Markov Random Field based Small Obstacle Discovery over Images

Suryansh Kumar*, M Siva Karthik*, K Madhava Krishna*

Abstract—Small obstacles of the order of $0.5 - 3cms$ and homogeneous scenes often pose a problem for indoor mobile robots. These obstacles cannot be clearly distinguished even with the state of the art depth sensors or laser range finders using existing vision based algorithms. With the advent of sophisticated image processing algorithms like SLIC [1] and LSD [9], it is possible to extract rich information from an image which led us to develop a novel architecture to detect very small obstacles on the floor using a monocular camera. This information is further processed using a Markov Random Field based graph cut formalism that precisely segments the floor and detects obstacles which are extremely low. We show robust and accurate obstacle detection and floor segmentation in diverse environments over a large variety of objects found indoors. In our case, low lying obstacles, changing floor patterns and extremely homogeneous environments are properly classified which leads to a drastic decrease in the number of obstacles that may not be classified by existing robotic vision algorithms.

I. INTRODUCTION

Indoor environments often consist of very small obstacles on the ground making it difficult for robots to navigate. Proliferation of indoor robots has increased the need for optimal ground segmentation and obstacle detection algorithms for effective navigation. To perform such tasks effectively, use of monocular vision systems is on the rise due to various reasons including low cost, low weight, portability, legacy of libraries, efficient process times and community support. Extremely low lying obstacles of the order of $0.5 - 2cms$ or those which are similar to floor appearance hinder the performance of existing monocular vision based floor segmentation or obstacle detection algorithms. Also, recent papers like [2] state that such low lying obstacles or virtual planes [10] pose a hindrance to the navigation of humanoids and other indoor mobile robots.

Current robotic vision algorithms depend either on the appearance of the floor or on the homography of the floor, but both of these approaches cannot cater to solve the task of identifying low-lying obstacles or those which share similar appearance as the floor. Appearance based floor segmentation models fail when the floor texture changes or the obstacle has an extremely similar appearance as the floor (figure 3). Also, Homography models under perform when obstacles are of the height of $0.5 - 2cms$ which virtually appear to be the ground plane (figures 3, 4 and 5).

We devise a novel pipeline that discovers and segments obstacles followed by precise floor segmentation in such difficult conditions as mentioned above. On careful observation, it is found that is is desirable to get a rough segmentation
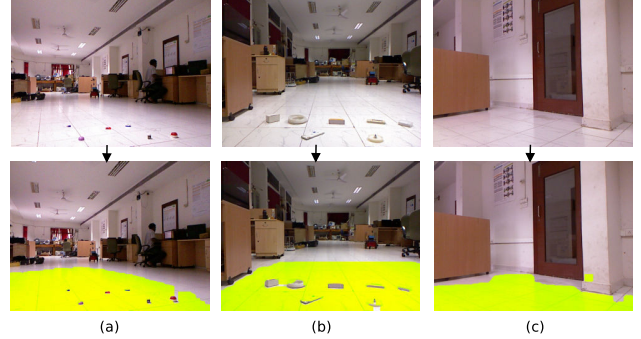
*All authors are with Robotics Research Center, IIIT-Hyderabad. K Madhava Krishna (mkrishna@iiit.ac.in)

Fig. 1. (a) Very low lying obstacles of the order of $0.5cms$ are segmented out. (b) Low lying obstacles which are highly homogeneous with the floor are sparsely segmented. (c) The bottom tile on the wall has the same appearance as floor (homogeneous) that makes it very difficult to segment. The pillar is segmented out till the bottom precisely.
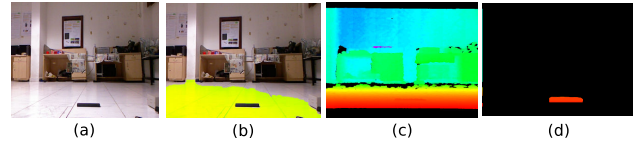

Fig. 2. (a) Very low lying obstacle of the order of $0.5cms$, a battery pack. (b) Segmented image using our algorithm. (c)The corresponding depth cloud of the scene using a Kinect. Note that the depth image of the object is non distinguishable from the background. (d) Depth cloud of the obstacle is segmented out since we know the precise object boundary.

of the scene into its constituent elements like floor, obstacles etc. This is achieved by SLIC Superpixelling based on [1] which would result in division of a given scene along the edges and contours of the floor and obstacles leading to numerous superpixels (figure 7). In addition to this, using Robust Line Segment Detector based on [8] we detect obstacle boundaries in the image that enables discoveries in highly homogeneous settings. A Markov Random Field based Graph Cut with the superpixels as its nodes, formulated based on the homography error of the superpixels and presence of object boundaries provides for optimal classification of scene into obstacle and floor areas. This would prove resourceful in extremely homogeneous environments as in figure 4 and 1(c) where there is dearth of key points for efficient homography estimation.

We show results in several challenging scenarios where there are extremely low lying obstacles as in figure 1(a), (b) and homogeneous environments like 1(b), (c) where existing algorithms under perform. In addition to being robust to changes in floor appearance, our pipeline is clearly able to segment out a large number of small obstacles like cups, boxes, small balls, bottles, bottle lids, etc. (Table I and II), that are commonly found in indoor environments.
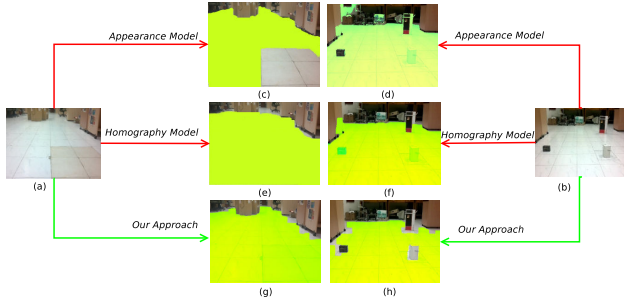
Fig. 3. Results when compared to the performance of Homography and Appearance models. (a) Has a patch that is not detected since it is not similar to the appearance of surrounding floor. (b) Right most image: Has a box which is similar in appearance to the floor. (c) Appearance approach misclassifies a part of floor obstacle due to different appearance. (d) Appearance model misclassifies the obstacle as floor due to similar appearance. (e) Homography model classifies the floor even if appearance changes. (f) Homography model fails when there is a low lying obstacle on the floor. (g) Our approach segments the floor when there is change in appearance. (h) Both the low lying obstacles and the floor which has a different appearance are precisely segmented out.

Obstacles of the height of about $0.5-3cms$ such as a battery pack, a very small box, a pen, small ball or a carpet are distinguished from the floor area. Also, drastic changes in the floor texture and appearance do not affect the performance of our algorithm.

It could be difficult for depth sensors such as Laser Range Finders and Kinect to detect very low lying obstacles based on depth values alone. This can be seen in the depth image of 2c where the battery pack is hardly distinguishable from the background floor. Discovery of such obstacles (figure 2) helps in precisely segmenting out the depth cloud pertaining to them for further uses, which is an added advantage of this algorithm.

II. RELATED WORKS

While obstacle detection shares close ties with segmentation based systems there are indeed prominent differences. Traditional appearance based segmentations could assign different labels to same floor area as appearance changes while end up giving the same floor label to obstacle regions due to strong appearance relationships with floor. Figure 3 depicts the two situations. In figure 3(c) appearance based approaches such as [4] segment the floor area into multiple parts due to changes in appearance, in figure 3(d), the algorithms fuse non floor parts into floor due to similarity in appearance for the image fed from right. Homography based solutions do alleviate a reasonable part of the above problem (figure 3(e)) but are not capable of discerning low lying obstacles (figures 5 and 3(f)).

The problem was first posed in [10] where the author summarizes the difficulty as virtual plane problem where homography solutions are not able to discern obstacles of small heights close to floor. This is depicted in figure 3(f), wherein homography based solutions are unable to segment low lying obstacles. The figure also shows the advantages of the proposed method in figures 4(d), 5(c) and 3(g), (h). Also figure 6 shows the homography error models for floor and
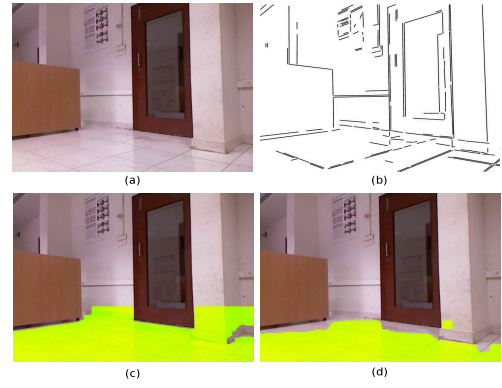

Fig. 4. (a) A highly homogeneous scene. Dearth of key points for homography estimators since a part of the floor tile merges into wall and wall appearance is very similar to floor (b) Vertical line segments detected pertaining to the pillar help in detecting it as obstacle (section III-C). (c) Pure homography based result. Considerable part of the wall detected as floor. (d) Current method discerns the bottom tile on the wall clearly.
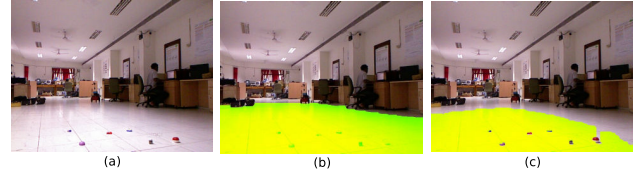

Fig. 5. (a) Extremely low lying obstacles of the height $0.5-3cms$ (b) Traditional homography algorithms cannot distinguish them from floor (c) Result of our algorithm. Low lying obstacles are clearly distinguished.

non floor regions which looks very similar at low heights, shown by the blue for floor areas and orange for non floor areas. It can be seen that it is very difficult to detect obstacles of very low height purely based on homography. There has been a fair mix of papers that have combined geometry and vision such as [6], [3], [7]. However these efforts do not address the problem of segmenting out obstacle regions from floors and do not show results on very small obstacles.

Specifically [6] combined appearance and homography cues in a Bayes Filter formulation to detect low lying obstacles. While effective detection of such obstacles were obtained, the results were hinged to the assumption that appearance of floor regions were repetitive while that of the low obstacles were not. The current formulation does not make any such assumptions about the appearance of floor and non floor regions, grounding its efficacy on purely geometric principles.
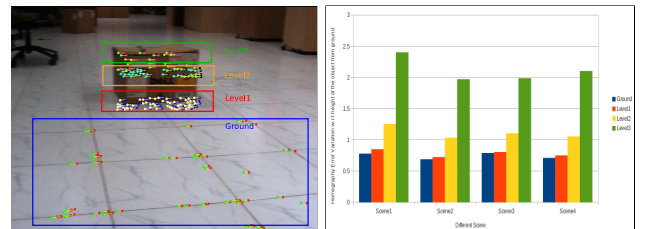

Fig. 6. Homography errors at different height are illustrated. Evidently, it is difficult to detect obstacles of very low height purely based on homography.

III. Our Approach

Our approach stems from the understanding that effective segmentation is possible by not considering temporal relations between tracked pixels but also due to relations that such pixels share with neighbors. Also, it is well known that graphical models developed over a group of pixels that share similar properties is computationally more viable than developing it over individual pixels. These considerations resulted in the pipeline we propose. What could greatly help in this regard is to group the image into different constituent elements which capture local features (figure 7).

The pipeline consists of three principal modules that enable such robust discoveries. Firstly SLIC superpixelling [1] is used in a way that the boundary of the superpixel coincides with the boundary of the small low obstacles. This enables isolating an area where the homography error terms appear distinct with respect to surrounding floor regions despite being at very low height. Figure 1 shows how homography based error terms combined with a graph cut provides for robust segmentation of obstacles at $0.5 - 3cms$ height. Such precise contouring of obstacles however is not possible when the obstacle and floor areas have extremely similar appearance (figure 4). For this we use state of the art LSD, the second module to develop error terms based on line segment features. Finally homography and segment based error terms are dovetailed into an MRF framework, the third module to provide for robust discovery of both low and homogeneous obstacles.

A. Pipeline

(Illustrated graphically in figure 8)

1) The first step of the algorithm is to estimate the ground plane homography H between two images I_n and I_{n-1} using the traditional method of detecting key points and tracking them in two corresponding images (figure 8(a)).

2) Further, image I_n is subjected to SLIC (figure 8(b)) based on [1] to produce a new image Sp_n which contains clusters of pixels called superpixels as seen in figure 7. The boundaries of these superpixels are generally aligned to those of the obstacles present in the image. Also, it is observed that small obstacles would be clustered in a single superpixel or into two ocassionally. A homogeneous environment would produce regular superpixels over the space (handled in the next point). III-B describes the role of superpixelling module in a detailed manner.

3) The same image I_n is subjected to Line Segment Detection(LSD) (figure 8(c)) based on [9] which would lead to detection of numerous line segments that constitute the scene (figure 8(e)). In all of these, the lines that are corresponding to the obstacles are of importance and are filtered as explained in section III-C. Now we have an image Ls_n (ex: figure 9(d)) which would consist of line segments pertaining to obstacles at various positions. The implications of using LSD and the process of selection of particular lines pertaining to objects is clearly detailed in III-C.
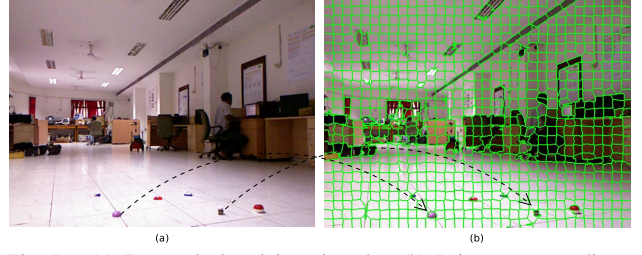

Fig. 7. (a) Extremely low lying obstacles. (b) Points corresponding to a single object clustered into either a single superpixel or a maximum of two.

4) A Markov Random field is formulated with its nodes as the superpixels previously obtained. The energy of a node is determined by two factors. One is the homography error averaged over all the feature tracks within the superpixel Sp_n (figure 8(d)), the other is the presence and deviation of an obstacle line in the Ls_n at the corresponding position of the superpixel Sp in Sp_n. The combined energy of the MRF is minimized using Graph Cut using the standard implementation by Kolmogorov [5] subject to fulfilling the sub modularity criteria. And hence we label each of the superpixels in Sp_n either as floor or as an obstacle as seen in figure 1(a), (b).

B. Superpixelling

Superpixelling using SLIC [1] can decompose an image into small clusters and the separation takes place at the boundaries of the constituent elements as seen in figure 7, 8(e). It is observed that pixels corresponding to small obstacles are clustered into a single superpixel which might ocassionally extend to two. Since we are formulating an MRF over the superpixels, we can use the consolidated homography error over all the tracked pixels of the superpixel (figure 8(d)) and use it in an efficient way for a graph cut which other wise will not prove fruitful in case only the pixels are considered. However SLIC superpixelling has a shortcoming. In case of extremely homogeneous environments, there are no clear boundaries among the floor and non floor regions. In such a case SLIC would produce regular, homogeneous superpixels which would not cater the need here. Section III-C tackles such issues.

C. Detection of line segments of obstacles

Extremely homogeneous environments (figure 4) deprive the homography estimators of key points which means an accurate homography for the floor cannot be calculated. We use LSD based on [9] which completely breaks down the image I_n into its numerous constituent line segments (figure 8(e)) . Even in extremely homogeneous environments, we observe that LSD clearly detects the line segments pertaining to obstacles. But added to these, there are other lines that are present. Obstacle lines that are near vertical as seen in figure 9(c) can be clearly detected. This is done by a two stage filtering. In the first stage, we eliminate all the lines that are not near vertical using a threshold of the angle they make with the horizontal. This leaves us with
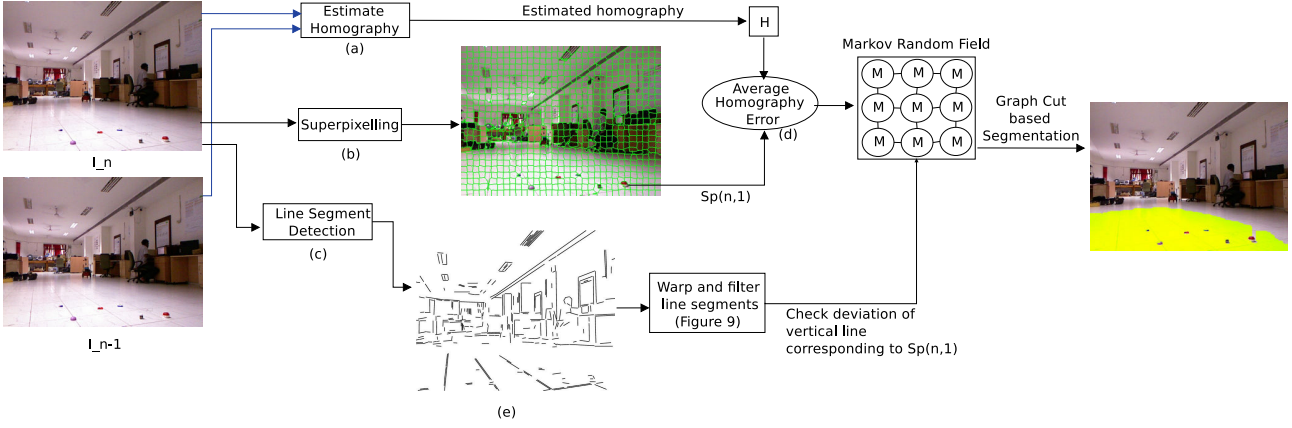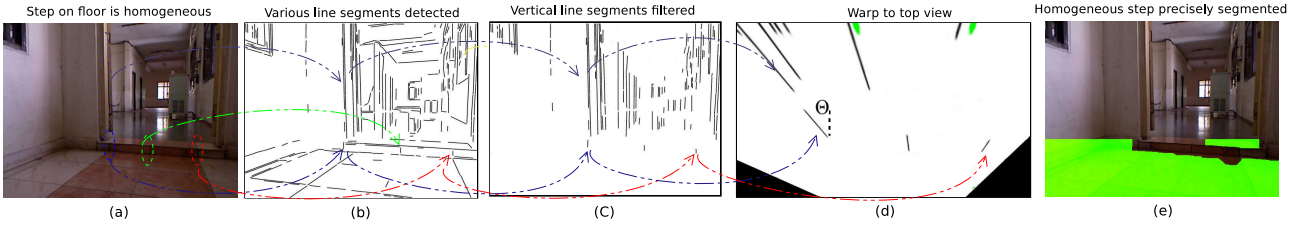
Fig. 8. The algorithm pipeline is illustrated.

Fig. 9. (a) A homogeneous scene where the step is the of the same color as the floor is. Hence it is difficult for traditional algorithms to segment it out. (b) Lines pertaining to different obstacles (the step, pillar beside, etc.) are detected. (c) Near vertical lines are selected for further stages. All the lines which are not near vertical will be dropped. As it can be seen, all the lines of the floor are dropped. (d) A warp of the scene is computed such that a view from the top is obtained. Lines pertaining to the obstacle, in this case the step have a substantial deviation from the vertical, after the warp. These add potential to the corresponding superpixels. (e) The whole step is clearly segmented out by our algorithm.

an image I'_n (figure 9(c)) with two kinds of lines. Obstacles lines which are vertical and floor lines of the similar kind. A warp of I'_n is considered in top view. An obstacle line would show significant deviation from the vertical line in the top view. While typically a floor line that appeared vertical in the perspective view would retain its verticality in the top down view. This holds as long as the camera only yaws but does not pitch or roll. For indoor planar environments this is an obvious situation (figure 9(d)). This warped image is further filtered for lines which have such substantial deviation. In the end, we get lines which are of the obstacles. So even in case of homogeneous scenes, we are able to extract information regarding the obstacles through lines if not through superpixelling and homography. We use these lines to serve in determining the energy of the corresponding superpixel in Sp_n. The presence of an obstacles vertical line in an area would accordingly increase the energy of the corresponding superpixel and hence the corresponding node in MRF. Figure 4(b) has lines pertaining to the pillar which is an obstacle. The lines contribute to the energy of the corresponding super pixels in Sp_n and hence act as subtle discriminators in case of homogeneous environments. There might be cases where the lines on floor might seep in at the end as well but because the floor homography is satisfied by them, they would be classified as floor finally.

D. Homography : Bootstrapping and further stages

The first initial estimate of the homography between I_2 and I_1 is done by a bootstrap process. A trapezoidal area on the floor which does not constitute an obstacle is selected and the homography is estimated. It is expected that clear floor is available for homography estimation in the boot strapping process. After the pipeline produces a segmented image of I_2 we have the actual floor area and hence we would be able to compute the floor homography between I_2 and I_3. This homography would be used to process I_3. And hence the frames I_{n-1} and I_n are used to compute the floor homography which would be used to process I_n.

E. The Markov Random Field Model

A Markov Random Field (MRF) is an undirected probabilistic graphical model to encode conditional dependencies among random variables. We pose our problem of segmenting small obstacles and floor in image in an MRF framework, and define an energy (cost) function such that its minimum corresponds to the target segmented image. In this framework, we represent superpixels of image as nodes in a Markov Random Field and associate a unary and pairwise cost of labeling these superpixels. We then solve the problem in an energy minimization framework where an MRF energy function ψ of following form is defined:

$$\psi(x,\theta,\xi) = \sum_i \psi_i(x_i,\theta_i,\xi_H) + \sum_{(i,j)\in\mathcal{N}} \psi_{ij}(x_i,x_j,\xi_H). \quad (1)$$

In Equation 1 $\psi_i(\cdot)$ represents unary term associated with i^{th} super-pixel and $\psi_{ij}(\cdot,\cdot)$ represent the smoothness term de-

fined over neighborhood system $\mathcal{N}$. Here $x = \{x_1, x_2, ..., x_n\}$ is the set of random variables corresponding to superpixels of image. Each of these random variables x_i takes a label x_i ε [0, 1] based on whether it is a floor or obstacle.

Each super-pixel is checked for vertical lines pertaining to the obstacle. If the vertical line is found in super-pixel, the change in angle with vertical axis by that edge in warped image is computed. For the ith super-pixel this angle is denoted as θ_i. If a super-pixel does not contain a vertical edge then its contribution to the unary term of 2 becomes zero. If ξ_H is the homography error associated with the super-pixel, then the unary term can be defined as,

$$\psi_i(x_i, \theta_i, \xi_H) = (\xi_H^2 + \lambda_1.\theta_i).\bar{x}_i + (\xi_H^2).x_i \qquad (2)$$

Here ξ_H is the average homography error in associated with each of the super-pixel using KLT feature detector and optical flow. λ_1 is a constant. For smoothness term we use Potts model, defined as follows,

$$\psi_{ij}(x, \xi_H) = \lambda_2. \sum_{(i,j)\varepsilon\mathcal{N}} (\xi_{Hi} - \xi_{Hj})^2, if \ x_i \neq x_j. \qquad (3)$$

where λ_2 determines the degree of smoothness. The smoothness term is added only if the neighbouring superpixel has a different label. Once unary and pairwise terms are defined, problem of segmenting small obstacles and floor is now to find the global minima of the energy function defined in Equation 1, i.e.,

$$x* = argmin_x \psi(x, \theta, \xi) \qquad (4)$$

The global minima of this energy function can be efficiently computed by graph cut. For this we construct a weighted graph $G = (V, E)$ where each vertex corresponds to an image super-pixel, and edges link adjacent superpixels. Two additional special vertices source (S) and target (T) are added to the graph. We then connect all the other vertices to them with weighted edges. The weights of edges are defined using definitions of unary and pairwise terms. The min cut of this graph corresponds to the global minima of the energy function. We use publicly available efficient implementation by Kolmogorov and Zabih *et al.* [5] for finding min cut of this graph.

IV. RESULTS

This section outlines the results when the algorithm has been put to test under various challenging scenarios. In all of the cases, the algorithm reliably detects and segments out the floor, demonstrating its robustness and adaptability. The challenges include the presence of small low lying obstacles, homogeneous floor and non floor regions, highly textured floor. Also, apart from different kinds of floor patterns, our method also faithfully segments out obstacles that are not necessarily manhattan in nature. These obstacles include a ball, a bottle, an amigo bot etc. A vast range of indoor objects having various shapes and sizes can be efficiently detected. These include small battery packs, low lying boxes, books, balls, marker pens etc. Following is the summary of the performance in various circumstances.
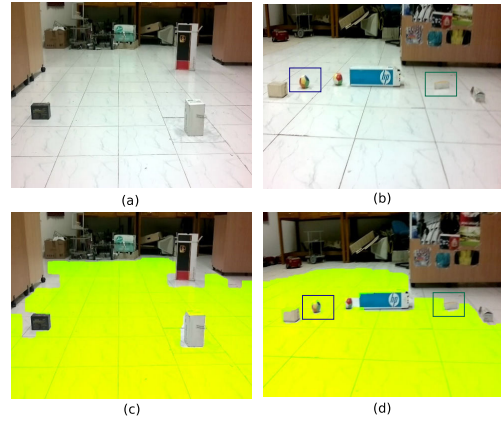

Fig. 10. (a) A general situation where obstacles of considerable height are present. The pipeline faithfully segments out the obstacles. (b) The height of the obstacles is now decreased. Small balls and boxes are segmented out.

A. Small low lying obstacle detection

Extremely low lying obstacles can pose serious problems for indoor robots. Obstacles in indoor environments exist in different shapes, sizes and textures. While an obstacle may be homogeneous to the floor which would not be detected by appearance models, another may be a very low lying one which blindfolds homography models. This section summarizes the results where various kinds of such obstacles are precisely segmented. Essentially, since the obstacles come up in one or two superpixels, the homography error over the superpixel is consolidated over the whole superpixel which is a node in the MRF. And hence the graph cut makes a clear distinction between the obstacle and the floor. We show promising results in various scenarios where obstacles of different kinds are detected precisely. Table I shows the heights and related figures of different obstacles that are detected and the corresponding images. As we pass on to subsequent images, height and size of the obstacles keeps decreasing, making it more and more complex for the algorithm and yet, the pipeline performs faithfully.

TABLE I

HEIGHTS AND DISTANCES(FROM CAMERA) OF OBSTACLES

Object/Obstacle	Height(cm)	Distance(m)	Figure
Pen	0.6	1.4	14(a)-(a1)
Pepper Mint Box	0.7	1.6	13(a)-(a2)
Transistor Battery Pack	1.5	1.65	13(a)-(a1)
Multimeter	1.8	1.6	12(b)-(b1)
Laptop Battery Pack	1.8	1.63	12(b)-(b2)
Wooden Plank	2	4	11(a)
Water Color Bottle	3	1.2	13(a)-(a3)
Mosquito Repellant Case	3.5	1.3	figure 13(a)(a4)
A ball	6	2.2	10(b)

B. Extremely homogeneous obstacles

As discussed earlier, when extremely homogeneous obstacles are present, it is generally difficult to make a good estimate of homography. In such a case if there are vertical lines present on the obstacle, they would rescue the performance of the algorithm. In figure 15(c), it can be seen that normal homographic methods or appearance methods
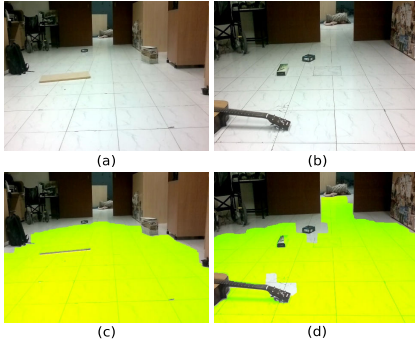
Fig. 11. (a) A plank of height 2*cms* is segmented out. Multiple superpixels are formed over the edge of the plank and each of them which form a node, combine the homography error of the points in them. (b) A guitar fret is precisely segmented.
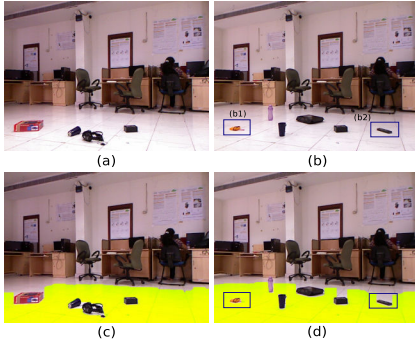

Fig. 12. (a) Low lying obstacles. (b) Points corresponding to a single object clustered into either a single superpixel or a maximum of two.

would fail in such a case where the floor appears to merge into the pillar. In such a situation, the vertical lines present on the pillar, that are detected by LSD would add potential to the energy of the corresponding superpixel in Sp_n and hence they are clearly distinguishable. Also figure 16 shows that the step which is homogeneous as well as low lying is clearly segmented by our algorithm. This is because of the line segments that are detected on the step, which add potential to the corresponding superpixel node. What could be a potential threat to robots in such case is easily handled by our pipeline. We are also able to segment out low lying homogeneous obstacles on the floor faithfully (figure 13(b), (d)). This could be a tough job for traditional homography or appearance based approaches. Table II gives the height of the homogeneous obstacles and their images.

TABLE II
HEIGHTS AND DISTACES(FROM CAMERA) OF HOMOGENEOUS
OBSTACLES

Object/Obstacle	Height(cm)	Distance(m)	Figure
Remote Control	0.8	1.2	13(b)-(b2)
A Notebook	1.5	1.15	14(b)-(b2)
Business Card Box	2	1.5	13(b)-(b6)
Small Wheel	2	1.15	13(b)-(b5)
Duster	3	1.4	13(b)-(b3)
A mesh of wire	3	1.4	13(b)-(b1)

C. Change in texture

Often it so happens that the texture of the floor that the robot is traversing on keeps changing. In this case the algorithm must be robust enough to sustain any changes in
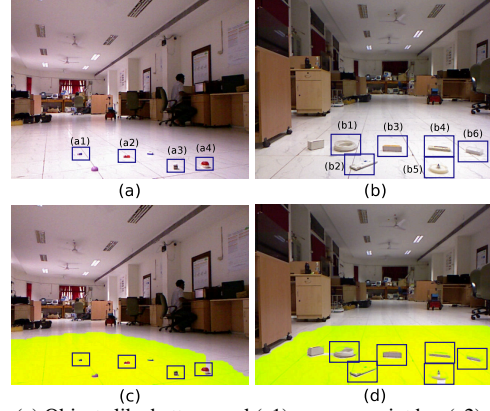

Fig. 13. (a) Objects like battery pack(a1), a peppermint box(a2), water color box(a3), mosquito repellant(a4) are accurately segmented. (b) Extremely low lying homogeneous objects are segmented.
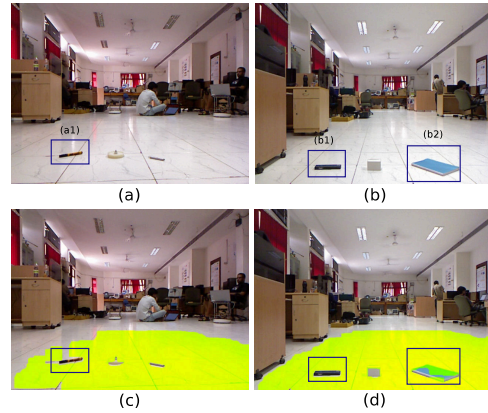

Fig. 14. (a) A pen is precisely classified as an object (b) A notebook which is very low lying is classified as object.
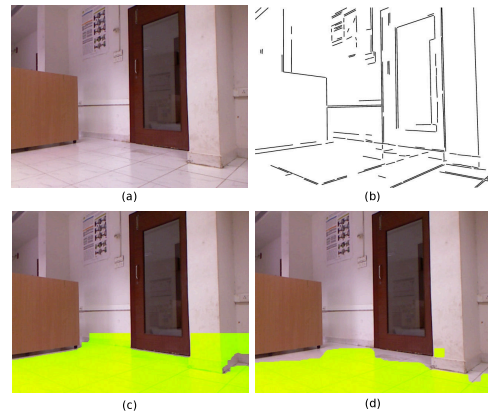

Fig. 15. (a) Bottom tile of the pillar is same as floor. (b) Vertical line segments detected pertaining to the pillar help in detecting it as obstacle as described in section III-C. (c) Pure homography based result. Considerable part of the wall detected as floor. (d) Our method clearly discerns the lower part of the pillar where a tile similar to the obstacles is present.

Fig. 16. (a) A scene where the step appears to be same as the floor and is low lying as well. (b) The step is clearly detected. The floor beyond the step is classified as floor which is expected.
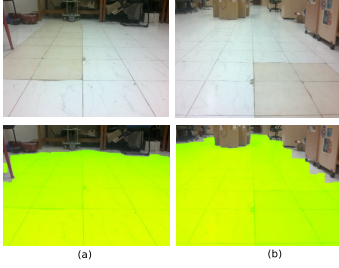

Fig. 17. (a) A visible change in floor texture can be seen. (b) While the floor appearance changes, obstacles are segmented out.

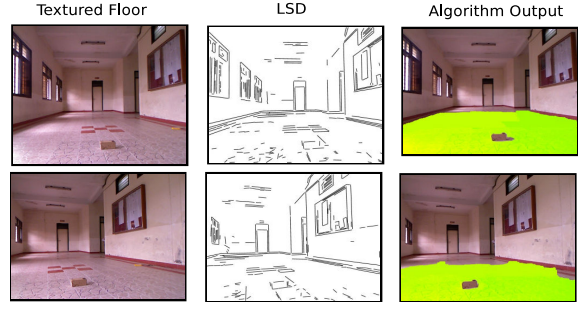Textured Floor   LSD   Algorithm Output

Fig. 18. A Highly textured floor is presented to the pipeline. Despite change in floor texture, the algorithm clearly segments out the floor. The lower part of the LSD image presents a lot of lines in various orientations. Despite their presence, the algorithm segments out the floor faithfully.
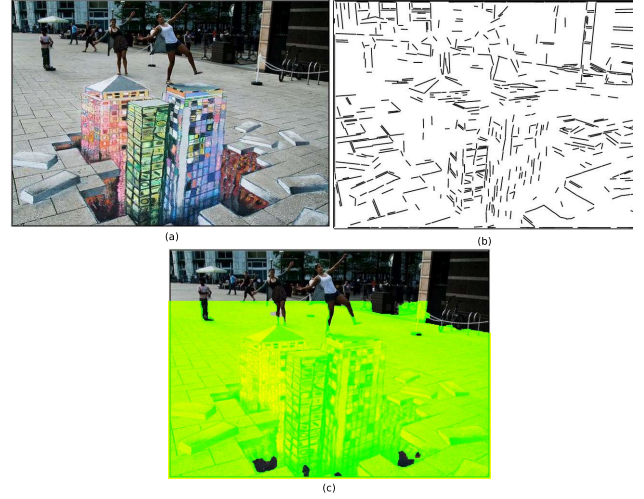

Fig. 19. (a) A 3D painting which could be deceptive to the human eye. The buildings in the front are drawings on the floor and not obstacles. Also since all of it is on the plane, it should segment out the painting as floor. (b) Numerous line segments, some of which are vertical on the floor might deceive the algorithm that there is an object there. But since the superpixels corresponding to those areas satisfy the homography, the potential of that superpixel node in MRF is quite low. Thus presence of vertical lines on the floor would not degrade the performance of the pipeline. (c) The whole of the floor is classified faithfully as expected.

the floor texture. Figures 17 and 18 show that the algorithm can efficiently perform even in cases where the floor texture is changing. Also, in figure 18 it can be seen that there are numerous lines detected in the floor due to the LSD. It might appear that the lines might deceive the algorithm of being associated with obstacle. But this does not happen since all of them are filtered off as specified in III-C, figure 9. In a case where there is a line segment which seeps in after all the process, the superpixel Sp which belongs to the floor would have negligible homography error. And hence a meager potential addition by the line in such a case would not make a difference (Figure 18). We present an interesting case of a 3D painting (figure 19) which generally deceives the human eye as though obstacles are present in the scene. Here the buildings in the front are actually drawings on the floor giving an illusion of standing on the floor. These need to be detected as floor while people behind ought to be classified as obstacles. Also, numerous number of line segments that pertain to the apparent obstacles in the painting might as well deceive the algorithm. But this does not happen. As mentioned earlier, the superpixels satisfy the homography in a clear manner and hence the potential that the lines add to the floor would not affect the performance. This we believe is a convincing case for our algorithm in that it is not in any ways confused by the deceiving segments.

D. Depth map segmentation : Consequent application

Laser Range Finders or RGB-Depth cameras like Kinect provide us with the depth data of the scene. In such a point cloud, small obstacles of $0.5 - 3cms$ in height would pose difficulty in segmenting out the depth points corresponding to them. Our algorithm provides for precise segmentation of the point cloud of such small obstacles. Since we have an accurate segmentation of small objects in the scene, we can select the corresponding area in the depth cloud of the scene

to obtain the depth details of the object. Figure 20 shows the result.

E. Discussion on results

Throughout the results, we demonstrated the performance of our pipeline over various conditions. It can be seen that efficient and precise segmentation of small obstacles is achieved. Thus, this methods provides for effective navigation in indoor environments using a monocular camera while there are numerous small obstacles present. Below, we present a discussion on how different modules contribute to the efficacy of the pipeline.

1) Superpixelling: As discussed earlier (III-B), superpixelling segments out the whole image into its constituent elements along their borders. And hence, small obstacles pop out as single superpixels and occasionally into two. Since we have the obstacles in superpixels, considering the overall homography error of a superpixel using its constituent tracked pixels substantially helps in precisely segmenting the obstacle using the MRF-
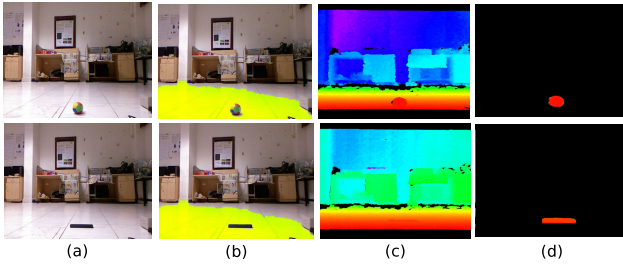
Fig. 20. (a) Very low lying obstacles: a ball, a battery pack. (b) Segmented image using our algorithm. (c)The corresponding depth cloud of the scene using a Kinect. (d) Depth cloud of the obstacle is segmented out since we know the precise object boundary.

Fig. 21. (a) The pipe, across the image and the the base of the chair are low lying and long. (b) Long obstacles are clearly segmented out.

Graph Cut formulation. In such cases, the obstacle is clearly segmented out despite the presence of a vertical line on them. The role of a line is discussed below.

2) Lines and detected line segments : In any given case, there are numerous line segments present in a given scene. Every node of the Markov Random Field is built by the potential contributed by homography error of the superpixel and the presence of a line. But the line essentially contributes in cases where there is extreme homogeneity. Homogeneous cases as seen in figure 15 and 16 pose problems in estimating the homography due dearth of key points. It can be seen that the floor apparently merges into the pillar (figure 15) and the step (figure 16). In such cases the vertical line segments of obstacles that are detected (described in III-C) in the scene come to the rescue. Line segments detected in section III-C (figure 9) add potential to the corresponding superpixel node and hence a clear distinction can be made in the Graph Cut process. It might some times appear that the lines present on the floor could deceive the algorithm of being associated with an obstacle. But the strict two stage filtering described in III-C helps in selecting only the vertical and near vertical lines of obstacles. If in case there is a floor line which seeps in through both the filters, the homography error of the superpixel of that corresponding line is extremely low, and hence it would be labeled as floor by the Graph Cut process. This can be seen in figures 18 and 19.

3) Further, there could be cases where the whole environment is so homogeneous that the Line Segment Detector is not able to detect near vertical lines of the obstacles. Such cases we believe are very rare and there are no current algorithms which cater to such problems.

4) Long Obstacles and Rotation: When an obstacle is long all of it is not in a single superpixel. However since the contours of superpixels on the obstacle coincide with the it, homography error terms provide sufficient potential for object discovery (figure 21). If rotations are abrupt, trackers tend to provide wrong correspondences. A wrong homography estimate can produce less than optimal performance. However if the rotation is smooth and not abrupt, the algorithm continues to do the needful.

V. CONCLUSIONS

The paper presents an efficient algorithm for segmenting out low lying obstacles of the order $0.5 - 3cms$ in height just by using a monocular commodity camera. In addition, extremely homogeneous environments where the obstacles appear to merge with the floor are dealt in an efficient manner. Consequently, this paper leads to a drastic increase in the kind of obstacles that can be detected by present obstacle detection and floor segmentation algorithms. In a homogeneous environment, a robot might ram into a non floor region (figure 16(b)) in case the pillar is not segmented out properly. Also, a humanoid might topple in case there are small obstacles lying on the floor (figure 13). This paper alleviates such issues and a considerable number of problems that are generally faced by current algorithms. Apart from the issues it addresses, the pipeline could be used for further applications to precisely collect the point cloud data, better navigation and path planning, etc. To the best of our knowledge there has not been a paper catered to this problem that shows similar results of such small and diverse obstacle discoveries.

REFERENCES

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. Slic superpixels. *École Polytechnique Fédéral de Lausssanne (EPFL), Tech. Rep*, 149300, 2010.

[2] C.-K. Chang, C. Siagian, and L. Itti. Mobile robot monocular vision navigation based on road region and boundary estimation. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 1043–1050. IEEE, 2012.

[3] E. Fazl-Ersi and J. Tsotsos. Region classification for robust floor detection in indoor environments. *Image Analysis and Recognition*, pages 717–726, 2009.

[4] P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.

[5] V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, 2004.

[6] S. Kumar, A. Dewan, and K. M. Krishna. A bayes filter based adaptive floor segmentation with homography and appearance cues. In *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing*, page 54. ACM, 2012.

[7] J. Rituerto, A. Murillo, and J. Kosecka. Label propagation in videos indoors with an incremental non-parametric model update. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 2383–2389. IEEE, 2011.

[8] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall. Lsd: A fast line segment detector with a false detection control. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(4):722–732, 2010.

[9] R. G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall. Lsd: A line segment detector. *URL http://www. ipol. im/pub/algo/gjmr line segment detector*, 2012.

[10] J. Zhou and B. Li. Robust ground plane detection with normalized homography in monocular sequences from a robot platform. In *Image Processing, 2006 IEEE International Conference on*, pages 3017–3020. IEEE, 2006.