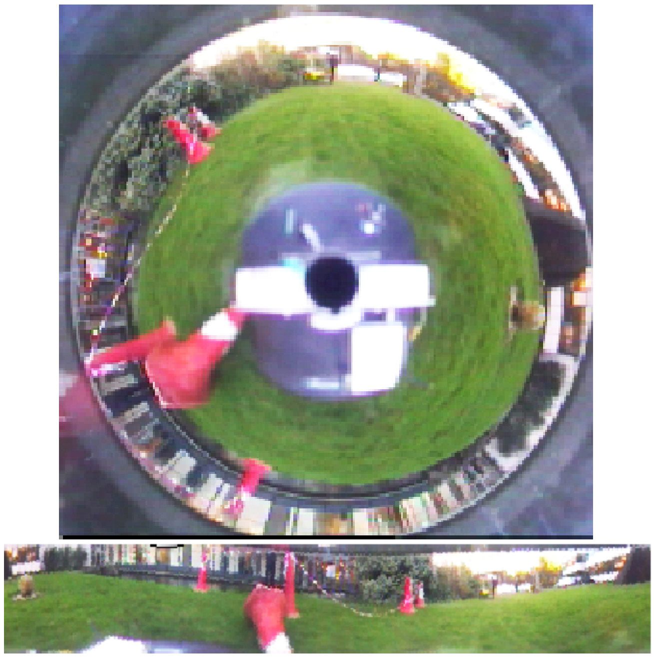


Figure 4: A panoramic image as grabbed and unwrapped

Images are then unwrapped, leaving out the pixels not corresponding to the environment. The unwrapping is done by scanning a line from the centre of the image around the image. 360 lines are used, producing on column in the panoramic images per degree of angle. Pixels of the unwrapped image can be taken as the nearest neighbour of the original image (not smooth version) or using bilinear interpolation between the four nearest neighbours in the original image (smooth version).

It is important to note that although the images could be unwrapped to produce undistorted panoramic images, we do not do that. The reasons are that the perspex tube attaching the mirror to the camera is not perfect, the process would need an accurate calibration and, more importantly, the image-based approach allows us not to use calibrated procedures.

Figure 5 shows the robot with the camera on top of it. Figure 6 shows a close-up of the omni-directional camera on which both the camera (bottom) and the mirror (top) are visible.

Visible above the omni-directional camera on Figure 5

[3] The heading is only relative to a reference, usually the heading of the robot when it started. However, if the heading at the beginning is known, then the current heading can be absolute, Section 5.

[4] The mirror comes from Neovision s.r.o., Czech Republic.

Figure 5: The mobile robot

Figure 6: Close-up of the omni-directional camera

is a small magnetic compass[5] used to evaluate our csystem. The compass is at the end of a 30cm long plastic tube above the mirror so that it is as far away as possible from any magnetic source on the robot (within reasonable limits). The compass has been accurately calibrated, both hardware and software driver, using a third traditional compass. However, we did notice that both indoors and outdoors, the compass was providing headings that were not necessarily predictable. Indoors, this was due to a large amount of metallic and electronic components (metallic mesh, trunking in the floor, CRT displays, etc.) as well as, to some extent, irregularities of the floor. Outdoors, the variations were mainly due to the very uneven terrain the robot was driving on (we did notice variations of up to almost 20° when the robot was not turning, Figure 11). Nevertheless, the compass provides a non-drifting heading.

Images are grabbed on the robot, unwrapped on the robot's computer and transmitted over TCP sockets over a wireless Ethernet connection to a laptop doing the remaining of the processing (the algorithm on Figure 3). The wireless Ethernet can transmit in one way at speeds of up to 1.5Mbps, which allows the transmission of about 1.2 frames per second. The compass is also read by the on-board computer and the values are also sent over TCP sockets over the same wireless connection.

[5]The model used is a CMPS03 from Devantech Ltd, England.

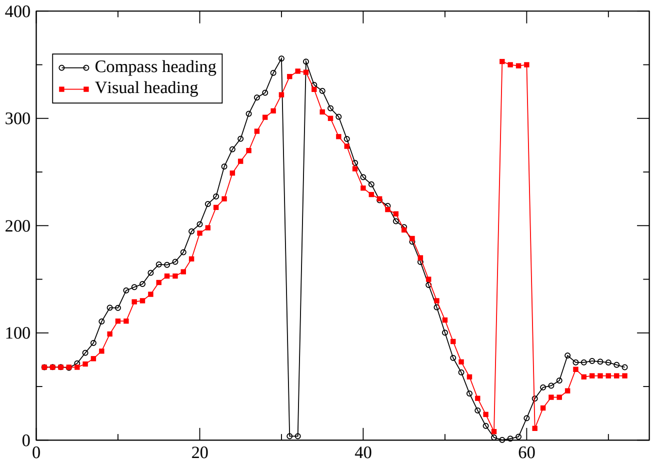
6 Results

We present four experiments, two indoors, two outdoors, with smooth images and not smooth images, Section 5. In each case, the visual heading is initialised with the value returned by the magnetic compass so that the two headings can be compared.

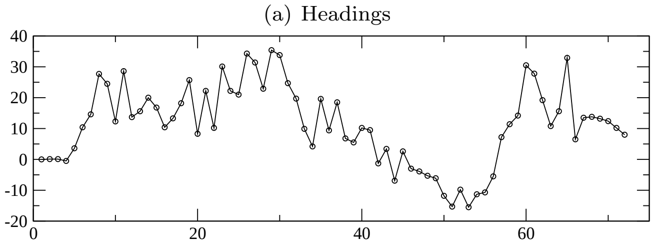
The indoors environment was a pen in our lab closed on three of its sides about 5m by 5m in area. The inside of the pen was free from obstacles but the lab was not arranged in any way for the experiments. The outdoors environment was a grassy area of about 20m by 20m delimited on three sides by a road, a path and bushes and containing a few trees in the middle, which implies that not everything is visible from everywhere. Moreover, cars and people where moving by the area during the experiments. The two environments present very different backgrounds: one is almost random clutter (the lab) while the other, though still presenting variations, is a lot more constant and smoothly varying, at least in the images.

The indoors experiments were ran so that the heading at the beginning and at the end were the same, around

Figure 7: Some images taken during the experiment in our lab



(a) Headings



(b) Errors

Figure 8: Results for the indoors experiment with smooth images

68° as returned by the magnetic compass.

6.1 *Indoors with smooth images*

Figure 7 shows typical images grabbed during the experiment. Figure 8(a) shows the headings, both from the compass and visual, as the robot was moving (the index of the x axis is the frame index). As can be seen, the two compasses agree quite closely, apart from two places, a few frames, where the large error is due to the two headings being on different sides of the zero mark. Figure 8(b) shows the error, corrected for when the headings are on opposite sides of the zero mark.

The maximum error is of about 35°, which seems to be important. However, the magnetic compass is not perfect due to different factors, Section 5. We have observed error of up to 20° in the heading returned by the magnetic compass and it is thus difficult to establish to what extent the error we observe is due to the visual compass or the magnetic compass[6].

The graphs show an almost systematic error when the robot turns right (heading increasing). Because the error does not increase, this is not due to an error in the evaluation of the heading. We think that the error is mainly due to the slowness of the wireless connection. The problem is that the image and the compass reading cannot be both downloaded at the same time. In other words, they do not correspond to the same physical spatial configuration of the robot. This is further discussed in Section 7.

It is important to notice that the heading returned by the visual compass does not drift. In particular. the heading at the beginning of the trajectory was 68°, as returned by the magnetic compass, and was 60○ at the end of the trajectory, while the magnetic compass was giving 68°, so an error of 8○.

6.2 *Indoors with not smooth images*

This experiment is more or less the same as the one in Section 6.1. The trajectory of the robot is not exactly the same but is similar. The difference is that we use not smooth unwrapped images, Section 5. The images are definitely different at the pixel level but the difference is not visible enough when printed on paper at these resolutions.

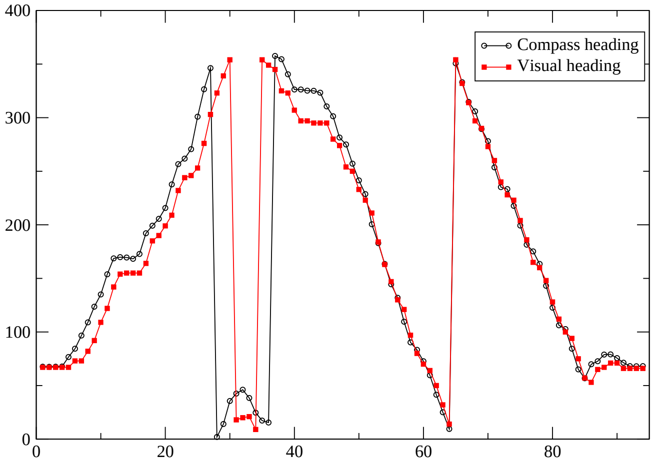
The graphs, Figures 9(a) and 9(b), show similar result as with smooth images. The maximum error is higher, about 50°, which seems to show that smooth images are better. This is however not very conclusive as this could be due to different trajectories, especially since the headings given by the visual compass was 67° at the beginning of the trajectory and 66° at the end (compared to respectively 67.6° and 68.1° given by the magnetic compass).

6.3 *Outdoors with smooth images*

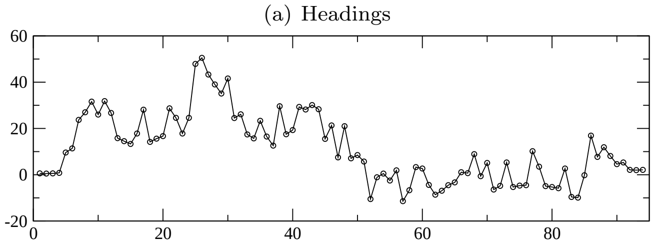
Figure 10 shows some typical views of the outdoors area we used for this experiment.

The graphs shown on Figure 11 show similar results, both in quality and in quantity. Two things need to be noted however, both about the performance of the magnetic compass. The first one is that the robot did not turn at all during the first 14 images. The visual compass accurately registers that. However, the magnetic compass shows non-negligible (up to almost 20°) but smooth variations. These are due to the fact that the ground had a large bump where the robot started, thus slowly changing the pitch and roll of the robot and compass. The second fact is present in different places

[6]We could have used a third compass as during the calibration procedure. However, for time and manipulation reasons, we did not perform this test.



(a) Headings



(b) Errors

Figure 9: Results for the indoors experiment with not smooth images

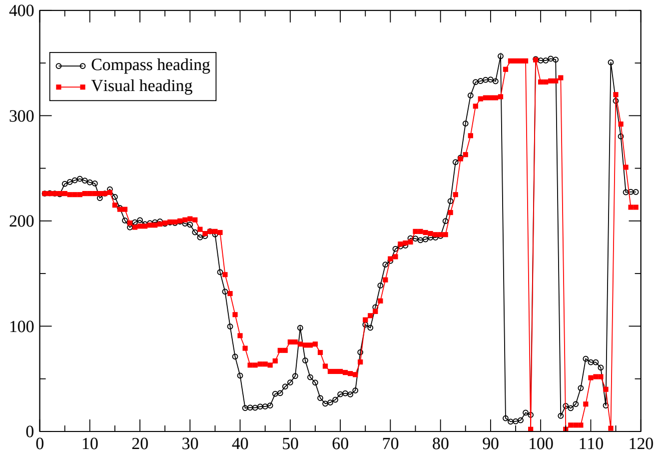
Figure 10: Some images taken during the outdoors experiment

on the graphs but particularly visible for image 52. A sudden change in pitch and roll of the robot gives a spurious magnetic compass reading.

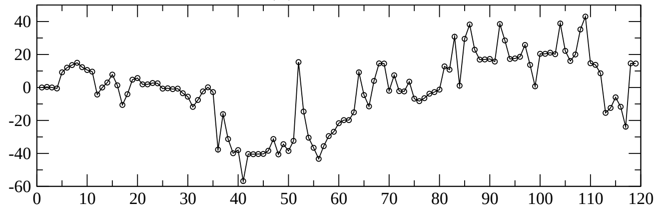
6.4 Outdoors with not smooth images

Figure 12 gives the graphs for the outdoors experiment with not smooth images. Again, they are both quantitatively and qualitatively similar to with smooth images, both maximum errors being of about 56°.

It is to be noted that this is only marginally worse that with the indoors experiments, the worse results probably being only due to the worse performance of the magnetic compass.

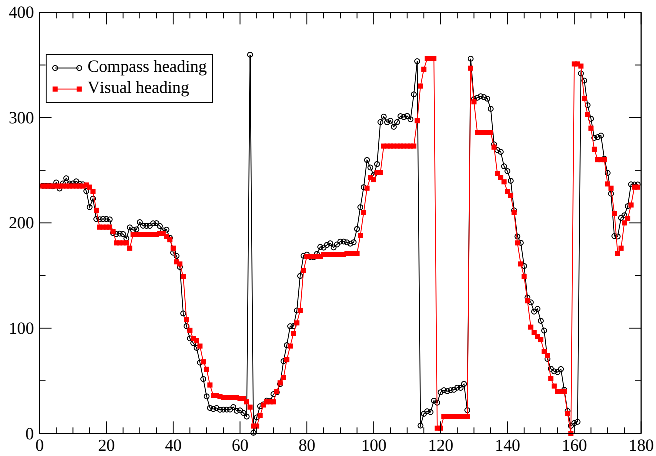


(a) Headings

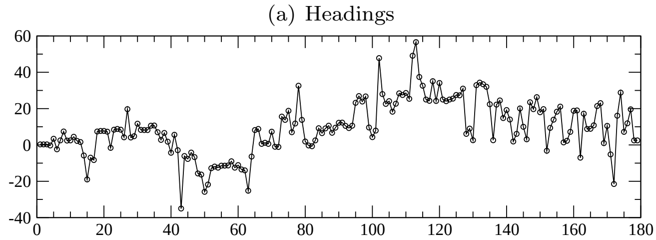


(b) Errors

Figure 11: Results for the outdoors experiment with smooth images



(a) Headings



(b) Errors

Figure 12: Results for the outdoors experiment with not smooth images

7 Conclusion and discussion

We have shown in this paper how to implement a visual compass. The compass works by estimating by how much the robot rotates between each pair of consecutive images grabbed while it is moving. The system is in its nature very similar to traditional odometry using wheel encoders. We have informally compared our visual compass to the heading information given by the robot's odometry, and needless to say, but we say it anyway, that the visual compass performs infinitely better.

We have shown both theoretically and practically that it makes sense to use the Euclidean distance between images in the image space, due to the fact that the consecutive images show a strong correlation.

The image-based approach obviously has limitations. The images need to have distinctive features. The presented experiments however show that both a random clutter and a more constant smoothly varying background perform similarly. However, the system would not work in places that are visually homogeneous or repetitive, as when surrounded by a chessboard[7].

The results we presented are promising. Although the maximum errors can go up to around 50°, the high errors are only peak values, the error being most often in the 20° error band. More importantly, it is clear that the visual compass does not drift. More experimentation should allow us to statistically characterise the performance of the visual compass.

We have also mentioned that most of the errors presented are due to the unreliable magnetic compass and to the fact that due to hardware limitations of our setup, it is not possible to measure both visual heading and magnetic heading at the same time (grabbing time to some extent, but mainly downloading time). A way around this would be to multi-thread the application and run it on the robot's computer.

Finally, we were able to process at speeds of about 1 frame per second, most of the time being taken by the downloading of the images. The code in its current state is not optimised at all. In particular, it always looks for the absolute minimum when determining the best shift, Figure 3. This can clearly be improved by only searching for a nearby local minimum of the distance function first in the same direction of shift as for the previous pair of images, on the grounds that the robot usually does not change its direction of rotation too quickly. Moreover, although the processing time is not negligible, it is anyway needed for higher level tasks performed by the robot, typically environment mapping and navigation (Neal and Labrosse, 2004b, Mitchell and Labrosse, 2004).

Despite all these possible improvements, we have shown a system that performs well enough to, for example, disambiguate a re-localisation process. As presented in (Neal and Labrosse, 2004b), we have noticed that blindly comparing in the image space panoramic images taken from very different positions in Cartesian space can lead to wrong matches. This system is an obvious solution to this problem.

[7] However, human beings and animals have similar limitations!

References

Bichsel, M. and Pentland, A. P. (1994). Human face recognition and the face image set's topology. *CVGIP: Image Understanding*, 59(2):254–261.

Collett, T. S. (1996). Insect navigation en route to the goal: multiple strategies for the use of landmarks. *Journal of Experimental Biology*, 199:227–235.

Fairchild, M. D. (1998). *Color Appearance Models*. Addison-Wesley.

Jogan, M. and Leonardis, A. (2001). Parametric eigenspace representations of panoramic images. In *Proceedings of the Workshop on Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing (NDT), International Conference on Advanced Robotics*, pages 31–36.

Judd, S. P. D. and Collett, T. S. (1998). Multiple stored views and landmark guidance in ants. *Nature*, 392:710–714.

Labrosse, F. and Willis, P. (2001). Towards continuous image representations. In *Proceedings of the International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, volume 1, pages 206–213, Plzen, Czech Republic.

Lu, H.-m. (1998). *Geometric Theory of Images*. PhD thesis, University of California, San Diego, CA, USA.

Lu, H.-m., Fainman, Y., and Hecht-Nielsen, R. (1998). Image manifolds. In *Proceedings of SPIE; Applications of Artificial Neural Networks in Image Processing III*, volume 3307, pages 52–63, San Jose, CA, USA.

Mitchell, T. and Labrosse, F. (2004). Visual homing: a purely appearance-based approach. In *Proceedings of Towards Autonomous Robotic Systems*, pages 101–108, University of Essex, Colchester, UK.

Neal, M. and Labrosse, F. (2004a). Generating appearance-based maps using an artificial immune system. In *Proceedings of the International Symposium on Robotics*, Paris, France.

Neal, M. and Labrosse, F. (2004b). Rotation-invariant appearance based maps for robot navigation using an artificial immune network algorithm. In *Proceedings of the Congress on Evolutionary Computation*, volume 1, pages 863–870, Portland, Oregon, USA.

Pajdla, T. and Hlaváč, V. (1999). Zero phase representation of panoramic images for image based localization. In *Proceedings of Computer Analysis of Images and Patterns*, pages 550–557, Ljubljana, Slovenia.

Regini, L., Tascini, G., and Zingaretti, P. (2002). Appearance-based robot navigation. In *Proceedings of the Workshop su agenti robotici, Associazione Italiana per l'Intelligenze Artificiale, VIII Convegno*, Siena, Italy.

Winters, N. and Santos-Victor, J. (2002). Information sampling for vision-based robot navigation. *Robotics and Autonomous Systems*, 41:145–159.