

# Information Sampling for vision-based robot navigation

Niall Winters<sup>a,b</sup>, José Santos-Victor<sup>b,\*</sup>

<sup>a</sup> Computer Vision and Robotics Group, Department of Computer Science, Trinity College, University of Dublin, Dublin 2, Ireland

<sup>b</sup> Instituto Superior Técnico, Instituto de Sistemas e Robótica, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal

## Abstract

This paper proposes a statistical, non-feature based, attention mechanism for a mobile robot, termed Information Sampling. The selected data may be a single pixel or a number scattered throughout an image. After ranking this data, we choose only the most discriminating to build a topological representation of the environment, obtained via Principal Component Analysis (PCA). Advantageously, using this approach, our robot gains the ability to make effective use of its perceptual capabilities and limited computational resources. Real world experimental results verify that vision-based navigation is possible using only a small number of discriminating image pixels.

© 2002 Elsevier Science B.V. All rights reserved.

*Keywords:* Omnidirectional vision; Vision-based navigation; Visual attention; Appearance-based methods

## 1. Introduction

This paper is concerned with the problem of vision-based navigation in structured environments. Vision generates a substantial amount of data and processing it can lead to a large computational load on a mobile robot. Thus, to help achieve the navigation task, it is of paramount importance that methods which rely only on the most relevant, in terms of position estimation, sensory input are developed. The ability to navigate using only this data is highly advantageous in that it allows the robot to maximize use of its limited computational resources. One may classify such input as a “landmark” and given their small size, each can be quickly memorized for future reference, thus mimicking the approach to navigation adopted by humans.

We present a statistical method, termed *Information Sampling*, to allow a robot to focus attention upon discriminating areas of an image set. Discriminating areas are defined as the pixels which vary significantly from one image to the next, i.e. those exhibiting the largest information change. Information Sampling is a component of a larger, *holistic*, navigation methodology [26], aimed at answering a number of fundamental questions raised by the application of vision to mobile robot navigation.

Our method was applied in an appearance-based context. In the field of Computer Vision, the use of appearance-based methods have become widespread in recent years. Successful applications have included handwritten character recognition [17], face characterization/recognition [24,25], visual inspection [19], object tracking/recognition [1,18], position estimation [4] and motion and gesture recognition [2,15]. In our case, visual environmental information was provided by an omnidirectional camera which gives a 360° view around the vertical. This work builds on our previous research with such a system [6,27,28].

\* Corresponding author. Tel.: +351-21-841-82-94; fax: +351-21-841-82-91.

*E-mail addresses:* niall@isr.ist.utl.pt (N. Winters), jasv@isr.ist.utl.pt (J. Santos-Victor).

As opposed to traditional position estimation techniques, which often relied upon precise metric measurements of position, appearance-based methods generate a qualitative estimate of position, essentially by image matching. For our experiments we used a topological [3,6,13] environmental representation.

At the training stage a map is built by capturing a sequence of images. Then, at runtime each acquired image is compared to the map and a qualitative estimate of position is determined. Given the large number of images involved, the data is compressed by applying Principal Component Analysis (PCA). The resulting low-dimensional eigenspace [18] implicitly represents the topological structure of the environment.

It is usual that entire images are used for matching. In order to overcome some well-known problems associated with this approach, recent work has attempted to determine *attentive regions* from the input images.

Ohba and Ikeuchi [20] proposed dividing each image into a number of smaller windows (which they termed eigenwindows). Eigenspace analysis was then applied to each window. This approach required storing a very large number of image windows and the chances of one window, acquired at runtime being matched to a number of images from the a priori set was high. As a solution, they used three criteria to eliminate the redundant windows, namely: detectability, uniqueness and reliability. Colin de Verdière and Crowley [5] reformulated the problem as a question of whether to use the set of eigenwindows, selected by a particular interest operator, or to use those windows selected from a predefined grid.

Discriminating regions are often defined as *features* in a single image, which together may form a landmark. Schmid and Mohr [23] used an interest operator [8] to determine where to compute local gray-value invariants in addressing the problem of image retrieval from a large database. In the area of mobile robotics, Knapek et al. [12] developed an approach to select salient landmarks from a single image which was strongly based on the work of Schmid and Mohr. Jugessur and Dudek [11] too utilized an interest point detector. Yeh and Kriegman [31] considered the problem of automatically selecting, from a set of 3D features, the set (landmark) which was most likely to be recognized in a single image.

We undertake a two-part approach to the navigation problem. Firstly, Information Sampling is applied to find the most discriminating data from the set of omnidirectional images. This data is then used to *directly* build a low-dimensional eigenspace representation of the environment for appearance-based navigation. Preliminary results were presented in [29,30]. Unlike the above feature-based approaches, our method does not exhibit the constraint of requiring the use of highly textured images. It is influenced by that of Rendas and Perrone [21] who addressed the problem of current mapping in coastal areas using a priori knowledge of the survey area.

In the field of robotics, appearance-based navigation using entire images is common. Perhaps one of the earliest examples was developed by Hong et al. [9]. The goal was for a robot to home to a given destination. Hancock and Judd [7] developed “Ratbot”, where localization was achieved by simply matching vertical bars from runtime images to those acquired a priori. A view sequence of images for navigation was proposed by Matsumoto et al. [16]. Zheng and Tsuji [32] presented a system which moved along a given route under human guidance and autonomously memorized a side-view of that route. These data were then used as a basis for route recognition. Ishiguro and Tsuji [10] used the frequency domain of the Fourier transform as an image-based memory of the environment. Maeda et al. [14] used the parametric eigenspace approach to image matching using images from a pan-and-tilt camera.

This paper is outlined as follows: Section 2 presents the Information Sampling method in detail. In Section 3, we show how Information Sampling relates to local appearance spaces. Section 4 presents our experimental results. We conclude and give possible future research directions in Section 5.

## 2. The Information Sampling method

Our approach requires the use of a priori images, independent of image type, from which we determined the most discriminating regions by applying *Information Sampling*. As a first step to explaining this process, we now outline the procedure for reconstructing an image, given only a small amount of data.

We assume that the images captured by the robot's camera can be modeled as a random vector  $I$ , characterized by a Gaussian distribution with mean  $\bar{I}$  and covariance  $\Sigma_I$ :

$$I \sim \mathcal{N}(\bar{I}, \Sigma_I) = p(I).$$

Usually, one can take an ensemble of images of the environment  $[I_1, \dots, I_m]$ , which can be utilized for computing  $\bar{I}$  and  $\Sigma_I$ , so that  $p(I)$  can be computed a priori. When the robot is navigating, we assume that the observations,  $d$ , consist of a selection of (noisy) image pixels (or subregions), rather than the entire image. Accordingly, the observation model can be expressed as:

$$d = SI + \eta, \quad (1)$$

where  $d$  stands for the observed data and the measurement noise,  $\eta$  is assumed to follow a Gaussian distribution with zero mean and covariance,  $\Sigma_n$ . We further assume that  $I$  and  $\eta$  are independent. The selection matrix,  $S$  is composed of a series of ones and zeros, the ones corresponding to the data points extracted from an image. We select a number of pixels to test by moving the set of ones in the selection matrix.

Having prior knowledge of  $I$ , in the form of a statistical distribution,  $p(I)$ , the problem now consists of estimating the (entire) image based on partial (noisy) observations of a few pixels,  $d$ . This problem can be formulated as a *Maximum a Posteriori* estimation of  $I$ . The posterior probability can be determined from Bayes rule as follows:

$$p(I|d) = \frac{p(d|I)p(I)}{p(d)}, \quad (2)$$

where  $p(d|I)$  is the likelihood of a pixel (or set of pixels) given a known image,  $I$ ; the prior distribution is denoted by  $p(I)$  and is assumed to have been learnt a priori. With this information we calculate the maximum a posteriori estimate of an image,  $\hat{I}_{\text{MAP}}$  [22] as follows:

$$\begin{aligned} \hat{I}_{\text{MAP}} &= \arg \max_I p(I|d), \\ \hat{I}_{\text{MAP}} &= (\Sigma_I^{-1} + S^T \Sigma_n^{-1} S)^{-1} (\Sigma_I^{-1} \bar{I} + S^T \Sigma_n^{-1} d). \end{aligned} \quad (3)$$

Thus,  $\hat{I}_{\text{MAP}}$  is the reconstructed image obtained using the pixel (or set of pixels),  $d$ . Notice that by

combining the prior image distribution with the statistical observation model, we can estimate the entire image based on the observation of a limited number of pixels.

### 2.1. Choosing the best data

Once we have reconstructed an image using the selected data, we can compute the error associated with this reconstruction. The error covariance matrix,  $\Sigma_{\text{error}}$  is given by:

$$\begin{aligned} \Sigma_{\text{error}} &= \text{Cov}(I - \hat{I}_{\text{MAP}}), \\ \Sigma_{\text{error}} &= (\Sigma_I^{-1} + S^T \Sigma_n^{-1} S)^{-1}. \end{aligned} \quad (4)$$

Of course, the quality of the estimate, and the “size” of  $\Sigma_{\text{error}}$  depend not only on the observation noise,  $\eta$  but also on the observed image pixels, as described by the selection matrix,  $S$ . Eq. (4) quantifies the quality of an estimate obtained using a particular set of image pixels. In theory, we can evaluate the *information content* of any individual image pixel or combination of pixels, simply by selecting an appropriate selection matrix,  $S$ , and determining the associated  $\Sigma_{\text{error}}$ .

This problem could be formulated as an experiment design process [22], in which we look for the optimal selection matrix  $S^*$  that minimizes (in some sense) the error covariance matrix. If we take the determinant of  $\Sigma_{\text{error}}$  as an indication of the “size” of the error, the optimal selection of image pixels would be given by:

$$S^* = \arg \min_S \{\det((\Sigma_I^{-1} + S^T \Sigma_n^{-1} S)^{-1})\}. \quad (5)$$

In practice, to avoid computing the inverse, we define the following equivalent optimization problem in terms of a modified uncertainty metric,  $U$ :

$$\begin{aligned} U &= -\log\{\det(\Sigma_I^{-1} + S^T \Sigma_n^{-1} S)\}, \\ S^* &= \arg \min_S U. \end{aligned} \quad (6)$$

Notice that the information criterion is based on the entire set of images and not, as with other methods, on an image-by-image basis. For instance, a highly textured image region would only be selected if it varied significantly from one image to the next.

So far, we have described *Information Sampling* as a process for (i) reconstructing an entire image from the

observation of a few (noisy) pixels and (ii) determining the *most relevant* image pixels,  $S^*$ , in the sense that they convey the most information about the image set.

Unfortunately, determining  $S^*$  is computationally impractical since we would have to compute  $\Sigma_{\text{error}}$  for all possible combinations of pixels scattered throughout the image. Instead, we partition the image into non-overlapping square windows of  $(l \times l)$  pixels. We term these regions *Information Windows*, denoted by  $\mathbf{w} = [w_1, \dots, w_n]$ .

## 2.2. Information Windows

Omnidirectional images were acquired at a resolution of  $768 \times 576$  pixels, filtered and subsampled to a resolution of  $128 \times 128$  pixels. These formed the database set,  $\mathbf{T}_{128}$ . In order to perform Information Sampling these images were further subsampled to a set of images,  $\mathbf{T}_{32}$ ,  $32 \times 32$  pixels in size. The reason for such a small image size relates to the complexity of determining the error covariance matrix,  $\Sigma_{\text{error}}$ , in Eq. (4).

To find the most discriminating Information Windows over  $\mathbf{T}_{32}$ , we found and ranked (see Section 2.3) the 16 *non-overlapping* windows of size  $8 \times 8$  pixels. We then calculated the equivalent  $32 \times 32$  Information Windows (in  $128 \times 128$  training images,  $\mathbf{T}_{128}$ ) to the  $8 \times 8$  windows (in the  $32 \times 32$  images,  $\mathbf{T}_{32}$ ).

Additionally, we ranked the 225 *overlapping* windows of size  $4 \times 4$  pixels in  $\mathbf{T}_{32}$ . The overlapping windows were generated by shifting each window in the horizontal and vertical direction by 2 pixels, thus generating an overlap of 50%. We then calculated the equivalent  $16 \times 16$  Information Windows (in  $128 \times 128$  training images,  $\mathbf{T}_{128}$ ) to the  $4 \times 4$  windows (in the  $32 \times 32$  images,  $\mathbf{T}_{32}$ ).

## 2.3. Search and rank

By using Eq. (6), we can rank *Information Windows* or combinations of such windows, in terms of their information content. Finding the set of pixels to select (from the a priori set of images) as the best information is highly computationally intensive. In order to overcome this problem, we implemented two greedy search algorithms: *Combinatorial Search* and *Simple Search*. These were used to find and rank the best Information Windows.

### 2.3.1. Combinatorial Search

We first search for the best Information Window. Then, the search for the next best window is made keeping the first window *fixed*, thus locating the best *pair* of windows. As the method continues it determines the best triplet of windows, etc. If we denote  $n$  as the number of windows within an image, this method requires the evaluation of Eq. (6),  $n!$  times. The method automatically groups the Information Window(s) into a single window, a pair of windows, a triplet of windows, etc. Notice that this method is not a true Combinatorial Search, which would require the evaluation of *all* possible combinations of windows.

### 2.3.2. Simple Search

This is a faster search algorithm. We rank each of the Information Windows *independently*. In this case, Eq. (6) has only to be evaluated  $n$  times. As distinct from Combinatorial Search, if we wish to group the best (pair, triplet, etc. of) window(s) we must do it manually based on the initial ranking.

## 3. Information Sampling for local appearance-based robot navigation

As previously noted, many appearance-based navigation systems relied upon the use of entire (possibly, omnidirectional) images to build a “global eigenspace” (or “global appearance space”) for navigation. Alternatively, if an eigenwindow approach was used, all windows must have first been projected before calculation of the most discriminating.

A key advantage of our approach is that we *only* require the most discriminating information obtained by Information Sampling to build a *local appearance space*. In this way, we directly project the best information from each image into the eigenspace, significantly reducing the number of projected windows. This immediately reduces the level of possible ambiguity. The local appearance space has an orthonormal basis of eigenvectors of size  $(l^2 \times 1)$ , where  $l$  is the length of the side of a square Information Window. It is our topological environmental representation.

We note here that, conceivably, navigation could be achieved by matching the reconstructed image,  $\hat{I}_{\text{MAP}}$  to the set of omnidirectional images, although this would be computationally expensive.

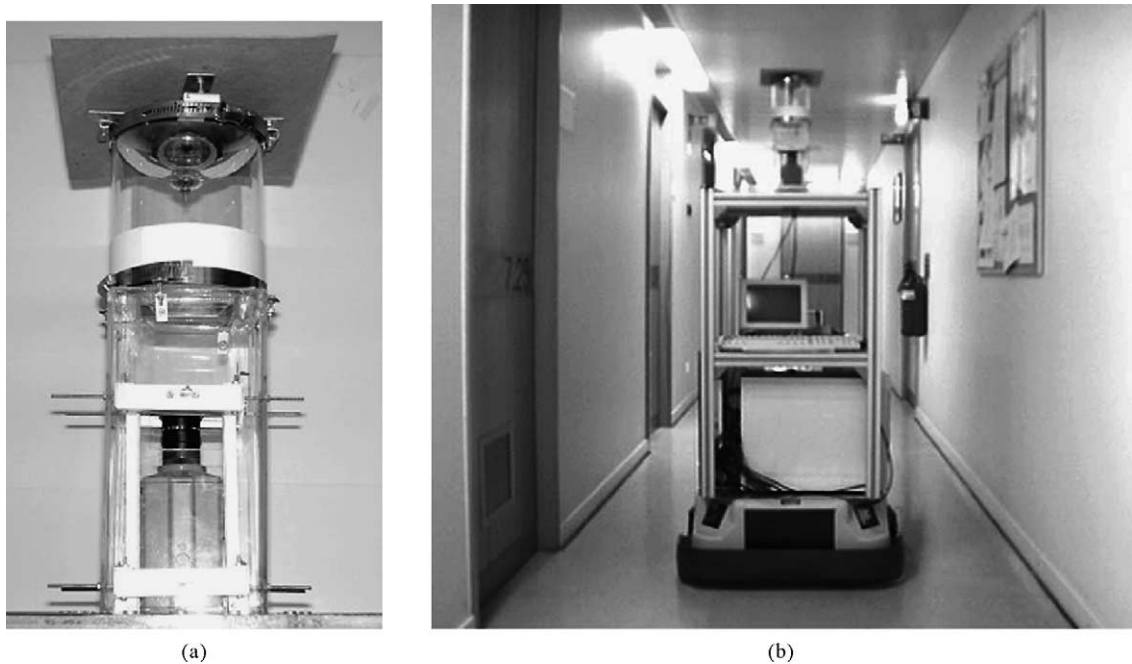


Fig. 1. (a) The omnidirectional camera with a spherical mirror and (b) the camera mounted on a Labmate mobile platform.

#### 4. Experimental results

For the experiments outlined in this paper, images acquired from an omnidirectional camera with a spherical mirror, built in-house at the Instituto de Sistemas e Robótica, Lisbon, were used. This camera was mounted on a Labmate mobile platform and images were captured as it traversed the environment. All processing was carried out using a Pentium III 350 MHz processor with 128MB RAM. The system is shown in Fig. 1. Real world experiments verify that by using Information Sampling to focus upon attentive regions, effective navigation is possible.

The experimental results are presented in three parts. The first concerns the ranking of Information Windows, the second, reconstructing images from Information Windows and the third, mobile robot navigation using only the Information Windows. Results were obtained for two sets of images: the first consisted of 89 omnidirectional images acquired every 10 cm and for the second, a set of 53 omnidirectional images, obtained every 20 cm were acquired in the same corridor but from a different starting position.

##### 4.1. Part I: window ranking

Fig. 2(a) shows the non-overlapping Information Windows available for selection and Fig. 2(b) these Information Windows, *individually ranked* from the most (number 1) to the least discriminating (number 16), using Simple Search. Fig. 3 shows the ranking when using panoramic images. The following example provides an intuitive idea of the Information Sampling method. All of the omnidirectional images in this paper show the robot in the center of each image. Any Information Window which contains the robot is not a discriminating one, and so, it follows that such a window should have a relatively low ranking. As shown in Fig. 2(b), this proves to be the case: the four Information Windows which contain the robot are ranked from numbers 10 to 15. Additionally, the four windows at the periphery of the image also have a low ranking, since they only contain a portion of the omnidirectional image. It should be noted that the corridor in which the a priori set of images were acquired had a number of offices on one side (the top half of the omnidirectional images) and only a single



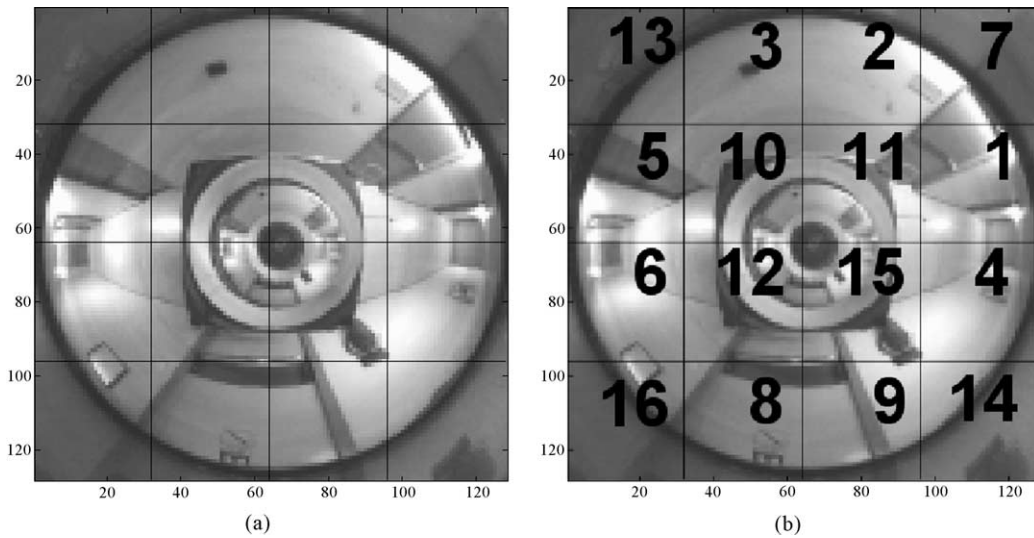


Fig. 2. Ranking results. (a) The 16 non-overlapping Information Windows. (b) Those windows ranked, according to the amount of information they contain, using Simple Search.

door and notice-board on the other (the bottom half of the omnidirectional images). Thus, as the robot travels down the corridor more information change occurs in the top half of the omnidirectional images. Again, this is borne out by the window ranking, where the three highest ranking Information Windows are all in the top half of the omnidirectional image.

Fig. 4(a) shows the Information Windows available for selection from the second set and Fig. 4(b) the window ranking when using *non-overlapping* windows. The four Information Windows which contain the robot are ranked from numbers 7 to 12. Additionally, the four windows at the periphery of the image have the lowest ranking. Significantly, the *same* highest ranking Information Window (window 8) is selected from both sets. The other Information Windows are ranked in a different order but note that Information Sampling chooses the same six top ranking windows from both sets.

To further test the technique, we found and ranked the 225 *overlapping* windows of  $16 \times 16$  pixels in size. The advantage of this approach is that we can focus upon smaller areas of the omnidirectional image. Naturally, relevant information contained within the image received a high ranking and, advantageously, portions of the image which were close to the background, and relevant, were highly ranked. As expected the dark background received the lowest ranking. An image of the 10 best overlapping Information Windows is shown in Fig. 5.

#### 4.1.1. Graphing the information content

Fig. 6 shows the graphs of the Information Windows, obtained from omnidirectional images, ranked using (a) Simple Search and (b) Combinatorial Search. In both cases, the  $x$ -axis corresponds to the window ranking, from 1 to 16 and the  $y$ -axis corresponds to the uncertainty metric,  $U$  calculated using Eq. (6).

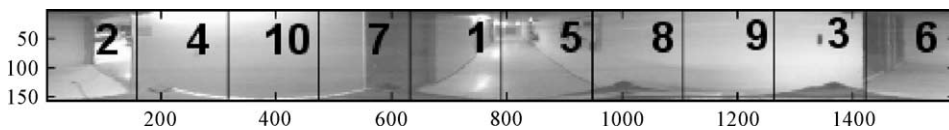


Fig. 3. The Information Windows obtained using Panoramic Images, ranked, according to the amount of information they contain, using Simple Search.

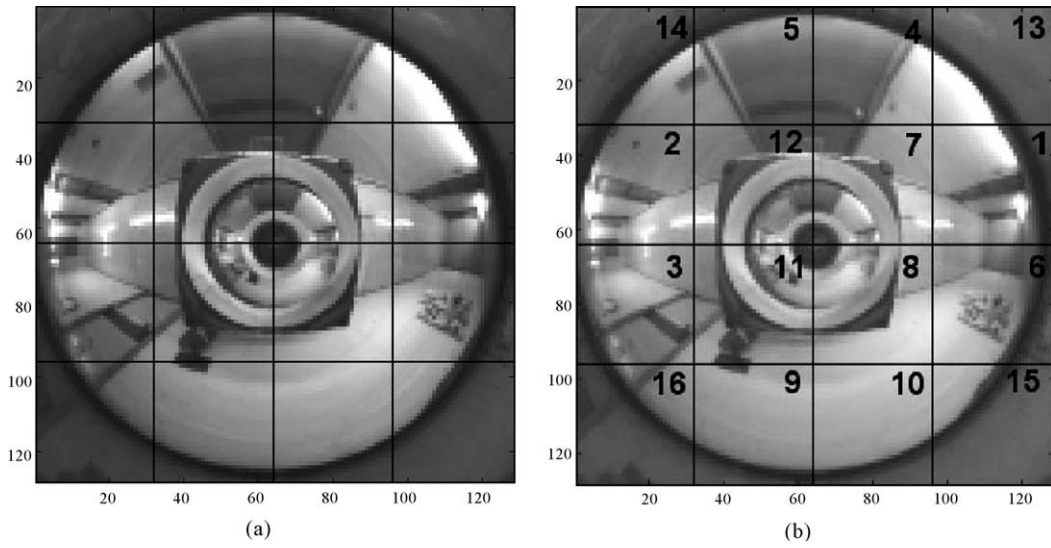


Fig. 4. Ranking results. (a) The 16 non-overlapping Information Windows. (b) The windows ranked, according to the amount of information they contain, using Simple Search.

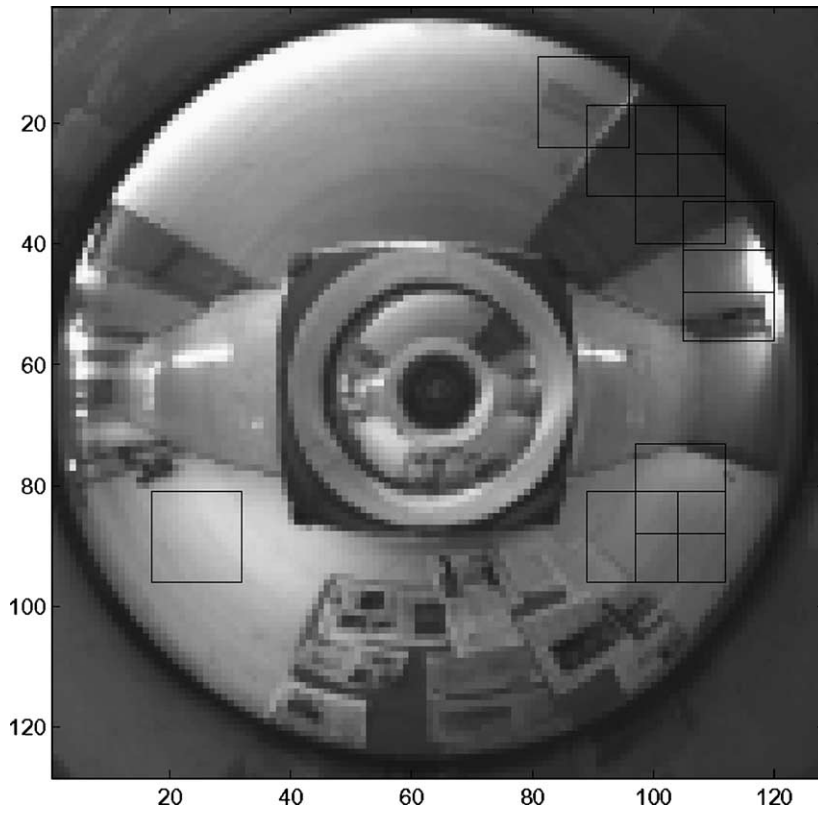


Fig. 5. The 10 best overlapping Information Windows.

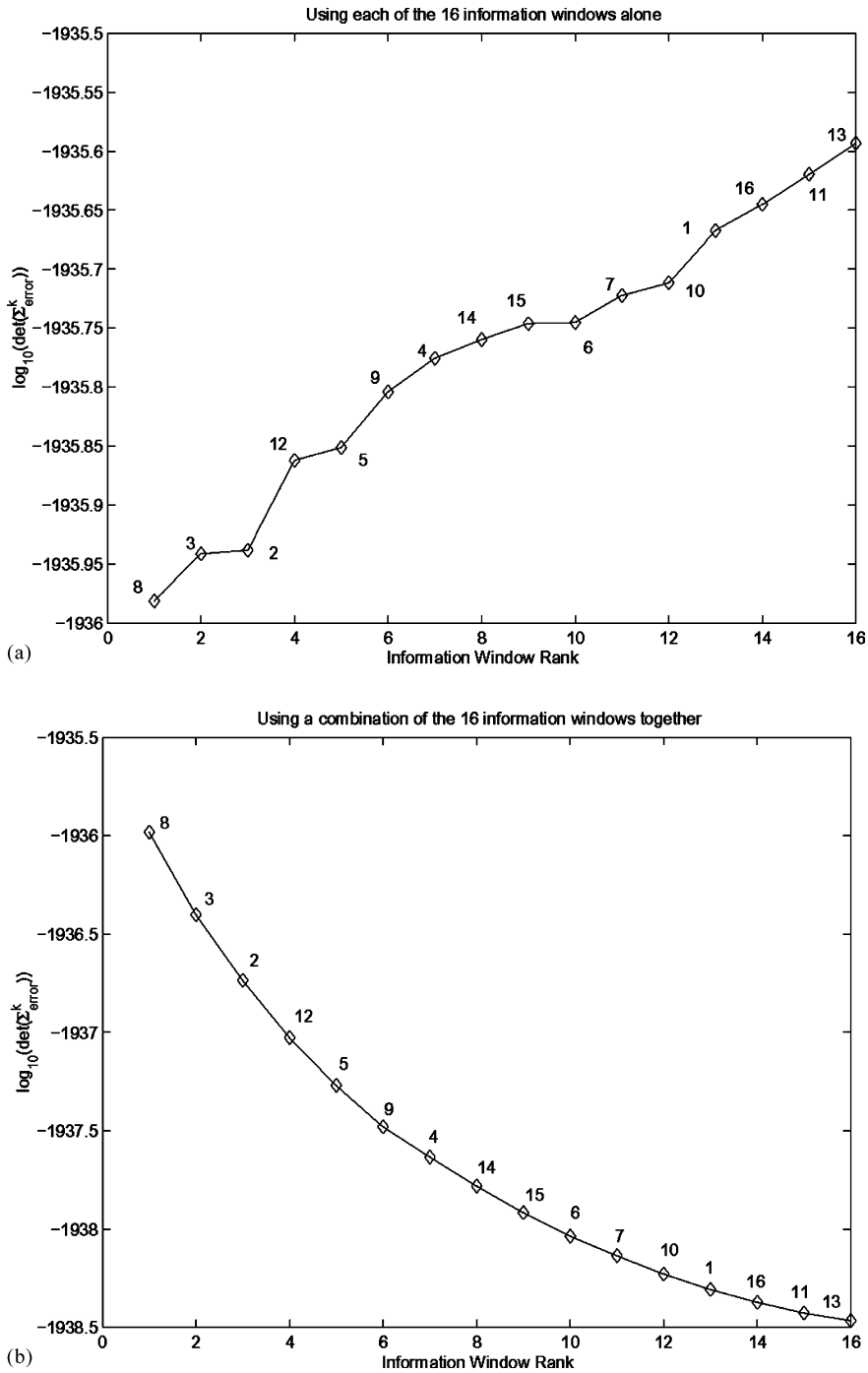


Fig. 6. Graphs of the data contained in each Information Window versus the window rank when using (a) Simple Search and (b) Combinatorial Search. The numbers along the graph line are the window numbers.



The numbers along the graph line correspond to the 16 non-overlapping Information Windows per omnidirectional image. For example, using Simple Search, Fig. 6(a) tells us that the eighth Information Window exhibits the lowest uncertainty value and so is individually ranked in first position, while the third window, having a higher uncertainty value, is individually ranked in second position, etc. Using Combinatorial Search Fig. 6(b) tells us that the eighth window is ranked in first position. This window is then fixed and the best pair of windows, in this case the eighth plus

the third, are found. Thus, the third window contains the next best amount of information and is ranked in second position. Using Combinatorial Search the next best window added at each stage matches the window rank chosen by Simple Search.

Combinatorial Search continues until all windows have been combined. As can be seen from Fig. 6(b) each combination of Information Windows exhibits a lower uncertainty measure than the previous one. Intuitively, this makes sense as the more information available, the better the image reconstruction (see

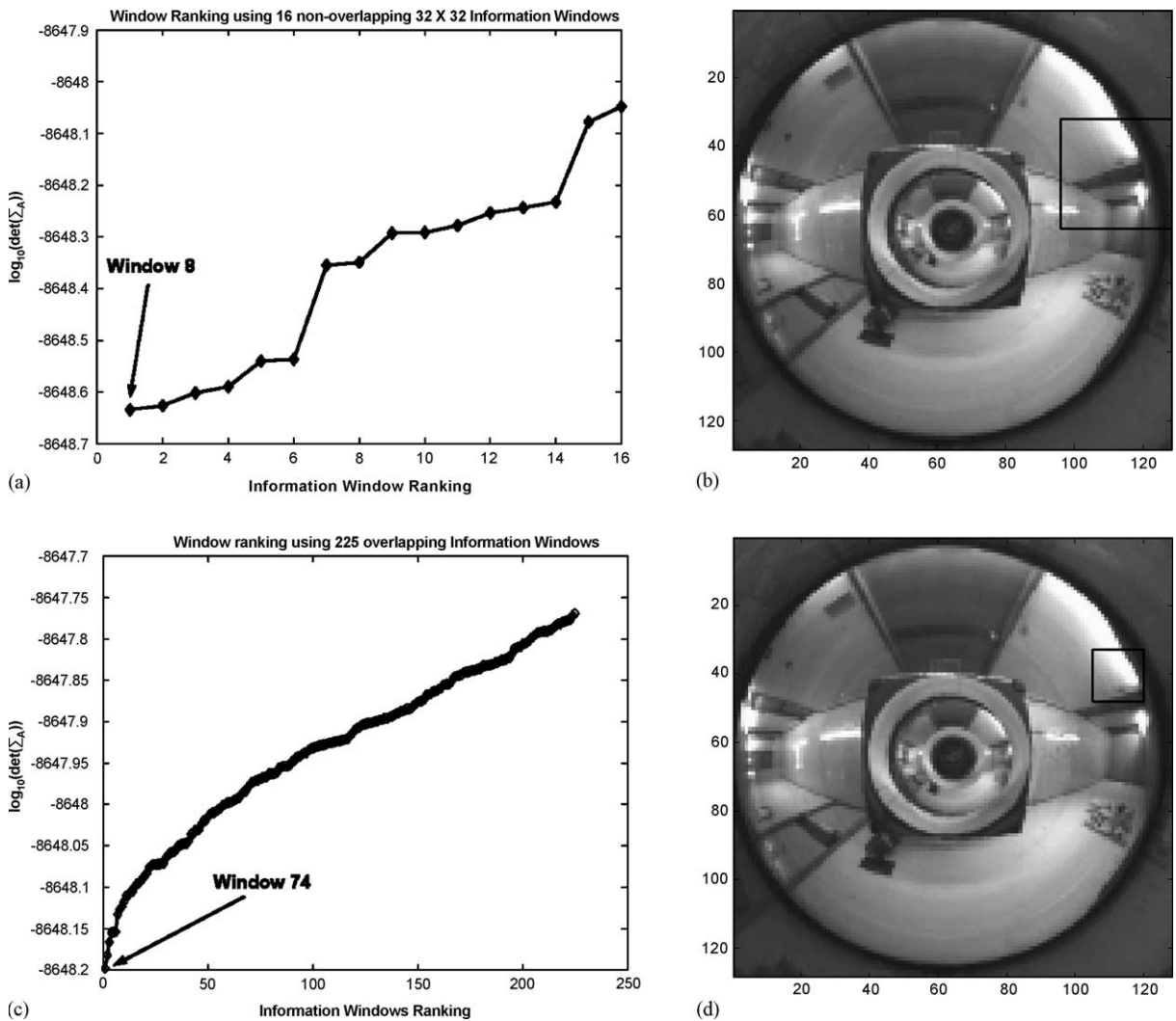


Fig. 7. Graphs of the information contained in each Information Window versus the window rank using (a) non-overlapping and (c) overlapping windows. The best (b) non-overlapping and (d) overlapping Information Window in an image.

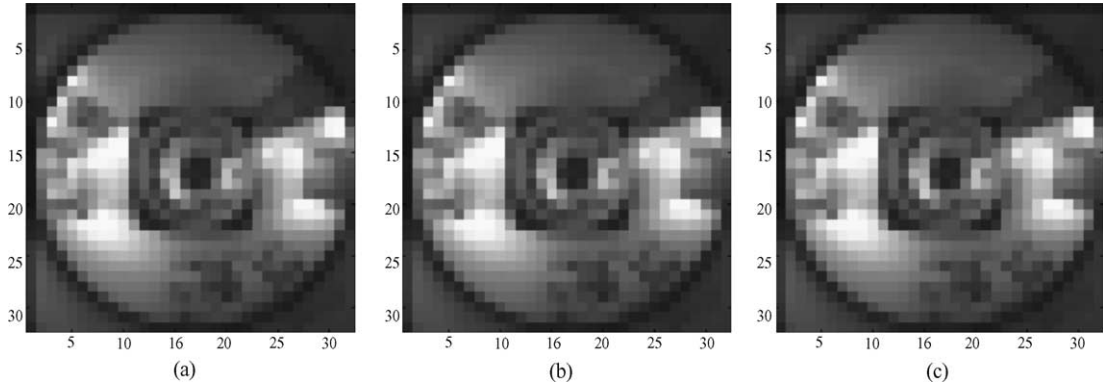


Fig. 8. (a) A  $32 \times 32$  omnidirectional image acquired at runtime. (b) Its reconstruction using the *most discriminating* Information Window. (c) Its reconstruction using all of the Information Windows. Each Information Window is  $8 \times 8$  pixels in size.

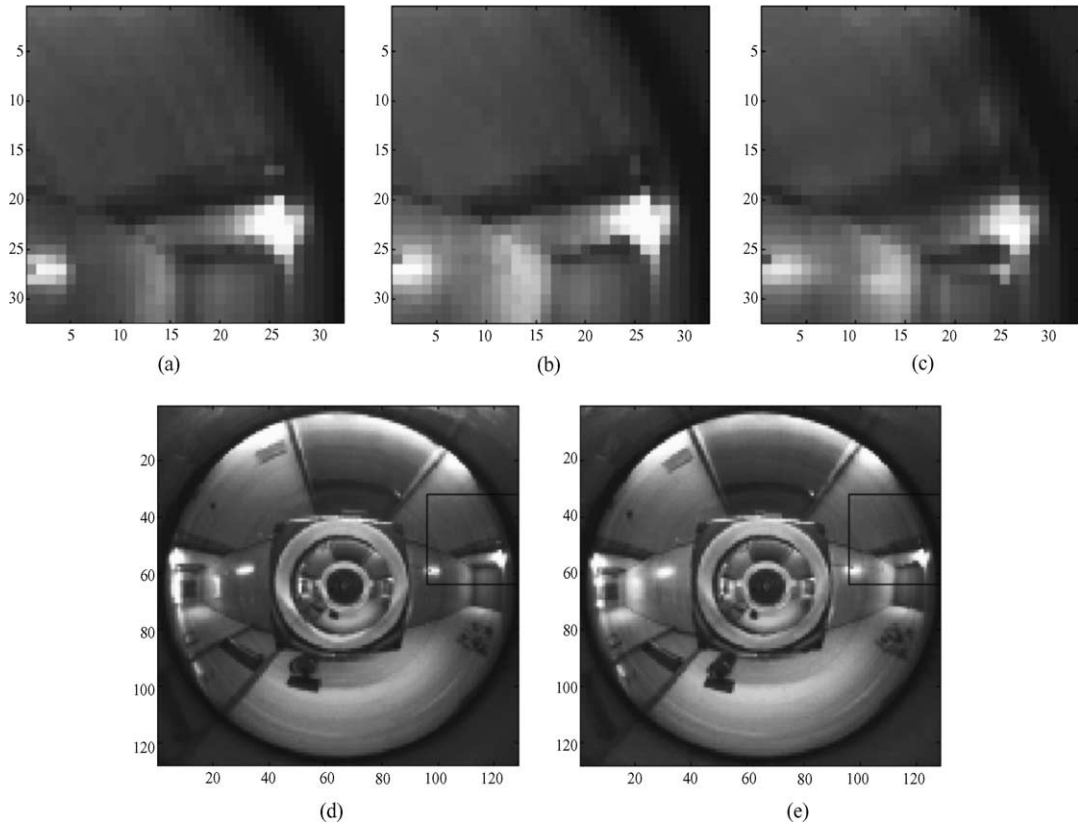


Fig. 9. Close-up of the  $32 \times 32$  Information Windows: (a) unknown, (b) closest and (c) reconstructed. The position of (d) the unknown and (e) the closest windows in their respective omnidirectional images.

Eq. (3)) should be. However, the payoff for using many Information Windows is not significant, as can be seen from the small drop in uncertainty. This result is also borne out by Fig. 8, as detailed in Section 4.2. Clearly, the fact that the highest ranking Information Window is not only the most relevant, but is the most relevant by a significant factor, is the reason why we need to use only it for navigation.

Fig. 7 shows the ranking results when using (a) non-overlapping and (c) overlapping windows from the second set of images. Again, when using non-overlapping Information Windows, the eighth exhibits the lowest uncertainty value and so is ranked in first position. When using (c) overlapping windows, window number 74 is ranked in first position. Fig. 7

shows the best (b) non-overlapping and (d) overlapping Information Windows in an omnidirectional image.

#### 4.2. Part II: image reconstruction

The reconstruction results were obtained using Information Windows of size  $8 \times 8$  pixels extracted from omnidirectional images of  $32 \times 32$  pixels in size. Fig. 8 shows an (a) omnidirectional image from the a priori set, (b) its reconstruction using only the most discriminating  $8 \times 8$  Information Window (i.e. number 8) and (c) its reconstruction using all of the Information Windows. Reconstruction was achieved using Eq. (3). As can be seen from the images, a good

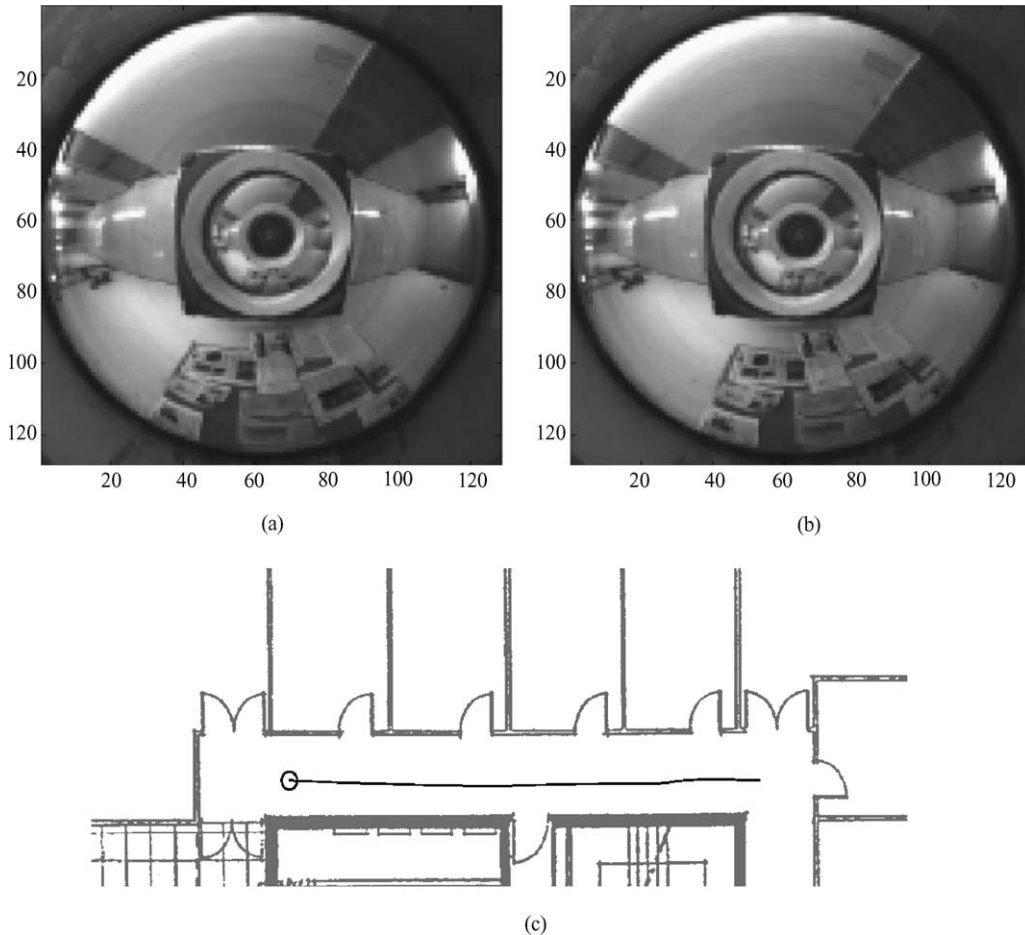


Fig. 10. (a) An unknown image, (b) its closest match and (c) the path traveled by the robot when using entire  $128 \times 128$  images.

reconstruction is obtained using only the best Information Window. This is an indication of the power of Information Sampling.

#### 4.3. Part III: navigation

As a first test of navigation, we ran our Labmate mobile robot in a structured environment. *Only* the best Information Window, from each image, was projected into the eigenspace. The images in Fig. 9(a)–(c) show the results obtained using windows of  $32 \times 32$  pixels in size. The top row shows (a) the most relevant Information Window from an unknown image, (b) its closest match from the a priori set of best Information Windows and (c) its reconstruction

using PCA. Fig. 9 shows (d) the best Information Window in the unknown  $128 \times 128$  image and (e) its closest match from the a priori set obtained by projecting only the most relevant Information Window. We note here that we could in principal, given enough computing power, use Eq. (3) to reconstruct a  $128 \times 128$  image using only the most relevant window.

To further test the applicability of the Information Sampling technique, we ran three more position estimation experiments. The first experiment used entire  $128 \times 128$  images for matching, the second the most informative *non-overlapping*  $32 \times 32$  Information Window and the third the 10 most informative *overlapping*  $16 \times 16$  Information Windows.

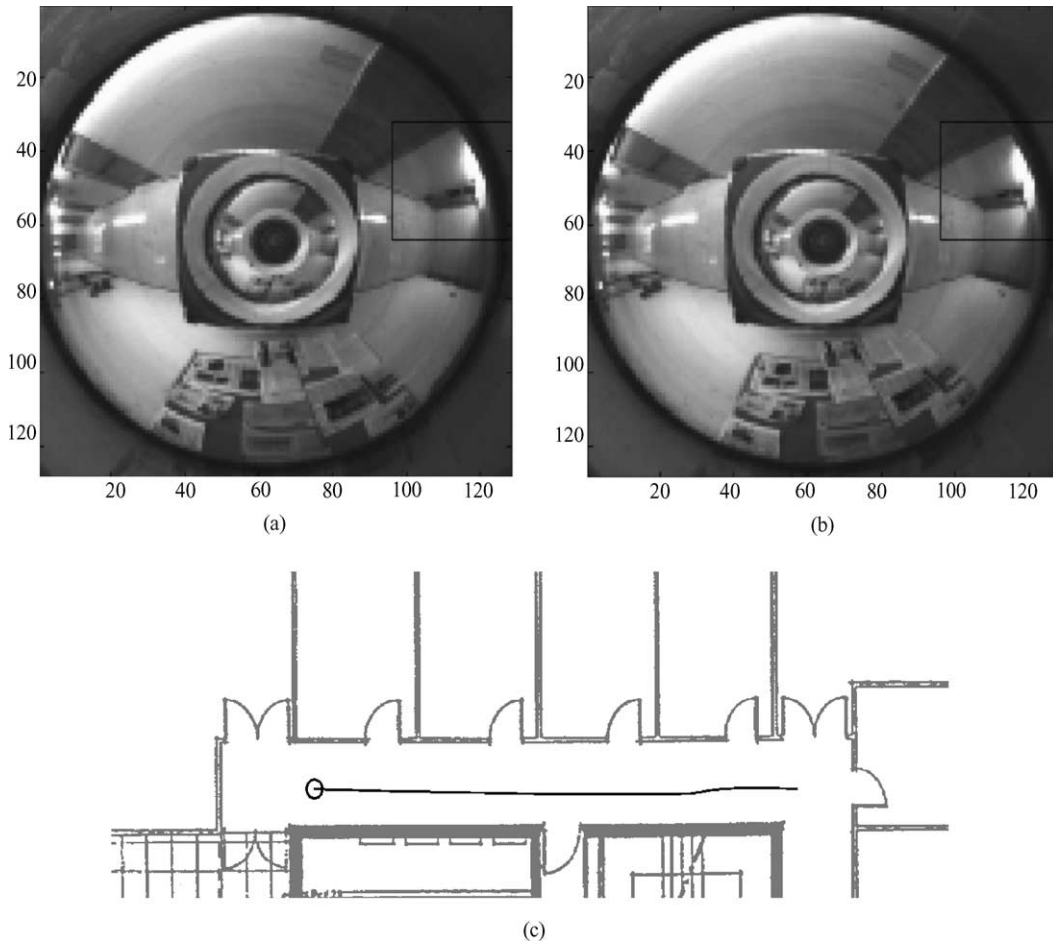


Fig. 11. (a) An unknown image, (b) its closest match and (c) the path traveled by the robot when using the best  $32 \times 32$  non-overlapping Information Window.

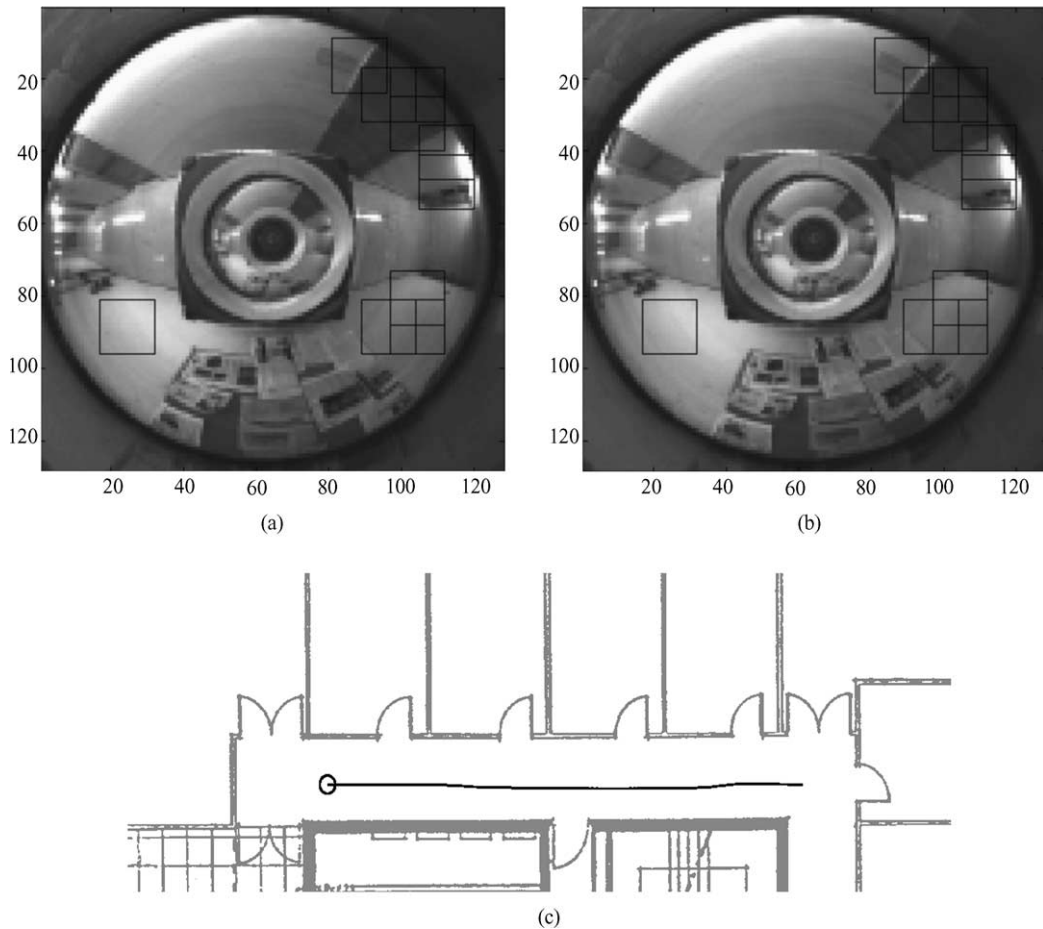


Fig. 12. (a) An unknown image, (b) its closest match and (c) the path traveled by the robot when using the 10 best  $16 \times 16$  overlapping Information Windows.

Fig. 10 shows (a) an unknown image, (b) its closest match from the a priori set of entire images and (c) the distance traveled ( $\approx 10.5$  m) by the robot under closed loop control. Figs. 11 and 12, respectively, show the same experiment but using *non-overlapping* and *overlapping* Information Windows. Significantly, in these latter cases, the number of pixels used for position estimation was only 6.25% of the total, when using non-overlapping windows, and just 10.93% in the case of overlapping windows, allowing the robot to maximize use of its limited computational resources. Clearly, vision-based navigation was successfully completed by using discriminating environmental information.

## 5. Conclusion and future work

In this paper we presented a statistical method, termed Information Sampling, which calculated the most discriminating pixels from an a priori image set. In terms of navigation this was highly advantageous as it allowed a mobile robot to successfully complete its task by building an environmental representation which made effective use of the robot's perceptual capabilities. It was noted that Information Sampling forms part of a larger methodology, which seeks to tackle vision-based navigation from a holistic viewpoint.

Our future work shall concentrate on building environmental representations *online*. Information

Sampling was a first step in this process, as it allowed for the construction of a highly compact environmental representation. Additionally, widening the scope of application to other areas of Computer Vision is envisioned.

## Acknowledgements

This work was partly funded by the European Union RTD—Future and Emerging Technologies (Project No.: IST-1999-29017), Omniviews.

## References

- [1] M.J. Black, A.D. Jepson, Eigentracking: robust matching and tracking of articulated objects using a view-based representation, *International Journal of Computer Vision* 26 (1) (1998) 63–84.
- [2] O. Chomat, J.L. Crowley, Recognizing motion using local appearance, in: *Proceedings of the Sixth International Symposium on Intelligent Robotic Systems*, Edinburgh, Scotland, 1998, pp. 271–279.
- [3] J.L. Crowley, Navigation for an intelligent mobile robot, *IEEE Journal on Robotics and Automation* RA-1 (1) (1985) 31–41.
- [4] J.L. Crowley, F. Pourraz, Continuity properties of the appearance manifold for mobile robot position estimation, *Image and Vision Computing* 17 (2001) 741–752.
- [5] V. Colin de Verdière, J.L. Crowley, Visual recognition using local appearance, in: *Proceedings of the Fifth European Conference on Computer Vision*, June 1998, pp. 640–654.
- [6] J. Gaspar, N. Winters, J. Santos-Victor, Vision-based navigation and environmental representations with an omnidirectional camera, *IEEE Transactions on Robotics and Automation* 16 (6) (2000) 890–898.
- [7] T. Hancock, S. Judd, Ratbot: robot navigation using simple visual algorithms, in: *Proceedings of the IEEE Regional Conference on Control Systems*, August 1993, pp. 181–184.
- [8] C. Harris, M. Stephens, A combined corner and edge detector, in: *Proceedings of Alvey Vision Conference*, 1988, pp. 147–151.
- [9] J. Hong, X. Tan, B. Pinette, R. Weiss, E.M. Riseman, Image-based homing, in: *Proceedings of the IEEE International Conference on Robotics and Automation*, Sacramento, CA, 1991, pp. 620–625.
- [10] H. Ishiguro, S. Tsuji, Image-based memory of environment, in: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Lisbon, Portugal, 1996, pp. 634–639.
- [11] D. Jugessur, G. Dudek, Local appearance for robust recognition, in: *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, June 2000, pp. 834–839.
- [12] M. Knapek, R. Swain Oropeza, D. Kriegman, Selecting promising landmarks, in: *Proceedings of the IEEE International Conference on Robotics and Automation*, San Francisco, CA, 2000, pp. 3771–3777.
- [13] K. Lynch, *The Image of the City*, MIT Press, Cambridge, MA, 1960.
- [14] S. Maeda, Y. Kuno, Y. Shirai, Active navigation vision based on eigenspace analysis, in: *Proceedings of the International Conference on Intelligent Robots and Systems*, Tucson, AZ, 1997, pp. 1018–1023.
- [15] J. Martin, J.L. Crowley, An appearance-based approach to gesture recognition, in: *Proceedings of the International Conference on Image Analysis and Processing*, September 1997.
- [16] Y. Matsumoto, M. Inaba, H. Inoue, Visual navigation using view-sequenced route representation, in: *Proceedings of the International Conference on Robotics and Automation*, Minneapolis, MN, 1996, pp. 83–88.
- [17] H. Murase, F. Kimura, F. Yoshimura, Y. Miyake, An improvement of the autocorrelation matrix in pattern matching method and its application to hand printed ‘hiragana’, *Transactions of IEICE J64-D* (3) (1981) 276–283.
- [18] H. Murase, S.K. Nayar, Visual learning and recognition of 3D objects from appearance, *International Journal of Computer Vision* 14 (1) (1995) 5–24.
- [19] S.K. Nayar, S.A. Nene, H. Murase, *Subspace methods for robot vision*, Technical Report CUCS-06-95, Columbia University, New York, 1995.
- [20] K. Ohba, K. Ikeuchi, Detectability, uniqueness and reliability of eigen windows for stable verification of partially occluded objects, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (9) (1997) 1043–1048.
- [21] M. Rendas, M. Perrone, Using field subspaces for on-line survey guidance, in: *Proceedings of Oceans 2000*, Providence, RI, June 2000.
- [22] A.P. Sage, J.L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill, New York, 1971.
- [23] C. Schmid, R. Mohr, Local grayvalue invariants for image retrieval, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (5) (1997) 530–535.
- [24] L. Sirovich, M. Kirby, Low dimensional procedure for the characterization of human faces, *Journal of the Optical Society of America* 4 (3) (1987) 519–524.
- [25] M.A. Turk, A.P. Pentland, Face recognition using eigenfaces, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1991, pp. 586–591.
- [26] N. Winters, A holistic approach to mobile robot navigation using omnidirectional vision, Ph.D. thesis, Trinity College, University of Dublin, Ireland, October 2001, <http://www.isr.ist.utl.pt/~niial>.
- [27] N. Winters, J. Gaspar, J. Santos-Victor, Omni-directional vision for robot navigation, in: *Proceedings of the First International IEEE Workshop on Omnidirectional Vision CVPR*, June 2000, pp. 21–28.
- [28] N. Winters, J. Santos-Victor, Omnidirectional visual navigation, in: *Proceedings of the Seventh International Symposium on Intelligent Robotic Systems*, July 1999, pp. 109–118.



- [29] N. Winters, J. Santos-Victor, Information sampling for optimal image data selection, in: Proceedings of the Ninth International Symposium on Intelligent Robotics Systems, Coimbra, Portugal, July 2001, pp. 249–257.
- [30] N. Winters, J. Santos-Victor, Visual attention-based robot navigation using information sampling, in: Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, HI, October 2001, pp. 1670–1675.
- [31] E. Yeh, D.J. Kriegman, Toward selecting and recognizing natural landmarks, Technical Report 9503, Yale University, New Haven, CT, 1995.
- [32] J.Y. Zheng, S. Tsuji, Panoramic representation for route recognition by a mobile robot, *International Journal of Computer Vision* 9 (1) (1992) 55–76.



**Niall Winters** received the Ph.D. degree in Computer Science from Trinity College, University of Dublin, Ireland in 2001 and the B.Sc. (D. Hons.) degree in Computer Science and Experimental Physics from the National University of Ireland, Maynooth in 1997. From 1997–2001, he was a member of the Computer Vision and robotics Group at Trinity College.

For the academic years 1998–1999 and 2000–2001 he was a Visiting Researcher at the VisLab—Computer

Vision Laboratory of the Institute of Systems and Robotics, Instituto Superior Técnico, Lisbon, Portugal. He is currently a Research Associate at the same institution. His primary research interests lie in the areas of omnidirectional vision, vision-based robot navigation and the application of technology to development and learning.



**José Santos-Victor** obtained the Licenciatura degree in Electrical and Computer Engineering from IST—Instituto Superior Técnico, Lisbon, Portugal, in 1988, the M.Sc. degree in 1991 and the Ph.D. in 1995 from the same institution, in the areas of Active Computer Vision and Robotics.

He is an Assistant Professor at the Department of Electrical and Computer Engineering of IST, and he is the scientific responsible for the participation of IST in various European and National research projects.

He is a researcher of the Institute of Systems and Robotics, where he coordinates the research at the VisLab—Computer and Robot Vision Laboratory. His research interests are in the areas of Computer and Robot Vision, Intelligent Control Systems, particularly in the relationship between visual perception and the control of action, in (land and underwater) mobile robots.