ECAI 2010
H. Coelho et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-606-5-1089

Knowledge-based adaptive thresholding from shadows

Paulo Santos[1] and **Hannah M. Dee**[2] and **Valquiria Fenelon**[3]

Abstract. This paper presents results of a mobile robot qualitative self-localisation experiment using information from cast shadows. We present results of self-localisation using two methods for obtaining the threshold automatically: in one method the images are segmented according to their grey-scale histograms, in the other the threshold is set according to a prediction about the robot's location, given a shadow-based map defined upon a qualitative spatial reasoning theory. To the best of our knowledge this is the first work that uses qualitative spatial representations both to perform egolocation and to calibrate a robot's interpretation of its perceptual input.

1 Introduction

Cast shadows as cues for depth perception have been used to enhance depictions of natural scenes since the Renaissance. Recent research within psychology suggests that the human perceptual system gives preferential treatment to information from shadows when inferring motion in depth and perceiving 3D scene layout. These studies suggest that information coming from shadows can override such basic notions as conservation of object size, rather than discard or distrust shadow information [3]. Shadows also contain information that are not used during passive perception, for instance, information about the presence and location of the light source and the caster; the intensity of the source; amongst others [1].

The contribution of this paper is the investigation of a qualitative self-localisation method using information from cast shadows, assuming a set of qualitative relations called Perceptual Qualitative Relations about shadows (PQRS). We discuss the experimental evaluation of this method using two techniques for obtaining the threshold automatically for segmenting each image picked out by a robot's camera: in one method the images are segmented according to its grey-scale histogram, in the other method the threshold is searched according to a prediction about the robot's location, given a shadow-based qualitative map.

PQRS [6] is a theory inspired by the idea that shadows provide the observer with the viewpoint of the light source, as they are a projection of the caster from it. Equivalently, we can say that every point in the shadow region is totally occluded by the caster from the viewpoint of the light source. This idea is developed by representing relations of occlusion and shadows within the scope of Qualitative Spatial Reasoning (QSR) field of research. The goal of QSR is to provide appropriate formalisms for representing and reasoning about spatial entities [2].

The basic part of PQRS is based on one particular QSR theory: the Region Occlusion Calculus (ROC) [5], which is itself built upon one of the best known QSR approaches: the Region Connection Calculus (RCC).

Using RCC relations, along with the primitive relation $TotallyOccludes(x, y, v)$ (which stands for "x totally occludes y with respect to the viewpoint v"), the Region Occlusion Calculus (ROC) represents the various possibilities of interposition relations between two arbitrary-shaped objects. In particular, with RCC and the primitive $TotallyOccludes/3$, it is possible to define occlusion relations for *non occlusion* ($NonOccludes/3$), *partial occlusion* ($PartiallyOccludes/3$) and *mutual occlusion* ($MutuallyOccludes/3$). In fact, [5] defines 20 such relations. However, considering the ROC relations between a caster o and its shadow s, from a viewpoint v, only the following relations have models in PQRS: $\{NonOccludesDC(o, s, v)$, $NonOccludesEC(o, s, v)$, $PartiallyOccludesPO(o, s, v)$, $PartiallyOccludesTPP(o, s, v)$, $TotallyOccludesTPPI(o, s, v)$, $TotallyOccludesEQ(o, s, v)$ and $TotallyOccludesNTPPI(o, s, v)\}$. Apart from the ROC relations inherited in PQRS, it assumes the primitive $Shadow(s, o, Scr, L)$ that represents that a shadow s is cast by a caster o, from the light source L.

Relative location

The formalism summarised above can be used to reason about shadows from arbitrary viewpoints: relating shadows with occlusion suggests the distinction of five regions defined from the lines of sight between the light source, the caster and its shadow. Figure 1 represents the regions used in the experiments of this paper, where L is the light source, O is the object (caster) and S is its shadow.

Therefore, any viewpoint v located on Region 1 (Fig. 1) will observe the shadow s and the object o as $NonOccludesDC(o, s, v)$; similarly, if v observes o and s from Region 3 it should see that $PartiallyOccludesPO(o, s, v)$ and from Region 5 that $TotallyOccludesNTPPI(o, s, v)$. Region 4 is the surface defined by the lines of sight from l tangential to o and s, from where v would observe $TotallyOccludesTPPI(o, s, v)$. In Region 2, v perceives object and shadow as $NonOccludesEC(o, s, v)$.

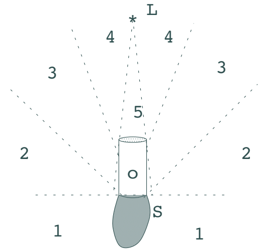

Figure 1. Regions implied by the observation of a shadow and its caster.

This idea for qualitative robot self-localisation using cast shadows was implemented on our Pioneer PeopleBot mobile robot using its monocular colour camera to obtain snapshots of objects and their

[1] Paulo Santos is with Elect. Eng. Dep., FEI, Sao Paulo, Brazil, email: psantos@fei.edu.br
[2] Hannah Dee is with School of Computing, University of Leeds, UK, email: hannah@comp.leeds.ac.uk.
[3] Valquiria Fenelon is with Elect. Eng. Dep., FEI, Sao Paulo, Brazil

shadows in an office-like environment (following the guidelines presented in [6]). Shadow detection was accomplished by first mapping the images captured by the camera into a HSV colour space. These images were then segmented by thresholding on V, whereby high values (light objects) are filtered out and low values (dark objects) are casters. Shadows are located within a value range in between light and dark objects. Morphological operators and the saturation value were used to filter noise (such as reflections of the light source on the object or background shadows). The robot was set to navigate through the room, stopping after a certain time interval to analyse its position with respect to the object-shadow locations according to the diagram shown in Figure 1. Shadow correspondence, which is the problem of matching each shadow to its caster, is solved in this work by assuming a simple heuristic: the shadow that is connected to an object's base is the shadow of this object. When there are various shadows connected to the object's base, the caster is associated with the shadow that is further away from the light source.

The ROC relations between a caster O and its shadow S are evaluated according to a threshold on the distance between the (top part of) the shadow when *Non Occlusion* holds. If the shadow is in some degree occluded by its caster, from the observer's viewpoint, the ROC relation is evaluated according to a percentage of the shadow that can be observed from behind the caster: $PartiallyOccludesPO(O, S, \nu)$ is interpreted when more than 10% of the shadow is observed; $TotallyOccludesTPPI(O, S, \nu)$ is assumed when less than (or equal to) 10% is still observed; and, $TotallyOccludesNTPPI(O, S, \nu)$ is concluded when no part of the shadow is seen from behind the caster.

2 Adaptive thresholds for segmentation

In this work we investigate the use of two distinct methods for automatically finding the best threshold for each given image: the traditional Otsu's method [4] and a threshold search related to the robot's prediction. The latter is the main contribution of the present paper.

Otsu's method [4] works by finding the threshold (t) that maximises the inter-class variance σ between two groups of pixels. Formula 1 expresses σ in terms of the threshold-dependent class probabilities ($\omega_1(t)$ and $\omega_2(t)$) and class means ($\mu_1(t)$ and $\mu_2(t)$) of groups 1 and 2.

$$\sigma^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2 \tag{1}$$

The second method for finding the best threshold uses the knowledge about the robot's previous location in order to make a prediction about it's current location. This procedure works as follows. The robot has to start in a known region. From this position the robot moves to another region (according to the diagram in Figure 1) in a moving action that is currently preprogrammed, but that still suffers from actuator noise. In this new position the robot captures a snapshot of the target object and uses it to decide on its location. If the location interpreted matches the prediction of its current position, then the robot moves on. If not, the robot varies the threshold until it finds a match between its predicted and interpreted positions, or fails otherwise.

3 Experiments

This section describes the results of the experiments on robot localisation with respect to the map in Figure 1. In these experiments the robot collected 1361 snapshots around the target object, which provides the frame of reference.

The baseline experiment uses fixed thresholds for image analysis chosen by experimentation within one of the camera views. The results obtained show a poor global performance of the system (47%) on localising the robot in every region. A high accuracy was obtained in the specific region used to calibrate the threshold (above 70% with respect to region 1), but within other regions the results where lower or equal to 50%. The poor performance outside of region 1 is because the foreground/background segmentation is not optimal for images obtained under other light conditions (i.e., the distinct position configurations between robot, caster and light produced by the agent's motion). In fact, by tweaking the thresholds, the system improved its performance in locating the robot on other regions, however this improvement came at the expense of losing accuracy on region 1.

The obvious approach for improving the poor results obtained by fixed-thresholding is to *adjust the thresholds for each snapshot taken*. The technique we have used to perform this adjustment is the Otsu method [4] (cf. Section 2). This should be able to automatically find the threshold for segmenting objects of interest (i.e. casters and their shadows) from background. The results obtained with a variable threshold method, surprisingly, were slightly worse than those obtained with a fixed threshold. For global localisation, the method answered correctly on 43% of the total 1361 snapshots. The localisation at region 1 was correct in 59% of the trials (decreasing from the 70% obtained with a fixed threshold), and the localisation accuracy on the other regions was below 50%. Investigation of the pixel value distributions indicated that the problem is that these distributions are not in general bi-modal, which increases the difficulty of searching for an appropriate threshold from the image histogram.

In our third set of results, the robot was set to vary the threshold until the interpretation of the target object and its shadow matches a robot's prediction of its location. The results obtained show that the system achieved an accuracy of around 90% in all regions. Thus the use of knowledge about shadow appearance, and reasoning based upon past location, can greatly assist in the refinement of a simple shadow-detection algorithm, outperforming also a traditional algorithm for adaptive thresholding.

4 Conclusion

This paper has demonstrated how the incorporation of qualitative spatial representation and *a priori* knowledge about shadow regions can be combined to enhance a simple shadow-detection algorithm based upon thresholding. Future work will involve the use of more sophisticated shadow detection algorithms, and the extension of the current snapshot-based system to one that analyses continuous video, and the inclusion of shadow reasoning within the perception-planning-action loop.

REFERENCES

[1] R. Casati, 'The shadow knows: a primer on the informational structure of cast shadows', *Perception*, **33**(11), 1385–1396, (2004).

[2] A. G. Cohn and J. Renz, 'Qualitative spatial representation and reasoning', in *Handbook of Knowledge Representation*, 551–596, Elsevier, (2008).

[3] P. Mamassian, D. C. Knill, and D. Kersten, 'The perception of cast shadows', *Trends in cognitive sciences*, **2(8)**, 288–295, (1998).

[4] N. Otsu, 'A threshold selection method from gray-level histograms', *IEEE Trans. Sys., Man., Cyber.*, (9), 62–66, (1979).

[5] D. Randell, M. Witkowski, and M. Shanahan, 'From images to bodies: Modeling and exploiting spatial occlusion and motion parallax', in *International Joint Conference on Artificial Intelligence*, pp. 57–63, Seattle, U.S., (2001).

[6] P. Santos, H.M. Dee, and V. Fenelon, 'Qualitative robot localisation using information from cast shadows', in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 220–225, (2009).