Towards Brain-based Sensor Fusion for Navigating Robots

Adam Jacobson and Michael Milford
School of Electrical Engineering and Computer Science
Queensland University of Technology
michael.milford@qut.edu.au

Abstract

Current state of the art robot mapping and navigation systems produce impressive performance under a narrow range of robot platform, sensor and environmental conditions, in contrast to animals such as rats that produce "good enough" maps that enable them to function under an incredible range of situations. In this paper we present a rat-inspired featureless sensor-fusion system that assesses the usefulness of multiple sensor modalities based on their utility and coherence for place recognition during a navigation task, without knowledge as to the type of sensor. We demonstrate the system on a Pioneer robot in indoor and outdoor environments with abrupt lighting changes. Through dynamic weighting of the sensors, the system is able to perform correct place recognition and mapping where the static sensor weighting approach fails.

1 Introduction

Current state of the art robot mapping and navigation systems generally display a mix of impressive performance and brittle specificity. On one particular robot platform or using offline data gathered from a particular sensing suite, these algorithms [Cummins and Newman, 2009, Davison, et al., 2007, Andreasson, et al., 2008, Paz, et al., 2008, Konolige and Agrawal, 2008] are able to generate maps that are accurate to the nearest few millimetres of large office building-sized areas or routes as long as 1000 km [Cummins and Newman, 2009] . The size and accuracy of these maps are arguably beyond the capabilities of what a human could generate, and are most likely superior in terms of metricity to the maps generated by animals.

Yet change the environment, robot platform (Fig. 1) or sensor suite slightly and the majority of these systems degrade in performance or fail completely. Vision-based systems that incorporate feature detectors [Lowe, 1999, Bay, et al., 2006] generally require the feature detector to repeatedly detect the same features in any one location of an environment, otherwise the mapping system will likely fail. Feature detection is usually environment specific, and a trained codebook that is applicable in urban areas may not apply to forested or desert environments. Some research has investigated place recognition over time when environments change [Glover, et al., 2010, Biber and Duckett, 2005, Kiyoshi, et al., 2011] or using sequences of image to improve performance [Milford and Wyeth, 2012, Maddern, et al., 2012, Milford, 2012] , but ultimately there are always going to be scenarios in which recognition by any one sensor modality fails. Range-based sensors can fail if dust or smoke fills the air, and some will even fail if sunlight directly impinges on the sensor. One obvious solution is to combine multiple sensors together to improve robustness [Castellanos, et al., 2002, Gutmann and Konolige, 1999] , but most current sensor fusion approaches have constraints such as the need for sensors to be able to detect common features such as edges or corners.

Figure 1: Robot platforms such as these Pioneer 3DX platforms come with a variety of sensing suites. Even identical sensor types can vary in their exact properties, requiring calibration before being used with many types of sensor processing algorithms.

Biological systems provide a stark contrast to robotic mapping systems – only days after birth, rats rapidly calibrate a wide range of sensors to learn and navigate diverse environments. The calibration process used by rats is not driven by genetic selection – and works in environments ranging from a pitch-black sewer in Cairo to a featureless arid desert in America.

This work was supported by an Australian Research Council Discovery Project DP1212775 to MM.

In this paper, we combine featureless sensor fusion algorithms inspired by how the rodent brain is thought to process multisensory data with the RatSLAM mapping system [Milford and Wyeth, 2010, Milford, 2008, Milford and Wyeth, 2008] , in order to provide a multi-sensor-based mapping capability in changing environments that is sensor naive, that is, independent of the number or type of sensors.

Our approach uses two techniques to weight sensor trustworthiness: spatial coherence and template expectation. Spatial coherence is assessed by examining the proximity of place recognition matches reported by each sensor using the spatial map provided by the RatSLAM system. Template expectation compares a template match reported by a sensor to the template that sensor is *expected* to match, given template matches by the other sensors. These two measures are integrated to continually provide sensor weightings representing the trustworthiness of each sensor, enabling the weighted fusion of sensor readings into super templates. Due to the complexity of the neural networks in the rodent brain, the sensor fusion algorithms mimic key high level characteristics of the rodent brain systems but are implemented as algorithms without any neural network machinery.

The paper proceeds as follows. In Section 2 we review the literature on robot and rat approaches to mapping changing environments. Section 3 presents our approach, briefly reviewing the OpenRatSLAM system, then describing in detail the sensor pre-processing, weighting and super template algorithms. In Section 4 we present the experimental setup, with results presented in Section 5. Section 6 discusses the outcome of the research and areas of future work.

2 Background

One of the dominant uses of sensors in robotics has been to enable mobile robots to map and navigate their environments. When a robot creates a map of its environment while simultaneously calculating its location within the map, the process is known as Simultaneous Localization And Mapping (SLAM) [Dissanayake, et al., 2001] . While external sensors or information sources such as visual or electric beacons, or a Global Positioning System (GPS), can be used to achieve robot navigation, most SLAM research has focused on the use of onboard robot sensors such as laser range finders, sonar, and cameras. The second approach is highly desirable in many applications because it does not require the environment to be modified, meaning a robot can in theory be introduced to an unknown environment and start mapping immediately. The reality, however, is very different. Range-based approaches that use laser range finder or sonar sensors work well in structured environments, but fail in unstructured or dynamic environments where geometry is not a reliable cue. Likewise, most vision-based SLAM approaches work well when the particular visual features the algorithm uses are reliably detected, but fail in a sufficiently different environment where those features do not exist, or when environmental conditions change, such as lights being turned out or day turning to night [Nuske, et al., 2009] .

In attempting to overcome the fragility of any one sensor modality across different environments and situations, multisensory approaches have been developed that perform sensor fusion at the level of landmarks [Castellanos, et al., 2002] or dense depth maps [Gutmann and Konolige, 1999] , typically using range and vision-based sensors. These approaches improve robustness but are still subject to the same limitations; selection of suitable features that are detectable by all sensor types, or environments with reliable, informative geometry. Furthermore they require sensor-sensor calibration to ensure features are represented in a common reference frame. The practical consequences of these limitations is that, with the exception of “dumb” vacuum cleaner robots, current autonomous robot systems are limited not only to one type of environment, but also to operation only when certain environmental assumptions are met.

Rodents, in contrast to robotic systems, are amongst the most flexible of animals in adapting to life in just about any environment. Rats are one of the most widely distributed animals in the world, but much of this distribution has come only in the last few centuries [Cornish, 1908] . Consequently, from an evolutionary perspective rats have not had any time to specialize in their new environments. “Wild” rats found in Alaska are genetically similar to “urban” rats that spend their time living in the walls of a house or sewer, as are burrow rats that never come above ground and rats living in an open air Asian market or Australian garbage dump. These rats are all equipped with the same set of sensors and the same neural circuits. Rats also live a very short life of less than a year in the wild and must therefore rapidly learn how to use their sensors so they can explore, forage for food and find safe pathways where they are protected from predators. To overcome these challenges, rats are able to, from only four days after birth, *learn* how to best sense, map and navigate in their environment [Golani, et al., 1981] . Furthermore, adult rats rapidly adapt to changes in the environment or to changes in their own sensing equipment during their adult life. Experiments using adult rats have revealed how easily and readily rats will adapt to changes such as the lights being turned out, or the removal of visual, odour or auditory cues. For example, adult rats are able to continue to localize in a square room even after all the lights have been turned out [Save, et al., 2000, Hafting, et al., 2005] . In such a situation they may rely on odour cues to localize. However, if even these odour cues are removed, they will continue to localize correctly [Save, et al., 2000] , by learning to rely only on their tactile whiskers and knowledge of the approximate room geometry.

Underpinning the capability of rats to flourish in such varied environments are a range of sophisticated neural functions that are performed in a brain region called the *entorhinal-hippocampal* formation, which consists of the *entorhinal cortex* (EC) and *hippocampus* [Nolan, et al., 2010] . These brain regions are thought to be involved in a number of critical functions; sensor fusion, sensor calibration, creation of spatio-sensory snapshots by fusing sensor and localization data, variable pattern discrimination to suit

perceptually distinctive or bland environments, and detection of inconsistent sensory input. Rats use these neural processing mechanisms immediately after birth, but also later in life if the environment changes (e.g. if the lights go out) or a sensor is damaged (e.g. if they lose their whiskers). In this paper, we focus on the detection of consistent and inconsistent sensory snapshots in order to weight each sensor modality based on current utility and enable place recognition and mapping across significant environmental change.

3 Approach

In this section we briefly describe the SLAM backend for this paper, and then provide details of the sensor pre-processing, sensor coherence and weighting schemes and the super template learning process.

3.1 OpenRatSLAM

OpenRatSLAM [Ball, et al., 2012] is a modular, open source version of the RatSLAM system. The system consists of three major components – the local view cells, the pose cells and the experience map (Figure 2). The local view cells encode distinct snapshots of the environment. In all previous implementations these snapshots were visual in nature, but in this paper the snapshots consist of super templates that combine information from the four sensing modalities. Significant modifications to the OpenRatSLAM codebase were made to enable multisensory integration, and these updates will be deployed in future OpenRatSLAM releases.

The pose cell network provides filtering of ambiguous perceptual data from the local view cells, while the experience map provides a graphical map. In this paper we use the experience map as a qualitative test of whether the mapping system is working correctly, by comparing the map topology to the physical layout of the environment. More detailed descriptions of the RatSLAM system can be found in [8, 9]. The experience map also forms a critical component of the sensor weighting schemes, by enabling the evaluation of whether sensors are reporting coherent place matches.

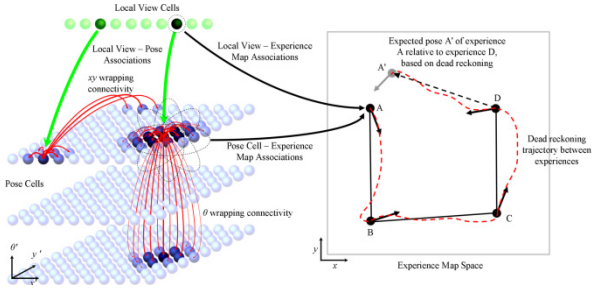


Figure 2: The major modules of the RatSLAM system reproduced from [Milford and Wyeth, 2010] . The local view cells represent learnt unique scenes in the environment. The pose cells represent the belief about the current pose. The experience map is a topological representation encoding the pose cells and local view cells in nodes and links.

3.2 Sensor Weighting Overview

The sensor weighting scheme is broadly based on a process of learning and recalling sensory snapshots or templates for each sensor individually, and assessing the coherence of reported template matches between sensors using the map provided by RatSLAM. A template in this context is simply an array comprising the sensor readings from a particular point in time. An optional additional template expectation check enables the weighting scheme to discriminate between sensors that are reporting correct loop closures but only to locations learnt since an environmental change, and sensors that are performing loop closures to locations learnt before the change.

3.3 Sensor Pre-Processing

The control and integration of multiple sensors is achieved under the ROS framework. As each sensor has the potential to report vastly different information, the data from each sensor is normalised to a range of (0, 1) by dividing by the maximum possible sensor reading. In addition the camera and depth images are resized to a resolution of 12×9 pixels, converted to greyscale and rearranged into a single vector (108 elements) rather than the two dimensional matrix usually associated with such sensors.

3.4 Template Comparison

We use a Sum of Absolute Differences (SAD) comparison to evaluate how similar both individual sensor templates and super templates are to those previously learnt. The difference score d between the current proposed template and a previous template j is given by:

$$d(j) = \frac{1}{s}\sum_{x=0}^{s-1}\left|p_x - p_x^j\right| \tag{1}$$

where p_x is the value of element x in the current template, p_x^j is the value of element x in a previous template j, n is the number of templates learned thus far and s is the length of the template. The best matching template b is the one with the lowest difference score, given by:

$$b = \arg\min_{0 \le j < n} d(j) \tag{2}$$

A threshold s_{thresh} determines whether the new template is considered novel ($m=0$) or familiar ($m=1$):

$$m = \begin{cases} 1, & d(b) \le s_{thresh} \\ 0, & d(b) > s_{thresh} \end{cases} \tag{3}$$

To keep track of whether a sensor is generally recognizing familiar sensory snapshots or not, we filter a recognition rate parameter r:

$$r = k_1 \times r' + (1 - k_1)q \tag{4}$$

where r' is the previous value of the recognition rate parameter and q is given by:

$$q = \begin{cases} 1 & |n - b| \ge \partial \\ 1 & m = 1 \\ 0 & \text{else} \end{cases} \tag{5}$$

where σ is a recency threshold given in Table I. We also define a parameter χ which tracks whether the majority of sensors are currently seeing familiar ($\chi=1$) or unfamiliar ($\chi=0$) visual input, calculated using the binary majority function [Knuth, 2008] :

$$\chi = \begin{cases} 1, & \frac{1}{2} + \sum_{i=1}^{n} \frac{1}{n}\left(e(i) - 0.5\right) \geq 0.5 \\ 0, & \text{else} \end{cases} \quad (6)$$

where $e()$ is given by:

$$e(k) = \begin{cases} 1, & r(k) \leq 0.8 \\ 0, & \text{else} \end{cases} \quad (7)$$

3.5 Individual Sensor Template Creation

Templates for each individual sensor are learned at a fixed rate regardless of whether the current sensory snapshot of that sensor is matched to a previous template or not. This forced learning simplifies the inter-sensor coherence calculations but has computational implications – these implications are discussed in the Results section and a proposed improvement is described in the Discussion.

3.6 Sensor Coherence

The first step in calculating the weighting for sensor i is to calculate its coherence with each of the other sensors j. The coherence measure c comprises two different measures of coherence –spatial coherence and template expectation similarity:

$$c(i,j) = d(i,j) m_i m_j v(i,j) \quad (8)$$

where d is a spatial distance function relating the place recognition matches reported by sensor i and j, m is the binary match outcome for each sensor (Equation 3), and v is the expected template similarity function. The sensor coherence measure is also filtered:

$$c_{new} = k_2 \times c' + (1 - k_2) c \quad (9)$$

The distance function is given by:

$$d(u,v) = \begin{cases} 1, & \sqrt{(x_u - x_v)^2 + (y_u - y_v)^2} \leq d_{thresh} \\ 0, & \text{else} \end{cases} \quad (10)$$

where x and y are the co-ordinates of the experience map node associated with templates u and v.

When two sensors report a template match, the expectation function uses the template match j reported for the second sensor to compare the *expected* template match j for the first sensor with the *actual* template match i reported for the first sensor:

$$v(i,j) = \begin{cases} 1, & \frac{1}{s} \sum_{x=0}^{s-1} \left| p_x^i - p_x^j \right| < s_{recall} \\ 0, & \text{else} \end{cases} \quad (11)$$

where s_{recall} is a threshold based on the maximum template difference score for which the expected and actual template matches are considered the same.

The expected template check is optional depending on the application and determines whether a sensor is allowed to regain trust in an environment after a sudden change by recognizing templates learnt after the change. An example would be a camera after the lights are turned out – if the expectation measure is omitted the camera will regain weighting once it starts learning and recognizing places within the changed environment, even if it is unable to recognize those camera templates learnt before the lights were turned out. In this paper we always use the expectation measure.

3.7 Individual Sensor Weighting

The final step in the sensor weighting process is to evaluate the inter-sensor coherence measures and calculate a scalar weighting value that represents the current level of trust in the individual sensor. The weighting w for sensor i is given by:

$$w(A) = \begin{cases} 1 & \beta = 1 \\ 1 & \beta = 0 \cap \chi = 1 \\ 0 & else \end{cases} \quad (12)$$

where:

$$\beta = \begin{cases} 1, & \frac{1}{2} + \sum_{i=1}^{n} \frac{1}{n}\left(c(i) - 0.5\right) \geq 0.5 \\ 0, & \text{else} \end{cases} \quad (13)$$

The weighting w for each sensor is also filtered:

$$w_{new} = k_3 \times w' + (1 - k_3) w \quad (14)$$

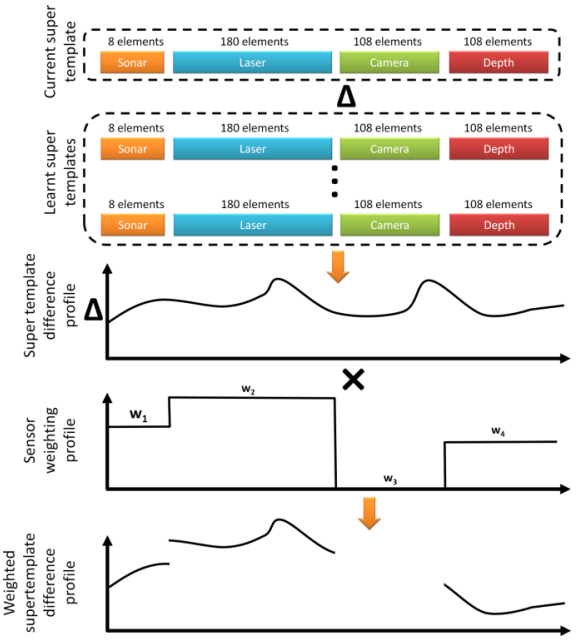


Figure 3: The super template comparison process. The absolute differences between the current super template and all learnt super templates are weighted by the sensor weighting profile **w**.

3.8 Super Templates

Super templates are formed by concatenating sensor data from each of the sensors into one vector (Figure 3). The current super template candidate is compared to all learnt super templates using a weighted version of the sum of absolute differences comparison:

$$D(j) = \frac{1}{s}\sum_{x=0}^{s-1} w_x \left| P_x - P_x^j \right| \quad (15)$$

where w_x is element number x in the overall weighting vector, formed by concatenating the weight vectors of each of the individual sensors:

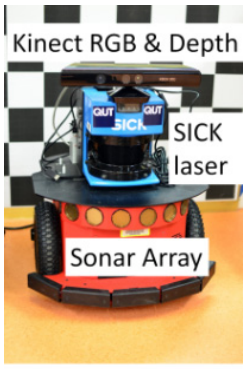
$$\mathbf{w} = \left[\mathbf{w_1} \quad \mathbf{w_2} \quad \ldots \quad \mathbf{w_n}\right] \quad (16)$$

Currently individual sensor weighting vectors are not normalized by the sensor vector size, meaning sensors that produce more data are more influential.

If the minimum matching difference *min(**D**)* exceeds a threshold *S* a new super template is learnt, otherwise the lowest scoring previous super template is deemed a match. A new super template is only learned if none of the current super templates are matched. The super templates are then used as the local view cell module in the RatSLAM model.

4 Experimental Setup

In this section we describe the robot platform (Fig. 4), testing environment and system parameters.



(a)

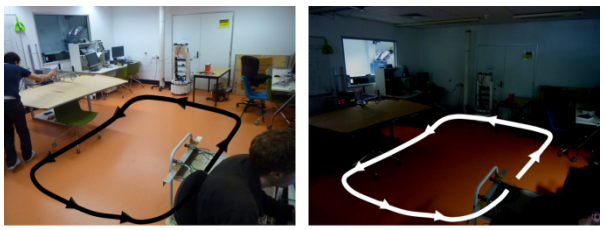
Figure 4: Pioneer 3DX robot with SICK laser, Microsoft Kinect and forward facing sonar ring.

4.1 Robot Setup

The testing platform was an *Adept Mobilerobots* Pioneer 3DX robot. The robot is equipped with a SICK laser, Microsoft Kinect utilising both RGB and Depth images and eight ultrasonic range sensors arranged in a ring on the front of the robot.

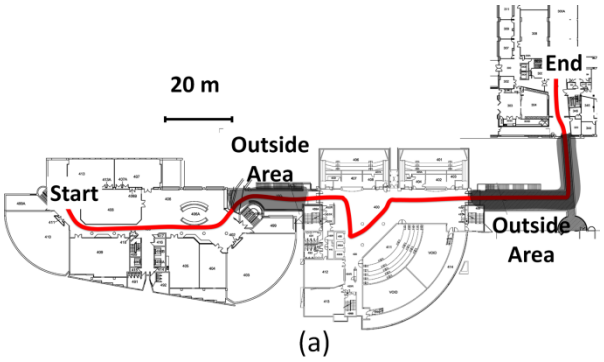
4.2 Testing Environment

Five laps of a laboratory environment were performed, with the lights on for the first two laps (Figure 5a) and off for the last three laps (Figure 5b). The ROS bag-file is available for readers to download and process at the following link: https://wiki.qut.edu.au/display/cyphy/Michael+Milford+Datasets+and+Downloads. We also present preliminary results for experiments in a larger campus environment. The robot was tele-operated through and in-between buildings along a mixed indoor-outdoor route approximately 200 metres long several times during day and night-time conditions (Figure 6).



(a) (b)

Figure 5: Laboratory testing environment with the lights (a) on and (b) off. Repeated laps were performed over the trajectory shown.



(a)

(b) (c)

Figure 6: The larger campus environment for which we present preliminary results. The route was traversed during both day- and night-time conditions.

4.3 System Parameters

The values of constants used for the experiment in this paper are given in Table 1. The recognition constants were chosen based on standard RatSLAM tuning processes outlined in [Ball, et al., 2012] . The filter values affect the responsiveness of the system and were chosen to generate reasonable smoothness in the weighting functions. In the discussion we talk about ways to automate parameter selection.

TABLE I
PARAMETER LIST

Parameter	Value	Description
k_1	0.95	Recognition rate filter
k_2	0.95	Sensor coherence filter
k_3	0.95	Sensor weighting filter
S	0.07	Super template matching threshold
σ	50 frames	Recency threshold
$s_{thresh\ kinect}$	0.07	Kinect depth matching threshold
$s_{thresh\ laser}$	0.07	Laser matching threshold
$s_{thresh\ sonar}$	0.15	Sonar matching threshold
$s_{thresh\ camera}$	0.15	Camera matching threshold
d_{thresh}	0.5 m	Location match threshold
s_{recall}	0.15	Recall check threshold

5 Results

In this section we present results showing sample templates and super templates, individual sensor template recognition performance, weighting profiles over time for each sensor, super template learning and recognition performance with and without the dynamic sensor weighting scheme, and experience maps produced with and without the weighting scheme. All computation was performed at real-time speed on an Intel dual core laptop which reached 60% CPU usage at the end of the experiment. Individual and super template comparison was the main consumer of computational resources and scaled linearly with the length of the dataset.

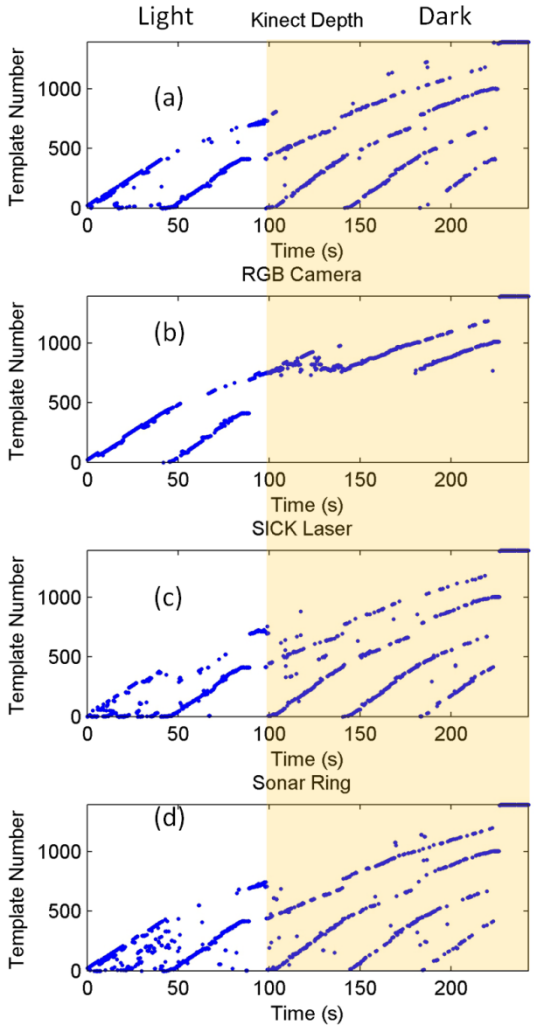


Figure 7: Individual sensor template recognition for the (a) kinect (b) camera (c) laser and (d) sonar sensors.

5.1 Individual Sensor Template Recognition

Figure 7 shows the template learning and recall for each of the individual sensors. After the first two laps the lights were turned out, indicated by the shaded area in the figure. On the third lap, the camera clearly fails to recall any of the templates learnt during the first two laps in the light, but then on the second and later laps in the dark starts to recognize previous laps in the dark. In contrast, the three range-based sensors all have varying degrees of success in recognizing templates from the first two loops, even after the lights are turned out, with the laser and kinect depth sensor performing better than the sonar. An example set of individual sensor templates and the resultant super template is shown in Figure 8. Note the normalization to a maximum sensor reading value of 1 – for the camera sensor this represents a white pixel, while for the depth sensors it represents an out of range reading.

5.2 Sensor Weighting

Overall, the laser is most consistently weighted at 1 apart from a brief drop at around 20 seconds (Figure 9). The sonar and kinect are generally heavily weighted, although both experience a significant period of down weighting around 15 to 45 seconds. Looking at the individual template graphs in Figure 7, all three range sensors report a varying number of false positive matches during this period in time.

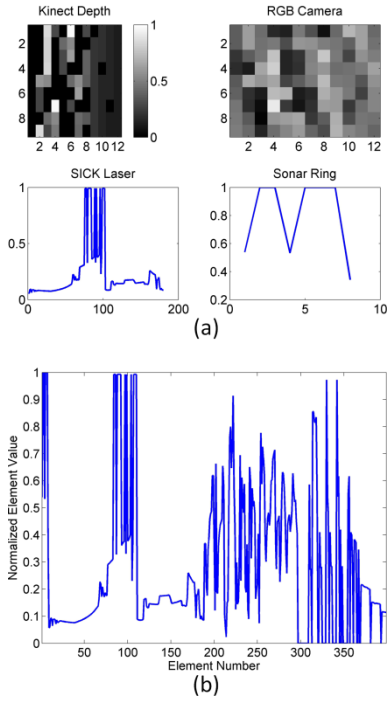


Figure 8: (a) Sample individual sensor templates and (b) the resultant super template formed by their concatenation.

The camera is variably weighted during the lights on segment of the dataset but is rapidly weighted to zero for almost the entirety of the lights out section of the dataset. During this same period all three range sensors are weighted at one. Near the end of the run, the camera starts sporadically being weighted above zero again – this is most likely due to periods of poor recall where the range sensors are not reporting matches either. When none of the sensors are reporting a match, Equation 12 weights all sensors towards one.

Ultimately, the suitability of the individual sensor weightings is a somewhat subjective affair. The next sections present the effects of the weighting scheme on super template recognition and experience mapping.

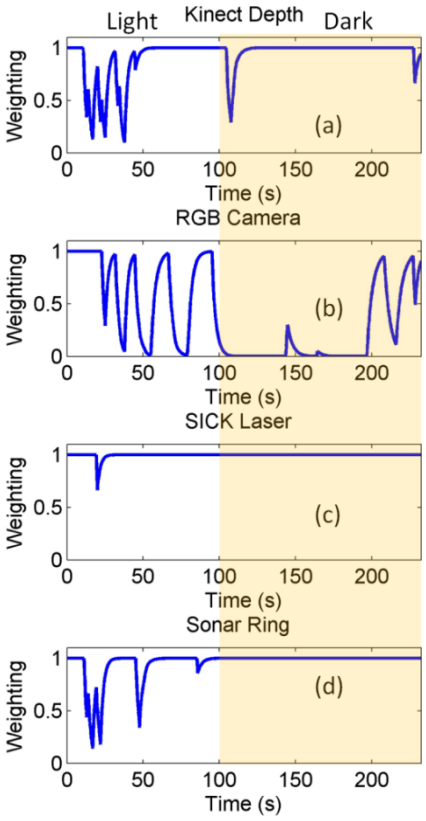


Figure 9: Dynamic sensor weighting for the (a) kinect (b) camera (c) laser and (d) sonar sensors.

5.3 Super Template Recognition

Figure 10 shows learning and recognition of super templates without the weighting scheme implemented. The second lap with the lights on is clearly recognized, but during the next 3 loops in the darkness, the recall rate drops to 0 percent. The second and third loop in the darkness is correctly matched back to the first loop in the darkness, but there are no loop closures back to the templates learnt during the lights on section.

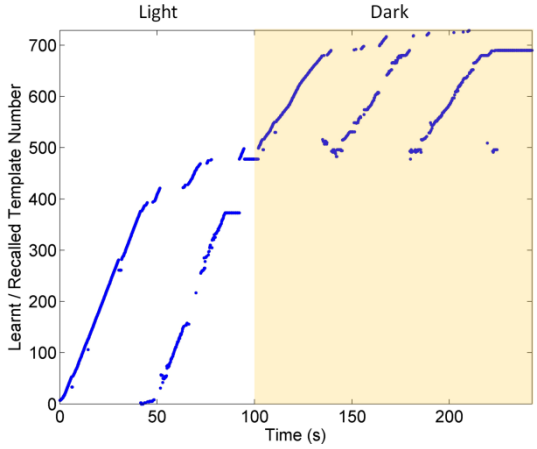


Figure 10:Unweighted super template recognition.

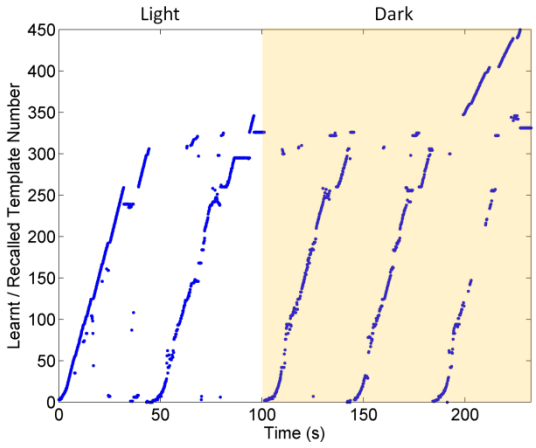


Figure 11: Weighted super template recognition.

Figure 11 shows the super template graph with the weighting scheme enabled. Immediately noticeable is the high rate of recall over all five loops of the environment, and that relatively few new super templates are learnt after the first loop of the environment. The super template matching algorithm reports many matches from the three laps in the dark back to the first loop with the lights on. The recall rate drops somewhat during the third loop in the darkness, which we attribute to variation in the robot path and the small degree of pose invariance in the template recognition system – the lower recall rates for each of the individual sensors during the last lap is also evident in Figure 7.

5.4 Experience Maps

Figure 12 shows the experience map for the laboratory environment without sensor weighting. The focus with the presentation is on the topology and connectivity of the map, rather than its metric accuracy.

There is little successful loop closure on repeated traverses of the environment. The loop closures that do occur are loop closures within the part of the map that was formed under the same environmental conditions, that is, loop closures within the lights on component of the map and within the lights off component but not between parts of the map formed during different conditions.

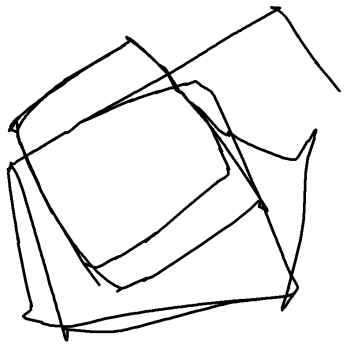


Figure 12: The unconnected experience map without dynamic sensor weighting.

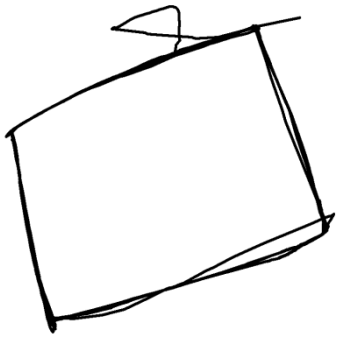


Figure 13: The coherent experience map with dynamic sensor weighting.

Figure 13 shows the experience map when the dynamic sensor weighting scheme shown in Figure 9 is implemented. The five loops of the environment overlap correctly apart from a short period of deviation from the trajectory, as discussed in Section 5.3.

5.5 Campus Environment

Figure 14 shows the super template matching performance with sensor weighting enabled during a night-time traverse of the campus environment after the robot had already performed a day-time traverse. The recall rate increased by 43% from 23% to 33%, and resulted in a correctly mapped second traverse of the environment (Figure 16), which was not possible without the weighting scheme (Figure 15).

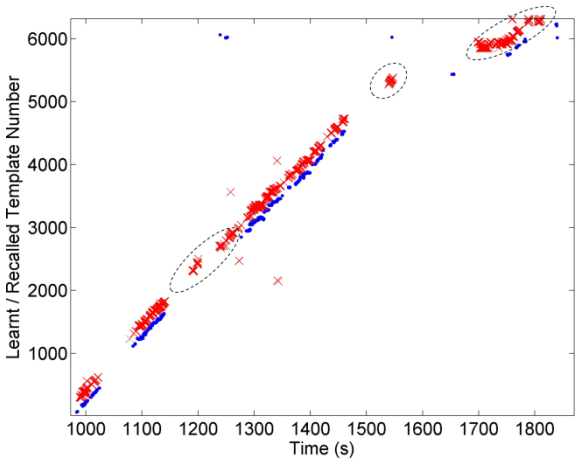


Figure 14:Unweighted (blue dots) and weighted (red crosses) super template recognition, with the weighted template matches artificially offset by 200 templates vertically for legibility reasons. Circled regions indicate regions where the weighting scheme enabled correct template matches where none were made before.

Figure 15: Experience map of the campus environment without sensor weighting. The map has been orientated to approximately correspond to the floor plan in Figure 6.

Figure 16: Experience map with sensor weighting, with the two traverses of the environment correctly linked, overlapping and representative of the topological layout of the environment shown in Figure 6.

6 Discussion and Future Work

The results presented in this paper suggest there is promise in the approach of a featureless approach to multisensory fusion for robot navigation which is naive to sensor type and number. By necessity, the fusion algorithms were mathematical rather than biological in nature, although they share some of the high level characteristics of the neural processes thought to occur in the hippocampus and entorhinal cortex brain regions. The strengths of the approach are that it is agnostic both to the type of sensors available and their number. Here we discuss areas of future work which will address some of the system's current limitations.

6.1 Neural Network Implementation

One line of future work will implement a full neural network model of the algorithms developed in this paper that can integrate with the existing RatSLAM components. The conversion to a neural rather than mathematical model will provide some beneficial computational characteristics such as enabling "The Race to Learn", a process by which new sensory input is gradually learnt in a robust manner by a neural loop called the trisynaptic loop [Nolan, et al., 2010] . There are a number of theorized neural processes that could facilitate autonomous tuning of template recognition thresholds and filter constants.

6.2 Computational Optimization

The individual template comparison process learns new templates at a fixed rate regardless of whether the data is novel or familiar, and hence scales linearly with the length of the dataset rather than the size of the environment. The fixed learning rate was chosen to simplify the coherence calculations by synchronizing learning of individual sensor templates. We will apply the super template learning system, which only learns a new template when the current sensory snapshot is unfamiliar, to the individual sensor template learning system and introduce an indexing system to cross-reference individual sensor templates, thereby ensuring compute scales only with the size of the environment.

6.3 Continual Template Recognition Threshold Tuning

To achieve truly sensor naive operation, a method for autonomous calibration of individual template recognition thresholds for each sensor modality is needed. In addition to the neural processes mentioned previously, we speculate that employing a similar process to that used for weighting, that is looking at sensor consensus regarding short-term recall rates, will enable thresholds to be adjusted until all sensors are producing comparable recall rates. This of course may

drive some sensors (for instance a camera in a dark room) into producing many false positives, but the weighing scheme should then downgrade that sensor's weighting.

6.4 Movement Behaviours for Sensor Calibration

While it is desirable from a practical perspective to calibrate sensors "on the fly" while the robot is already navigating, there are likely to be scenarios where the robot will need to perform a special calibration procedure before normal operations can resume. While the use of an initial set of calibration movements is not a novel concept in robotics and computer vision [Klein and Murray, 2007] , current techniques are geared towards tuning feature-based geometric systems or determining camera characteristics such as distortion or stereo baseline distance. Rat pups display a stereotypical sequence of repetitive warm up behaviours that increase in complexity in order to calibrate their sensory and mapping systems [Golani, et al., 1981] . We will develop movement behaviours inspired by rat pup head rotation and foot pivoting behaviour. Combined with assumptions about the relative trustworthiness of self-motion cues, such motion will provide a means of testing various sensors when their evaluation during normal navigation behaviour is too challenging.

References

[M. Cummins and P. Newman, 2009] M. Cummins and P. Newman, "Highly scalable appearance-only SLAM - FAB-MAP 2.0," in *Robotics: Science and Systems*, Seattle, United States, 2009.

[A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, 2007] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 29, pp. 1052-1067, 2007.

[H. Andreasson, T. Duckett, and A. Lilienthal, 2008] H. Andreasson, T. Duckett, and A. Lilienthal, "A Minimalistic Approach to Appearance-Based Visual SLAM," *IEEE Transactions on Robotics,* vol. 24, pp. 1-11, 2008.

[L. M. Paz, P. Pinies, J. D. Tardos, and J. Neira, 2008] L. M. Paz, P. Pinies, J. D. Tardos, and J. Neira, "Large-Scale 6-DOF SLAM With Stereo-in-Hand," *IEEE Transactions on Robotics,* vol. 24, pp. 946-957, 2008.

[K. Konolige and M. Agrawal, 2008] K. Konolige and M. Agrawal, "FrameSLAM: From Bundle Adjustment to Real-Time Visual Mapping," *IEEE Transactions on Robotics,* vol. 24, pp. 1066-1077, 2008.

[D. G. Lowe, 1999] D. G. Lowe, "Object recognition from local scale-invariant features," in *International Conference on Computer Vision*, Kerkyra, Greence, 1999.

[H. Bay, T. Tuytelaars, and L. Van Gool, 2006] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in *Computer Vision – ECCV 2006*, ed, 2006, pp. 404-417.

[A. J. Glover, W. P. Maddern, M. J. Milford, and G. F. Wyeth, 2010] A. J. Glover, W. P. Maddern, M. J. Milford, and G. F. Wyeth, "FAB-MAP + RatSLAM: Appearance-based SLAM for Multiple Times of Day," in *International Conference on Robotics and Automation*, Anchorage, United States, 2010.

[P. Biber and T. Duckett, 2005] P. Biber and T. Duckett, "Dynamic Maps for Long-Term Operation of Mobile Service Robots," in *Robotics: Science and Systems*, Cambridge, United States, 2005.

[I. Kiyoshi, Y. Tomoaki, and T. Masahiro, 2011] I. Kiyoshi, Y. Tomoaki, and T. Masahiro, "A High Dynamic Range Vision Approach to Outdoor Localization," in *International Conference on Robotics and Automation*, Shanghai, China, 2011.

[M. Milford and G. Wyeth, 2012] M. Milford and G. Wyeth, "SeqSLAM: Visual Route-Based Navigation for Sunny Summer Days and Stormy Winter Nights," in *IEEE International Conference on Robotics and Automation*, St Paul, United States, 2012.

[W. Maddern, M. Milford, and G. Wyeth, 2012] W. Maddern, M. Milford, and G. Wyeth, "CAT-SLAM: probabilistic localisation and mapping using a continuous appearance-based trajectory," *The International Journal of Robotics Research,* vol. 31, pp. 429-451, 2012.

[M. Milford, 2012] M. Milford, "Visual Route Recognition with a Handful of Bits," in *Robotics: Science and Systems VIII*, Sydney, Australia, 2012.

[J. A. Castellanos, J. Neira, and J. D. Tardós, 2002] J. A. Castellanos, J. Neira, and J. D. Tardós, "Multisensor fusion for simultaneous localization and map building," *IEEE Transactions on Robotics and Automation,* vol. 17, pp. 908-914, 2002.

[J. S. Gutmann and K. Konolige, 1999] J. S. Gutmann and K. Konolige, "Incremental mapping of large cyclic environments," in *Conference on Intelligent Robots and Applications*, 1999.

[M. Milford and G. Wyeth, 2010] M. Milford and G. Wyeth, "Persistent Navigation and Mapping using a Biologically Inspired SLAM System," *International Journal of Robotics Research,* vol. 29, pp. 1131-1153, 2010.

[M. J. Milford, 2008] M. J. Milford, *Robot Navigation from Nature: Simultaneous Localisation, Mapping, and Path Planning Based on Hippocampal Models* vol. 41. Berlin-Heidelberg: Springer-Verlag, 2008.

[M. Milford and G. Wyeth, 2008] M. Milford and G. Wyeth, "Mapping a Suburb with a Single Camera using a Biologically Inspired SLAM System," *IEEE Transactions on Robotics,* vol. 24, pp. 1038-1053, 2008.

[G. Dissanayake, P. M. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, 2001] G. Dissanayake, P. M. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localisation and map building (SLAM) problem," *IEEE Transactions on Robotics and Automation,* vol. 17, pp. 229-241, June, 2001 2001.

[S. Nuske, J. Robert, and G. Wyeth, 2009] S. Nuske, J. Robert, and G. Wyeth, "Robust outdoor visual localization using a three-dimensional-edge map," *Journal of Field Robotics,* vol. 26, pp. 728-756,

2009.

[C. J. Cornish, 1908] C. J. Cornish, *The Standard Library of Natural History* vol. 1: The University Society, 1908.

[I. Golani, G. Bronchti, D. Moualem, and P. Teitelbaum, 1981] I. Golani, G. Bronchti, D. Moualem, and P. Teitelbaum, ""Warm-up" along dimensions of movement in the ontogeny of exploration in rats and other infant mammals," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 78, pp. 7226-7229, 1981.

[E. Save, L. Nerad, and B. Poucet, 2000] E. Save, L. Nerad, and B. Poucet, "Contribution of multiple sensory information to place field stability in hippocampal place cells," *Hippocampus,* vol. 10, pp. 64-76, 2000.

[T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser, 2005] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser, "Microstructure of a spatial map in the entorhinal cortex," *Nature,* vol. 11, pp. 801-806, 2005.

[C. R. Nolan, G. Wyeth, M. Milford, and J. Wiles, 2010] C. R. Nolan, G. Wyeth, M. Milford, and J. Wiles, "The race to learn: spike timing and STDP can coordinate learning and recall in CA3," *Hippocampus,* 2010.

[D. Ball, S. Heath, M. Milford, and G. Wyeth, 2012] D. Ball, S. Heath, M. Milford, and G. Wyeth. (2012). *OpenRatSLAM*. Available: http://code.google.com/p/ratslam/

[D. E. Knuth, 2008] D. E. Knuth, *Introduction to combinatorial algorithms and Boolean functions*: Addison-Wesley, 2008.

[G. Klein and D. Murray, 2007] G. Klein and D. Murray, "Parallel tracking and mapping for small AR workspaces," in *International Symposium on Mixed and Augmented Reality*, Nara, Japan, 2007.