IEEE International Conference on Robotics and Automation, San Francisco, CA, April 2000, pp. 1023-1029. Best Vision Paper Award.

Appearance-Based Place Recognition for Topological Localization

Iwan Ulrich and Illah Nourbakhsh

The Robotics Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
iwan@ri.cmu.edu, illah@ri.cmu.edu

ABSTRACT

This paper presents a new appearance-based place recognition system for topological localization. The method uses a panoramic vision system to sense the environment. Color images are classified in real-time based on nearest-neighbor learning, image histogram matching, and a simple voting scheme. The system has been evaluated with eight cross-sequence tests in four unmodified environments, three indoors and one outdoors. In all eight cases, the system successfully tracked the mobile robot's position. The system correctly classified between 87% and 98% of the input color images. For the remaining images, the system was either momentarily confused or uncertain, but never classified an image incorrectly.

1. INTRODUCTION

Localization is a fundamental problem in mobile robotics. Most mobile robots must be able to locate themselves in their environment in order to accomplish their tasks. Since mobile robot localization is a prerequisite for most applications, research has been very active in this field [4].

Current localization methods can be classified as being *geometric*, *topological*, or *hybrid*. Geometric approaches typically use a two-dimensional grid as a map representation. They attempt to keep track of the robot's exact position with respect to the map's coordinate system. Topological approaches use an adjacency graph as a map representation. They attempt to determine the node of the graph that corresponds to the robot's location. Hybrid methods combine geometric and topological approaches.

Most of the recent work on mobile robot localization focuses on geometric localization. In general, these geometric approaches are based on either *map matching* or *landmark detection*. Most map matching systems rely on an extended Kalman filter that combines information from *intrinsic* sensors with information from *extrinsic* sensors to determine the current position [2,9,16]. Good statistical models of the sensors and their uncertainties must be provided to the Kalman filter. Situations where the robot travels over a bump (e.g., a cable lying on the floor) are difficult to model and can thus lead to unpredictable results.

Landmark localization systems rely on either *artificial* or *natural* features of the environment. Artificial landmarks are easier to detect reliably than natural landmarks. However, artificial landmarks require modifications of the environment, such that systems based on natural landmarks are often preferred. Various features have been used as natural landmarks: corners [8], doors [3,22], overhead lights [22], air diffusers in ceilings [1], and distinctive buildings [20]. Because most of the landmark-based localization systems are tailored for specific environments, they can rarely be easily applied to different environments.

Although research in geometric localization has been very active and has made recent improvements, we are not aware of any system that works well both indoors and outdoors. In contrast, this paper presents a novel topological localization method that performs well in a variety of man-made indoor *and* outdoor environments. Our localization system uses color vision, works in real-time, and can easily be trained in new environments. The next section outlines the basic approach of the algorithm and describes the motivation for the use of a panoramic color vision system. Section three shows the similarities between the problems of place recognition and image retrieval, and describes the multiple advantages of histograms. Section four explains how nearest-neighbor learning can be applied to topological localization. Section five describes in detail the place recognition algorithm. Section six presents experimental results from four different environments. Section seven discusses possible future improvements. Section eight briefly describes related work in topological localization, and the paper ends with a conclusion in section nine.

2. TOPOLOGICAL APPROACH

Unlike the majority of current localization research, our algorithm uses a topological representation of the environment. In a topological approach, maps are represented by adjacency graphs. Nodes of the graph represent locations, while arcs represent the adjacency

relationships between the locations. Figure 1 shows an example of a topological map for an apartment with eight rooms. In the current implementation, adjacency maps are manually created with a map editor. Creating such maps takes very little time, as no geometric information is required, i.e., there is no need to measure the dimensions of the environment. Although the map's geometric layout is ignored by the localization system, it is useful for visualization and debugging.

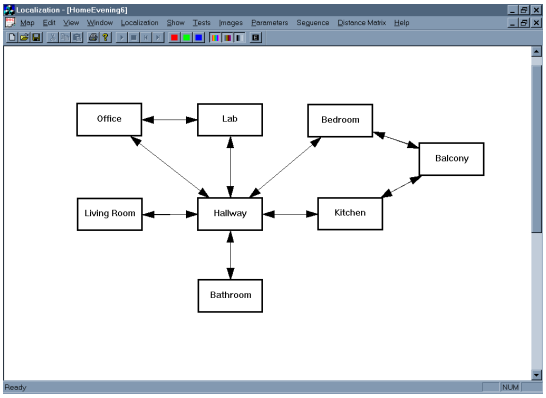

Figure 1: Adjacency map of an apartment.

A robust localization system requires an extrinsic sensor that provides rich information in order to allow the system to reliably distinguish between adjacent locations. For this reason, we use a passive color vision camera as our extrinsic sensor. Because many places can easily be distinguished by their color appearance, we expect that color images provide sufficient information without the need for range data from additional sensors such as stereo, sonars, or a laser rangefinder.

Our vision-based localization system must be trained before it can be used in an environment. During training, representative images are captured from the environment and associated with the corresponding locations. During operation, the input image is compared to the map's reference images. The location whose reference image best matches the input image is then considered to be the currently visible location. Therefore, the goal of the training stage is to assign a set of representative images to each location. The current training process is straightforward and consists of two steps. In the first step, the robot is simply driven through the environment. During this process, the localization system captures a sequence of images at a rate of 1 Hz and saves them onto the harddisk. In the second step, the user labels the captured images with the names of the corresponding locations. Due to the sequential nature of the image sequence and the easy-to-use interface of our sequence editor, the image labeling process takes only a few minutes.

The images captured during the training stage need to be representative of the environment. In order to keep the number of necessary reference images low, we use an *Omnicam*, which provides panoramic images of the environment [15]. An example image is shown in Figure 2. The most important advantage of the Omnicam system over a system with a regular lens is that it provides rotational invariance in the field of view, i.e., the same area of the environment is captured independent of the camera's orientation. Another advantage is the large field of view, which further reduces the number of images necessary to represent a location. The large field of view also makes the system robust to small changes in the environment.

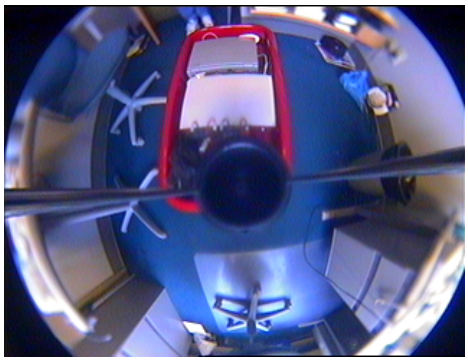

Figure 2: Example image acquired by Omnicam system.

3. IMAGE RETRIEVAL

The key element of our topological localization method is the *place recognition* module. Assuming that the previously learned set of images is representative of the locations, the goal of the place recognition module is to determine the reference image that is most similar in appearance to the current input image. A similar problem is faced by researchers in the field of image retrieval, which has been very active during the last ten years due to the creation of large image databases [7]. The goal of an image retrieval system is to find images in the database that are similar in appearance to an input query.

Due to the similarity of the underlying problem, our localization approach greatly benefits from existing image matching techniques developed in the field of image retrieval. Although the heart of the problem is similar, the two problems have some interesting differences. Image retrieval systems need to compare a query image with the images of an entire database, which can exceed millions of images. The place recognition system needs to compare the current image only with images taken at the currently believed location and its immediate neighbors. The adjacency relationship usually limits the number of reference images to less than one hundred, which results in a fast matching process.

However, while a few bad matches are not a problem in image retrieval, a single bad match could lead the robot localization system to become lost. Confident bad matches must therefore be strictly avoided for the localization task.

Most image retrieval systems rely on histograms for the matching process. Histograms have many attractive features [19]. First, histograms provide a compact representation of an image, and thus require less memory space. In our system, a raw input image requires 231 kilobytes, while a single histogram only uses 1 kilobyte of memory. This reduction of memory requirements also makes the comparison of two histograms much faster than the comparison of two raw images. Second, image histograms are invariant to rotation of the image around the optical axis, which is especially attractive when combined with a panoramic vision system. Due to the rotational invariance, images taken at a certain position with different camera orientations have almost identical histograms. Therefore, a single image acquired at a certain position is representative for all images acquired with different orientations at the same position. This property greatly reduces the number of reference images, and thus further improves the speed of the place recognition system, because the input image needs to be compared to fewer reference images. Third, histograms are not very sensitive to small translations, which further reduces the number of reference images. In summary, the combined use of histograms and the Omnicam results in a system that requires little memory and performs in real-time.

4. NEAREST-NEIGHBOR LEARNING

Our localization method is based on the concept of nearest-neighbor learning. In the context of topological localization, the task of the place recognition module is to determine the reference image that best matches the input image. As our algorithm uses several color bands, the nearest-neighbor algorithm is applied to each color band separately. The results of the nearest-neighbor algorithms are then combined with a simple scheme based on unanimous voting.

Due to the previously described advantages of histograms over raw images, our nearest-neighbor algorithm operates in histogram space instead of image space. The histograms for the reference images are built off-line, while the histograms for the input image are built on-line. In our current implementation, the color images are first transformed into the HLS (hue, luminance, saturation) and the normalized RGB color spaces. One-dimensional histograms are then built for each individual band. To reduce the effects of noise, all histograms are low-pass filtered with an average filter.

Each image is represented by three one-dimensional histograms for the HLS color bands. Depending on which color space performs better in a given environment, each image is also represented by three one-dimensional histograms for either the RGB or the normalized RGB color bands. Each image is thus represented by six one-dimensional histograms. At the end of the off-line initialization, a location with n reference images has n reference histograms for each of the six color bands.

5. PLACE RECOGNITION

During operation, an image is acquired and then classified by the place recognition function, which consists of the following four steps:

1. Build color histograms of the acquired input image.
2. For each color band, determine the best match for each candidate location.
3. For each color band, determine location vote and confidence.
4. Classify image based on unanimous voting.

The first step is identical to the off-line initialization of the reference histograms. It transforms the input image into six one-dimensional histograms.

In the second step, the histograms are matched to the reference histograms. Due to the adjacency map, the search can be limited to a small number of locations. Only the currently believed location and its immediate neighbors need to be considered. For each of these candidate locations and for each of the six color bands, the algorithm then determines the reference histogram that best matches the input histogram. In order to measure how well two histograms match, we need a histogram similarity or dissimilarity metric. A good overview of such metrics is given by Rubner et al. [18].

To determine which metric best suits our application, we have investigated a number of bin-by-bin measures: L_1 distance, L_2 distance, histogram intersection, Jeffrey divergence, and χ^2 statistics. We have also tested a number of cross-bin measures: quadratic-form distance, match distance, and Kolmogorov-Smirnov distance. Briefly summarizing our tests, we consistently obtained the best results with the Jeffrey divergence, closely followed by the χ^2 statistics. We have thus chosen to use the Jeffrey divergence as our distance metric. Given two histograms H and K, with h_i and k_i denoting the histogram entries, the Jeffrey divergence is defined as follows:

$$d(H,K) = \sum_i \left(h_i \log \frac{2h_i}{h_i + k_i} + k_i \log \frac{2k_i}{h_i + k_i} \right) \qquad (1)$$

For each candidate location and each color band, the distance between the input histogram and the reference histograms is thus computed using the Jeffrey divergence. For each candidate location, the minimum distance is then determined for each band.

In the third step, each color band individually decides for which candidate location it votes. Each band simply votes for the candidate location with the smallest minimum matching distance. Each color band b also computes a confidence measure c_b of its vote. The confidence measure c_b is computed as follows:

$$c_b = 1 - \frac{d_m}{\min\limits_{i \neq m}\{d_i\}} \tag{2}$$

with: d_m: minimum matching distance of all candidate locations.
d_i: minimum matching distance of all other candidate locations.

Confidence values range between 0 and 1. The higher the confidence value, the more confident a band is. The confidence value is high if a candidate location matches the input image much better than any other candidate location. The confidence value is low if the second best candidate location matches the input image similarly well as the best candidate location. If no candidate location matches the input image well, a high confidence value is unlikely. Thus, if a color band is unable to reliably classify an input image, its uncertainty is reflected by a low confidence value.

In the fourth step, the six votes from the color bands are combined to classify the input image. The localization system output consists of the image classification, a confidence value, and a confidence type. The confidence type can either be *confident*, *uncertain*, or *confused*. Only confident classifications are used by the system to keep track of the location.

Each color band b has its own confidence threshold τ_b. In order to cast a confident vote, the band's confidence value must be above its confidence threshold. The votes from the confident color bands are then combined with a simple scheme based on unanimous voting. There are three possible cases:

i) The confident bands are unanimous and vote for the same location. In this case, the total confidence value is computed as defined in (3). If the total confidence is above the action threshold, then the system is *confident* about its classification. If the confidence value is below the action threshold, then the system is *uncertain*. The action threshold is currently set at 0.1 for all transitions. This parameter could be set individually for each transition, but so far this was unnecessary, as we achieved good results with a global threshold value.

$$c_{total} = \sum_{b=1}^{6} (c_b - \tau_b) \tag{3}$$

ii) The confident bands are not unanimous and vote for different locations. In this case, the classification status is *confused*, and the confidence value is equal to zero. Not surprisingly, confused votes happen most often near transitions.

iii) None of the bands is sufficiently confident to cast a vote. In this case, the classification status is *uncertain*, and the confidence value is equal to zero.

The higher the number of confident classifications the better. If there are not enough confident classifications, the system might miss a transition and could become lost. The system also becomes lost if an incorrect confident classification occurs. If the system's status is confused or uncertain, then the system does not classify the input image. For its internal belief, the system assumes that it is still at the same location as at the last confident classification.

It is important to note that our current system does not rely on odometry information and odometry uncertainty models. While odometry is very useful for short-distance navigation, we chose not to use odometry information for the task of localization at this time. Therefore, errors in odometry do not affect our system at all.

6. EXPERIMENTAL RESULTS

For experimental verification, we implemented our passive localization system on an all-terrain Radio Flyer wagon. The wagon is pulled manually, a process that is much easier than controlling an actively driven vehicle. The wagon carries a laptop computer with a Pentium II processor clocked at 333 MHz, a Hitachi KPD-50 color CCD camera with an Omnicam lens system, and batteries to power the camera. The camera is connected to the laptop with a PCMCIA framegrabber from MRT. The entire system is shown in Figure 3.

We have used the wagon as an experimental platform to test our localization system in a variety of environments. Table 1 summarizes the performance of our localization system in four different environments, three indoors and one outdoors. All tests were performed in unmodified environments. Each environment was evaluated twice with a cross-validation test. A cross-validation test is based on two different sequences of images, acquired at different times in the same

environment. One of the sequences is used as training data, while the other sequence is used as test data. In all the tests, the initial location is assumed to be known at the start. The localization system's task consists of classifying the test images, and keeping track of the vehicle's position. Although the tests are performed off-line for a better analysis, the results would be identical in an on-line test.

Tests #1 and #2 were performed in an apartment with eight rooms. Tests #3 and #4 were performed on the first floor of Smith Hall at the Robotics Institute. Tests #5 and #6 were performed on the second floor of Smith Hall. Tests #7 and #8 were performed outdoors on the Carnegie Mellon University campus, on a route leading from Doherty Hall to the Football Field, a distance of about 250 meters. The adjacency map for the first two tests was used as an example in Figure 1, while the adjacency maps for the other six tests are shown in Figure 4. The size of a location can vary substantially, ranging from 1-2 m^2 for a small indoor location to more than 100 m^2 for a large outdoor location.

For each test, the voting thresholds for the individual bands were optimized by setting their values equal to the confidence values of the highest incorrect votes. The action threshold was set to 0.1 for all transitions. All tests used the HLS color space, and either the RGB or the normalized RGB color space.

Figure 3: Experimental platform.

All eight tests can be classified as being very successful, as our localization system was able to continually track the location of the wagon, without ever missing a transition. The percentages of correct confident classifications range from 87.5% to 97.7%. Most importantly, none of the confident classifications were incorrect. The total number of images that were classified was 1307. The number of images ranges between 131 and 231 for each of the individual test sequences.

Due to the topological representation where locations are defined as rooms, our algorithm does not need a fast sampling rate to keep track of the robot's location. We expect that an update rate of about 2 Hz

Test Number	#1	#2	#3	#4	#5	#6	#7	#8
Confident Correct	**97.7 %**	**97.7 %**	**93.3 %**	**87.5 %**	**93.8 %**	**97.4 %**	**96.4 %**	**90.7 %**
Confident Wrong	**0.0 %**	**0.0 %**	**0.0 %**	**0.0 %**	**0.0 %**	**0.0 %**	**0.0 %**	**0.0 %**
Uncertain Location	0.8 %	0.0 %	3.4 %	6.5 %	2.5 %	1.7 %	1.4 %	6.0 %
Uncertain Transition	0.0 %	0.0 %	2.2 %	5.4 %	1.2 %	0.0 %	0.7 %	2.7 %
Confused Location	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
Confused Transition	1.5 %	2.3 %	1.1 %	0.5 %	2.5 %	0.9 %	1.4 %	0.7 %
Average Confidence	1.753	2.716	1.149	0.666	1.105	1.528	0.964	1.330

Table 1: Results of cross-validation tests

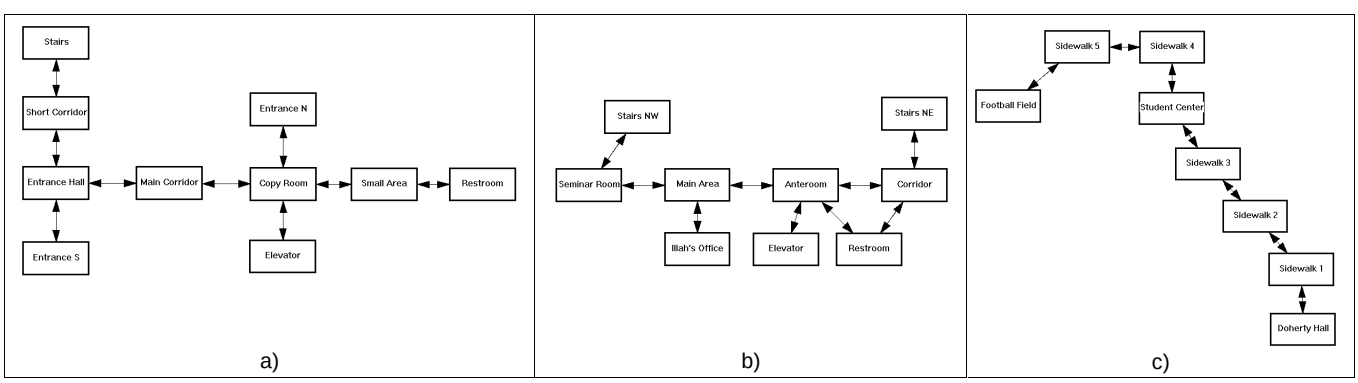

Figure 4: Adjacency maps of experiments: a) first floor of Smith Hall, b) second floor of Smith Hall, c) campus route

should be sufficient for most environments. Because our system applies fast image retrieval methods to rather small databases, the system easily achieves the required update rate of 2 Hz. In fact, with the current equipment, the algorithm needs less than 250 ms to compare an input image to 100 reference images, thus leaving a good amount of processing time for other modules.

Our current system is robust to changes that represent less than a few hundred pixels in an image. The more image pixels are modified by the change in the environment, the more affected our system is.

7. FUTURE WORK

Although the achieved results are promising, we have many ideas of how to further improve the current system. A straightforward task consists of investigating other promising color spaces and determining how higher dimensional histograms perform. In addition to color, we could also experiment with texture as an attribute.

Our localization system differentiates between locations based on their appearance. As object appearance depends heavily on illumination conditions, our system is very sensitive to changes in illumination. A modified algorithm that is capable of compensating for varying illumination would be very beneficial to the system in many environments. We expect that placing a reference chart into the camera's visual field will help us implement a relatively simple color constancy algorithm.

The ideal system will be capable of creating topological maps on its own. Automatic generation of maps would have many advantages. First, such a system would be very user-friendly. To train the system, it would only need to be led through the environment once. This training run could be performed under human control or by an autonomous exploration system. Second, allowing the system to define its own locations would permit it to achieve an even better performance. In the current system, where locations are defined by the user, an inappropriate choice can have negative effects on the system's performance. Third, automatic map generation would make it practical for each band to have its own adjacency map. Because different bands are sensitive to different features in the environment, it would be optimal to have individual maps for each band. Because creating individual maps by hand is not practical, automatic map generation is a requirement for the use of individual maps. Based on additional experiments, we believe that automatic map generation could indeed be feasible. This would not only make the system extremely user-friendly, but it would also greatly improve its performance by allowing the system to define its own locations and to create individual maps for the different bands.

8. RELATED WORK

The concept of topological maps for mobile robots was pioneered by Kuipers in 1978 [11]. Since then, several approaches of topological localization have been developed. The remainder of this section briefly describes a number of topological approaches based on how they define a place and what kind of sensor they use to perform the critical task of place recognition.

In Kuipers and Byun's system, a place is a well-defined position in space that maximizes a certain criterion in terms of sensor input [12]. Their system makes no assumptions about the structure of the environment. However, their system was only evaluated in simulations.

In 1992, Kortenkamp et al. defined the concept of *gateways*, places that mark the transition between two adjacent spaces in the environment [10]. In experiments with a mobile robot equipped with both sonars and vision, their system was able to recognize seven out of seven places. In a similar approach, Fu et al. proposed a system that detects corridor intersections based on sonar data [6]. However, they only tested their system in simulations. Kunz et al. developed Inductobeast, which defines places as intersections and transitions between hallways and open areas [13]. Inductobeast uses sonars to perform wall and corridor following, and then builds topological maps from encoder history. While these gateway approaches performed well in both simulations and real experiments, they are all restricted to highly structured and orthogonal environments.

Mataric's robot Toto uses sonars and a flux-gate compass to detect straight walls and corridors, which are used as landmarks [14]. The regions near these landmarks define the places of Toto's purely topological map. Although Toto is not restricted to orthogonal environments, it requires a large proportion of straight walls in its environment.

Courtney and Jain use a hybrid representation of the environment [5]. They represent the workspace as a set of grid-based maps interconnected via topological relations. Their system uses sonars, lateral motion vision, and infrared sensors to perform recognition of rooms and doors. To evaluate their system, they collected ten maps from each room. In leave-one-out validation tests, they obtained high recognition rates. However, their rate of misclassifications is not negligible. Moreover, in order to recognize a room, the system must first create a map of the room, which is a time-consuming process.

Taylor and Kriegman also use a hybrid representation of the environment. Local geometric maps are connected by a global topological map [21]. Each node defines the visibility region of an artificial landmark. Landmarks consist of bar-codes placed at a well-defined height.

Similar to the geometric localization methods described in the introduction, none of these topological methods claim to perform well both indoors and outdoors. In order to achieve reliable place recognition in a variety of environments, we strongly believe that an information-rich sensor like color vision is necessary. In prior work, we evaluated the potential of color-based localization with a vision-based transition detector [17]. Due to the promising preliminary results, we continued this research direction. While our previous method was based on a Naïve Bayes classifier, the current method is based on nearest-neighbor learning. The main advantage of the current method is that it does not require a meta-level training process to adjust individual transition thresholds.

9. CONCLUSION

This paper presented a new method for topological localization. The system has been tested in four different environments, indoors as well as outdoors. In all four environments, the system successfully tracked the mobile robot's position. The system correctly classified between 87.5% and 97.7% of the input color images. For the remaining images, the system was either momentarily confused or uncertain, but never classified an image incorrectly.

ACKNOWLEDGEMENT

Thanks to Martial Hebert at the Robotics Institute for lending us the Omnicam used in this project.

REFERENCES

[1] Abe, Y., Shikano, M., Fukuda, T., Arai, F., and Tanaka, Y., "Vision Based Navigation System for Autonomous Mobile Robot with Global Matching", *IEEE International Conference on Robotics and Automation*, May 1999, pp. 1299-1304.

[2] Arras, K.O., and Tomatis, N., "Improving Robustness and Precision in Mobile Robot Localization by Using Laser Range Finding and Monocular Vision", *3rd European Workshop on Advanced Mobile Robots*, September 1999.

[3] Asensio, J.R., Montiel, J.M.M., and Montano, L., "Goal Directed Reactive Robot Navigation with Relocation Using Laser and Vision", *IEEE International Conference on Robotics and Automation*, May 1999, pp. 2905-2910.

[4] Borenstein, J., Everet, H.R., and Feng, L., "Navigating Mobile Robots", A.K. Peters, Wellesley, MA, 1996.

[5] Courtney, J.D., and Jain, A.K., "Mobile robot localization via classification of multisensor maps", *IEEE International Conference on Robotics and Automation*, May 1994, pp. 1672-1678.

[6] Fu, D.D., Hammond, K.J., and Swain, M.J., "Navigation for everyday life", *TR-96-03*, University of Chicago, 1996.

[7] Idris, F., and Panchanathan, S., "Review of Image and Video Indexing Techniques", *Journal of Visual Communication and Image Representation*, Vol. 8, No. 2, June 1997, pp. 146-166.

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

[9] Jensfelt, P., and Christensen, H.I., "Laser Based Pose Tracking", *IEEE International Conference on Robotics and Automation*, May 1999, pp. 2994-3000.

[10] Kortenkamp, D., Baker, L.D., and Weymouth, T., "Using Gateways to Build a Route Map", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1992, pp. 2209-2214.

[11] Kuipers, B.J., "Modeling spatial knowledge", *Cognitive Science*, Vol. 2, No. 2, 1978, pp. 129-153.

[12] Kuipers, B.J., and Byun, Y.-T., "A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations", *Robotics and Autonomous Systems*, Vol. 8, 1981, pp. 47-63.

[13] Kunz, C., Willeke, T., and Nourbakhsh, I.R., "Automatic Mapping of Dynamic Office Environments", *Autonomous Robots*, Vol. 7, No. 2, September 1999, pp. 131-142.

[14] Mataric, M.J., "Integration of Representation Into Goal-Driven Behavior-Based Robots", *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 3, June 1992, pp. 304-312.

[15] Nayar, S.K., "Catadioptric Omnidirectional Camera", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1997, pp. 482-488.

[16] Pérez, J.A., Castellanos, J.A., Montiel, J.M.M., Neira, J., and Tardós, J.D., "Continuous Mobile Robot Localization: Vision vs. Laser", *IEEE International Conference on Robotics and Automation*, May 1999, pp. 2917-2923.

[17] Radhakrishnan, D., and Nourbakhsh, I.R., "Topological Robot Localization by Training a Vision-based Transition Detector", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 1999, pp. 468-473.

[18] Rubner, Y., Tomasi, C., and Guibas, L.J., "The Earth Mover's Distance as a Metric for Image Retrieval", *STAN-CS-TN-98-86*, Stanford University, 1998.

[19] Swain, M.J., and Ballard, D.H., "Color Indexing", *International Journal of Computer Vision*, Vol. 7, No. 1, 1991, pp. 11-32.

[20] Takeuchi, Y., and Hebert, M., "Evaluation of Image-Based Landmark Recognition Techniques", *CMU-RI-TR-98-20*, Carnegie Mellon University, 1998.

[21] Taylor, C.J., and Kriegman, D.J., "Vision-Based Motion Planning and Exploration Algorithms for Mobile Robots", *IEEE Transactions on Robotics and Automation*, Vol. 14, No. 3, June 1998, pp. 417-426.

[22] Thrun, S., "Finding Landmarks for Mobile Robot Navigation", *IEEE International Conference on Robotics and Automation*, May 1998, pp. 958-963.