# IMPROVED OCCUPANCY GRID LEARNING
## *The ConForM Approach to Occupancy Grid Mapping*

Thomas Collins, J. J. Collins and Conor Ryan
*Department of Computer Science and Information Systems, University of Limerick, Ireland*

Keywords: Robotics, Occupancy Grids, Machine Learning, Artificial Intelligence.

Abstract: A central requirement for the development of robotic systems, that are capable of autonomous operation in non-specific environments, is the ability to create maps of their operating locale. The creation of these maps is a non trivial process as the robot has to interpret the findings of its sensors so as to make deductions regarding the state of its environment. Current approaches fall into two broad categories: on-line and offline. An on-line approach is characterised by its ability to construct a map as the robot traverses its operating environment, however this comes at the cost of representational clarity. An offline approach on the other hand requires all sensory data to be gathered before processing begins but is capable of creating more accurate maps. In this paper we present a new means of constructing occupancy grid maps which addresses this problem.

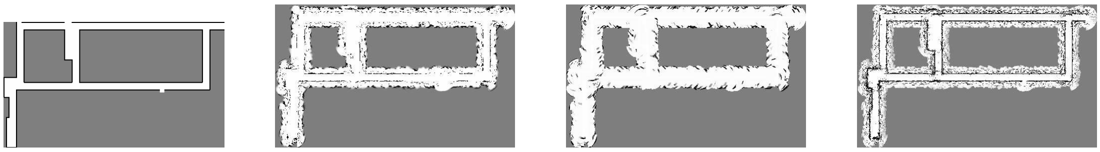
## 1 INTRODUCTION

In recent times Occupancy Grids have become the dominant paradigm for environmental modelling in mobile robotics (D. Kortenkamp and Murphy, 1998). An Occupancy Grid is a tessellated 2D grid in which each cell stores fine grained qualitative information regarding which areas of a robots operating environment are occupied and which are empty (Moravec and Elfes, 1985; Elfes, 1989). Specifically, each individual cell in the grid records a certainty factor relating to the confidence that the particular cell is occupied. Such maps are extremely useful for mobile robotic applications as they facilitate tasks such as navigation, path planning, localisation and collision avoidance (Borenstein and Koren, 1991; Dissanayake et al., 2001).
Currently in the Occupancy Grid mapping domain there are two broad approaches: on-line and off-line. The on-line approach is characterised by traditional paradigms such as those from Moravec (Moravec and Elfes, 1985), Matthies (Matthies and Elfes, 1988) and Konolige (Konolige, 1997). The off-line approach has emerged from a more recent paradigm from Thrun (Thrun, 2003). The on-line approach is capable of generating maps in real-time as the robot operates. However these maps often contain inconsistencies such as over estimation of occupied or free space which is undesirable. The off-line approach on the other hand, is capable of generating more consistent maps but cannot do so in real time. These diametric approaches give rise to a mode versus clarity dilemma.
In this paper we introduce and empirically evaluate a novel robotic mapping framework called ConForM (**Con**textual **For**ward **M**odelling) which solves this dilemma through combining the beneficial aspects of both existing approaches. Results from empirical evaluations we have undertaken show that ConForM provides maps that are of better quality than existing paradigms.

## 2 ON-LINE VS. OFFLINE OPERATION: THE ROBOTIC MAPPING DILEMMA

Two types of model are available for sensory interpretation in robotic mapping. These are the *Inverse* and the *Forward* models (Thrun, 2003). An inverse model attempts to describe an environment by trans-


Collins T., J. Collins J. and Ryan C. (2007).
IMPROVED OCCUPANCY GRID LEARNING - The ConForM Approach to Occupancy Grid Mapping.
In *Proceedings of the Fourth International Conference on Informatics in Control, Automation and Robotics*, pages 492-497
DOI: 10.5220/0001651904920497



(a) Ideal Map (b) Inverse model: Moravec and Elfes 1985 (c) Inverse model: Thrun 1993 (d) Forward model: Thrun 2001

Figure 1: Illustrating map generation using inverse/forward sensory models. Overall environmental size: `44m x 35m`. Corridor width: `1.5m`.

lating from effects (sensory measurements) to causes (obstacles). The forward model describes the characteristics from causes to effects. The inverse model is associated with on-line real-time paradigms such as those mentioned previously and the forward with the offline, non real-time approach.

Traditional approaches using inverse sensor models are prone to generating maps that are inconsistent with the operational data from which they were constructed (Thrun, 2003). This is because such techniques decompose the high-dimensional mapping problem into a number of one-dimensional estimation problems, one for each cell in the map. In doing so they do not consider the dependencies that exist between these cells. The forward sensory model addresses this deficit by considering the dependencies that exist between neighbouring grid cells thereby generating more consistent maps.

Figure 1 presents some illustrative maps. Each paradigm used identical sensory data in generating the maps shown. As can be seen the map generated by the forward model is more compatible with the ideal map. This demonstrates the problem currently inherent in the domain which we are addressing. That is, the dilemma of selecting an on-line paradigm that yield maps of lower accuracy versus an off-line paradigm which produces better quality maps.

## 3 SPECULARITY AND REDUNDANT INFORMATION IN ROBOTIC MappINg

In addition to the type of sensory model used by a mapping paradigm two other issues have a direct correlation on the quality of map produced. These are *Specular Reflection* and *Redundant Information* (Murphy, 2000; Konolige, 1997).

- Specular Reflection: generally occurs when a sonar beam hits a smooth surface and is reflected off the surface at an obtuse angle. This results in either no reading being returned to the sensor or an erroneous reading being returned that has bounced off many surfaces.
- Redundant Information: commonly arises when the robot has been in the same pose for a period of time and hence its sensors report multiple identical readings from that pose.

## 4 THE CONFORM APPROACH TO ROBOTIC MAPPING

ConForM has two distinct aspects. These are:

1. The explicit modelling of sensory data to deal with the specular and/or redundant information.
2. The use of an on-line forward sensory model to translate the sensory readings into occupancy values for inclusion in the grid map.

### 4.1 Conform: Dealing With Specular Readings

ConForM's treatment of the problem of specularity is novel as we consider it from two perspectives. The first is labelled *Acceptability/Agreeability* and the second *Trait Verification*. At each time-step *Acceptability/Agreeability* consider solely the set of readings currently received and evaluates each with respect to its neighbouring readings. *Trait verification* on the other hand takes a wider perspective by evaluating readings in relation to the current perceived state of the environment.

#### 4.1.1 Acceptability and Agreeability

**Acceptability:** Consider a reading $s$ and let us assume that it reports a range reading with a distance of $d$. As operating environments are formed from regular features and as the perceptual fields of neighbouring sensors generally overlap we can assess the consistency of a particular reading by evaluating its probabilistic profile in relation to its neighbours. A reading whose measurement is corrupted by Gaussian noise of zero mean and variance $\sigma^2$ has the following probability distribution where $m$ is the map as illustrated in equation 1. This is based on the standard specification of a sensory model (Elfes, 1989).

$$p(s_t|m) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(d)^2}{\sigma^2}} \tag{1}$$

Strictly speaking $m$ is the local map corresponding to the current perceptual field and therefore a sub set of the overall map that is produced.
Now consider the readings $s^{-1}$ and $s^{+1}$ the neighbouring readings on either side of the reading $s$. The probabilistic profile of these readings are used to support or refute the reading $s$. If reporting an obstacle each will have an associated distance $d^{-1}$ and $d^{+1}$. Therefore we can calculate the probability distribution for these readings using equation 1. These distributions are compared to determine if the readings are consistent. This is accomplished by translating the reading $s$ to the position of $s^{-1}$. Upper and lower bound profiles for $s$ are calculated at this position through scaling the original distance to the point of interest $d$ by the amount of translation required and also taking cognisance of the natural error range of the sensor. If the readings are reporting on the same environmental conditions the reading $s$ will be encompassed by the determined bounds. If this is so the reading is deemed as being acceptable and subsequently allowed to progress for further consideration. An identical procedure is utilised when considering the reading $s^{+1}$. A reading $s$ is discarded only when both acceptability tests indicate that it is unacceptable.
**Agreeability:** The sister concept of acceptability is *Agreeability*. It considers readings that report free space. It is similar to *Acceptability* in that we evaluate a reading in terms of its neighbours. Robotic sensors are good at accurately reporting free space meaning that we can use a direct comparison method with free space readings as it is the detection of an obstacle or not which is important, not the actual difference in any distance reported. Therefore when determining agreement, for efficiency, we do not construct probabilistic profiles for the readings. Rather we use the ranges reported instead. If one of a readings immediate neighbours is not in agreement with the reading itself we allow the reading $s$ to proceed to the next stage of the process where it will be checked in the context of the generated map, using *Trait Verification*. If neither of $s$'s immediate neighbours report a free-space reading then the reading is discarded.

#### 4.1.2 Trait Verification

Agreeability and acceptability deal with specular readings in a bottom up fashion at the local level. Specifically this is in the context of a single reading set. As outlined above there are cases when the reliability of readings cannot be determined from purely considering the local view of the reading set from which they originated. Therefore we also need to consider the top down, global, perspective which takes into account the environmental features determined to date and recorded in the map being constructed. This is the basis of the *Trait Verification*.
In its operation Trait Verification makes use of the fact that environments contain structural regularities and symmetries such as walls that can be approximated using line segments. This is used as a basis for the construction of two environmental views:

- $V$: A temporary sonar view which consists of traits, or line segments, that can be estimated from the current set of sensory readings.
- $L$: A local view which contains a history of the line segments estimated from past sensory readings. Line segments are maintained for an area covering four times the perceptual field of the robot along the path the robot has traversed.

$L$ is used to form a hypothesis as to the *probable* state of the environment from the robots current perspective. This is accomplished by extending $L$ to cover the current location of the robot using the historical perspectives as a reference point.
Following this $L$ and $V$ are reconciled. Firstly, certainty values in the range $0 \rightarrow 1$ are calculated for the readings that give rise to traits in $V$. This is accomplished through use of standard singular displacement specifications presented in (Elfes, 1989).
Having determined certainty values in the readings, $V$ and $L$ are reconciled. Two courses of action are applicable, depending on whether or not sufficient state was available for $L$'s construction.
If enough state was not present to provide four perceptual lengths centred on the oath traversed by the robot, $v_i$'s attributes are considered. $v_i$ is a trait in $V$ and its attributes relate to the reading(s) that gave rise to the trait. For example the certainty associated with the reading(s) or whether the reading(s) were previously flagged as potentially erroneous. If the reading was flagged as potentially erroneous from the *Acceptability/Agreeability* and *Trait Verification* steps or the reading certainty is below a determined threshold and there is not an equivalent trait in $L$, where in this case $L$ has a size equivalent to maximum perceptual range available, the reading is discarded.
If sufficient state was available $L$ and $V$ are compared directly. If traits coincide in both views the readings that gave rise to those traits are accepted, provided that they have not been flagged as possibly erroneous. If they have been flagged the attributes of the trait $v_i$

in $V$ are considered. If two or more sensors agree on the existence of the trait then the flagged reading is accepted. If the trait was detected by a single sensor the certainty value associated with that reading is consulted. If the certainty is below the threshold the reading is rejected. Otherwise it is accepted. If a trait occurs solely in $V$ and not in $L$ then the attributes of the trait are considered. If the flagged status and confidence value of the reading(s) that gave rise to the trait are acceptable, the reading is allowed to proceed for further utilisation. The problem of a reading relating to a trait solely in $L$ and not in $V$ is dealt with in the same manner.

### 4.2 Conform: Addressing Redundant Information

To deal with the problem of redundant information ConForM makes use of pose buckets (Konolige, 1997). With pose buckets a map has a dual representation where each cell represents both the occupancy of the area and the pose of readings that have affected that cell. Therefore a record is maintained stating whether a reading from a given distance and angle has affected a particular cell. This means that the first reading received from a specific pose will be utilised, and all following readings from that pose for this cell are discarded, as they merely duplicate information already in the model.

### 4.3 Conform: Sensor Model

As per the original formulation, ConForM's forward model it also based on optimisation using the EM algorithm (Dempster et al., 1977). It is a mixture model, which accounts for the potential causes of a reading (Thrun, 2003). A measurement may correspond to the detection of an obstacle somewhere in the perceptual field of the sensor, failure to detect any obstacle thereby reporting freespace, or indeed, a random fluctuation of a sensor. Each case has an associated probability. The model convolves these potential causes and associated Gaussian noise into an amalgamated probability distribution which is subsequently optimised by the EM algorithm to determine the most likely cause of the received reading.

Our model differs from the original in that operates on-line and in real-time. The on-line and real-time use of the EM algorithm in ConForM is facilitated through a two step approach. The first step consists of explicitly dealing with potentially erroneous or redundant information through *Acceptability/Agreeability, Trait Verification* and *Pose Buckets*. As such the readings available for the second stage encompass more accurately the true state of the perceived environment meaning that EM can be applied to a search space that is tractable during real-time operation.

**Using the EM algorithm to determine a map**

1. *Initialisation*: Unlike traditional occupancy grid mapping algorithms using inverse sensor models EM does not estimate posteriors. Therefore maps resulting from EM are discrete with each cell being either occupied or empty. As such the cells in the map being constructed are initialised to an occupancy of 0.5.
2. *E-step*: The E-Step calculates the expectations for the potential causes of readings conditioned on the map $m$ and the current set of readings $S$.
3. *M-step*: The M-step assumes all expectations are fixed and calculates the most likely map based on these expectations. The probability distributions calculated in the E-Step encapsulate all potential causes of the readings in $S$ when determining a new map $m$. Maximisation of these distributions are performed by hill climbing in the space of all maps. The search is terminated when the target function is no longer increasing.
4. *Incorporating Uncertainty*: EM calculates only a single map not an entire posterior. An approximation which conditions the posterior on the map generated by EM is utilised to incorporate uncertainty into the map, thereby providing useful information for real-time operation.
5. Finally we integrate the map generated by EM into the overall map using a Bayesian based integration.

## 5 EMPIRICAL EVALUATION

Real world and simulated environments were used to empirically evaluate ConForM. The simulator used was the Saphira architecture with the associated Pioneer simulator. For simulated experiments odometry error was turned off so that wheel slippage would not be a factor thus allowing us to focus on evaluating the performance of the mapping paradigms in large cyclic environments such as those illustrated earlier. For real world experimentation we used relatively small office environments purely for the reason that wheel slippage and thus odometric error is minimal over such short distances.

## 5.1 Benchmarking Technique

To evaluate the maps generated during our experiments we use an extensible suite of benchmarks which allow for the empirical evaluation of map building paradigms (Collins et al., 2004; Collins et al., 2005).

1. *Correlation*: As a generated map is similar to an image it is possible to use a technique from image analysis known as *Baron's cross correlation coefficient* (Baron, 1981) as a basis for evaluating the map.
2. *Map Score*: This is a technique which calculates the difference between a generated map and an ideal map of the environment (Martin and Moravec, 1996).
3. *Map Score of Occupied Cells* This metric is similar to the previous one but only tests those cells in the map that are occupied.
4. *Path Based Analysis*: To fully evaluate a generated map its usefulness to a mobile robot must be considered.
   - The degree to which the paths created in the generated map would cause the robot to collide with an obstacle in the real world, and are therefore invalid. *False Positives*.
   - The degree to which the robot should be able to plan a path from one position to the another using the generated map, but cannot. *False Negatives*.

### 5.1.1 Determining an Overall Score

To allow an overall score to be determined we have developed an amalgamation technique which can be used to rank the overall performance of mapping paradigms relative to each other as outlined in equation 2.

$$C_{\mathtt{map}\in M} = \frac{D_{\mathtt{map}} + P_{\mathtt{map}}}{2} \tag{2}$$

$$D_{\mathtt{map}} = \frac{(1 - \text{MapScore}_{all}) + (1 - \text{MapScore}_{occ}) + B_n}{300}$$

$$P_{\mathtt{map}} = 1 - \frac{(\text{FP}) + (\text{FN})}{200}$$

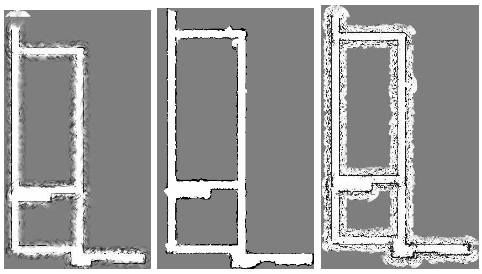
$C_{\mathtt{map}\in M}$ is the overall classification score obtained, $M$ is the set of maps generated in an experiment, $\mathtt{map}$ is a particular map within the set of maps $M$. MapScore$_{all}$ and MapScore$_{all}$ are the normalised result from the *Map Score* metrics, $B_n$ is the normalised *Correlation* result. FP is the normalised *False Positive* result and FN is the normalised *False Negative* result.

## 5.2 Results

In determining the performance of ConForM we empirically evaluated it in relation to its peer mapping paradigms, the original Forward Modelling paradigm of Thrun (Thrun, 2003) and an on-line paradigm from Konolige (Konolige, 1997) which has proven to have the best performance of the inverse model based paradigms (Collins et al., 2005).
Benchmarking consisted of completing a number of data acquisition runs in the environments and using this data in conjunction with the mapping paradigms to generate the grid maps. Our experiment used four differing environments, two simulated and two real world, with three data acquisition runs being completed per environment. Therefore the results presented here are derived from evaluating a total of thirty six individual grid maps. Table 5.2 presents the amalgamated score for the mapping paradigms obtained using the benchmarks outlined above. A larger evaluation recently completed and to be reported on, which consisted of ten differing environments and 3600 individual maps, reported trends consistent with those outlined here.

Table 1: Evaluating the ConForM approach to robotic mapping.

| **Mapping Paradigm** | **Result** |
|---|---|
| Moravec and Elfes 1985 | 0.67 |
| Matthies and Elfes 1988 | 0.65 |
| Konolige 1997 | 0.76 |
| Thrun 2001 | 0.84 |
| ConForM | 0.87 |



(a) Konolige 1997 (b) ConForM (c) Thrun 2003

Figure 2: Illustrative maps from the ConForM evaluation.

## 5.3 Analysis

Overall the results show that ConForM has outperformed the other approaches. ConForM outperforms the inverse model based approaches because of its improved sensor model and the manner in which it tackles the problem of specularity in addition to its use of pose buckets. In dealing with specularity, the multi-faceted approach consisting of *Acceptability and Agreeability* and *Trait Verification* is capable of a finer reading set analysis when compared to inverse model based approaches. This also has the knock-on effect of making the operation of the pose buckets more accurate as they suffer less from the problem of spurious readings giving rise to false hypothesis regarding the perceived state of the environment.

ConForM outperforms the original Forward Modelling approach because of its pro-active approach to the problems of specularity and redundant information. That original approach addressed the problems of seemingly conflicting information through the EM algorithm. The likelihood of the reading was evaluated in a global context meaning that some localised accuracy may be sacrificed. In ConForM the Forward Model used considers the local perspective meaning that it is capable of capturing and retaining more subtle characteristics that may be dismissed in the offline approach.

## 6 SUMMARY

Overall ConForM overcomes the problems inherent in traditional approaches such as the need for assumption of cell independence or the need for offline operation. It also overcomes the issue of the existing forward model approach not being applicable in an on-line context. In addition it generates maps that are more consistent then existing approaches. The areas for further consideration and research in relation to ConForM include refining the threshold used with *trait verification*, investigating the use of EM as a basis for refining already generated portions of the map and investigating alternative EM formulations such as Bayesian based approximations.

## REFERENCES

Baron, R. J. (1981). Mechanisms of human facial recognition. In *International journal of man machine studies*, volume 15, pages 137–178.

Borenstein, J. and Koren, Y. (1991). The vector field histogram - fast obstacle avoidance for mobile robots. *IEEE Transactions on Robotics and Automation*, 7(3):278–288.

Collins, J., O'Sullivan, S., Mansfield, M., Eaton, M., and Haskett, D. (2004). Developing an extensible benchmarking framework for map building paradigms. In *Proceedings Ninth International Symposium on Artificial Life and Robots*, Oita, Japan.

Collins, T., Collins, J., O'Sullivan, S., and Mansfield, M. (2005). Evaluating techniques for resolving redundant information and specularity in occupancy grids. In *AI 2005: Advances in Artificial Intelligence*, pages 235–244.

D. Kortenkamp, R. B. and Murphy, R. (1998). *AI-based Mobile Robots: Case studies of successful robot systems*. MIT Press, Cambridge, MA.

Dempster, A., Laird, A., and Rubin, D. (1977). Maximum likelihood from incomplete data via the em algorithm. In *Journal of the Royal Statistical Society*, volume 39 of *B*, pages 1–38.

Dissanayake, G., Newman, P., Clark, S., Durrant-Whyte, H., and Csorba, M. (2001). A solution to the simmultaneous localisation and map building (slam) problem. In *IEEE Transactions of Robotics and Automation*, volume 17, pages 229–241. IEEE.

Elfes, A. (1989). *Occupancy Grids: A Probabilistic Framework for Robot Perception and Navigation*. PhD thesis, CMU.

Konolige, K. (1997). Improved occupancy grids for map building. *Autonomous Robots*, (4):351–367.

Martin, M. C. and Moravec, H. (1996). Robot evidence grids. Technical Report CMU-RI-TR-96-06, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

Matthies, L. and Elfes, A. (1988). Integration of sonar and stereo range data using a grid-based representation. In *Proceedings of the 1988 IEEE International Conference on Robotics and Automation*.

Moravec, H. and Elfes, A. (1985). High resolution maps from wide angle sonar. In *Proceedings of the 1985 IEEE International Conference on Robotics and Automation*.

Murphy, R. R. (2000). *Introduction to AI Robotics*. MIT Press.

Thrun, S. (2003). Learning occupancy grids with forward models. *Autonomous Robots*, 15:111–127.