Aberystwyth University

Knowledge-based adaptive thresholding from qualitative robot localisation using cast shadows

Santos, Paulo E.; Fenelon, Valquiria; Dee, Hannah

Publication date:
2010

Citation for published version (APA):
Santos, P. E., Fenelon, V., & Dee, H. (2010). *Knowledge-based adaptive thresholding from qualitative robot localisation using cast shadows*. Paper presented at Spatial temporal dynamics workshop, ECAI, .
http://hdl.handle.net/2160/5899

Knowledge-based adaptive thresholding from qualitative robot localisation using cast shadows

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

Abstract. This paper presents results of a mobile robot qualitative self-localisation experiment using information from cast shadows. Shadow detection was accomplished by mapping the images from the robot's monocular colour camera into a HSV colour space and then thresholding on V. 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. This map-related threshold search is the main contribution of the present work, and 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 [11]. 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 [18, 6, 20]. Casati in [3] points out that cast 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; the caster's shape; the screen texture; and the distance between the caster and the screen.

Whilst psychologists have demonstrated the centrality of shadows to our own perception of depth, size and motion, much work in computer vision and robotics starts from the premise that shadows are sources of noise rather than information. The present work falls within the small but growing area of research which aims to use shadows not as sources of noise, but as sources of information. This requires not only a model of the kinds of information that shadows can purvey, but also a robust and accurate shadow detection system. Researchers within both computer vision and robotics have been working in this area – many engaged in shadow suppression in videos from fixed cameras, but some engaged in the more challenging task of shadow identification, localisation and use.

The contribution of this paper is the investigation of a qualitative self-localisation method using information from cast shadows. 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.

This paper is organised as follows. Section 2 outlines related research from within both computer vision and robotics. Section 3 describes the theory upon which the work is based - the Perceptual Qualitative Relations about Shadows (PQRS), which formalises the problem of shadow reasoning and egolocation within a qualitative spatial reasoning context. The adaptive thresholding methods considered in this work are presented in Section 4, and the experiments are described in Section 5. Discussions are drawn on Section 6 and Section 7 concludes this paper.

Throughout this paper, constants are written in upper-case letters and variables in lower case, unless explicitly stated otherwise.

2 Related research

When considering the task of segmentation of moving objects from a static background, shadows are a frequent source of false positives [10, 21] and therefore shadow *suppression* is a major research area. In this context, shadow detection in computer vision almost always involves some model of the colour of the *screen* or in computer vision terminology *background*, and detection is performed using a model of shadows characterising them as *roughly the same colour as the background, but darker'*. Perhaps the simplest shadow detection method proposed is that of [34], in which a grey-scale image is simply thresholded and the darker pixels are labelled *shadow*; however this approach fails on complex images and in situations where lighting changes due to either environmental effects or egomotion. Prati in [27] provides an overview and a taxonomy of early shadow-detection techniques, dividing them into *model-based* and *non-model-based*, however this categorisation does not apply well to more recent works, many of which can be thought of as ensemble methods [21, 26].

Cucchiara et al. in [10] take as their starting point detected moving objects (and a background model). The pixel values of moving objects are converted to the HSV (Hue, Saturation and Value) colour space, and then these are investigated to determine whether they are real moving objects or merely shadow pixels. This is accomplished by considering observed and background values of all three HSV components, considering the difference between foreground and background values for H and S, and the ratio of the two V values. This captures the intuitive observations that shadows are about the same hue as the same part of the scene unshadowed, slightly more saturated, and darker. A similar approach based upon the observation of colour changes in cast shadows is presented in

[1] Paulo Santos is with Electrical Engineering Dep., FEI, S. Paulo, Brazil, email: psantos@fei.edu.br
[2] Valquiria Fenelon is with Escola Politécnica, USP, S. Paulo, Brazil
[3] Hannah Dee is with School of Computing, University of Leeds, UK

[30]. Stauder et al. in [32] use assumptions about the background (*it will dominate the scene*), the nature of shadows and luminance (*shadows are darker and tend to have uniform shading*) and the presence of moving and static edges. Other methods for shadow filtering are described in [24, 35, 14], an overview of such approaches is left for future work.

Within the last two or three years, models inspired by the physics of light have become more prominent. These systems, rather than simply observe that "shadows are a bit darker", consider the nature of reflectance and the effect of lighting changes on perceived colour. Martel-Brisson and Zaccarin [22] take a simplified reflectance model and use it to learn the way in which colours change when shaded, and Huang and Chen [15] have also incorporated a richer, physics-based colour model for shadow detection based upon the work of Maxwell et al. [23]. Maxwell presents a bi-illuminant dichromatic reflection model, which enables the separation of the effects of lighting (direct and ambient) from the effects of surface reflectance. Huang and Chen [15] simplify this model in several ways, such as assuming that the ambient illumination is constant, which enables them to implement shadow detection based upon the simplified model in video analysis.

There are a few systems within computer vision that use cast shadows as sources of information rather than noise. [2] use known 3D locations and their cast shadows to perform camera calibration and light location (using known casters and screen to tell about the light source); [4] uses the moving shadows cast by known vertical objects (flagpoles, the side of buildings) to determine the 3D shape of objects on the ground (using the shadow to tell about the shape of the screen). Balan et al. [1] use shadows as a source of information for detailed human pose recognition: they show that using a single shadow from a fixed light source can provide a similar disambiguation effect as using additional cameras.

In robotics, the story is similar. Fitzpatrick and Torres-Jara in [13], inspired by the research reported in [5], track the position of a robotic arm and its shadow cast on a table to derive an estimate of the time of contact between the arm and the table. Shadows are detected in this work using a combination of two methods: in the first, a background model of the workspace is built without the arm and then used to determine light changes when the arm is within the camera view. The second method compares subsequent frames in order to detect moving regions of light change. The authors motivate their work pointing out that depth from shadows and stereopsis may work as complementary cues for robot perception, while the latter is limited to surfaces rich in textures, the former works well in smooth (or even reflective) surfaces. Cheah et al. [7] present a novel controller for a robot manipulator, providing a solution to the problem of trajectory control in the presence of kinematic and dynamic uncertainty. In order to evaluate their results, an industrial robot arm was controlled using the visual observation of the trajectory of its own shadow. Lee and colleagues [17] use cast shadows inside pipes to detect landmarks: by fitting bright lights to the front of their pipe inspection robot, they can determine when a pipe bends by detecting cast shadows.

Information from shadows are also considered in unmanned autonomous planetary exploration. Tompkins et al. [33] describe an autonomous path planning system that takes into account various conditions of the robot's state, including particularities of the terrain and lighting. In this context, the information about shadows cast by terrain irregularities allows the rover to plan a trajectory that maximises the trade-off between the exposure of the solar cells to sun light and the limited resources in planetary missions. Kunii and Gotoh [16] propose a *Shadow Range Finder* system that uses the shadow cast by a robot arm on the surface of a terrain in order to obtain depth information around target objects. In planetary explorations this type of system may provide low-cost, energy-saving, sensors for the analysis of the terrain surrounding rock samples of interest.

More recently, Santos et al. [31] describe an initial representation of cast shadows in terms of a spatial logic formalising occlusion relations. This representation, called Perceptual Qualitative relations about Shadows (PQRS), is used in a mobile robot self-localisation procedure in office-like environments. The present paper builds upon this representation and, therefore, the next section describes it in more detail.

3 Perceptual qualitative relations about shadows (PQRS)

Perceptual Qualitative Relations about Shadows (PQRS) [31] 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, which is part of the artificial intelligence sub-area known as Knowledge Representation and Reasoning [12].

The goal of QSR is to provide appropriate formalisms for representing and reasoning about spatial entities, such as part-whole relations, connectivity, orientation, line segments, size and distance, amongst others [9, 8].

PQRS assumes a static light source, denoted by L, situated above the observer (in agreement to recent research on the psychophysics of perception [19]). It is also assumed that the scenes are observed from an egocentric point of view (v), and that shadows are cast on a single screen Scr which does not need to be flat.

The basic part of PQRS is based on one particular QSR theory: the Region Occlusion Calculus (ROC) [29], which is itself built upon one of the best known QSR approaches: the Region Connection Calculus (RCC) [28]. RCC is a first-order axiomatisation of spatial relations based on a reflexive, symmetric and non-transitive dyadic primitive relation of *connectivity* ($C/2$) between two regions. Informally, assuming two regions x and y, the relation $C(x, y)$, read as *"x is connected with y"*, is true if and only if the closures of x and y have at least one point in common.

Assuming the $C/2$ relation, some mereotopological relations between two spatial regions can be defined, such as *disconnected from* (DC), *equal to* (EQ), *overlaps* (O); *part of* (P) ; *partially overlaps* (PO); *proper part of* (PP); *externally connected* (EC) and *tangential* or *non-tangential proper part* (resp. (TTP) and ($NTTP$)).

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, [29] 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: $\{NonOccludesDC(o, s, v)$, $NonOccludesEC(o, s, v)$, $PartiallyOccludesPO(o, s, v)$, $PartiallyOccludesTPP(o, s, v)$, $TotallyOccludes$-$TPPI(o, s, v)$, $TotallyOccludesEQ(o, s, v)$ and $Totally$-

$OccludesNTPPI(o, s, v)\}$. Figure 1 represents these relations, where the dashed object is the caster and the blank is its shadow.

The Region Occlusion Calculus makes a distinction between the occupancy regions of bodies and their images (or projections) from the viewpoint of an observer by assuming the function $region(x)$, which maps a body x to its occupancy region, and the function $image(x, \nu)$ that maps a body x (and the viewpoint ν) to the body's image. Therefore, given two bodies X and Y and a viewpoint ν, the statement $PartiallyOccludesTPP(X, Y, v)$ is defined as $PartiallyOccludes(X, Y)$ *and* $TPP(image(X), image(Y))$.

It is worth pointing out also that the "I" in the relations $TotallyOccludesTPPI(o, s, v)$ and $TotallyOccludesNTPPI(o, s, v)$ represents the inverse of TPP and PP, resp.; so, for instance, $TotallyOccludesTPPI(o, s, v)$, means that the caster o totally occludes its shadow s, but s is the tangential proper part of o.

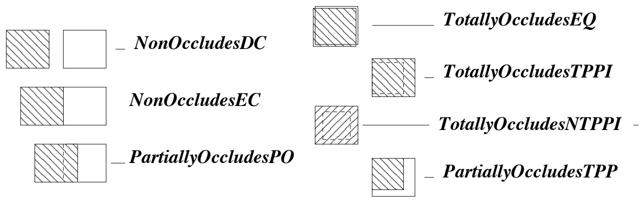


Figure 1. The ROC relations that are part of PQRS.

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. The axiom constraining the $Shadow/4$ relation is represented by Formula 1 below.

$$\begin{aligned} Shadow(s, o, Scr, L) \leftrightarrow PO(r(s), r(Scr)) \wedge \\ TotallyOccludes(o, s, L) \wedge \\ \neg\exists o' TotallyOccludes(o', o, L). \end{aligned} \tag{1}$$

The axiom represented in Formula 1 states that the shadow of a caster o is the region in a screen Scr that is totally occluded by o from the light source viewpoint L.

3.1 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 (or the top-half part of the latter if it is cast on the floor), as represented in Figure 2. Therefore, any viewpoint v located on Region 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 $NonOccludesEC(o, s, v)$. Regions 2 and 4 are in fact *boundaries* separating regions 1 and 3, and between 3 and 5 respectively. Therefore, it is virtually impossible for a robot to locate itself on them. In the real robot environment, however, regions 2 and 4 are extended assuming an interval of uncertainty around these boundaries. Figure 3 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.

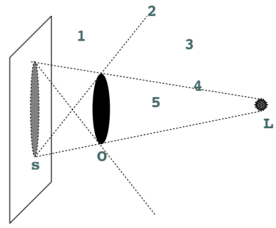


Figure 2. Distinct regions implied by the observation of a shadow and its caster. It is worth noting that, in this figure, regions 2 and 4 are zero-width boundaries.

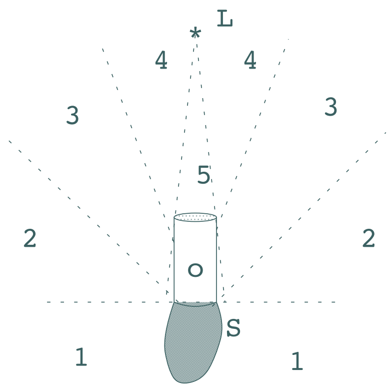
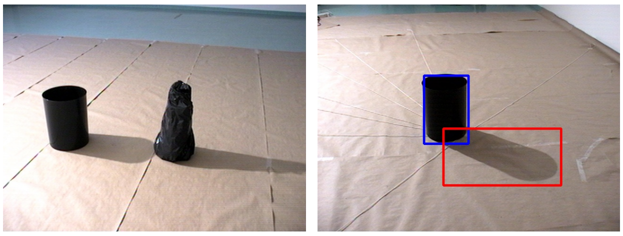


Figure 3. Regions implied by the observation of a shadow and its caster.

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 shadows in an office-like environment (following the guidelines presented in [31]). 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 3. One example of the snapshots used in this work is shown in Figure 4(b). Shadow correspondence, which is the problem of matching each shadow to its caster [18, 20], 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 (Fig. 4(a) shows an example of such situation).

Given a threshold Th, a $Scene$ and a viewpoint ν, Algorithm 1 summarises the method for self-localisation described in this section.

In Algorithm 1 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:



(a) Two shadows (b) Example segmented image

Figure 4. (a) two shadows in one object's base and (b) example of a segmented image

Algorithm 1 PERCEPTION_ACTION(Th, Scene, ν)

segment $Scene$ using the threshold Th to obtain a caster O and its shadow S
if $NonOccludesDC(O, S, \nu)$ **then**
  robot is on region 1
else if $NonOccludesEC(O, S, \nu)$ **then**
  robot is on region 2
else if $PartiallyOccludesPO(O, S, \nu)$ **then**
  robot is on region 3
else if $TotallyOccludesTPPI(O, S, \nu)$ **then**
  robot is on region 4
else if $TotallyOccludesNTPPI(O, S, \nu)$ **then**
  robot is on region 5
else
  FAIL
end if

$PartiallyOccludesPO(O, S, \nu)$ 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.

The material presented up to this point is discussed in greater details in [31]. The remainder of this paper is completely original.

4 Adaptive thresholds for foreground/background 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 [25] and a threshold search related to the robot's prediction. The latter is the main contribution of the present paper.

Otsu's method [25] works by finding the threshold (t) that maximises the inter-class variance σ between two groups of pixels. Formula 2 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.

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

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 3) 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. This method is summarised in Algorithm 2 below.

In the pseudocode THRESHOLD AND POSITION (Algorithm 2), the function MOVING_ACTION(s_0, v, dir, I) gives the prediction of the robot's position s_i after its motion from the position s_0, with speed v, direction dir and for a time interval I; the function PERCEPTION_ACTION($th, Scene, \nu$) outputs the perceived robot's position according to the observed PQRS relation, for a threshold th, a $Scene$ and a viewpoint ν (as discussed in Section 3.1), th_{aux} is an auxiliary variable for threshold and s_0, s_i and sj are variables for the robot's position. The constant $Step$ is used to update the threshold from its minimum (Th_{min}) to its maximum (Th_{max}) values (in this work these constants were set at $Step = 5, Th_{min} = 40$ and $Th_{max} = 230$).

Algorithm 2 THRESHOLD AND POSITION(th_0, t_0, $scene$)

s_i = MOVING_ACTION(s_0, v, dir, I)
s_j = PERCEPTION_ACTION($th_0, Scene, \nu$)
if ($s_i == s_j$) **then**
  return (th_0, s_i)
else
  $th_{aux} = Th_{min}$
  while (($si \neq sj$) and ($th_{aux} < Th_{max}$)) **do**
    $th_{aux} = th_{aux} + Step$
    s_j =PERCEPTION_ACTION($th_{aux}, Scene, \nu$)
  end while
  if ($th_{aux} > Th_{max}$) **then**
    return FAIL
  else
    return (th_{aux}, s_j)
  end if
end if

To the best of our knowledge this is the first paper where the segmentation threshold is obtained as a result of the robot's prediction of its location according to a qualitative map. In this way, we use the PQRS theory not only for robot self-localisation based upon shadow perception, but also for the refinement of the shadow perception itself. The next section presents an empirical evaluation of this technique.

5 Experiments

This section describes the results of the experiments on robot localisation with respect to the map in Figure 3. In these experiments the robot collected 1361 snapshots around the target object, which provides the frame of reference (e.g. the black bucket in Figure 4(b)). This target was always within camera view, but not necessarily at its centre. We allowed up to three objects within the robot's field of view. Due to the narrow view of the robot, and the use of a single dominant light source, localisation estimates with respect to each object do not contradict one another.

The baseline experiment uses fixed thresholds for image analysis chosen by experimentation within one of the camera views. These results are represented in Table 1, which shows 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.

Table 1. Fixed thresholds

Region	n. of images	correct answers	correct answers (%)
1	320	235	73
2	436	119	27
3	438	222	50
4	111	44	40
5	56	21	38
Global	1361	641	47

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 [25] (cf. Section 4). 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 are represented in Table 2.

Table 2. Adaptive threshold using the Otsu method

Region	n. of images	correct answers	correct answers (%)
1	320	190	59
2	436	195	45
3	438	157	36
4	111	27	24
5	56	14	25
Global	1361	583	43

Table 2 shows that the results with a variable threshold method, surprisingly, were slightly worse than those obtained with a fixed threshold (Table 1). 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 (using Algorithm 2, as explained in Section 4). The results obtained are represented in Table 3, which shows that the system achieved an accuracy of around 90% in all regions. Thus the incorporation 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.

6 Discussion and open issues

In this work we investigated robot self-localisation using qualitatively distinct regions defined of a visual observation of cast shadows. Central to this problem is the segmentation of casters and shadows from the background, which was accomplished here by thresholding on value (V) on the HSV colour space. In the present paper we proposed a new strategy for calibrating this threshold, where the prediction about the robot's location is used to search for a match between the interpreted position (as given from visual observation) and its predicted location. In order to evaluate this method, we presented three sets of experiments whereby different ways of defining the threshold were tried. In the first set of experiments a hand-coded fixed threshold was used, in the second set of experiments we used Otsu's method [25] in order to find the threshold values from the image histogram. Finally, the third set of experiments presents the results of applying our proposed method for matching the prediction with the observation.

Table 3. Knowledge-based adaptive threshold

Region	n. of images	correct answers	correct answers (%)
1	320	297	93
2	436	385	88
3	438	410	94
4	111	102	92
5	56	48	86
Global	1361	1242	91

The intuition behind the experiments with fixed thresholds was to provide a lower-bound for the evaluation of our idea, since (as we hypothesised) nothing could perform worse than a hand-coded threshold. Experiments with Otsu's method were then to set the standard, as this is one of the most traditional methods for adaptive thresholding. However, it turned out that Otsu's performance was in fact approximately as accurate as that of using the fixed threshold. This is due to the fact that we chose for the first set of experiments the best threshold we could find, after a number of trials where the value was changed by hand. Otsu's method, however, had to deal with arbitrary images, where it had to maximise a value that is dependent on an *a priori* hypothesis of bi-modal pixel distribution. This was not the case in some of the snapshots taken by the robot: a great number of them suffered from the effect of reflections of the light source on the caster; moreover, from some angles, there was a negative gradient of luminosity just behind the object. These problems caused Otsu's method to perform worse than using a fixed threshold.

In contrast, the method for calibrating the threshold using the prediction about the robot's location performed as well as could be expected, obtaining an accuracy of around 90% with respect to our dataset containing 1361 snapshots of the robot's environment. However, this method is totally dependent on the capability of the robot's actuators on generating accurate predictions for the robot's future location, given a moving action. In this paper, the robot's motion was completely pre-programmed in order to minimise the actuator's noise, this gave us the guarantee that we were only evaluating the localisation procedures. We leave for a future work applying this framework on a system that has a path planning module, so that it can be verified how the qualitative localisation procedure proposed in this paper is affected by errors in the planning-acting-sensing cycle. Evaluating the ideas put forward in this paper on a more complex scenario is also a desirable future goal.

Also subject for our long term investigations is the complete exploration of the knowledge content of shadows, as described in [3], in order to create a robotic system that is capable of perceiving (and interpreting) shadows in a similar fashion to humans. The reason for pursuing this goal resides in our hypothesis that the human perceptual system, by preferring shadow information over other depth cues (even when these cues contradict each other [20]), is in fact saving processing time. Investigating how this could be accomplished in robotic systems is a major motivation of this work.

Although this work explores only a qualitative theory about space, this choice does not preclude the use of quantitative or statistical methods. Rather, we believe that qualitative methods in robotics should complement the traditional numerical algorithms, providing another processing level where it is possible to extract information from the knowledge level.

7 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 incorporation of more sophisticated shadow detection algorithms, and the extension of the current snapshot-based system to one which incorporates continuous video, and the inclusion of shadow reasoning within the perception-planning-action loop.

A number of questions have been raised by this work, and we consider these questions in themselves to be a useful contribution. For example, how can shadows improve object localisation when contrasted to object-based methods? Under what conditions can shadows be effectively exploited? How can we combine predictive shadow-based localisation with predictive localisation based upon object pose? These are all questions which we hope to consider in more depth in future work.

REFERENCES

[1] A. O. Balan, M. J. Black, H. Haussecker, and L. Sigal, 'Shining a light on human pose: On shadows, shading and the estimation of pose and shape', *Proc. International Conference on Computer Vision (ICCV)*, (2007).

[2] Xiaochun Cao and Hassan Foroosh, 'Camera calibration and light source orientation from solar shadows', *Proc. Computer Vision and Pattern Recognition (CVPR)*, **105**(1), 60–72, (2007).

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

[4] Y. Caspi and M. Werman, 'Vertical parallax from moving shadows', in *Proc. Computer Vision and Pattern Recognition (CVPR)*, New York, USA, (2006).

[5] U. Castiello, D. Lusher, C. Burton, and P. Disler, 'Shadows in the brain', *Journal of Cognitive Neuroscience*, **15:6**, 862–872, (2003).

[6] P. Cavanagh, 'The artist as neuroscientist', *Nature*, **434**, 301–307, (2005).

[7] C.-C Cheah, C. Liu, and J.-J. E. Slotine, 'Adaptive tracking control for robots with unknown kinematic and dynamic properties', *I. J. Robotic Res.*, **25**(3), 283–296, (2006).

[8] A. G. Cohn and S. M. Hazarika, 'Qualitative spatial representation and reasoning: An overview', *Fundamenta Informaticae*, **46**(1-2), 1–29, (2001).

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

[10] R. Cucchiara, C. Grana, G. Neri, M. Piccardi, and A. Prati, 'The sakbot system for moving object detection and tracking', in *Video-based Surveillance Systems: Computer Vision and Distributed Processing (Part II - Detection and Tracking)*, 145–158, Kluwer Academic Publishers, (2001).

[11] T. da Costa Kauffmann, 'The perspective of shadows: The history of the theory of shadow projection', *Journal of the Warburg and Courtauld Institutes*, **38**, 258–287, (1979).

[12] R. Davis, H. Shrobe, and P. Szolovits, 'What is a knowledge representation?', *AI Magazine*, **14**(1), 17–33, (1993).

[13] P.M. Fitzpatrick and E.R. Torres-Jara, 'The power of the dark side: using cast shadows for visually-guided touching', in *Proc. of the 4th IEEE/RAS International Conference on Humanoid Robots*, pp. 437–449, (2004).

[14] J.-W. Hsieh, W.-F. Hu, C.-J. Chang, and J.-S. Chen, 'Shadow elimination for effective moving object detection by Gaussian shadow modeling', *Image and Vision Computing*, **21(6)**, 505–516, (2003).

[15] J. B. Huang and C. S. Chen, 'Moving cast shadow detection using physics-based features', in *Proc. Computer Vision and Pattern Recognition*, (2009).

[16] Y. Kunii and T. Gotoh, 'Evaluation of Shadow Range Finder: SRF for Planetary Surface Exploration', in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2573–2578, (2003).

[17] J-S. Lee, S-G. Roh, D. W. Kim, H. Moon, and H. R. Choi, 'In-pipe robot navigation based upon the landmark recognition system using shadow images', in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1857–1862, (2009).

[18] P. Mamassian, 'Impossible shadows and the shadow correspondence problem', *Perception*, **33**, 1279–1290, (2004).

[19] P. Mamassian and R. Goutcher, 'Prior knowledge on the illumination position', *Cognition*, **81**(1), B1–B9, (2001).

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

[21] N. Martel-Brisson and A. Zaccarin, 'Learning and removing cast shadows through a multidistribution approach', *IEEE transactions on Pattern Analysis and Machine Intelligence (PAMI)*, **29 (7)**, 1134–1146, (2007).

[22] N. Martel-Brisson and A. Zaccarin, 'Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation', in *Proc. Computer Vision and Pattern Recognition*, (2008).

[23] B. Maxwell, R. Friedhoff, and C. Smith, 'A bi-illuminant dichromatic reflection model for understanding images', in *Proc. Computer Vision and Pattern Recognition*, (2008).

[24] S. Nadimi and B. Bhanu, 'Physical models for moving shadow and object detection in video', *IEEE transactions on Pattern Analysis and Machine Intelligence*, **26(8)**, 1079– 1087, (2004).

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

[26] F. Porikli and J. Thornton, 'Shadow flow: A recursive method to learn moving cast shadows', in *Proc. International Conference on Computer Vision (ICCV)*, (2005).

[27] A. Prati, I. Mikic, M.M. Trivedi, and R. Cucchiara, 'Detecting moving shadows: algorithms and evaluation', *IEEE transactions on Pattern Analysis and Machine Intelligence*, **25(7)**, 918–923, (2003).

[28] D. Randell, Z. Cui, and A. Cohn, 'A spatial logic based on regions and connection', in *International Conference on Knowledge Representation and Reasoning*, pp. 165–176, Cambridge, U.S., (1992).

[29] 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).

[30] Elena Salvador, Andrea Cavallaro, and Touradj Ebrahimi, 'Cast shadow segmentation using invariant color features', *Computer Vision and Image Understanding (CVIU)*, **95**(2), 238–259, (2004).

[31] 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).

[32] J. Stauder, R. Mech, and J. Ostermann, 'Detection of moving cast shadows for object segmentation', *IEEE Transactions on multimedia*, **1(1)**, 65–76, (1999).

[33] P. Tompkins, A. Stentz, and W. L. Whittaker, 'Automated surface mission planning considering terrain, shadows, resources and time', in *Proceedings of the 6th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS '01), Montreal, Canada.*, (June 2001).

[34] A. Troccoli and P. K. Allen, 'A shadow-based method for model registration', in *Computer Vision and Pattern Recognition Workshops*, (2004).

[35] Y. Wang, K.-F. Loe, T. Tan, and J.-K. Wu, 'A dynamic Hidden Markov Random Field Model for foreground and shadow segmentation', in *Proc. IEEE Workshop on Applications of Computer Vision*, (2005).