A Hybrid Approach to Topological Mobile Robot Localization*

Paul Blaer and Peter K. Allen
Computer Science Department
Columbia University
New York, NY 10027
Email: {pblaer, allen}@cs.columbia.edu

***Abstract*— We present a hybrid method for localizing a mobile robot in a complex environment. The method combines the use of multiresolution histograms with a signal strength analysis of existing wireless networks. We tested this localization procedure on the campus of Columbia University with our mobile robot, the Autonomous Vehicle for Exploration and Navigation of Urban Environments. Our results indicate that localization accuracy is significantly improved when five levels of resolution are used instead of one in color histogramming. We also find that incorporating wireless signal strengths into the method further improves reliability and helps to resolve ambiguities which arise when different regions have similar visual appearances.**

I. Introduction

Localizing a mobile robot in a complex environment is a complicated and difficult problem. Localization can be accomplished through either geometric or topological methods. In this paper we present a fast and robust method for topological localization using a combination of techniques. An analysis of multiresolution color histograms is combined with a signal strength analysis of an existing wireless ethernet network to provide an accurate estimate of the region in which the robot is currently located.

This topological localization is part of our Autonomous Vehicle for Exploration and Navigation of Urban Environments, *AVENUE* system [1]. The ultimate goal of the AVENUE system is to autonomously model an urban site. The system plans a path to a desired viewpoint, navigates the mobile robot to that viewpoint, acquires images and three-dimensional range scans of the building, and then plans for the next viewpoint. Topological localization, however, is not sufficient for geometrically localizing the robot. Our approach is to use a coarse-fine localization in which the topological localization described in this paper feeds a very precise vision-based system which uses prominent linear features on buildings to determine the robot's exact location. The vision-based fine localization is described in [2], [3].

The topological localization builds upon our earlier work [4] in which the system attempts to match omnidirectional images acquired from the robot to a pre-existing database of images using color histograms. This method is fast and rotation invariant, but suffers somewhat from sensitivity to outdoor lighting changes. We have decreased this sensitivity by incorporating multiresolution histograms. Such histograms, unlike normal histograms, encode some spatial information in addition to color composition [5].

Even with the improved histogram matching, the system still has some difficulty distinguishing between very similar looking topological regions. A secondary discriminator is therefore necessary. Drawing from our other previous work [6], we have chosen to utilize information from wireless ethernet networks, which are becoming very common in urban environments. A profile of the signal strengths of nearby access points is constructed and then used for matching with an existing database.

Our paper is organized as follows. In the next section, we indicate previous and related work. We then describe in section III our robot's equipment. In section IV, we detail the process of constructing the combined multiresolution histograms and wireless signal strength database. We then describe the matching procedure. In section V we discuss the results of our localization system during a test run on the Columbia University campus. In the concluding remarks of section VI, we summarize our results and suggest additional possible uses of our method for the AVENUE project.

II. Related Work

Topological maps for general navigation were originally introduced for use in mobile robotics by [7]. Many localization methods involve the use of computer vision to detect the transition between regions [8]. Recently a number of researchers have used omnidirectional imaging systems [9] to perform robot localization. Cassinis et al. [10] used omnidirectional imaging for self-localization, but they relied on artificially colored landmarks in the scene. Winters et al. [11] also studied a number of robot navigation techniques utilizing omnidirectional vision.

The vision component of our current work most closely resembles that of Ulrich and Nourbakhsh [12], who studied outdoor topological localization of a mobile robot using color histograms of omnidirectional images. The primary distinction between their work and the vision part of our work is our use of multiresolution histograms. Sablak and Boult [13] also studied the use of just the histogram peaks from omnidirectional images for indoor room recognition. Gross et al. [14] used the Monte Carlo Localization method on omnidirectional images with a reference-based method to control for variance

Fig. 1. The ATRV-2 Based AVENUE Mobile Robot (left). A sample laser scan taken with the AVENUE system on the Columbia University campus (right). The hole at the center of the scan is where the scanner was positioned.

in the luminance and color in the scene over changing lighting conditions.

The concept of using color histograms as a method of matching two images was pioneered by Swain and Ballard [15]. A number of different metrics for finding the distance between histograms have been explored [16]–[18]. Hadjidemetriou et al. suggest the use of multiresolution histograms in texture classification and recognition in [5].

The use of existing 802.11b wireless network signals as a means of locating a user was originally presented in Microsoft Research's RADAR project [19]. The Microsoft group collected the signal data manually in an indoor environment and then used this information for estimating the position of a user at a later time. Other groups have also made use of manually-obtained 802.11b signals for indoor localization [20]. We have extended the work of these groups by having our mobile robot autonomously construct the database, while covering a much larger outdoor urban environment.

There have also been a number of systems [21] based on the characteristics of cellular signals and designed for geolocating cellular telephone users in outdoor environments. In addition, there have been attempts to use RF based networks, as in the Daedalus project [22], to localize a user in an outdoor area.

Other approaches include simultaneous localization and map building [23]–[26], probabilistic approaches [26] and [27], and Monte Carlo localization [28].

III. The Platform

Our mobile robot, AVENUE, has as its base unit the ATRV-2 model (see Fig. 1) manufactured by Real World Interfaces, now part of iRobot. The base unit has an on-board computer, odometry from wheel encoders, and a set of sonar units located around the perimeter of the robot. In addition to these base features, we have added additional sensors including a differential GPS unit, a laser range scanner, a camera mounted on a pan-tilt unit, an omnidirectional camera, a digital compass, and two 802.11b wireless network cards. Figure 1 also shows a sample laser scan taken with this system on the Columbia University campus.

The sensor we use to acquire color histograms is an omnidirectional camera manufactured by Remote Reality (see top of Fig. 2). It is important to note that the ground plane around almost all of the buildings in our environment has the same brick pattern. As a result, aiming the omni–camera upward (that is, with the mirror facing down at the ground) is not an option, because all of the regions would look essentially the same. We must therefore aim the camera downward (mirror upward) in order to obtain a good view of the upper portions and tops of all buildings (see bottom of Fig. 2). In addition, the camera is mounted on top of the robot's superstructure so that as little of the robot as possible is in the camera's field of view.

Communication with the networks' base stations is accomplished through an omnidirectional antenna which is mounted on the highest point of the robot and which is connected to the PCMCIA wireless network card in the on-board computer. Software on the robot's on-board computer polls this wireless card and returns a list of access points that are in range together with the strengths of the signals measured in dBm. This on-board computer also handles the image acquisition and image processing at the same time.

Our experiments were run in an outdoor environment, specifically the northern half of the Morningside Heights Campus of Columbia University (see Fig. 4). There is an

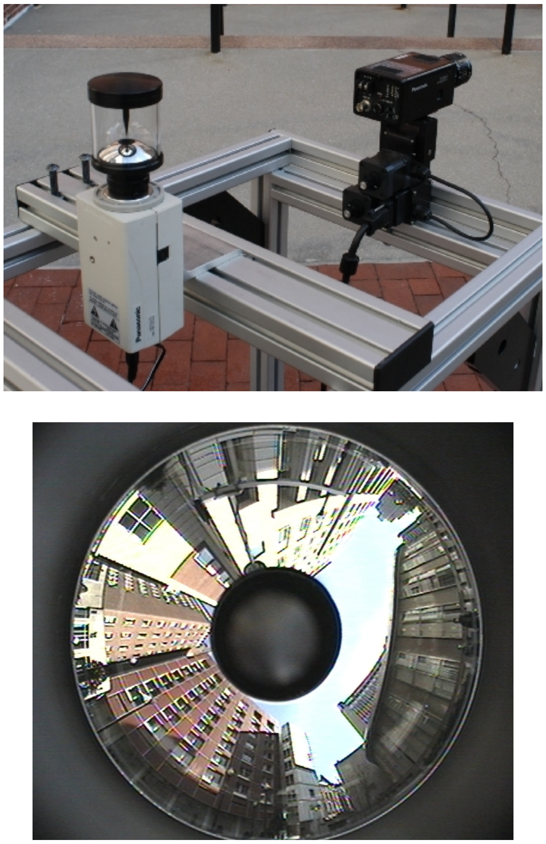

Fig. 2. Our robot's omnicamera (top) and a typical image from that camera (bottom).

extensive wireless network already installed on the campus, so we simply used the existing infrastructure. Our method did not rely on any knowledge of the exact location of the access points.

IV. Localization System

A. Multiresolution Histograms

Our method involves constructing a database of reference images taken throughout the various known regions that the robot will be exploring at a later time. Each reference image is then reduced to three multiresolution histograms, using the red, green, and blue color bands separately. We compute a multiresolution histogram for each image at full resolution, as well as at $1/2$, $1/4$, $1/8$, and $1/16$ resolutions. Down sampling of the image is accomplished by first convolving the original image with a 5x5 Gaussian kernel to blur the image. Then the blurred image is sub-sampled down to the lower resolution. The resulting multiresolution histogram is a set of five 256-bucket sub-histograms. Each bucket contains the number of pixels in the image at a specific intensity. Since the lower resolution sub-histograms have fewer total number of pixels across all buckets, the 5 sub-histograms for a given reference image are normalized such that the sum across all 256 buckets is the same for each resolution level. This prevents the matching metric from being dominated by the highest resolution sub-histogram.

When the robot is exploring the same database regions at a later time; it will take an image, convert it to a set of three multiresolution histograms, and attempt to match those histograms against the existing database. The database itself is divided into a set of characteristic regions. The goal is to determine in which specific physical region the robot is currently located.

The images themselves, both for the database and for the later unknowns, are taken with the robot's on-board omnicamera. The images are taken at a resolution of 640x480 with a color depth of 3 bytes per pixel. We use an omnidirectional camera instead of a standard camera because it allows our method to be rotation invariant. Images taken from the same location but with a different orientation will differ only by a simple rotation. Since the histogram only takes into account the colors of the pixels and not their position within the image, two histograms taken from the same location but from a different orientation will essentially be the same.

Relying solely on color has its drawbacks since lighting conditions change over time, especially outdoors, and can cause large variations of color in a scene. As a result, we needed a method that would implicitly consider some spatial information in addition to color. Multiresolution histograms provided us with this additional information. At every level of resolution we continued to look only at color histograms, so that rotation invariance was maintained.

The process of blurring our images changes the histograms significantly. Because the blurring combines the effect of adjacent pixels, the new histogram is dependent on the physical location of each pixel. As a result, two dissimilar scenes that would normally have similar histograms could now have very different histograms at lower resolution levels. As an extreme example (suggested in [5]), consider one image that consists of alternating pixels of intensity 0 and 255 and a second image that has half of its pixels in one solid block of intensity 0 and the other half in a solid block of intensity 255. The histograms of these two images would be identical. However, if you blur both images, the alternating pixels would average to gray whereas the solid blocks would remain mostly white and black with only the boundary between them becoming gray (see Fig. 3).

The rotation invariance of the histograms allows us to reduce the size of our database considerably, because only one image at a given physical location is needed to get a complete picture of the surrounding area. In addition, by using multiresolution histograms, we embed some information about the geometry of the scene into the histogram, which can help overcome large variations in color caused by different lighting conditions.

B. Wireless Signal Strengths

As the AVENUE robot travels through its environment, a program running on the robot accesses the primary wireless card and returns a vector of information. For each access point that the robot can detect, we record its unique hardware address and the signal strength (measured in dBm).

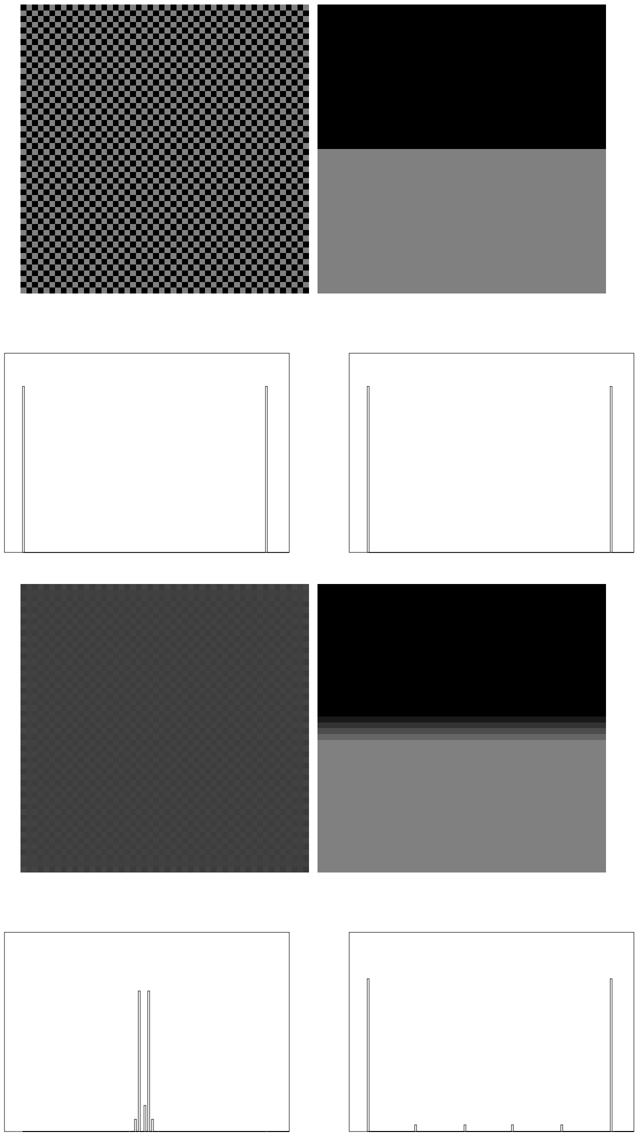

Fig. 3. An example illustrating the usefulness of multiresolution histograms. The top row shows 2 very different scenes at full resolution, with their corresponding histograms in the second row. The third row shows these scenes blurred to $1/2$ resolution, with their corresponding histograms in the last row.

Ultimately we will have to compare the wireless signal strength vectors in a fast and meaningful way. This is difficult if we only store the access points that are visible from a given location because we never see all of them at once. To deal with cases in which there are access points that appear in one vector but not in the other, we assume that the other vector has that access point in its list but with a strength of zero.

C. Constructing the Database

The robot collects simultaneous readings from the omnicamera and the wireless card. An entry in the database is created consisting of the 3 computed multiresolution histograms of the image along with a record of the access points visible and their corresponding signal strengths.

Driving the robot straight through the center of each region does not give us enough variation in the database to identify all potential positions reliably. The proximity of a building or other structures has a large effect on the images that the robot acquires. We therefore build up a more comprehensive database by having the robot zigzag through the test environment. This allows us to obtain representative images of a given region from a variety of positions within that region. Although this does increase the size of the database, it is not a major problem because the database is stored as a series of histograms, not images, and the comparison between each of the 256-bucket histograms is extremely fast.

D. Matching against the Database

At this point, our software has a collection of records grouped together according to their geographical region. We now use this database for matching an unknown location's reading with one of the database entries.

Each entry in the database has a total of 15 histograms. There are 3 multiresolution histograms, one for each separate color band (r, g, and b). Each of these multiresolution histograms has a separate sub-histogram for 5 resolution levels (r_k, g_k, and b_k, for $k = 1, ..., 5$). There are 256 buckets in each of these 15 histograms.

To compare the histograms of an unknown location X with those of a database entry E, we use the following histogram difference metric $hd(X, E)$. Let $r_{k,i}(X \textit{ or } E)$, $g_{k,i}(X \textit{ or } E)$, and $b_{k,i}(X \textit{ or } E)$ denote the number of pixels in the i^{th} bucket ($i = 1, ..., 256$) of the red, green, and blue sub-histograms of X or E at the k^{th} resolution ($k = 1, ..., 5$). Then,

$$hd(X, E) = \sum_{k=1}^{5} \sum_{i=1}^{256} (R_{k,i} + G_{k,i} + B_{k,i})$$

where

$$\begin{aligned} R_{k,i} &= |r_{k,i}(X) - r_{k,i}(E)| \\ G_{k,i} &= |g_{k,i}(X) - g_{k,i}(E)| \\ B_{k,i} &= |b_{k,i}(X) - b_{k,i}(E)| \, . \end{aligned}$$

Because histograms at each resolution level have been normalized such that the total number of pixels across all buckets is the same, each of these 15 differences have been weighted equally. We then renormalize this difference metric to a value between 0 and 1 by dividing by its maximum possible value.

Each database entry also has a list of visible access points p along with a signal strength measure for each of them. The signal strengths are approximately within the range of -80dBm to -20dbm. We renormalize these strengths to be in a quality range between 1 and 50. By going through all entries in the database, we can find all the visible access points throughout our test region. For each entry in the database, we explicitly add a strength quality of 0 for each unobserved access point, in addition to the visible access point readings. As a result, each entry will contain a list of N

signal strengths, where N is the total number of access points observed throughout our test region.

To compare the access point strengths of an unknown location X with those of a database entry E, we use the following access point difference metric $apd(X,E)$. Let $p_n(X \ or \ E)$ denote the renormalized strength of the n^{th} access point ($n = 1, ..., N$) as observed at X or E. Then,

$$apd(X,E) = \sum_{n=1}^{N} |p_n(X) - p_n(E)| \, .$$

As with hd, we renormalize this difference metric to a value between 0 and 1 by dividing by its maximum possible value.

In order to combine the histograms and signal strengths in the most effective way, we used a weighted sum of the two renormalized difference metrics defined above to give our final difference metric:

$$D = (w_h)(hd) + (w_a)(apd)$$

with $w_h + w_a = 1$. We tried several different possible weightings and found the best overall results with $w_h = 0.8$ and $w_a = 0.2$.

This weighted sum D is used as our final measure of the difference between an unknown location's reading and a database entry. For an unknown location, D is computed for all entries in the database, and the entry that has the smallest value of D is found. We choose the region of that database entry as the region of the unknown. Furthermore, the actual value of D could be used as a measure of our confidence in the estimated location.

V. Experimental Results

We divided our outdoor test area into 13 regions according to which buildings were most prominent. Our goal was to get the robot to localize itself to one of these regions. The regions spanned the northern half of the Columbia campus (see Fig. 4). Approximately 40 combined readings (both image and wireless signal strength) were taken in each of the regions. Histograms were computed and the database was constructed. Another set of readings was taken on a different day to be used as unknowns to compare against the database. In all, 100 "unknown" readings were taken per region.

To judge the effects of multiresolution histogramming and the additional wireless ethernet data, matching was done in four ways. In the first, only the full-resolution color histograms of the original image were used for matching. In the second test, only the multiresolution histograms were used. For the third test, only wireless readings were used. Finally, we used the entire combination of multiresolution histograms and wireless data described in the previous section. For each of the four methods, the "unknown" readings were compared against the database. A success was recorded if the method classified the "unknown" reading in the correct region. Table I contains the success rates for each method in each region.

In the initial test, just using simple histograms, we had an overall success rate of 65%. The use of multiresolution histograms increased that success rate to 83%. When using the multiresolution histograms, most of the individual regions had consistent success rates; however, of particular note were regions 4 and 6. These two regions had substantially lower success rates (65% and 62%, respectively). The two regions are physically very similar as they are on opposite sides of a mostly symmetric building. The system often confused these two regions. Testing the wireless data alone gave an overall success rate of 72%. Problems arose with the wireless data in regions where one access point was mounted in a central location and significantly covered several of our regions, making them harder to distinguish. It would have been possible to redefine our regions according to the influence of a particular access point, but that would have had very negative effects on the vision-based method.

TABLE I

Success Rates of the Localization Experiments

Method	A	B	C	D
Region 1	85%	93%	73%	94%
2	62%	91%	70%	92%
3	70%	95%	74%	95%
4	48%	65%	66%	83%
5	60%	77%	68%	90%
6	52%	62%	75%	82%
7	57%	79%	73%	87%
8	68%	82%	69%	85%
9	66%	93%	73%	93%
10	75%	86%	75%	87%
11	70%	86%	71%	89%
12	61%	81%	73%	88%
13	71%	89%	76%	92%
Average	65%	83%	72%	89%

The percent of successful classifications for 13 regions in the map of Fig. 4. Statistics are given for the four methods: (A) simple histograms, (B) multiresolution histograms, (C) wireless signal strengths, and (D) combined method with both multiresolution histograms and wireless signal strengths.

When combining the multiresolution histograms with the wireless signal strengths, the latter took a secondary role. As noted in the previous section, we chose to weight the signal strength vector less than the histograms. Nevertheless, the signal strengths did contribute a noticeable improvement in our results. The overall success rate with the combined system was 89%. Most notably, regions 4 and 6 had success rates much closer to those of the other regions. These two regions had very similar visual appearances but, because of their different locations, had very different access points visible. Our system had the most difficulty in nearby regions for which both the visual scene and the detectable access points were very similar.

VI. Conclusion and Future Work

The hybrid method we have presented is able to quickly and accurately localize the robot into a correct region. Our results indicate that localization accuracy is significantly improved when five levels of resolution are used instead of one in color histogramming. Single resolution histograms become less useful in classification and matching when the environment is subject to variable lighting conditions. The

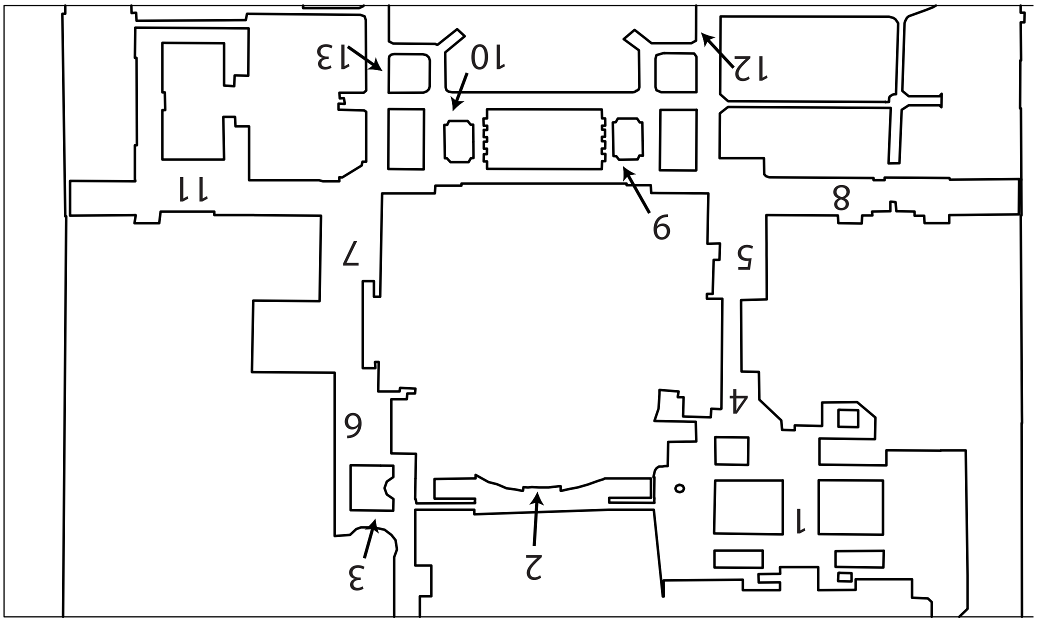

Fig. 4. The two dimensional map of the northern half of the Columbia University campus. The 13 regions we used for our test cases are indicated on the map.

multiresolution histograms help in this situation and provide additional information about spatial relationships in the scene.

We also find that incorporating wireless signal strengths into the method further improves reliability and helps to resolve ambiguities which arise when different regions have similar visual appearances. Physically distant regions will often be covered by different wireless access points, thus giving us another clue as to the mobile robot's current location.

Ultimately, we need to localize the robot exactly. In a very few cases, the robot's region is incorrectly chosen. However, our coarse-fine localization method is able to recover from some of these errors. For the fine-level, we use a method based on camera pose estimation to predict the exact location of the mobile robot [3] assuming we are in the correct region. This method uses the coarse position information from the topological localization to then visually find nearby buildings. We then identify prominent linear features in the scene and match them with a reduced model of those buildings, yielding a pose estimation of the robot. If in fact the wrong region is chosen, the fine matching procedure will report no matches. At this point we can choose to either use the second best matching region as the robot's estimated position and repeat the vision-based fine localization or perturb the robot's position and repeat the topological localization. In essence, the fine-based localization can serve as a feedback confirmation method for the coarse localization. The combination of the two systems will allow us to accurately localize our robot within its test environment without any artificial landmarks or pre-existing knowledge about its position.

References

[1] P. K. Allen, I. Stamos, A. Gueorguiev, E. Gold, and P. Blaer, "Avenue: Automated site modeling in urban environments," in *3DIM, Quebec City*, May 2001, pp. 357–364.

[2] A. Georgiev and P. K. Allen, "Vision for mobile robot localization in urban environments," in *Proc. of IEEE Int. Conference on Intelligent Robots and Systems*, October 2002, pp. 472 – 477.

[3] A. Georgiev and P. K. Allen, "Localization methods for a mobile robot in urban environments," *Transactions on Robotics*, vol. 20, no. 5, October 2004.

[4] P. Blaer and P. K. Allen, "Topological mobile robot localization using fast vision techniques," in *IEEE ICRA*, May 2002, pp. 1031–1036.

[5] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, "Spatial information in multiresolution histograms," in *IEEE CVPR*, 2001, pp. I–702–I–709.

[6] P. Blaer and P. K. Allen, "Topbot: automated network topology detection with a mobile robot," in *IEEE ICRA*, September 2003, pp. 1582–1587.

[7] B. Kuipers and Y.-T. Byun, "A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations," *Journal of Robotics and Autonomous Systems*, vol. 8, pp. 47–63, 1991.

[8] D. Radhakrishnan and I. Nourbakhsh, "Topological robot localization by training a vision-based transition detector," in *IEEE IROS*, October 1999, pp. 468–473.

[9] S. Nayar, "Omnidirectional video camera," in *Proc. DARPA IUW*, May 12-14 1997, pp. 235–242.

[10] R. Cassinis, D. Grana, and A. Rizzi, "Self-localization using an omnidirectional image sensor," in *International Symposium on Intelligent Robotic Systems*, July 1996, pp. 215–222.

[11] N. Winters, J. Gaspar, G. Lacey, and J. Santos-Victor, "Omni-directional vision for robot navigation," in *IEEE Workshop on Omnidirectional Vision*, June 12 2000, pp. 21–28.

[12] I. Ulrich and I. Nourbakhsh, "Appearance-based recognition for topological localization," in *IEEE ICRA*, April 24-28 2000, pp. 1023–1029.

[13] S. Sablak and T. E. Boult, "Multilevel color histogram representation of color images by peaks for omni-camera," in *IASTED International Conference Signal and Image Processing*, October 1999.

[14] H.-M. Gross, A. Koenig, C. Schroeter, and H.-J. Boehme, "Omnivision-based probabilistic self-localization for a mobile shipping assistant continued," in *Proc. of IEEE Int. Conference on Intelligent Robots and Systems*, October 2003, pp. 1505 – 1511.

[15] M. Swain and D. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[16] J. Hafner and H. S. Sawhney, "Efficient color histogram indexing for quadratic form distance functions," *IEEE PAMI*, vol. 17, no. 7, pp. 729–735, July 1995.

[17] M. Stricker and M. Orengo, "Similarity of color images," in *SPIE Conference on Storage and Retrieval for Image and Video Databases III*, vol. 2420, February 1995, pp. 381–392.
[18] M. Werman, S. Peleg, and A. Rosenfeld, "A distance metric for multi-dimensional histograms," in *CVGP*, 1985, vol. 32, pp. 328–336.
[19] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *INFOCOM (2)*, 2000, pp. 775–784.
[20] A. M. Ladd, K. Bekris, A. Rudys, G. Marceau, L. E. Kavraki, and D. S. Wallach, "Robotics-based location sensing using wireless ethernet," in *MOBICOM*, 2002.
[21] S. Tekinay, "Wireless geolocation systems and services." in *IEEE Communications Magazine*, April 1998.
[22] T. D. Hodes, R. H. Katz, E. S. Schreiber, and L. Rowe, "Composable ad hoc mobile services for universal interaction," in *MOBICOM*, September 1997, pp. 1–12.
[23] J. A. Castellanos, J. M. Martinez, J. Neira, and J. D. Tardos, "Simultaneous map building and localization for mobile robots: A multisensor fusion approach," in *IEEE ICRA*, 1998, pp. 1244–1249.
[24] H. Durrant-Whyte, M. Dissanayake, and P. Gibbens, "Toward deployment of large scale simultaneous localization and map building (SLAM) systems," in *Proc. of Int. Simp. on Robotics Research*, 1999, pp. 121–127.
[25] J. Leonard and H. J. S. Feder, "A computationally efficient method for large-scale concurrent mapping and localization," in *Proc. of Int. Simp. on Robotics Research*, 1999, pp. 128–135.
[26] S. Thrun, W. Burgard, and D. Fox, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Autonomous Robots*, vol. 5, pp. 253–271, 1998.
[27] R. Simmons and S. Koenig, "Probabilistic robot navigation in partially observable environments," in *IJCAI*, 1995, pp. 1080–1087.
[28] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte Carlo localization for mobile robots," in *IEEE ICRA*, 1999, pp. 1322–1328.