Dual-Feature Warping-based Motion Model Estimation (ICCV2015)

Shiwei Li, Lu Yuan, Jian Sun, Long Quan

Motivation (limitation)

1. Misalignments due to the lack of reliable keypoints

2. Distortion artifact induced by flexible warps due to insufficient corresponding information

(a) keypoints

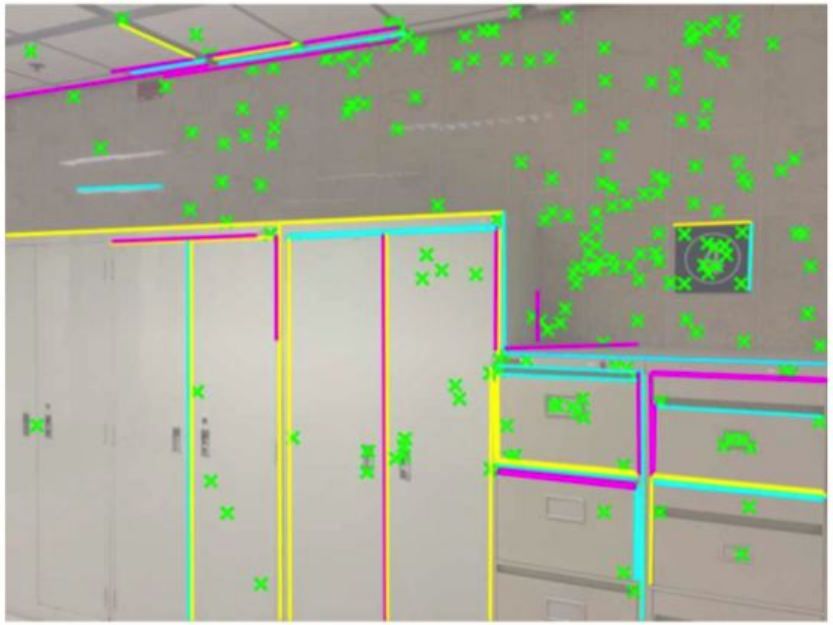
(b) homography by keypoints

(c) result of APAP [27]

Motivation (solution)

1. Dual-features (keypoints & lines)

2. Line preserving flexible warp



(d) dual features

(e) homography by dual features

(f) our final dual-feature warp

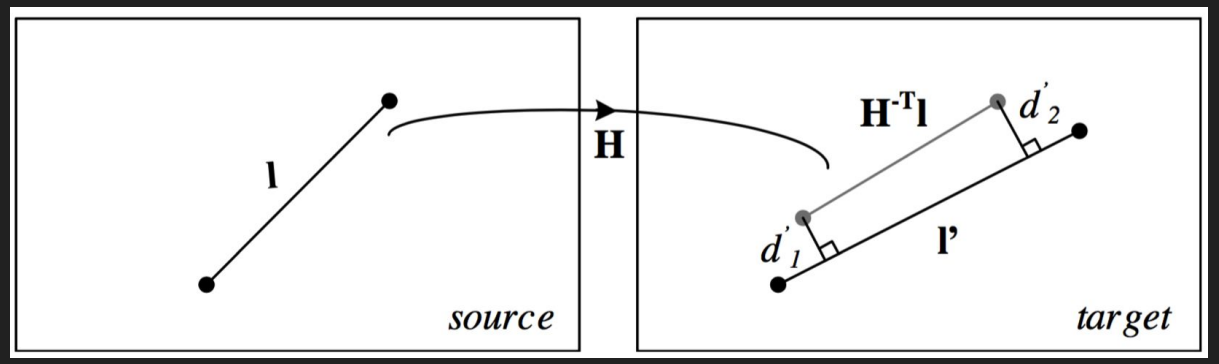
Dual-features (keypoints & lines)

Parameterization

Endpoints: $\hat{p}^0, \hat{p}^1$

Line transformation: $I' = \mathcal{M} \circ I$

Line to line distance: $d(I, I') = \sqrt{d^2(\hat{p}^0, I') + d^2(\hat{p}^1, I')}$



Energy function

$$\widehat{\mathcal{M}} = \underset{\mathcal{M}}{\operatorname{argmin}}\left(\sum_i d^2(\hat{p}_i, p_i') + \sum_j d^2(\hat{I}_j, I_j')\right)$$

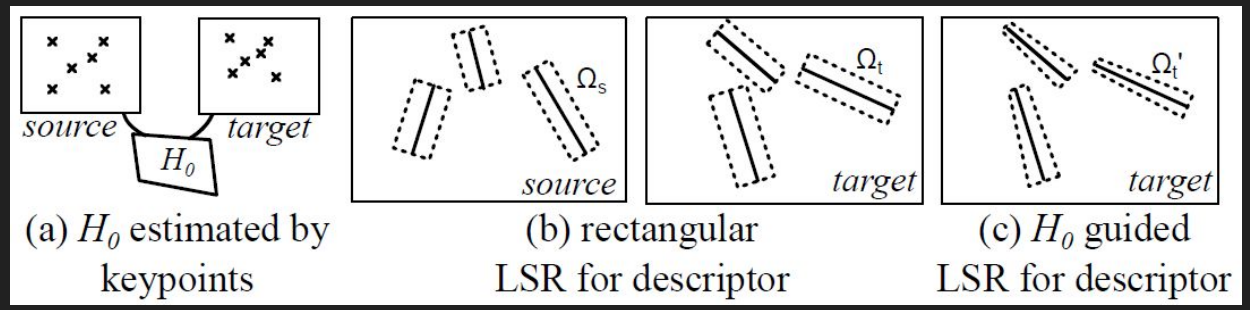
$$\hat{p}_i = \mathcal{M} \circ p_i \qquad \hat{I}_j = \mathcal{M} \circ I_j$$

Dual-features (keypoints & lines)

Line detection and matching

1. Detection: EDLine [1]

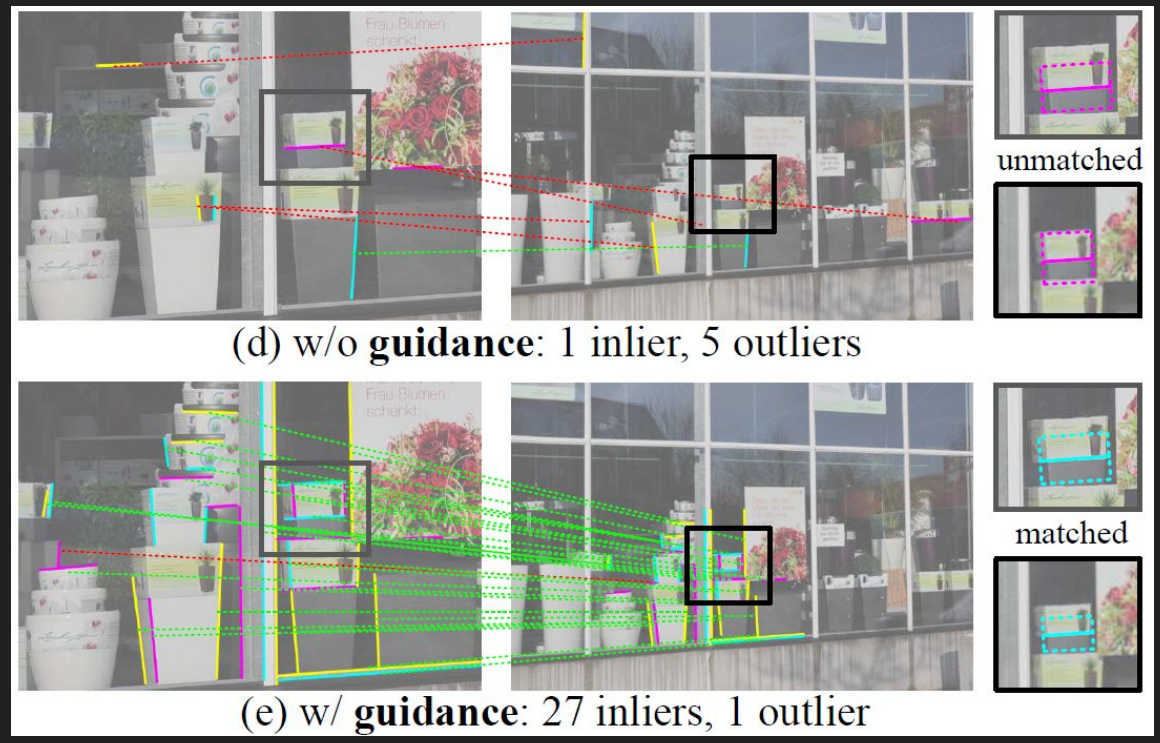
2. Matching: LSR (Line Supporting Region) & MSLD [2]



(a) H_0 estimated by keypoints

(b) rectangular LSR for descriptor

(c) H_0 guided LSR for descriptor



(d) w/o **guidance**: 1 inlier, 5 outliers

(e) w/ **guidance**: 27 inliers, 1 outlier

[1] C. Akinlar and C. Topal. Edlines: A real-time line segment detector with a false detection control. Pattern Recognition Letters., 32(13):1633–1642, 2011.

[2] Z. Wang, F. Wu, and Z. Hu. Msld: A robust descriptor for line matching. Pattern Recognition Letters., 42(5):941–953, 2009.

Warping-based Motion Model

Dual-feature Homography Estimation

1. Geometric Distance:

$$\min\left(\sum_i \|p_i' - \hat{p}_i\|^2 + \sum_j \frac{\left|I_i'^T \hat{p}_j^0\right|^2 + \left|I_i'^T \hat{p}_j^1\right|^2}{a'^2 + b'^2}\right)$$

2. Algebraic Distance:

$$\begin{aligned}\hat{H} &= \underset{H}{\operatorname{argmin}}\left(\sum\nolimits_i \|p_i' \times Hp_i\|^2 + \sum\nolimits_j \left\|I_j'^T Hp_j^{0,1}\right\|^2\right) \\ &= \underset{H}{\operatorname{argmin}}\left(\textstyle\sum_i \|A_i h\|^2 + \sum_j \left\|B_j h\right\|^2\right)\end{aligned}$$

$$A_i = \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i'x_i & -x_i'y_i & -x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -y_i'x_i & -y_i'y_i & -y_i' \end{bmatrix} \quad B_i = \lambda_i \begin{bmatrix} a_j'u_j^0 & a_j'v_j^0 & a_j' & b_j'u_j^0 & b_j'v_j^0 & b_j' & c_j'u_j^0 & c_j'u_j^0 & c_j' \\ a_j'u_j^1 & a_j'v_j^1 & a_j' & b_j'u_j^1 & b_j'v_j^1 & b_j' & c_j'u_j^1 & c_j'v_j^1 & c_j' \end{bmatrix}$$

Warping-based Motion Model

Dual-feature Homography Estimation

1. Balancing

Algebraic residual:

$\|A_i h\| \sim d(Hp_i, p_i') \cdot w_i$ (i.e., the geometric distance multiplying w_i)

$\|B_i h\| = \lambda_i {I_j'}^T H p_j^{0,1} \sim d(\hat{p}_j^{0,1}, I_j') \cdot w_j^{0,1} \Rightarrow \lambda_i = \frac{1}{\sqrt{{a_j'}^2 + {b_j'}^2}}$

Warping-based Motion Model

Dual-feature Homography Estimation

2. Normalization (Hartley's normalization)

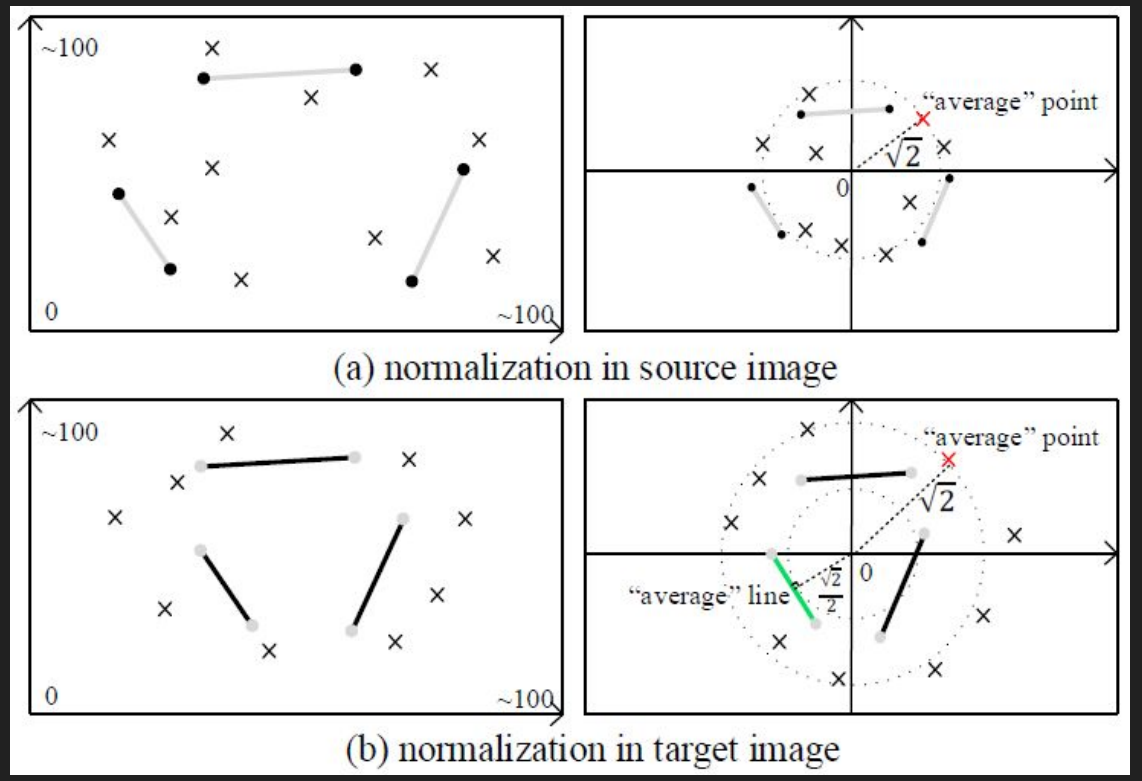
2.1 Source image: (x, y, u, v) only points and endpoints

Points & Endpoints: "average point" is (1, 1, 1)

2.2 Target image: (x', y', a', b', c') points and lines

Point: "average point" is (1, 1, 1)

Line: $\left[\frac{a'}{c'}, \frac{b'}{c'}, 1\right] \sim 1$ "average distance" to origin is $\frac{1}{\sqrt{\left(\frac{a'}{c'}\right)^2 + \left(\frac{b'}{c'}\right)^2}} = \frac{1}{\sqrt{2}}$



Warping-based Motion Model

Dual-feature Homography Estimation

3. Robust Estimation（RANSAC）

Fitting error for point: $d(\hat{p}, p') = \sqrt{\|p' - \hat{p}\|^2}$

Fitting error for line: $d(\hat{l}, l') = \sqrt{\frac{|l'^T \cdot \hat{p}^0|^2 + |l'^T \cdot \hat{p}^1|^2}{a'^2 + b'^2}}$

The inliers of dual-feature will be further used for local warp estimation

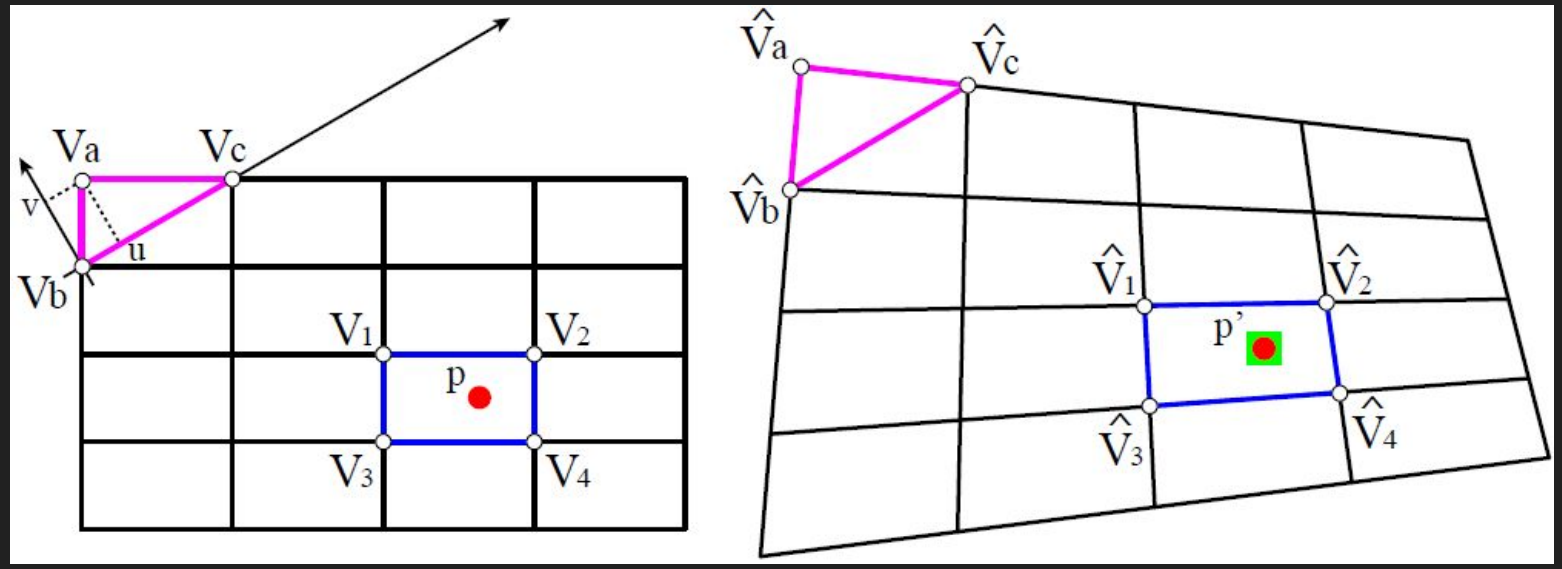
Warping-based Motion Model

Dual-feature Local Warps

Content-preserving warp（CPW）: Data term + Smoothness term

$$E_{point}(V) = \sum_i \|V_{p_i} w_{p_i} - p_i'\|^2$$

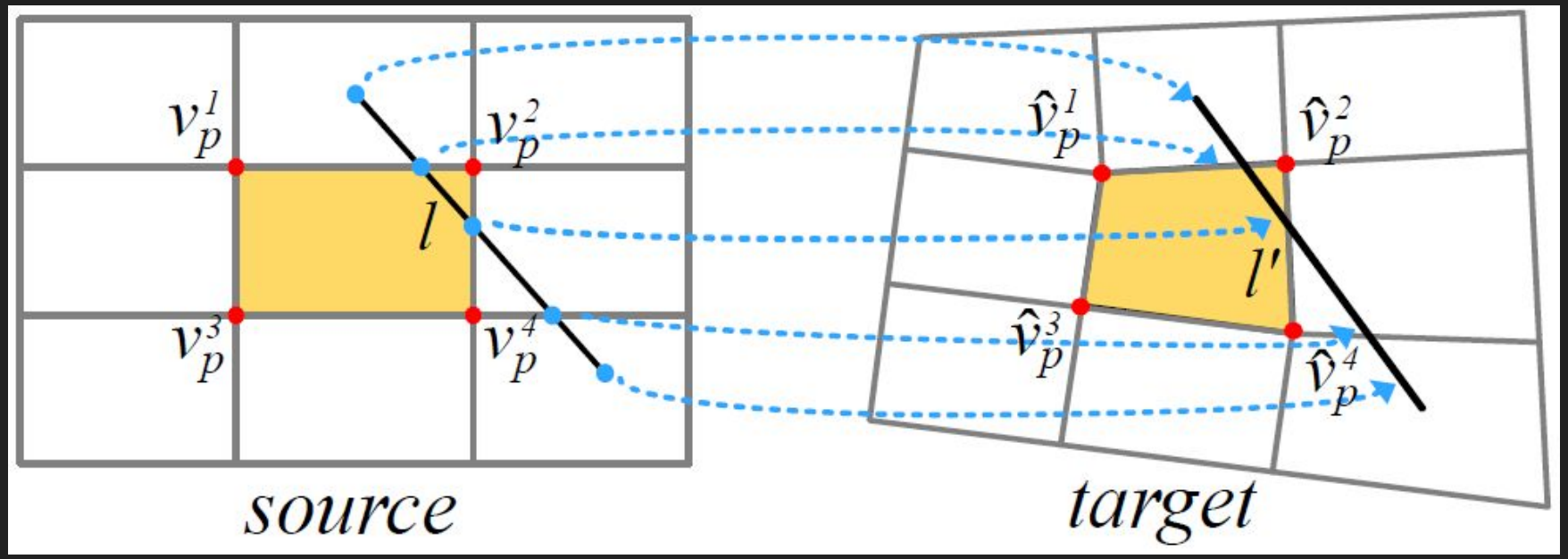
$$E_{smoothness}(V) = \sum_i \|V_a - (V_b + u(V_c - V_b) + vR_{90}(V_c - V_b))p_i'\|^2$$



Warping-based Motion Model

Dual-feature Local Warps

Line term: $$E_{line}(V) = \sum_{i,k} \left\| I_j'^T \cdot V_{p_k} w_{p_k} / \left(\sqrt{a_j'^2 + b_j'^2} \right) \right\|^2$$



Warping-based Motion Model

Dual-feature Local Warps

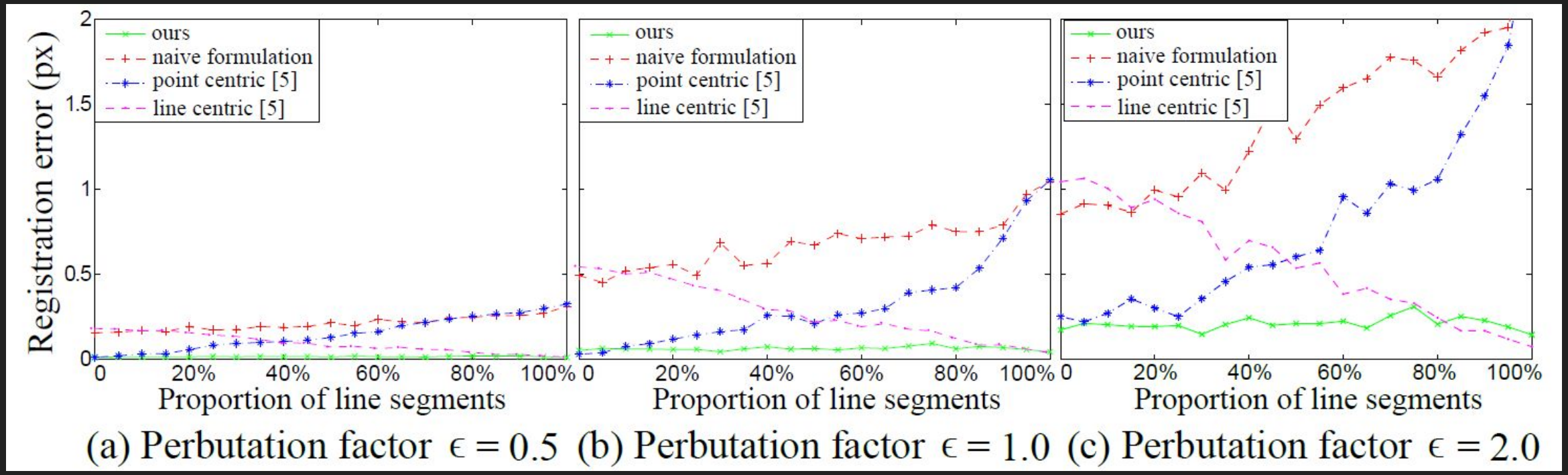
Total energy: $E(V) = E_{point}(V) + E_{line}(V) + \alpha E_{smoothness}(V)$

(a) keypoint feature (b) keypoint-based homography (c) keypoint-based warp (w/o E_{line}) (d) line segment feature (e) our dual-feature warp (w/ E_{line})

Quantitative Evaluation

Numerical stability

Synthesized data: perturbation on point and line features



Quantitative Evaluation

Numerical stability

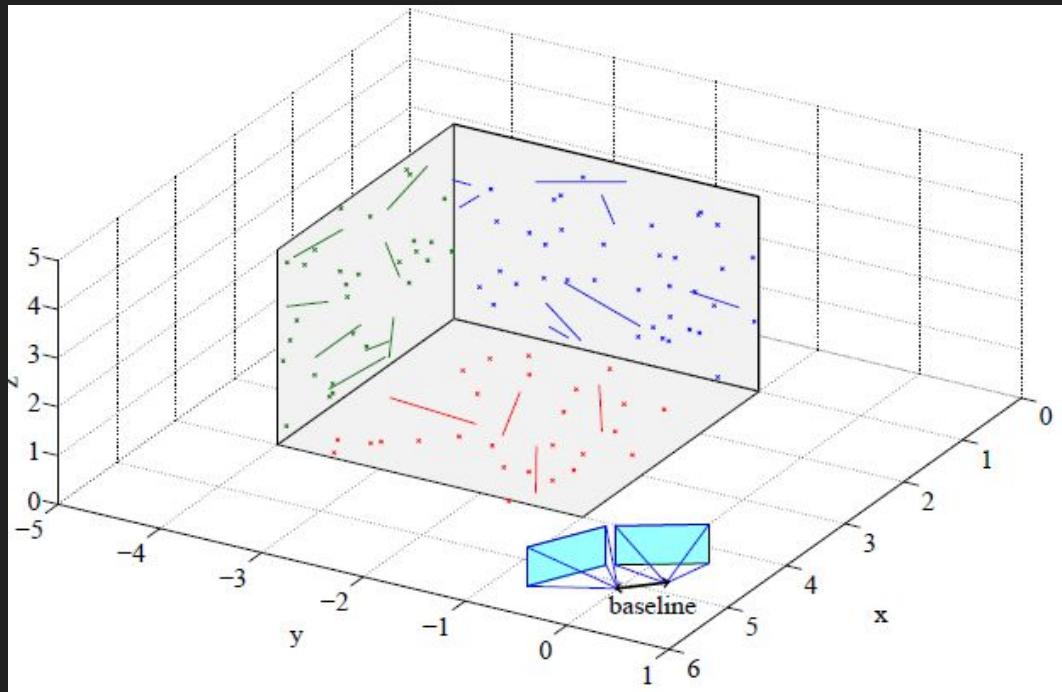
Condition number of the matrix for SVD decomposition

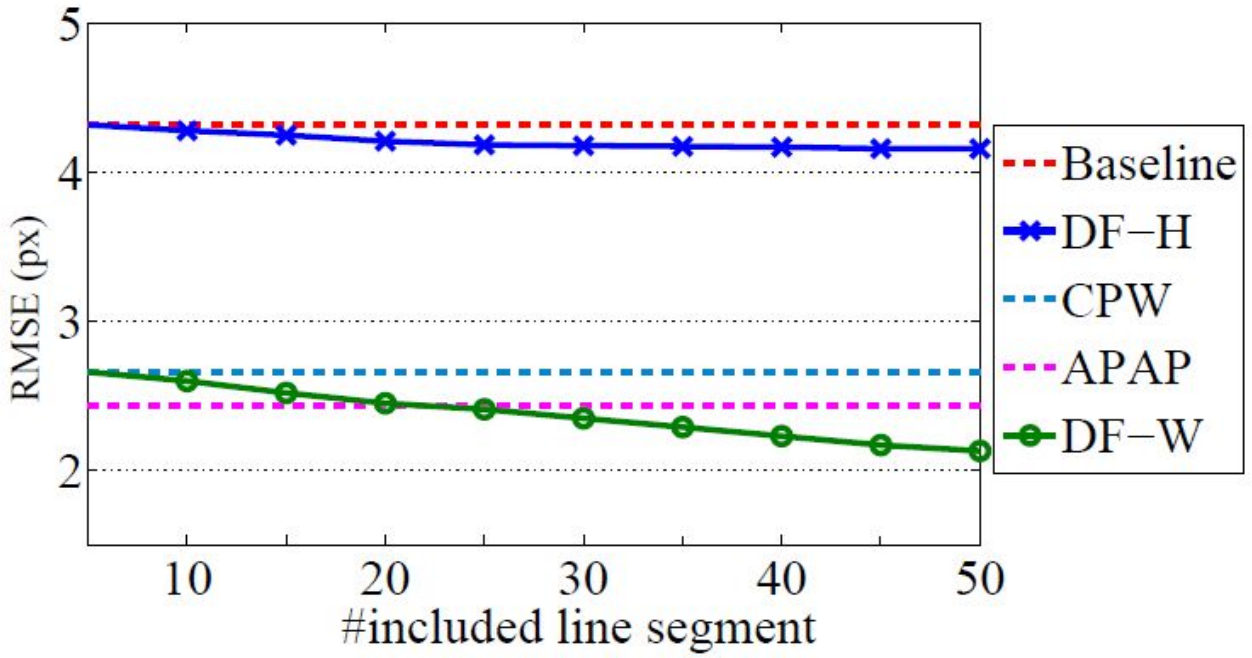
ratio	naive formulation	point centric	line centric	ours
0%	2.12E+6	**1.04E+3**	2.12E+6	**1.04E+3**
20%	2.28E+6	2.87E+3	1.32E+6	**1.05E+3**
40%	2.60E+6	5.18E+4	6.78E+5	**1.12E+3**
60%	3.11E+6	5.30E+4	4.20E+4	**1.19E+3**
80%	3.20E+6	1.34E+5	3.26E+3	**1.30E+3**
100%	8.75E+6	8.75E+6	1.62E+3	**1.49E+3**

Quantitative Evaluation

Translational camera motion

Synthesized data (points & lines) + translational cameras





Quantitative Evaluation

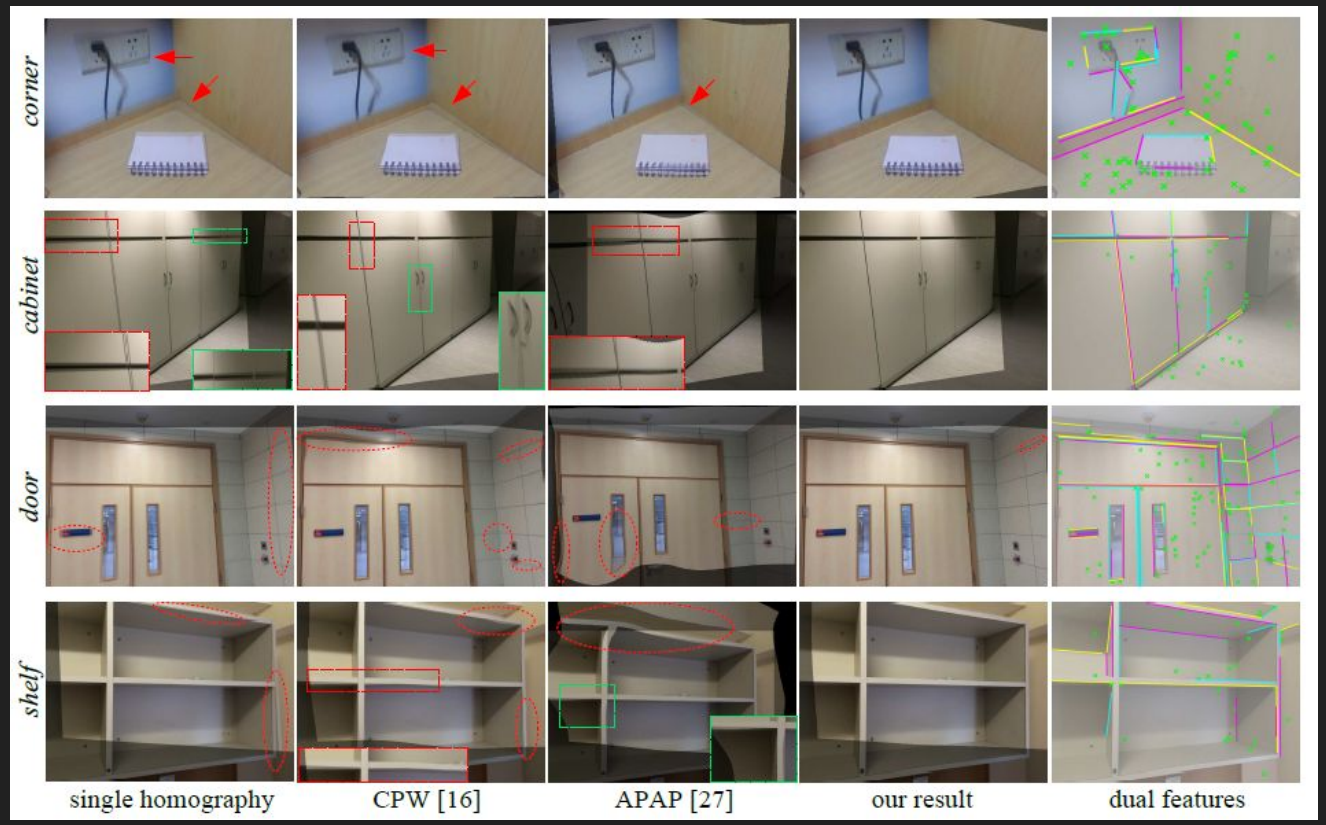
Real Images

$$EMSE(I_i, I_j) = \sqrt{\frac{1}{N}\sum_{\pi}\left(1 - NCC(x_i, x_j)\right)^2}$$

model	homography model		warping-based model		
method	Baseline	DF-H	CPW	APAP	**DF-W**
four	12.78	6.12	7.42	6.92	**2.36**
door	14.47	8.31	4.89	7.37	**3.50**
shelf	8.62	3.04	6.28	8.76	**1.54**
window	9.90	6.94	7.46	5.78	**4.94**
cabinet	6.75	3.72	3.48	4.55	**2.63**
roof	4.84	4.28	5.68	7.82	**2.25**
desk	16.94	12.71	10.67	6.17	**4.89**
corner	10.02	4.34	8.67	6.84	**1.44**
park	21.73	12.61	16.87	11.07	**8.18**
car	3.08	2.77	2.65	**2.07**	2.13
bridge	11.37	7.70	8.47	7.95	**6.60**
girl	8.76	7.82	9.17	5.20	**4.81**
villa	16.23	13.38	7.58	6.72	**5.20**
road	8.17	6.38	6.48	**2.28**	4.59
rotation	2.57	2.28	1.37	1.12	**1.06**
bench	11.5	7.18	8.97	**4.01**	7.12

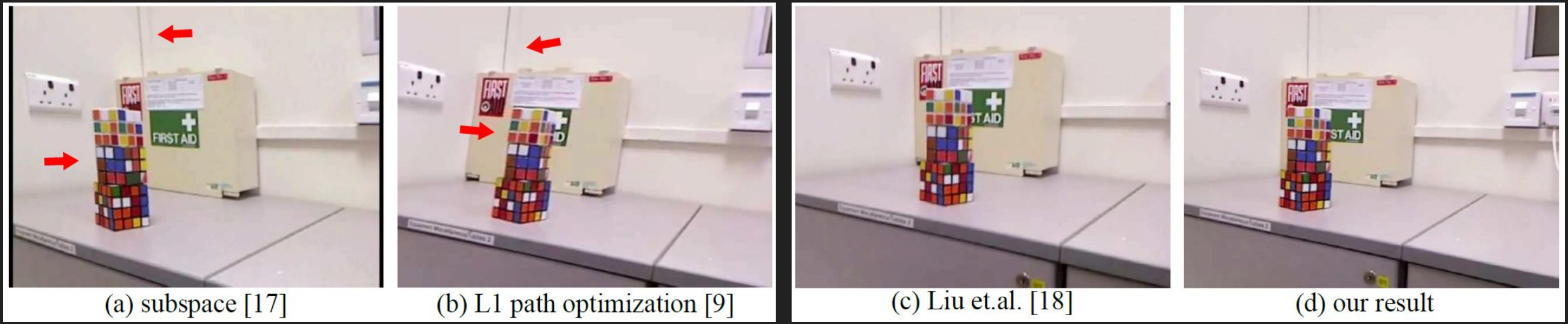
Qualitative Evaluation

Image Stitching (indoor environments)



Qualitative Evaluation

Video Stabilization



(a) subspace [17] (b) L1 path optimization [9] (c) Liu et.al. [18] (d) our result