Derivation of an Expression for the Mean Information Maximum Value

Jose-Luis Blanco
Dept. of System Engineering and Automation
University of Malaga
Málaga, Spain
jlblanco@ctima.uma.es

Technical Report
February, 13, 2006
http://www.isa.uma.es

***Abstract* – In mobile robots SLAM based on Bayesian approaches, Rao-Blackwellized particle filters (RBPF) enable the efficient estimation of the posterior belief over robot poses and the map. These particle filters have recently received great attention by the research community in mapping and exploration tasks. A central issue in many of those works is that of measuring the uncertainty at a given instant of time. Recently we proposed a new measurement of the *certainty* of a RBPF, named the *Expected Map Mean Information* (EMMI), which was shown to be better suited than other estimators for detecting certainty changes as those produced after closing a large loop. Our method first builds an *expected map* (EM) which condenses all the current map hypotheses and then computes its *mean information* (MI) – an entropy derived measurement that quantifies the inconsistencies in the EM. The MI presents attractive features that sets it apart of the classic direct entropy application. For example, it is largely insensitive to grid maps size and resolution. In fact, while the bare entropy of a map increases without bound when resolution increases, the MI tends towards an asymptotical limit. Unfortunately, this limit can not be obtained for the general case. The purpose of this technical report is to derive an expression for this limit in a very specific case and to provide numerical examples where the convergence of MI could be appreciated.**

I. The Mean Information

We are interested in a measurement of the map information *certainty*, in particular when maps are represented as occupancy grids. For that purpose, in the following we derive an information measurement which stresses the *certainty* in the occupancy likelihood of cells in the grid in opposition to measuring the information *quantity* (which is related to the number of observed cells).

Information theory establishes that the information associated with a random variable is related to its entropy [2]. Consider the entropy of a single cell in the grid,

$$\begin{aligned} H\left(p\left(m_{xy}\right)\right) &= -p\left(m_{xy}\right)\log p\left(m_{xy}\right) \\ &\quad -\left(1-p\left(m_{xy}\right)\right)\log\left(1-p\left(m_{xy}\right)\right) \end{aligned} \tag{1}$$

which is the entropy of a discrete random variable with two possible outcomes, i.e. a Bernoulli distribution. Assuming independence between cells the entropy for the whole map turns into:

$$H\left(p(m)\right) = \sum_{x,y} H\left(p\left(m_{xy}\right)\right) \tag{2}$$

This estimation of the entropy is widely used as a measurement of the information in the map ([1],[3]). However, it exhibits the following limitations:

1. Its absolute value depends on the grid size (the rectangular limits of the map) instead of the actually observed area. Notice that, from Eq.(1), each unobserved cell ($p(m_{xy})=0.5$) will contribute to the global entropy.
2. It depends also on the grid resolution, since it settles (together with the map limits) the total number of cells in the map. This means that the entropy of any map with unobserved areas (all maps in practice) increases without bound when resolution increases.

To overcome these drawbacks we employ the information (I) of a map instead of its entropy:

$$\begin{aligned} I(p(m_{xy})) &= 1 - H(p(m_{xy})) \quad \text{(bits)} \\ I(m) &= \sum_{x,y} I(p(m_{xy})) \qquad \text{(bits)} \end{aligned} \tag{3}$$

where the entropy $H(\cdot)$ is computed by substituting the natural logarithms in Eq.(1) by base-2 ones. As a result we obtain a natural unit for information: *bits*. Noticeably the maximum information value (1 bit) is given to a certainly occupied/ free cell while the minimum value (0 bits) is associated to any unobserved cell[1]. The entropy dependency on the grid size is therefore avoided: the limits of the map become irrelevant since all unobserved cells now contribute with a null information value. Thus, the resulting measurement is a more practical quantifier of the information carried by the map than the direct application of the entropy.

However, in this work we are not interested in the absolute information contained in the map, but in its *certainty*. To effectively reflect the uncertainty in a map we introduce the *mean information* (MI) of a map m, defined as

$$\bar{I}(m) = \frac{I(m)}{N_{obs}} \quad \text{(bits/cell)} \tag{4}$$

where N_{obs} represents the number of observed cells in the map, i.e. cells with an occupancy likelihood different to 0.5. Notice that MI delivers bounded values in the range [0,1]. As an illustrative example of how the MI fits our purposes please consider the occupancy grids in Fig. 1.

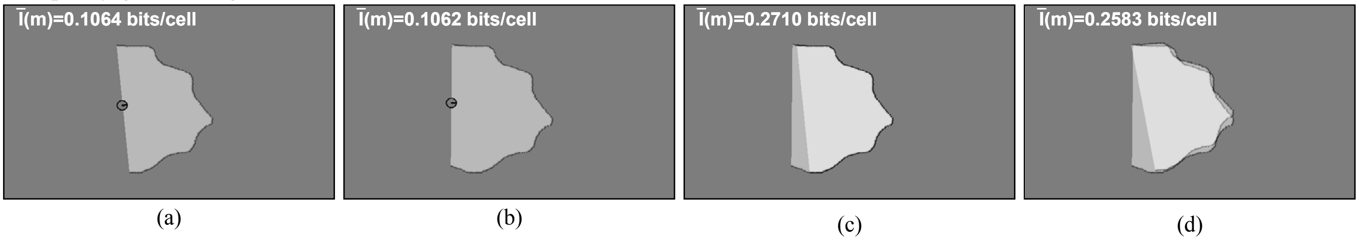


Fig. 1 (a)-(b) Scans and MI values for their associated maps. (c) When both scans are consistently fused in the same grid, the information value increases. (d) Inconsistencies due to poor robot localization decrease the quality of the fused map, which is confirmed by a lower MI value.

The first pair of maps represents two observations of the same environment from slightly different poses. In Fig. 1(c) both are merged into the map with the correct alignment: each observation confirms the occupancy values of the overlapped cells and, as desired, the resultant MI value is greater than in maps where only one observation is inserted. When both observations are misaligned the consequent inconsistencies reduce the MI value, as shown in Fig. 1(d). Notice that in fact, the misaligned map contains more observed cells than the previous case (the total area of the misaligned mixture is bigger than that of the perfectly aligned maps), but the certainty of those cells is lower, in terms of the mean information.

We can highlight the following properties of MI which set it apart from classical entropy-based measurements:

1. An empty map (containing only unobserved cells) has a null mean information value.
2. It is mostly independent on the grid resolution for practical cell sizes. As a graphical demonstration of this property, please refer to Fig. 2(a)-(c), where the same observations are inserted into grids of different resolutions. As expected, the resulting MI of those maps increases as we consider finer-grained maps. But as the resolution increases the MI asymptotically tends towards a maximum value, as appreciated in Fig. 2(d). This presents a great difference with the direct entropy behavior, which in that case tends to infinite. The asymptotical value of MI depends on the inverse sensor model, the specific environment being mapped, and other factors. As an example, in the next section we derive a theoretical expression for this limit for the case of a given synthetic scenario.
3. Higher values are obtained when observations are respectively well aligned.

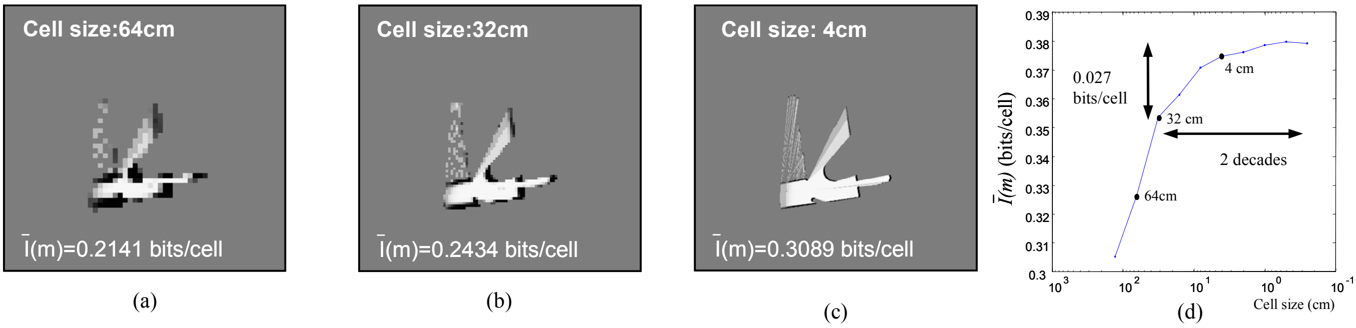


Fig. 2 (a)-(c) The occupancy grids obtained from mapping the same environment with different resolutions. (d) The mean information (MI) is plotted for different grid resolutions, from 1m to 2.5mm. It is remarkable the small increase in information produced by increasing the resolution of a map in the highlighted range, which coincides with commonly used resolutions. Therefore, the MI can be regarded as a resolution independent indicator in practice.

---

[1] Notice that the measurement in *bits* is only a stochastic indicator not related to the actual storage space occupied by the map in memory.

II. DERIVATION OF THE MI MAXIMUM VALUE

As the resolution of an occupancy grid increases, its MI tends towards a maximum value. Unfortunately, an analytical expression for this value can not be derived in the general case since it depends on the inverse sensor model, the number of observations that are merged in the map and the environment. For a particular case, however, we can derive demonstrating the convergence of MI and therefore its high independence on cells size.

Assume a robot at the center of a circular environment of radius R, which is being observed by a 360º field-of-view radial range sensor, and that only one observation (z_1) is carried out. Because of the circular symmetry, we consider the circular grid map $p(m_\rho)$ instead of the classical $p(m_{xy})$, where ρ is the distance from the origin (where the robot is initially) to locations stated by $\langle x,y \rangle$. Thus, the map contents will be assumed to be the same in all directions θ for the whole range $]-\pi,\pi]$. At first we have no prior information about the map, thus $p(m_\rho)=0.5$ for all ρ values. If we consider a range sensor whose measurements are corrupted with additive Gaussian noise with standard deviation of σ, its inverse sensor model would be like the one plotted in Fig. 3, defined as the density:

$$p(m_\rho|z)=\begin{cases} e^{-\left(\frac{\rho-z}{\sigma\sqrt{2}}\right)^2} & \rho \le z+\sigma\sqrt{\ln 4} \\ 0.5 & \text{otherwise} \end{cases} \tag{5}$$

where z represents the sensed range.

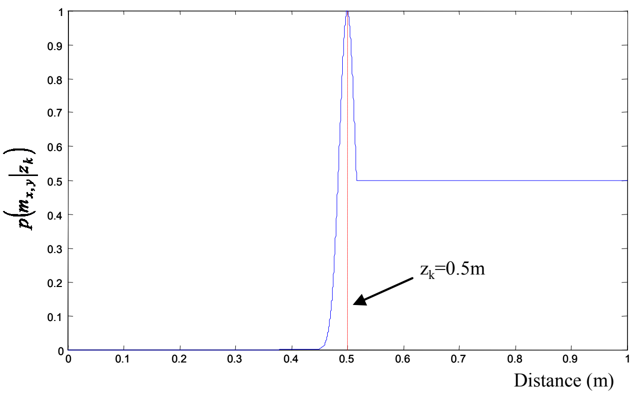


Fig. 3 The probabilistic inverse model of a sensor estimates the occupancy of a given cell, given that some reading has been obtained. In the figure this density is plotted for a laser range finder (and for a measurement of z_k=0.5m). The distribution is highly peaked due to the sensor high accuracy.

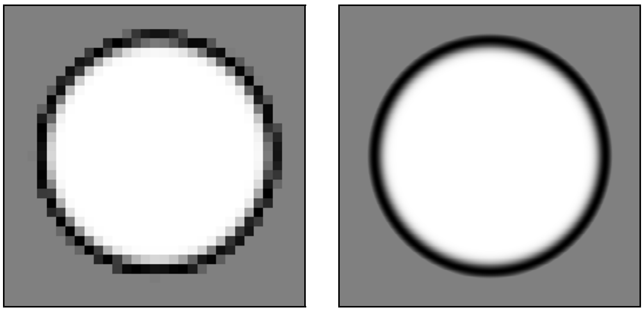
For just one observation, the Bayesian estimation of the map becomes simply $p(m_\rho|z_1)=p(m_\rho|z)$ where z is the only measurement. If we make the cell size tends toward zero, Eq.(4) for the MI becomes an integral over the map (which no more is a discrete grid but a continuous surface):

$$\bar{I}(m)=\frac{\iint_{\theta,\rho} 1-H(m_\rho)d\rho\, d\theta}{\iint_{\theta,\rho} Obs(m_\rho)d\rho\, d\theta} \tag{6}$$

where the binary function $Obs(m_p)$ takes the value of 1 for those positions which have been observed, and 0 otherwise. Because of the circular symmetry, the integration over θ leads to constant values in both integrals and can be neglected. Taking this into account in Eq. (6), and substituting the integration limits we end up with:

$$\bar{I}(m)=\frac{\int_0^{z+\sigma\sqrt{\ln 4}} 1-H\left(\exp\left(-\left(\frac{\rho-z}{\sigma\sqrt{2}}\right)^2\right)\right)d\rho}{z+\sigma\sqrt{\ln 4}} \tag{7}$$

The resultant expression for the continuous case has no analytical primitive, but precise values can yet be obtained by numerical integration. To demonstrate the convergence of the grid-based MI measurement towards Eq.(7) we have generated grids for many resolutions and R and σ values, obtaining maps as those in Fig. 4.



(a) (b)

Fig. 4 Two grids generated in the presented experiment, for resolutions of 16cm (a) and 1cm (b). Here we have set R=2, σ=0.20.

In Fig. 5 some experimental results are shown with their respective parameter values. It is clear from these results that not only the MI converges as the resolution increases, but that the theoretically derived expression gives us that limit value for the specific case of the discussed synthetic environment.

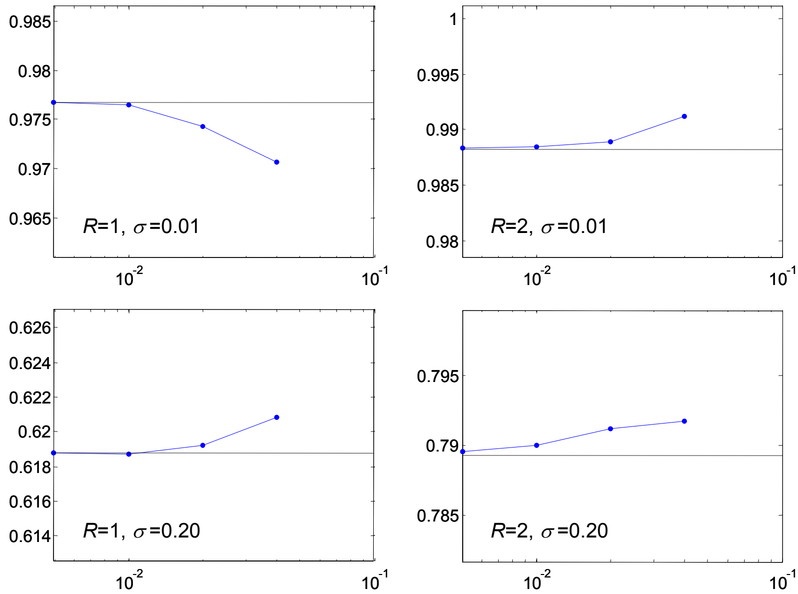


Fig. 5 The MI values for the synthetic scenario discussed in the text are shown as continuous plots, together with the theoretical limit (derived in this paper) as a dashed line. It can be seen how the MI values gets closer to the computed values as the resolution gets more fine-grained.

REFERENCES

[1] F. Bourgault, A.A. Makarenko, S.B. Williams, "Information based adaptive robotic exploration", in *IROS02*, pp. 540-545, 2002.

[2] C.E. Shannon, "A mathematical theory of communication", in *Bell System Technical Journal*, v. 27, pp.379-423 and 623-656, Jul. and October, 1948.

[3] C. Stachniss, G. Grisetti, W. Burgard, "Information Gain-based Exploration using Rao-Blackwellized Particle Filters", in *Proc. of Robotics: Science and Systems (RSS)*, 2005.