A Low-level Framework for a Probabilistic Treatment of the Topological Description of a Robot Mission

F. Ferreira and V. Santos
Department of Mechanical Engineering
University of Aveiro, 3810-193 Aveiro, Portugal
Email: cfferreira,vsantos@mec.ua.pt

J. Dias
Institute of Systems and Robotics
University of Coimbra, 3030-290 Coimbra, Portugal
Email: jorge@isr.uc.pt

***Abstract*— This article describes a mathematical basis required to integrate features obtained for perception for topological navigation. It is intended for application to navigation in an environment that is not mapped, but in which a mission is described in the form of a semantic description of the perception stimulus that the robot is expected to encounter. The need to integrate features from different sensors led to the use of an uncertainty estimate employed in information theory; binary entropy. By using entropy, the features are ranked in order of decreasing uncertainty. This article describes the state of the work in an as yet preliminary stage, but appears promising for application to navigation using topological information. It also offers interesting perspectives on commonly used sensory data such as local intensity image features.**

***Index Terms*— Binary entropy, Autonomous navigation, Topological Features.**

I. Introduction

Mapping applications, using either metric or topological methods, or both, have improved immeasurably to the present day. Single robots [1][2] or teams [3] are now able to effectively use spaces meant for humans and even replace the humans in applications in these spaces.

Although much work has been done on mapping of indoor environments, a large set of applications that require navigation in large indoor and outdoor environments have still not been satisfactorily resolved; for example, applications in which semantic descriptions or navigation cues, placed for human users of the space, are given to the robot, instead of requiring the robot to map the entire environment. Such applications would involve a topological description for a single path through the environment in terms of the landmarks that will be encountered. The robot would selectively integrate indicated and other landmarks and features based on their marginal information contribution.

Researchers have been tackling with the more complex problem in which a topological representation of the complete environment is provided and the robot must perform a given motion subject to certain constraints. Markov models have justifiably been used as representations of the physical motion of the robot in the environment. The uncertainty in the pose of the robot and in the perceptual capabilities of the robot have resulted in the modeling of the position [given-the-perception] as a Hidden-Markov-Model. When the requirement of autonomous navigation and path planning (in greater or smaller degree) is added to localization problem, tools dealing with the stochastic treatment of Markov processes have been applied. Kaelbling [4] explains the need for and the means to model navigation using a topological mapping as a Partially Observable Markov process, with a special emphasis on the motion. Simmons [5], uses a similar approach emphasizing and developing a model of the pose of the robot instead. This article describes work that borrows the Markovian means of handling the Bayesian inference when perception is performed using the topological approach. Its focus, however, is on the information content of the features used in the perception process.

This article and the work on the robot platforms which it describes aims for the integration of different sensors including vision, stereo vision and a laser range finder. Presently, a partial system has been developed to demonstrate the framework using only a laser rangefinder (LRF). The addition of other sensory systems will use an approach similar to the one adopted for the LRF.

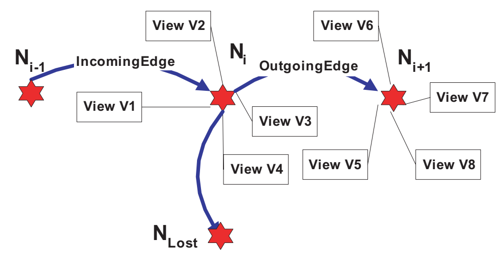

Fig. 1. Excerpt of a topological map indicated the edges and nodes

At this stage, a brief description of the terminology utilized in this work is in order. The **Mission** refers to a sequence of movements that the robot must perform to get from a starting place to a final destination. The robot will depend on features obtained from sensing the environment and associated with various places while completing the mission; **Views** refer to collections of features at discrete instances. For example, views can be composed of SIFT features in images [6][7], doors [8], lines and corners in a laser range image; each feature is arranged in a feature 'space' defined independently for each type of sensor. **Affordances**, a reference to the term 'affordance' coined by J J Gibson [9], refers to the possibilities of motion that the environment provides locally. **Behaviors** refer to motion that follows a particular control law.

The mission that a user presents to the robot is built into a graph description. The path indicates the topological path that the robot must travel, see Fig. 1. The topological description of the mission takes the form of a path in the Graph $G = \langle N, E, I \rangle$ where N indicates the set of **nodes**, E the set of **edges** and I the set of incidental relationships that unite nodes and edges. Each node is an 4-tuple $N_i = \langle V, TV, M, TM \rangle$ where V indexes named views, TV the topology of the views, M the affordances at the node, and TM the topology of these affordances. Each edge $E_k = \langle M, R, D \rangle$ is denoted as a triple where M defines a single or a sequence of behaviors, R denotes the reason for termination of the edge and D represents the distance covered since the beginning of the edge. The incidence relationships that join nodes to edges are denoted by $I_k = \langle b_i, N_i, N_j, b_j \rangle$ which indicate the initiating behaviors, starting node, ending node and ending behavior in that order. The path in the topological graph is the exhaustive set of edges from the first node, $N_{initial}$, to the final node N_{final}.

Before a robot can perform a mission autonomously, the user leads the robot through the environment in an environment-familiarization phase. This environment-familiarization phase is completed by alternating between the user choosing behaviors (at a node) and the robot executing these behaviors. As the robot travels with the behaviors specified, it automatically detects new possibilities or its inability to perform the current behavior: resulting in the creation of a new node and a new set of behavior choices for the user.

The Views containing the features and the actions associated with taking the robot from a Node to the next one are recorded in a 'perception string'. During autonomous navigation, successful identification of previously 'recorded' features allows for an estimation of the current position of the robot in the mission. A key problem remains, that of evaluating the 'quality' of the prediction and a comparison of the quality across different features obtained by different sensors.

To tackle this problem, the information obtained in the environment-familiarization stage is propagated backwards, along the path, to obtain a measure of the importance of each feature to each node and edge in the path. This measure, defined in terms of the conditional entropy of the feature with respect to some persistant feature such as the distance covered or the time elapsed, is then used to obtain an ordering of the features denoting their 'quality'. The use of 'good' features along the mission coupled with information gained from motion should improve the chances of the robot successfully staying on the path and completing the mission. It shall also allow the use of long-viewing sensors such as vision for localization by explicitly quantifying the quality of the information that these sensors provide.

II. Uncertainty of robot Position along the Mission

The method described in this article is inserted in a larger work that aims at the integration of sensory data obtained from a number of sensors. The work is to be implemented on

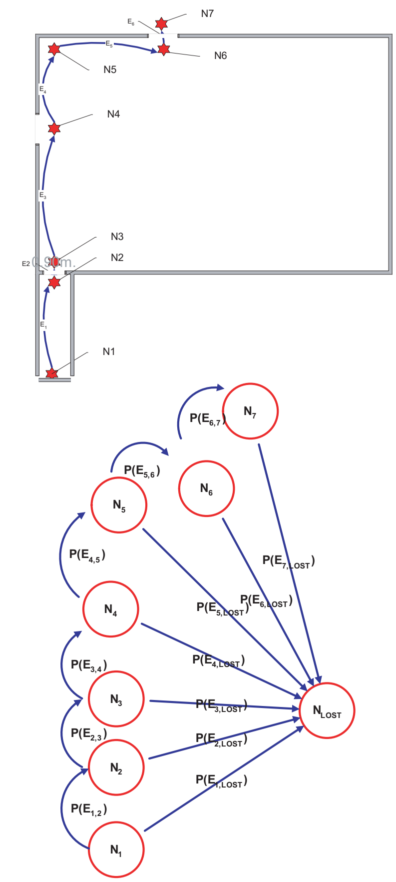

Fig. 2. Layout of the Path and a topological representation of the mission

two, very different robot platforms equipped with different sensors. A modular sensor approach is required to fuse features from different sensors without a calibration of the complete sensor setup; the latter being a solution that would be difficult to port to another robot platform.

The uncertainty in the robot location has been defined in terms of the uncertainty of moving from one node to the other. Accordingly, the robot is localized if at any Node N_i, the correct OutGoingEdge is initiated given that the InComingEdge was correctly initiated. All the possibilities that make the robot perform some motion that does not correspond to the OutGoingEdge edge lead the robot to places that are not on the path, and are collectively called the 'Lost' node (shown in the topological representation in Fig. 2 as N_{LOST}). Thus, if an edge is selected at a node, then either

1) It will arrive at the next node N_{i+1} on the mission.
2) It will get off the path, denoted by reaching a lost node N_{LOST}.

The probability of getting from one node to another depends on the number of different affordances at a node. It also depends on the uncertainty of the behavior once it is chosen, though, in this exercise this source of uncertainty has been neglected.

Let the topological map of the positions that the robot can take be composed of a random variable X which takes values in the set $1, 2, 3, \ldots, n$ corresponding to the robot occupying each of the nodes $N_1, N_2, \ldots, N_n$ respectively, i.e. $X = i$ when $N = N_i$.

During navigation, another Node N_{LOST} is added to account for the robot being at some unknown position as seen in Fig. 2. Without loss of generality, we can allow X to take the value of -1 when the robot is at N_{LOST}. The robot is said to be localized when $1 \geq X \geq n$ and lost when $X = -1$.

The change in position as the robot executes the mission can be modeled as a Markov process. A single-step transition probability matrix denotes the probability of moving successfully to the next node in the mission or getting 'lost' at each node. It depends on the layout of the environment and the the reliability of the edge traversing behaviors. During autonomous navigation the robot is not certain of detecting that it has arrived at a node nor is it certain of correctly identifying the node. The single-step transition probability transition matrix in the case of navigation is shown in Fig. 3. If the next node is always identified (never skipped), then this matrix indicates the chances of the robot moving one step-at-a-time along the mission and the probability of getting off the path at each step. In case the detection of a node is not deterministic, there is a non-zero probability that the robot skips one or more nodes ahead, and this is obtained by multiplying appropriate rows of the matrix.

$$
\begin{array}{c}
\\ X{=}1 \\ 2 \\ 3 \\ \vdots \\ n-1 \\ n \\ -1
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} X{=}1 & 2 & 3 & \ldots & n-1 & n & -1 \end{array} \\
\left[\begin{array}{ccccccc}
0 & p_1 & 0 & \ldots & 0 & 0 & (1-p_1) \\
0 & 0 & p_2 & \ldots & 0 & 0 & (1-p_2) \\
0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 0 & p_{n-1} & (1-p_{n-1}) \\
0 & 0 & 0 & \ldots & 0 & 1 & 0 \\
0 & 0 & 0 & \ldots & 0 & 0 & 1
\end{array}\right]
\end{array}
$$

Fig. 3. Single-step transition probability matrix in the case where the robot is navigating autonomously.

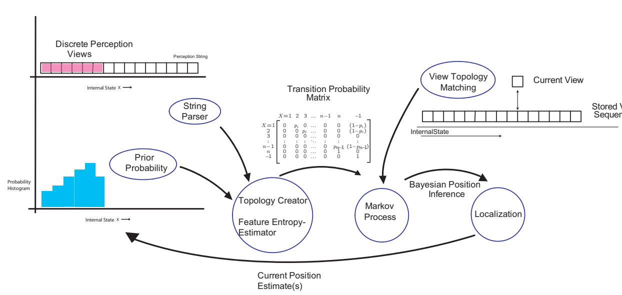

Fig. 4. Schematic of the Complete procedure adopted for localization.

The localization itself is performed by maintaining one or more hypothesis of the current robot position. A single action is associated with each position, see Fig. 4. The action associated with the winning node is chosen whatever the strategy used to pick this winning node. It must be emphasized that the focus of this article lies solely in the study of the information content of the features.

III. Uncertainty in the conditional Estimation of the Position given a Feature

Let Y be a random variable which denotes the state of perception of a single feature. Y can take the values $y_1, y_2, \ldots, y_m$ denoting 'm' distinct and detectable states that the feature can take.

Y is correlated with the position X of the robot. The correlation may not be perfect however (the function is not injective), indicating that the same value of Y can be obtained at more than one position (X).

The aim of using the feature Y in any process of perception is to estimate the position that the robot occupies given the Y equals some value y_t, that is to obtain $P(X|Y = y_t)$. For an efficient use of the features, the probability distribution of $Y = y_1, y_2, \ldots, y_m$, of $X = x_1, x_2, \ldots, x_n$ and the relationship between Y and X, i.e. $P(X|Y = y_t)$, must be captured in some way.

The approach that has been described here attempts to evaluate the uncertainty in the probability distribution for this conditional relationship, $P(X|Y = y_t)$. This uncertainty has been modelled in terms of the *binary entropy* of the distribution. As might be expected, the effect of the uncertainty in the position of the robot itself, and in the relation between perception and position as captured in the environment familiarization stage, and the values that Y can take, all contribute to the uncertainty of the above relation.

The Entropy of the random variable was described by Shannon [10] in terms of the probability distribution as shown in (1).

$$H(X) = -\sum_x p(X = x) Log_2(p(X = x)) \tag{1}$$

This value of the entropy is expressed in the number of bits and takes significance as a measure of the uncertainty of the distribution (see [11] for an excellent introduction).

We can write the joint entropy of X and Y in terms of conditional and independent entropy as in (2),

$$H(X, Y) = \quad H(X) + H(Y|X) = H(Y) + H(X|Y) \tag{2}$$

By combining the above two equations we can write the expression (3), which includes the **mutual information** term $I(Y; X)$.

$$
\begin{aligned}
H(X|Y) &= \quad H(X) + H(Y|X) - H(Y) \\
&= \quad H(X) - (H(Y) - H(Y|X)) \\
&= \quad H(X) - I(Y; X)
\end{aligned} \tag{3}
$$

A couple of inferences about the desirable behavior of each of the above terms can be immediately obtained from (3).

1) The entropy $H(X)$, of the random variable X. During autonomous navigation, the transition matrix now includes an extra node N_{LOST} and the probabilities are less concentrated. Thus $H(X) > 0$ during navigation as compared to the case when it is equal to zero

during the environment-familiarization phase. A change in the probability distributions for X changes both the conditional entropy of the feature and entropy $H(X)$. The change in the value of X is still a Markovian process though.

2) The entropy $H(Y)$ of the random variable Y. This term measures the uncertainty of the values taken by Y. A higher entropy is a desirable characteristic as this term is a subtractive component. A random variable Y which can take as large a number of distinct states with a uniform probability distribution is desirable, as it will result in a higher entropy.
3) The entropy $H(Y|X)$ of Y given X. This entropy must be minimized which is to say that the features should be as correlated to each of the values of X as possible (or Y must be an injective function X). In various localization and navigation algorithms developed till now, this term has been reduced by limiting the range over which the an environmental property affects the perception. Thus, sensing is limited to the near vicinity of the pose of the robot and the perception is recorded at positions that are sufficiently distinct from each other.

IV. Evaluating the Entropy terms

A. *Evaluating* $H(X)$

The entropy of the distribution of X is a function of the total number of nodes in the path and of the number of nodes that are yet to be covered to complete the mission. Since changes in X are Markovian, only the position of the last node need be maintained to estimate the new position, i.e. if X changes from $1 \rightarrow 2 \rightarrow \ldots \rightarrow i \ldots \rightarrow j$ then the probability of the transition from i to j is given by (4).

$$\begin{aligned} P(X = j|X = 1, X = 2, \ldots X = i) \\ = \quad P(X = j|X = i) \end{aligned} \tag{4}$$

The entropy of the random variable X can be evaluating using (5). Two different approaches might be used to obtain the probability distribution required to evaluate $H(X)$; not maintaining a prior estimate of its position (assuming all values of X as equally likely at all moments) or maintaining an estimate of its last known position before it moved. Greater the number of nodes left to be covered and lesser the precision with which the robot's position is known, greater is the entropy.

$$H(X = 1) = \quad \sum_j H(X = i|X = j), \forall i \neq j \tag{5}$$

Case: Robot has no prior estimate of its position. If the robot has no estimate of its current position, then, each of the nodes (including the lost node N_{lost}) is equally likely and the entropy is given by (6). The affordances of the environment are not considered and, hence, the transition probability matrix of Fig. 3 is not considered.

$$H(X)_{\text{no prior estimate}} = \quad \frac{1}{n+1} log_2(n + 1) \tag{6}$$

Case: Robot has an estimate of where it started out from. Although the robot moves along the edge that will take it to the next node, it is not sure of detecting this next node and might fail to identify this and other nodes. If the probability of correctly identifying a node 'i' is denoted by $P(X_{detect\_i})$, then, using the transition probability matrix in Fig. 3, the entropy $H(X)$ might be calculated using the recursive equation of the form in (7).

$$\begin{aligned} H(X|X = i) = \quad & p_i log_2(p_i) P(X_{detect\_i}) + \\ & (1 - p_i) log_2(1 - p_i) + \\ & (1 - P(X_{detect\_i})) \times (H(X|X_{i+1})) \end{aligned} \tag{7}$$

Thus, we see that as the number of nodes travelled increases, the entropy reduces, not as a result of the perception, but because of the number of nodes yet to be reached are reduced.

Also, in a more robust set-up, if multiple hypothesis of the position are to be maintained, using $P(X|Y)$, then a separate different value of $H(X|X_i)$ for each estimate of the starting node N_i must be maintained.

B. *Evaluating* $H(Y|X)$

The conditional entropy of Y given X is calculated according to (8).

$$\begin{aligned} H(Y|X) = \quad & P(X = 1) \times \\ & \textstyle\sum_k P(Y = y_k|X = 1) log_2(P(Y = y_k|X = 1)) \\ & + P(X = 2) \times \\ & \textstyle\sum_k P(Y = y_k|X = 2) log_2(P(Y = y_k|X = 2)) \\ & + \ldots \\ & + P(X = n) \times \\ & \textstyle\sum_k P(Y = y_k|X = n) log_2(P(Y = y_k|X = n) \\ = \quad & \textstyle\sum_i \{ P(X = i) \times \\ & \textstyle\sum_k \left[\frac{-\delta_{X=i,Y=y_k}}{|Y=y_k|} log_2 \left(\frac{\delta_{X=i,Y=y_k}}{|Y=y_k|} \right) \right] \} \end{aligned} \tag{8}$$

where $\delta_{X=i,Y=y_k}$ is the Kronecker delta function and is equal to '1' if feature Y takes value y_k when $X = i$ and 0 otherwise, and $|Y = y_k|$ is the total number of times that Y takes the value y_k over all the nodes $X = 1, ...n$.

In case there is uncertainty (type 1 and type 2 estimation error) in the detection of features, this error can be included by substituting the Kronecker delta function by the probability $P(Y = y_j|X = i)$ for all y_j.

There is still the question of the perception at the lost node, N_{LOST}. Since it cannot be included during the environment-familiarization phase, we have no data with which to explicitly estimate $H(Y|X = -1)$. Instead, we can assume that each outcome of Y is likely with, the *same probability distribution* that it has over the mission, i.e. over $X = 1, ..n$.

$$\begin{aligned} H(Y|X = -1) = \quad & P(X = -1) \times \\ & \textstyle\sum_k \left[\frac{-P(Y=y_k)}{m} log_2 \left(\frac{P(Y=y_k)}{m} \right) \right] \end{aligned} \tag{9}$$

C. Evaluating $H(Y)$

The entropy of the random variable Y is obtained from the probability distribution from all the values that Y can take over the entire mission as shown in (10).

$$\begin{aligned} H(Y) = \ & \textstyle\sum_k P(Y = y_k | X = 1, 2, \ldots, n) \times \\ & log_2(P(Y = y_k | X = 1, 2, \ldots, n)) \end{aligned} \tag{10}$$

The more the number of states that a feature can take, greater is the information that the feature brings to the perception process. Thus a feature that can take only one of two states brings with it the least amount of information.

The probability distribution of the states themselves also affects the term $H(Y)$. Other factors being equal, the greatest amount of information is carried by a feature in which the states are equally likely.

V. The Software Architecture and the Perception Process

In Fig. 5 the architectural layout of the setup is shown with the layers of software that seek to implement reactive behavior at the bottom, and those concerned with deliberative behavior at the top. The communication between objects occurs using TCP sockets. The Sensors have been developed as independent objects that service requests as they arrive.

The perception process consists of self-contained 'sensor classes' that are entrusted with the task of extracting features from the current sensor data. These features are matched against a previously created database of features and provide the localization modules with a list of feature IDs that correspond to the current features being observed. The database of features can be explicitly specified as in the case of the Laser Range Finder sensor that detects doors, walls and corners, or it might be built automatically as in the case of vision, where the database of conspicuous local image features (described by their SIFT descriptors[6]) are collected during a previous data-familiarization phase. It has been considered that the database does not change during the course of the mission, which itself consists of traversing a path completely and just once. However, new features might be added to the database for use in a subsequent mission.

VI. Experiments

A mission whose layout is depicted in Fig. 2 was simulated, with the topological map of the path shown in the image at the bottom of the same figure. The aim of this exercise is to verify the evolution of the conditional entropy of the position of the robot given each of the features. More specifically, the conditional entropy will be compared for different features and under two different assumptions i.e. given no prior probability distribution and, given a prior probability of the position. Four features were compared,

1) $Y1 =$ A Door; can take one of three states, no door in sight, door in sight and next to a door.
2) $Y2 =$ A Corner; can take one of three states, no corner in sight, corner in sight and next to a corner.

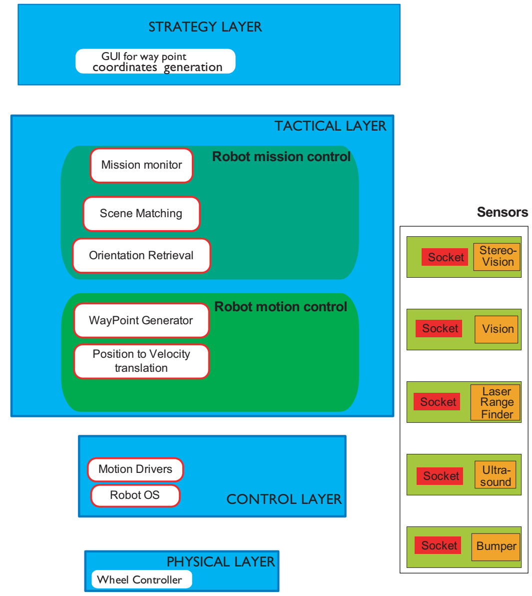

Fig. 5. Layered architecture with low-level reactive layers, intermediate 'drivers' and high-level layers concerned with topological localization.

3) $Y3 =$ A Unique marker at End Point; can take one of two states, unique marker is not visible, unique marker is visible. Local SIFT features [6][7] in an intensity image can be represented using this type of feature.

The values taken by each of the above features at each of the nodes of the path shown in Fig. 2 is shown in the table in Fig. 6. The entropy values, calculated using the formulae in the earlier section are presented in the table in Fig. 7 and the final $H(X|Y)$ values are plotted in the graphs shown in Fig. 8. It can be seen, from Fig. 8, that the conditional entropy

Feature \ Position	X = 1	X = 2	X = 3	X = 4	X = 5	X = 6	X = 7
Y = A door = {No_Door=1,See_Door=2,At_Door=3}	2	3	3	3	1	3	3
Y = A corner = {No_Corner=1,See_Corner=2,AtCorner=3}	1	1	1	2	3	2	1
Y = EndMarker = {NoEndMarker=1,SeeEndMarker=2}	1	1	1	1	1	1	2

Fig. 6. Values taken by the random variables Y_1 to Y_3 during the environment-familiarization phase in the mission from Fig. 2.

Cond. Entropy\ Initial Posn Est	X = 1	X =2	X = 3	X = 4	X = 5	X = 6
H(Y1|Xinitial)	0.58	0.70	0.74	0.77	0.76	0.82
H(Y1)	1.38	1.38	1.38	1.38	1.38	1.38
H(Y2|Xinitial)	0.58	0.70	0.75	0.74	0.78	0.82
H(Y2)	1.38	1.38	1.38	1.38	1.38	1.38
H(Y3|Xinitial)	0.19	0.19	0.19	0.20	0.21	0.26
H(Y3)	0.59	0.59	0.59	0.59	0.59	0.59
H(X|Xinitial)	1.86	1.72	1.60	1.50	1.34	0.90

Fig. 7. Calculated values of the conditional entropy for the simulation. The values are plotted in the lower graph in Fig. 8.

of the robot position given a prior position of the robot is

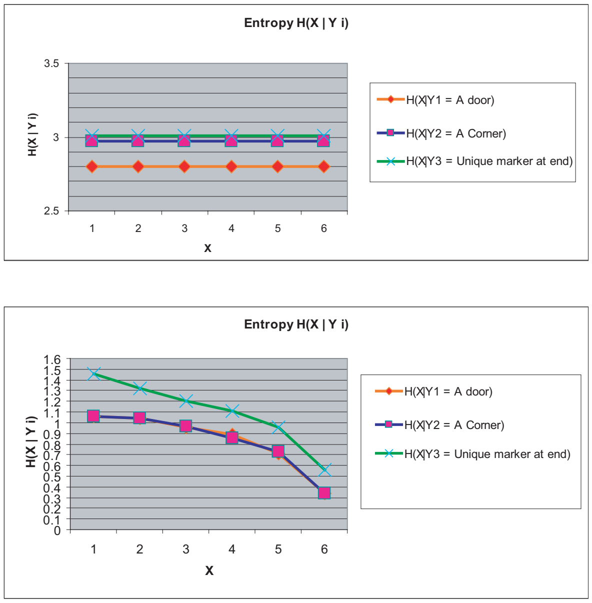

Fig. 8. Conditional Entropy of the estimate of the position of the robot without (top) and, given an estimate of the position of the robot and that the robot has moved (bottom).

always lower than if no prior probability is included. While the uncertainty, H(X), contributes most to the conditional entropy depicted in the graphs in Fig. 8, it affects all the features in a similar way allowing us to make a comparison between the entropy of each of the features.

Features one and two are similar in nature despite depicting different landmarks. The number of states that the features can take could be increased, thereby improving the discerning ability of the features. Also, an aspect that has not been touched upon is the cross-correlation between features. It can be seen that the entropy of feature one and two are similar but a compound feature, a function of these features, might offer a much lower entropy given the larger number of possible states and also the lack of correlation between the states of feature one and two. A complete enumeration of the joint states of these two features would result in the creation of another feature with more states and, hence, greater entropy.

In the case of feature three, the unique marker at the end point, it can be seen that while it is of relatively little use along most of the mission, it dramatically improves in 'quality' as the end point approaches. While the relative entropy $H(Y|X)$, of this feature is very low, the entropy $H(Y)$ of the variable is also low, resulting in a poor 'quality' feature over most of the mission. Its usefulness increases greatly in the vicinity of the position where it is expected. Thus, maintaining a prior estimate of the position has a considerable influence on the usefulness of this type of feature, the type of feature that we intend to use extensively in our work.

VII. Summary and Conclusions

A mathematical formulation that can evaluate the uncertainty of the estimate of the position of the robot as provided by a single feature was developed. The different sources of uncertainty that contribute to the uncertainty of this conditional probability are identified and evaluated using consistent reasoning. The method allows the use of very different types of feature and even allows the detection of features to be probabilistic. The approach can be extended to any final scheme for localization ranging from a heuristic method that chooses the most likely position to a more elaborate system using filters and/or the cross-correlation of the conditional probabilities $P(X|Y)$.

While the experimental validation of this exercise was only partially complete, the exercise provides a basis with which to evaluate the sensory information required to perform navigation along the path and maintain a high probability of completing, successfully, the mission.

This article describes work that is being carried out to evaluate the information that is provided for the inference of the position change in an autonomous navigation system. Future work includes the integration of the information provided by two or more features (detected by the same sensor or by different sensors).

VIII. Acknowledgements

The first author is a holder or a scholarship from the Instituto de Investigação of the University of Aveiro. The authors wish to thank four anonymous reviewers whose comments helped clarify the text and aided the comparison of the method with work performed by other researchers.

References

[1] Emilio Remolina and Benjamin Kuipers. Towards a general theory of topological maps. *Artificial Intelligence*, (152):47–104, 2004.

[2] Sebastian Thrun, Maren Bennewitz, Wolfram Burgard, Armin B. Cremers, Frank Dallaert, Dieter Fox, and Dirk Hahnel. Minerva: A tour-guide robot that learns. In *KI - Kunstliche Intelligenz*, 1999.

[3] Kurt Konolige, Dieter Fox, Charlie Ortiz, Andrew Agno, Michael Eriksen, Benson Limketkai, Jonathan Ko, Benoit Morisset, Dirk Schulz, Benjamin Stewart, and Regis Vincent. Centibots: Very large scale distributed robotic teams. In *Proceedings ISER2004*, 2004.

[4] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. Technical Report CS-96-08, 1996.

[5] Reid Simmons and Sven Koenig. Probabilistic robot navigation in partially observable environments. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1080–1087, 1995.

[6] David G. Lowe. Distinctive image features-from scale-invariant keypoints. *International Journal of Computer Vision*, 2003.

[7] Filipe Ferreira, Jorge Dias, and Vitor Santos. Convergence of methods for multi-scale image analysis for object and image recollection. In *8TH Conference On Intelligent Autonomous Systems (IAS-8)*, 2004.

[8] Zhiyu Xiang, Vitor M. F. Santos, and Jilin Liu. Locating and crossing doors and narrow passages for a mobile robot. In *First Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO)*, 2004.

[9] James J. Gibson. *The Senses Considered as Perceptual Systems*. Greenwood Press, 1983.

[10] Claude Elwood Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:623–656, October 1948.

[11] Thomas Cover and Joy Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. Wiley InterScience, 1991.