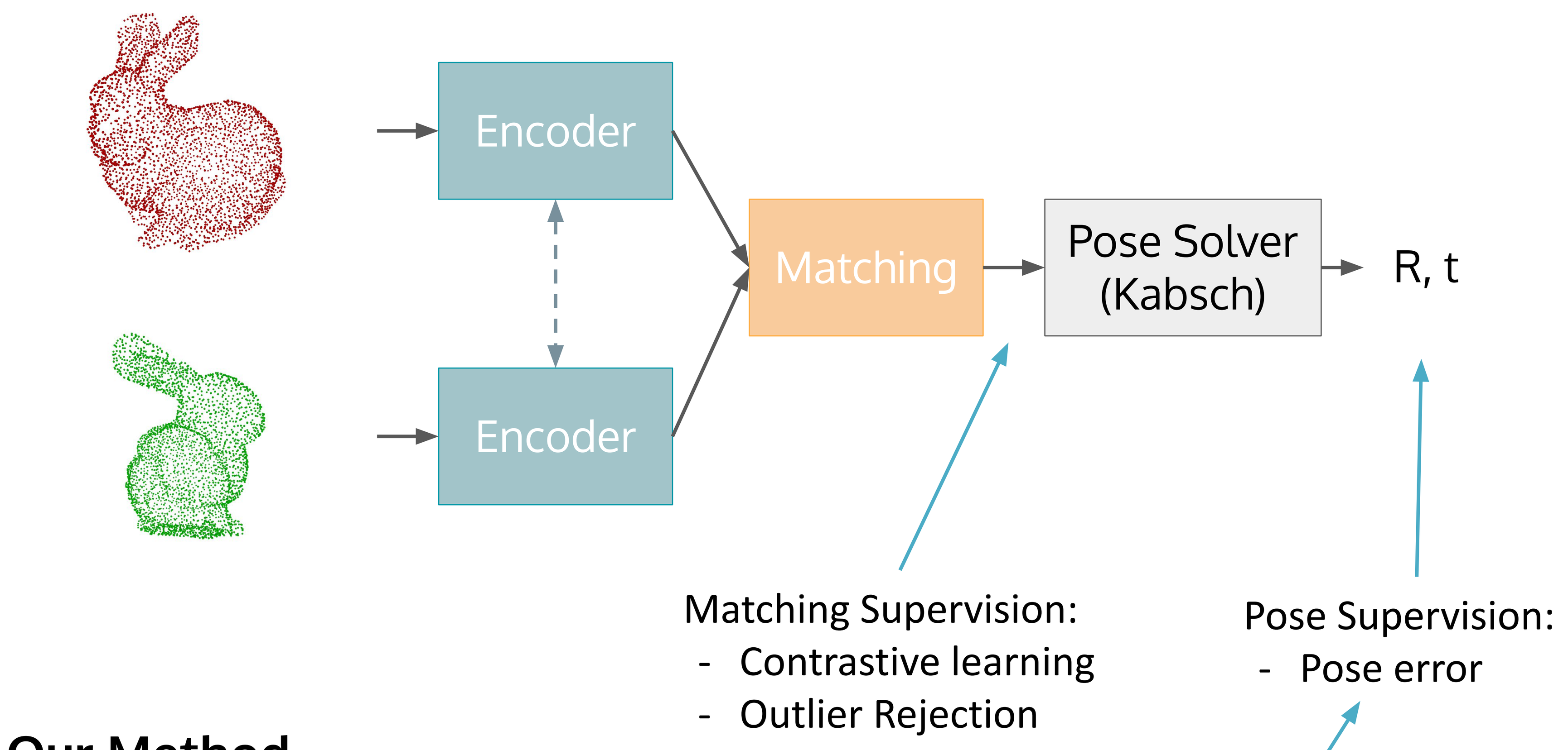
(Just) A Spoonful of Refinements Helps the Registration Error Go Down

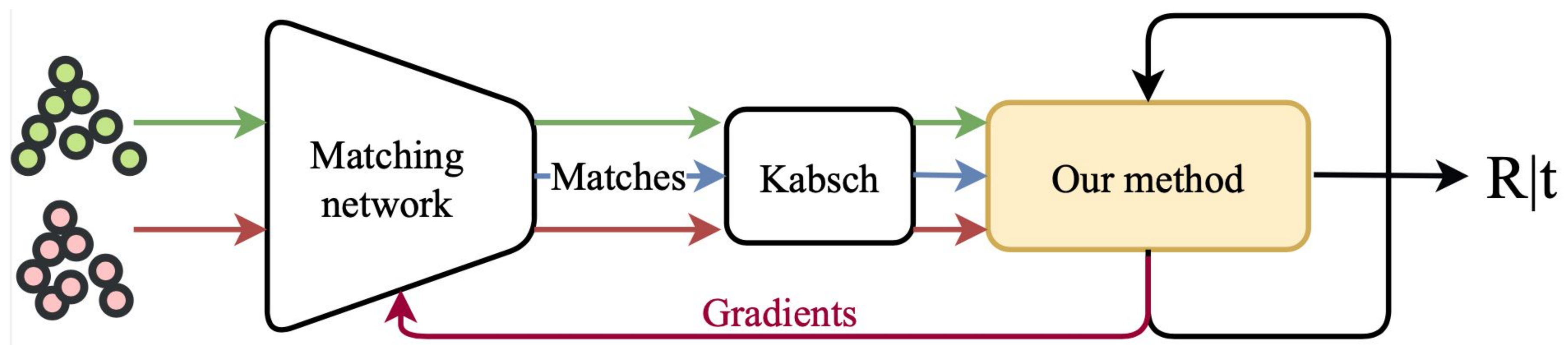
Sérgio Agostinho   Aljoša Ošep   Alessio Del Bue   Laura Leal-Taixé

ICCV 2021 OCTOBER 11-17 VIRTUAL

A traditional correspondence-based pipeline

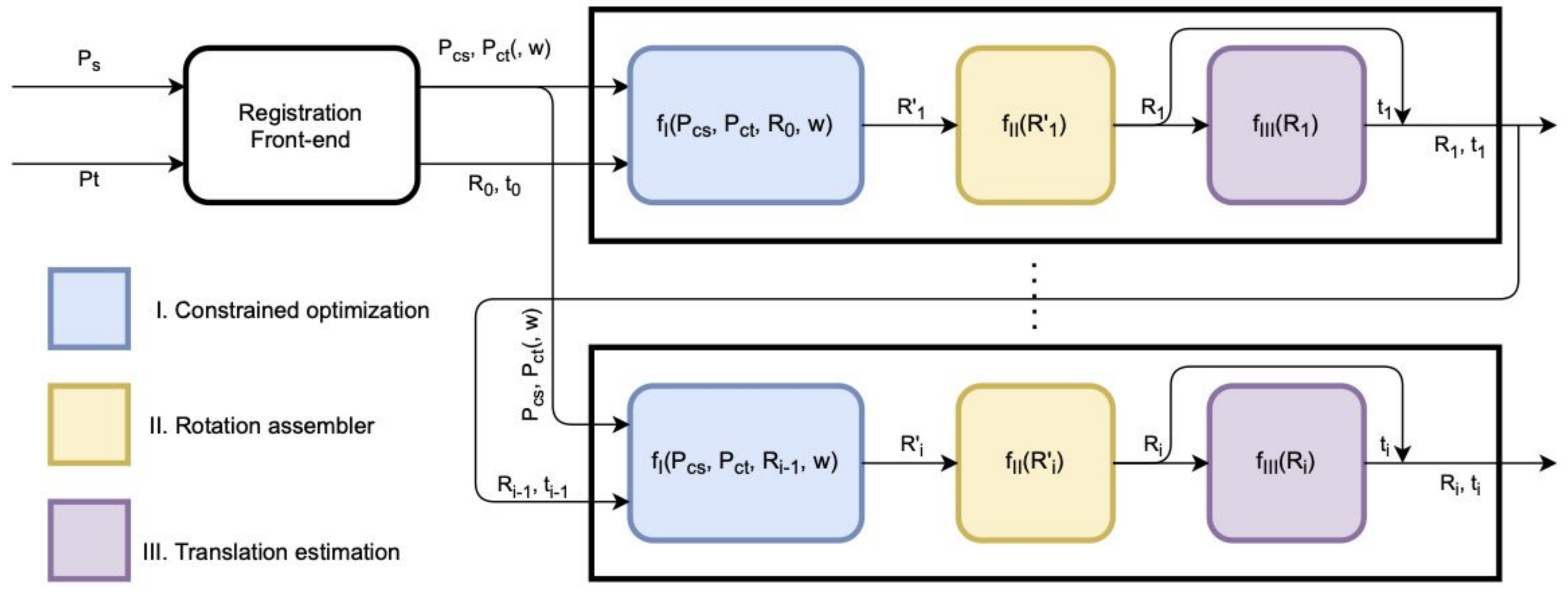


Our Method



- Parameter-free iterative layer
- New pose per iteration, augmenting pose supervision
- Only added during training
- Implicitly helps the matching network to produce better matches

What does it do? It takes an estimate produced by Kabsch and performs a number of iterative updates.



What happens?

1. The estimates produced by our layer diverge will occasionally diverge from Kabsch's estimate.
2. This divergence is paired with a higher penalty in the loss, conditioning the network to avoid the set of matches it provided.
3. This is beneficial for the host network, encouraging it to learn better correspondences.

Stage I - Constrained Optimization

$$\underset{\mathtt{R}}{\arg\min} \quad \sum_{i=1}^{N} w_i \|\tilde{\mathbf{p}}_{t_i} - \mathtt{R}\tilde{\mathbf{p}}_{s_i}\|\|^2$$

$$\text{s.t.} \qquad \mathtt{R}^\top\mathtt{R} = \mathtt{I}$$

$$\cancel{\det(\mathtt{R}) = 1}$$

linearize

Stage II - Rotation Assembler

Input:

$$\mathtt{R}' = \begin{bmatrix} | & | & | \\ \mathbf{r}'_1 & \mathbf{r}'_2 & \mathbf{r}'_3 \\ | & | & | \end{bmatrix}$$

Output:

$$\mathbf{r}_1 = \frac{\mathbf{r}'_1}{\|\mathbf{r}'_1\|}$$

$$\mathbf{r}_2 = \frac{(\mathtt{I} - \mathbf{r}_1\mathbf{r}_1^\top)\mathbf{r}'_2}{\|(\mathtt{I} - \mathbf{r}_1\mathbf{r}_1^\top)\mathbf{r}'_2\|}$$

$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$$

Stage III - Estimate translation

$$\mathbf{t} = \frac{\sum_{i=1}^{N} w_i(\mathbf{p}_{t_i} - \mathtt{R}\mathbf{p}_{s_i})}{\sum_{i=1}^{N} w_i} = \bar{\mathbf{p}_t} - \mathtt{R}\bar{\mathbf{p}_s}$$

Changes to the Loss

$$\mathcal{L} = \|\mathtt{R}^\top\mathtt{R}_{gt} - \mathtt{I}\|^2 + \|\mathbf{t} - \mathbf{t}_{gt}\|^2 + \lambda\|\theta\|^2$$

$$\mathcal{L} = \frac{1}{N_r+1}\sum_{i=1}^{N_r+1} \|\mathtt{R}_i^\top\mathtt{R}_{gt} - \mathtt{I}\|^2 + \frac{1}{N_r+1}\sum_{i=1}^{N_r+1} \|\mathbf{t}_i - \mathbf{t}_{gt}\|^2 + \lambda\|\theta\|^2$$

Results

- Objects rotated [0°, 45°] and translated [−0.5, 0.5] in each axis
- 5 iterations of our method

DCP on ModelNet40 - Unseen Categories

Model	RMSE(R)°	MAE(R)°	RMSE(t)	MAE(t)
ICP	29.876431	23.626110	0.293266	0.251916
Go-ICP [41]	13.865736	2.914169	0.022154	0.006219
FGR [47]	9.848997	1.445460	0.013503	0.002231
PointNetLK [13]	17.502113	5.280545	0.028007	0.007203
DCP-v2	3.150191	2.007210	0.005039	0.003703
DCP-v2 + ours	**2.051713**	**1.431898**	**0.004543**	**0.003333**

One half of the total object categories is used training and the other half is only used for evaluation.

RPM-Net on ModelNet40 - 70% surface overlap

Method	Anisotropic err. (Rot.)	Anisotropic err. (Trans.)	Isotropic err. (Rot.)	Isotropic err. (Trans.)	$\tilde{CD}$
ICP	13.719	0.132	27.250	0.280	0.0153
RPM	9.771	0.092	19.551	0.212	0.0081
FGR	19.266	0.090	30.839	0.192	0.0119
PointNetLK	15.931	0.142	29.725	0.297	0.0235
DCP-v2	6.380	0.083	12.607	0.169	0.0113
RPM-Net	0.893	0.0087	1.712	0.018	**0.00085**
RPM-Net + Ours	**0.826**	0.0081	1.575	0.017	**0.00085**
RPM-Net†	0.993	0.0087	1.861	0.018	0.00099
RPM-Net + Ours†	0.872	**0.0074**	**1.554**	**0.015**	0.00088

Best improvements once RPM-Net is trained without a loss term encouraging inliers †

Official Implementation

Soon!

Conclusions

- Parameter-free layer for correspondence-based registration networks.
- It implicitly improves matching quality, only through pose supervision.
- It is a guilt free addition because it does not hinder registration performance.