

# Revising Stereo Vision Maps in Particle Filter Based SLAM using Localisation Confidence and Sample History

Simon Thompson and Satoshi Kagami

Digital Human Research Center

National Institute of Advanced Industrial Science and Technology

Aomi, Koto-ku, Japan

Email: simon.thompson@aist.go.jp

## Abstract

A method for the backwards revision in time of maps built using stereo vision is presented. Stereo vision data is noisy and inaccurate and can produce inconsistencies between sensor data and existing maps. This results in noisy and inaccurate maps. Gaussian approximations of particle filter based localisation distributions are used to produce localisation confidence measures which can be used to trigger revision events. Remapping occurs when localisation confidence passes a threshold because of re-encountering known objects in the map. Remapping itself consists of extracting the pose history from the most likely particle and incorporating past sensor data from the defined path. The proposed method is implemented and applied to a real world experiment.

**Keywords:** mapping, localisation, stereo vision, revision

## 1 Introduction

A key problem in mobile robotics is that of constructing representations of the environment while estimating ego-motion. Robots moving through an environment must form a representation of that environment through integration of sensor data and to simultaneously be able to localise themselves within the map representation. This is the well known Simultaneous Localisation and Mapping (SLAM) problem. A difficult challenge within this problem is dealing with accumulation of position error especially in environments which contain loops.

### 1.1 Revision of Maps

A popular approach to the SLAM problem is to separate the mapping and localisation steps and perform them in succession. This approach is usually applied to grid based maps as it allows for a large amount of features to be updated. Thrun [1] applied a particle filter state estimation technique to solve the localisation task and subsequently used Expectation Maximisation (EM) to produce the most likely map. By performing EM over the entire (or sub-sampled) past history of sensor data, loops can be accurately described by the mapping process. However this process is computationally expensive and not suitable for real time applications. The approach was later refined to estimate the full posterior of robot position distribution function[2] and adapted to support incremental mapping. Identifying the deviation between the instantaneous and the full posterior maximum likelihood allows for the detection

of and correction (via revision) of accumulated errors in robot position.

This approach is reliant on the quality of sensor data being received, and is also dependent on the structure of the environment. The accuracy and wide range (both in angular field of view and depth) of laser range finders allow for accurate registration between two consecutive samples, thus producing very constrained error accumulation in the current localisation estimate. The local position estimate is therefore very accurate and reliable. The structure of the environment in reported studies is also very regular and the system does not experience large periods where no informative sensor data is available. Detection of a difference between the instantaneous and full posterior position estimates can then be reliably attributed to cumulative errors in the sensor data registration and not on high levels of local noise.

### 1.2 Stereo Vision Sensors

Stereo vision sensors however contain more noise, produce inaccurate estimates and be limited in their angular field of view and their depth perception. Performing particle filter based position estimation using stereo vision produces a greater accumulation of position error compared to using laser range finders as the sensor data cannot constrain the growth as regularly.

While Expectation Maximisation over the full posterior [2] could be applicable in this case, the authors consider it undesirable to fully compute the most likely full

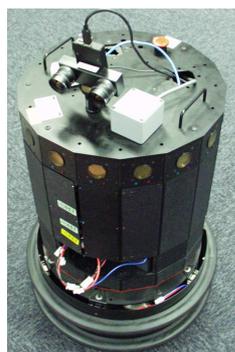


Figure 1: The Nomad 200 mobile robot equipped with a sonar sensor ring and a stereo vision system.

posterior every time the visual localisation produces a noisy result. Using laser range finders such a difference can be reliably attributed to a loop in the environment and as such should happen relatively rarely.

This paper presents a method to detect errors in map construction due to the noise in vision based localisation and revise the map without resorting to Expectation Maximisation. The next section briefly describes the robot platform. Past work in map construction [3], localisation and confidence estimation [4] is then introduced. Then the method of map revision based on localisation confidence and the history of the particle set is described. Throughout, data captured from a real world experiment is used to illustrate the method.

## 2 Robot Platform

A Nomad 200 mobile robot was equipped with a stereo vision system, mounted on top of the robot and facing forward and down. The robot system is shown in Figure 1.

The stereo vision system is calibrated in relation to the ground plane. The system captures  $320 \times 240$  pixel images from the left and right camera, which are used to produce a disparity map. The disparity map is calculated using greyscale images although colour images are available for display purposes. Disparity map generation is based on correlation template matching and is described in [5]. The points in the disparity map are then projected onto the ground plane coordinate system and discretised into a grid cell representation. An example of stereo vision data is given in Figure 2.

The robot executed a short path through a large room and stereo vision and odometry data was logged. The limitations in field of view of the vision system ensured that the robot received no visual localisation cues over a short section of the path. The captured path is used to illustrate and validate the methods described in this paper.

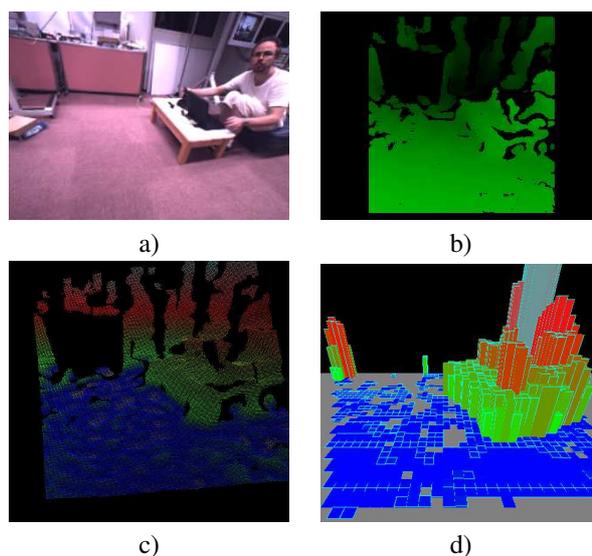


Figure 2: Example of stereo vision sensor: a) camera image, b) disparity map, c) data points projected onto ground plane, and d) associated grid cell occupancy probabilities.

## 3 Map Construction

A mobile robot mapping system was previously developed which fused sonar and stereo vision sensor data to produce maps of complex environments [3]. This system used a 2.5D map representation to estimate the height of obstacles in the environment. The map representation is an occupancy probability annotated with a height measurement and variance measure.

## 4 Particle Filter Localisation

A particle filter uses samples, or particles, to represent location hypotheses and propagates (through resampling) likely samples through time while discarding unlikely hypotheses. The particle set thus forms an approximation of the Probability Density Function of the robot position.

The Condensation algorithm [6] [7] is an example of the particle filter method, and is summarised below. A set of sampled states and their associated probabilities are used to approximate the probability distribution over a temporal sequence of observations and actions. At each iteration of the algorithm, the following steps are applied:

1. Resample: randomly reselect (with replacement) particles in the set based on their measured likelihood to form a new particle distribution.
2. Predict: predict particle poses by applying the current action and the motion model.
3. Measure: evaluate the likelihood of making the current observation from the predicted poses.

Particles condense around likely hypotheses, allowing for the formation of multiple modes in the approximated PDF. The robot location is usually assumed to be in the position of the most likely particle, which is susceptible to fluctuation in the presence of multimodal observation or prior probability densities. This representation also makes it difficult to form a measure of uncertainty regarding the localisation estimate, unlike the Gaussian distribution in Kalman filtering approaches.

#### 4.1 Stereo Vision Sensor Model

A sensor model is used to evacuate the likelihood of observing the current data from a hypothesised location. Typically this involves matching the sensor data with that predicted by the hypothesised pose and the current map. In order to facilitate matching, stereo vision data is reduced to a range profile annotated with height measurements to form a range/height profile. Current stereo vision data is first transformed into a 2.5D local map as shown in Figure 2 and then a range profile is extracted containing the distance to the nearest obstacles over the 40 degree field of view. The height of the detected obstacles is also recorded. The details of the sensor model and the matching process are given in [4].

### 5 Confidence in Localisation Estimate

A measure of the confidence in localisation estimates is desirable. Typically in particle filter approaches, determination of such a measure is difficult because of the arbitrariness of possible particle distributions. In such cases, it is expedient to assume the most probable particle as representative of the global maximum of the approximated PDF. When attempting to measure the confidence of the estimate, it becomes necessary to make further assumptions about the shape of the distribution.

In [4] it was proposed to form a Gaussian approximation of the uncertainty in the distribution while still selecting the most likely particle as the current localisation estimate. Particles are still diffused, measured and resampled as before, but after each iteration a Gaussian approximation of the distribution of the particle set is formed. This has the following benefits:

- The particle set can still form arbitrary distributions: propagation of multi-modal hypotheses is possible.
- The uncertainty in the Gaussian approximation is small when the particles are distributed tightly around a single hypothesis, and large in the case of a loose distribution. In addition tight distributions about multiple hypotheses result in a large estimated uncertainty.

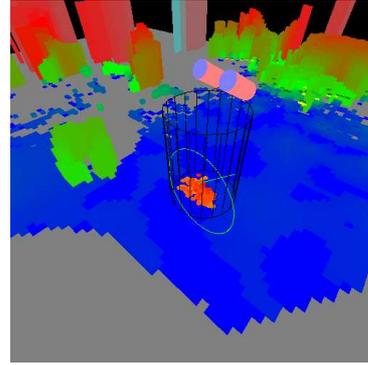


Figure 3: An example of Gaussian approximation of particle distribution

#### 5.1 Approximating Gaussian Uncertainty

The Gaussian uncertainty  $\Sigma_S$  within the particle set can be calculated using a weighted least square of differences:

$$\Sigma_S = \begin{pmatrix} \frac{\sum_i^N R_i (\hat{x}-x_i)^2}{\sum_i^N R_i} & \frac{\sum_i^N R_i (\hat{x}-x_i)(\hat{y}-y_i)}{\sum_i^N R_i} & \frac{\sum_i^N R_i (\hat{x}-x_i)(\hat{\theta}-\theta_i)}{\sum_i^N R_i} \\ \frac{\sum_i^N R_i (\hat{x}-x_i)(\hat{y}-y_i)}{\sum_i^N R_i} & \frac{\sum_i^N R_i (\hat{y}-y_i)^2}{\sum_i^N R_i} & \frac{\sum_i^N R_i (\hat{y}-y_i)(\hat{\theta}-\theta_i)}{\sum_i^N R_i} \\ \frac{\sum_i^N R_i (\hat{x}-x_i)(\hat{\theta}-\theta_i)}{\sum_i^N R_i} & \frac{\sum_i^N R_i (\hat{y}-y_i)(\hat{\theta}-\theta_i)}{\sum_i^N R_i} & \frac{\sum_i^N R_i (\hat{\theta}-\theta_i)^2}{\sum_i^N R_i} \end{pmatrix} \quad (1)$$

In [4] the variance in the  $x$  and  $y$  estimates were used to estimate uncertainty. Here, the well known process of using the eigenvectors of the covariance matrix to calculate the principle axes of the uncertainty ellipse is used. An example of the approximation of a Gaussian uncertainty ellipse is shown in Figure 3. The green ellipse shows the approximated uncertainty distribution while the orange/red points show the underlying particle set.

#### 5.2 Confidence Measure

The Gaussian uncertainty approximation can be used to form a measure of belief in the localisation estimate [4]. We form a measure of localisation belief  $P(\beta)$  for each sensor modality comprised of confidence in the translational,  $P(\beta_T)$ , and the rotational,  $P(\beta_R)$ , components of the position estimate:

$$P(\beta) = P(\beta_T)P(\beta_R) \quad (2)$$

where

$$P(\beta_T) = 1 - \frac{1}{1 + e^{-(\max(\Sigma_S^x, \Sigma_S^y) - C_T) * k_T}} \quad (3)$$

$$P(\beta_R) = 1 - \frac{1}{1 + e^{-(\Sigma_S^\theta - C_R) * k_R}} \quad (4)$$

and  $C_T, C_R, k_T$  and  $k_R$  describe sigmoids reflecting the sensors sensitivity to uncertainty in translation and rotation. The values were experimentally determined (Vision:  $C_T = 30, C_R = \pi/180, k_T = 0.1, k_R = 100$ ).

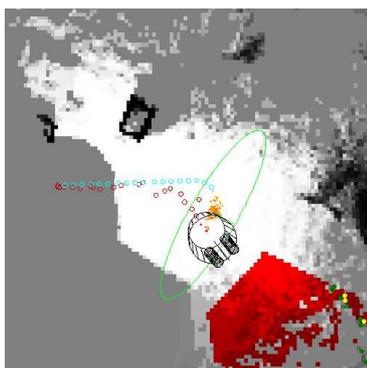


Figure 4: Noise in estimated path due to limited visual localisation cues

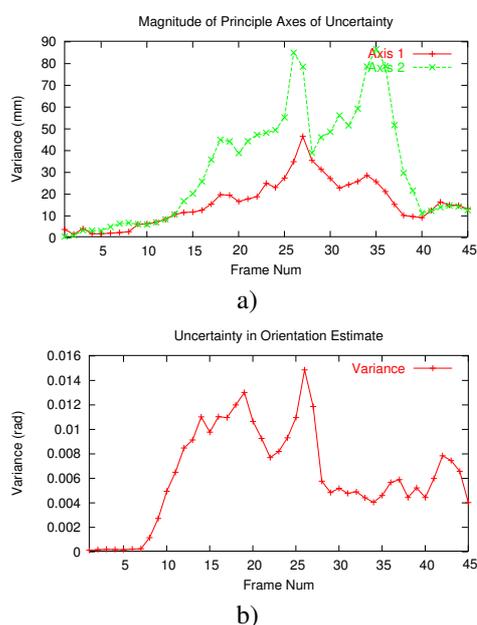


Figure 5: Magnitude of uncertainty over the robot path

### 5.3 Application of Confidence Measure to Mapping Task

A measure of the confidence in the localisation estimate can be used to inform the mapping task. Sensor data integrated into a map when the robot has a strong belief in its localisation estimate should lead to consistent and accurate maps. Conversely sensor data integrated at periods of low localisation confidence could introduce noise into the map. To take into account the notion of localisation belief during mapping the occupation probabilities produced by the map sensor models [3] are weighted with the confidence measures  $P(\beta)$ .

Over the experimental path, the magnitude of the principle axes of the uncertainty ellipse as well as the variance in the orientation estimate was calculated. The results are shown in Figure 5 a) and b) respectively.

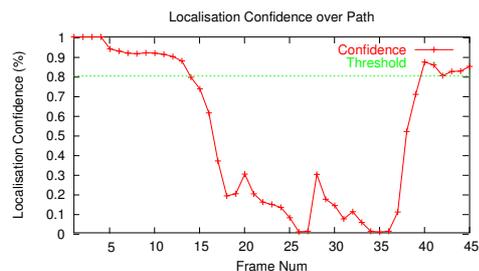


Figure 6: Localisation confidence measure over the robot path

Using the iterative calculation of uncertainty, the localisation confidence measure described above was calculated and is shown in Figure 6. The measure drops off sharply after the robot loses sight of all localisation cues. As this depresses the integration of subsequent objects into the map, the confidence measure stays low until the robot re-encounters previously (to the lose in confidence) seen objects. A threshold measure used to determine when map revision is needed is shown as a dotted line.

## 6 Map Revision

Map revision is needed when the current localisation estimate is no longer consistent with the history of data integrated into the map. An example of this is when closing a loop when navigating in an environment containing cycles. A related problem is that of dealing with noisy or sparse localisation cues, where a high noise to signal ratio can lead to inconsistencies in the map. The presence of noise, or the absence of significant signal for a period of time can be thought of as a special case of closing the loop.

Accumulated position error integrated into the map results in inconsistencies when comparing the estimated map with current sensor data. By revising position estimates over the loop and reintegrating data into the map, these inconsistencies can be dealt with. The task of revising maps backward in time is comprised of two parts: detecting that revision is needed, and performing the revision itself.

### 6.1 Detecting Re-localisation

The measure of localisation confidence described above can be used as a trigger for map revision. If the localisation confidence has previously been below a threshold level, and subsequently rises above this threshold, a revision of the map is in order. In comparison to [2] this method triggers map revision when re-localisation has occurred and not just when an inconsistency has been detected. As mentioned above, stereo vision sensor data can contain large amounts of noise and inconsistencies in localisation estimates

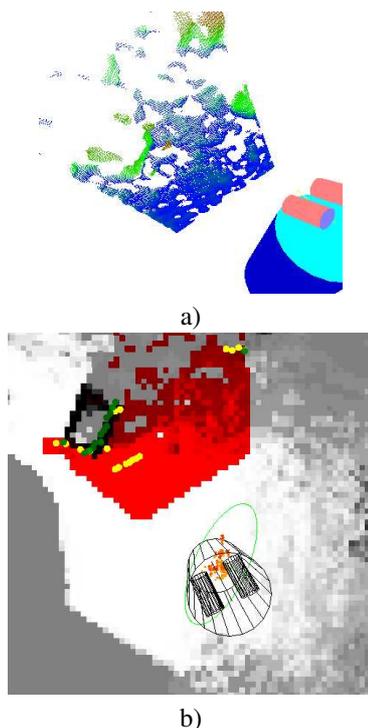


Figure 7: Example of noisy sensor data and effect on uncertainty

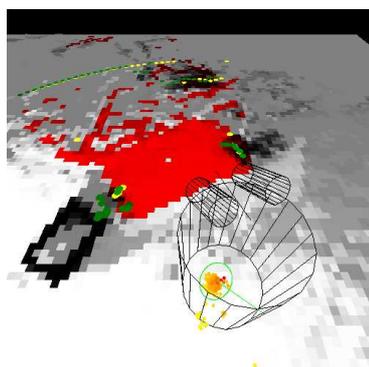


Figure 8: Uncertainty condenses after sensing known objects

can be caused by noisy sensor data as well as by accumulated errors in the map. The current approach delays revision of the map until the position error has been eliminated due to re-localisation from well known objects. This re-localisation is enabled by the reduction in occupancy probability of objects sensed under high position uncertainty.

Figure 7 shows an example of noise in the sensor data contributing to an inconsistency between the expected view and the current sensor view. Parts of the box in the depth image are calculated to be closer to the robot than they are because of characteristics of the depth map generation process. Such noise could trigger map revision while uncertainty as robot position remains high.

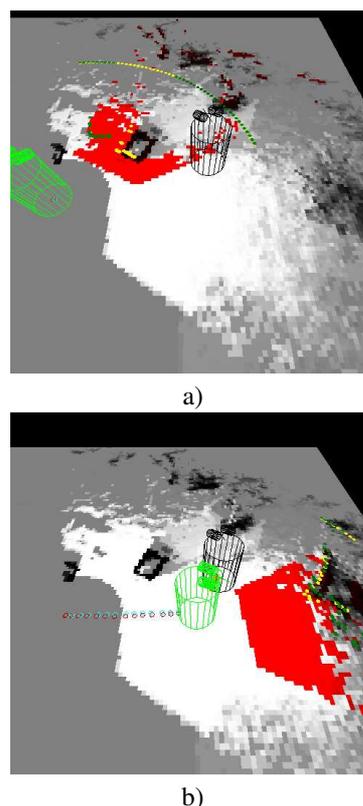


Figure 9: Map revision in progress: a) beginning; b) after 20 frames

In the current system map revision is triggered by a rise in localisation confidence over the threshold hold level (as shown in Figure 6). In effect, this means that map revision is triggered when previously mapped (under high localisation confidence) objects are re-encountered and high confidence restored. Figure 8 shows an example of when the uncertainty has condensed and the rise in localisation confidence has triggered a map revision.

## 6.2 Path Estimation and Remapping

Once a map revision has been triggered the system must estimate the path the robot traveled and then revise the map accordingly. Each particle in the particle set describing robot position has a history attached to it. Given that when navigating under high uncertainty sensor information does not influence the resampling of particles then particle propagation is essentially random, while diffusion occurs according to the the motion model. This process preserves essentially smooth paths among the particle trajectories. Upon re-localisation, the path described by the history of the most likely particle is used to perform map revision. This of course is not the most optimal path, but it does have the nice property of not requiring any extra calculation.

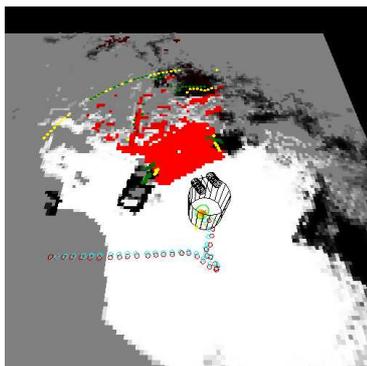


Figure 10: The map and robot position estimate after revision

To revise the map then, the robot replays the sensor data from each pose in the chosen path incrementally updating the map. Figure 9 shows map revision in process: a) revision triggered, black wire-frame shows current robot position, green frame the first position in the chosen path. The sensor data from that pose is shown in red; b) revision after 20 frames. Odometry and estimated position are shown by the light blue and maroon circles respectively.

Figure 10 shows the final state of the system after revision. Note the reduction in noise in estimated position when compared to Figure 4. Also of interest is the increase in occupancy of the object to the right of the map.

## 7 Conclusions and Further Work

This paper has presented a method for performing map revision during SLAM. It is targeted toward the need for frequent revision due to high amounts of noise inherent in stereo vision data. A localisation confidence measure was calculated from a Gaussian approximation of particle spread and used to trigger map revision events. This delays map revision until confident re-localisation occurs.

As it stands this approach would seem to be suitable for revisions of a short length given it tends to inhibit map integration as uncertainty grows. A method specifically aimed at closing long loops in the environment such as Expectation Maximisation [2] is necessary.

However, over short “loops” caused by limited sensor field of view or excessive noise, the system provides a nice simple way of revising maps backward in time.

## 8 Acknowledgments

Research supported by the Japan Society for the Promotion of Science (JSPS)

## References

- [1] S. Thrun, W. Burgard, and D. Fox. A probabilistic approach to concurrent mapping and localisation for mobile robots. *Machine Learning and Autonomous Robots (joint issue)*, 31(5):1–25, 1998.
- [2] S. Thrun, W. Burgard, and D. Fox. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3d mapping. In *Proc. of the 2001 IEEE/RSJ International Conference on Robots and Automation*, 2000.
- [3] S. Thompson and S. Kagami. Stereo vision and sonar sensor fusion for 2.5d map generation. In *The 8th Conference on Intelligent Autonomous System, IAS8*, 2004.
- [4] S. Thompson and S. Kagami. Stereo vision and sonar sensor based view registration for 2.5 dimensional map generation. In *Proc. of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2004.
- [5] S. Kagami, K. Okada, M. Inaba, and H. Inoue. Design and implementation of onbody real-time depthmap generation system. In *IEEE Intl. Conf. on Robotics and Automation*, pages 1441–1446, 2000.
- [6] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.
- [7] F. Dellaert, D. Fox, W. Burgard, and S. Thrun. Monte carlo localization for mobile robots. In *Proceedings of 1999 International Conference on Robotics and Automation*, 1999.