ROBUST METHODS FOR ROBOT LOCALIZATION UNDER CHANGING ILLUMINATION CONDITIONS

Comparison of Different Filtering Techniques

Lorenzo Fernández Rojo, Luis Payá, Oscar Reinoso, Arturo Gil and Miguel Juliá
Departamento de Ingeniería de Sistemas Industriales, Miguel Hernández University
Avda. de la Universidad s/n. 03202, Elche (Alicante), Spain

Keywords: Omnidirectional vision, Robot mapping, Appearance-based methods, Robust localization and illumination effects filtering.

Abstract: The use of omnidirectional systems provides us with rich visual information that allows us to create appearance-based dense maps. This map can be composed of several panoramic images taken from different positions in the environment. When the map contains only visual information, it will depend heavily on the conditions of the environment lighting. Therefore we get different visual information depending on the time of day when the map is created, the state of artificial lighting in the environment, or any other circumstance that causes a change in the illumination of the scene. To obtain a robust map against changes in the illumination of the environment we apply different filters on the panoramic images. After that, we use some compression methods that allow us to reduce the amount of information stored. We have conducted a comprehensive experimentation to study which type of filter best adapts to changing lighting conditions.

1 INTRODUCTION

When a robot or a team of robots have to carry out a task in a given environment, in most cases, an internal representation is required to allow the robot to estimate its initial position and orientation, and navigate to the target points. Omnidirectional vision systems are commonly used at this kind of applications due to their low cost and the amount of information they provide. When working in unstructured environments where the creation of appropriate models of recognition can be an arduous chore, it is useful to use appearance-based approaches that offer a systematic and intuitive way to construct the map. The main problem such approaches present is the high computational cost because they do not extract relevant information from images, using the image as a whole.

To alleviate the high computational cost, several researchers have shown how it is possible to use a representation of the environment in a lower order subspace, using compression techniques. A widely extended method is PCA (Principal Components Analysis). One example is the database created in (Kröse, Bunschoten, Hagen, Terwijn and Vlassis, 2004). Uenoara and Kanade (1998) studied the problem of rotation in the plane in which the robot moves, using a set of rotated images. Jogan and Leonardis (2000) applied these concepts to an appearance-based map of an environment. Other related works (Menegatti, Maeda and Ishiguro, 2004) defined the concept of Fourier Signature and (Rossi, Ranganathan, Dellaert and Menegatti, 2008) used the Spherical Fourier Transform of omnidirectional images, using the Discrete Fourier Transform to compress the information. Appearance-based techniques constitute a basis framework to other robotics applications, as in route-following, as Payá, Reinoso, Gil and Sogorb (2008) show.

The appearance of an image will depend, in general, on the appearance of the objects that appear on it. Adini, Moses and Ullman (1997) show the influence of the illumination of the scene in a process of facial recognition. An individual cannot be recognized if there is a substantial change of lighting in the scene. Murase and Nayar (1994) use an appearance-based approach to avoid the problems of illumination variation. With this aim many views of the object are generated under different lighting conditions. Faraid and Adelson (1999) show that it is possible to separate the effects of reflections and

Fernández Rojo L., Payá L., Reinoso O., Gil A. and Juliá M. (2010).
ROBUST METHODS FOR ROBOT LOCALIZATION UNDER CHANGING ILLUMINATION CONDITIONS - Comparison of Different Filtering Techniques.
In *Proceedings of the 2nd International Conference on Agents and Artificial Intelligence - Artificial Intelligence*, pages 223-228
DOI: 10.5220/0002705002230228

illumination using ICA (Independent Component Analysis). Other researchers (Bischof, Wildenauer and Leonardis, 2004) have shown how to mitigate the effects of lighting on the appearance of an object, using gradient filter banks. The approach consists in implementing a series of filters before building the linear subspace using PCA. Other works (De Araújo, Maia, D´Angelo and D´Angelo, 2006) make use of homomorfic filters banks to separate the components of luminance and reflectance. This way it is possible to filter these components separately, reducing significantly the dependence of image appearance with respect to changes in lighting.

In this paper we present a methodology to build an appearance-based dense map. Several kinds of filters and compression techniques have been tested to make the map robust against changes in lighting conditions.

The work is structured as follows. Section 2 introduces some filtering techniques to eliminate the dependence on changes in the lighting. Section 3 presents some compression techniques to reduce the computational cost. In section 4 the method to build the map and how to obtain the position of the robot is detailed. We show the results of experiments carried out in section 5. Finally, in section 6, we present the conclusions of the work.

2 FILTERING TECHNIQUES USING PANORAMIC IMAGES

The appearance of an object in an image can vary strongly depending on the kind and level of illumination of the scene. When we work with the appearance of panoramic images, it is necessary to take into account the fact that appearance is influenced both by the position and shape of the objects and the lighting conditions. It is therefore necessary to implement a mechanism that allows us to work independently of the lighting conditions of the environment.

Several researchers have studied how to get invariance with respect to the illumination of the scene in object recognition tasks. We have separated the different methods in two fields. The first one is related to the application of a bank of gradient (first derivative) or Laplacian (second derivative) filters. The second one consists in performing a homomorfic filtering of the image separating the luminance from the reflectance component.

2.1 Edge Detector

The main advantage of using a representation of the image edges resides mainly in the fact that we obtain a compact representation and that, in most cases, it is insensitive to changes in the lighting on the objects of the image.

An edge detection filtering can be carried out through the Prewitt gradient filter, based on the estimation of the modulus of the gradient using two masks of size 3x3 (h_1 in the x-axis and h_2 in the y-axis):

$$h_1 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad h_2 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \tag{1}$$

An evolution of the Prewitt Filter is the Sobel filter that, apart from estimating the value of the modulus of the gradient, produces a smoothing of the image that may be beneficial, taking into account the noisy behaviour that the estimations based on the derivation of the image may present:

$$h_1 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad h_2 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{2}$$

Another method for detecting edges is the Laplacian of Gaussian operator, which combines the effect of a Gaussian smoothing with the improvement in the location of the edge (cross of 0 for the second derivative). In this case it is only necessary to apply a mask:

$$h_2 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \tag{3}$$

2.2 Homomorfic Filter

The Homomorfic filter can separate the components of luminance and reflectance of an image (Gonzalez and Woods, 1993). Thus it is possible to build a filter for each component separately, allowing us to control the contribution of each component on the image appearance. It is possible to separate the luminance from the reflectance component by applying the Neperian logarithm operator on the image:

$$\begin{aligned} f(x,y) &= i(x,y) \times r(x,y) \\ z(x,y) &= \ln(f(x,y)) \\ z(x,y) &= \ln(i(x,y)) + \ln(r(x,y)) \end{aligned} \tag{4}$$

Once the components are separated, the 2D Discrete Fourier Transform is computed. It is at this point that we can filter the image in the frequency domain:

$$\begin{aligned} &\Im(z(x,y)) = \Im(\ln(i(x,y))) + \Im(\ln(r(x,y))) \\ &\Im(z'(x,y)) = \Im(z(x,y)) \cdot \mathrm{H}(u,v) \end{aligned} \quad (5)$$

It will be necessary to perform the inverse process to obtain the filtered image in the spatial domain.

The low frequency components are associated with the illumination of the image and the high frequency ones with the reflectance of the image. So, to reduce the effects of changes in the illumination of the image, a high pass filter could be applied. We build this high pass filter from a low pass one in the next way:

$$\begin{aligned} &\mathrm{H}'_{hp}(u,v) = 1 - \mathrm{H}_{lp}(u,v) \\ &\mathrm{H}_{hp}(u,v) = (\alpha_h - \alpha_l) \cdot \mathrm{H}'_{hp}(u,v) + \alpha_l \end{aligned} \quad (6)$$

We have used two families of filters, Butterworth and Gaussian. Fig. 1 shows the Homomorfic Filter Transfer Function from a Butterworth filter. The transfer functions are as follows:

$$\begin{aligned} &D(u,v) = (u^2 + v^2)^{1/2} \\ &\mathrm{H}_{Butt}(u,v) = \frac{1}{1 + \left[D(u,v) / D_0 \right]^{2n}} \\ &\mathrm{H}_{Gauss}(u,v) = \exp\left(-\left[D(u,v) / D_0 \right]^2 \right) \end{aligned} \quad (7)$$

3 COMPRESSION TECHNIQUES USING PANORAMIC IMAGES

The map created is composed of a set of panoramic images of the environment. To reduce the computational cost, it is necessary to extract the most relevant information from the set of panoramic images. In this section, some techniques that allow us to obtain this information are outlined.

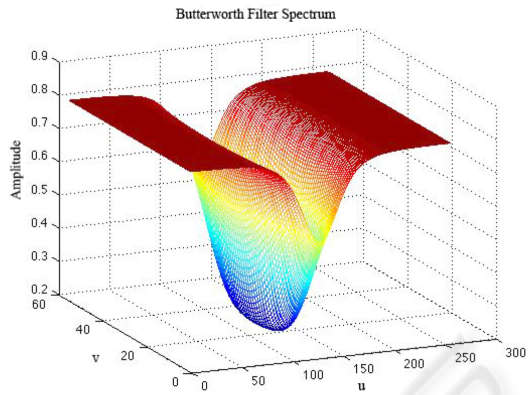

Figure 1: Transfer Function of Homomorfic Filter with Butterworth Filter.

3.1 PCA-based Techniques

As (Kirby, 2000) shows, it is possible to transform each image from a set of N images with M pixels each, $\vec{x}^j \in \Re^{Mx1}; j = 1 \ldots N$, into a vector with the K PCA features that contain the most relevant information, $\vec{p}^j \in \Re^{Kx1}; j = 1 \ldots N$, $K \leq N$. However, when we build the database in this way, it only contains information about the direction that the robot had when each image was captured, but not for all the possible orientations. Jogan and Leonardis (2000) present a method to include this orientation information, with the uniqueness that it is only necessary to acquire an image per position, and Payá, Fernández, Reinoso, Gil and Úbeda (2009) make use of it in a robot localization task, comparing to other techniques.

In brief, to construct the covariance matrix C, we obtain Q rotations from each image of the map. As we work with panoramic images, the covariance matrix of our data matrix $\mathbf{X} \in \Re^{Mx(QxN)}$, shall consist of a set of N blocks of size Q x Q:

$$\mathbf{X} = \left[\mathbf{X}^1 \middle| \mathbf{X}^2 \middle| \ldots \middle| \mathbf{X}^N \right]$$
$$\Rightarrow \quad C = X^T X = \begin{bmatrix} X^{11} & X^{12} & \ldots & X^{1N} \\ X^{21} & X^{22} & \ldots & X^{2N} \\ \vdots & \vdots & \ddots & \vdots \\ X^{N1} & X^{N2} & \ldots & X^{NN} \end{bmatrix} \quad (8)$$

The covariance matrix is composed of circulant blocks. This fact allows us to perform the SVD decomposition of C through Q decompositions of order N, thus reducing the computational cost of the compression.

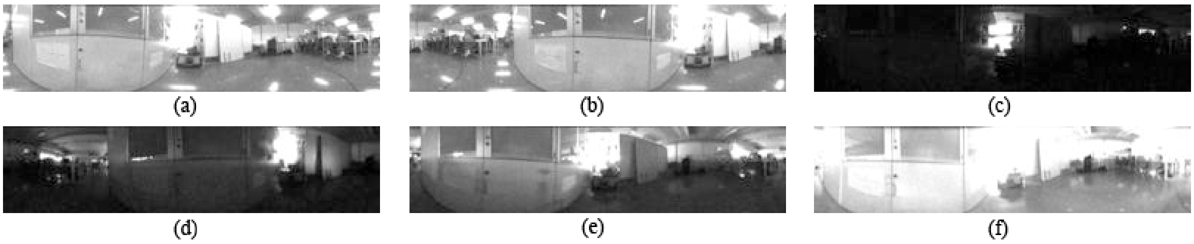

Figure 2: Sets of test images. (a) Test 1 (9:00, artificial light), (b) Test 2 (9:00, artificial light, 90 degrees rotation), (c) Test 3 (18:00, no light), (d) Test 4 (11:00, natural light, 90 degrees rotation), (e) Test 5 (13:00, daylight) and (f) Test 6 (16:00, daylight).

3.2 Fourier-based Techniques

When we have an image *f(x,y)* with N_y rows and N_x columns, we can obtain the relevant information of the image by applying the Discrete Fourier Transform. There are several possibilities, such as to implement the 2D Discrete Fourier Transform (Payá *et al*, 2009), (Rossi *et al*, 2008), the Spherical Fourier Transform of omnidirectional images or the Fourier Signature of the panoramic image (Menegatti *et al*, 2004).

The Fourier signature exploits better the invariance to ground-plane rotations in panoramic images. This transformation consists in expanding each row of the panoramic image $\{a_n\} = \{a_0, a_1, \ldots, a_{Ny-1}\}$ using the Discrete Fourier Transform into the sequence of complex numbers $\{A_n\} = \{A_0, A_1, \ldots, A_{Ny-1}\}$. The most important information is concentrated in the low frequency components of each row. It is possible to prove that if each row of the original image is represented by the sequence $\{a_n\}$ and each row of the rotated image by $\{a_{n-q}\}$ (being q the amount of shift), when the Fourier Transform of the shifted sequence is computed, we obtain the same amplitudes A_k than in the non-shifted sequence, and there is only a phase change, proportional to the amount of shift q (eq. 9).

$$\Im\left[\{a_{n-q}\}\right] = A_k e^{-j\cdot\frac{2\pi qk}{N_y}}; \qquad k = 0, \ldots, N_y - 1 \tag{9}$$

4 MAP BUILDING AND LOCALIZATION

In this section, we expose in general terms, how a dense map can be built, and how the location and orientation of the robot in it can be computed. (Payá *et al*, 2009) evidence that use of the Fourier signature of the image clearly outperforms PCA both in time consumption and in localization accuracy. Therefore to create the map and retrieval of the location, we use only the Fourier Signature of the image.

4.1 Map Building

To perform the experiment, we have captured a set of 101 omnidirectional images on a predefined grid of 40x40 cm in an indoor environment. We work with panoramic images with a size of 56x256 pixels. Once we have all the panoramic images, we used the Fourier signature described in the previous section. To test the validity of the maps constructed, we have captured several test images in some half-way points among those stored in the map. We have captured several sets of test images with changing illumination conditions and changing the position of some objects (Fig. 2). Fig. 3 (a) shows a bird's eye view of the grid used to capture the images to construct the map and an example of panoramic images.

4.2 Localization and Orientation Recovering

The objective is to calculate the position and orientation of the robot in the points where the test images where taken, under different lighting conditions, using only the visual information stored in the map.

To calculate the position and orientation of the robot for each test image, we calculate the Fourier transform (using the Fourier signature) and then, we calculate the Euclidean distance of the power spectrum of the test image with respect to the spectra stored in the map. The corresponding position of the robot is extracted as the best matching. Furthermore, the orientation is calculated with eq. 9.

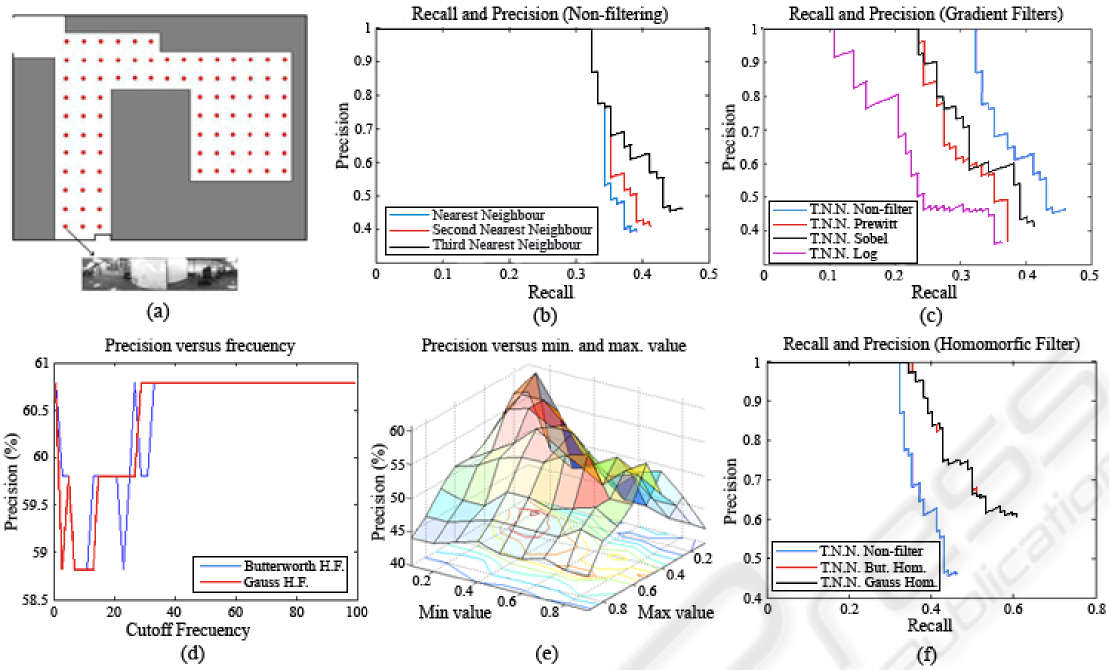

Figure 3: (a) Grid used to capture the set of training images, (b) Precision versus Recall localization without filtering, (c) Precision versus Recall localization using Gradient Filters, (d) Precision in terms of frequency of the filters, (e) Precision in terms of maximum and minimum value of the filter, (f) Precision versus Recall localization using Homomorfic Filters.

5 EXPERIMENTAL RESULTS

In this section the results obtained from experiments are presented. To perform the experiments, we have constructed the map by taking a total of 101 panoramic images. We have 6 sets with 17 test images each, taken at different times of day and under different conditions of illumination.

We will use *Recall and Precision* charts (Gil, Martinez, Ballesta and Reinoso, 2009) to compare the different methods of filtering. The parameters are defined as follows:

$$recall = \frac{\#correct\ \ matches\ \ retrieved}{\#total\ \ correct\ \ matches}$$
$$precission = \frac{\#correct\ \ matches\ \ retrieved}{\#matches\ \ retrieved} \tag{10}$$

For the data association we use the minimum Euclidean distance, through the descriptors Nearest Neighbour (N.N.), Second Nearest Neighbour (S.N.N.) and Third Nearest Neighbour (T.N.N.). Fig. 3 (b) shows the results obtained when we perform a localization process without prior filtering. We can observe the improvement that occurs when we use the descriptor T.N.N.

In Fig. 3 (c) we can observe how worse results are obtained when we apply gradient-based filters. It reduces the accuracy from 46.08% (no filter) to 41.18% (Sobel), 37.25% (Prewitt), or, in the worst case to 36.27% (Laplacian).

When working with homomorfic filters the parameters of the filter need to be adjusted previously. The homomorfic filter built using a Butterworth filter depends mainly on the cut-off frequency, the order of the filter and the maximum and minimum value of the filter. When we build the filter from a Gaussian filter, the most important parameters are the cut-off frequency and maximum and minimum value of the filter. As we can see in Fig. 3 (d) and Fig. 3 (e), both filters are more dependent on the maximum and minimum values, that the cut-off frequency.

After exhaustive tests, the optimal values for the parameters are a cut-off frequency of 50 Hz, Butterworth filter of order 3, homomorfic filter maximum value equal to 0.21 and minimum value equal to 0.20. Fig. 3 (f) shows how we can improve the accuracy of the location within the map, applying a homomorfic filter to it. In this case we have passed from an accuracy of 46.08% (no filter) to an accuracy of 60.78% with the homomorfic filter. We can see how the results obtained with the Gaussian filter are almost identical to those obtained using the Butterworth filter.

6 CONCLUSIONS

In this work, we have presented some methods for the creation of robust dense maps of real environments, using an appearance-based approach from previously filtered panoramic images.

We have presented two possible methods for filtering against illumination changes in the environment. As shown, the application of the first method (edge detection), not only does not improve but also worsens the results. On the other hand, applying a homomorfic filter on the panoramic image significantly improves the localization. Very similar results are obtained when constructing the homomorfic filter using a Gaussian filter or using a Butterworth filter. Furthermore, we have tuned the parameters of the filters to obtain a robust location against changes in illumination.

We have built the database by applying a compression of the visual information. We have used the Fourier signature due to the fact that it presents better results in terms of amount of memory and computation times needed to build the database. It is also important the fact that it presents orientation invariance and it allow us to compute the robot orientation. Finally, an important property is that the Fourier transform is an inherently incremental method. These properties make it possible to be applied in future works where robots have to add new information to the map and localize simultaneously in real time.

This work opens the door to the use of appearance-based methods with applications in mobile robots. As we have shown, the map created is robust against changes of lighting conditions, and it permits thus to recover the location and orientation of the robot in the map even if there are changes in the illumination of the scene.

ACKNOWLEDGEMENTS

This work has been supported by the Spanish government through the project DPI2007-61197. ‘Sistemas de percepción visual móvil y cooperativo como soporte para la realización de tareas con redes de robots’.

REFERENCES

Adini, Y., Moses, I., Ullman, S., 1997. Face recognition: the problem of compensating for changes in illumination direction. In *IEEE Trans.Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 721-732.

Bischof, H., Wildenauer, H., Leonardis, A., 2004. Illumination intensitive recognition using eigenspaces. In *ELSEVIER Computer Vision and Image Understanding,* Vol. 95, No. 1, pp. 86-104.

De Araujo, V., Maia, R., D’Angelo, M., D’Angelo, G., 2006. Automatic Plate Detection using genetic algorithm. In *Proc. World Scientific and Engineering Academy and Society,* pp. 43-48.

Faraid, H., Adelson, E., 1999. Separating reflections and lighting using independent components análisis. In *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition.* Vol. 1, pp. 1262-1267.

Gil, A., Martinez, O., Ballesta, M., Reinoso, O., 2009. A comparative evaluation of interest point detectors and local descriptors for visual SLAM. In *SPRINGER Machine Vision and Applications*.

Gonzalez, R.C., Woods, R.E., 1993. Digital Image Processing. Ed. Addison Wesley.

Jogan, M., Leonardis, A., 2000. Robust Localization Using Eigenspace of Spinning-Images. In *Proc. IEEE Workshop on Omnidirectional Vision*, Hilton Head Island, USA, pp. 37-44, IEEE.

Kröse, B., Bunschoten, R., Hagen, S., Terwijn, B. Vlassis, N., 2004. Household robots: Look and learn. In *IEEE Robotics & Automation magazine*. Vol. 11, No. 4, pp. 45-52.

Menegatti, E.; Maeda, T. Ishiguro, H., 2004. Image-based memory for robot navigation using properties of omnidirectional images. In *Robotics and Autonomous Systems*. Vol. 47, No. 4, pp. 251-276.

Murase, H., Nayar, S., 1994. Illumination planning for object recognition using parametric eigenspaces. In IEEE Trans. Pattern Anal. Mach. Intell., Vol 16, pp. 1219-1227.

Payá, L., Reinoso, O., Gil, A., Sogorb. J., 2008. Multi-robot route following using omnidirectional vision and appearance-based representation of the environment. In *Lecture Notes in Artificial Intelligence.* Vol. 5271, pp. 680-687 Springer.

Payá, L., Fernández, L., Reinoso, O., Gil, O., Úbeda, D., 2009. Appearance-based dense maps creation. Comparison of compression techniques with panoramic images. In *Int. Conf. on Informatic in Control, Automation and Robotic*, Vol. 6, pp. 238-246.

Rossi, F., Ranganathan, A., Dellaert, F., Menegatti, E., 2008. Toward topological localization with spherical Fourier transform and uncalibrated camera. In *Proc. Int. Conf. on Simulation, Modeling and Programming for Autonomous Robots*. Venice (Italy), pp. 319-330.

Ueonara, M., Kanade, T, 1998. Optimal approximation of uniformly rotated images: relationship between Karhunen-Loeve expansion and Discrete Cosine Transform. In *IEEE Trans. Image Processing*. Vol. 7, No. 1, pp. 116-119.