A Qualitative-Probabilistic Approach to Autonomous Mobile Robot
Self Localisation and Self Vision Calibration

Valquiria Fenelon Pereira, Fabio Gagliardi Cozman
Department of Mechanical Engineering
Escola Politécnica da Universidade de São Paulo
São Paulo, SP – Brazil
vfenelon@usp.br, fgcozman@usp.br

Paulo Eduardo Santos, Murilo Fernandes Martins
Department of Electrical Engineering
Centro Universitário da FEI
São Bernardo do Campo, SP – Brazil
psantos@fei.edu.br, murilo@ieee.org

Abstract—Typically, the spatial features of a robot’s environment are specified using metric coordinates, and well-known mobile robot localisation techniques are used to track the exact robot position. In this paper, a qualitative-probabilistic approach is proposed to address the problem of mobile robot localisation. This approach combines a recently proposed logic theory called Perceptual Qualitative Reasoning about Shadows (PQRS) with a Bayesian filter. The approach herein proposed was systematically evaluated through experiments using a mobile robot in a real environment, where the sequential prediction and measurement steps of the Bayesian filter are used to both self-localisation and self-calibration of the robot’s vision system from the observation of object’s and their shadows. The results demonstrate that the qualitative-probabilistic approach effectively improves the accuracy of robot localisation, keeping the vision system well calibrated so that shadows can be properly detected.

Keywords—Qualitative Spatial Reasoning; Bayesian Filtering; Mobile Robot; Self-localisation.

I. INTRODUCTION

Navigation is one of the challenges for autonomous mobile robots. Four distinct skills compose navigation: perception, localisation, cognition, and motion control [19]. In this paper, the goal is to perform localisation under uncertainty using a qualitative spatial reasoning approach with information from images that contain shadows. The localisation problem is characterised as position tracking assuming that an initial qualitative robot pose is given. The environment is static, with one object and fixed light source, where the mobile robot is able to move while it captures frames from its camera. Previous studies have addressed the problem of qualitative region localisation using information from shadows. We address the localisation problem using a combination of qualitative logic and probabilistic techniques.

We employ the theory of Perceptual Qualitative Relations about Shadows (PQRS), proposed in [17], to reason about relations between cast shadow and objects. This particular formalisation allows the robot to infer its qualitative position w.r.t. the fixed light source and the object. Although the formal theory can describe all possible locations the robot might be, a major issue arises when the vision system extracts the shadow from the scenes: the camera is rather noisy and therefore a threshold parameter must be fine-tuned in order to correctly detect the shadows. A method to overcome this issue was proposed in [8], which takes into account beliefs about previous locations.

The present paper extends that previous work by combining the qualitative beliefs defined by PQRS with a Bayesian filter. In this approach, the qualitative beliefs represent the prediction, whilst the extracted object and shadow observed in the image represent the measurement update. Additionally, the filter is able to make predictions not only about the robot’s localisation, but also about the best threshold value to be used with the next image to be captured by the robot. Our empirical results clearly demonstrate that the proposed approach considerably enhances the accuracy in localisation when compared to [8].

This paper is organised as follows. Section II outlines related work, discussing the state of the art. Then, the PQRS approach used in robot self-localisation is briefly introduced in Section III. Next, the proposed probabilistic approach to PQRS with a Bayesian filter approach is detailed in Sections IV and V. Finally, Section VI describes the experiments performed, results and discussions, while Section VII concludes this paper and presents possible future directions.

II. RELATED WORK

Renaissance painters have always exploited the effect of shadows in the human depth perception [3]. However, only in the 21st century researchers begun investigating the cognitive processes involved in utilising shadows as cues for external world perception [5], [14]. Shadows carry information about presence, location, as well as intensity of the light source. Moreover, shadows provide information about the caster (object), its shape and texture. In addition, the distance between caster and screen can be hypothesised given whether or not the caster and the shadow appear to be in contact with each other [12]. The information about shadows enhances detection techniques in Geographic Information Systems, even improving the detection of buildings [2]. In [4], a system is described that determines the 3D
The first idea of qualitative self-localisation was presented by [13], where a topological map was built by spaces bounded by sets of lines connecting pairs of point-wise landmarks. Later, several efforts [10], [18], [21] developed spatial representation behind this idea. In particular, Foglia-roni et al. [10] considered extended convex objects to be handled as landmarks and the decomposition of navigable space was based on the model of occlusion and visibility. Following similar ideas, in our previous work [8], [17] we have proposed a qualitative spatial theory named Perceptual Qualitative Relations about Shadows (PQRS), which has been used in the task of robot self-localisation. Since the present work builds upon this theory, PQRS is presented in more detail in the next section.

III. PERCEPTUAL QUALITATIVE RELATIONS ABOUT SHADOWS (PQRS)

The PQRS theory [8], [17] provides a formalism that can be used to reason about shadows from a particular viewpoint. The theory relates shadows with occlusion using seven occlusion relations from the Region Occlusion Calculus (ROC) [16], here defined w.r.t. the light source, the caster and its shadow. The set of relations for a caster \( o \), its shadow \( s \) and a viewpoint \( \nu \) defined by PQRS includes the Region Connection Calculus (RCC) [15] and a subset of ROC, as follows:

\[
\begin{align*}
\text{NonOccludesDC}(o,s,\nu) \\
\text{NonOccludesEC}(o,s,\nu) \\
\text{PartiallyOccludesPO}(o,s,\nu) \\
\text{PartiallyOccludesTPP}(o,s,\nu) \\
\text{TotallyOccludesTPPI}(o,s,\nu) \\
\text{TotallyOccludesEQ}(o,s,\nu) \\
\text{TotallyOccludesNTTP}(o,s,\nu)
\end{align*}
\]

In Formula 1, the shadow of a caster \( o \) is the region on a screen \( Scr \) which is totally occluded by \( o \) from the viewpoint of the light source \( L \) (provided that no other object \( o' \) is occluding \( o \) from \( L \)).

Informally, there are three relations between shadows and objects. The first relation, free of occlusions (NonOccludes), is defined when the viewer \( \nu \) can see the whole shadow cast, which can be seen in the image as being disconnected (DC) from or externally connected (EC) to the caster. The second, partial occlusion (PartiallyOccludes), is defined when the cast covers part of the shadow cast and hence the shadow cast appears in the image as overlapped by the caster. Finally, the third relation, total occlusion (TotallyOccludes) is defined when, from the viewpoint \( \nu \), only the caster can be observed in the image, since the shadow is totally occluded by caster.

The PQRS theory inherits a central concept from the RCC8 calculus [15]: the continuous transitions between relations are represented by the Conceptual Neighbourhood Diagram (CND), depicted in Figure 2 [11]. Continuous transitions are represented by arrows between neighbour relations.

Relations expressed by PQRS divide the space into five distinct regions obtained from the distinctions of the perceived relations between a caster and its shadow’s top. This qualitative map is shown in Figure 3, where, for each region, the robot perceives a distinct relation \( \nu \) can see the whole shadow cast, which can be seen in the image as being disconnected (DC) from or externally connected (EC) to the caster. The second, partial occlusion (PartiallyOccludes), is defined when the cast covers part of the shadow cast and hence the shadow cast appears in the image as overlapped by the caster. Finally, the third relation, total occlusion (TotallyOccludes) is defined when, from the viewpoint \( \nu \), only the caster can be observed in the image, since the shadow is totally occluded by caster.

The PQRS theory inherits a central concept from the RCC8 calculus [15]: the continuous transitions between relations are represented by the Conceptual Neighbourhood Diagram (CND), depicted in Figure 2 [11]. Continuous transitions are represented by arrows between neighbour relations.

Relations expressed by PQRS divide the space into five distinct regions obtained from the distinctions of the perceived relations between a caster and its shadow’s top. This qualitative map is shown in Figure 3, where, for each region, the robot perceives a distinct relation between the object-shadow. Hence, any viewpoint \( \nu \) located in Region\(_1\) observes the shadow’s top \( top(s) \) and the object caster \( o \) as NonOccludesDC\((o, top(s), \nu)\).

Similarly, in Region\(_2\), \( \nu \) perceives object-shadow as

![Figure 1: Representation of PQRS relations.](image1)

![Figure 2: Conceptual Neighbourhood Diagram of RCC8.](image2)
Figure 3: The representation of the qualitative map with five distinct regions, showing both sides (right, left) and regions (1, 2, 3, 4 and 5). The lines of sight between the light L, the caster O and its shadows define boundaries between regions.

NonOccludesEC(o,top(s),ν). When ν is located in Region3, it perceives PartiallyOccludesPO(o,top(s),ν), while in Region4 the shadow appears as a stripe on the side of its caster and is thus perceived as TotallyOccludesTPPI(o,top(s),ν). Finally, in Region5 the shadow is behind the object and therefore ν perceives TotallyOccludesNTPPI(o,top(s),ν).

As we can observe, a motion in the environment, if represented by the qualitative map, follows the continuous transition in the Conceptual Neighbourhood Diagram. As the robot navigates from, e.g., Region1 to Region5, it observes the relation sequence EC, DC, PO, TPPI, NTPPI between the object and its shadow. This continuous change on the robot’s perceived object-shadow relation is a qualitative description of the robot’s motion. The changes in perception occur frequently while the observer moves.

IV. PROBABILISTIC INFERENCE

The definition of PQRS is built upon a classical first-order logic language and, therefore, it is not able to handle sensor uncertainty. On the other hand, the robot’s vision system, besides being noisy in general, is dependent on the threshold, which is a key parameter to object-shadow segmentation. Thus, inferences within PQRS are dependent on the image segmentation. In order to achieve a good segmentation, it is necessary to find appropriate thresholds as the robot changes its position. In this work, we extend the PQRS theory by combining it with a Bayesian filter. This filter is used to provide an estimation of the robot’s location and the best shadow-caster segmentation threshold, given the robot’s current state (i.e., the robot’s location w.r.t. a region in the map shown in Figure 3).

We refer to belief distributions as the posterior probabilities over state variables, conditioned on the available data. We use the notation $bel(s_t) := P(s_t|e_{0:t})$; that is, the belief $bel(s_t)$ is the probability of a state ($s_t$) given all evidence ($e_{0:t}$) up to an instant ($t$). In this work, the state $s_t$ indicates the region where the robot is located at instant time $t$, and ($e_{0:t}$) represents evidence denoted in terms of PQRS relations from instant 0 to instant $t$. This evidence is obtained from the shadow-caster images. Thus, we need to find the value of $s_t$, and also of $Th$ (threshold used to extract shadows from images) that maximises $bel(s_t)$.

In order to compute the beliefs, we use:

$$ bel(s_t) = P(s_t, Th|e_{0:t}) $$
$$ = P(s_t, Th|e_{0:t-1}, e_t) $$
$$ = \eta P(e_t, Th|s_t, e_{0:t-1}) P(s_t|e_{0:t-1}), $$

where $\eta$ is a normalisation constant. Using the Markov hypothesis, the next state is independent of earlier measurements $e_{0:t-1}$; hence:

$$ bel(s_t) = \eta P(e_t, Th|s_t) P(s_t|e_{0:t-1}). $$

This expression is the measurement update [20]. Its first term $P(e_t, Th|s_t)$ is, in fact, the image model, which incorporates the new evidence $e_t$. Its second term (known as posterior belief $P(s_t|e_{0:t-1}) = bel(s_t)$) must be calculated before incorporating the evidence $e_t$.

The posterior belief $bel(s_t)$ represents a one-step prediction of the next state $s_t$, obtained by conditioning $s_t$ on the previous state $s_{t-1}$. Consequently,

$$ bel(s_t) = \sum_{s_{t-1}} P(s_t|s_{t-1}, e_{0:t-1}) P(s_{t-1}|e_{0:t-1}) $$
$$ = \sum_{s_{t-1}} P(s_t|s_{t-1}) P(s_{t-1}|e_{0:t-1}). $$

In the previous expression, the term $P(s_t|s_{t-1})$ is the state transition model. In this work the transition model $P(s_t|s_{t-1})$ refers to qualitative motion in one of the five regions in Figure 3; motions can be either clockwise or counter-clockwise. The state $s_t$ is a Region$i$ where $i = 1, 2, 3, 4, 5$. Table I depicts the transition model $P(s_t|s_{t-1})$.

We then obtain $P(s_{t-1}|e_{0:t-1})$ by conditioning on $Th$:

$$ P(s_{t-1}|e_{0:t-1}) = \sum_{Th} P(s_{t-1}, Th|e_{0:t-1}). $$

Therefore, the posterior belief (or prediction) is the combination of previous expressions:

$$ bel(s_t) = \sum_{s_{t-1}} P(s_t|s_{t-1}) \sum_{Th} P(s_{t-1}, Th|e_{0:t-1}). $$

Finally, the belief is given by:

$$ bel(s_t) = \eta P(e_t, Th|s_t) * bel(s_t). $$

In the next section, we explain the self-localisation and self-calibration algorithms, which combine the probabilistic belief formula given above with the PQRS theory.
V. THE SELF-LOCALISATION, SELF-CALIBRATION ALGORITHM

The robot estimates its relative location using a monocular colour camera. From its viewpoint, the robot perceives a relation between shadow and object. The PQRS theory then gives the robot the ability to infer its position. A light intensity threshold segments the image into object and shadow. We use a morphological operator along with saturation properties to perform segmentation; after that, the algorithm looks for the shadow connected to the object’s base. Figure 4 shows an output of this segmentation procedure.

Algorithm 1 initialises the system with a uniformly distributed uncertainty about the robot’s position (line 1). Line 2 runs a threshold calibration, assuming that a good shadow-caster recognition for the first frames was provided. The frame sequence is represented by the discrete variable $t$.

Subsequently, Algorithm 1 repeatedly calls Algorithm 2 (PERCEPTION-ACTION), which returns the evidence with maximum probability from $P(E)$. In Algorithm 2, PQRS relations are evaluated by some detected features, such as the degree of connectivity between the side of the shadow’s top and the caster, the degree of connectivity between shadow-object, the shadow’s bounding-box size and also by the relative position of the object’s bounding box. The algorithm returns the relation with the greatest number of features perceived. This evidence can be either $DC$, $EC$, $PO$, $TPPI$ or $NTPPI$. If the object has no shadow, only the relation $NTPPI$ is analysed. Then, given the evidences, line 5 of Algorithm 1 calculates probabilities for all five regions (here represented by $S_t$) and a threshold range $[Th-5, Th+5]$ respecting the boundaries $0 \leq Th \leq 255$. Then, line 6 returns the beliefs about the current region ($s_t$) and the threshold ($Th$) to be used with the next image. The last two lines of Algorithm 1 run a Bayesian prediction step, where the posterior belief $\overline{bel}(S_{t+1})$ is calculated in line 8, which takes into account the current beliefs for all states (given the evidence), as calculated in line 7. The probability of the next states (given the current state) is the motion model given by Table I.

In Algorithm 1 we have abbreviated the PQRS relations as follows: $NODC(O, top(S), \nu)$ represents $NonOccludesDC(o, s, \nu)$; $NOEC(O, top(S), \nu)$ represents $PartillyOccludesPO(o, s, \nu)$; $TPPI(O, top(S), \nu)$ represents $TotallyOccludesTPPI(o, s, \nu)$ and $NTPPI(O, top(S), \nu)$ represents $TotallyOccludesNTPPI(o, s, \nu)$.

The Bayesian filter handles two probability models, the sensor model and the transition model. In Algorithm 1 our sensor model is the image model (line 5). The motion model represents the transition between regions in the map and is used for the calculation of the posterior belief (line 8). Next, we describe how the image and motion models were designed.

A. The image model

The image model is represented by $P(e, Th|s_t)$, which indicates what we should expect from sensor data when in a given state. In this work, each evidence is linked to a region and this idea is formalised within the PQRS theory. However, the evidences depend on the threshold used in the vision system. As a result, we need a model that indicates the frequency of which a threshold gives the right evidence to each region. In order to obtain that, we constructed a model with a set of images captured by the robot. In this set,
the target object-shadow was always in the scenes observed. This guarantees that evidence was always available to the vision system at least for one given threshold during the evaluation of the image model. For each threshold value, the set of images was analysed. The results obtained form a matrix with information about the percentage of right answers found per threshold.

### B. The motion model

The transition, or motion model, represents the changes between the robot’s states, given a moving action. Considering the qualitative map (Figure 3), and steady-speed clockwise motion around the target object-shadow, the probability of the changes between regions is mostly dependent on the regions’ sizes. This is represented in Table I. The multiplying factor $w$ represents perceived shadow features that modulate the probability for a state change. This can be understood by taking into account that, in this work, the robot’s motion is inferred from image changes w.r.t. shadow features. Thus, when the robot approximates a borderline region, the detection of shadow features related to the next region should increase the probability of a state change.

### VI. Experiments and Results

In our experiments an Adept MobileRobots Peoplebot mobile robot collected snapshots in an office-like environment. In order to make the results comparable to [8], we used the same data as in [8], where the robot collected snapshots around a target object (i.e., the black bucket in Figure 4). It is worth highlighting that the target was not always within the camera’s field of view. Images were represented by the Hue-Saturation-Value (HSV) colour space, and the shadow-object detection was performed with a threshold filter on V (light intensity).

Algorithm 1 uses the belief about the robot’s location and what is perceived from the scene. This algorithm returns both the robot state and the threshold. The robot was teleoperated and could start in any of the regions indicated in Figure 3. While the robot navigated around the object, the operator marked the real region (gold standard) to compare with the robot’s answers. The lines on the floor (Figure 4) were drawn to help the operator to distinguish the correct robot’s location w.r.t. the map in Figure 3.

The approach presented in this paper was applied to 587 snapshots of the target object-shadow. The results of the localisation procedure (in the map shown in Figure 3) are presented in a confusion matrix (Table II). Each column represents our system’s outputs while the rows represent the actual region in which the robot was located (according to the gold standard).

In Table III, rows refer to the true location for each image (regions 1, 2, 3, 4 and 5). The column “# images” is the number of images captured for each region. The results from [8] are presented in column “Knowledge based (%).” The results shown in column “Belief based” refer to the percentage of right regions returned by the approach proposed in this paper. The global performance of the current experiments presented an increase from 58% (obtained in [8]) to 80%. On the boundary between two regions (where the shadow-caster features give equal evidence for two locations) the system chooses the first region as the robot’s location. This choice reflects positively on Regions 1, 2 and 4, but on Regions 3 and 5 the accuracy decreased in comparison to our early experiments.

Table II shows that the highest number of false negatives for Region3 was related to locating the robot in region 2 (21% of the cases). Part of this error is due to the borderline problem explained in Section V-B. Another issue is that the herein proposed system never detected when the robot was located in Region3. This poor performance when in Region5 is due to the motion model, where the relation Totally Occludes was not considered in the composition of the weight $w$ (because the weight is calculated on the existence of shadows). We are currently investigating possible solutions to the borderline problem and to the issue of detecting Region5.

### VII. Conclusion

In this paper we have proposed an approach to robot self-localisation using information about shadows in images, where we have combined a probabilistic Bayesian filter with a qualitative spatial representation. Our method uses an adaptive threshold with a Bayesian filter to produce
better results than a purely qualitative localisation approach. Moreover, considering that the dataset used was composed of 30% of images where self-localisation was impossible (since the target object was not present in these data points), our algorithm had a performance of around 80% over all images. This suggests the robustness of a probabilistic approach to infer qualitative facts about space. Future work shall consider an extension of the method proposed in this paper in order to allow robot self-localisation within an environment with various objects.

ACKNOWLEDGMENTS

Valquiria Fenelon is supported by a scholarship from CAPES. Paulo E. Santos acknowledges support from FAPESP grant 2012/04089-3 and CNPq, grant PQ2 - 303331/2011-9. Murilo F. Martins acknowledges support from FAPESP grant 2012/12640-1. Fabio Cozman is partially supported by CNPq.

REFERENCES


[18] Christoph Schlieder, Representing visible locations for qualitative navigation, Qualitative Reasoning and Decision Technologies, CIMNE, 1993, pp. 523–532.

