Mobile Robot Localization through Identifying Spatial Relations from Detected Corners

Sergio Almansa-Valverde[1], José Carlos Castillo[1],
Antonio Fernández-Caballero[1,2], José Manuel Cuadra Troncoso[3],
and Javier Acevedo-Rodríguez[4]

[1] Instituto de Investigación en Informática de Albacete (I3A), n&aIS Group,
Campus Universitario s/n, 02071-Albacete, Spain
[2] Departamento de Sistemas Informáticos, Universidad de Castilla-La Mancha,
Campus Universitario s/n, 02071-Albacete, Spain
[3] Departamento de Inteligencia Artificial, E.T.S.I. Informática,
Universidad Nacional de Educación a Distancia, 28040-Madrid, Spain
[4] Universidad de Alcalá,
Departamento de Teoría de la Señal y Comunicaciones,
Alcalá de Henares, Spain

Abstract. In this paper, the Harris corner detection algorithm is applied to images captured by a time-of-flight (ToF) camera. In this case, the ToF camera mounted on a mobile robot is exploited as a gray-scale camera for localization purposes. Indeed, the gray-scale image represents distances for the purpose of finding good features to be tracked. These features, which actually are points in the space, form the basis of the spatial relations used in the localization algorithm. The approach to the localization problem is based on the computation of the spatial relations existing among the corners detected. The current spatial relations are matched with the relations gotten during previous navigation.

1 Introduction

A mobile robot must possess the capacity of self-localization while navigating in an environment [1]. An appearance-based approach for place recognition involves matching scenes based on selected features observed within the current local map or sensor view. The combination of a location and descriptor vector is termed a key point [3]. Place recognition then becomes a matter of identifying places by associating key points, or deciding that a place has not previously been seen. In a paper [16] the performance of a variety of corner (point) detecting algorithms for feature tracking applications is assessed. The overall observation of the results suggest that the Harris corner detector [9] is very suitable for tracking features in long sequences. Also, the overall empirical results revealed that the Kanade-Lucas-Tomasi (KLT) [17] and Harris detectors provided the best quality corners (qualitatively and quantitatively). In a recent review [12], detectors and local descriptors of local features for computer vision are described in a comprehensive way.

Time-of-flight (ToF) range cameras give depth information per pixel which make them ideal for background foreground segmentation, as in general the depth defines the subject from background in a much more basic way than the light intensity does [8], [14], [6]. Intensity images are on the other hand affected by colors, lighting, reflections and shadows in almost every normal scenario [13]. Thanks to the larger vertical field of view of ToF cameras, difficult obstacles (like tables) are better detected by a ToF camera than by a 2D laser scanner [18]. Most of the applications extract planar regions using both intensity and depth images. In a paper [15] different methods are explored to improve pose estimation. The normal of the extracted planes is also used [10] to detect badly conditioned plane detection, as horizontal planes in a staircase. Also a corner filtering scheme combining both the intensity and depth image of a ToF camera has been proposed [7].

Our approach uses the Harris corner detection algorithm, similarly to [5], [4], mainly as its computational cost is lower than other approaches like SIFT [11]. In this case, the ToF camera is exploited as a gray-scale camera for localization. The ToF camera allows to create a gray-scale image representing distances for the purpose of finding good features to be tracked. These features, which actually are points in the space, form the basis of the spatial relations used in the localization algorithm. The approach to the localization problem is based on the computation of the spatial relations existing among the aforementioned points in the space. The current spatial relations are matched with the ones created in previous iterations of the algorithm during autonomous navigation. The robot position is estimated in accordance to the localization information provided through the identified spatial relations.

2   Mobile Robot Localization Algorithm

To estimate the robot position in the environment, only the information provided by the ToF camera is used. In this case, the images provided by the ToF camera are considered as traditional images. That is, the depth information provided is considered as gray-level values of a traditional image. From these image pixel values, a series of characteristic points (or corners) are extracted to create spatial relations to be placed in the map that the robot has been previously given. When new spatial relations are created from detected corners, they are compared to the previously created ones and identified to calculate their location. A flow diagram of the localization algorithm is presented in Fig. 1.

2.1   Detection of Corners

To exploit the visual characteristics of the ToF camera, a gray-scale image is created from the distance information, I_d. The ToF camera provides as output a matrix where each position represents the three-dimensional coordinates (x, y, z) (in meters) of a system. Here the camera is the origin of coordinates, x varies along the horizontal axis, y varies along the vertical axis and z is the distance
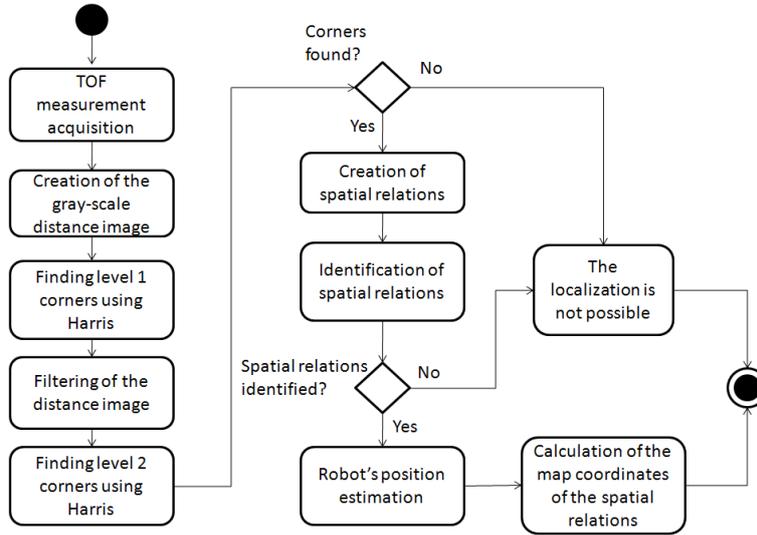
Fig. 1. Flow diagram of the localization algorithm

from the plane defined by the x and y axes. In this case, the gray level of each image pixel depends on the z value (depth) of the coordinates provided by the ToF camera for this point. Therefore, the image represents the scene using darker gray levels for near objects, and lighter levels for distant objects. Creating an image representation of the distance measurements enables the use of corner detection algorithm. Concretely, the Harris algorithm has been used for this purpose.

Applying Harris algorithm to the distance image results in a list of corners. Sometimes the difference between distances (and therefore, between gray levels) in the image is too slight, which makes it difficult to find enough significant corners. A hierarchy of two levels of corners is implemented to tackle this problem. The first level of corners is composed by those found in the initial distance image, I_d. Usually not many corners are found due to the low contrast of these images, although the gotten corners are quite resistant to noise. Some filters are performed on image I_d to get the second level of corners. A first filter equalizes the image histogram to enhance the contrast. After that, as noise is also enhanced with the equalization, a smoothing Gaussian is applied to the equalized result, obtaining I_f. In order to implement Gaussian smoothing, a transformation matrix is created from the values of the Gaussian distribution, as follows:

$$G(x,y) = \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

where x and y are the coordinates of the value in the Gaussian distribution and σ is the standard deviation. Matrix $G(x,y)$ is applied to every pixel of the image, setting the new value of each pixel to a weighted average of its neighboring pixels.

Using Harris algorithm on I_f returns more corners than on I_d, but they will be less resistant to noise. Hence, spatial relations based on corners belonging to level 1 are more trusted. So, they have higher priority than those based on corners belonging to level 2. But the last ones are indispensable for a correct localization as level 1 rarely contains enough corners to achieve a good localization. Fig. 2a and Fig. 2b show the corners found at levels 1 and 2, respectively. Each corner possesses the coordinates with respect to the image, $(x_{ToF}, y_{ToF}, z_{ToF})$ from the camera.
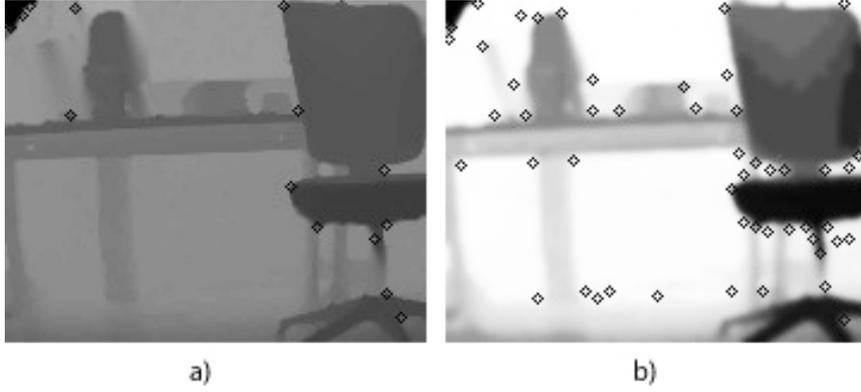

a)   b)

Fig. 2. Hierarchy of corners. (a) Level 1 corners on non-filtered image I_d. (b) Level 2 corners on filtered image I_f.

2.2   Creation of Spatial Relations

Once the corners are extracted, spatial relations connecting every pair of points from the same level are established. Some information is associated to these spatial relations in order to define and to identify them. The spatial relations information contains the following attributes: priority, according to the level of the corners belonging to the spatial relation, and distance in meters between the two points, N, calculated as follows:

$$N = \sqrt{(x_{ToF_1} - x_{ToF_2})^2 + (y_{ToF_1} - y_{ToF_2})^2 + (z_{ToF_1} - z_{ToF_2})^2} \qquad (2)$$

where the first corner has $(x_{ToF_1}, y_{ToF_1}, z_{ToF_1})$ as camera coordinates, and the second one has coordinates $(x_{ToF_2}, y_{ToF_2}, z_{ToF_2})$. The third and fourth attributes are the slopes between x and y coordinates and between x and z coordinates of the vector connecting the two corners in the space. The slope between x and y is calculated as follows:

$$S_y = \frac{x_{ToF_1} - x_{ToF_2}}{y_{ToF_1} - y_{ToF_2}} \qquad (3)$$

And the slope S_z between x and z is calculated in a similar manner. The inclination of the vector in the space is represented with these two slope values

(the slope between y and z, S_x, can be calculated from the other two slopes). There is enough information to identify spatial relations comparing the last three attributes.

2.3   Identification of Spatial Relations

The next step in the localization algorithm is the identification of new spatial relations to calculate their location in the robot's map (that is, if some points have known positions in the map, new points can also be placed if there are spatial relations between them), to finally estimate the robot position. For this purpose, the attributes of distance and slope are compared to identify the spatial relations.

When identifying a spatial relation it is not necessary to find another identical one, but similarity tolerance values τ_{si} and τ_{sl} in the distance between the points and in the slopes, respectively, have been included. These tolerance values are mandatory because the ToF camera usually provides noisy distance information; so, different measurements of the same scene would result on slightly different distance information, thus affecting corners and spatial relations.

Being R_1 a spatial relation obtained from two points of the current observation and R_2 a spatial relation from two previously observed points, the matching between R_1 and R_2 can be formulated as follows:

$$R_1 = R_2 \Leftrightarrow \begin{cases} N_1 - N_2 < \tau_{si} \\ S_{y1} - S_{y2} < \tau_{sl} \\ S_{z1} - S_{z2} < \tau_{sl} \end{cases} \qquad (4)$$

being N_1, S_{y1}, S_{z1} the attributes previously explained of R_1, and N_2, S_{y2}, S_{z2} the attributes of R_2.

The number of spatial relations grows with the size of the environment, raising the probability of creating very similar spatial relations. This might cause erroneous identification of spatial relations and, in consequence, lead to an erroneous robot localization. To minimize identification errors, the number of identifications is restricted to a maximum of 10, although the number of spatial relations is not limited. The robot position is estimated multiple times using them one by one. Next, Chebyshev's inequality [2] is applied to the estimated robot positions to discard outliers caused by erroneous identifications. Finally, the median value (calculated from the rest of estimations) constitutes the final robot pose estimation. Chebyshev's inequality assures that in a data sample almost all the values are close to its mean value. It defines intervals depending on the parameter k to fix the deviation tolerance, and formulates that:

$$P(Q - \mu > k\sigma) \leq \frac{1}{k^2} \qquad (5)$$

being Q a random variable of mean μ and typical deviation σ. The right side of the equation represents the highest possible percentage of values in the sample

that will not belong to the interval. For example, for $k = 1.6$ the Chebyshev's inequality guarantees that at least 60% of the values will belong to the interval, which means that these values are close enough to the rest of the sample values.

After the spatial relations -created at time instant t- are identified by the comparison with the previously created ones, two important pieces of information are available. Firstly, there are the ToF coordinates of the corners belonging to the new spatial relation, which represent a connection between the position of the robot and the corners. And secondly, the map coordinates of the corners belonging to the previously created spatial relation equivalent to the new one, are present. As the spatial relations are considered equivalent, the new spatial relation has the same map coordinates than the older one.

3   Data and Results

In this section, the results and configuration of the experiments are presented. The test environment consists of a laboratory containing several obstacles (mainly composed of tables and several objects under them) as shown in Fig. 3. Also notice the presence of chairs in the scenario. The ToF camera used in the research and experiments is a Mesa Imaging camera model SR4000.

Fig. 3. Laboratory test scenario partial view

There are two factors which have complicated the localization performance. First, the presence of the aforementioned obstacles in the scenario, and second, the height interval that determines which objects will be included. Fig. 4 presents the ToF observation made at a random iteration. It is presented with the approximate height interval used in the experiment and the corners found by Harris algorithm. In the figure, it is shown that several objects (including the wall) are found within the same height interval.
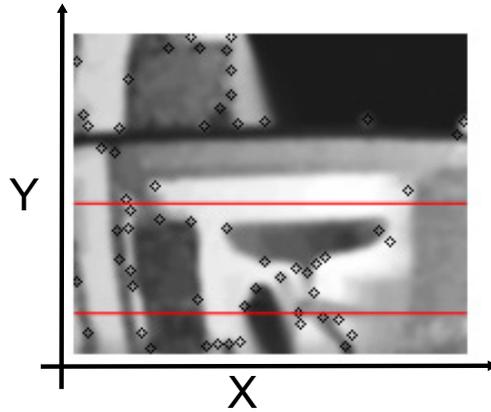
Fig. 4. Different obstacles detected with the selected height interval

To understand the difference between the representation of a group of obstacles in the same direction and the representation of only one obstacle (the most intuitive case would be a wall), Fig. 5 shows a comparison of the representation of multiple objects (see Fig.5a), and the representation of a wall (as shown in Fig. 5.b). The figure also presents the ToF images and the relative position of the robot.
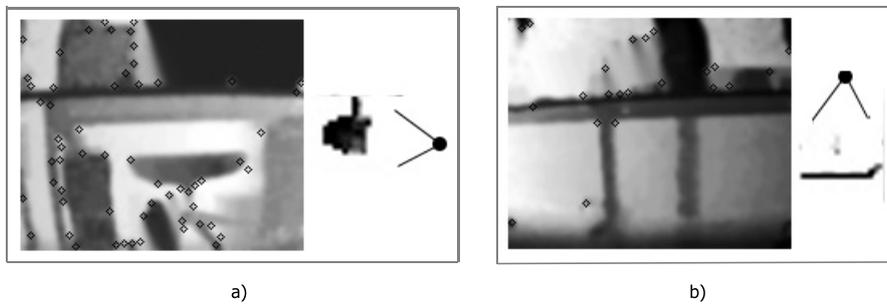

a)   b)

Fig. 5. Comparison of the representation of different kinds of obstacles. (a) Left: ToF image of several obstacles at different heights. Right: representation in the map (coordinates $(-2, 1)$) and position of the robot. (b) Left: ToF image of a single obstacle (a wall). Right: representation in the map (coordinates $(1, -3)$) and position of the robot.

ToF images are presented in Fig. 4 and Fig. 5, all of them with the respective corners found by the Harris algorithm. As explained before, spatial relations are created from each pair of corners at each level. Fig. 6 shows the spatial relations created for only one corner in a random ToF image captured during this test. Of course, spatial relations from the rest of corners and from the other level corners
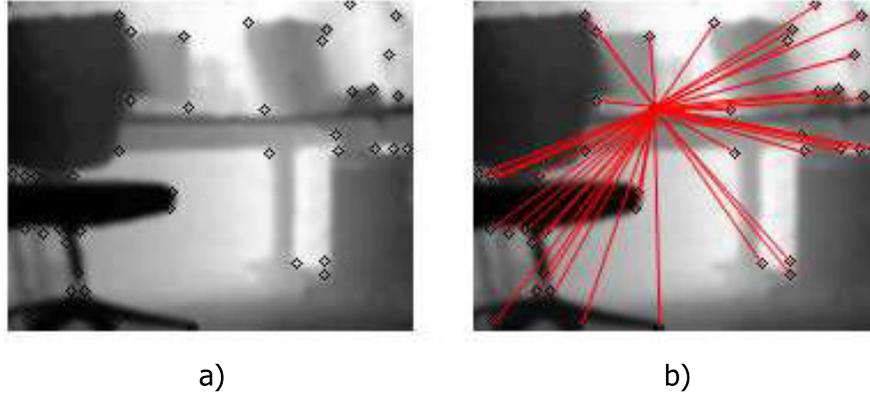
a)   b)

Fig. 6. Example of spatial relations. (a) ToF image with second level corners. (b) Spatial relations corresponding to a random corner.

are also created. Thus, the localization is based on the identification of these spatial relations, having a group of at most 10 identifications per iteration.

The information of the localization for one of the iterations of the current test is presented next. Initially the robot coordinates are $(0, 0, 0)$. In this case, the robot has been moved to coordinates $(0, -0.4)$ and has an orientation of -1.047 radians (or 330 degrees). Firstly, the comparison between new and previous spatial relations' attributes is carried out to achieve the identifications. Table 1 presents the results of the identification of the spatial relations for the current iteration. As aforesaid, S_y is defined as the slope between x and y, S_z as the slope between x and z, and N as the distance between the points (measured in meters). Each row of the table presents, on the one hand the attributes of the spatial relations observed in the current iteration and, on the other one the attributes of the spatial relations previously detected and stored, and compared to the first ones.

Table 1. Identification of spatial relations

#	Observed spatial relation			Stored spatial relation			Level
	S_{y1}	S_{z1}	N_1	S_{y2}	S_{z2}	N_2	
1	0.148803	0.045169	0.638492	0.148148	0.046868	0.639393	1
2	0.010794	0.026259	0.977325	0.014683	0.025786	0.975971	1
3	0.341993	0.471470	0.499058	0.338043	0.471988	0.495946	1
4	1.024206	0.832313	0.740300	1.022245	0.835643	0.737851	1
5	0.257143	0.257885	0.620442	0.257143	0.259601	0.618227	1
6	0.562114	0.330744	0.558678	0.562114	0.331833	0.557799	1
7	0.257143	0.274846	0.602269	0.257143	0.271577	0.605315	1
8	0.470173	0.456345	0.532023	0.469658	0.457272	0.531060	1
9	0.017528	0.003643	0.550909	0.018519	0.008466	0.546147	1
10	0.100000	0.103997	0.370625	0.100000	0.099922	0.371421	1

All the identifications in the table fulfill equation (4) as it can be checked by comparing the attributes of the spatial relations in the table. There are two significant pieces of information in the table. First, all of the identified spatial relations belong to the first level, and second, the maximum number of spatial relations has been identified. This information indicates, a priori, that the localization would be reliable. For example, a localization starting from a shorter group of identifications, with spatial relation from the second level would be less reliable.

4   Conclusions

In this paper, a localization algorithm for mobile robots has been introduced. The Harris corner detection algorithm is applied to images captured by a ToF camera mounted on the mobile robot. In this case, the ToF camera is exploited as a gray-scale camera. The gray-scale image represents distances for the purpose of finding good features to be tracked. These features form the basis of the spatial relations used in the localization algorithm. The approach to the localization problem is based on the computation of the spatial relations existing among the corners detected. The current spatial relations are matched with the relations gotten during previous navigation.

The paper has explained the three steps present in the algorithm, namely, "detection of corners", "creation of spatial relations", and "identification of spatial relations". An experiment in a complex laboratory has shown the goodness of the proposal. Our future work consists now in taking a next step towards simultaneous localization and mapping (SLAM).

Acknowledgements

References

1. Begum, M., Mann, G.K.I., Gosine, R.G.: Integrated fuzzy logic and genetic algorithmic approach for simultaneous localization and mapping of mobile robots. Applied Soft Computing 8, 150–165 (2008)
2. Bennett, G.: Probability inequalities for the sum of independent random variables. Journal of the American Statistical Association 57(297), 33–45 (1962)
3. Bosse, M., Zlot, R.: Keypoint design and evaluation for place recognition in 2D lidar maps. Robotics and Autonomous Systems 57, 1211–1224 (2009)
4. Böhm, J.: Orientation of image sequences in a point-based environment model. In: Sixth International Conference on 3-D Digital Imaging and Modeling, pp. 233–240 (2007)

5. Davison, A.J., Murray, D.W.: Simultaneous localization and map-building using active vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 865–880 (2002)

6. Fernández-Caballero, A., López, M.T., Mira, J., Delgado, A.E., López-Valles, J.M., Fernández, M.A.: Modelling the stereovision-correspondence-analysis task by lateral inhibition in accumulative computation problem-solving method. Expert Systems with Applications 33(4), 955–967 (2007)

7. Gemeiner, P., Jojic, P., Vincze, M.: Selecting good corners for structure and motion recovery using a time-of-flight camera. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5711–5716 (2009)

8. Guethmundsson, S.A., Pardas, M., Casas, J.R., Sveinsson, J.R., Aanaes, H., Larsen, R.: Improved 3D reconstruction in smart-room environments using ToF imaging. Computer Vision and Image Understanding 114, 1376–1384 (2010)

9. Harris, C., Stephens, M.: A combined corner and edge detector. In: The Fourth Alvey Vision Conference, pp. 147–151 (1988)

10. Hedge, G., Ye, C.: Extraction of planar features from Swissranger SR-3000 range images by a clustering method using normalized cuts. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4034–4039 (2009)

11. Klippenstein, J., Zhang, H.: Quantitative evaluation of feature extractors for visual SLAM. In: Proceedings of the Fourth Canadian Conference on Computer and Robot Vision, pp. 157–164 (2007)

12. Li, J., Allinson, N.M.: A comprehensive review of current local features for computer vision. Neurocomputing 71, 1771–1787 (2008)

13. López, M.T., Fernández-Caballero, A., Mira, J., Delgado, A.E., Fernández, M.A.: Algorithmic lateral inhibition method in dynamic and selective visual attention task: Application to moving objects detection and labelling. Expert Systems with Applications 31(3), 570–594 (2006)

14. López-Valles, J.M., Fernández, M.A., Fernández-Caballero, A.: Stereovision depth analysis by two-dimensional motion charge memories. Pattern Recognition Letters 28(1), 20–30 (2007)

15. May, S., Droeschel, D., Holz, D., Fuchs, S., Malis, E., Nüchter, A., Hertzberg, J.: Three-dimensional mapping with time-of-flight cameras. Journal of Field Robotics 26(11-12), 934–965 (2009)

16. Tissainayagam, P., Suter, D.: Assessing the performance of corner detectors for point feature tracking applications. Image and Vision Computing 22, 663–679 (2004)

17. Tomasi, C., Kanade, T.: Detection and tracking of point features. Carnegie Mellon University Technical Report CMU-CS-91-132 (1991)

18. Weingarten, J.W., Gruener, G., Siegwart, R.: A state-of-the-art 3D sensor for robot navigation. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 3, pp. 2155–2160 (2004)