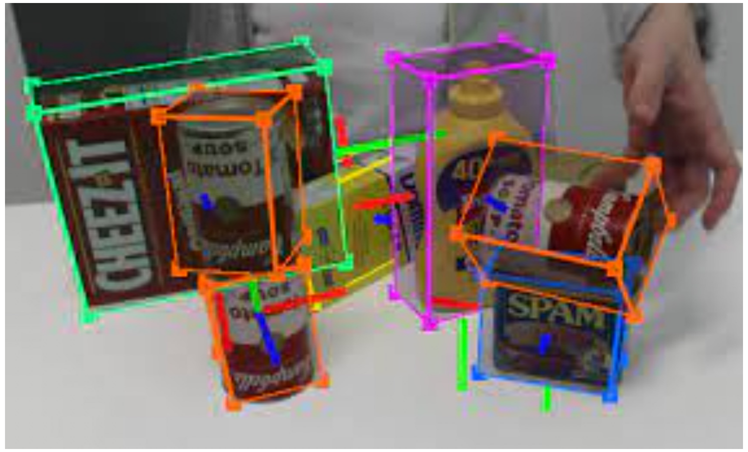
Robust Outlier Rejection for 3D Registration with Variational Bayes

Haobo Jiang[1], Zheng Dang[2], Zhen Wei[2], Jin Xie[1], Jian Yang[1], and Mathieu Salzmann[2]

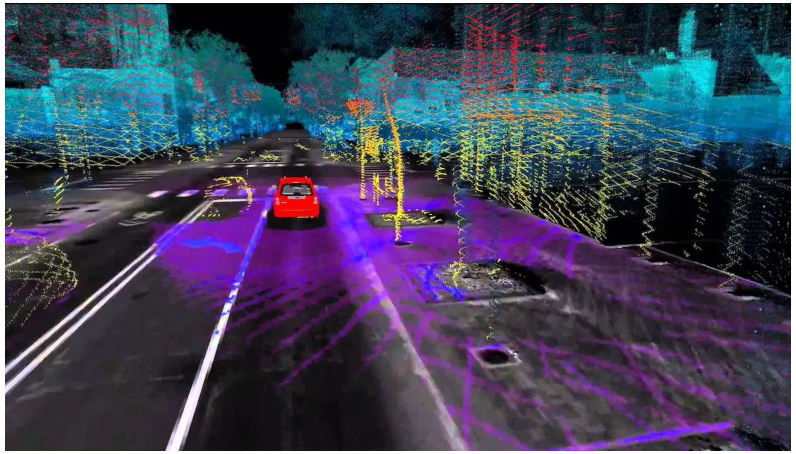
[1]PCA Lab, Nanjing University of Science and Technology, China
[2]CVLab, EPFL, Switzerland

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

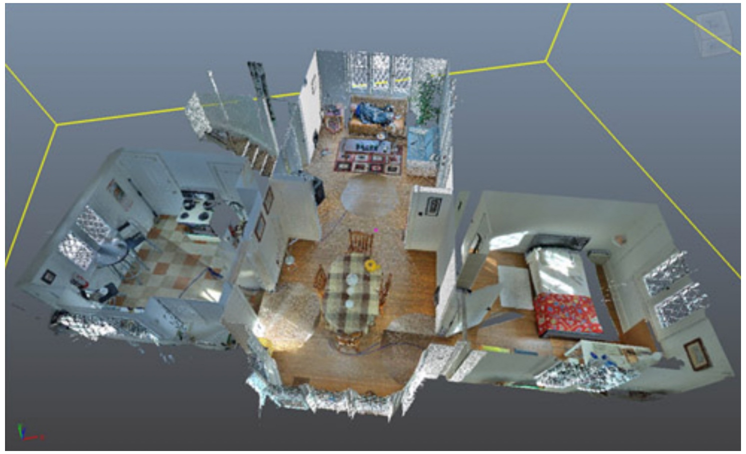
➢ Application



Object pose estimation



Lidar SLAM



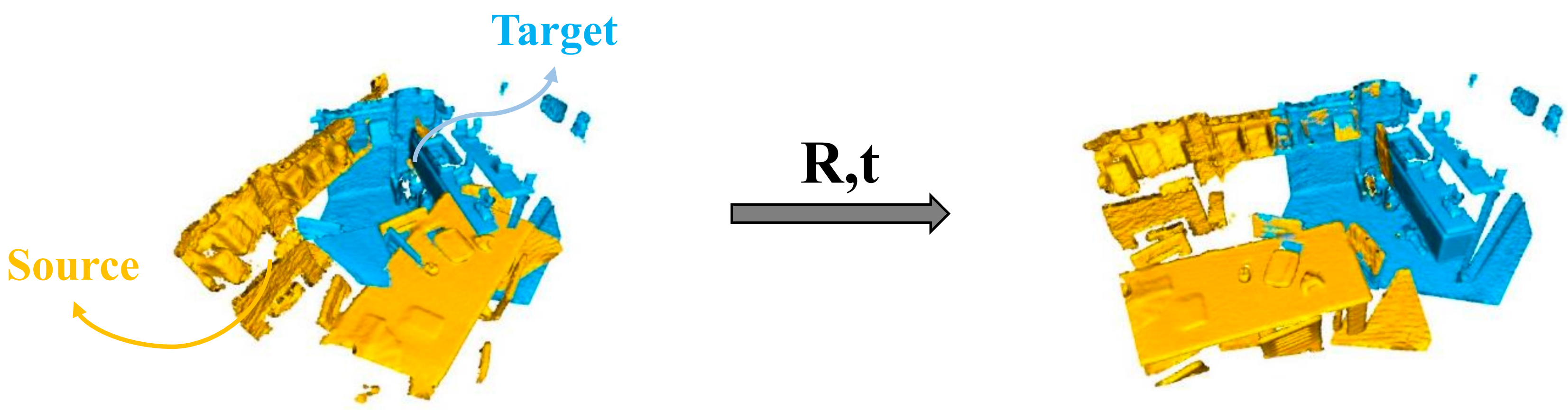
3D scene reconstruction

Robust Outlier Rejection for 3D Registration with Variational Bayes

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

➢ *Background* - Point Cloud Registration

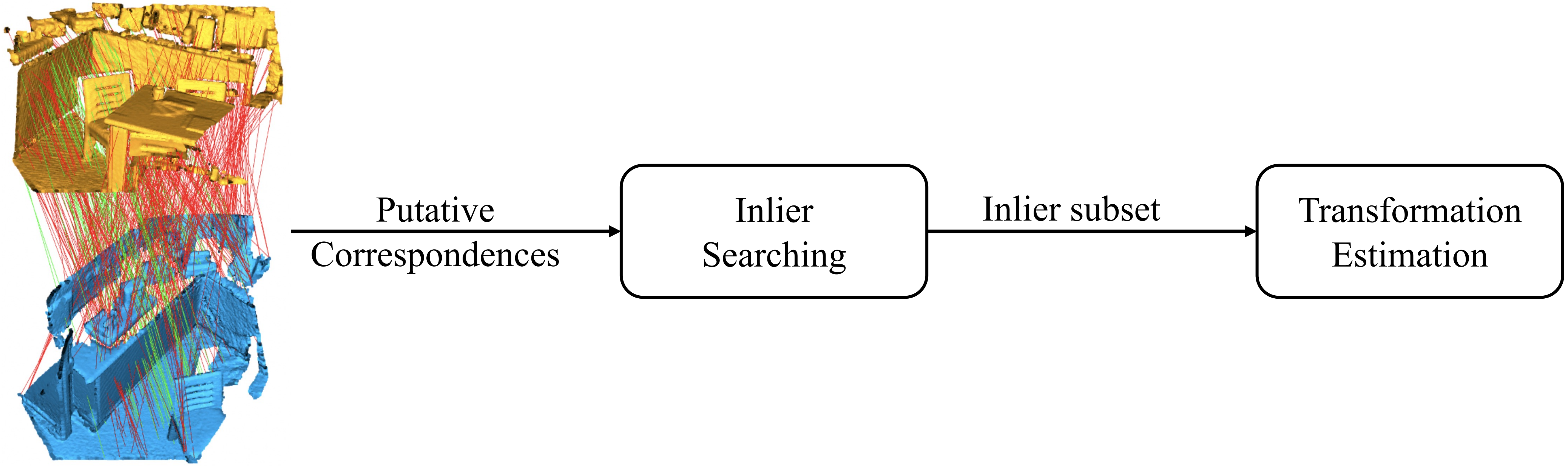
- Estimate rigid transformation (R + t) to align overlap region between the source and target



Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ ***Background*** **- Feature descriptor-based registration pipeline**

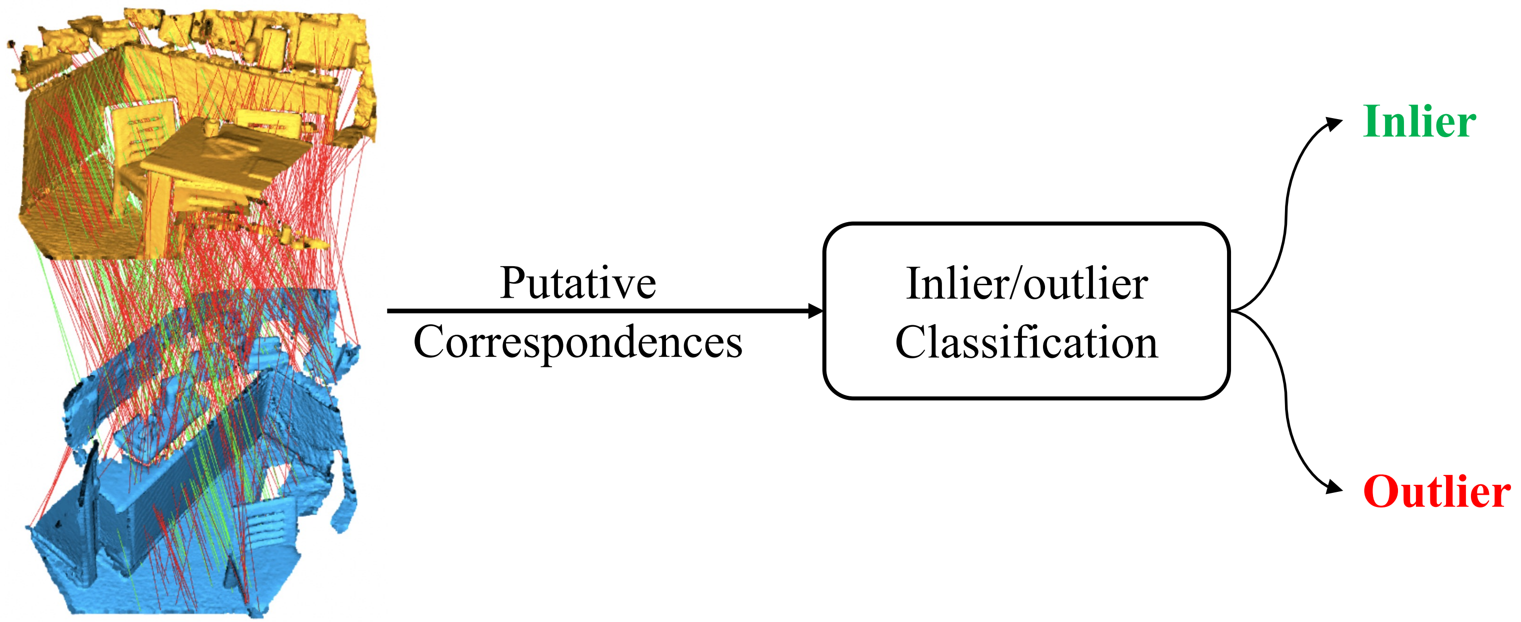
- Construct correspondence (inlier + outlier) with feature similarity
- Search inlier subset for optimal transformation estimation



Robust Outlier Rejection for 3D Registration with Variational Bayes

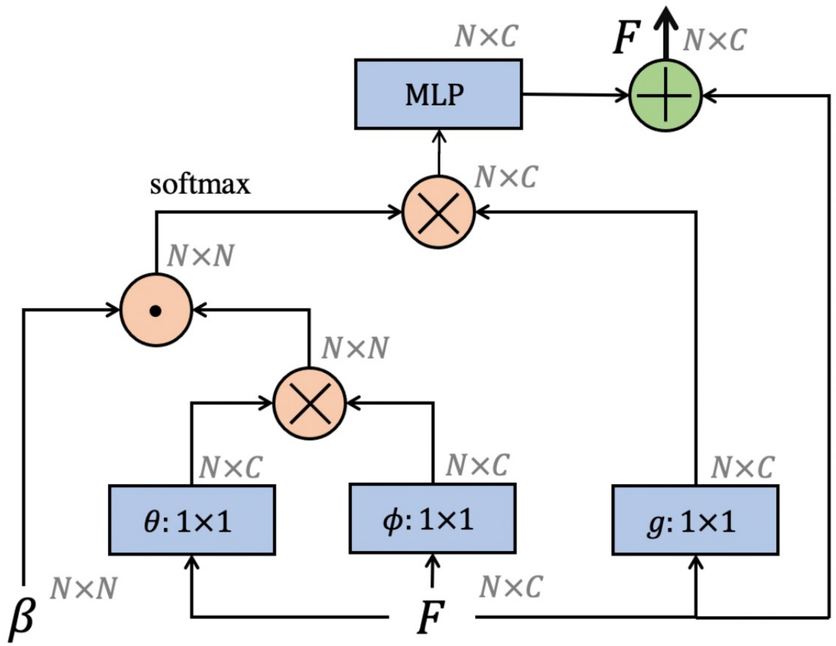
➢ *Background* - Deep outlier rejection methods

- Formulate outlier rejection as inlier/outlier classification
- **Core**: Learning discriminative inlier/outlier feature representations



Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ ***Baseline*** **- Spatial Consistency-guided Non-local Network (SCNonlocal)** [1]



- Inlier/outlier feature aggregation with long-range dependencies

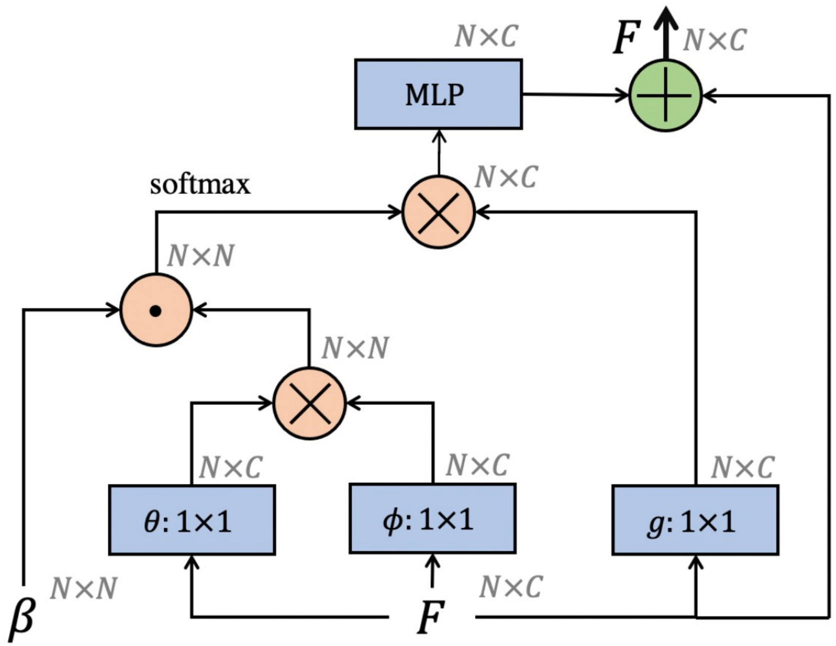
$$\boldsymbol{f_i} = \boldsymbol{f_i} + \mathrm{MLP}\left(\sum_{j}^{|C|} \mathrm{softmax}_j(\boldsymbol{\alpha\beta})g(\boldsymbol{f_j})\right)$$

- Correct long-range dependencies with **spatial consistency**

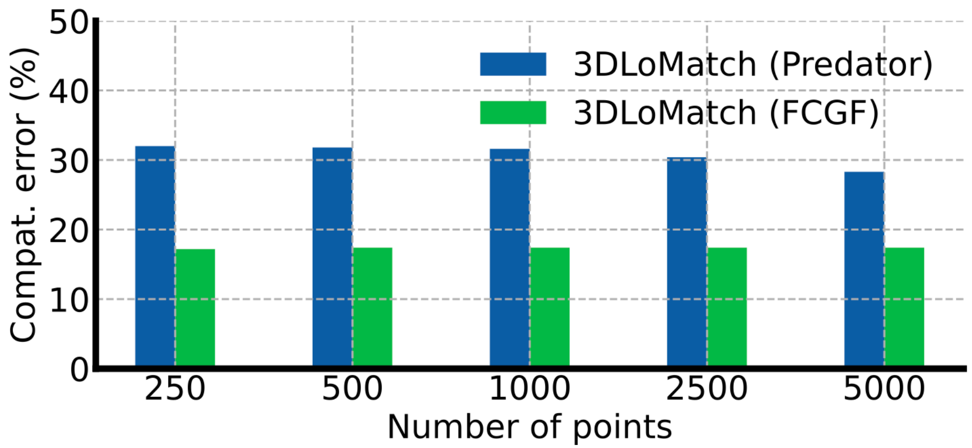
[1] PointDSC: Robust Point Cloud Registration using Deep Spatial Consistency. CVPR'2021.

Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ ***Baseline* - Spatial Consistency-guided Non-local Network (SCNonlocal)** [1]



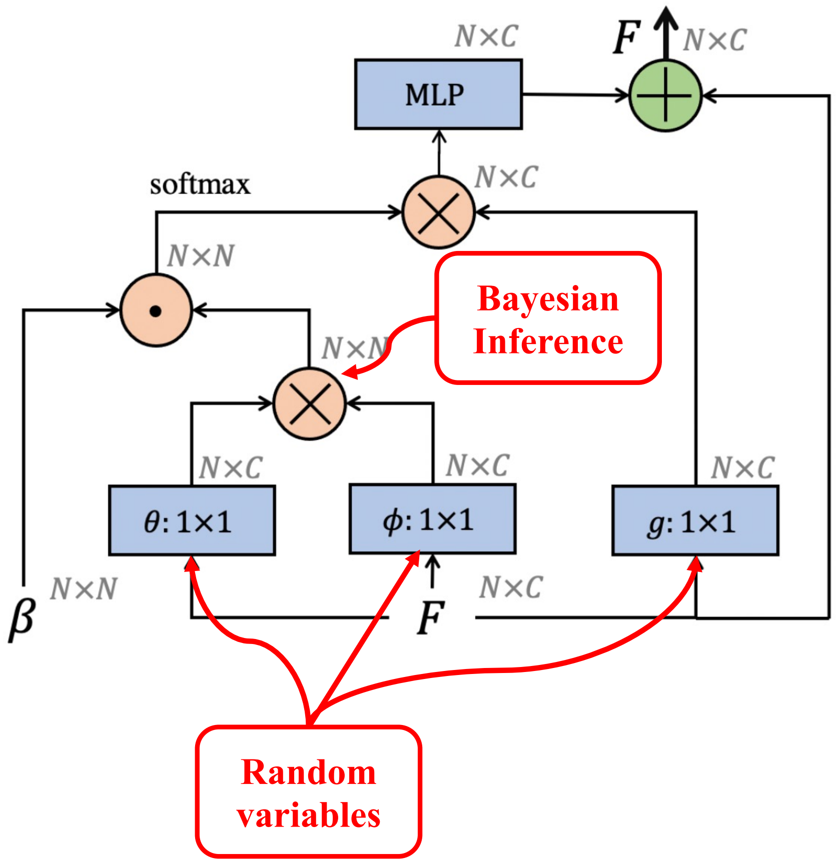
- **Limitation**
- High ratios of wrong spatial consistencies => Mislead attention map
- Lack of uncertainty modeling ability (repetitive, geometry-less…)



[1] PointDSC: Robust Point Cloud Registration using Deep Spatial Consistency. CVPR'2021.

Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ *Approach* - Variational Bayesian-based Non-local Network (VBNonlocal)



- **Solutions**
  - Lack of uncertainty modeling ability (repetitive, geometry-less…)

=> Inject random variables into query/key/value features

  - High ratios of wrong spatial consistencies => Mislead attention map

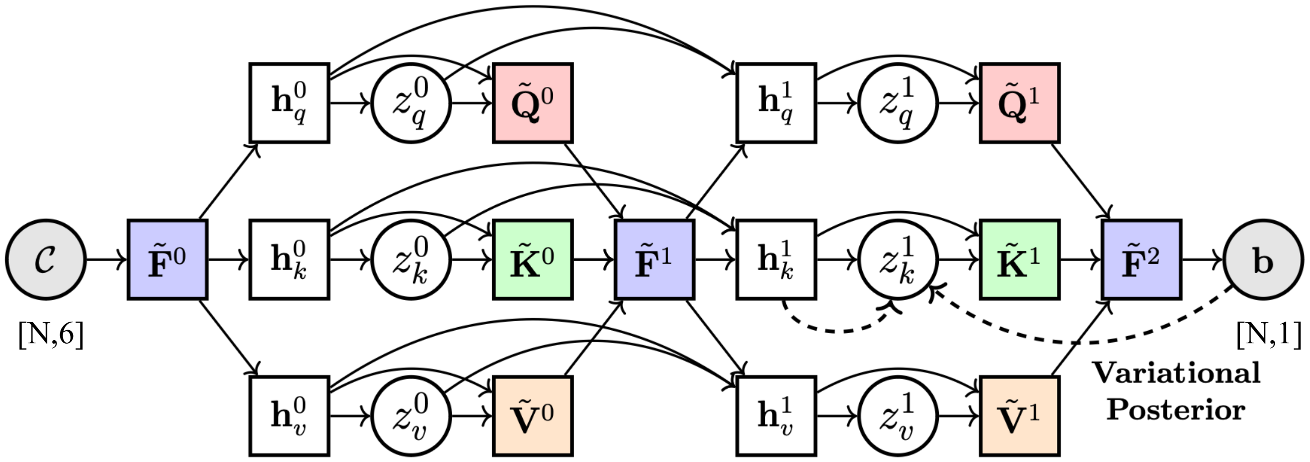
=> Bayesian-driven long-range dependencies

[1] PointDSC: Robust Point Cloud Registration using Deep Spatial Consistency. CVPR’2021.

Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ *Approach* - Variational Bayesian-based Non-local Network (VBNonlocal)

- Why VBNonlocal can achieve more discriminative long-range dependencies?
  - Variational posterior is label dependent => more discriminative
  - Label-dependent posterior to guide the prior distribution in training phase
  - Sample more discriminative feature from the learned prior in test phase

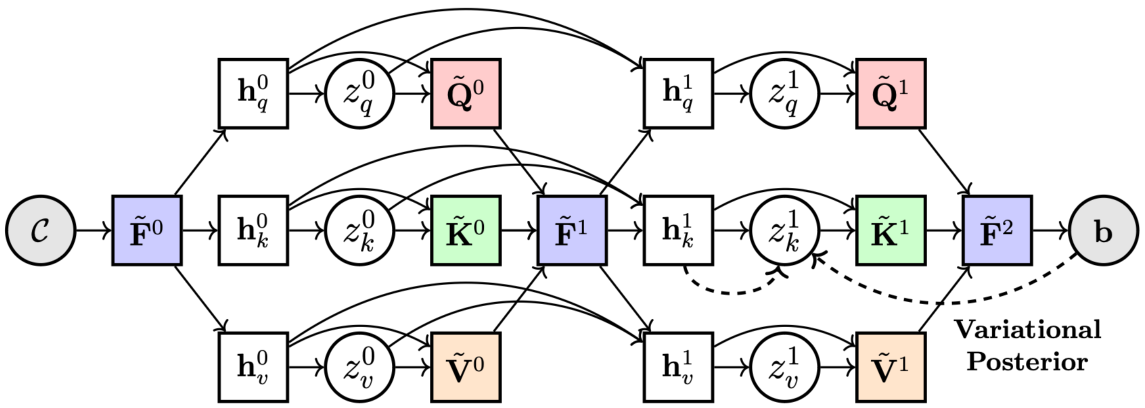


Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ ***Approach*** **- Variational Bayesian-based Non-local Network (VBNonlocal)**

(1) Hidden Q/K/V feat

$$\mathbf{h}_{q,i}^{l} = \mathrm{GRU}_q(\mathbf{h}_{q,i}^{l-1}, [\mathbf{z}_{q,i}^{l-1}, \tilde{\mathbf{F}}_i^{l-1}]),$$
$$\mathbf{h}_{k,i}^{l} = \mathrm{GRU}_k(\mathbf{h}_{k,i}^{l-1}, [\mathbf{z}_{k,i}^{l-1}, \tilde{\mathbf{F}}_i^{l-1}]),$$
$$\mathbf{h}_{v,i}^{l} = \mathrm{GRU}_v(\mathbf{h}_{v,i}^{l-1}, [\mathbf{z}_{v,i}^{l-1}, \tilde{\mathbf{F}}_i^{l-1}]),$$



$$\tilde{\mathbf{Q}}_i^l = f_\theta^q([\mathbf{z}_{q,i}^l, \mathbf{h}_{q,i}^l]),$$
$$\tilde{\mathbf{K}}_i^l = f_\theta^k([\mathbf{z}_{k,i}^l, \mathbf{h}_{k,i}^l])$$
$$\tilde{\mathbf{V}}_i^l = f_\theta^v([\mathbf{z}_{v,i}^l, \mathbf{h}_{v,i}^l])$$

(3) Bayesian-driven Q/K/V feat.

$$\mathbf{z}_{q,i}^l \sim p_\theta(z_{q,i}^l \mid \mathbf{h}_{q,i}^l)$$
$$\mathbf{z}_{k,i}^l \sim p_\theta(z_{k,i}^l \mid \mathbf{h}_{k,i}^l)$$
$$\mathbf{z}_{v,i}^l \sim p_\theta(z_{v,i}^l \mid \mathbf{h}_{v,i}^l)$$

(2) Prior Q/K/V feat. distrib.

KL

$$\mathbf{z}_{q,i}^l \sim q_\phi(\mathbf{z}_{q,i}^l \mid [\mathbf{h}_{q,i}^l, [b_i]_{\times k}])$$
$$\mathbf{z}_{k,i}^l \sim q_\phi(\mathbf{z}_{k,i}^l \mid \mathbf{h}_{k,i}^l, [b_i]_{\times k}])$$
$$\mathbf{z}_{v,i}^l \sim q_\phi(\mathbf{z}_{v,i}^l \mid \mathbf{h}_{v,i}^l, [b_i]_{\times k}]),$$

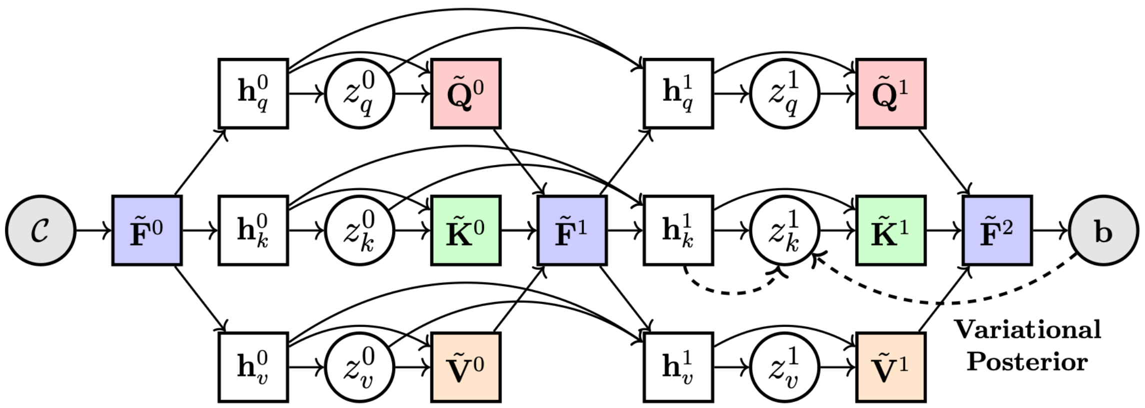
Posterior Q/K/V feat. distrib.

Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ *Approach* - Variational Bayesian-based Non-local Network (VBNonlocal)

(1) Hidden Q/K/V feat

$$\mathbf{h}^l_{q,i} = \mathrm{GRU}_q(\mathbf{h}^{l-1}_{q,i}, [\mathbf{z}^{l-1}_{q,i}, \tilde{\mathbf{F}}^{l-1}_i]),$$
$$\mathbf{h}^l_{k,i} = \mathrm{GRU}_k(\mathbf{h}^{l-1}_{k,i}, [\mathbf{z}^{l-1}_{k,i}, \tilde{\mathbf{F}}^{l-1}_i]),$$
$$\mathbf{h}^l_{v,i} = \mathrm{GRU}_v(\mathbf{h}^{l-1}_{v,i}, [\mathbf{z}^{l-1}_{v,i}, \tilde{\mathbf{F}}^{l-1}_i]),$$



$$\tilde{\mathbf{Q}}^l_i = f^q_\theta([\mathbf{z}^l_{q,i}, \mathbf{h}^l_{q,i}]),$$
$$\tilde{\mathbf{K}}^l_i = f^k_\theta([\mathbf{z}^l_{k,i}, \mathbf{h}^l_{k,i}])$$
$$\tilde{\mathbf{V}}^l_i = f^v_\theta([\mathbf{z}^l_{v,i}, \mathbf{h}^l_{v,i}])$$

(3) Bayesian-driven Q/K/V feat.

$$\mathbf{z}^l_{q,i} \sim p_\theta(z^l_{q,i} \mid \mathbf{h}^l_{q,i})$$
$$\mathbf{z}^l_{k,i} \sim p_\theta(z^l_{k,i} \mid \mathbf{h}^l_{k,i})$$
$$\mathbf{z}^l_{v,i} \sim p_\theta(z^l_{v,i} \mid \mathbf{h}^l_{v,i})$$

(2) Prior Q/K/V feat. distrib.

KL

$$\mathbf{z}^l_{q,i} \sim q_\phi(\mathbf{z}^l_{q,i} \mid [\mathbf{h}^l_{q,i}, [b_i]_{\times k}])$$
$$\mathbf{z}^l_{k,i} \sim q_\phi(\mathbf{z}^l_{k,i} \mid \mathbf{h}^l_{k,i}, [b_i]_{\times k}])$$
$$\mathbf{z}^l_{v,i} \sim q_\phi(\mathbf{z}^l_{v,i} \mid \mathbf{h}^l_{v,i}, [b_i]_{\times k}]),$$

Posterior Q/K/V feat. distrib.

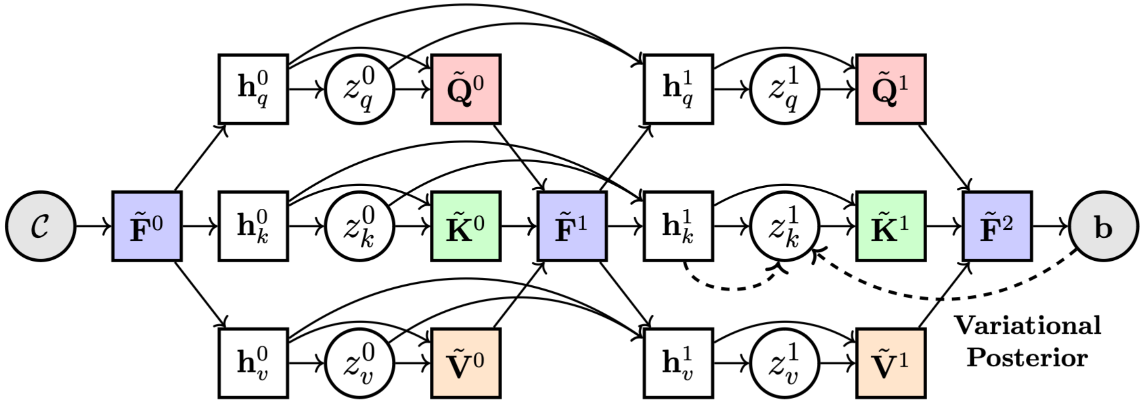
Robust Outlier Rejection for 3D Registration with Variational Bayes

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

➢ *Approach* - Variational Bayesian-based Non-local Network (VBNonlocal)

(1) Hidden Q/K/V feat

$$
\begin{aligned}
\mathbf{h}_{q,i}^{l} &= \mathrm{GRU}_q(\mathbf{h}_{q,i}^{l-1}, [\mathbf{z}_{q,i}^{l-1}, \tilde{\mathbf{F}}_i^{l-1}]),\\
\mathbf{h}_{k,i}^{l} &= \mathrm{GRU}_k(\mathbf{h}_{k,i}^{l-1}, [\mathbf{z}_{k,i}^{l-1}, \tilde{\mathbf{F}}_i^{l-1}]),\\
\mathbf{h}_{v,i}^{l} &= \mathrm{GRU}_v(\mathbf{h}_{v,i}^{l-1}, [\mathbf{z}_{v,i}^{l-1}, \tilde{\mathbf{F}}_i^{l-1}]),
\end{aligned}
$$



$$
\begin{aligned}
\tilde{\mathbf{Q}}_i^l &= f_\theta^q([\mathbf{z}_{q,i}^l, \mathbf{h}_{q,i}^l]),\\
\tilde{\mathbf{K}}_i^l &= f_\theta^k([\mathbf{z}_{k,i}^l, \mathbf{h}_{k,i}^l])\\
\tilde{\mathbf{V}}_i^l &= f_\theta^v([\mathbf{z}_{v,i}^l, \mathbf{h}_{v,i}^l])
\end{aligned}
$$

(3) Bayesian-driven Q/K/V feat.

$$
\begin{aligned}
\mathbf{z}_{q,i}^l &\sim p_\theta(z_{q,i}^l \mid \mathbf{h}_{q,i}^l)\\
\mathbf{z}_{k,i}^l &\sim p_\theta(z_{k,i}^l \mid \mathbf{h}_{k,i}^l)\\
\mathbf{z}_{v,i}^l &\sim p_\theta(z_{v,i}^l \mid \mathbf{h}_{v,i}^l)
\end{aligned}
$$

(2) Prior Q/K/V feat. distrib.

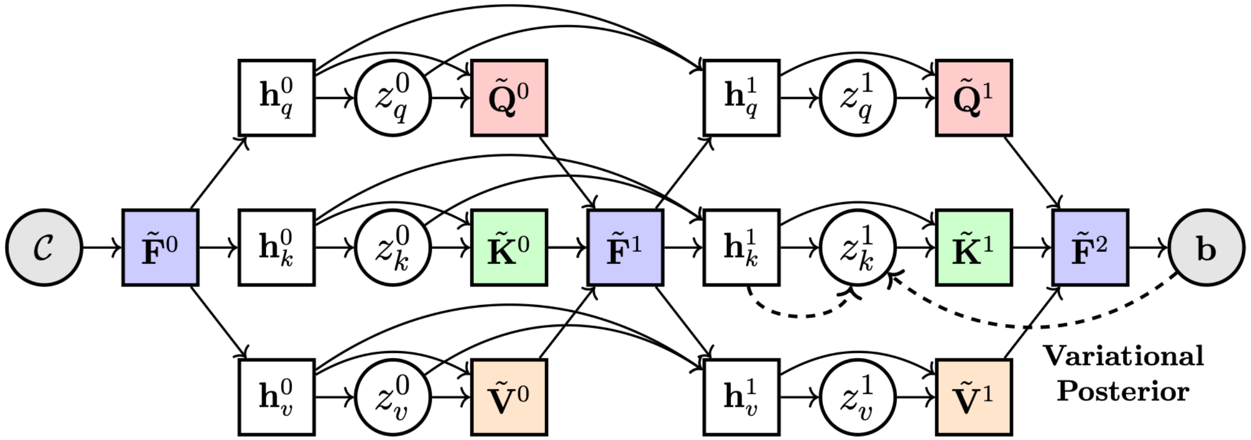
KL

$$
\begin{aligned}
\mathbf{z}_{q,i}^l &\sim q_\phi(\mathbf{z}_{q,i}^l \mid [\mathbf{h}_{q,i}^l, [b_i]_{\times k}])\\
\mathbf{z}_{k,i}^l &\sim q_\phi(\mathbf{z}_{k,i}^l \mid \mathbf{h}_{k,i}^l, [b_i]_{\times k}])\\
\mathbf{z}_{v,i}^l &\sim q_\phi(\mathbf{z}_{v,i}^l \mid \mathbf{h}_{v,i}^l, [b_i]_{\times k}]),
\end{aligned}
$$

Posterior Q/K/V feat. distrib.

Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ ***Approach* – Variational low bound**



$$\mathrm{ELBO}(\theta, \phi) = \mathbb{E}_{\prod_{l=0}^{L-1} q_\phi(z^l_{q,k,v} | \mathbf{h}^l_{q,k,v}, \mathbf{b})} \left[\ln y_\theta(\mathbf{b} \mid \tilde{\mathbf{F}}^L) \right] -$$

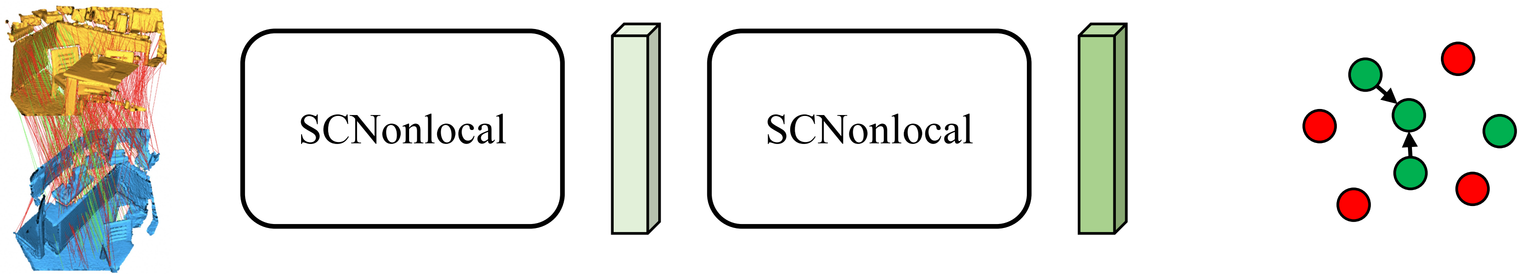
$$\sum_{l=0}^{L-1} \mathbb{E}_{q_\phi} \left[\mathrm{D_{KL}} \left(q_\phi(z^l_{q,k,v} \mid \mathbf{h}^l_{q,k,v}, \mathbf{b}) || p_\theta(z^l_{q,k,v} \mid \mathbf{h}^l_{q,k,v}) \right) \right]$$

Robust Outlier Rejection for 3D Registration with Variational Bayes

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

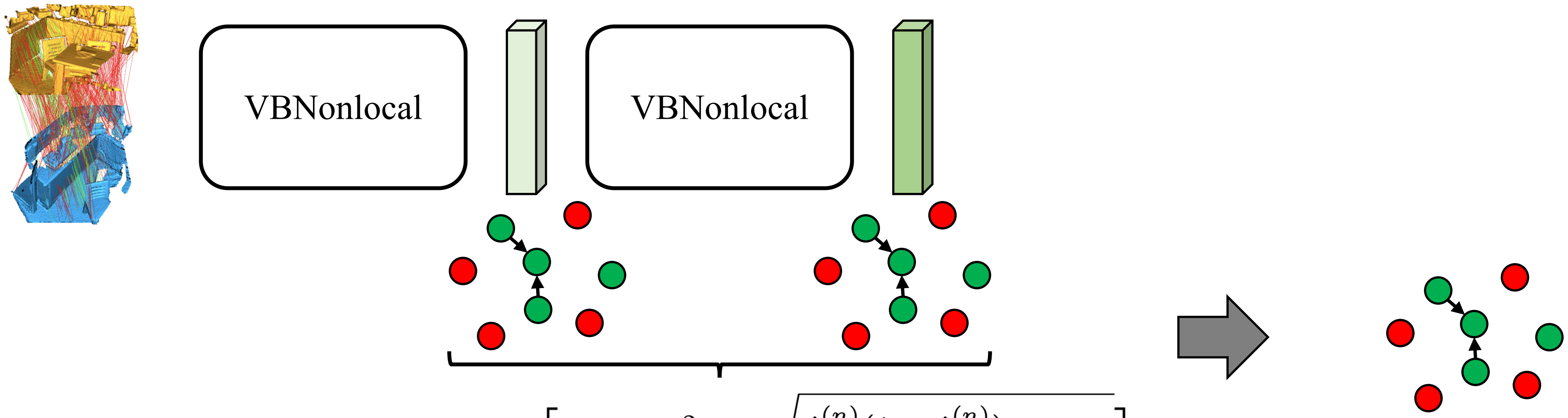
➢ *Approach* - Voting-based Inlier Searching

PointDSC:



Feature similarity-based Inlier clustering

Ours:



Wilson score $$\mathbf{W}_{i,j}^{n} = \frac{1}{1+\frac{z^2}{n}}\left[\hat{p}_{i,j}^{(n)} + \frac{z^2}{2n} - z\sqrt{\frac{\hat{p}_{i,j}^{(n)}(1-\hat{p}_{i,j}^{(n)})}{n} + \frac{z^2}{4n^2}}\right]$$

Wilson score-based clustered inliers

Robust Outlier Rejection for 3D Registration with Variational Bayes

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

➢ Experimental Results - *Comparison with SOTA Methods*

- 3DMatch & KITTI

Models	**3DMatch** (FCGF)			**3DMatch** (FPFH)			**KITTI** (FCGF)			**KITTI** (FPFH)			
	RR(↑)	RE(↓)	TE(↓)	RR(↑)	RE(↓)	TE(↓)	RR(↑)	RE(↓)	TE(↓)	RR(↑)	RE(↓)	TE(↓)	Sec.
FGR [52]	79.17	2.93	8.56	41.10	4.05	10.09	96.58	0.38	22.30	1.26	1.69	47.18	1.39
SM [26]	86.57	2.29	7.07	55.82	2.94	8.13	96.58	0.50	19.88	75.50	0.66	15.01	0.02
RANSAC [16]	91.50	2.49	7.54	73.57	3.55	10.04	97.66	**0.28**	22.61	89.37	1.22	25.88	6.43
TEASER++ [45]	85.77	2.73	8.66	75.48	2.48	7.31	83.24	0.84	**12.48**	64.14	1.04	14.85	0.07
DGR [10]	91.30	2.40	7.48	69.13	3.78	10.80	95.14	0.43	23.28	73.69	1.67	34.74	1.36
DHVR [25]	89.40	2.19	6.95	67.10	2.56	7.67	–	–	–	–	–	–	0.40
SC2_PCR [9]	93.10	**2.04**	6.53	**83.92**	2.09	6.66	97.48	0.33	20.66	97.84	0.39	9.09	0.09
PointDSC [2]	92.42	2.05	**6.49**	77.51	**2.08**	**6.51**	97.66	0.47	20.88	98.20	0.58	7.27	0.11
VBReg	**93.53**	**2.04**	**6.49**	82.75	2.14	6.77	**98.02**	0.32	20.91	**98.92**	**0.32**	**7.17**	0.22

Table 1. Quantitative comparison on 3DMatch [50] and KITTI [19] benchmark datasets with descriptors FCGF and FPFH. The registration speed is achieved by computing the averaged time cost on 3DMatch with FCGF descriptor.

Robust Outlier Rejection for 3D Registration with Variational Bayes

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

➢ Experimental Results - *Comparison with SOTA Methods*

- 3DLoMatch

Feature	Model	5000	2500	1000	500	250
FCGF	FGR [52]	18.6	19.4	16.9	16.0	12.4
	SM [26]	32.4	31.3	31.4	28.0	23.5
	RANSAC [16]	37.6	37.2	35.9	32.1	25.9
	TEASER++ [45]	42.8	42.4	39.5	34.5	25.7
	DHVR [25]	50.4	49.6	46.4	41.0	34.6
	SC2_PCR [9]	57.4	56.5	51.8	46.4	**36.2**
	TR_DE [8]	49.5	50.4	48.4	43.4	34.3
	PointDSC [2]	55.8	52.6	46.8	37.7	26.7
	VBReg	**58.3**	**56.8**	**52.9**	**47.2**	34.5
Predator	FGR [52]	36.4	38.2	39.7	39.6	38.0
	SM [26]	53.8	55.1	55.4	54.5	50.2
	RANSAC [16]	62.3	62.8	62.4	61.5	58.2
	TEASER++ [45]	62.9	62.6	61.9	59.0	56.7
	DHVR [25]	67.2	67.3	66.1	64.6	60.5
	SC2_PCR [9]	69.5	69.5	68.6	65.2	62.0
	TR_DE [8]	64.0	64.8	61.7	58.8	56.5
	PointDSC [2]	68.1	67.3	66.5	63.4	60.5
	VBReg	**69.9**	**69.8**	**68.7**	**66.4**	**63.0**

Table 2. Registration recall (*RR*) with different numbers of points on 3DLoMatch benchmark dataset [21].

Robust Outlier Rejection for 3D Registration with Variational Bayes

➢ Experimental Results - *Ablation studies*

Model	3DMatch FCGF	3DMatch FPFH	3DLoMatch (FCGF) 5000	3DLoMatch (FCGF) 2500	3DLoMatch (FCGF) 1000	3DLoMatch (FCGF) 500	3DLoMatch (FCGF) 250	3DLoMatch (Predator) 5000	3DLoMatch (Predator) 2500	3DLoMatch (Predator) 1000	3DLoMatch (Predator) 500	3DLoMatch (Predator) 250	Sec.
PointDSC *w/* SCNonlocal^{xyz}	92.42	77.51	55.8	52.6	46.8	37.7	26.7	68.1	67.3	66.5	63.4	60.5	0.11
PointDSC *w/* VBNonlocal^{xyz}	93.04	81.21	57.4	55.1	50.9	41.9	27.6	69.9	68.2	68.7	65.2	61.5	0.17
PointDSC *w/* SCNonlocal^{feat}	92.36	77.76	54.6	50.6	44.9	36.8	25.4	69.2	68.6	67.9	63.5	59.9	0.13
PointDSC *w/* VBNonlocal^{feat}	92.98	80.96	56.8	55.1	50.9	42.5	27.5	69.6	69.0	68.0	65.2	60.9	0.18
PointDSC *w/* SCNonlocal^{cls}	92.98	78.99	54.1	52.2	46.0	38.7	27.7	67.6	66.9	67.2	63.7	60.2	0.11
PointDSC *w/* VBNonlocal^{feat}+Vote	93.16	81.45	**58.2**	55.9	52.6	44.8	30.5	70.2	69.1	68.7	66.4	63.2	0.20
PointDSC *w/* VBNonlocal^{feat}+Vote+CS	**93.53**	**82.62**	58.1	**56.8**	**52.9**	**47.6**	**34.6**	**70.5**	**69.7**	**69.5**	**66.7**	**64.9**	0.22
Iteration times $L = 6$	93.16	81.64	**58.3**	56.5	**53.5**	**48.5**	**35.7**	70.3	69.1	68.9	66.4	64.0	0.19
Iteration times $L = 9$	93.41	**83.12**	57.9	**56.8**	**53.5**	48.0	33.6	69.9	69.5	68.6	66.3	64.7	0.20
Iteration times $L = 12$*	**93.53**	82.62	58.1	**56.8**	52.9	47.6	34.6	**70.5**	**69.7**	**69.5**	**66.7**	**64.9**	0.22
Random feat. dim. $\tilde{d} = 32$	**93.53**	**82.62**	**59.0**	**57.8**	**54.4**	46.7	33.3	70.1	69.1	69.0	66.8	**65.4**	0.20
Random feat. dim. $\tilde{d} = 64$	93.41	81.95	58.3	56.0	53.3	47.2	**34.6**	70.1	**69.9**	69.0	**67.2**	64.9	0.21
Random feat. dim. $\tilde{d} = 128$*	**93.53**	**82.62**	58.1	56.8	52.9	**47.6**	**34.6**	**70.5**	69.7	**69.5**	66.7	64.9	0.22

Table 3. Ablation studies on 3DMatch [48] and 3DLoMatch [19] datasets. *SCNonlocal*: Spatial consistency-guided non-local network; *VBNonlocal*: Variational Bayesian-based non-local network; *Vote*: Voting-based inlier searching; *CS*: Conservative seed selection.

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Thanks!