The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Robust Online Map Merging System using Laser Scan Matching and Omnidirectional Vision

Fredy Tungadi, Wen Lik Dennis Lui, Lindsay Kleeman and Ray Jarvis

***Abstract*— This paper describes a probabilistic online map merging system for a single mobile robot. It performs intermittent exploration by fusing laser scan matching and omnidirectional vision. Moreover, it can also be adapted to a multi-robot system for large scale environments. Map merging is achieved by means of a probabilistic Haar-based place recognition system using omnidirectional images and is capable of discriminating new and previously visited locations in the current or previously collected maps. This dramatically reduces the search space for laser scan matching. The combination of laser range finding and omnidirectional vision is very attractive because they reinforce one another when there is sufficient structure and visual information in the environment. In other cases, they complement one another, leading to improved robustness of the system. This is the first system to combine a probabilistic Haar-based place recognition system using omnidirectional images with laser range finding to merge maps. The proposed system is also algorithmically simple, efficient and does not require any offline processing. Experimental results of the approach clearly illustrate that the proposed system can perform both online map merging and exploration robustly using a single robot configuration in a real indoor lab environment.**

I. INTRODUCTION

The problem of map merging is an important but difficult problem in mobile robotics. It addresses the issue of merging independent maps collected by a team of mobile robots or merging partial maps collected by a single robot on different runs into a globally consistent map. The multi-robot system is more suitable for tackling large-scale environments since multiple robots can cooperate to explore the same environment. However, the same result can be achieved by using a single robot which performs intermittent explorations at different times. This eliminates the complexity of a multi robot system but with the tradeoff that exploration and mapping will take more time to complete. There are numerous applications ranging from autonomous home vacuum cleaners, lawn mowers, scout robots, search and rescue robots. Nonetheless, this problem has not been receiving as much attention as the SLAM problem.

The introduction of probabilistic frameworks and specifically laser-based probabilistic SLAM systems have achieved tremendous success in solving the SLAM problem. It is not surprising that many map merging systems in the literature are based on laser scan matching techniques [17], [18], [25]. This is because map merging systems are normally required to be capable of performing SLAM. Of course, the current state of the art is not only restricted to laser-based SLAM systems. The availability of cheap computing power and potential of vision systems have made the realisation of a visual SLAM approach practical. A popular approach is to identify and track distinctive landmarks/features in the environment. Unfortunately, there are no complete geometric map merging systems using vision as its primary sensor. In [14], two mobile robots each equipped with trinocular stereovision can localise themselves relative to each other at different times by detecting corresponding landmarks between the two maps. However, the maps are not merged into a global map. Another system described in [10] proposes multi-robot map building by taking visual measurements (with landmarks detected using SIFT [19]) from multiple robots and building a common global map for all the robots simultaneously instead of performing map merging. Nevertheless, this approach is only suitable if the relative starting positions of the robots are approximately known in advance.

As for laser-based map merging systems, there are many interesting approaches such as the virtual robot approach described in [1] which treats local laser scans from multiple robots as range measurements to the virtual robot and derives its odometry by registering similar structures in local maps. In another work described in [24], a hierarchical Bayesian approach is proposed to capture the structure of the environment (built using laser scans) using a hidden Markov process that represents transitions between views of the environment. The proposed system learns this structural model via an offline learning process and is able to predict what a typical view will look like if the robot moves out of a previously explored environment. This subsequently allows the system to estimate the probability of the robot being in a new or previously visited location. As the robot moves into a new location, the learnt model will be adapted and refined.

Unfortunately, these methods may fail because they rely heavily on distinct geometric laser scan feature descriptors for place recognition such as corners and edges which may not always be unique. As there are many geometrically/structurally similar places with different colour texture and/or visual features in typical man-made environments (e.g. supermarket), it is common to find laser range measurements being fused with visual data in order to improve the robustness of current SLAM systems. Of course, the difficulty which requires additional sensing modes to complement laser range finding arises mostly from the very limited sensing range of the employed laser range finder. If a SICK or longer range Hokuyo were used, the need for these other sensing

Fredy Tungadi, Wen Lik Dennis Lui, Lindsay Kleeman and Ray Jarvis are with the Intelligent Robotics Research Centre, Department of Electrical and Computer Systems Engineering, Monash University, Clayton Campus, Australia {fredy.tungadi, wen.lui, lindsay.kleeman, ray.jarvis}@eng.monash.edu.au

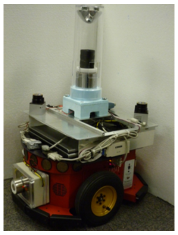

Fig. 1. The ActivMedia Pioneer P3-DX with two Hokuyo laser range finders and an omnidirectional vision system.

modes may be drastically reduced (also depends on the scale of the environment). For the case where the sensing range of the laser range finder is sufficient, our proposed system is able to reduce the time required to perform scan matching. In other cases, the proposed system, which combines laser scan matching and omnidirectional vision, can further improve its robustness in environments where laser scan matching is ambiguous.

It is also common to find that loop closing techniques in SLAM can also be easily adapted to solve the map merging problem since these problems overlap to a certain extent. Both problems attempt to identify whether the robot is at a previously visited location with the additional extension to merge corresponding maps into a globally consistent one for the map merging problem. This paper shares the same idea with the work by Ho and Newman [12], in the sense that both visual and laser range information is required for a robust place recognition system which can be used to solve both the loop closing and map merging problems. One of the main limitations of the visual loop closure detection system is the lack of a probabilistic framework to allow the system to degrade gracefully when uncertainty becomes prevalent. However, this limitation has been resolved by Cummins and Newman [6] by proposing a generalised probabilistic framework for an appearance-based localisation and mapping system. Unfortunately, the proposed method requires a one-off learning process to build its visual dictionary for the bag-of-words paradigm which may take hours to perform. Nevertheless, this method still performs well (suboptimal) when a standard dictionary is used instead.

To the best of our knowledge, this is the first work which combines a probabilistic Haar-based place recognition system using omnidirectional vision with a SLAM system based on laser scan matching for map merging. Instead of the pan-tilt camera system used in [12], [6], an omnidirectional vision system is used to alleviate the *windowing* problem with perspective cameras and the time required to produce an image with 360^o FOV is significantly reduced although the effective resolution of the image is lower. Furthermore, the proposed approach is algorithmically simple, efficient and does not require any offline processing, and hence the whole process of map merging can be executed in real time.

Since the proposed system is tested on a single robot configuration, one or more previously collected maps can be loaded into the system. Subsequently, the robot performs SLAM by associating consecutive laser scans via an EKF framework. These laser scans or local maps will also be associated with an omnidirectional image which describes the appearance of the location. At the same time, exploration is performed autonomously using the Voronoi Loop Exploration strategy described in [27]. This is found to be highly beneficial for this map merging system due to the usage of Voronoi graphs which ensures that the robot traverses close to the previous path, such that loop closing and map merging opportunities will not be missed (will be thoroughly discussed in Section III-B). The system will proceed with map merging via scan matching only if the place recognition system detects a previously explored environment. Finally, the maps are merged if scan matching succeeds.

The rest of the paper is organised as follows: A brief description of the research platform is provided in Section II. This is then followed by the introduction of the SLAM and probabilistic Haar-based place recognition system in Section III and IV respectively and how these are integrated to tackle the map merging problem in Section V. This approach is then validated by the experimental results in Section VI followed by discussion and possible future extensions in Section VII. Finally, conclusions are presented in Section VIII.

II. SYSTEM OVERVIEW

Our main research platform is the two wheeled differential drive ActivMedia Pioneer P3-DX [22] illustrated in Fig. 1. It is equipped with two Hokuyo URG-04LX laser range finders, each with a maximum range of 4m, mounted on the front and rear of the robot covering a full 360^o of the plane. In addition, an omnidirectional vision system made up of a PixeLink camera looking upwards to an equiangular mirror designed by Chahl and Srinivasan [5] is mounted onto the centre of the robot.

III. AUTONOMOUS EXPLORATION AND SCAN MATCHING SLAM

The SLAM and exploration algorithm are based on a previous work described in [27], except that the Advanced Sonar sensors [16] have been replaced with the omnidirectional vision system for the purpose of place recognition. The Advanced Sonar sensors were previously used for producing range and bearing measurements to small corner cube type targets naturally occurring in doorjambs and other corridor wall features not resolvable by the laser. These measurements were used to assist the scan matching in localising the robot in corridor environments. For this reason, in this work, the environment is assumed not to be made up of relatively long corridors in which the localisation of the robot can be compromised by the unavailability of advanced sonar sensors. Of course, the need of advanced sonar sensing in [27] also arises from the limited range and resolution of the employed laser range finder. However, in environments where even the sensing range of current laser range finder is not sufficient, these additional sensing modes will become invaluable. It is also important to note that in this work, the

SLAM implementation is based on laser only and that the vision is only used to recognise previously visited locations for map merging. This issue will be further discussed in Section VII. In addition, the inclusion of the autonomous exploration algorithm is to allow a fully autonomous operation for SLAM and online map merging.

A. SLAM

The SLAM algorithm is implemented using the EKF (Extended Kalman Filter) framework described by A. Davison [7] where all map features are included in the SLAM state vector and updated on each observation step. The prediction model based on odometry is derived as in [15], where the error model assumes error sources are additive white noise on the wheel separation as well as the left and right wheel distance measurements.

As in [26], landmarks are defined by a template of the two raw laser measurements which are collected approximately every 0.7 metre of robot travel and observed by a process of scan matching. There are many available approaches for scan matching. However in this paper, Polar Scan Matching (PSM) [8] has been chosen because of its fast convergence. It operates in the laser scanner's polar coordinate system and therefore eliminates the need to search for correspondence by simply matching the points by their bearing.

The augmented state vector containing both the state of the robot (θ_v, x_v, y_v) and the state of landmark locations is defined as follows,

$$X = [\theta_v, x_v, y_v, L_1, L_2, ..., L_n] \tag{1}$$

where the landmarks detected by the lasers are represented by the pose of the centroid of the reference scans in the global coordinate frame, $L_i = (x_{Li}, y_{Li}, \theta_{Li})$. The observation model for the pose of a landmark coordinate frame with respect to the robot is calculated as follows,

$$H_L(t) = \begin{bmatrix} x_{hi}(t) \\ y_{hi}(t) \\ \phi_{hi}(t) \end{bmatrix} = \begin{bmatrix} (x_{Li}(t) - x_v(t))cos(\theta_v(t)) + (y_{Li}(t) - y_v(t))sin(\theta_v(t)) \\ -(x_{Li}(t) - x_v(t))sin(\theta_v(t)) + (y_{Li}(t) - y_v(t))cos(\theta_v(t)) \\ \phi_{Li}(t) - \theta_v(t) \end{bmatrix} \tag{2}$$

A heuristic error estimation approach as described in [26] has been used for correctly providing the error covariance of the scan matching to the EKF in corridor environments.

B. Autonomous Exploration

As described in [27], the exploration algorithm takes advantage of the characteristic of the Voronoi Graph to enable the robot to strategically explore the environment in both a loop closing fashion and safe manner. It is safe, because the line segments in the Voronoi Graph are always equidistant from nearby obstacles in the environment and hence it naturally makes the robot to travel in the path with maximum clearance from obstacles. For gaps that are too small for traversal, the perpendicular distance from line segments of the Voronoi Graph to the nearest obstacle which is smaller than the radius of the robot is removed. Moreover, the Voronoi Graph also exhibits loop-paths that can help in guiding the robot to close loops during exploration in order to maintain the SLAM consistency.

Fortune's plane sweep algorithm [9], which provides a simple $O(nlogn)$ solution, has been used to generate the Voronoi Graph. The Voronoi Graph is converted into an undirected-weighted graph structure to allow it to be used for path planning and exploration.

The exploration algorithm works by periodically extracting loop-paths using the loop-path extraction algorithm [27] and subsequently executing these loop-paths in order of size (small to large) until all loop-paths have been executed. Then, a graph-based exploration technique is used to fully map the environment. This strategy ensures a stable partial map creation before the robot travels further to explore the rest of the environment.

The usage of the Voronoi Graph for exploration also ensures that the robot is close to the previous track when it revisits an explored part of the environment. This is beneficial for loop closing, place recognition and map merging techniques based on the visual appearance of the location because visual information will be significantly different when the robot is too far off track. This is because a new location will not match previous locations and thus resulting in lost opportunities for loop closing. Of course, this is not possible if new obstacles are present.

IV. PROBABILISTIC HAAR-BASED PLACE RECOGNITION

A. Place Recognition using Haar Wavelets

The place recognition system is based on the image retrieval system described in [11] and is originally proposed by Jacobs et al. [13]. Ho and Jarvis [11] have successfully adapted this algorithmically simple, efficient and yet robust framework for mobile robot localisation and illustrated its robustness against lighting variation and occlusion. In the original system [13], RGB images are converted into YIQ color space, decomposed using the standard 2D Haar decomposition technique, and the top 60 coefficients (quantised magnitudes and locations) are retained as the image signature. Subsequently, whenever a query image is presented to the system, it will be decomposed, quantised and a weighted score (depending on the location of the coefficient) is calculated.

Ho and Jarvis [11] customised this to work with panoramic images by downsampling the original unwarped image to a size of 512 x 128 and retaining the coefficients within a bounding box of size 64 x 16 originating from the (0,0) coordinate of the decomposed image. The magnitude of these coefficients are quantised and conveniently stored into a bit array, which significantly reduces the memory footprint for each image signature (location of coefficient is not required). Since the Haar wavelets vary rotationally, the unwarped panoramic image is column-wise shifted every 10 degrees

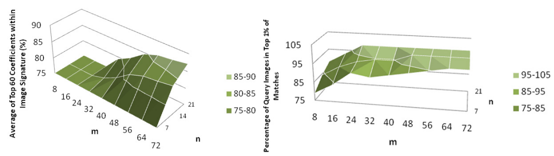

Fig. 2. (Left) Average of top 60 coefficients within bounding box of size m by n and (Right) average top 1% matches using bounding box size *m* by *n*.

equivalent in pixels and decomposed, quantised and stored in the database.

In Fig. 2 (Left), a total of 335 panoramic images were used to find the average number of top 60 coefficients within the bounding box of size *m* by *n* and Fig. 2 (Right) shows the effect of matching accuracy with different bounding box sizes using 57 query images for a database with 2052 image signatures.With these quantitative results, an image signature of size 56 x 14 was chosen instead which contains an average of 82.8% of the top 60 coefficients and performing at an average of 98.2% to rank the correct image signature in the database in the top 1% of all returned matches for the 57 query images. An independent evaluation using the Haar wavelet for place recognition can be found in [20].

B. Bayesian Place Recognition System

The proposed system is based on the probabilistic, incremental and online loop closing detection framework described in [2] which employs the discrete Bayes filter. However, several modifications were made due to some significant differences in system characteristics. Firstly, only a single image from the omnidirectional vision system is associated with the corresponding laser scan in our system instead of using a continuous stream of video images. Then, Haar coefficients of the omnidirectional images are used to discriminate between the images instead of the bag-of-words approach using perspective cameras. Since each image is associated with a particular laser scan, we maintain the topological relationship between these images in a bidirectional graph structure.

Problem Definition: Given a set of nodes, $p = 1, 2, ..., t$, (previously collected, current or combination of both) each associated with an omnidirectional image in the sequence $I_1, I_2, ...I_t$, compute the probability of the robot being in a previously explored node and the probability of it being in an unmapped location. The variable S_t is then used to describe the hypothesis that at time *t*, the robot is at a previously visited node if S_t = *j*, where *j* denotes an existing node's index, or a new location if *j=-1*. In a Bayesian framework, this is equivalent to searching for the past image I_k where the index k is derived by using the following expression,

$$k = \underset{j=-1,\dots,t-e}{\operatorname{argmax}}\ p\left(S_t = j|I^t\right) \tag{3}$$

where $I^t = I_1, ...I_t$. Subsequently, by using Bayes rule and the Markov assumption, the full posterior is decomposed into,

$$p\left(S_t|I^t\right) = \eta p\left(I_t|S_t\right) p\left(S_t|I^{t-1}\right) \tag{4}$$

where η is the normalisation factor and $p\left(I_t|S_t\right)$ is the likelihood $\mathcal{L}\left(S_t|I_t\right)$ of S_t given image I_t. By marginalising the belief distribution, $p\left(S_t|I^{t-1}\right)$, it can then be written as,

$$p\left(S_t|I^{t-1}\right) = \sum_{k=-1}^{t-e} \underbrace{p\left(S_t|S_{t-1}=k\right)}_{state\ transition} \underbrace{p\left(S_{t-1}=k|I^{t-1}\right)}_{prior\ posterior} \tag{5}$$

where *e* is the number of most recently visited nodes to be excluded (set to 1 in our experiments) and the prior initialized to $p(S_{t-1} = -1|I^{t-1}) = 1$. This decomposition is simply elegant because the state transition model can be used to maintain temporal coherency and reduce transient errors due to perceptual aliasing and the prior posterior is readily available. The following will describe how the state transition is modeled and the likelihood $\mathcal{L}\left(S_t|I_t\right)$ is computed.

C. State Transition Model

The state transition probabilities are modeled according to our system characteristics. Given that $e \geq t$ and $e = 0$ if $e < t$, the state transition probabilities can be described as follows,

- p$(S_t = -1|S_{t-1} = -1) = \frac{2.0}{t-e+1.0}$ if the system believes it was at a new location in the previous time step ($SI = -1$) or $\frac{1.5}{t-e+0.5}$ otherwise: describes the probability that the robot is at a new location at time *t* given that it was previously at a new location at time *t-1*.
- p$(S_t = j|S_{t-1} = -1) = \frac{1-p(S_t=-1|S_{t-1}=-1)}{t-e}$, $j \in (0, t-e)$: describes the probability that the robot is currently at node *j* given that the robot was previously at a new location at time *t-1*.
- p$(S_t = -1|S_{t-1} = k) = 1 - \frac{N_k+1}{N_k+2}$ if the system believes it was at a new location in the previous time step ($SI = -1$) or 0.2 otherwise, $k \in (0, t-e)$: describes the probability that the robot is currently at a new location given that the robot was previously at node *k* at time *t-1*.
- p$(S_t = j|S_{t-1} = k)$, where $j \in (0, t-e)$ and $k \in (0, t-e)$: describes the probability of the robot currently at node *j* given that the robot was previously at node *k* at time *t-1* and can be expressed as,

$$\frac{1.6}{N_k+2.0} : SI > -1, j = k \tag{6}$$

$$\frac{2N_k+2.0}{(N_k+2.0)^2} : SI = -1, j = k \tag{7}$$

$$\frac{0.8-\frac{1.6}{N_k+2.0}}{N_k} : SI > -1, j \in \text{neighbours of } k \tag{8}$$

$$\frac{0.8-\frac{2N_k+2.0}{(N_k+2.0)^2}}{N_k} : SI = -1, j \in \text{neighbours of } k \tag{9}$$

where N_k represents the total number of neighbours of node *k* (can be found using the graph which maintains the topological relationship between the images) and SI is the selected index (largest probability) from the prior posterior,

$p\left(S_{t-1}=k|I^{t-1}\right)$, which represents the actual state of the system being at an unexplored location if $SI = -1$ or previously visited location if $SI > -1$ in the previous time step.

D. Likelihood Voting Scheme

The likelihood is computed by finding a subset $H_t \subseteq I^{t-e}$ of images whose score, D_i ($i \in (-1, t-e)$), is smaller than the threshold computed using the mean of all scores, μ_a, minus its standard deviation, σ (more negative Haar scores being a better match). At the same time, the mean of all inliers, μ_{in}, which represents the average score of those which is larger than the threshold, is computed. Subsequently, $\mathcal{L}\left(S_t|I_t\right)$ is expressed as,

$$\mathcal{L}\left(S_t|I_t\right) = \begin{cases} \frac{(D_i-\mu_{in})^A+\mu_{in}}{\mu_{in}} & : D_i \leq \mu_a - \sigma \\ 1.00 & : \text{otherwise} \end{cases} \tag{10}$$

where A can be normally expressed as,

$$A = \frac{log_{10}(15/p\left(S_t=j|S_{t-1}=-1\right))}{log_{10}(15)} \tag{11}$$

For the case when $i=-1$ and $SI > -1$, A is expressed as,

$$A = \frac{log_{10}(15/p\left(S_t=j|S_{t-1}=-1\right))}{log_{10}(15)} - 0.3 \tag{12}$$

In [2], a virtual image, I_{-1}, is created and maintained such that it is statistically more likely for this virtual image to match the incoming query image if the robot is currently at an unexplored location. However, for our system, the score of this virtual image is calculated by subtracting the mean, μ_a, with 2 times the standard deviation, σ. As such, $\mathcal{L}\left(S_t=-1|I_t\right)$ is expressed as,

$$\mathcal{L}\left(S_t=-1|I_t\right) = \mu_a - 2.0\sigma \tag{13}$$

Finally, the full posterior, $p\left(S_t|I^t\right)$, is computed and subsequently normalised.

V. MAP MERGING

The probabilistic Haar-based place recognition system decides whether the current location of the robot already exists in the current and/or the previously built maps. When the system returns a match, one of the nodes in the current or previously built maps will have a larger probability than the others. This matching node is then utilized to find the associated laser reference scan. Since Haar wavelets vary rotationally, the system can also provide the approximate relative orientation of the current robot's position with respect to the reference image/laser scan up to a resolution of 10^o. This is because any unwarped panoramic image to be stored into the database is column-wise shifted every 10^o equivalent in pixels. This relative orientation is used in scan matching to improve the convergence time and success rate.

Since each match returned by the place recognition system is validated by laser scan matching, the overall robustness of the system improves. Scan matching is performed by firstly transforming the coordinate frame of the current reference scan to the matching reference scan and then PSM is performed with the suggested orientation from the probabilistic place recognition system by iteratively minimising the sum of square range residual. The output of scan matching is the relative pose of the current reference scan with respect to the previous matching reference scan. Once the match is validated by scan matching, it can mean either loop closing in the current map or map merging with previously built maps. This is decided based on the index value returned by the place recognition system. It is loop closing if the returned index value refers to a node which exists in the current map or map merging if the returned index refers to a node which exists in the previously built maps.

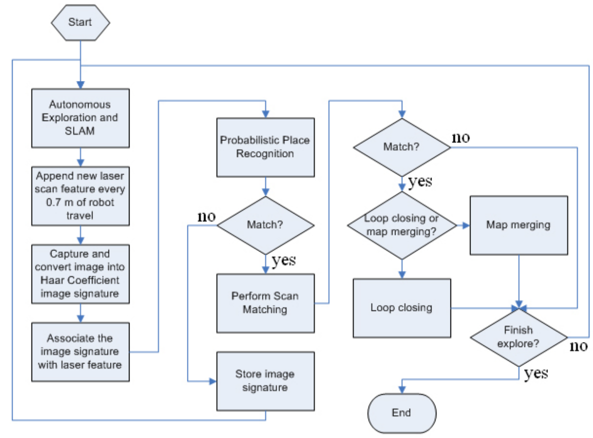

Fig. 3. Flowchart of the online map merging algorithm.

For the case of loop closing, the current map is updated with the scan matching result, using the standard EKF update equation [7]. Since loop closure is detected, there is no need to append any new reference scan features into the SLAM state vector. In contrast, for the case of map merging, the relative pose resulting from the scan matching is used to find the relative transformation matrix of the current map with the previously built map. The transformation matrix is then used to transform the entire current reference scan to the previous map reference frame. However, it does not end here, since the two maps are yet to be correlated and combined into one state vector. Subsequently, in order to maintain the correct correlation between the maps, each of the pose is appended into the SLAM state vector one by one as in [7] using the following equations,

$$X_{new} = [x_v, L_1, ..., L_n, L_i^{new}] \tag{14}$$

$$P_{new} = \begin{bmatrix} P_{xvxv} & P_{xvL1} & \cdots & P_{xvLn} & P_{xvxv}(D_L)^T \\ P_{L1xv} & P_{L1L1} & \cdots & P_{L1Ln} & P_{L1xv}(D_L)^T \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P_{Lnxv} & P_{LnL1} & \cdots & P_{LnLn} & P_{Lnxv}(D_L)^T \\ D_L P_{xvxv} & D_L P_{xvL1} & \cdots & D_L P_{xvLn} & P_{L_i^{new} L_i^{new}} \end{bmatrix} \tag{15}$$

where

$$D_L = (\frac{\partial L_i^{new}}{\partial x_v}) \quad (16)$$

$$P_{L_i^{new} L_i^{new}} = D_L P_{xvxv} (D_L)^T + (\frac{\partial L_i^{new}}{\partial H}) R (\frac{\partial L_i^{new}}{\partial H})^T \quad (17)$$

L_i^{new} is the ith pose of the reference scan from the previous map, x_v is the current pose of the robot, H is the measurement of the new landmark L_i^{new} and R is the estimated covariance of H. In addition, the graphs which maintain the topological structure of these maps for the place recognition system are merged accordingly.

The overall operation of the autonomous exploration and SLAM with online map merging is summarised in the flowchart of Fig. 3.

VI. EXPERIMENTAL RESULTS

Experiments were conducted in indoor lab environments at Monash University. In these experiments, the mobile robot creates a partial map of the environment before it is moved to a random location. The system can either be turned off/reset (with the partial map preloaded into the system prior to execution) or remain switched on while it is being moved from one location to the other. However, in these experiments, the former case is illustrated in order to simulate a typical indoor mobile robot which is more likely to be switched on and off at different times and days for exploration and mapping rather than a one off continuous process.

The omnidirectional vision system is capable of producing images at resolutions of 1280 x 1024. However, in these experiments, we are only using images at resolutions of 640 x 512. The robot performs autonomous exploration and SLAM, and creates unique reference laser scans at intervals of approximately 0.7m which are associated with the corresponding omnidirectional images that describe the appearance of the locations. In the first experiment, the mobile robot creates a partial map of lab G15 (Fig. 4 (Left)). It is then restarted in a random location and preloaded with the previously collected partial map and heads towards the opposite direction for exploration (Fig. 4 (Right)). Subsequently, the robot detects that it is at a location in the previously collected partial map and performs map merging in Fig. 5. The robot continues to traverse the environment and Fig. 5(b) shows the locations where loop closing and false negatives were detected by the place recognition system. No false positives were reported by the system. A video of this experiment can be found at
`http://www.youtube.com/watch?v=LymgfkVpwLs`

For the second experiment, we test the robustness of the algorithm by introducing a more challenging environment to the system. In this experiment, the robot initially explores and maps the entire lab G15 (Fig. 6 (Left)). Then, it is restarted and randomly placed in lab G10 (neighbour lab to G15) preloaded with the map of lab G15 (Fig. 6 (Right)). Then, the robot revisits some nodes in lab G10 before it heads out to the path which leads to lab G15. Subsequently, the robot detects that it is at a location in the preloaded map of lab G15 and performs map merging in Fig. 7. The robot continues to traverse the environment and Fig. 7(b) shows the locations where loop closing and false negatives were detected before and after map merging is performed. Similarly, no false positives were reported by the system in this experiment. A video of this experiment can be found at
`http://www.youtube.com/watch?v=dQemNJX3kAY`

The third experiment is similar to the second experiment except that the robot is required to explore a larger environment and perform 3 map merging instead of 1. The robot is preloaded with 4 non-overlapping maps as shown in Fig. 8. Then, it is placed at a random location in lab G10 and autonomously explores the environment. It is not possible to merge with the map of the High Voltage lab (Fig. 8(c)) since the robot has been deliberately denied access to it. Nevertheless, this map is still loaded into the system to create more ambiguitiy. Experimental results are shown in Fig. 9-11. Fig. 11 shows the locations where loop closing, false negatives and positives were detected before and after map merging is performed. The two false negatives were rejected using laser scan matching and a video of this experiment can be found at `http://www.youtube.com/watch?v=GXlWLdit5TM`

The indices of the full posterior probability shown in Fig. 5(c) and Fig. 7(c) still correspond to the indices of the reference laser scans as shown in Fig. 5(b) and Fig. 7(b). However, in the videos, when false negatives or map merging occurs, it can be observed that these indices do not correspond anymore. Nonetheless, these relationships are properly maintained internally in the system.

VII. DISCUSSION

From the experimental results, it is clearly shown that the system can robustly perform map merging in challenging environments such as geometrically similar corners and junctions. The process of detecting loop closure and online map merging is also made more efficient since the search area for scan matching is significantly reduced when the place recognition system provides it with the matching node and an approximate relative orientation with respect to its reference. Fig. 12 compares the time required to perform an exhaustive scan matching as opposed to image querying on a laptop with a 1.6GHz AMD processor and 1Gb RAM.

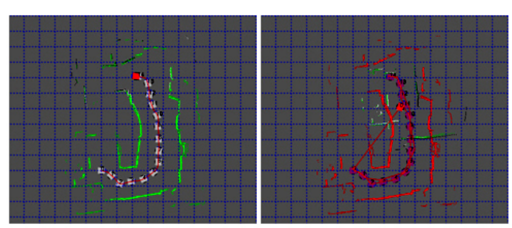

Fig. 4. (Left) Partial map of Lab G15 and (Right) Mobile robot loads the partial map (red) and starts at an unknown location in Lab G15. Grid size is 1x1m.

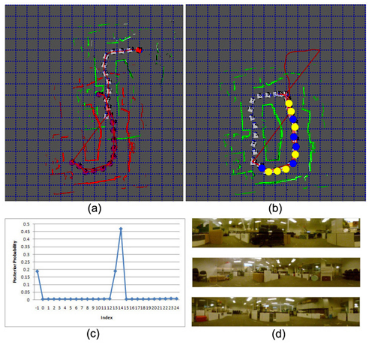

Fig. 5. (a) Map merging detected at this location. (b) Merged map. The probabilistic place recognition system detects accurate loop closures at locations of yellow circles and false negatives at locations of blue circles after map merging is performed. (c) Posterior probability at the merging pose. (d) Examples of unwarped omnidirectional images of the environment. Grid size is 1x1m.

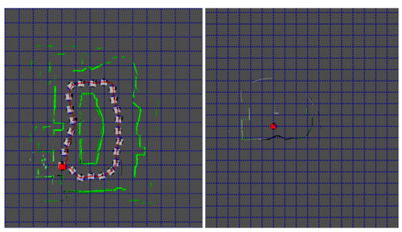

Fig. 6. (Left) Map of Lab G15 and (Right) Mobile robot loads the previously collected map (not shown here) and starts at an unknown location in Lab G10. Grid size is 1x1m.

Furthermore, image querying can be further optimised by replacing the current image querying technique described in [13] with a kd-tree implementation.

As illustrated in the videos, the system builds up its belief that it is indeed in a previously explored environment (depending on how discriminative the image matching scores are) and dampens the effect of perceptual aliasing which may lead to false positives by utilizing the state transition model (described in Section IV-C). Of course, there are instances when this belief builds up rather quickly if the scores are highly discriminative and this is not surprising since discriminative image matching scores provide a strong indication that this is indeed a good match. Nevertheless, the overall adaptation of the probabilistic framework proposed by Angeli et al. [2] has been successful although there are many differences in terms of system characteristics that makes the original model being unusable if it were to be directly applied into our system. It is also more risky to decide a place has been visited when it has not (false positive) than to wrongly decide it has not been visited when it has (false negative). Although this issue has been partly addressed in the current framework but this can be further improved and extended by incorporating the likelihood ratio test (min. Bayes risk) [23] such that less risk is taken when loop closing/map merging is decided.

However, further extension of the current system will be required to tackle the extreme case of a perfect corridor environment (long and straight corridors without any geo-

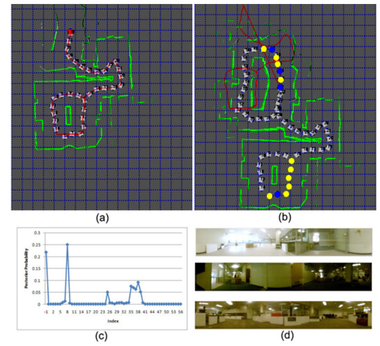

Fig. 7. (a) Map merging detected at this location. (b) Merged map. The probabilistic place recognition system detects accurate loop closures at locations of yellow circles and false negatives at locations of blue circles before and after map merging is performed. (c) Posterior probability at the merging pose. (d) Examples of unwarped omnidirectional images of the environment. Grid size is 1x1m.

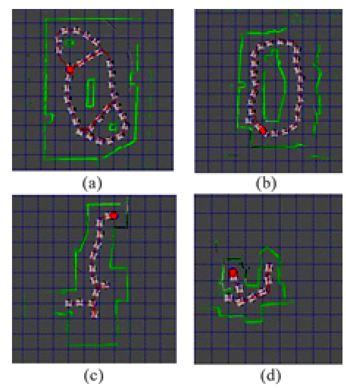

Fig. 8. Partial maps created for the third experiment. (a) Map of Lab G10 (rearranged), (b) Map of Lab G15, (c) Map of High Voltage Lab and (d) Map of Room G13. Grid size is 1x1m.

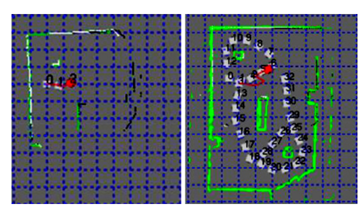

Fig. 9. The first merging in the third experiment. (Left) Map merging detected at this location. (Right) Merged map. Grid size is 1x1m.

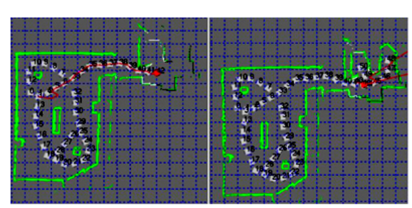

Fig. 10. The second merging in the third experiment. (Left) Map merging detected at this location. (Right) Merged map. Grid size is 1x1m.

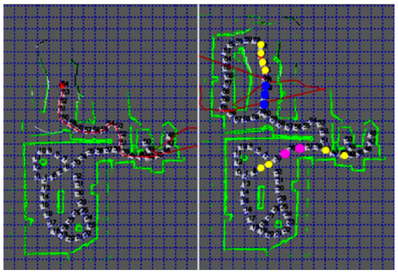

Fig. 11. The third merging in the third experiment. (Left) Map merging detected at this location. (Right) Merged map. The probabilistic place recognition system detects accurate loop closures at locations of yellow circles, false negatives at locations of blue circles and false positives at locations of magenta circles before and after map merging is performed. Grid size is 1x1m.

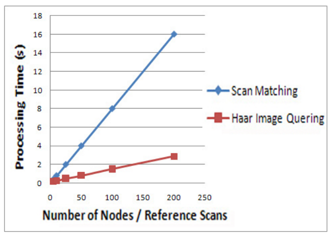

Fig. 12. Comparison of image querying and scan matching processing times (exhaustive search through all reference scans/nodes).

metrical differences which is briefly mentioned in Section III). With the current system, the place recognition system is able to detect that map merging should take place even when it is in a long corridor. However, scan matching will fail in such situations because of these perfectly straight and aligned corridors. To resolve this issue, we intend to extend this system by calibrating the omnidirectional vision system using the proposed calibration method in [21] in order to compute the 3D location of distinctive SURF features [3] in the unwarped panoramic images. With these landmarks, map merging is possible even when scan matching fails.

VIII. CONCLUSIONS

This is the first system to combine a probabilistic Haar-based place recognition system using omnidirectional images with laser ranging to perform map merging. By using an omnidirectional vision system instead of a pan-tilt camera unit, it alleviates the *windowing* problem commonly encountered with perspective cameras and the time required to produce an image with 360^o FOV is significantly reduced. In conclusion, we have experimentally validated the proposed system is able to perform online map merging more robustly and efficiently. With the intention to extend the system to track the 3D location of reliable visual landmarks, we strongly believe that we are one step closer to solving the map merging problem.

REFERENCES

[1] N. Adluru, L.J. Latecki, M. Sobel and R. Lakaemper, "Merging Maps of Multiple Robots", *in IEEE International Conference on Pattern Recognition*, 2008, pp 1-4.

[2] A. Angeli, S. Doncieux, J.-A. Meyer and D. Filliat, "Real-Time Visual Loop-Closure Detection", *in IEEE International Conference on Robotics and Automation*, 2008 pp 1842-1847.

[3] H. Bay, A. Ess, T. Tuytelaars, L Van Gool, "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding*, vol. 110, no. 3, 2008, pp 346-359.

[4] S. Carpin, A. Birk and V. Jucikas, "On map merging", *Robotics and Autonomous Systems*, vol. 53, 2005, pp 1-14.

[5] J.S. Chahl and M.V. Srinivasan, "Reflective Surfaces for Panoramic Imaging", *Applied Optics*, vol. 36, no. 31, 1997, pp 8275-8285.

[6] M. Cummins and P. Newman, "FAB-MAP Probabilistic Localization and Mapping in the Space of Appearance", *International Journal of Robotics Research*, vol. 27, no. 6, 2008, pp 647-665.

[7] A. Davison, "Mobile Robot Navigation using Active Vision", *PhD Thesis, University of Oxford*, 1998.

[8] A. Diosi and L. Kleeman, "Fast Laser Scan Matching Using Polar Coordinates", *International Journal of Robotics Research*, vol. 26, 2007, pp 1125-1153.

[9] S. Fortune, "Voronoi Diagrams and Delaunay Triangulations", *Handbook of Discrete and Computational Geometry, CRC Press, Inc., Boca Raton, FL*, 1997.

[10] A. Gil, Ó. Reinoso, M. Ballesta, M. Juliá, "Multi-robot Visual SLAM Using a Rao-Blackwellized Particle Filter", *Robotics and Autonomous Systems*, vol. 58, no. 1, 2010, pp 68-80.

[11] N. Ho and R. Jarvis, "Vision Based Global Localisation Using a 3D Environmental Model Created by a Laser Range Scanner", *in IEEE International Conference on Intelligent Robots and Systems*, 2008, pp 2964-2969.

[12] K. Ho and P. Newman, "Detecting Loop Closure with Scene Sequences", *International Journal of Computer Vision*, vol. 74, no. 3, 2007, pp 261-286.

[13] C.E. Jacobs, A. Finkelstein and D.H. Salesin, "Fast Multiresolution Image Querying", *Computer Graphics*, 29 (Annual Conference Series), 1995, pp 277-289.

[14] C. Jennings, D. Murray and J. Little, "Cooperative Robot Localization with Vision-based Mapping", *in IEEE International Conference on Robotics and Automation*, 1999, pp 2659-2665.

[15] L. Kleeman, "Advanced Sonar and Odometry Error Modelling for Simultaneous Localisation and Map Building", *in International Conference on Intelligent Robots and Systems*, 2003, pp 699-704.

[16] L. Kleeman, "On-The-Fly Classifying Sonar with Accurate Range and Bearing Estimations", *in IEEE International Conference on Intelligent Robots and Systems*, 2002, pp 178-183.

[17] K. Konolige, D. Fox, B. Limketkai, J. Ko and B. Stewart, "Map Merging for Distributed Robot Navigation", *in IEEE International Conference on Intelligent Robots and Systems*, 2003.

[18] A. León, R. Barea, L.M. Bergasa, E. López, M. Ocaña and D. Schleicher, "SLAM and Map Merging", *Journal of Physical Agents*, vol. 3, no. 1, 2009, pp 13-23.

[19] D.G. Lowe, "Distinctive Image Features from Scale Invariant Keypoints", *International Journal of Computer Vision*, vol. 60, no. 2, 2005, pp 91-110.

[20] W.L.D. Lui and R. Jarvis, "A Pure Vision-based Approach to Topological SLAM", *IEEE International Conference on Intelligent Robots and Systems*, 2010.

[21] W.L.D. Lui and R. Jarvis, "Eye-Full Tower: A GPU-based Variable Multibaseline Omnidirectional Stereovision System with Automatic Baseline Selection for Outdoor Mobile Robot Navigation", *Robotics and Autonomous Systems*, vol. 58, no. 6, 2010, pp.747-761.

[22] Mobile Robots, Online, 2010, http://www.mobilerobots.com.

[23] R.J. Radke, S. Andra, O. Al-Kohafi and B. Roysam, "Image Change Detection Algorithms: A Systematic Survey", *IEEE Transactions on Image Processing*, vol. 14, no. 3, 2005, pp 294-307.

[24] B. Stewart, J. Ko, D. Fox and K. Konolige, "The Revisiting Problem in Mobile Robot Map Building: A Hierarchical Bayesian Approach", *in Proceedings on the Conference of Uncertainty in Artificial Intelligence*, 2003.

[25] S. Thrun, "A Probabilistic On-Line Mapping Algorithm for Teams of Mobile Robots", *International Journal of Robotics Research*, vol. 20, no. 5, 2001, pp 335-363.

[26] F. Tungadi and L. Kleeman, "Multiple Laser Polar Scan Matching with Application to SLAM", *in Australasian Conference on Robotics and Automation*, 2007.

[27] F. Tungadi and L. Kleeman, "Loop Exploration for SLAM with Fusion of Advanced Sonar Features and Laser Polar Scan Matching", *in IEEE International Conference on Intelligent Robots and Systems*, 2009, pp 388-394.