- Exploring High-Level Plane Primitives for Indoor 3D Reconstruction with a Hand-held RGB-D Camera, ACCV, 2013.
- Dense Planar SLAM, ISMAR (International Symposium on Mixed and Augmented Reality), 2014.
- Simultaneous Localization and Mapping with Infinite Planes, ICRA, 2015.

Exploring High-Level Plane Primitives, ACCV, 2013.

Contributions

- a robust pair-wise matching algorithm across frames via matching of both extracted planes and RGB image visual features (SIFT).
- incorporate plane correspondences (in addition to visual feature correspondence) to the Bundle Adjustment (BA).

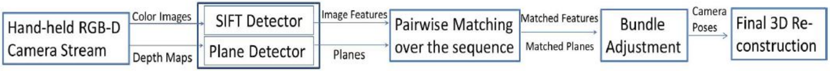


Figure: The flow chart.

Planar Surface Extraction from Depth Map

- Voting algorithm (Hough transform).
- Finally, we assign each pixel to one of the detected planes, or as a non-plane if the distance to all planes is too large.
- the convex hull $\{\mathbf{v}_i\}_{i=1}^{K}$ of a plane segment is found to indicate its boundary.

Robust Pair-Wise Matching

- initial feature match set is computed by checking the similarity of SIFT descriptors.
- Plane matching hypothesis
- Run RANSAC on one plane matching hypothesis and the feature match set.

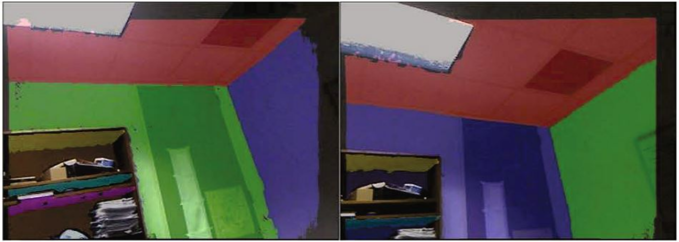
Robust Pair-Wise Matching

- Plane matching hypothesis
  - matching criterion: relative plane angle and plane appearance similarity.
  - plane appearance similarity
    - $\sum_i \min\left(h_1^{HS}(i), h_2^{HS}(i)\right) + \sum_i \min\left(h_1^{I}(i), h_2^{I}(i)\right)$
    - a joint histogram of hue-saturation h^{HS} - color information.
    - an intensity histogram h^{I} - texture information.

Exploring High-Level Plane Primitives, ACCV, 2013.

Robust Pair-Wise Matching

- Plane matching hypothesis
  - a plane matching hypothesis - a subset of planes in one frame and the matching planes in the other frame.
  - To eliminate some hypotheses, constrain the rotation angle within a threshold, given the practical assumption that two nearby frames should not rotate too much.



Exploring High-Level Plane Primitives, ACCV, 2013.

Robust Pair-Wise Matching

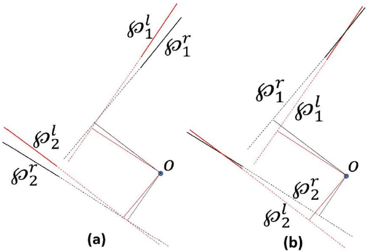
- Run RANSAC on one plane matching hypothesis and the feature match set
  - Randomly Sample Matched Pairs.
    - 3 planes
    - 3 point features
    - 2 planes with 1 point feature
    - 2 point features with 1 plane
  - Calculating the transformation from pairs of matches.
    - consider n pairs of matched planes $S = \{< P^l_i, P^r_j >\}$
    - and m pairs of matched features $T = \{< f^l_i, f^r_j >\}$
    - a transformation $< R, T >$ is estimated s.t. the overall distance between matched items should be minimized

$$\min \sum_{<P^l_i, P^r_j> \in S} D^2_{pln}\left(Q(R, T, P^l_i), P^r_j\right) + \sum_{<f^l_i, f^r_j> \in T} D^2_{pt}(Rf^l_i + T, f^r_j)$$

Exploring High-Level Plane Primitives, ACCV, 2013.

Robust Pair-Wise Matching

- Run RANSAC on one plane matching hypothesis and the feature match set
  - Distance between plane segments
    - The closeness on planes parameters shown in (a) does not equal to the closeness of plane segments shown in (b).
    - The solid line segments denote the plane segments, and O is the origin of the world coordinate.



Exploring High-Level Plane Primitives, ACCV, 2013.

Robust Pair-Wise Matching

- Run RANSAC on one plane matching hypothesis and the feature match set
  - Distance between plane segments
    - measure the distances from the boundary points (convex hull) of one plane segment to its matched plane.
    - Instead of measuring Euclidean distance between plane parameters.
    - distance between plane segments $< P_i^l, P_j^r >$ is defined as

$$D_{pln}^2 = \sum_{k=1}^{K_1} \omega_{i,k} ||\mathbf{n}_j^T R \mathbf{v}_{i,k} + \mathbf{n}_j^T T - d_j||_2^2 + \sum_{k=1}^{K_2} \omega_{j,k} ||\mathbf{n}_i^T R^T \mathbf{v}_{j,k} - \mathbf{n}_i^T R^T T - d_i||_2^2$$

Robust Pair-Wise Matching

- Run RANSAC on one plane matching hypothesis and the feature match set
  - For each transformation candidate, count how many other matching pairs fit this transformation.
    - point features - Euclidean distance.
    - plane segments - D_{pln} and overlap in the image space.

Exploring High-Level Plane Primitives, ACCV, 2013.

Extended Bundle Adjustment of Feature Points and Planes

- feature track - a set of linked features $\{f_k^i\}_{i\in C_k}$, corresponding to the same 3D point $\mathbf{p}_k$ in the world coordinate system.
- plane track - a set of linked planes $\{P_j^i\}_{i\in D_j}$, corresponding to the same world plane $\mathscr{Q}_j$.
- Problem statement
  - M plane tracks $\left\{\{P_j^i\}_{i\in D_j}\right\}_{j=1}^{M}$ and K feature tracks $\left\{\{f_k^i\}_{i\in C_k}\right\}_{k=1}^{K}$.
  - unknown - camera poses $\{R_i, T_i\}_{i=1}^{N}$, plane parameters $\{\mathbf{n}_j, d_j\}_{j=1}^{M}$ and point locations $\{\mathbf{p}_k\}_{k=1}^{K}$.

Exploring High-Level Plane Primitives, ACCV, 2013.

Extended Bundle Adjustment of Feature Points and Planes

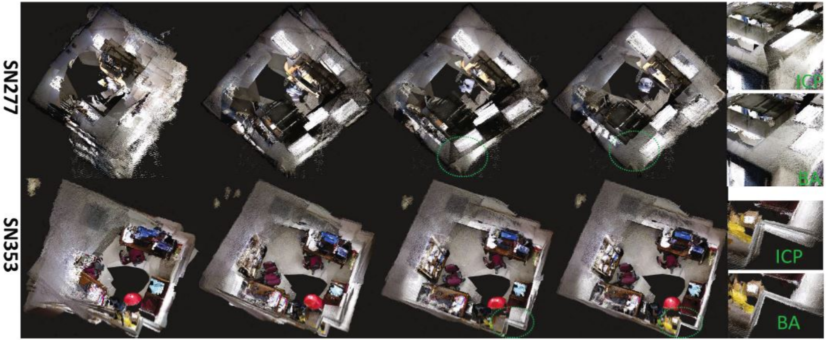
- Cost Function

$$\frac{c}{N_{pln}} \sum_{i,j|i \in D_j} c_j^i D_{pln}^2 \left(Q(R_i, T_i, \mathcal{Q}_i), P_j^i \right) + \frac{1-c}{N_{pt}} \sum_{i,k|i \in C_k} D_{pt}^2 \left(Q(R_i, T_i, \mathbf{p}_k), f_k^i \right)$$

- Plane Track Refinement
  - some large planes tend to have several disjoint plane tracks.
  - merge planes in the world space that are close enough to each other.
  - delete a detected plane from its track if its distance to the corresponding world plane is beyond a threshold.
  - delete the whole track if more than half the planes do not fit the corresponding world plane.

Exploring High-Level Plane Primitives, ACCV, 2013.

Experiments



(a) Bundler (b) ICP (c)ICP with Error Re-Distribution (d) Our method (e) details

Exploring High-Level Plane Primitives, ACCV, 2013.

Experiments

- Quantitative Measurement of Errors
  - the relative angles between some planes in the room, such as walls, ceilings and floors, are known (zero angle or right angle).
  - these angles serve as the ground truth for the measured angles between the world planes.

dataset	**point proj error**(cm)	**plane proj error**(cm)	**zero angles**($^\circ$)	**right angles**($^\circ$)
SN353	1.46± 1.37	1.74± 2.03	-0.60± 2.95	89.97± 2.14
SN277	2.01± 1.83	1.68± 1.58	-0.22± 0.80	89.94± 1.70
FB220	2.38± 1.63	2.36± 2.15	-1.17± 2.41	89.96± 2.42
Lab	1.91± 1.77	1.91± 1.66	-0.59± 4.27	89.61± 2.11

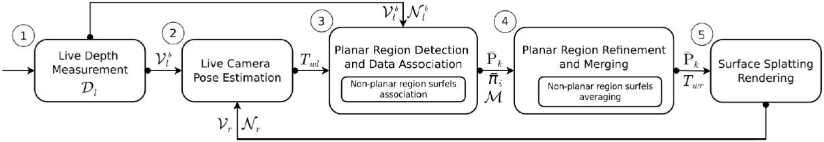
Experiments

- Running Times (PC with 3.0 CPU Hz)
    - plane extraction - 2.5~3s
    - SIFT feature extraction - 1.5s
    - whole BA procedure - 5~20min on a dataset

Dense Planar SLAM, ISMAR, 2014.

System Overview

- densely map the environment with surfels [Keller2013]
- label each surfel in the 3D map either with one of the plane labels, or no label if it is not part of any plane.
- Associated planes are converted into the same world reference frame and refined with a running average.
- Overlapping modelled planes with similar properties are merged together to incrementally extend areas.



Dense Planar SLAM, ISMAR, 2014.

Mapping with Planes

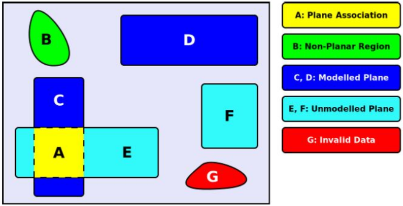
- map representation - a set of k unstructured surfels $\bar{\mathbf{P}}_k$
  - position $\bar{\mathbf{v}}_k \in \mathbb{R}^3$
  - normal $\bar{\mathbf{n}}_k \in \mathbb{R}^3$
  - radius $\bar{r}_k \in \mathbb{R}$
  - confidence $\bar{c}_k \in \mathbb{R}$
  - timestamp $\bar{t}_k \in \mathbb{N}$
  - (additional) plane ID $\bar{o}_k = i, i = 1, ..., p \in \mathbb{N}$
- associate modeled surfels with measurements, producing data-associated pairs

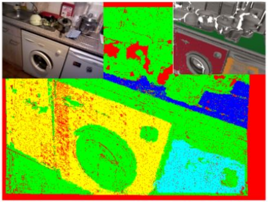
  $a_{surfels} = \{(i,j)\}; i = 1, ..., k; j = 1, ..., w \times h$

Dense Planar SLAM, ISMAR, 2014.

Mapping with Planes

- Planar Region Detection
  - connected component labelling [Dellencourt1992]
  - label map $\mathscr{L}(\mathbf{u}) = i; i = 1, \ldots, q \in \mathbb{N}$
- Data-Association with Planes
  - (A) a modelled plane and a measured plane intersect - $a_{plane} = \{(i,j)\}; i = 1, \ldots, p; j = 1, \ldots, q$
  - (B) modelled surfels lack a planar measurement.
  - (C,D) modelled planar region surfels lack planar measurement.
  - (E,F) unmodelled (measured) planar regions.
  - (G) invalid data.





Dense Planar SLAM, ISMAR, 2014.

Mapping with Planes

- Planar Region Refinement and Merging
  - a modelled plane $\bar{\pi}_i$ is refined with the associated measured plane π_i using a running average.

$$\mathbf{n}_i \leftarrow \frac{\omega \mathbf{n}_i + R\mathbf{n}_j}{\omega + 1}, d_i \leftarrow \frac{\omega d_i + (-R\mathbf{n}_j \cdot t + d_j)}{\omega + 1}, \omega \leftarrow \omega + 1$$

Dense Planar SLAM, ISMAR, 2014.

Map Compression

- compress planar regions whenever they become nonvisible (i.e. outside the view frustum).
- if the planar region does not intersect all of the 6 planes enclosing the frustum.

Dense Planar SLAM, ISMAR, 2014.

Map Compression

- Compression
  - perform an additional PCA step.
  - estimate the major x-y axis of the extended plane.
  - representing the plane as a binary image
    - virtual image of dimensions $w_{vi} \times h_{vi}$
    - compute the compressed index of a surfel
    - $\mathbf{v}_c = \mathbf{v} - \hat{\mathbf{v}}$
    - $\mathbf{v}_p = (\mathbf{x}_{axis} \cdot \mathbf{v}_c, \mathbf{y}_{axis} \cdot \mathbf{v}_c)^T$
    - $\mathbf{v}_{vi} = round(\mathbf{v}_p \times 1000)$
    - $\mathbf{v}_o = \mathbf{v}_{vi} + (w_{vi}, h_{vi})^T / 2$
    - $index = \mathbf{v}_{o(y)} \times w_{vi} + \mathbf{v}_{o(x)}$

Dense Planar SLAM, ISMAR, 2014.

Results

- Synthetic scenes

Error	trajectory-0		trajectory-1	
	non-planar	planar	non-planar	planar
RMSE	0.254134	**0.246437**	0.018997	**0.016940**
Mean	0.222794	0.218559	0.016906	0.015043
Median	0.179679	0.182547	0.014714	0.016449
Std	0.122258	0.113857	0.008666	0.007789
Min	0.055287	0.070420	0.003252	0.002228
Max	0.728229	0.645284	0.032742	0.028430

Dense Planar SLAM, ISMAR, 2014.

Results

- Real-world scenes

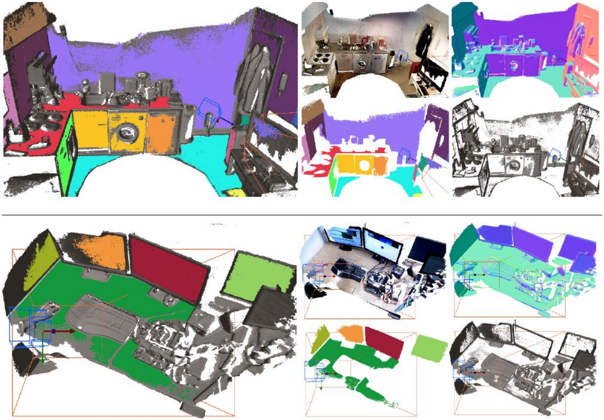


Figure: Real scene reconstruction of an apartment (top) and desktop (bottom). (left) Displaying both planar and non-planar regions surfels. (right) In clockwise order: Colour output, Normal Map, Non-Planar region surfels only, Planar region surfels only.

SLAM with Infinite Planes, ICRA, 2015.

Contributions

- usually presented by the overparametrized representation of an infinite plane.
- a commonly used minimal representation using spherical coordinates for the plane normal suffers from singularities.
- introduce a minimal representation for the homogeneous parametrization of infinite planes suitable for least-squares estimation with Gauss-Newton methods and related incremental solvers.

Mapping With Infinite Planes

- State and Plane Representation
  - sensor pose $x = (\mathbf{t}, \mathbf{q}) \in \mathbb{R}^3 \times S^3$
  - transformation matrix

$$T_{gx} = \begin{pmatrix} R(\mathbf{q}) & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

  - a point in projective space is represented by homogeneous coordinates $\mathbf{p} = (p_1, p_2, p_3, p_4)^T \in \mathbb{P}^3$ where its corresponding Euclidean coordinates for $p_4 \neq 0$ are $(p_1/p_4, p_2/p_4, p_3/p_4)^T \in \mathbb{R}^3$
  - transform from local to global frame

$$\mathbf{p}_g = T_{gx}\mathbf{p}_x$$

SLAM with Infinite Planes, ICRA, 2015.

Mapping With Infinite Planes

- State and Plane Representation
  - plane represented in projective space [Hartley2003]

$$\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T \in \mathbb{P}^3$$

  - a point $\mathbf{p} \in \mathbb{P}^3$ lies on the plane iff

$$\pi^T \mathbf{p} = 0$$

  - transform from local to global frame

$$\pi_g = T_{gx}^{-T} \pi_x$$

Mapping With Infinite Planes

- Minimal Representation
  - overparametrization - information matrix becomes rank-deficient and cannot be inverted as needed for Gauss-Newton type optimization.
  - only 3DoF in plane parameters - its orientation α, β and its orthogonal distance d from the origin.
  - if represented by (α, β, d), there are singularities.

Mapping With Infinite Planes

- Minimal Representation
    - find a minimal representation by restricting the ambiguity in the homogeneous representation.
    - normalize the vector π to lie on the unit sphere of $\mathbb{R}^4$ as $\pi' = \pi/||\pi|| \in S^3$
    - then use the element ω of Lie algebra $\mathfrak{su}(2)$ of S^3 as the minimal representation.
    - π is updated by an increment ω using the exponential map

$$\pi' = \exp(\omega)\pi$$

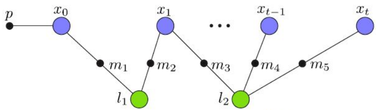
$$\exp(\omega) = \begin{pmatrix} \frac{1}{2} sinc\left(\frac{1}{2}||\omega||\right)\omega \\ \cos\left(\frac{1}{2}||\omega||\right) \end{pmatrix}$$

SLAM with Infinite Planes, ICRA, 2015.

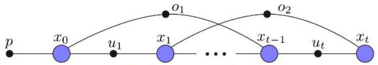
Mapping With Infinite Planes

- SLAM Formulation
  - estimate the sensor poses $x_0, \ldots x_t$ and planes $\pi_1, \ldots, \pi_m$ given the plane measurements.
  - use a factor graph as a graphical model [Kaess2012]

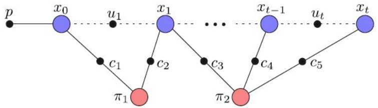
$$G = (F, \Theta, E)$$



(a) Structure-from-motion (SFM)



(b) Pose graph SLAM



(c) Planar SLAM

- factor nodes $f_i \in F$ - measurements
- variable nodes $\theta_j \in \Theta$ - poses and planes
- edge $e_{ij} \in E$ - connect factor and variable nodes

SLAM with Infinite Planes, ICRA, 2015.

Mapping With Infinite Planes

- SLAM Formulation
  - find variable assignment Θ^*

$$\Theta^* = \arg\max_{\Theta} \prod_i f_i(\Theta_i)$$

  - for Gaussian measurement models

$$f_i(\Theta_i) \propto \exp\left(-\frac{1}{2}||h_i(\Theta_i) - z_i||^2_{\Sigma_i}\right)$$

Mapping With Infinite Planes

- Plane Measurement Model
  - plane measurement

$$\pi_x = T_{gx}^{-T} \pi \oplus \mathbf{v}, \mathbf{v} \sim N(0, \Sigma)$$

$$p(\hat{x}, \hat{\pi} | \tilde{\pi}_x) = \frac{1}{\sqrt{(2\pi)^3 |\Sigma|}} \exp\left(-\frac{1}{2} || h(T_{g\hat{x}}, \hat{\pi}) \ominus \tilde{\pi}_x ||_{\Sigma}^2 \right)$$

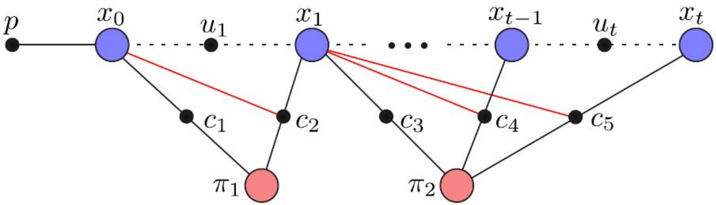
  - cost function

$$c_{x\pi}(\hat{x}, \hat{\pi}) = || h(T_{g\hat{x}}, \hat{\pi}) \ominus \tilde{\pi}_x ||_{\Sigma}^2$$

SLAM with Infinite Planes, ICRA, 2015.

Mapping With Infinite Planes

- Relative Formulation



SLAM with Infinite Planes, ICRA, 2015.

Evaluation

-

TABLE I: Batch optimization, comparing relative and absolute formulation as well as overparametrization and our minimal representation for a simulated sequence (76 poses, 31 planes, 450 plane measurements).

	Gauss-Newton	Levenberg-Marquardt	Powell's Dog-Leg
Overpar., absolute	not possible	76 it (475ms)	not possible
Overpar., relative	not possible	15 it (208ms)	not possible
Minimal, absolute	diverged	58 it (340ms)	17 it (126ms)
Minimal, relative	**5 it** (105ms)	**5 it** (106ms)	**7 it** (129ms)

Experimental Results

- ASUS Xtion Pro Live sensor at 640×480 resolution.
- laptop computer with i7-3920XM 2.9GHz CPU. No GPU is used.
- multi-threaded, with separate threads for plane detection, graph optimization, and visualization.
- runs at 15 frames per second.

SLAM with Infinite Planes, ICRA, 2015.

Experimental Results

Length of sequence	$58s$
Number of planes	81
Number of poses	868
Number of plane observations	$5,934$
Time for normal computation	$34ms$
Time for plane segmentation	$33ms$
Time for data association	$< 1ms$
Time for graph optimization	$24ms$

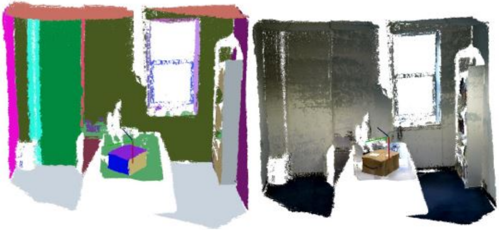


Fig. 6: Office sequence. (left) The 3D model with a random color assigned to each plane. (right) The 3D model with color from the input images.

REFERENCE

M. Keller, D. Lefloch, M. Lambers, S. Izadi, T.Weyrich, and A. Kolb. Real-time 3D Reconstruction in Dynamic Scenes using Point-based Fusion. In Proc. of Joint 3DIM/3DPVT Conference (3DV), June 2013.

M. B. Dillencourt, H. Samet, and M. Tamminen. A General Approach to Connected-component Labeling for Arbitrary Image Representations. Journal of the ACM, 39(2):253280, Apr. 1992.

R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge University Press, 2003, second Edition.

M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, iSAM2: Incremental smoothing and mapping using the Bayes tree, The International Journal of Robotics Research, vol. 31, pp. 217236, Feb. 2012. [9] T. Lee, S. Lim, S. Lee,

C. Feng, Y. Taguchi, and V. R. Kamat, Fast plane extraction in organized point clouds using agglomerative hierarchical clustering, in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2014.