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

Floor Segmentation of Omnidirectional Images for Mobile Robot Visual Navigation

Luis Felipe Posada, Krishna Kumar Narayanan, Frank Hoffmann and Torsten Bertram

Abstract— **This paper describes a novel approach for purely vision based mobile robot navigation. The visual obstacle avoidance and corridor following behavior rely on the segmentation of the traversable floor region in the omnidirectional robocentric view. The image processing employs a supervised approach in which the segmentation optimal with respect to the appearance of the local environment is determined by cross validation over 3D scans captured by a photonic mixer device (PMD) camera. The range data in the front view provides the seeds and validation data to supervise the appearance based segmentation in the omniview. Segmentation relies on histogram backprojection which maintains separate appearance models for floor, obstacles and background. A naive Bayes classifier predicts the occupancy of the robots local environment by fusing the evidence provided by different segmentations and models. The classification error is analyzed on ground truth data generated by a PMD camera and manually segmented scenes. The scheme is highly robust with respect to ambiguous and misleading visual appearances of obstacles and floor, thus enabling the robot to navigate safely in unstructured environments of diverse appearance, texture and illumination. The proposed vision algorithm and the navigation behavior demonstrate a robust performance in extensive robotic experiments across several hours of autonomous operation.**

I. Introduction

Computer vision has received substantial recognition in robotics in recent years [1]. Many approaches are inspired by visual perception in animals, which behaviors often rely on vision as the key perceptual stimulus. This development is further promoted by the increasing availability of powerful computer vision systems at affordable costs.

Mobile robot navigation conventionally relies on range sensors which provide robust proximity information to nearby obstacles. Vision, compared to range sensors, is much more informative as it allows the distinction among objects that are relevant for navigation, such as floor, obstacles, walls, corridors and doors. Notwithstanding, in spite of the large amount of information in an image, pure autonomous visual navigation still remains a considerable challenge since:

- raw image data is rather complex, which implies the need to segment and group objects, to extract their relevant features and to aggregate those in a meaningful manner,
- navigation behaviors usually rely on depth information of the scene not provided by a single 2D image,
- visual cues such as intensity, texture, optical flow are highly context dependent and
- indoor environments vary substantially in visual appearance and lighting conditions

Visual navigation has been investigated by numerous researchers from both fields robotics and computer vision [1]. Vision has been widely used for map based navigation in the context of vision based localisation and map building. Mapless robot navigation is primarily concerned with reactive behaviors that realize a tight coupling between action and visual cues derived from the segmentation of the image, the observation of features or landmarks traceable across multiple frames or optical flow. The authors in [2], [3] propose color histograms to model the appearance of objects. Obstacles are classified as those regions which differ significantly in their appearance from the floor in the bottom part of the image. The method in [3] is superior in extracting the ground plane accurately, even in the presence of obstacles in the reference region. This robust segmentation is achieved by an adaptive scheme that continuously updates the floor model based on the recently traversed terrain. However, both systems present drawbacks, since the robot is not able to continue navigation if the appearance of the floor changes abruptly, for example markers or carpets on an otherwise homogeneous floor. The approach presented by [4] relies on color histograms, thus is also sensitive to ambiguous texture or color.

The idea of segmenting the floor ground plane for visual navigation has also been adopted by other researchers. Martin designs the vision system by means of evolutionary algorithms in order to estimate the depth of free space along different directions thus mimicking a conventional proximity sensor [5]. The authors in [6] detect ground patches by tracking corner points using planar homographies. The approach in [7] segments the ground plane by estimating plane normals from motion fields. However, both methods assume either texture or movement and the field of view is restricted by the opening angle of the monocular camera.

Photonic mixer device (PMD) cameras provide depth information in addition to an intensity image and unlike stereo vision do not require texture or contrast. Although the limited field of view of the 3D camera is sufficient for detecting frontal obstacles, it is in general too narrow to navigate robustly in confined spaces. These scenarios require behaviors with an omnidirectional perception of the local environment, such as corridor centering, goal point homing or door traversal [8].

Our approach employs a PMD camera to obtain the ground truth segmentation in a narrow frontal field of view. The accuracy of alternative segmentation schemes in conjunction

Luis Felipe Posada, Krishna Kumar Narayanan, Frank Hoffmann and Torsten Bertram are with the Institute of Control Theory and Systems Engineering, Technische Universität Dortmund, 44227, Dortmund, Germany. felipe.posada@tu-dortmund.de

978-1-4244-6676-4/10/$25.00 ©2010 IEEE

with alternative visual cues is evaluated by cross validation on the frontal view. The combination that performs best in the current context is then applied to segment the omnidirectional view providing 360^0 depth information. Moreover, the paper presents a naive Bayes classifier to detect the local free space by fusing visual information provided by multiple segmentations and models.

Recently, self-supervised detection of traversable terrain has become an active field of research in outdoor robotics. The so called near-to-far online learning has been successfully implemented by several researchers [9], [10], [11]. Many of these approaches use laser range finder scans to detect flat drivable areas in front of the vehicle to provide ground truth classification. This data is used to train a model to identify traversable areas of off-road terrain beyond the training region in the vehicles immediate vicinity.

Our approach albeit similar differs in several important aspects, namely: a) PMD camera instead of laser scanner; b) segmentation of the range image in addition to the 2D image; c) multiple representations of the appearance based floor model; d) multiple segmentation algorithms rather than one; e) obstacle and background modelling and f) mixture of experts approach in which the range information is used to validate and select the optimal segmentation and not merely to update the model.

The main contribution of this paper is the presentation of a new robust segmentation method for visual navigation, which confronts unstructured environments under diverse appearance, texture and illumination conditions. The novelty is a mixture of experts approach [12] in which the best segmentation algorithm in conjunction with the best cue is determined from ground truth segmentation of floor and obstacles provided by a PMD camera. This labeled data is used to initialize and validate alternative seed based segmentation algorithms in the omnidirectional view. Two reactive visual behaviors that rely on the segmented omniview are designed: obstacle avoidance and corridor centering.

II. Floor Segmentation

The Pioneer 3DX mobile robot is equipped with a PMD camera and an omnidirectional camera. The PMD camera provides 3D measurements at a 204x204 pixel resolution across a 40° x 40° field of view. The omnidirectional sensor consists of a CCD camera with a hyperbolic mirror and a vertical field of view of 75° directed towards the bottom in order to capture the floor.

The overall system architecture is shown in Fig. 1. The supervised proposed segmentation proceeds in four stages: a) range image segmentation into planar regions; b) projection of these regions into the omnidirectional view; c) omniview segmentation; d) validation and selection of optimal segmentation and cue.

A 3D scan of the frontal region is generated from the depth information of the PMD camera. The 3D scan is segmented into planar surfaces by means of RANSAC. These segments are then classified according to the orientation of the surface normal, distance to the camera and area into three categories:

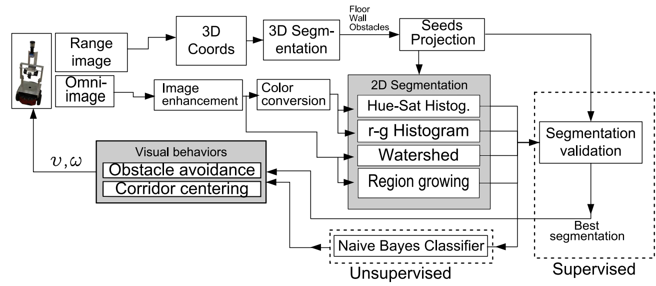
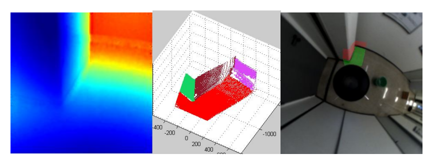


Fig. 1. System architecture

ground considered as free space, walls and obstacles. Scan points that do not belong to a surface are labeled as obstacles. These labeled regions are projected as ground truth segmentation into the omnidirectional view. This ground truth allows the selection of the visual cue and segmentation method that is optimal with respect to the appearance and illumination of the local environment. Part of the 3D scan segmentation is utilized as seeds for the subsequent 2D appearance based segmentation, the remaining data is used to validate the performance of alternative segmentations and cues.

RANSAC is applied in an iterative manner, such that the scan points that belong to the plane with most inliers are removed from the data set. The next plane is estimated from the residual points until the number of points belonging to the best surface model falls below a threshold. The classified 3D points are projected into the omnidirectional view as shown in Fig. 2. It illustrates the range image in which brightness indicates proximity, the estimated and labeled planes in 3D space, and the projection of these labels onto the corresponding sector in the omnidirectional view.



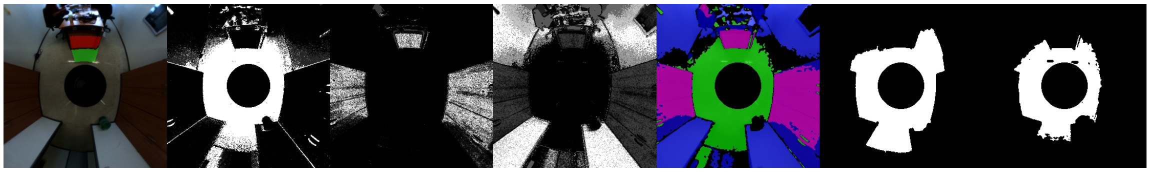
(a) (b) (c)

Fig. 2. Seeds projection: (a) Range image, (b) Planes fitted from the 3D data with RANSAC, (c) Seeds projected into the omnidirectional view.

A key aspect of marker based segmentation algorithms is the selection of initial seeds. The more informative the set of seeds, the better the final segmentation result. In the following we propose four alternative segmentation schemes: segmentation based on color histograms backprojection in two different color spaces, watershed algorithm and model based region growing method. Although we acknowledge the potential utility of other appearance features such as texture [13], only color is considered as it is difficult to extract reliable texture due to the low resolution of the omni-image.

A. Histogram Backprojection

Color histograms provide a compact representation of object appearance that offers the advantages a) to allow a discrete representation of multi modal distributions, b) are less sensitive to changes in viewpoint and scale, and (c)



(a) (b) (c) (d) (e) (f) (g)

Fig. 3. Example of histogram backprojection using three models: (a) input image with floor and obstacle seeds projected, (b) backprojection using H-S floor model, (c) backprojection using H-S obstacle model, (d) backprojection using H-S background model, (e) label image combining the 3 H-S backprojections, (f) watershed, (g) region growing

are partially robust with respect to occlusion. Although, a histogram representation ignores shape and texture, it is suitable for ground floor segmentation in visual navigation [2], [3], [4]. These previous approaches utilize 1D histograms in order to estimate the model of the free space. However, in our experience 2D histograms proved to be more robust and accurate in particular in scenes where floor and obstacle are similar in appearance.

The color histogram models are based on the HSV (Hue-Saturation-Value) and rgb (normalized RGB) color spaces. These color spaces are more robust with respect to changes in intensity, illumination and view point compared to other color representations. Instead of merely building an explicit color model of the floor, our approach maintains additional models of the obstacles and the background. The omnidirectional view extends about $15°$ beyond the horizon line such that pixels above the horizon line certainly do not belong to the floor. Thus the background model is estimated from those pixels which radial distance from the image center is above the corresponding horizon line. The obstacle model is generated from pixels which the 3D segmentation either labeled as obstacle or wall. Those regions that belong to the horizontal floor plane provide the data to maintain the floor appearance model. In order to improve generalization the 2D histograms are quantized into 32-32 bins to reduce the dimensionality and are then averaged and normalized.

Histogram backprojection [14] finds areas in the image which appearance matches with a reference model. We adopt this technique in order to recognize the floor area according to the histogram reference model obtained from the ground truth 3D segmentation. The normalized histogram reference model M in a 2D color space, is compared with the normalized histogram H of the current frame I and the backprojection image B is constructed by

$$B(u, v) = min\left(\frac{M(c(u, v))}{H(c(u, v))}, 1\right) \quad (1)$$

in which $c(u, v)$ denotes the 2D color of pixel (u, v). The backprojection image in the range between $[0, 1]$ reflects the degree of similarity between the area and the reference model. Those colors that occur more frequently in the model than in the overall current scene are likely to belong to the floor. Fig. 3 shows some examples of backprojection.

To cope with dynamic changes in the environment, the histogram models are updated using an exponential moving average (EMA):

$$\bar{M}(x, y)_t = \alpha M(x, y)_t + (1 - \alpha)\bar{M}(x, y)_{t-1} \quad (2)$$

in which M_t denotes the histogram of the current frame and $\bar{M}_{t-1}$ the previous average. The smoothing factor α is set to 0.6 giving more importance to recent observations. The EMA allows a smooth transition of the model under slowly varying conditions. However, in fast changing conditions, the histogram models change very fast and instead of updating, we allow to memorize new models of significant difference. The maximum number of models stored in a short term memory is limited to five instances for obstacle and two for floor and background in order to meet computational constraints. Histogram intersection determines if the current histogram overlaps with any of the models stored in the short term memory:

$$d(M_1, M_2) = \sum_{x,y} \min(M_1(x, y), M_2(x, y)) \quad (3)$$

in which a score of one corresponds to an exact match and zero to a complete mismatch. In practice, a new model is initialized if the histogram intersection between the current frames and the most similar stored models falls below a threshold of 0.4. New histogram models replace the oldest model in case the maximum storage capacity is exceeded. Models that do not match with any of the thirty most recent frames are automatically discarded.

B. Watershed

Watershed segmentation [15] is understood by interpreting the gradient image as elevation information. High magnitude gradients correspond to watershed lines. Pixels in the elevation image are attracted to their local minimum, thus forming so called basins. In marker based watershed, water is flooded evenly from each marker to flood the basins. The process stops once the water level reaches the highest peak in the landscape. Basins connected with the same original marker are merged into a single region. In order to prevent unbounded growth of floor regions, the border pixels corresponding to pixels above the horizon are considered as non floor. An example of marker controlled watershed segmentation is shown in Fig. 3f.

C. Region Growing

Region growing [16] starts with a set of seed points or subregions and neighboring pixels are added based on their similarity in an incremental fashion to each region. In our case, the similarity criterion is based on color, pixels that have similar R, G and B values as the neighboring pixels are added to the region. Region growing works well, if the regions are separated by sharp edges in intensity or color. Nevertheless, it is highly sensitive to variations in intensity and illumination. A result of region growing segmentation is shown in Fig. 3g.

D. Supervised Segmentation

It is apparent that no single segmentation scheme alone provides a robust accurate segmentation across all scenes. The idea is therefore to validate the accuracy of segmentation on labeled ground truth data and then to select the method best suited for the current floor texture and illumination. For this purpose the ground truth segmentation data obtained from the 3D scans is partitioned into training and validation data. The training data is used to build the reference model for the histogram backprojection and to provide the initial seeds for marker based watershed and region growing. A false negative is constituted by a floor pixel incorrectly classified as obstacle and a false positive is an obstacle pixel incorrectly classified as floor. The segmentation accuracy is determined based on the false positive rate f_p, number of false positives divided by the total number of negatives, and false negative rate f_n on the validation data. Both rates are aggregated into a total classification error

$$f = (2f_p + f_n)/3 \tag{4}$$

in which false positives are weighted twice as strong, since an obstacle miss is potentially more severe in the context of obstacle avoidance.

The segmentation validation utilizes two-fold cross validation, in which the roles of training and validation data set are reversed and the classification error is averaged over both folds. The segmentation method with lowest aggregated classification error is applied to the entire omnidirectional view. In order to save computational resources the segmentation validation and selection is only repeated every fifth frame. The final segmentation is filtered by a 5x5 median filter in order to eliminate isolated pixels and noise.

E. Naive Bayes Classification

The next step is to achieve robust floor segmentation without supervision and validation of the segmentation by depth data. The 3D data is only used offline to compute the likelihood of features and class priors.

Instead of selecting the optimal among the four alternative segmentations based on cross validation, the naive Bayes classifier labels pixels as either occupied or free space. The features are constituted by the similarity values of the backprojection images for the three different models, floor, obstacle and background in conjunction with the cues of hue, saturation and rgb color. The likelihood $p(f_i|C_j)$ of a continuous feature f_i given the class C_j is modeled by a Gaussian distribution. The likelihoods of the data and the class priors are computed from the observed frequencies of classes and features in the training data. The training data consists of the backprojection similarity with different models of almost two million pixels captured from 500 images of which the true class label is established from the PMD depth information and 3D segmentation. The class priors and feature probability distributions are approximated with relative frequencies from the ground truth provided by the PMD camera. The classification performance is validated on additional 500 images with ground truth obtained from PMD data and 30 images in which the actual floor area is segmented by hand. The naive Bayes classifier computes the a posteriori probability of the classes $C = \{Floor, Obstacle\}$ according to the likelihood of the conditionally independent features:

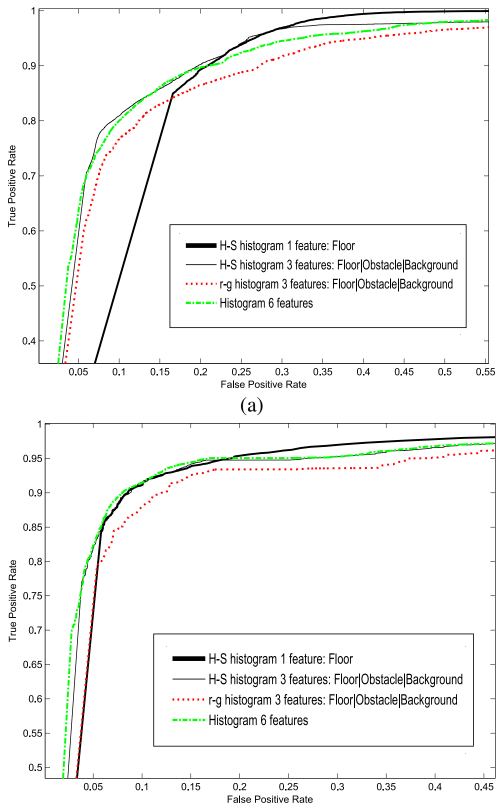
$$p(C_j|f_i) = \frac{1}{Z} p(C_j) \prod_{i=1}^{n} p(f_i|C_j) \tag{5}$$

in which the normalisation factor Z is the evidence of the features f_i. The ultimate decision boundary depends on the application specific relative costs of false positive and false negative classifications. Fig. 4 shows the receiver operator characteristic (ROC) curve of the naive Bayes classifier for different feature sets. The upper ROC curve refers to test data obtained from PMD depth information, the lower plot depicts the classification error on the manually segmented and labelled image set. It is apparent that hue and saturation histogram backprojection are superior to r-g histogram backprojection. In fact the ROC curve for only hue and saturation histogram backprojection almost coincides with the performance of the Bayes classifier that uses r-g segmentation in addition. It becomes also apparent that maintaining separate models for floor, obstacle and background results in a much better classification than classification based on the segmented floor model alone. The classification performance is similar across the PMD validation and the hand labeled data set. In fact the classification error on the hand labelled data is even lower. We attribute this phenomenon to the fact that the depth based segmentation is less accurate than hand segmentation. Thus the test data set itself contains a small fraction of incorrect samples that contribute to the classification error.

III. Visual Behaviors

The corridor centering behavior is supposed to align the robots heading with the orientation of the walls and centering the robot in the middle of the corridor. The robots heading error and lateral offset are reflected in the distribution of floor area in robocentric coordinates. For this purpose the segmented floor region is characterized by image moments [17]. From these moments the distribution of the free space is approximated by an ellipse with semi major axis a, semi minor axis b, and orientation φ as shown in Fig. 5a. Considering a look ahead point A located on the major axis at a distance d from the centroid, the angles α and β are easily computed.

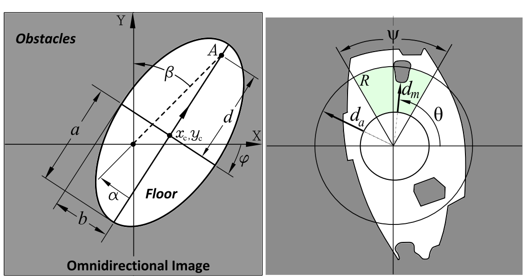
In the experiments the distance d in pixels corresponds to a distance of 2m. The turn rate to align the robot with the



(a)

(b)

Fig. 4. Naive Bayes classifier ROC curves (a) for PMD ground truth data set, (b) for hand label ground truth data set



(a) Corridor centering (b) Obstacle avoidance

Fig. 5. Visual behavior features

corridor middle line results from a stabilizing error feedback proportional control law:

$$\begin{aligned} v &= v_0 \\ \omega &= -K_a\alpha - K_b\beta \end{aligned} \tag{6}$$

in which the gains K_a and K_b are tuned to achieve a smooth convergence. The robots translational velocity for corridor following is constant v_0.

The obstacle avoidance behavior monitors the frontal region of the omnidirectional image. The avoidance region R is defined by the aperture angle ψ and the activation distance d_a as illustrated in Fig. 5b. If a non floor pixel denoting an obstacle is detected inside R, the behavior is activated. The perception is aggregated into the distance d_m and heading θ of the closest obstacle. The perception is mapped onto a

TABLE I

COMPARISON OF THE SEGMENTATION SCHEMES UNDER TWO DATA SETS

	PMD ground truth		Hand labeled	
Method	FPR	TPR	FPR	TPR
H-S histogram features: F	0.173	0.855	0.089	0.905
H-S histogram features: F|O|	0.123	0.845	0.094	0.971
H-S histogram features: F|O|B	0.137	0.845	0.096	0.911
r-g Histogram features: F	0.210	0.828	0.118	0.923
r-g Histogram features: F|O	0.145	0.836	0.105	0.870
r-g Histogram features: F|O|B	0.147	0.825	0.112	0.896
Histogram all 6 features	0.158	0.867	0.085	0.904
Watershed	0.021	0.948	0.140	0.749
Region growing	0.097	0.932	0.149	0.676

F: Floor, O: Obstacle, B: Background

motor action

$$\begin{aligned} v &= v_a d_m / d_a \\ \omega &= sgn(\omega_0) \cdot \omega_a \cdot \sin\theta \end{aligned} \tag{7}$$

The sign of the initial turn rate ω_0 depends on whether the obstacle is located in the right or left half plane, the robot then turns in the opposite direction. A turn flag $sgn(\omega_0)$ memorizes this turning direction which is maintained throughout the entire avoidance manœuvre. The magnitude of the turn rate is modulated by the relative heading of the obstacle. The constants v_a and ω_a denote the avoidance translational and rotational velocities.

IV. EXPERIMENTAL RESULTS

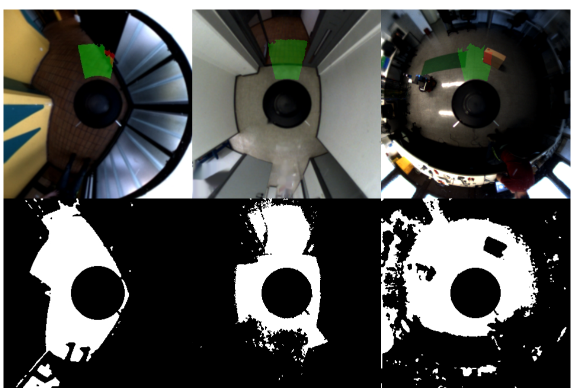
Table I compares the true positive and false negative rates of watershed and region growing seeded with one percent of labeled pixels as seeds with the unsupervised classification errors of the backprojection schemes. The results confirm the conclusions obtained from the ROC curve analysis, namely that three models outperform the single floor model and that hue and saturation are more reliable cues compared to color. Using color cues in addition to hue and saturation does not improve the accuracy. On both test sets the hue saturation classifier with three models achieves true positive (floor classified as floor) rates between $0.85 - 0.9$ if one accepts a false positive (obstacle classified as floor) rate between $0.1 - 0.13$.

The classification performance of watershed and region growing is superior to the Bayes classifier on the PMD test set, with watershed outperforming region growing. It is not surprising that watershed and region growing achieve high classification rates on the PMD test set, as the seeds stem from the same narrow frontal region as the test data. The true generalized classification error of watershed and region growing becomes apparent on the manually labeled data set, in which the test data is uniformly distributed across the omnidirectional view. In this case the classification performance is clearly inferior in comparison to the Bayes classifier. The dependence of watershed and region growing on a representative set of seeds is a clear disadvantage over more robust histogram backprojection.

The Pioneer robot is equipped with a 1.8 GHz dual-core standard laptop for running the computer vision algorithms.

Considering unoptimized code, the complete processing time of a 320x320 pixels cropped omnidirectional image including the validation step, takes between 150 to 180ms. One complete 2D histogram backprojection segmentation takes about 15ms, watershed segmentation takes 10ms, and region growing takes on average 30ms.

Figure 6 shows the results of four prototypical scenarios together with the corresponding floor segmentation results. Figure 6a, shows the segmentation of a tiled floor with strong sun reflection entering from the windows at the right side. The resulting segmentation is obtained from naive Bayes classification with the complete set of features. The ability of our scheme to adapt to new unseen floor surfaces is illustrated in Figure 6b in which the floor color and texture abruptly change behind the door. Despite this ambiguity in appearance the entire floor is correctly segmented. The final figure shows a situation of imperfect lighting conditions, a flat carpet to the left and an obstacle to the right. R-g histogram backprojection performs best as in this particular scenario color constitutes the most discriminating cue able to generalize ground floor, carpet and obstacle beyond the seeded regions. Notice, that the regions with light reflections are still correctly segmented due to the masking of saturated areas.



(a) (b) (c)

Fig. 6. Segmentation Examples

The proposed scheme is able to navigate the mobile robot safely through environments over prolonged periods of operation. Several test runs were conducted, some of them taking more than 20 minutes, which accumulate to several hours of validation under realistic conditions. The enclosed videos show examples of navigation from the omnidirectional perspective.

V. Conclusion

This paper describes a novel approach for purely visual segmentation of free space for mobile robot navigation. The omnidirectional floor segmentation employs a supervised mixture of experts approach which determines the segmentation method that is optimal for the current situation by cross validation over ground truth data provided by a PMD camera.

An unsupervised scheme is also proposed, in which a naive Bayes classifier fuses the evidence provided by different segmentation methods and features. The consideration of multiple appearance models for floor, obstacles and background considerably improves the segmentation performance.

The visual obstacle avoidance and corridor following utilize the distribution of local free space to control the robots motion. The robot is able to operate safely in unstructured environments of diverse appearance, texture and illumination. The proposed segmentation and visual behaviors were successfully tested and validated in robotics runs over several hours.

References

[1] F. Bonin-Font, A. Ortiz, and G. Oliver. Visual navigation for mobile robots: A survey. *Journal of Intelligent Robotics System*, 53(3):263–296, 2008.

[2] L.M. Lorigo, R.A. Brooks, and W.E.L. Grimsou. Visually-guided obstacle avoidance in unstructured environments. In *Proc. IEEE/RSJ IROS'97*, volume 1, pages 373–379, Sep 1997.

[3] I. Ulrich and I. Nourbakhsh. Appearance-based obstacle detection with monocular color vision. In *Proc. AAAI 2000*, Austin, TX, 2000.

[4] S. Lenser and M. Veloso. Visual sonar: fast obstacle avoidance using monocular vision. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2003).*, volume 1, pages 886–891, Oct. 2003.

[5] C. M. Martin. Evolving visual sonar: Depth from monocular images. *Pattern Recognition Letters*, 27(11):1174–1180, 2006.

[6] N. Pears and B. Liang. Ground plane segmentation for mobile robot visual navigation. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, volume 3, pages 1513–1518, 2001.

[7] Young Kim and Hakil Kim. Layered ground floor detection for vision-based mobile robot navigation. In *Proc. IEEE Int. Conf. on Robotics and Automation*, volume 1, pages 13–18, April 2004.

[8] L. F. Posada, T. Nierobisch, F. Hoffmann, and T. Bertram. Image signal processing for visual door passing with an omnidirectional camera. In *Proc. Int. Conf. on Computer Vision Theory and Applications*, pages 472–479, 2009.

[9] H. Dahlkamp, A. Kaehler, D. Stavens, S. Thrun, and Bradski G. Self-supervised monocular road detection in desert terrain. In *Proc. of the Robotics Science and Systems Conference*, 2006.

[10] G. Grudic, J. Mulligan, M. Otte, and A. Bates. Online learning of multiple perceptual models for navigation in unknown terrain. In *6th International Conference on Field and Service Robotics*, 2007.

[11] Dongshin Kim, Jie Sun, Sang Min, Oh James, M. Rehg, and Aaron F. Bobick. Traversability classification using unsupervised on-line visual learning for outdoor robot navigation. In *In Proc. of Int. Conf. on Robotics and Automation (ICRA)*, 2006.

[12] R. Polikar. Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, 6(3):21–45, 2006.

[13] Morten Rufus Blas, Motilal Agrawal, Aravind Sundaresan, and Kurt Konolige. Fast color/texture segmentation for outdoor robots. In *IROS*, pages 4078–4085, 2008.

[14] Michael J. Swain and Dana H. Ballard. Color indexing. *International Journal of Computer Vision*, 7:11–32, 1991.

[15] Jos B. T. M. Roerdink and Arnold Meijster. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae*, 41(1-2):187–228, 2001.

[16] Linda G. Shapiro and George C. Stockman. *Computer Vision*. Prentice-Hall, 2001.

[17] J. Prokop and Anthony P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. In *Graphical Models and Image Processing*, pages 438–460, 1992.