# Functional Room Detection and Modeling using Stereo Imagery in Domestic Environments

Karthik Mahesh Varadarajan and Markus Vincze

***Abstract*—In situ functional detection and classification of rooms in indoor environments is an important aspect in the training and map building stages of deployment of robotic assistants in home environments. Traditional place learning methods do not perform functional room or unit identification. Explicit user labeling of places as well as map editing is required for practical applications. Alternative place learning schemes use feature based methods to detect typical objects and hypothesize room functionality/ learn places based on localization of these objects in the map. Besides extensive user intervention, these methods are unsuitable in dynamic environments or in unoccupied/ unfurnished homes. Furthermore, traditional indoor 3D structural environment modeling algorithms employ schemes such as clustering of dense point clouds for parameterization and identification of the 3D surfaces. RANSAC based plane fitting, extensions to feature based stereo, half-plane detection, real-plane or facade reconstruction, plane sweeping etc. have been used for 3D environment reconstruction. Noise in the range data, especially in low texture regions and accidental line/plane grouping under lack of cues for visibility tests can hamper efficiency of practical systems. In order to counter these issues, we propose a novel framework fusing 2D local and global features such as edges, texture and regions, with geometry information obtained from range data for reliable 3D indoor scene representation. The algorithm is shown to perform superior to RANSAC based plane fitting approaches. Functional room boundary detection and modeling is carried out using cues from the number of detected doorways and open boundaries. By avoiding the use of feature based place learning, robustness and versatility of the scheme is improved.**

## I. Introduction

Functional detection and classification of rooms in indoor environments is an important aspect in the training and map building stages of deployment of robotic assistants in home environments. Conventional methods for environment learning such as SLAM do not perform semantic functional room or unit identification. Furthermore, explicit user labeling of places as well as map editing is required for practical applications. There are also several different methods for place learning. These include schemes that use feature based methods to detect typical objects and hypothesize room functionality/ learn places based on localization of these objects in the map [31]. The RobotVision @ ImageCLEF challenge [30] is specifically targeted at place learning. Besides extensive user intervention, all these methods are unsuitable in dynamic environments or in unoccupied/ unfurnished homes. Other semantic cues based exploratory algorithms for map building have been presented in [32] and [33], but lack the direction towards the development of functional semantic definitions for rooms in domestic environments.

With regard to indoor 3D structural modeling, traditional methods employ schemes such as clustering of dense point clouds for parameterization and identification of the 3D surfaces. RANSAC based plane fitting [1] is one common approach in this regard. Alternatively, extensions to feature based stereo have also been used, half-plane detection [2,3,4], real-plane or facade reconstruction [5], plane sweeping etc. have been proposed. Recent efforts at plane grouping based on PCA and visibility tests include [6] and [7]. Other techniques include line grouping [8] and model based recognition [9]. However, the performance of most of these techniques rapidly degrade in the presence of high amounts of noise (in range data such as stereo) under conditions of low illumination and in regions of low-texture or sparse features. Furthermore, accidental line/plane grouping (due to shelves/ cupboards), especially under lack of cues for visibility tests, presence of depth edges or discontinuities that are not visible in the 2D image and difficulty in adaptively estimating metrics for clustering can hamper efficiency of practical systems for door/doorway detection. On the other hand, traditional laser [10] or panoramic camera based [11,12] (multi-view) room modeling and doorway detection systems (often using piecewise planar modeling [13], triangulation [14] or space carving [15]) are often impractical for cost-effective domestic robots. Moreover, machine learning based door recognition systems (usually from only 2D images) such as [16,17,18], perform poorly in cluttered scenes (especially with floor reflectance) and when the door is open, is viewed partially or the doorway is structurally similar to an arch, lacking the actual door or door frame. Depth based doorway detection is more practical in such cases and also provides cues for place learning.

In order to resolve these challenges, we extend out framework presented in [29]. We fuse 2D local and global features such as edges, textures and regions, with geometry data obtained from pixel-wise dense stereo for reliable 3D

Manuscript received April 2 2011. The research leading to these results has received partial funding from the European Community's Sixth Framework Programme (FP6/2003-2006) under grant agreement no FP6-2006-IST-6-045350 (robots@home).

Karthik Mahesh Varadarajan and Markus Vincze are with the Vienna University of Technology, Automation and Control Institute, Gusshausstrasse 30 / E376, A-1040 Vienna, Austria (email: {kv,mv} @acin.tuwien.ac.at).

indoor scene representation. The strength of the approach is derived from novel depth diffusion and segmentation algorithms resulting in superior surface characterization. Unlike earlier schemes, our methods enable identification of depth edges critical to surface isolation, while rendering visibility tests superfluous.

The proposed framework follows a three tier process – detection of walls, followed by the enclosing room and finally doorways. Walls and wall-like surfaces are detected using 2D edge, texture and region features. The 3D surfaces corresponding to the walls are then generated using piecewise depth diffusion techniques followed by depth segmentation to identify intra - wall depth discontinuities. The room model is built by selecting wall-like surfaces to fit approximate cuboidal constraints. Doorways in the room are estimated by clustering of the dense stereo data pixels that do not conform to the concave room hypothesis. Finally, room functionality is determined based on number of doorways, room area and topology.

## II. Overview

This paper offers a number of novel contributions. The main contributions are listed below. Firstly, this paper presents an innovative framework for functionality based room boundary detection. Secondly, this paper offers a scheme for room functionality determination and place learning based on structural features of the room. Thirdly and most significantly, this paper offers an integrated framework for complete functional room modeling from stereo range data.

The images used for evaluating the developed algorithms have been obtained in an indoor environment, from an experimental robot at a height of about 1m from the ground plane. A high dynamic range monochrome stereo camera is used to estimate the range images along with a centrally mounted inexpensive color camera. Note that the algorithms presented here are well suited for fusion of data from distinct color and range (stereo or otherwise) sensor systems. In order to simplify the algorithmic framework, it has been assumed that the fixed pose of camera and its height above the ground plane are known accurately, thereby establishing the approximate ground plane in any scene without further processing. The approach for 3D room reconstruction and doorway detection presented in this paper follows a 3 stage modeling pipeline comprising of wall modeling, room modeling and doorway modeling. The various assumptions for the modeling/ hypothesis at each stage are:

### *A. Wall Modeling*

Walls are typically characterized by

1. Homogeneous regions or areas with regular texture, usually with high numeric intensity values.
2. Largest single color regions in a given scene, especially with no large occluding obstacles in the vicinity.
3. Hold pixels with the farthest visible range information on planes parallel to the ground plane.
4. Frequent loss of homogeneity in color values owing to lighting and shading effects.

### *B. Room Modeling*

Rooms are characterized by

1. Combination of walls approximating a cuboid.
2. Largest and most consistent of all possible cuboids in the scene (helps exclude walls internal to the room).
3. Often encompasses all extreme range pixels in the horizontal dimensions (along the image width).
4. Room fitting can be reduced in most cases (based on assumptions of known floor and ceiling) to fitting of a maximum of just three (largest) vertical walls.

### *C. Doorway Modeling*

Doorways are characterized by

1. External outliers (or exclave points in range images) to the room model, that can be grouped to form regions with size bounds similar to that of typical doorways.
2. These outliers should be at a jump discontinuity to the modeled room surfaces.
3. Floors are typically uniform across doorways.

## III. Algorithm

The algorithmic pipeline presented in this paper follows from the above sequence of modeling. The framework has three main sections. The first section deals with color image processing, wherein after pre-processing, reflectance image gradients are extracted from the 2D image and segmentation (along with region selection) is carried to identify walls and wall-like regions. The second section details dense stereo depth data processing in a number of steps that include denoising, piecewise diffusion to reconstruct depth surfaces and depth segmentation to identify intra-object depth bounds. The last section deals with functional indoor structure generation by fitting planes to the wall-like surfaces and grouping them to find room boundaries followed by room functionality hypothesis. This 3D reconstruction of the room leads to detection of doorways and other negative spaces.

### **Color Image Processing**

### *A. Color Pre-processing*

The color image, obtained from the centrally located camera is rectified and used as the reference image. As a pre-processing step, the noise in the color image is reduced using a bilateral filter that preserves salient gradient values and hence sharp edges that are crucial for algorithms in the following stages of processing, including 2D segmentation.

### *B. Intrinsic Reflectance Gradients Extraction*

The gradients of the filtered color image are estimated and these gradients are decomposed into shading and reflectance components. The shading component captures the lighting and shadows in the scene while the reflectance component captures the distinction in the material surfaces. This step is helpful to eliminate the highlights and shadow patterns created by light fixtures typically mounted on walls. The algorithm we employ is based on the intrinsic image extraction algorithm developed by Weiss [19] and extended by Tappen [20]. In the presented framework, gradients in the intensity channel of the color image are classified as 'shading' or 'reflectance' gradients by modeling an asymptotic linear color variation across neighboring pixels. The formulation for intrinsic image extraction [20] is

$$I(x,y) = S(x,y) \text{ x } R(x,y) \quad (1)$$

where $S(x,y)$ is the shading image, $R(x,y)$ is the reflectance image and $I(x,y)$is the input image defined in the dimensions $x$and $y$. Using a logarithmic transformation and applying multiple scale selective gradient/ derivative filters $f_x, f_y$ we have the gradient images $F_x$ and $F_y$, the $(x,y)$ components of which can be classified as shading if the color

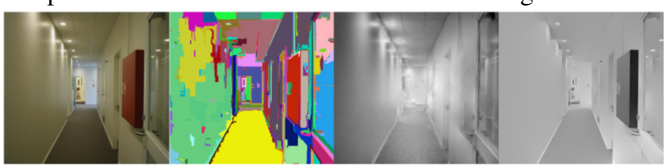

**Figure 1**. Intrinsic Image Extraction and Segmentation (A) Input color image (B) Segmentation using the standard Felzenszwalb-Huttenlocher (FH) graph based algorithm – demonstrates high clutter in regions of the left wall with lighting changes (C) Shading intrinsic image (D) Reflectance intrinsic image – note that C and D (obtained by inversion of input image gradients classified as shading or reflectance respectively)

pixels satisfy the constraints $c_{x+1} = \alpha c_x$ and $c_{y+1} = \alpha c_y$ respectively and as reflectance otherwise. The component images can be reconstructed as

**Figure 2.** (A) Segmentation on the input image using a low complexity multi-scale full gradient edge analysis scheme (B) Segmentation on the input image with the same scheme using reflectance-only gradients – shows superior performance in wall regions affected by lighting changes in comparison with full-gradient image segmentation schemes such as the graph based FH. Similar values of gradient and region size thresholds were used for all three segmentation scenarios. (C) Results using FH (mislabeled pixels: 70292) (D) Results using our framework (mislabeled pixels: 32069)

$$C(x,y) = g * [(f_x(-x,-y) * F_{cx}) + (f_y(-x,-y) * F_{cy})] \quad (2)$$

where, * represents convolution, $F_{cx}$and $F_{cy}$are component (shading/ reflectance) gradients and $g$ is obtained from

$$g * [(f_x(-x,-y) * f_x(x,y)) + (f_y(-x,-y) * f_y(x,y))] = \delta \quad (3)$$

The shading and reflectance components as defined by equation (2) are shown in Fig. 1C and 1D. In our framework, the reflectance image gradients $\mathbf{F_{cx}}$ and $\mathbf{F_{cy}}$ are used directly in the segmentation process. One possible disadvantage of using intrinsic gradients is that gradients at edges pertaining to surface orientation changes in walls and other structures may not be captured in the reflectance component. However, this is limitation is overcome in the presented framework as the additional step of depth segmentation detects these gradients, from the depth image.

### *C. 2D Reflectance Gradient based Segmentation and Region Isolation using Texture Analysis*

Using the gradients $F_{cx}$ and $F_{cy}$ obtained in the previous stage, segmentation is carried out using a low complexity multi-scale edge analysis scheme. The scheme links edges found at various scales (by analysis of reflectance only gradients) using proximity and similarity measures to form enclosed regions or segments. The choice of the segmentation algorithm is based on the goal of meeting real-time constraints for deployments on robots, which excludes the possibility of using algorithms like the Felzenszwalb-Huttenlocher (FH) graph based algorithm. It can be seen from Fig. 2C and 2D that the output of the proposed 'reflectance gradient only' segmentation approach is superior to traditional full-gradient algorithms like FH (with gradients as grid graph edge weights) given the given context of wall detection with lighting and shading changes. The segmented regions are then subjected to a region selection algorithm to select walls and wall-like structural surfaces that are expected to support the room model. The characteristic features of walls such as low texture, high homogeneity, large pixel spans and representations using high gray-scale intensity values are used in region selection. The current framework employs 2 levels of thresholds (hard and soft) on measures of entropy (E), homogeneity (H), uniformity energy (U), correlation (R), contrast (C) and other constraints based on the Grey Level Co-occurrence Matrix (GLCM) to select wall-like surfaces. Estimated soft threshold values, along with the assigned confidence values of the measures on condition conformance (in brackets) for the two-class separation (positive wall classification) are H>0.99 (1.0), C<0.0275 (1.0), R>0.9 (0.9), U>0.6 (0.3) and E<5.5 (0.8). The hard threshold values are H>0.96 (0.8), C<0.0475 (0.7), R>0.85 (0.6), U>0.3 (0.1), E<7.0 (0.5). The surface is classified as a wall if the aggregate confidence value exceeds 3.0 out of a maximum of 4.0. On a representative data set of 80 image chips of various material textures found indoors, such as wood, tile, brick, rock, vegetation, carpet, cloth, curtain, steel, bronze, tree bark, granite etc., besides painted wall surfaces, the simple classifier achieved a classification rate of 95%, with a wall detection rate of 97.87%. Thresholds on pixel spans of the surfaces (> $I_w$*$I_h$/15, where, $I_w$ is image width and $I_h$ is image height) and average gray-scale intensity (> 100/255), further help reduce the detected segments to the set of primary room surfaces. The region masks thus obtained (Fig. 2D) are used for piecewise isotropic depth diffusion.

**Stereo Depth Image Processing**

*D. Depth Pre-processing*

Depth pre-processing involves noise removal done using a novel sparse de-noising algorithm, employing iterative hysteresis filtering & morphological reconstruction [29].

*E. Depth Diffusion*

Since the input depth map from the range sensor can be quite sparse, it is required to convert it into a dense cloud for reliable and coherent surface estimation. This step is important since the span of the surfaces (in terms of pixels) is crucial for reliable weighting in the fitting and room reconstruction process and for outlier rejection. Diffusion of depth values is carried out using a Piecewise Isotropic Laplacian Partial Differential Heat Linear Equation (PDE) Solver that operates only in regions identified by masks obtained in step C. By combining Multi-grid and Iterative Back Substitution (IBS) schemes to solve the PDE equation, rapid convergence is obtained.

The PDE representing the flow of heat in a 2 dimensional isotropic medium [21] is given by

$$\frac{\partial u(\mathrm{r},t)}{\partial t} = c\left(\frac{\partial^2 u(\mathrm{r},t)}{\partial x^2} + \frac{\partial^2 u(\mathrm{r},t)}{\partial y^2}\right) \quad (4)$$

where, $u(\mathrm{r},t)$ represents the heat measured in the two dimensional space r(*x*,*y*) at time *t*.

Traditional isotropic diffusion solvers smooth out edge regions, while direct application of anisotropic diffusion to depth data smoothens depth edges in regions where image gradients are weak (e.g. wall intersections). In the piecewise isotropic diffusion solver, the calculation of the forward and backward substitution modules is suppressed for known depth pixels, thereby propagating and preserving segment boundaries as well as depth edges across iterations (Fig. 4).

While the IBS solution is reasonably fast (of the order of 0.5 sec on a 3.2 GHz single core PC with 512 MB RAM, for a 320x240 depth image), the convergence rates are to be further enhanced for real-time operation on resource constrained systems. In our framework, we use a variant of the multi-grid approach (that employs solutions to equation systems at smaller scales as pre-conditioners for higher scales) to speed-up calculations of the IBS.

*F. Depth Segmentation*

An additional step of depth segmentation is necessary to detect depth discontinuities and hence surface boundaries that are not captured in 2D edge segmentation. A good example is the case of a discontinuity in a wall surface as a result of a pillar or column like structure or a depth edge created at the intersection of two wall faces of a room. Since the faces of the room are expected to be of the same color, it is possible that a reliable edge is not detected at the junction of these faces or at locations of surface orientation changes on a column during color processing. As explained in the previous section, our novel diffusion step renders these edges detectable and regions separable to segmentation. In our approach, the simple, low-complexity multi-scale edge detection and linking approach explained in section C is sufficient for intra-object (here intra-wall) depth discontinuity. This approach is chosen to keep computation requirements low. This also removes noisy depth surfaces.

**Functional Room Detection**

*G. Surface Fitting*

The detected wall-like depth segments are then fit to planar surfaces. This process helps parameterize the depth surfaces, rendering surface orientation analysis easier. All surfaces that do not conform to planar constraints are eliminated based on quality of fit. Since walls are expected to satisfy Manhattan constraints and the floor plane is approximated to be perpendicular to the image plane, all depth surfaces that are not perpendicular to the floor plane (within tolerance limits) are also excluded from further analysis.

The plane fitting is carried out using Iteratively Re-weighted Least Squares Robust Linear Regression. To prevent the effect of propagation of errors to the 3D planar coordinates - *X* & *Y* (ideally independent variables) from the depth coordinate $Z$ (ideally dependent) during point cloud estimation; the 3D fitting is carried out using a reprojected equation (12) in the image plane. The equation is solved using a transformation of variable (1/Z) to a temporary reciprocal variable, with *x* and *y* being image coordinates.

$$\frac{1}{Z} = \left(\frac{A}{f_x D}\right)x + \left(-\frac{B}{f_y D}\right)y + \left(\frac{C}{D} - \frac{A(1-C_x)}{f_x D} + \frac{B(1-C_y)}{f_y D}\right) (5)$$

where, *A*, *B*, *C* and *D* are the true plane equation coefficients in the 3D world, $f_x$ and $f_y$ are camera focal lengths, while $(C_x, C_y)$ is the principal point (The additional negative sign is due to an inverted *Y* reference system).

*H. Functional Room Boundaries Detection*

The depth surface planes are projected onto the ZX plane and PCA is used to find the principal orientation of each wall like surface. The planar surfaces are then classified by orientation and mapped to a cuboidal structure (with angular deformation – to permit some deviation). Planes that do not support the cuboid hypothesis are rejected. The choice in the usage of a convex shape (cuboid) is significant. In the case of domestic environments, it can be expected that barriers in the environment represent boundaries of functional separation. One example would be the case of the kitchen, which is usually separated from the hall by a simple barrier and not a doorway. Furthermore, dining halls typically sections of the hall that are separated from the functional hall area by an open boundary, i.e. while the functional hall area forms a convex shape such as a cuboid or rectangle (on the plan), the dining area forms a separate cuboid adjoining it. Planes are ranked based on consistency, span, texture content and the

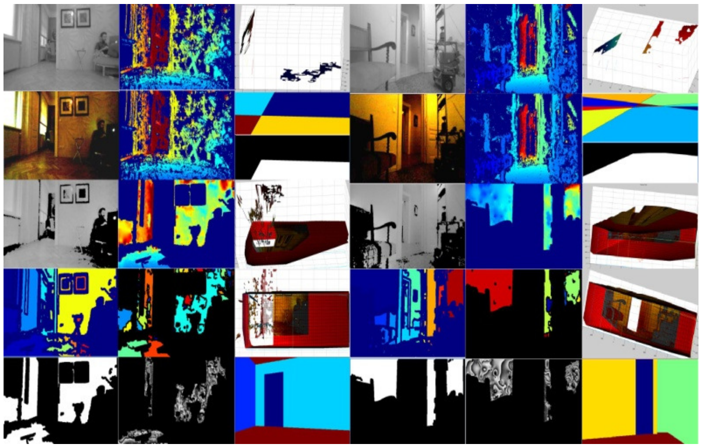

**Figure 3.** Complete Algorithmic Pipeline for 2 input scenes (columns 1-3 and 4-6) (Top to bottom) Left Row – Color image processing: (A) 2D input monochrome left image (B) 2D color image (C) Reflectance image (D) Image segmentation output (E) Region selection output (of wall-like structures) Middle Row - (A) Input depth map (B) Filtered depth map (C) Depth diffusion output (D) Depth segmentation output (E) Masked depth map Right Row - (A) Surface fitting results (near top view) (B) Surface categorization & PCA based boundary detection (C) 3D Reconstruction (top view (D) 3D Reconstruction (front view) (E) 3D map reprojected to camera plane

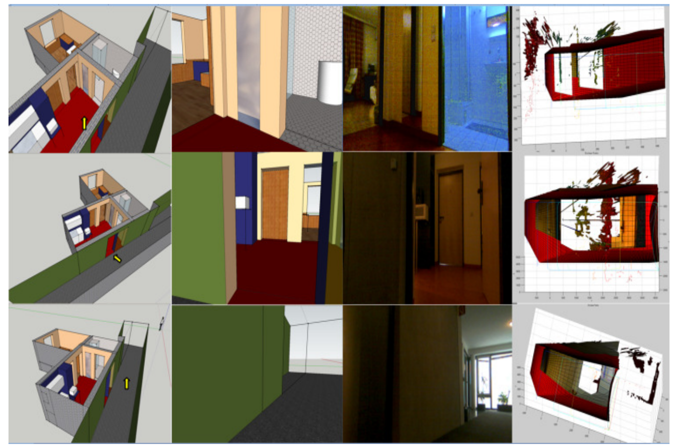

**Figure 4.** 3D Room reconstructions for 3 scenes (row-wise) (Left to right) (A) 3D ground truth model of room along with camera viewpoint (B) Synthetic image from camera viewpoint (C) Input scene (D) Generated 3D model of the room with doorways denoted by cavities in the model leading to 3D data points from the room observed through the doorway. For the 3 cases, the functional labels output by the system are (1) hall and adjoining rooms as bedroom and bath (2) unknown (3) corridor

degree of meeting Manhattan constraints. Higher ranked planes are preferred (in a rule based framework) for the room boundary establishment. This permits use of wall-like surfaces (doors at consistent orientations and cupboards, with suitable contention resolution) for approximating the room structure whenever the current camera viewpoint does not contain significant wall surfaces. The active room sector is also identified by the scheme. This framework renders any visibility tests and constraints superfluous. This is because the depth diffusion uses the values of all know depth pixels (noise suppressed) to build the depth surfaces, irrespective of the curvature and the depth segmentation step breaks those surfaces that would not have satisfied visibility constraints. Also, the PCA ensures that the most consistent and visible surfaces are used in the room modeling process (Fig. 3).

### *I. 3D Room Reconstruction and Doorway Detection*

Using the detected room sector map, height of the camera above ground and standard room height measurements, the 3D structure of the room is reconstructed. Doorways are detected by clustering depth pixels that do not support the concave room structure hypothesis. Typical measurements of doors are used to improve localization of doorways. These doorways are modeled as open regions in the 3D representation with the exclave pixels (those belonging to the room seen through the doorway) as sparse 3D points. Bounds on the depth discontinuity ranges (between current room boundaries and exclave pixels) help identify pixel surfaces in rooms beyond doorways making the scheme robust to presence of cupboards and other enclosures, leading to high recognition rates for true doorways (Fig. 3 and 4).

### *J. Room Functionality Hypothesis (Place Learning)*

Based on the number of doorways and the topology of the scene, it is possible to hypothesize the functionality of the room in the domestic environment. As stated earlier, the functional room boundary detection module estimates rooms or room sections that are expected to provide functional separation from other areas in the domestic environment. The categorization of these functional areas is then carried out using size of the area, number of doorways, number and size of open boundaries. Large areas with more than one door are categorized as halls. Corridors are identified based on parallel walls with low numerical values for the distance between the walls and open boundaries in orthogonal directions. Kitchens/ dining halls are identified as areas adjoining living area with open boundaries leading to the hall (categorized based on distance of open boundary). Thresholds on areas and depth regions with a single doorway help categorize into bedrooms and bath. It should be noted that it is also possible to yield room functionality labels to areas that are observed beyond the doorways based on this scheme. While the current framework uses a single-shot image frame processing to yield a response on the room functionality type, it can be extended to use metric maps of SLAM or topological maps to refine the hypothesis on the room type. The reliability of the single-shot scheme declines when the robot is close to the doorway. Multi-frame processing or full 3D analysis based on the metric maps can yield reliable solutions in such cases. Results of room functional analysis are presented for various scenarios in Fig. 4. For the 3 cases in Fig. 4, the output labels obtained from the system are (1) hall and adjoining rooms as bedroom and bath – the classification between bedroom and bath has been made based on the depth of points through the doorway (2) unknown – since the doorway is too close to the camera (3) corridor – since 2 parallel wall surfaces are obtained with open boundaries orthogonal to the walls.

## IV. ANALYSIS

The results presented in the previous section demonstrate the robustness of the framework. This framework outperforms the traditional RANSAC based plane fitting and room boundary detection algorithm that uses the output of our surface segmentation approach. The pixel mislabeling error is 5 times higher for RANSAC. Our algorithm is shown to be robust for a variety of complex scenes. Fig. 3, 4 and 5 describes robust and consistent performance in large environments with multiple doorways and heavy clutter such as cupboards and closets. Results are presented in Figure 5 in the form of reprojections of the generated 3D models on to the image plane. Dark blue and red regions denote doorways and the floor, while walls are labeled in other colors. The reprojection error, measured here in terms of a rough metric - number of mislabeled pixels (with respect to manually labeled doorway, floor and wall region pixels) is about 5% in typical scenes and exceptions are due to large and dynamic occlusions (such as humans).

## V. CONCLUSION AND FUTURE WORK

In this paper, we have demonstrated an innovative framework for functionality based room boundary detection and place learning in domestic environments. This paper also offers an integrated framework for complete functional room modeling from stereo range data. The strength of the approach is derived from the novel depth diffusion and segmentation algorithms that result in better surface characterization as opposed to traditional feature based stereo or RANSAC plane fitting approaches. It should also be noted that in the context of indoor 3D room reconstruction, the presented framework is (a) highly efficient with extremely sparse range data (b) preserves and detects depth edges in regions where there are no visible edges in the color data and (c) handles shadows and specular highlights effectively, unlike other related sensor fusion schemes like [27] and [28]. While simple semantic features based on doorways and area have been used for room functional analysis, improvements to the model based on other features forms future work. Integration of the system into a full SLAM framework is ongoing work.

## REFERENCES


[1] A. Bartoli, "A random sampling strategy for piecewise planar scene segmentation", Computer Vision and Image Understanding, 2007.

[2] C. Baillard, C. Schmid, et al., "Automatic line matching and 3D reconstruction of buildings from multiple views", ISPRS 1999.

[3] C. Baillard and A. Zisserman."A plane-sweep strategy for the 3D reconstruction of buildings from multiple images". ISPRS, 2000.

[4] C. Schmid and A. Zisserman."Automatic line matching across views". CVPR, 1997.

[5] S.C.Lee and R. Nevatia."Interactive 3D Building Modeling Using a Hierarchical Representation", ICCV2003.

[6] JK. Lee, S. Ahn et al., "A Prospective Algorithm for Real Plane Identification from 3D Point Clouds and 2D Edges", ICHIT '08.

[7] JK. Lee, S. Ahn et al., "Visibility-Based Test Scene Understanding by Real Plane Search", Advances in Visual Computing, Springer, 2008.

[8] C. Guerra et al."Line-based object recognition using Hausdorff distance from range images to molecular secondary structures" IVC05

[9] SH. Chang, S. Lee, D. Moon, et.al,"Model based 3D Object Recognition using Line Features", ICAR 2007.

[10] P. Biber, et al., "3D Modeling of Indoor Environments by a Mobile Robot with a Laser Scanner and Panoramic Camera", IROS 2004.

[11] P. Doubek, T. Svoboda, "Reliable 3D reconstruction from a few catadioptric images", OMNIVIS 2000.

[12] M. Nevado, et al. "Obtaining 3D models of indoor environments with a mobile robot by estimating local surface directions", RAS 2004.

[13] A. Dick, R. Cipolla et al. "Combining Single View Recognition & Multiple View Stereo for Architectural Scenes", ICCV 2001

[14] DD. Morris, T. Kanade, "Image Consistent Surface Triangulation", CVPR 2000.

[15] K Kutulakos, "Approximate N View Stereo", - Lecture Notes in Computer Science, 2000

[16] AC. Murillo, J. Kosecka, et al.,"Visual door detection integrating appearance and shape cues", RAS 2008.

[17] Z. Chen, S.T. Birchfield, "Visual Detection of Lintel-Occluded Doors from a Single Image", CVPRW 2008.

[18] R. Munoz-Salinas, E. Aguirre, M. Garcıa-Silvente, "Detection of doors using a genetic visual fuzzy system for mobile robots", Autonomous Robots, 2006.

[19] Y. Weiss, "Deriving intrinsic images from image sequences", ICCV, 2001.

[20] MF. Tappen, WT. Freeman, EH. Adelson, "Recovering intrinsic images from a single image", PAMI 2005.

[21] P. Perona and J. Malik, "Scale-space and Edge Detection Using Anisotropic Diffusion. PAMI,Vol.12, No.7,pp.629-639,1990.

[22] H. Wei, S. Zabuawala, KM. Varadarajan, et.al. "Adaptive pattern-based image compression for ultra-low bandwidth weapon seeker image communication", SPIE 2009.

[23] Aldo Dall'Osso, "An iterative back substitution algorithm for the solution of tridiagonal matrix systems with fringes", JCAM, 2003.

[24] A. Hoover, G. Jean-Baptiste, et al., "A Comparison of Range Image Segmentation Algorithms", PAMI, 1996.

[25] J Yin, JR Cooperstock, "Improving depth maps by nonlinear diffusion", WSCG, 2004.

[26] H Zimmer, A esBruhn, et al. "PDE-based anisotropic disparity-driven stereo vision", VMV 2008.

[27] J. Diebel and S. Thrun, "An application of Markov random fields to ranage sensing", NIPS 2005.

[28] J. Dolson, et.al. "Upsampling range data in dynamic environments", CVPR2010.

[29] KM. Varadarajan, M. Vincze, "3D Room Modeling and Doorway Detection from Indoor Stereo Imagery using Feature guided Piecewise Depth Diffusion", IROS 2010.

[30] RobotVision @ ImageCLEF http://www.imageclef.org/2010/robot

[31] Jensfelt, P. Ekvall, S. Kragic, D. Aarno, D., "Augmenting SLAM with Object Detection in a Service Robot Framework", ROMAN 2006.

[32] L. Shi, S. Kodagoda and G. Dissanayake, "Environment Classification and Semantic Grid Map Building Based on Laser Range Finder Data", SPME 2010.

[33] Aravindhan K Krishnan and K Madhava Krishna, "A Visual Exploration Algorithm using Semantic Cues that Constructs Image based Hybrid Maps", SPME 2010.


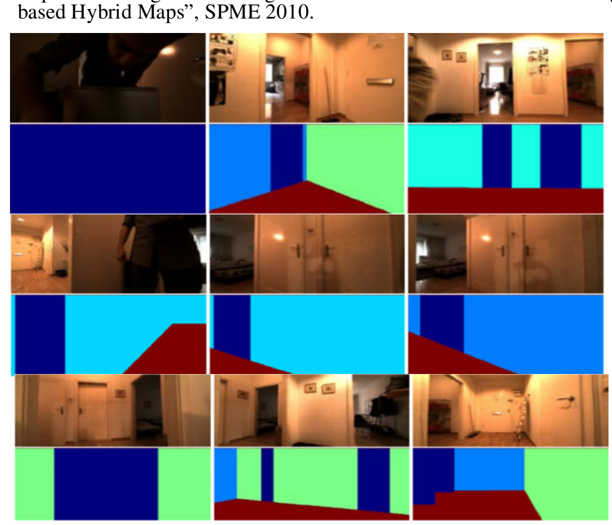

**Figure 5**. Results from test environment (Top rows) Input scenes (Bottom rows) 3D model reprojected on to the image plane. The percentage of mislabeled pixels (w.r.t to manual segmentation) was 5%.