Introducing A Novel Vision Based Obstacle Avoidance Technique for Navigation of Autonomous Mobile Robots

Mostafa Sharifi, Prof. XiaoQi Chen, *Senior IEEE Member*
Department of Mechanical Engineering
University of Canterbury
Christchurch, New Zealand
mostafa.sharifi@pg.canterbury.ac.nz, xiaoqi.chen@canterbury.ac.nz

***Abstract*— This paper introduces a novel vision based obstacle avoidance technique for indoor navigation of autonomous mobile robots. The indoor environment is considered as office environment with homogenous surfaces. In this technique, a color image taken by a monocular vision camera is clustered by mean-shift algorithm, then the clustered image is classified by a novel classification technique based on graph partitioning theory. The classified image includes meaningful information such as floor, walls and obstacles for robot to navigate around office environment. The simulation results show the effectiveness of proposed technique for further real-time implementation and experiments.**

Keywords—mobile robotics; vision-based navigation; appearance based; image classification; graph partitioning

I. INTRODUCTION

Robot navigation is the process of defining a safe and suitable path for a mobile robot to travel from starting to the target point [1]. There are different sensors to provide navigation information for mobile robots. Generally, popular sensors include ultrasonic sensors, laser rangefinders, and vision sensors. Ultrasonic sensors are cheap but they suffer from having a low range and resolution. Laser rangefinders provide better resolution but they are expensive. Vision sensors provide better performance in comparison to other types of sensors for indoor navigation of mobile robots because of low cost, low power and high resolution and capturing ability. They are sensitive to the light variations but it will not be a problem as long as it is used for indoor and structured environment with proper lighting conditions. Vision sensors provide detailed information of the environment which can be used for navigation of mobile robots to follow the path and avoid the obstacles [2].

Vision-based obstacle avoidance techniques are divided into two main groups: vision based apparent motion techniques and vision based appearance techniques. Vision based apparent motion techniques work based on the idea of controlling robot using the optical flow which is the pattern of apparent motions of objects, surfaces and other features in an image between camera and the scene [3]. Appearance based techniques are based on basic image processing techniques and consist of operation on image pixels to classify them as path (ground) and obstacles (e.g. walls). These techniques suffer from sensitivity to the illumination and light variations of operated environment [4]. All Appearance based techniques require three assumptions that are practical for different types of environment. These assumptions are [5]:

a) Dissimilarity in appearance between obstacles and ground

b) flat ground

c) No hanging obstacles.

Both groups are easy to implement and also have good performance for real-time applications, however the major problem of apparent motions techniques is motion discontinuity caused by movement of objects with respect to other objects or background.

Several research have been done using appearance-based techniques to solve the problem of navigation and obstacle avoidance for mobile robots. Ulrich and Nourbakhsh [6] developed an appearance based technique to classify each pixel from a single passive colour camera into ground or obstacle by comparing each image histogram to the reference area histogram. Booji, O. et al in [7] developed an appearance based topological map building system for navigation of mobile robots. This technique uses epipolar geometry and a planar constraint to compute required information for navigation of robot. Zhang and Kleeman in [8] presented an appearance-based visual navigation technique working based on the comparison between images and a sequence of reference images obtained during the human-guide route-teaching phase. The comparison is done using cross-correlation in the Fourier domain to compensate the orientation difference. Zhang Y. et al in [9] developed an appearance-based navigation technique using omnidirectional images captured by an omnidirectional camera. The robot is able to use the information from these image to localize itself in the environment.

This paper introduces a novel vision appearance based obstacle avoidance technique for indoor navigation of mobile robots. This technique is free from any reference map or images of the operating environment. This paper is organized as follows. In section 2, the methodology for proposed technique is described. In section 3, simulation results and

978-1-4799-8389-6/15/$31.00 ©2015 IEEE

discussion on the proposed technique and results are presented and finally in section 4, the conclusion of the whole paper is provided.

II. Proposed Methodology

This vision-based obstacle avoidance technique is an appearance-based technique which classifies the input colour image from a monocular vision camera into defined classes such as ground, walls and obstacles. The methodology of proposed technique is as follows. The algorithm consists following stages as shown in Fig. 1.

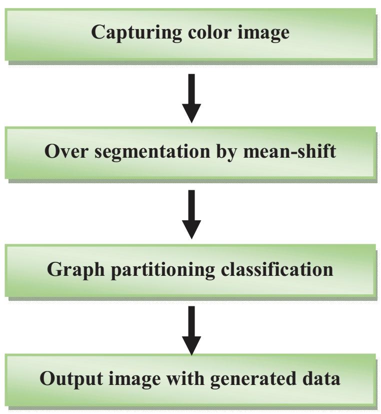


Fig. 1. Flow diagram of the proposed approach

A. Image Accusiotion and Laboratory Equipment

The images used in simulation are taken by a monocular vision camera with resolution of 640 × 480. The considered environment was Mechatronics research lab at the mechanical engineering department, University of Canterbury. The computer which has been used is configured with Intel® Core™ i7-4770 CPU @ 3.40 GHz and 16 GB RAM memory. The simulation platform is Matlab R2014a.

B. Over-Segmentation by Mean-Shift Algorithm

In image segmentation and classification, the process of over segmentation prepares image to be processed faster in next stages. In graph-based image processing techniques, each pixel of image is considered as one graph node and each node should be processed which eventually causes high computational cost. With over segmentation using clustering techniques, image is segmented into several numbers of clusters. In this case, each cluster is considered as one node and the graph size will be decreased which it causes low computational cost and faster speed.

Mean Shift (MS) algorithm is one of the clustering algorithms which is operating in feature space. In this algorithm, a nonparametric kernel density function converts image into feature space to model the features. Then it tries to find the modes density function and assign each point to the specific modes [10]. The formulation of MS algorithm to approximate the kernel density at point x in a dataset of $\{x_i\}_{i=1}^{n} = R^D$ are as follows:

$$\hat{f} = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i) \tag{1}$$

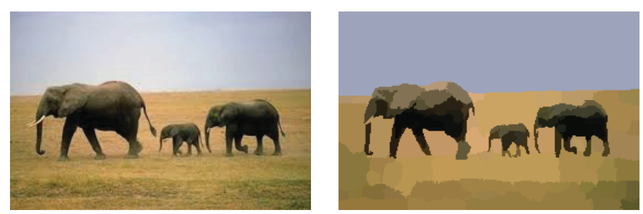
By a Given kernel k , bandwidth parameter h , kernel density estimation would be:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} k\left(\left\|\frac{x - x_i}{h}\right\|^2\right) \tag{2}$$

And finally the mean shift vector is given by:

$$m(x) = \frac{\sum_{i=1}^{n} x_i k\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{k\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x \tag{3}$$

Fig. 2 shows an example of image clustering performed by MS algorithm.



a b

Fig. 2. Results by MS algorithm: a) Sample image b) Clustered image by MS algorithm

C. Graph Partitioning Based Image Classification

In graph theory, data can be represented in the form of a graph $G = (V, E)$, with V vertices and E edges. Such a graph can be partitioned into smaller parts using graph partitioning techniques. Different graph partitioning techniques are available which work based on different methods. Spectral graph partitioning is one of the partitioning techniques which has been favored and considered vastly to solve many problems especially in the area of image processing and computer vision. The applications includes image segmentation [10-11], image retrieval [12], object tracking and recognition [13-14], and etc.

In spectral graph image partitioning, image is represented as a weighted graph where, each pixel is a node in graph and adjacent nodes are connected together by a link. There is a weight value for each link represents the similarity or dissimilarity between two adjacent nodes (pixels). Graph partitioning tries to partition the graph nodes into smaller groups with high similarity for intragroup nodes and low similarity for intergroup nodes.

For a graph $G=(V,E,W)$ where V is the set of nodes, E is the set of linked edges between nodes and W is the affinity weight matrix where $w(u,v)$ is the weight value between the nodes u and v as the $(u,v)^{th}$ element in W matrix [15]. For this graph, graph partitioning can be done by removing the edges with the least weight values. Therefore the graph will be partitioned into dissimilar components. The total value of the removed edges weights between two sets of nodes is defined as $cut(A,B)$:

$$cut(A,B)=\sum_{e\in A,v\in B} w(u,v) \tag{4}$$

Achieving the minimum $cut(A,B)$ is the best cut which the graph will be partitioned into two disjoint regions. Finding minimum cut value in a graph requires especial computational techniques. One of the spectral graph partitioning techniques is Normalized-Cut (N-Cut) developed by Shi and Malik [16] for image segmentation. This technique considers the dissimilarity between the different groups as well as the similarity within the groups. The formulation for N-Cut is:

$$Ncut(A,B)=\frac{cut(A,B)}{assoc(A,V)}+\frac{cut(A,B)}{assoc(B,V)} \tag{5}$$

Where $assoc(A,V)$ and $assoc(B,V)$ are the total weights of connected nodes in part A and B to all the other nodes in graph. Therefore, with maximizing intergroup association value and minimizing intragroup association value a better cut will be obtained. This process is NP complete problem and required approximation techniques to achieve. Shi and Malik provided a good approximation technique for N-Cut by representing as a generalized eigenvalue problem.

The proposed graph based image classification is based on the technique we developed previously and described completely in [17]. This technique uses N-Cut procedure but with a different formulation for creating weigh matrix. With this difference, N-Cut procedure provides image classification rather than image segmentation. Weight value $w(u,v)$ between nodes u and v is defined by:

$$w(u,v)=\begin{cases} e^{-\left[\frac{\|F(u)-F(v)\|_2^2}{d_I}\right]} & \text{if u and v are adjacent, otherwise} \\ 0 & \end{cases} \tag{6}$$

Where $F(u)=\{L(u),u(u),v(u)\}$ is the color vector of node u in luv color space, and d_I is a positive scaling factor defining the sensitivity of the $w(u,v)$ to the color differences between u and. As it can be seen in (6), the weight value for those nodes which are not adjacent are considered as zero. With this formulation, graph partition achieves image segmentation for the defined number of segments (i.e. $k=4$, k is the number of segmented regions). By removing adjacency condition, every node will be connected to all the other nodes by a weighted edge and the representation of the graph will be as a complete graph. The formulation for this case is:

$$w(u,v)=e^{-\left[\frac{\|F(u)-F(v)\|_2^2}{d_I}\right]} \tag{7}$$

With this procedure, graph partitioning achieves image classification and k will be the number of desired classes. Fig. 3 shows the segmentation and classification by this technique. As it can be seen, this technique classifies the sample image into three classes (Fig. 3 (d)).

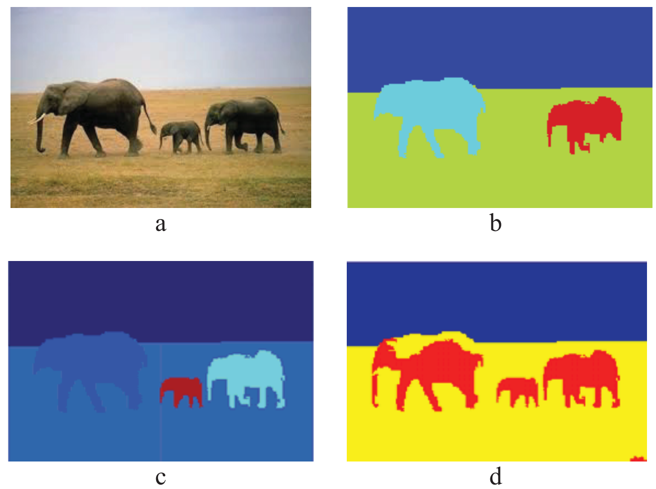


Fig. 3. Simulation results of segmentation and classification for the proposed technique: a) sample image, b) segmentation k = 4, c) segmentation k = 5, d) classification k = 3.

III. Simulation Results and Discussion

In this section, the simulation results are presented to evaluate the performance of proposed approach used for indoor navigation of autonomous mobile robots. In order to test the performance, first series of simulations have been done on some sample images taken by a monocular VGA camera from the office environment in Mechatronics Research Lab, University of Canterbury. The second series of simulations have been done on images from [18] to compare the results with one of the recent related techniques.

The simulation results for first set of sample images are shown in Fig. 4. As it can be seen on sample images, the proposed techniques classifies images into the ground and walls very good. This information can be used to extract the navigation path for robot to navigate inside the office environment. It is tried to find a specific type of environment where there is not much dissimilarity in appearance between the ground and walls to challenge the proposed technique.

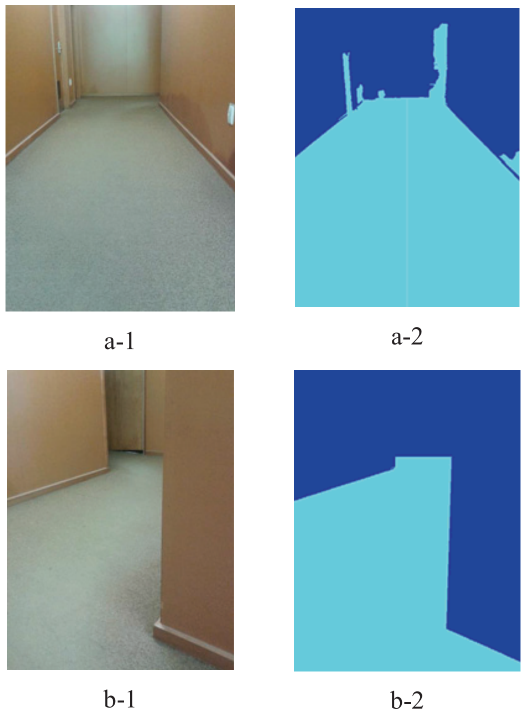


Fig. 4. Simulation results on the images captured by a VGA monocular vision camera

To compare the proposed technique with one of the recent techniques, the simulation has been done on some images from the technique proposed in [18], a vision-based ground-plane classification technique for autonomous indoor navigation of mobile robots, developed by Low and Manzanera. This technique has three main phases including initialisation phase, operation phase and update phase. Initialisation phase includes a training process to create a world model consists of classified ground/object space which is constructed by a Distributed-Fusion process. In operation phase, Markov Random Field (MRF) classifies a new image using world model while update phase is working in parallel with operation phase to update and feedback classification confidence to the learning model. This method is considered as a learning-based technique but because of using visual cues to improve classification, it is also considered as an adaptive self-supervised system.

The simulation results to compare two techniques are shown in Fig. 5. In Fig. 5, a-1 and a-2 are the two sample images used and classified using MRF in [18] and b-1 and b-2 are the results as well. These two images have been tested by our technique which the results can be seen as c-1 and c-2 in Fig.5 as well. As it can be seen, for the first image of the corridor, our proposed approach classifies the ground and walls correctly (Fig. 5-c-1) as well as the result by the other technique.

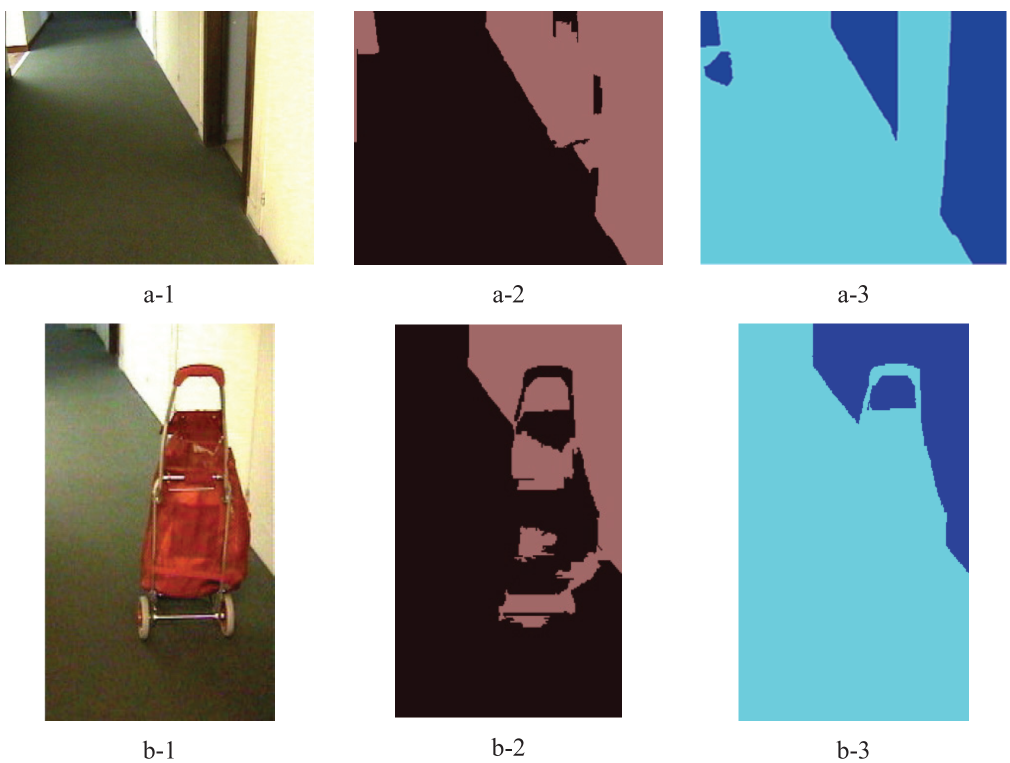


Fig. 5. Simulation and comparison of results on the images from [18]

For the image with trolley as an obstacle, it can be seen that in both techniques trolley is classified incorrectly. These results show a bad classification with unseen obstacle. In order to cope with this problem the obstacle should be considered as an extra class, thus there will be three classes as ground, wall and obstacle. The simulation result when classification is done based on three mentioned classes is shown in Fig. 6.

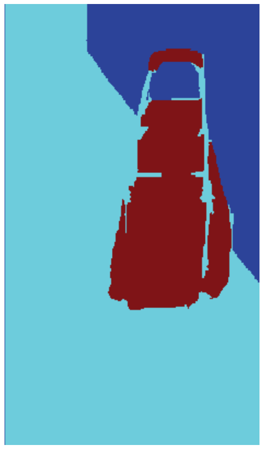


Fig. 6. Simulation result of classification in the presence of obstacle into three classes

It can be seen that the extra class defines the obstacle properly. To compare the results, a histogram matching has been done to quantify the difference between result obtained by each technique and a desired classification done by hand. The comparison details are shown in Table. 1. The numbers provided in the table are pairwise distances between histogram vectors of two compared images. Smaller number shows the better match between two images and zero would be the complete match.

TABLE I. COMPARISON RESULTS OF PROPOSED TECHNIQUE AND THE TECHNIQUE DEVELOPED IN [18]

	Low and Manzanera [18]	**Proposed technique (2 classes)**	**Proposed technique(3 classes)**
Figure 5-1	0.0263	0.1389	-
Figure 5-2	0.9079	1.0051	-
Figure 6	-	-	0.6539

As it can be seen in the above table, the technique developed in [18] provides slightly better results (Fig. 5) compare to our proposed technique but both of them do not produce good results in the presence of obstacle (Fig 5-2). Adding an extra class as an obstacle decreases the error and provides better results compare to have only two classes (Fig. 6). Also it is worth to mention that the technique in [18] is considered as a supervised technique while our proposed technique is completely unsupervised and it only depends on determining the number of classes (k).

As it has been described previously, in appearance based navigation three assumptions should be considered to have the practical and desired outcomes. These assumptions are dissimilarity in appearance between obstacles and ground, flat ground and no hanging obstacles. These conditions must be provided by the operation environment. The classified regions information can be used by mobile robot to navigate in the operating environment. Based on the obtained results, our proposed approach provides satisfactory results compared to similar techniques to be used for indoor navigation of mobile robots.

IV. CONCLUSION

A novel vision-based obstacle avoidance technique for indoor navigation of mobile robots is introduced in this paper. This vision appearance-based technique includes mean-shift algorithm for clustering and a novel graph-based classification to classify images captured by a VGA monocular vision camera from the homogenous office environment. The classified information can be used by mobile robot to move around the office environment. The simulation results show that the proposed technique is useful and practical to apply on mobile robots. The future work will be focused on implementing the proposed technique on a mobile robot for real-time experiments.

REFERENCES

[1] Federico, C.; Ollero, A., Intelligent mobile robot navigation. Vol. 16. Springer, 2005.

[2] Mehmet Serdar, G.; Bicker, R., "Vision based obstacle avoidance techniques." Recent Advances in Mobile Robotics, 83-108, 2012.

[3] Mehmet Serdar, G.; Bicker, R., "Optical flow based system design for mobile robots." In Robotics Automation and Mechatronics (RAM), 2010 IEEE Conference on, pp. 545-550. IEEE, 2010.

[4] Desouza, G.N.; Kak, A.C., "Vision for mobile robot navigation: a survey," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.24, no.2, pp.237,267, Feb 2002.

[5] Saitoh, T., Tada, N., & Konishi, R., "Indoor mobile robot navigation by center following based on monocular vision," Computer Vision, 352-366, 2008.

[6] Ulrich, I.; Nourbakhsh, I. R., "Appearance-Based Obstacle Detection with Monocular Color Vision," In Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence, AAAI Press 866-871, 2000.

[7] Booij, O.; Terwijn, B.; Zivkovic, Z.; Krose, B., "Navigation using an appearance based topological map," Robotics and Automation, 2007 IEEE International Conference on , vol., no., pp.3927,3932, 10-14 April 2007.

[8] Zhang, Alan M.; and Kleeman, L., "Robust appearance based visual route following for navigation in large-scale outdoor environments." The International Journal of Robotics Research 28, no. 3, 331-356, 2009.

[9] Zhang, Y.; Cao, J.; Lou, L.; Su B., "Appearance-based mobile robot navigation using omnidirectional camera," Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on , vol., no., pp.2357,2361, 29-31 May 2012.

[10] Park, A.; Kim, K.; Min, S.; Yun, S.; Jung, K., "Graph Cuts-Based Automatic Color Image Segmentation," Digital Image Computing: Techniques and Applications (DICTA), 2008 , vol., no., pp.564,571, 1-3 Dec. 2008.

[11] Yi, F.; Moon, I., "Image segmentation: A survey of graph-cut methods," Systems and Informatics (ICSAI), 2012 International Conference on , vol., no., pp.1936,1941, 19-20 May 2012.

[12] Kumar, A.; Jinman Kim; Dagan Feng; Fulham, M., "Graph-based retrieval of PET-CT images using vector space embedding," Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on , vol., no., pp.413,416, 20-22 June 2013.

[13] Gomila, C.; Meyer, F., "Graph-based object tracking," Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on , vol.2, no., pp.II,41-4 vol.3, 14-17 Sept. 2003.

[14] Hingway, S.P.; Bhurchandi, K.M., "A Simple Graph Theoretic Approach for Object Recognition," Emerging Trends in Engineering and Technology (ICETET), 2011 4th International Conference on , vol., no., pp.200,205, 18-20 Nov. 2011.

[15] Boykov, Y.; Kolmogorov, V., "An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.26, no.9, pp.1124,1137, Sept. 2004.

[16] Shi, J.; Malik, J., "Normalized cuts and image segmentation," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.22, no.8, pp.888,905, Aug 2000.

[17] Sharifi, M.; Kiani, K.; Kheyrkhahan, M., "A Graph-Based Image Segmentation Approach for Image Classification and Its Application on SAR Images," Przegląd Elektrotechniczny 89, 202-205, 2013.

[18] Low, T.; Manzanera, A., "Ground-plane classification for robot navigation: Combining multiple cues toward a visual-based learning system," Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on , vol., no., pp.994,999, 7-10 Dec. 2010