Adaptive Threshold for Robust Segmentation of Mobile Robots from Visual Information of their own Movement

C. Losada, M. Mazo, S. Palazuelos, F. Redondo

Departamento de Electrónica.
Universidad de Alcalá.
losada@depeca.uah.es

***Abstract*: In this work, a solution for robust motion segmentation of mobile robots is presented. Motion segmentation is obtained from the images acquired by a calibrated camera which is located in a fixed position in the environment where the robots are moving, and without incorporating invasive landmarks on board the robots. The proposal is based on the minimization of an objective function that depends on three groups of variables: the segmentation boundaries, the 3D rigid motion parameters (components of linear and angular velocity) and depth (distance to the camera). For the objective function minimization, we use a greedy algorithm which, after initialization, consists of three iterative steps. The accuracy in the results and also the processing time are closely related to the initial values of the involved variables. GPCA technique is used for curve initialization, comparing the reconstruction error with a threshold. Two approaches (fixed and adaptive) are proposed to set that threshold. The experimental tests carried out have proved that the proposed adaptive threshold increases, notably, the robustness of the system against lighting changes.**

I. INTRODUCTION

Motion segmentation is a basic task for the analysis of image sequences. It can be defined as the task of partitioning an image sequence into regions where the movement in the image plane is coherent along the sequence, without having prior knowledge about which pixels of the image move according to a specific motion model. Visual motion analysis in image sequences has a wide range of applications in computer vision, such as visual surveillance, mobile robotics, traffic monitoring, etc.

There are several methods for motion segmentation in an image sequence. Classical approaches to 2D motion segmentation are based on separating the image flow in different regions by looking for flow discontinuities [1][2] but as well as the aperture problem, these techniques have trouble dealing with noisy flow estimations. Traditionally, motion segmentation has been connected with motion detection, where each region corresponds to a particular motion model that explains temporal changes in that region [3][4]. Alternative approaches for motion segmentation are based on clustering [5][6].

Finally, the works [7] and [8] propose to obtain 3D rigid motion segmentation through the minimization of an objective function. This objective function depends on three groups of variables which are related with motion parameters. Minimization is carried out using a greedy algorithm which consists of three iterative steps. After initialization, these steps are repeated until convergence. As shown in [7] and [8], this method allows to obtain 3D motion segmentation. However, this algorithm presents several limitations. The accuracy of the results and also the number of iterations until convergence are closely related to the initialization, and the constants that weight the contribution of each term to the objective function. Moreover, the proposal in [8] is a slow algorithm. The number of iterations until convergence is between 700 and 10.000.

This work has been structured as follows: section II presents the proposed system for 3D motion segmentation and 3D localization of mobile robots. In this section we explain the need to initialize the involved variables (segmentation contours and depth). Curve initialization process is described in section III. In section IV, the proposed adaptive threshold is presented. Section V shows several experimental results and, finally, section VI presents the main conclusions.

II. PROPOSED SYSTEM

This paper proposes a robust system that allows the motion segmentation and 3D localization of mobile robots. Motion segmentation is obtained from the images acquired by a calibrated camera which is located in a fixed position in the environment. Fig. 1 shows a general block diagram of the proposed system.

Using the work in [7] as a starting point, the motion segmentation and 3D localization are obtained through the minimization of an objective function. This objective function (which is shown in equation (1)) is obtained from the 3D brightness constraint for rigid objects defined in the work of Sekkati and Mitiche [7].

$$E\left[\{\gamma_k\}_{k=1}^{N-1},\{\mathbf{v}_k\}_{k=1}^{N},\{\boldsymbol{\omega}_k\}_{k=1}^{N},Z\right]=$$
$$=\sum_{k=1}^{N}\left[\int_{\Omega_k}\psi_k^2(\mathbf{x})d\mathbf{x}+\mu\int_{\Omega_k}g\left(\|\nabla Z\|\right)d\mathbf{x}\right]+\sum_{k=1}^{N-1}\lambda\oint_{\gamma_k}ds \qquad (1)$$

In equation (1) ψ_k is the 3D brightness constraint for the pixels inside the curve k (γ_k); λ and μ are positive, real constants to weigh the contribution of the terms in equation (1) and $\nabla=(\partial_x,\partial_y)$ is the spatial gradient operator. This objective function depends on three groups of variables:

1. A set of N-1 curves which defines the N-1 mobile robot segmentation boundaries $\{\gamma_k\}_{k=1}^{n-1}$. These curves divide the image in N regions $\{\Omega_k\}_{k=1}^{N-1}$ (where Ω_k represents the set of points of the image inside the curve k γ_k).
2. The components of linear and angular velocity of (N-1) mobile robots and background $\{\mathbf{v}_k,\boldsymbol{\omega}_k\}_{k=1}^{n}$.
3. The depth (distance from each 3D point (**P**) to the camera) which is the Z_c coordinate of point **P.**

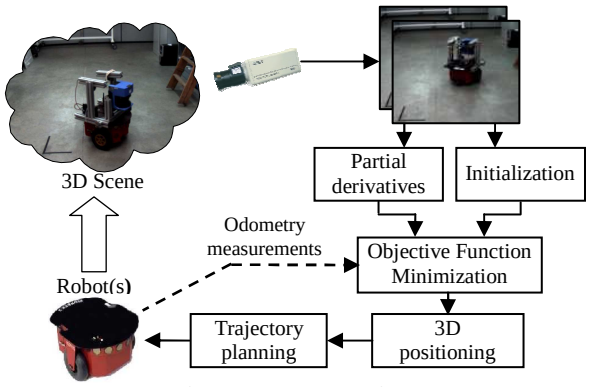


Fig. 1. Block diagram of the proposed system for motion segmentation of mobile robots.

The objective function minimization is carried out using a greedy algorithm which consists of three iterative steps. After curve and depth initialization, the three steps are repeated until algorithm convergence. The accuracy of the results and also the processing time are closely related to initial curves and depth.

In this work, depth (Z_c) is initialized using the measurements given by the odometric sensors on board the robots, if available. Otherwise, depth is initialized to a constant over the image domain.

Regarding the curves, there are several alternatives for curve initialization. In [7], circumferences are used for curve initialization. But this approach requires a high number of iterations until convergence. In order to reduce the number of iterations (and, therefore, the processing time) of the segmentation algorithm in this work, GPCA (Generalized Principal Components Analysis) technique is used [9][10] to obtain the initial set of N-1 curves which define the segmentation boundaries. Curve initialization algorithm is described in section III.

After motion segmentation, robot identification is carried out by comparing the estimated linear and angular velocities in each region to the velocities measured by the odometric sensors on board the robot. Then, 3D position is obtained by projecting the points inside the robot from the image plane to the 3D space using the estimated depth (Z_c)

III. Curve Initialization

Curve initialization using GPCA [9] consists of two different stages. In the first stage, GPCA is used to obtain the background model from N_i images $\{\mathbf{I}_j\}_{j=1}^{N_i}$ without any mobile robot. Background model is represented using two GPCA transformation matrices (**L** and **R**). This stage is carried out only once (because the camera is located in a fixed position) and can be executed off-line.

In the second stage, each image is compared to the background model (that has been obtained previously in the off-line stage). It allows to detect which pixels in the image belong to mobile robots (or other objects) that have appear in the scene after obtaining the background model. In this stage, each image is projected to the GPCA space using the matrices **L** and **R** which have been obtained previously in the off-line stage. The transformation equation is shown in (2). Then, the original image is reconstructed using equation (3). In the two equations, **M** is the mean of the N_i images used to obtain the background model.

$$\mathbf{I}_T=\mathbf{L}^T(\mathbf{I}-\mathbf{M})\mathbf{R} \qquad (2)$$

$$\mathbf{I}_R=\mathbf{L}\mathbf{I}_T\mathbf{R}^T+\mathbf{M} \qquad (3)$$

The reconstruction error is defined as the difference between the reconstructed image (I_R) and the original image (I). Even though this error can be calculated subtracting the images pixel-to-pixel, this approach is not robust against noise. Therefore, we use a set of pixels (windows) around each pixel in the image and we obtain the reconstruction error for these windows. In this case, square windows with dimensions qxq pixels have been defined. These windows are called $\boldsymbol{\Phi}_{wi}$ in the original image an $\hat{\boldsymbol{\Phi}}_{wi}$ in the reconstructed image. The reconstruction error associated to the central pixel in the window (whose coordinates are (w, i)) is calculated using equation (4).

$$\varepsilon_{wi}=\left\|\boldsymbol{\Phi}_{wi}-\hat{\boldsymbol{\Phi}}_{wi}\right\| \qquad (4)$$

Pixels whose reconstruction error (calculated using equation (4)) is higher than a threshold are candidate to belong to a mobile robot, because in those pixels there is an important difference between the actual image and the background model. A block diagram that includes the different stages involved in curve initialization using GPCA is shown in Fig. 2.

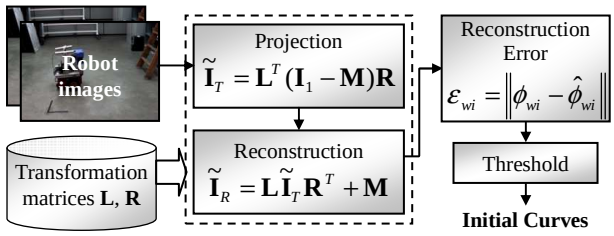
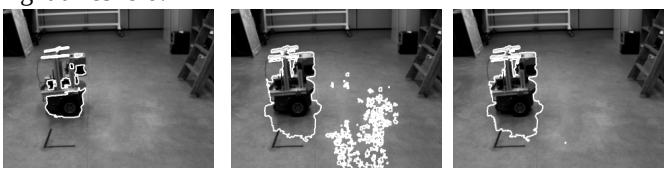


Fig. 2. Block diagram of the proposed method for curve initialization using GPCA

In order to decide which pixels belong to mobile robots, it is necessary to set correctly the threshold, because it determinates the quality of initial curves. If the threshold is too high, some pixels on the robot can be considered as background pixels, as is shown in Fig. 3.a. On the contrary, if the threshold is too low, background pixels can be included as pixels of the mobile robot (this situation is shown in Fig. 3.b). Finally, Fig. 3.c depicts the initial curve obtained using the right threshold.



(a) (b) (c)

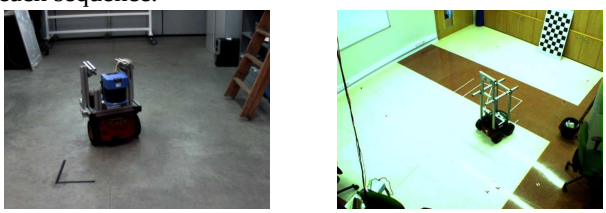
Fig. 3. Initial curves obtained for different reconstruction error thresholds (a) Threshold too high. (b) Threshold too low. (c) Right threshold.

Robot shadow is removed in a subsequent stage after projecting the image to a lighting invariant space which is described by Finlayson et al in [11].

IV. Threshold of the Reconstruction Error

Since the importance of using a correct threshold has been proved, in this work two approaches to set the threshold are proposed. In both proposals, the threshold is obtained as a function of the current image and N_i background image features.

In order to set the threshold, two different image sequences containing a mobile robot have been used. This image sequences have been acquired using a calibrated camera, located in a fixed position. Fig. 4 shows one image belonging to each sequence.



(a) (b)

Fig. 4. Images belonging to the sequences used in the experimental tests. (a) Sequence 1. (b) Sequence 2

A. Fixed threshold

The value of the fixed threshold depends on the value of the reconstruction error (4) calculated over the background pixels. It is realistic to assume that this reconstruction error has similar values in background images and in images that contain mobile robots. Under this assumption, threshold is set up as a function of the reconstruction error calculated for the N_i background images. These background images are the ones that have been used to obtain the background model using GPCA. For that purpose, after obtaining the GPCA transformation matrices (**L** and **R**) that define the background model, each of the N_i background images is projected (using equation (2)) and reconstructed (using equation (3)). Then, the reconstruction error is calculated using equation (4). The maximum recovery error (for the N_i background images) is obtained for each pixel with coordinates (i, j). This maximum error for each image pixel is defined in equation (5) for an image with dimensions mxn. Finally, the maximum reconstruction error ($E^2_{r,\,fondo}$) is obtained using equation (6).

$$\varepsilon^2_{r,background}(i,j) = \max\left(\left\{\varepsilon^2_{r,k}(i,j)\right\}_{k=1}^{N_i}\right) \quad \begin{matrix} i=1,2,\ldots,m \\ j=1,2,\ldots,n \end{matrix} \tag{5}$$

$$E^2_{r,background} = \max\left(\left\{\varepsilon^2_{r,background}(i,j)\right\}_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}\right) \tag{6}$$

In a first approximation, the threshold is related to the maximum reconstruction error, obtained from the N_i background images, through a constant K. It means that the fixed threshold is proportional to the maximum reconstruction error as shown in equation (7).

$$U_1 = K \cdot E^2_{r,background} \quad K \in \Re,\, K \geq 1 \tag{7}$$

Several experiments have been carried out to set the optimum value of K for any sequence (which belong the images shown in Fig. 6.a and b). In these experimental tests, 25 images of each sequence have been chosen. Then, for each image, the optimum value of K has been obtained manually. We have considered a value of K optimum when, using the threshold defined in equation (7) all the pixels on the mobile robot are inside the obtained initial curves, but background pixels are not included in them. Table I shows the optimum values of K, and also the obtained threshold (using equation (7)) for 5 images which belong to sequences 1 and 2.

TABLE I
OPTIMUM VALUE OF CONSTANT K AND FIXED THRESHOLD OBTAINED FOR 5 IMAGES OF THE SEQUENCES 1 AND 2.

Image	Sequence 1 (Fig. 6Fig. 4.a)		Sequence 2 (Fig. 4.b)	
	K	Threshold	K	Threshold
100	6.15	598.95	5.1	437.84
200	5.05	491.82	1.75	150.24
300	1.95	189.91	9.05	776.94
400	2.60	253.19	9.7	832.75
500	4.75	462.56	4.35	373.45

The curves which have been obtained manually are compared with the curves obtained for different values of K. To compare the curves, the percentage of pixels correctly initialized has been measured. In the experimental tests, we have proved that a value of K=10 is suitable for both sequences. Table II shows the percentage of pixels that have been correctly initialized for 3 images belonging to each sequence. It can be seen that the percentage is higher than 99% for all the test images.

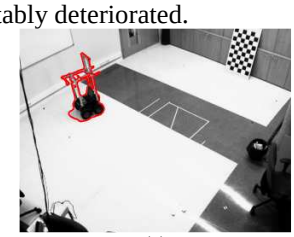
It is worth highlighting that, if lighting is controlled, the value of K is not critical, because the reconstruction error in pixels that belong to the mobile robot is much higher than in background pixels. However, the fixed threshold (defined in equation (7)) is not robust against lighting changes. This is because, as lighting conditions change, the intensity of the image pixels also changes. For that reason, the reconstruction

error in the background pixels increases, and the difference between the reconstruction error of the pixels on the robot and of background pixels decreases.

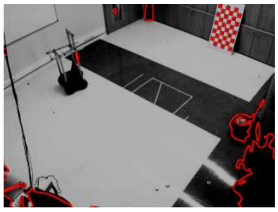
TABLE II
PERCENTAGE OF PIXELS THAT HAS BEEN CORRECTLY INITIALIZED USING THE THRESHOLD DEFINED IN EQUATION (7) WITH K=10.

Image	% de pixels correctly initialized	
	Sequence 1 (Fig. 4.a)	Sequence 2 (Fig. 4.b)
120	99.4486 %	99.9808 %
250	98.5384 %	99.9414 %
400	99.4264 %	99.8675 %

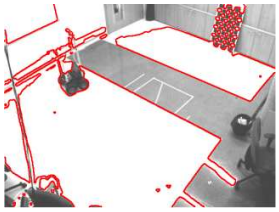
Fig. 5 shows three images with different lighting conditions. It can be seen that, if lighting decreases (Fig. 5.b) or increases (Fig. 5.c) the initial curves calculated with a fixed threshold are notably deteriorated.



(a)



(b)



(c)

Fig. 5. Initial curves obtained for an image that belongs to sequence 1. (a) Original image. (b) Original image after subtracting a value of 50 to each pixel in the original image. (c) Original image after adding a value of 50 to each pixel in the original image.

B. Adaptive threshold

In order to improve the system robustness against lighting changes we propose to adjust the threshold if there are important changes in the intensity of the images.

As the fixed threshold, the proposed adaptive threshold is proportional to the maximum reconstruction error of the background images (equation (5)). However, to obtain the adaptive threshold, this maximum error is multiplied by a function that depends on the absolute value of the difference between the mean intensity of the actual image and the mean intensity of the N_i background images. The adaptive threshold is defined in equation (8).

$$U_2 = \left(\alpha + \beta \left| I - I_{mean}^{background} \right|\right) \cdot \mathrm{E}^2_{r,background} \qquad (8)$$
$$\alpha, \beta \in \Re,\ \alpha, \beta \geq 0$$

In equation (8), α and β are real positive constants. The optimum values of these constants have been obtained experimentally. The initial curves have been obtained manually and are compared to the curves obtained with different values of α and β for images with lighting changes in order to set the best values for α and β. We evaluate the percentage of pixels correctly initialized, and also the range of intensity values for which the initialized curves are right. We consider that a curve is right if the percentage of pixels correctly initialized is greater than 98%. The results obtained for some values of α and β are shown in Table III. All the results shown in this table have been obtained using the images of the sequence 2 (Fig. 4.b).

TABLE III
POSSIBLE VARIATIONS IN THE INTENSITY OF THE ORIGINAL IMAGE WITHOUT DETERIORATING THE INITIAL CURVES FOR DIFFERENT VALUES OF α AND β.

	β = 1.0	β = 1.2	β = 1.4	β = 1.6	β = 1.8
α=0.6	70.6	72.4	75.2	72.8	81.8
α=0.8	66.6	72.0	69.4	72.8	81.2
α = 1	70.6	85.0	53.8	43.2	38.4
α = 3	59.4	55.0	42.4	41.8	36.2
α = 5	52.0	45.2	43.0	35.4	35.2
α = 7	39.4	36.0	41.2	33.8	29.0
α = 9	32.4	35.2	40.2	33.4	28.6
α = 11	32.2	35.6	39.6	32.6	27.8

As it was previously mentioned, shadows inside the initial curves are removed in a subsequent stage. For this purpose, the method proposed in [11] is used.

V. EXPERIMENTAL RESULTS

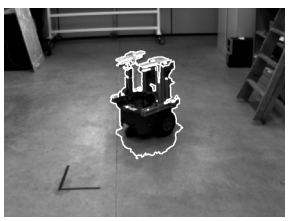
In order to validate the proposal in this work, several experimental tests have been carried out. In these experiments four different image sequences have been used. All these image sequences have been acquired by a calibrated camera located in a fixed position. Image sequences 1 and 2 have been previously shown in Fig. 4. Fig. 6 shows one image belonging to sequences 3 and 4. It is worth pointing out that image sequences 3 and 4 have been acquired by the same camera than image 2, but under different lighting conditions.

(a)

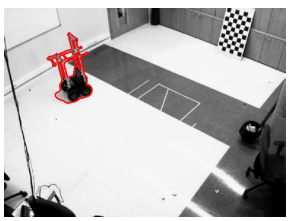
(b)

Fig. 6. Images belonging to the sequences used in the experimental tests. (a) Sequence 3. (b) Sequence 4.

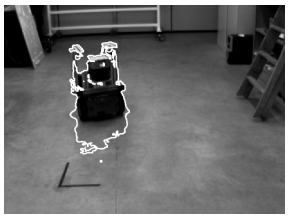
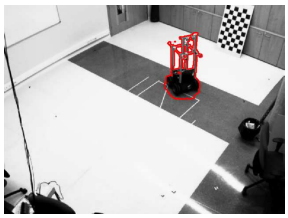
Fig. 7 shows the initial curves obtained by applying the fixed threshold defined in equation (7) to the reconstruction error to two images belonging to sequences 1 and 2 (Fig. 6.a y b).



(a) Image 120 of the sequence 1



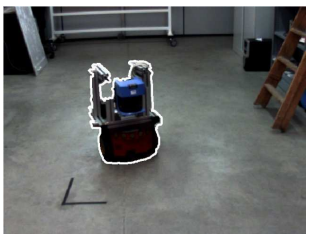
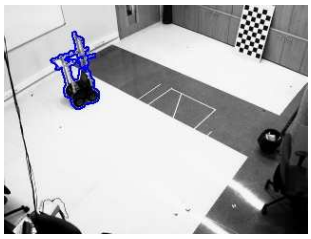
(b) Image 120 of the sequence 2

(c) Image 250 of the sequence 1 (d) Image 250 of the sequence 2

Fig. 7. Initial curves obtained using the proposed, fixed threshold defined in equation (7) to the reconstruction error.

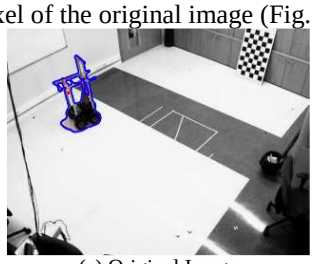
Fig. 7 shows that the initial curves obtained using the fixed threshold are close to the real contours of the mobile robots. As it was previously mentioned, shadows inside the initial curves are removed in a subsequent stage using the method proposed in [11]. Fig. 8 shows the initial curves obtained after removing the shadows. It is worth highlighting that, even including the shadows, the initial curves obtained using GPCA are closer to the real contours than the circumferences proposed in [7]. It allows to carry out the motion segmentation with a lower number of iterations [10].

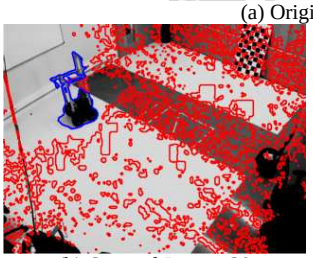
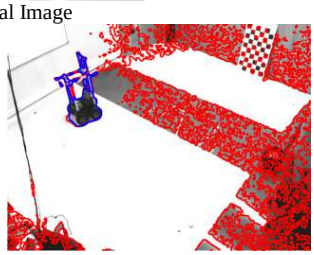
(a) (b)

Fig. 8. Initial curves obtained after removing the shadows (a) Image 120 of the sequence 1 (Fig. 6.a) (b) Image 250 of the sequence 2 (Fig. 6.b)

Fig. 9 shows the initial curves obtained for the image 120 of the sequence 2 (Fig. 6.b). Lighting changes have been simulated by adding (Fig. 9.b) or subtracting (Fig. 9.c) a value of 30 to each pixel of the original image (Fig. 9.a).



(a) Original Image

(b) Original Image -30 (c) Original Image +30

Fig. 9. Initial curves obtained for the image 120 of the sequence 2 (Fig. 6.b). Obtained using the fixed threshold (7) (indicated in red color), and obtained using the adaptive threshold (8) (indicated in blue color).

In Fig. 9 we can observe that the proposed adaptive threshold (equation (8)) has a better behavior against lighting changes than the fixed threshold (equation (7)).

The initial curves obtained using the fixed threshold (equation (7)), and using the adaptive threshold (equation (8)) have been compared with the curves obtained manually. For these experiments we have used 25 images belonging to sequences 1 and 2. Lighting changes have been simulated by adding or subtracting a constant value to all the pixels in the original image. Fig. 10 shows the percentage of pixels correctly initialized for each value added between -50 and +50. The results obtained using the fixed threshold are indicated in black color and the results obtained using the adaptive threshold are indicated in red.

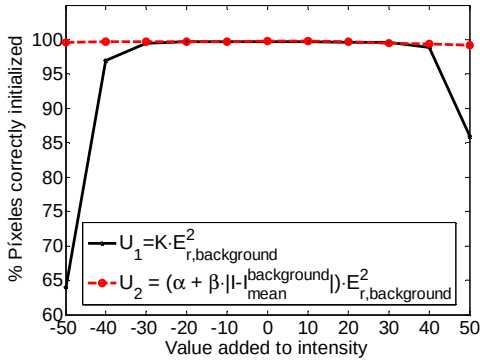
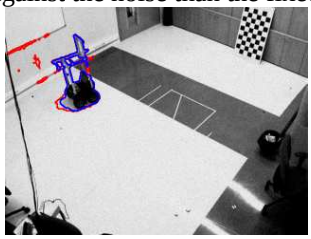
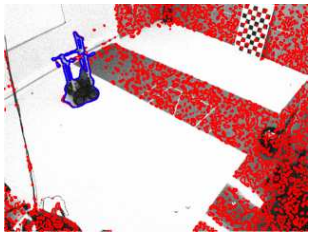


Fig. 10. Percentage of pixels correctly initialized. Using the fixed threshold (7) (in black) and using the adaptive threshold (8) (in red).

In Fig. 10, we can observe that if the value added (or subtracted) to the intensity of the original image is lower than 25, both thresholds have similar results. However, if the value added to the intensity is higher than 25, the percentage of pixels correctly initialized using the fixed threshold (equation (7)) decreases notably, whereas the adaptive threshold allows to obtain the right initial curves (the percentage of pixels correctly initialized is greater 99%) even for a value 50 added/subtracted to all the pixels of the original image.

Moreover, several experimental tests have been carried out to determine the effect of noise in the images. In that sense, the original images have been modified by adding to each pixel a random value between 0 and 30 (Fig. 11.a) or 0 and 50 (Fig. 11.b). The initial curves are shown in Fig. 11. This figure shows that the adaptive threshold is more robust against the noise than the fixed one.

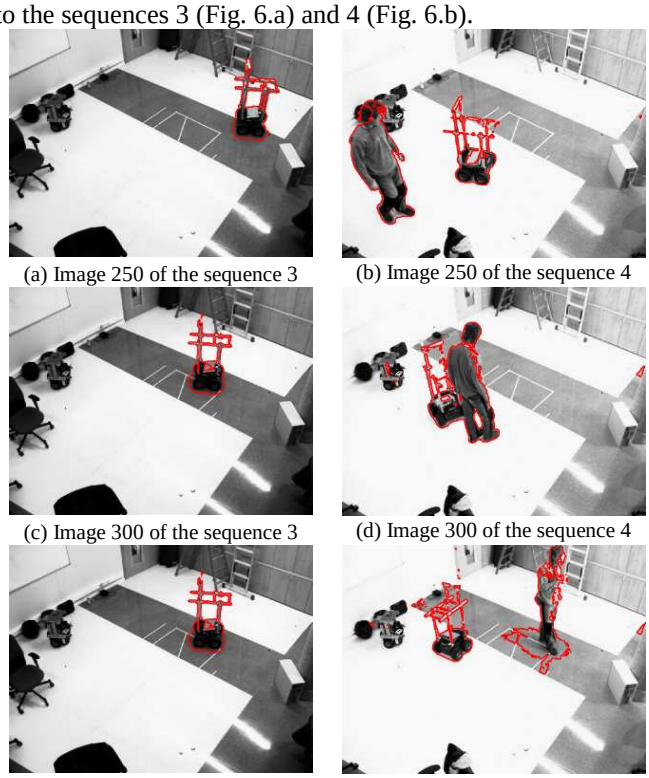



(a) (b)

Fig. 11. Initial curves obtained applying the fixed threshold (indicated in red) and the adaptive threshold (in blue) to the reconstruction error in the presence of noise. (a) After the addition of a random value between 0 and -30 to each pixel. (b) After the addition of a random value between 0 and 50 to each pixel.

In order to evaluate the robustness of the adaptive threshold proposed in this work, the images belonging to the sequences 3 and 4 (which are shown in Fig. 6.a and b) have been used. These image sequences have been acquired by the same camera than the sequence 2 (Fig. 4.b), but with different lighting conditions. The images of sequence 4,show a mobile robot and a person moving in the intelligent space. α and β values used for this sequences are the same values that have been previously set using the sequence 2.

Fig. 12 shows the initial curves for some images belonging to the sequences 3 (Fig. 6.a) and 4 (Fig. 6.b).



(a) Image 250 of the sequence 3 (b) Image 250 of the sequence 4

(c) Image 300 of the sequence 3 (d) Image 300 of the sequence 4

(e) Image 350 of the sequence 3 (f) Image 350 of the sequence 4

Fig. 12. Initial curves obtained for different images of the sequences 3 and 4. These curves have been obtained using the adaptive threshold proposed in this work and defined in equation (8).

Images in Fig. 12 show that the initial curves obtained using the proposed threshold are close to the real contour of the mobile robot in spite of the changes in the lighting conditions. This threshold increases the robustness of the motion segmentation algorithm against lighting changes.

VI. Conclusions

This work has presented a system for 3D motion segmentation, and 3D positioning of mobile robots using a calibrated camera that is placed in a fixed position in the environment.

Motion segmentation and positioning is obtained minimizing an objective function. Before the minimization it is necessary to initialize the *N*-1 curves that define the boundaries of the segmentation and also the depth. In this work, initial curves are obtained using the GPCA technique.

Two approaches have been proposed to set the threshold of GPCA reconstruction error in the curve initialization algorithm.

Several experimental tests have been carried out. These experiments have proved that the fixed threshold proposed works correctly in controlled environments where lighting conditions do not change. The proposed adaptive threshold increases notably the system robustness against lighting changes with respect to the fixed threshold. It has been proved with four different image sequences that have been acquired in different lighting conditions.

Acknowledgments

The work described in this paper was supported by the Ministry of Science and Technology (MEC) under RESELAI project (REF-TIN2006-14896-C02-01).

References

[1] Black, M. & Anandan, P.; Robust dynamic motion estimation over time. Proc. of the IEEE International Conference on Computer Vision and Pattern Recognition, pp. 296-302. June 1991.

[2] Yan, H.; Tjahjadi, T.; Multiple motion segmentation through a highly robust estimator. IEEE International Conference on Systems, Man and Cybernetics, pp. 3082-3087. Oct. 2004

[3] Darrell, T.; Pentland, A. “Robust estimation of a multi-layered motion representation”. Proc. of the IEEE Workshop on Visual Motion, pp. 173-178. 1991.

[4] Weiss, Y. A unified mixture framework for motion segmentation: incorporating spatial coherence and estimating the number of models. Proc. of the IEEE International Conference on Computer Vision and Pattern Recognition. pp. 321-326. 1996.

[5] Costeira, J. & Kanade, T. A multibody factorization method for independently moving objects. International Journal of Computer Vision. Vol. 29, N. 3, pp. 159–179. September 1998.

[6] Kanatani, K. Motion Segmentation by Subspace Separation and Model Selection. Proceedings of the 8th IEEE International Conference on Computer Vision 2001. Vol. 2, pp. 586-591.

[7] Sekkati, H. & Mitiche, A. Concurrent 3D Motion Segmentation and 3D Interpretation of Temporal Sequences of Monocular Images. IEEE Trans. on Image Processing, Vol. 15, No. 3, pp. 641-653, 2006.

[8] Sekkati, H. & Mitiche, A. Joint Optical Flow Estimation, Segmentation, and Interpretation with Level Sets. Computer Vision and Image Understanding, Vol. 103, No. 2, pp. 89-100, 2006.

[9] Ye, Jieping; Janardan, Ravi; Li, Qi. GPCA: an efficient dimension reduction scheme for image compression and retrieval. Proc. of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining. pp: 354-363. 2004.

[10] Losada, C., Mazo, M., Palazuelos, S. & Redondo, F. Segmentación robots móviles en espacios inteligentes utilizando técnicas GPCA y minimización de funciones de energía. XV Seminario Anual de Automática, Electrónica Industrial e Instrumentación (SAAEI'08). ISBN: 13:978-84-96997-04-2. 2008.

[11] Finlayson, G.D., Drew, M.S. & C.Lu. “Intrinsic Images by Entropy Minimization”. Proc. 8th European Conf. on Computer Vision. Page(s): 582-595. 2004