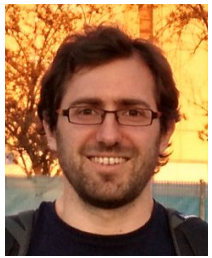
Neighbourhood Consensus Networks

Ignacio Rocco



Mircea Cimpoi

Relja Arandjelović

Akihiko Torii

Tomas Pajdla

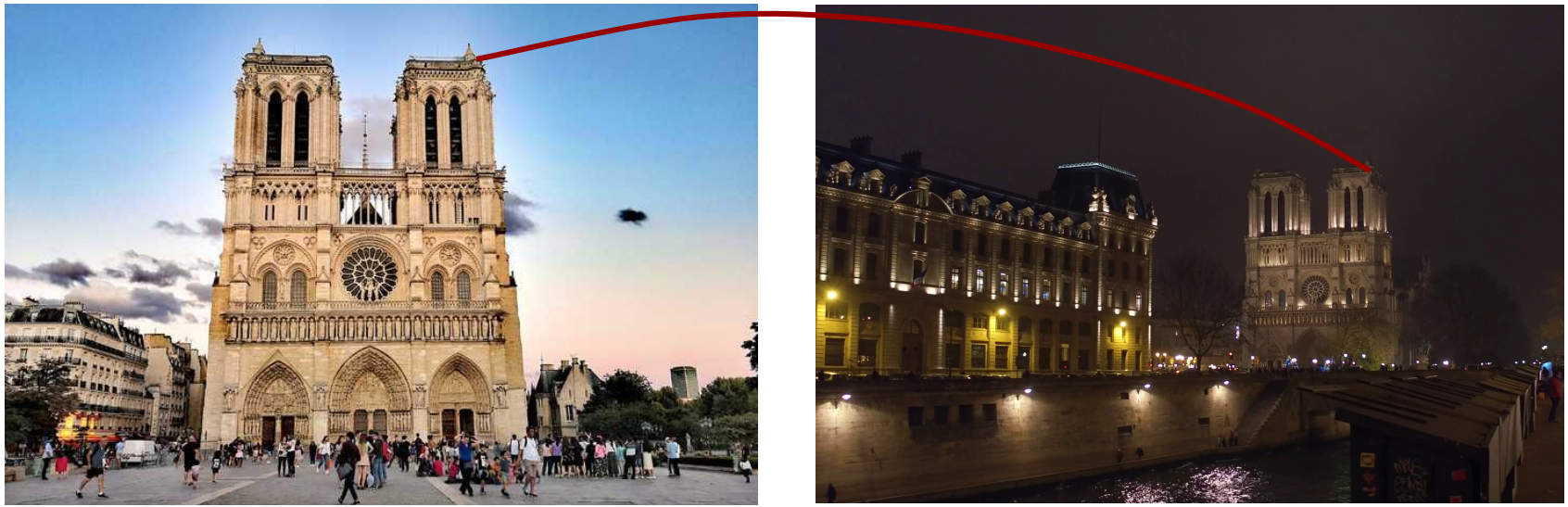
Josef Sivic

32nd Conference on Neural Information Processing Systems (NeurIPS), Montréal, 2018

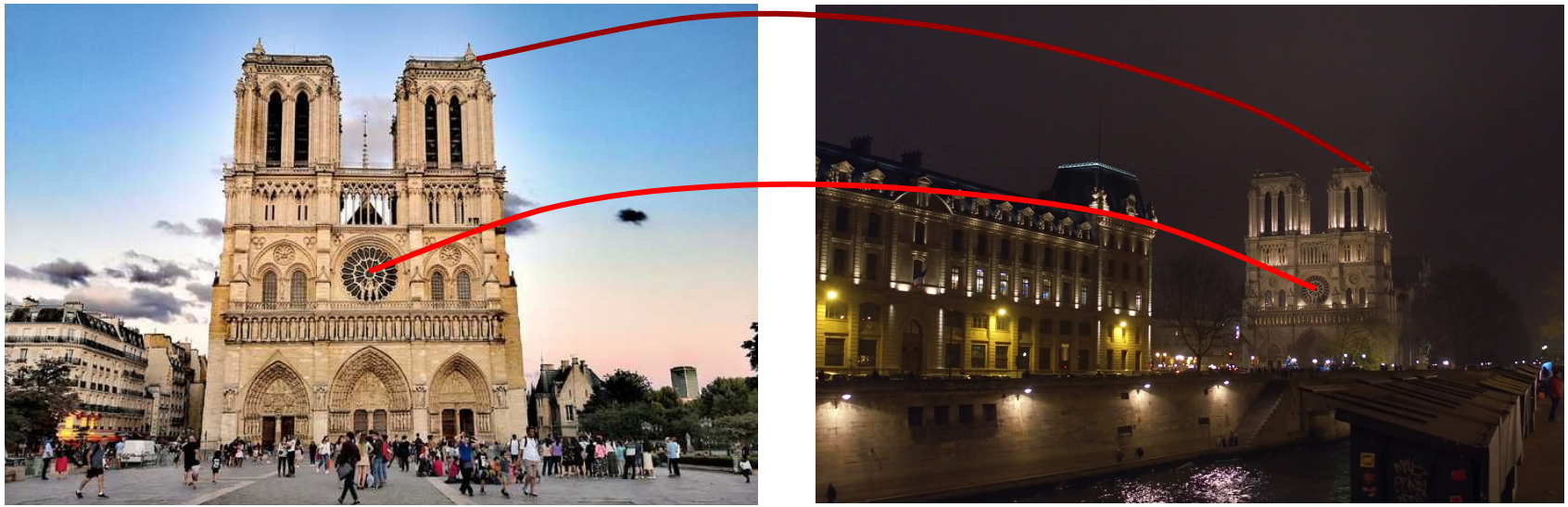
Task: *find pixel-level image correspondences*

Task: *find pixel-level image correspondences*

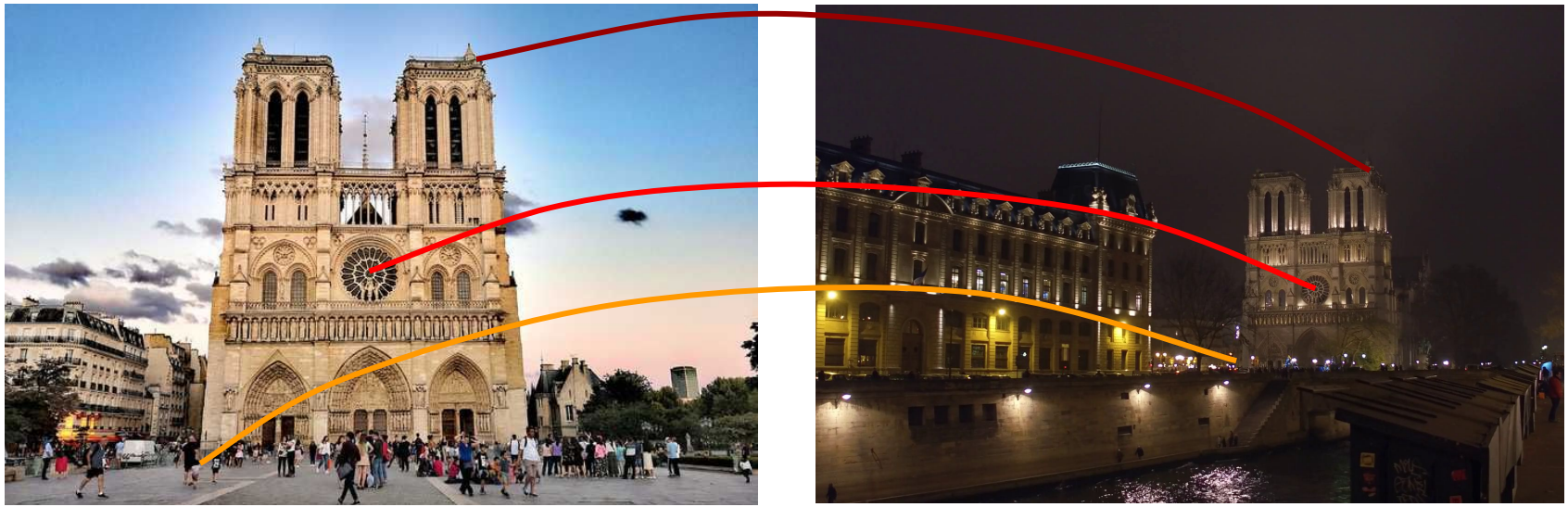
Task: *find pixel-level image correspondences*



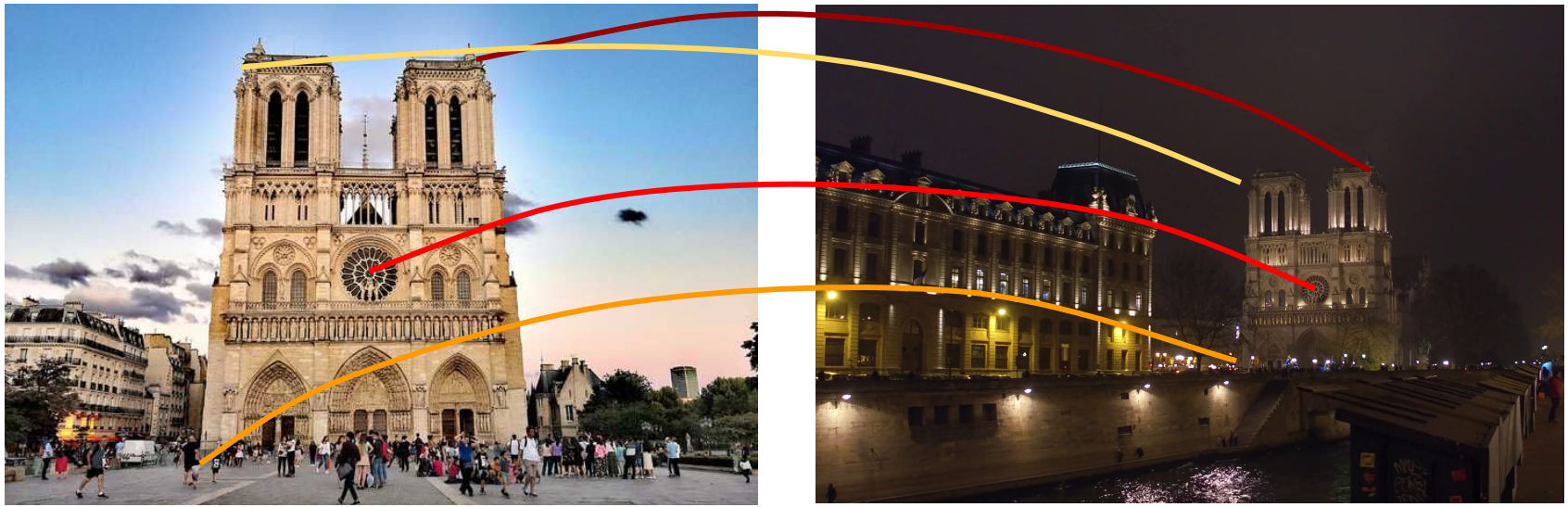
Task: *find pixel-level image correspondences*



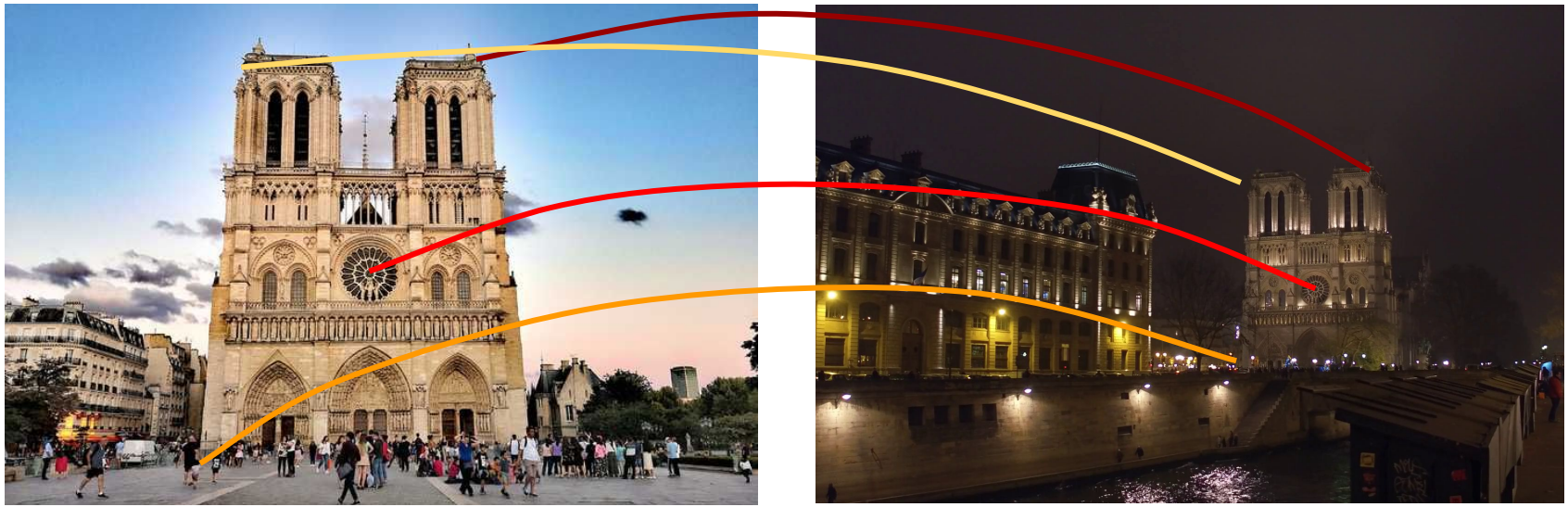
Task: *find pixel-level image correspondences*



Task: *find pixel-level image correspondences*



Task: *find pixel-level image correspondences*



day-night matching

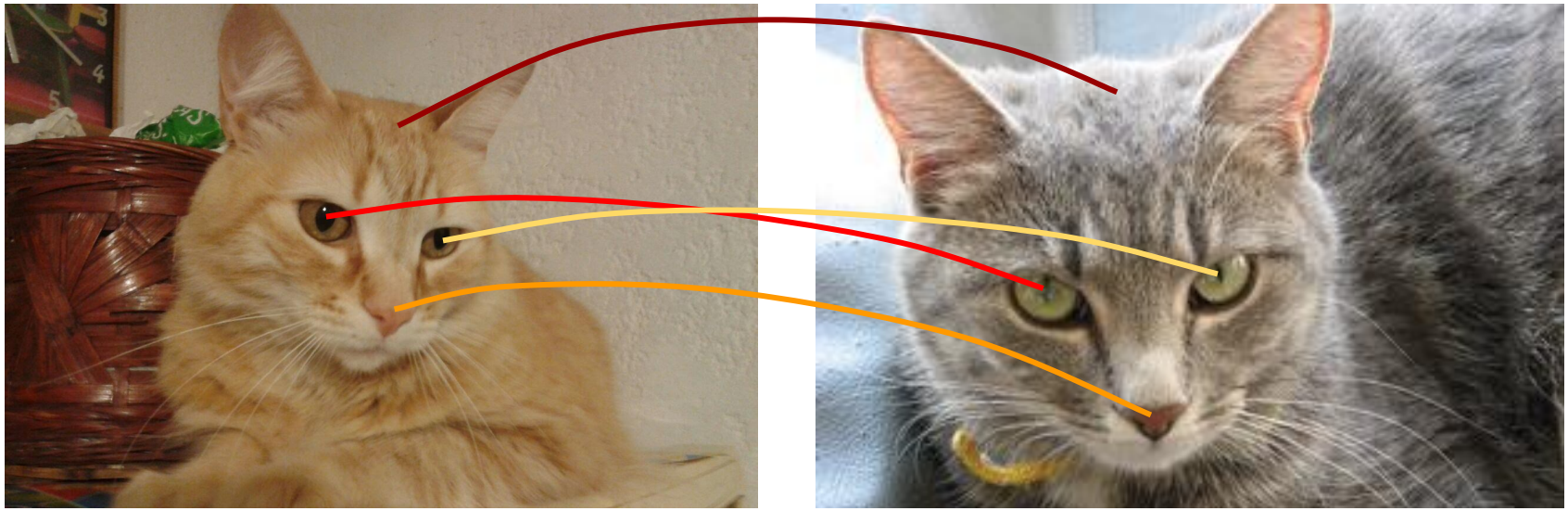
Task: *find pixel-level image correspondences*

changes across time

Task: *find pixel-level image correspondences*

repetitive structures

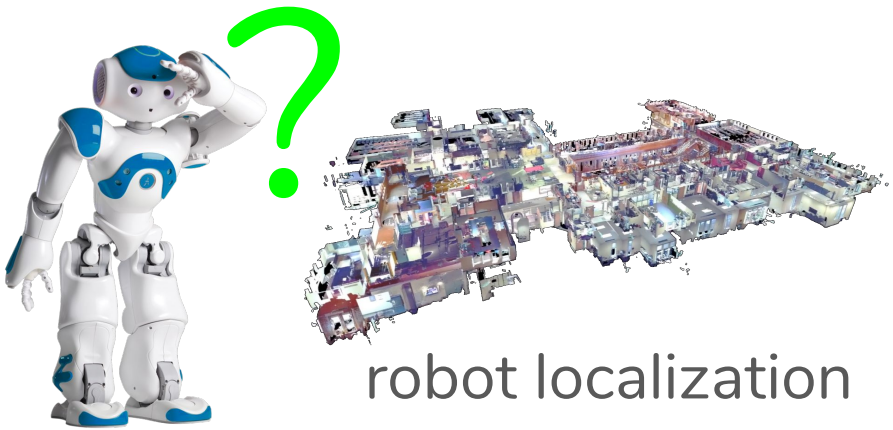
Task: *find pixel-level image correspondences*



intra-class variation

Motivation:

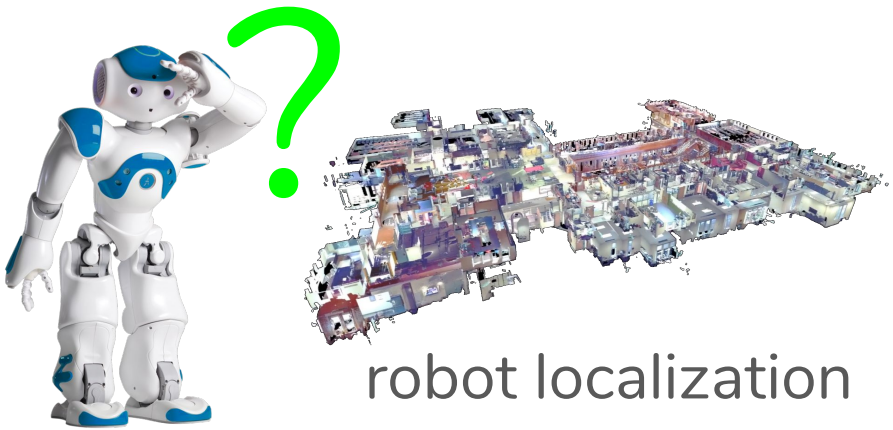
Motivation:



robot localization

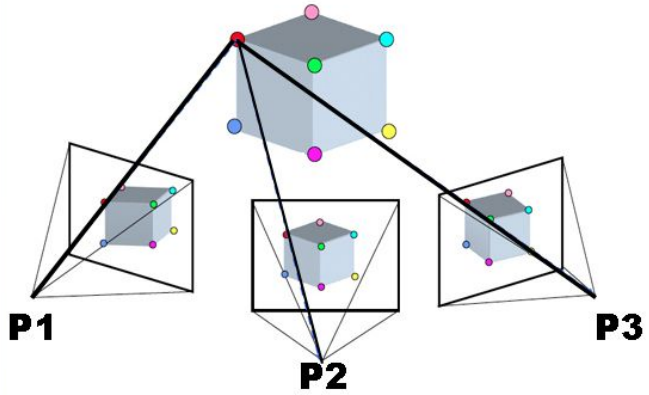
Taira et al. CVPR 2018

Motivation:



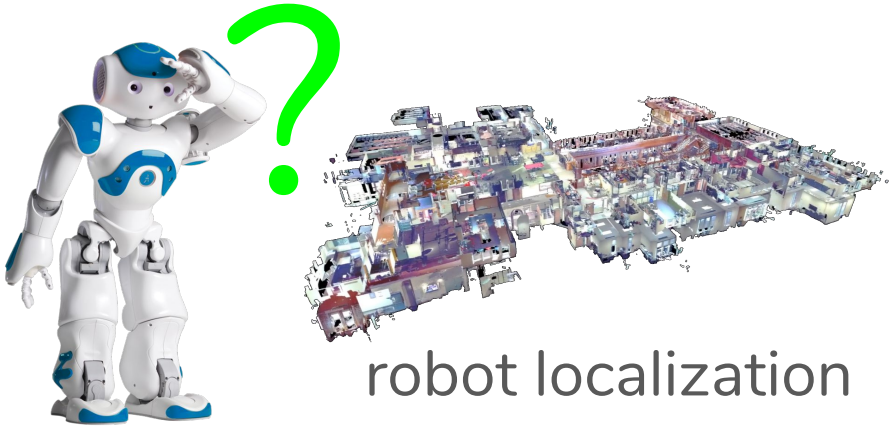
robot localization

3d reconstruction



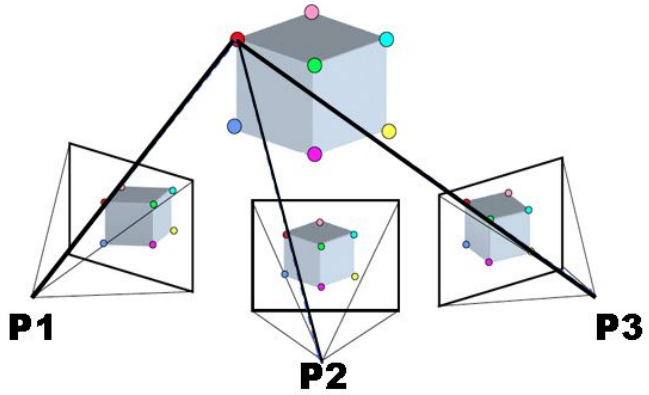
Taira et al. CVPR 2018
Agarwal et al. CACM 2011

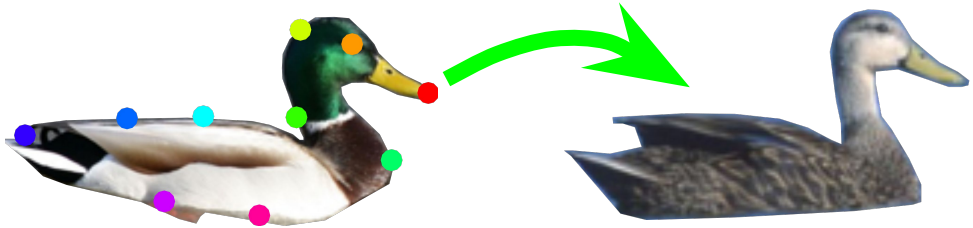
Motivation:



robot localization

3d reconstruction



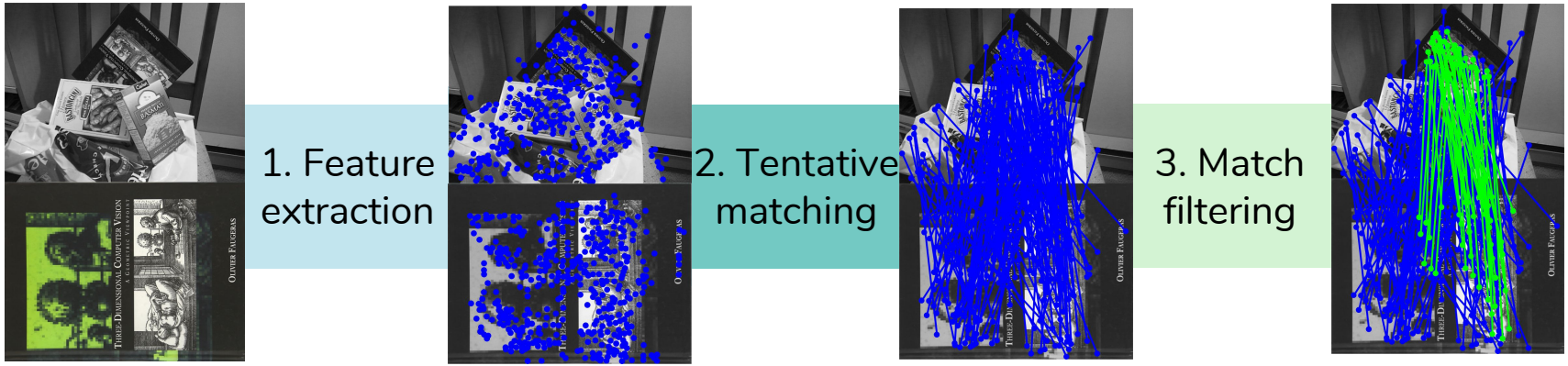


annotation transfer

Taira et al. CVPR 2018
Agarwal et al. CACM 2011
Rocco et al. CVPR 2017

Review - *classical pipeline*

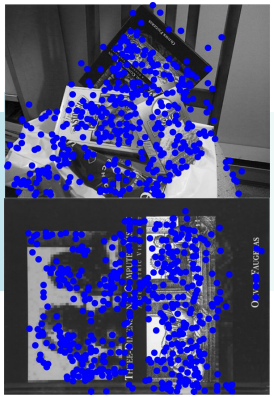
Review - *classical pipeline*



Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

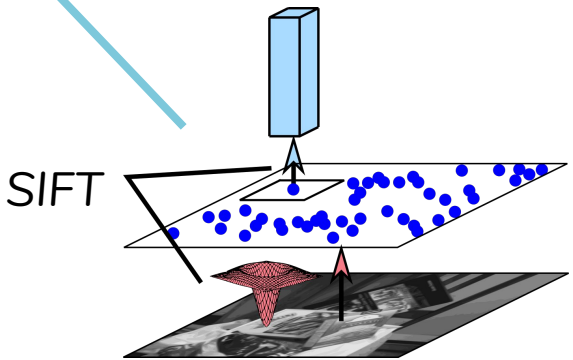
Review - *classical pipeline*

1. Feature extraction



2. Tentative matching

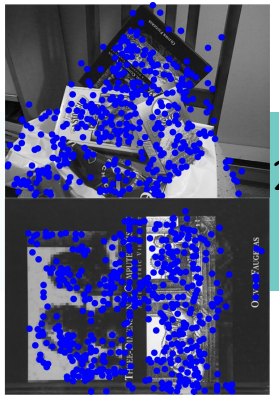
3. Match filtering



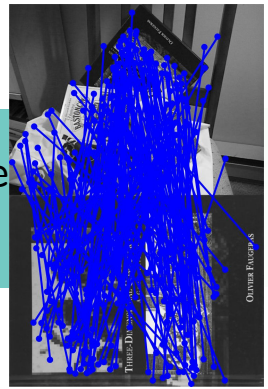
Lowe et al. IJCV 2004

Review - *classical pipeline*

1. Feature extraction

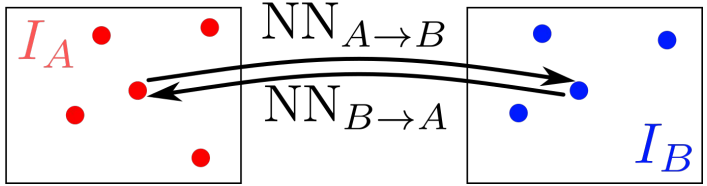


2. Tentative matching



3. Match filtering

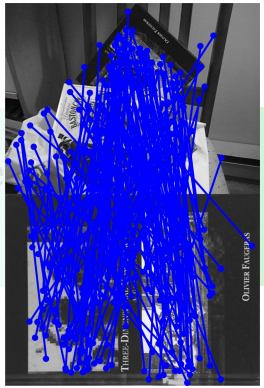
Mutual matches



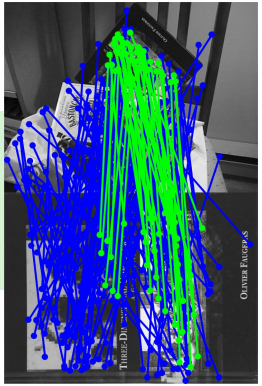
Review - *classical pipeline*

1. Feature extraction

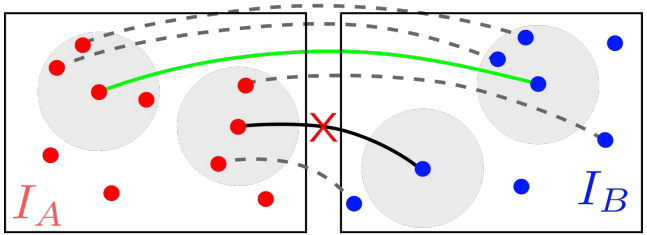
2. Tentative matching



3. Match filtering



Neighbourhood consensus

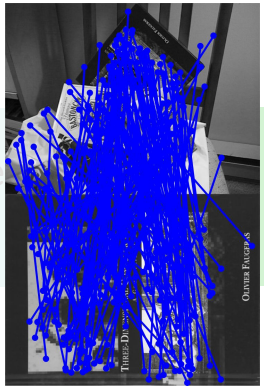


Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

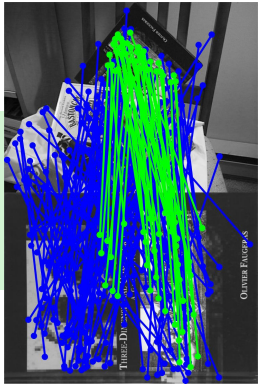
Review - *classical pipeline*

1. Feature extraction

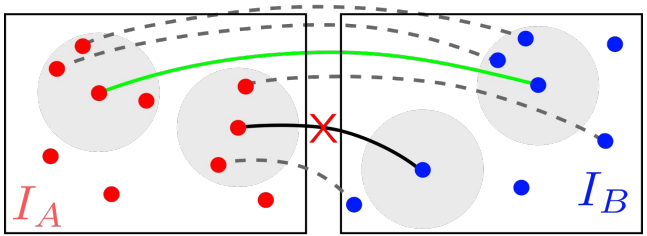
2. Tentative matching



3. Match filtering



Neighbourhood consensus

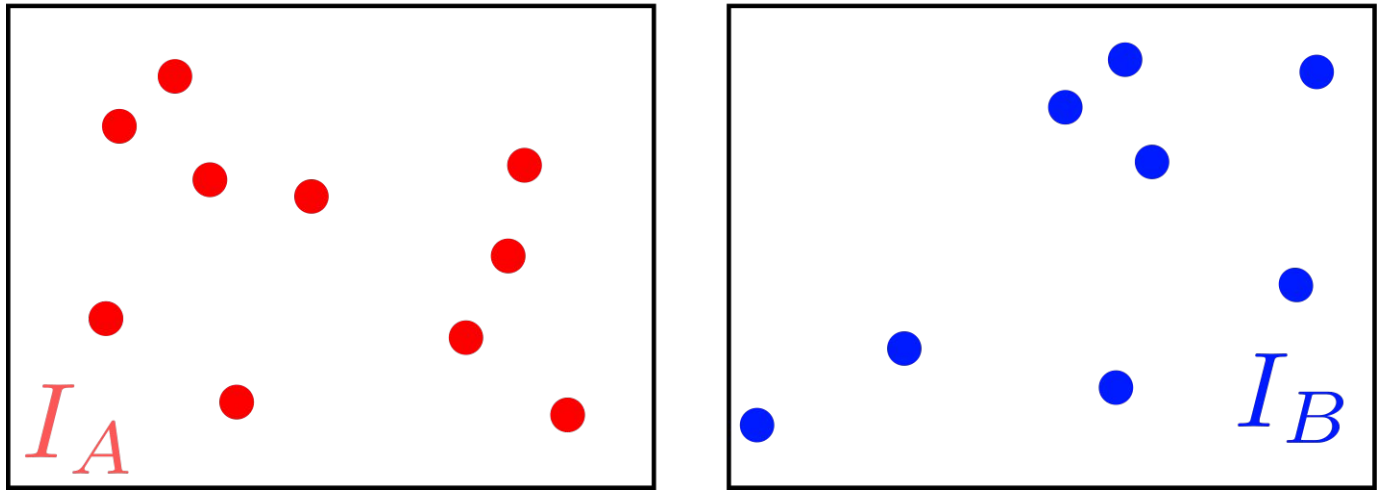


Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus

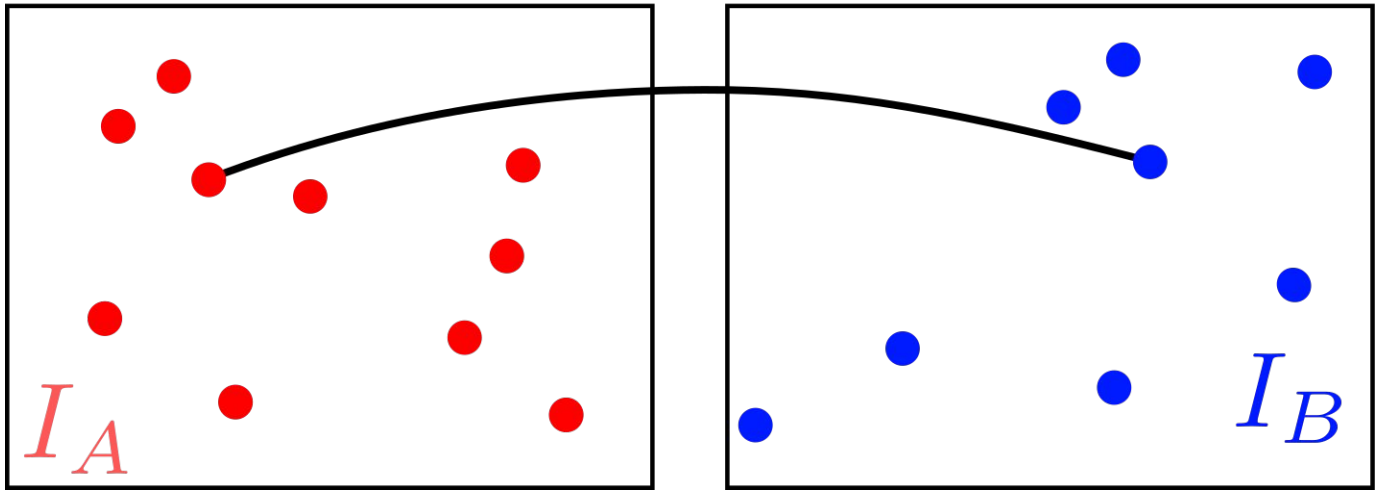
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



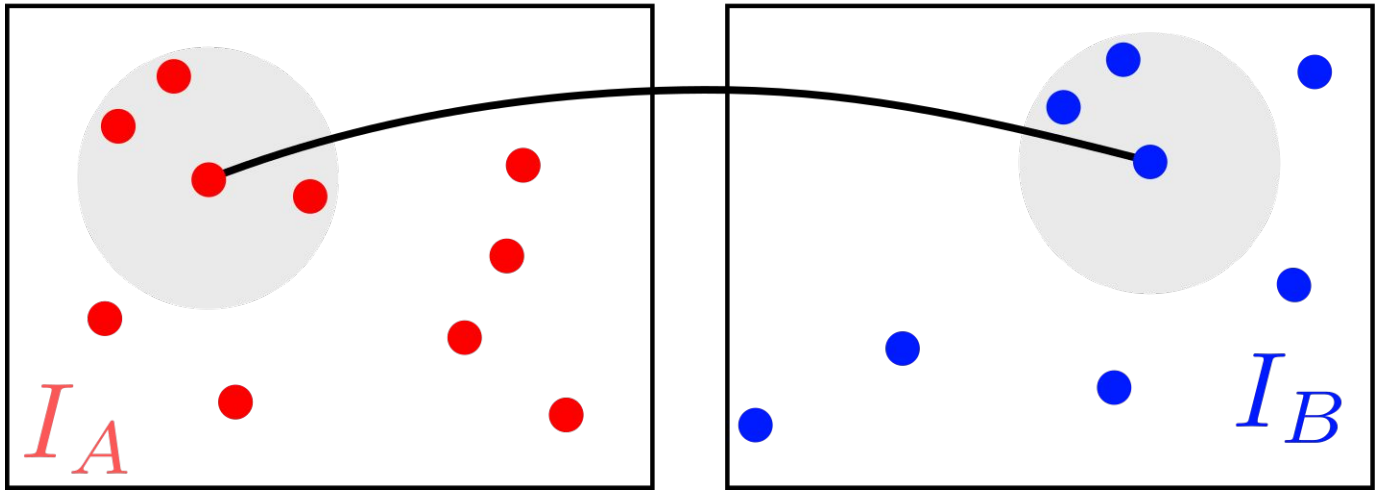
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



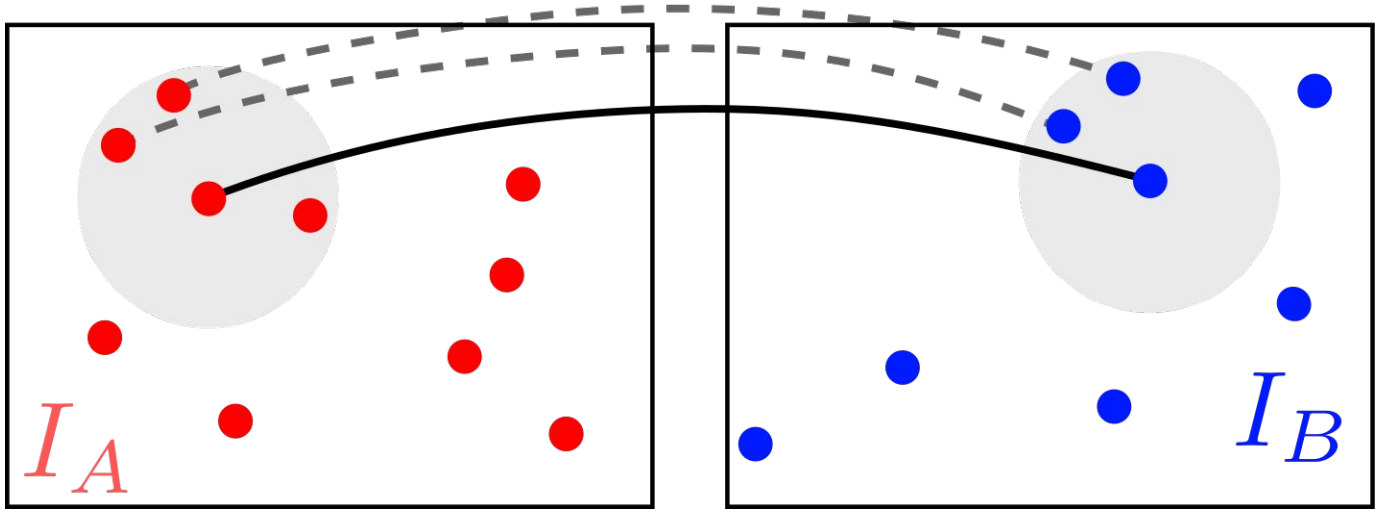
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



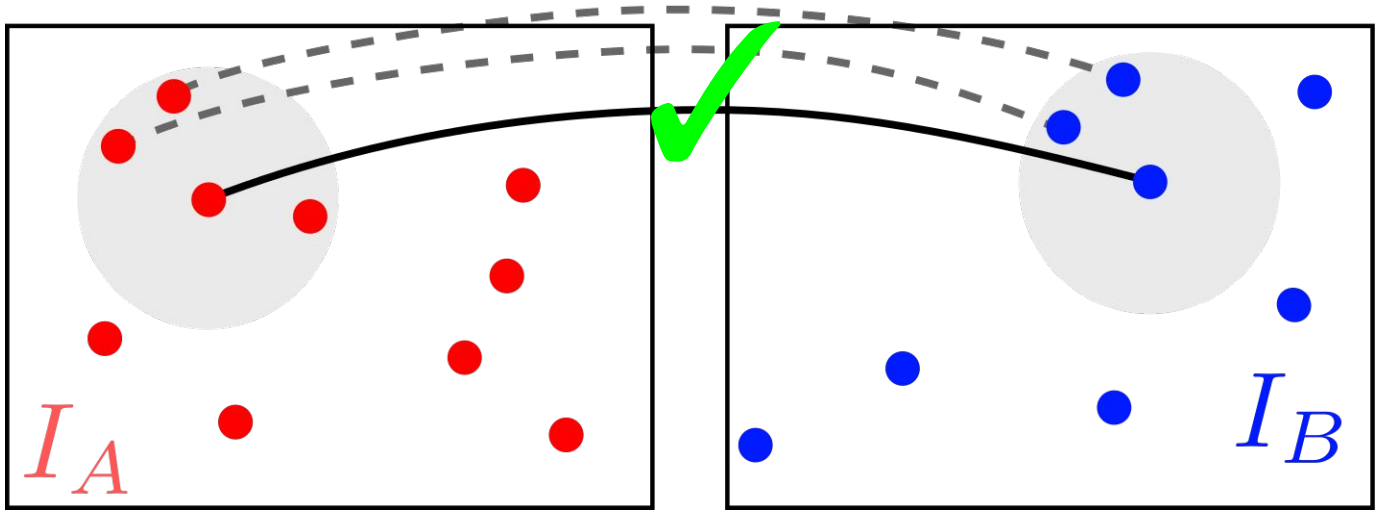
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



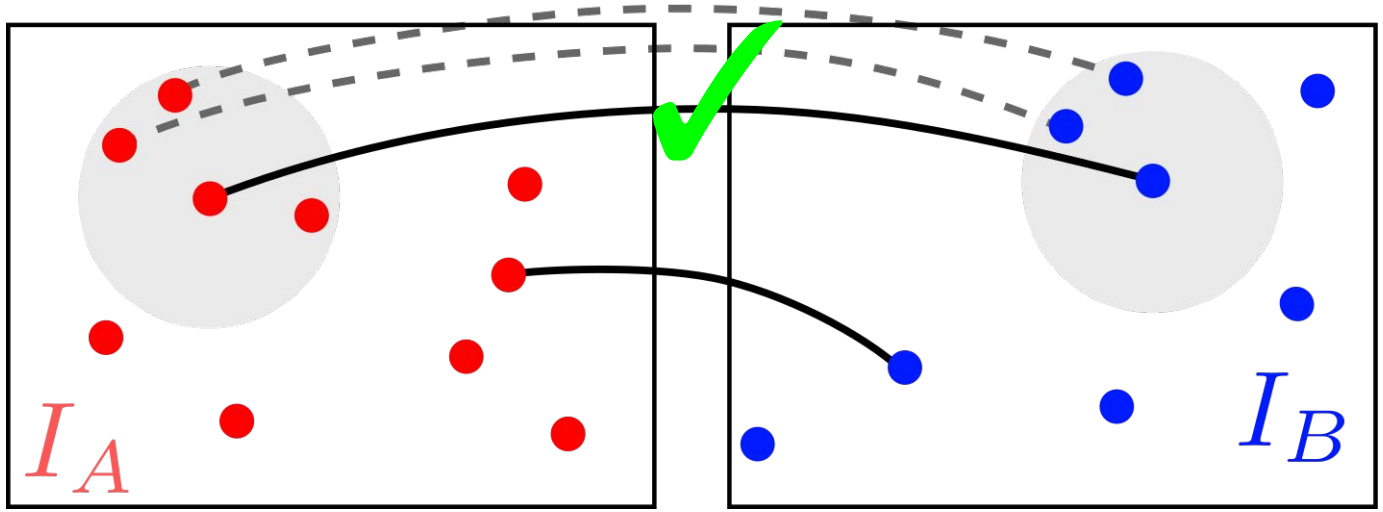
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



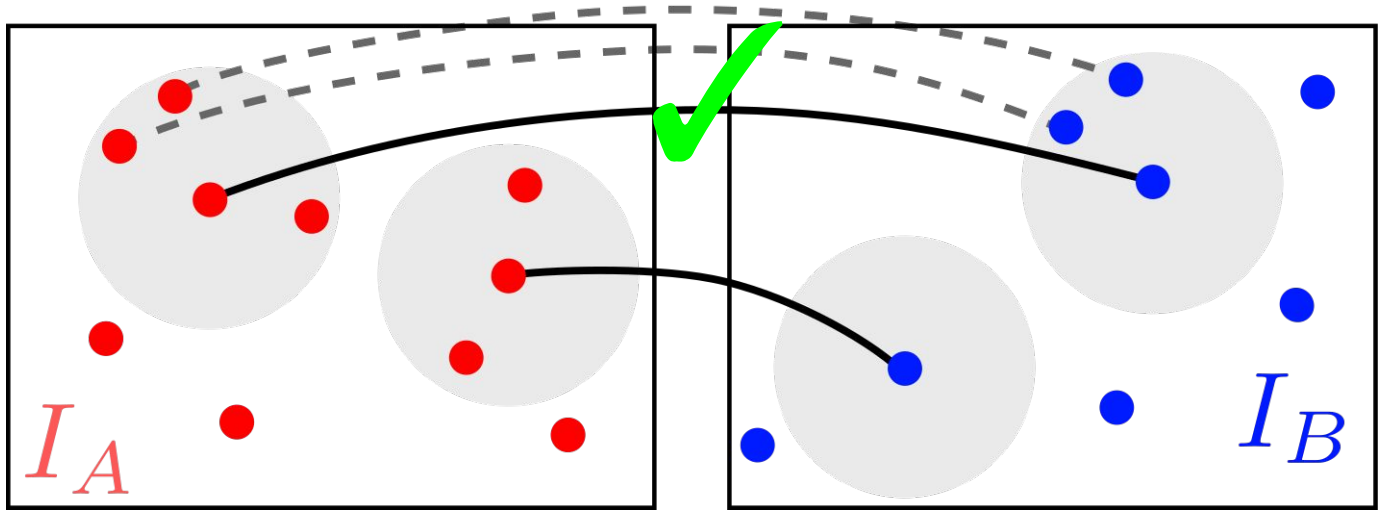
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



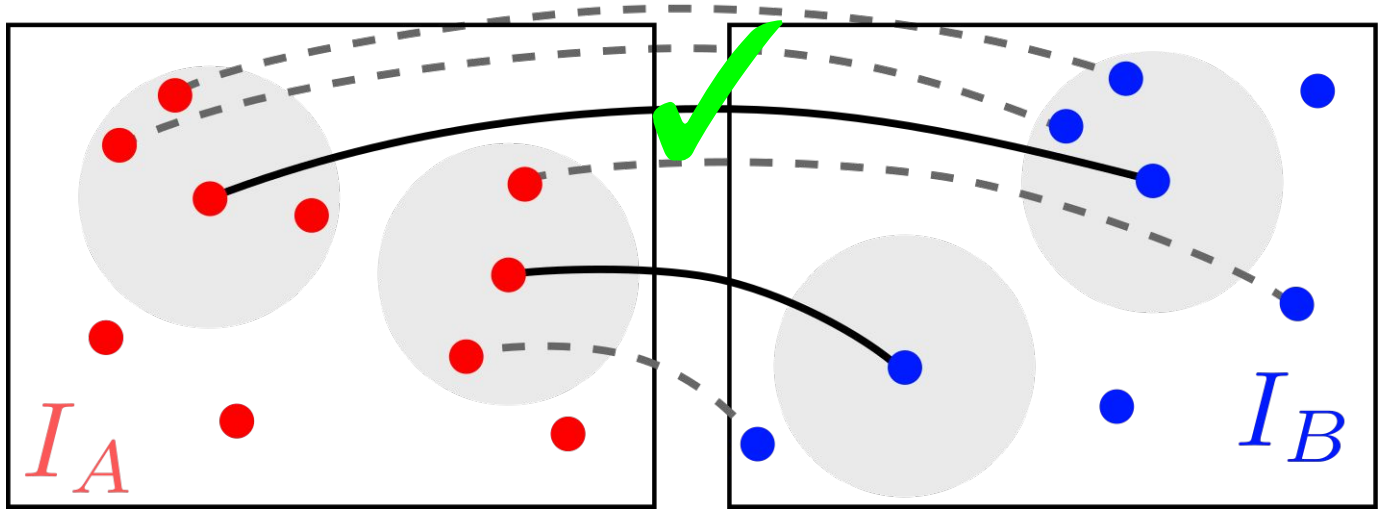
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



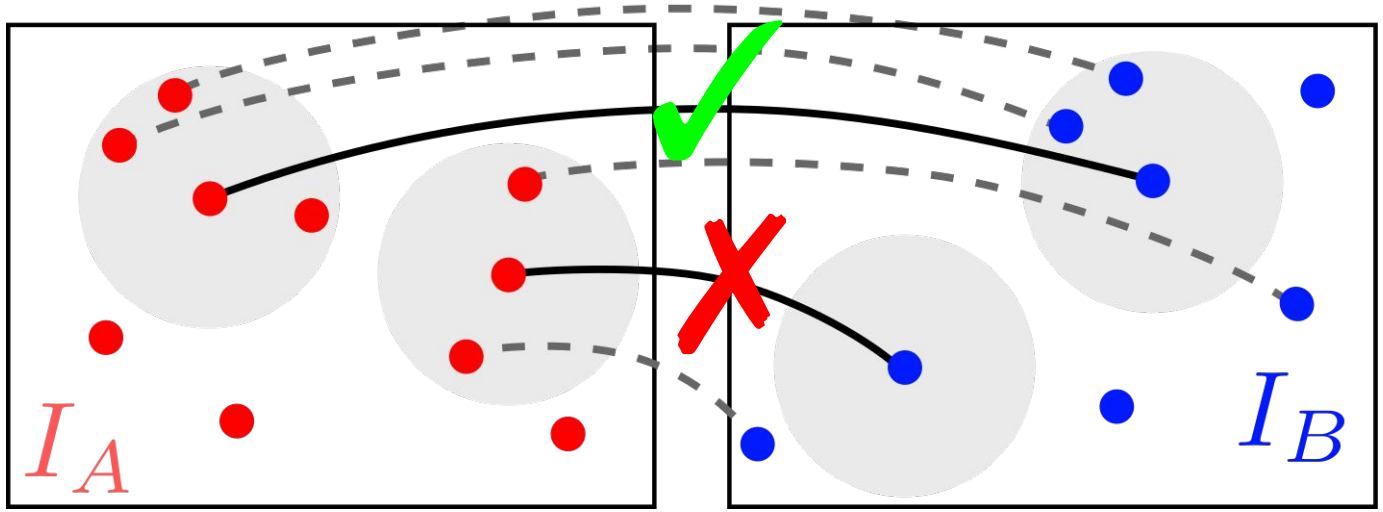
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Neighbourhood consensus



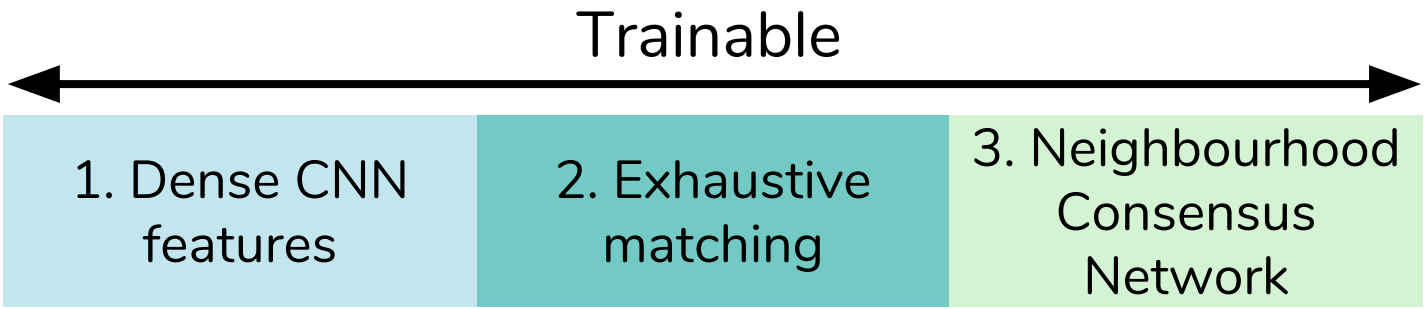
Schmid et al. PAMI 1997
Schaffalitzky et al. CIVR 2002
Sivic et al. ICCV 2003

Proposed method

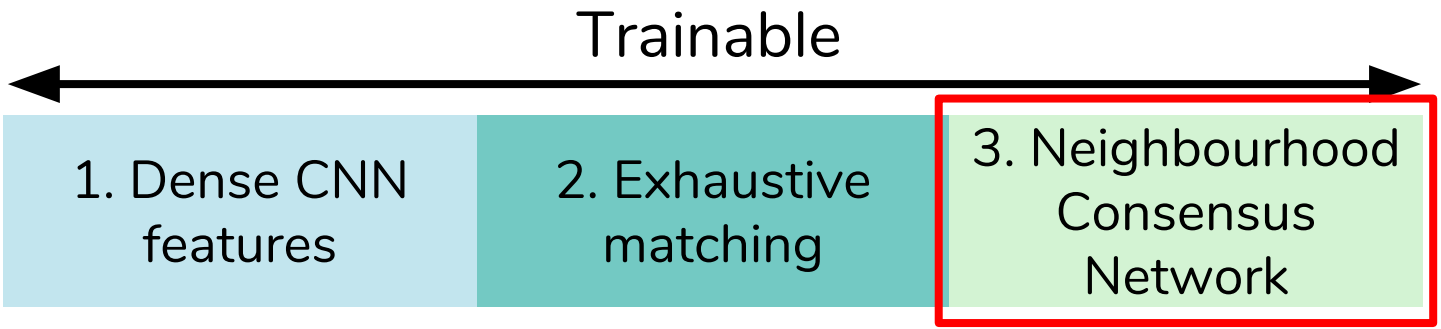
Proposed method

1. Dense CNN features	2. Exhaustive matching	3. Neighbourhood Consensus Network

Proposed method



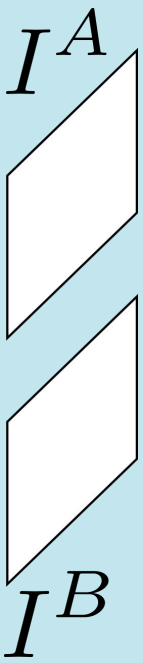
Proposed method



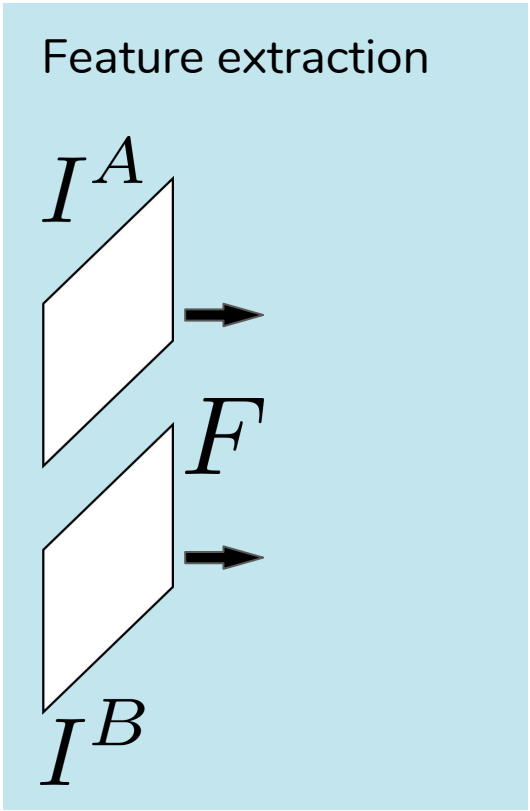
Proposed method

Proposed method

Feature extraction

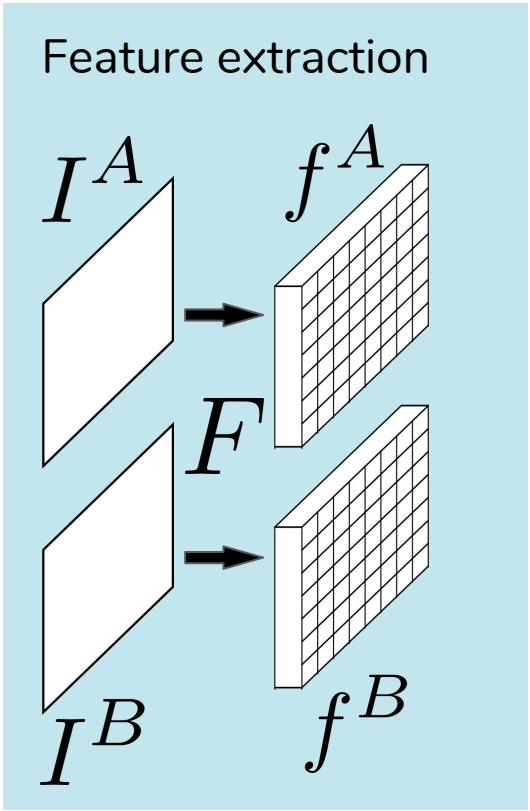


Proposed method



F: feature extraction CNN

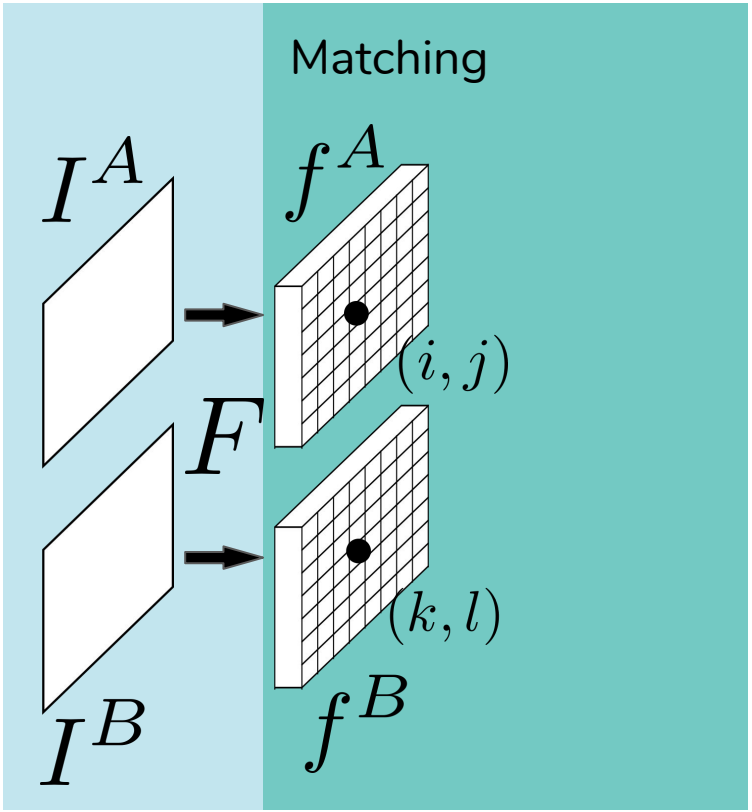
Proposed method



F: feature extraction CNN

f^A, f^B: dense feature maps

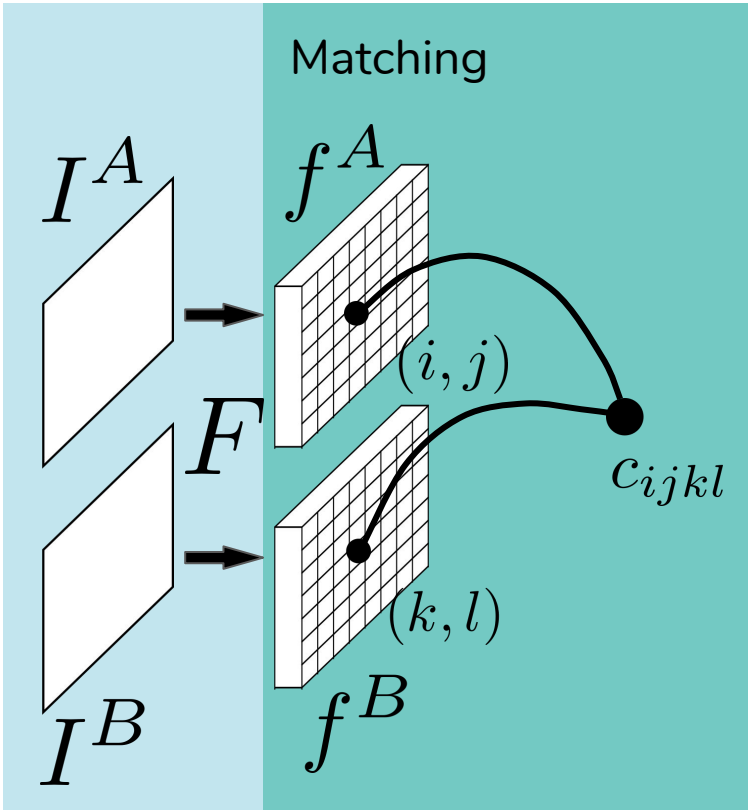
Proposed method



F: feature extraction CNN

f^A, f^B: dense feature maps

Proposed method



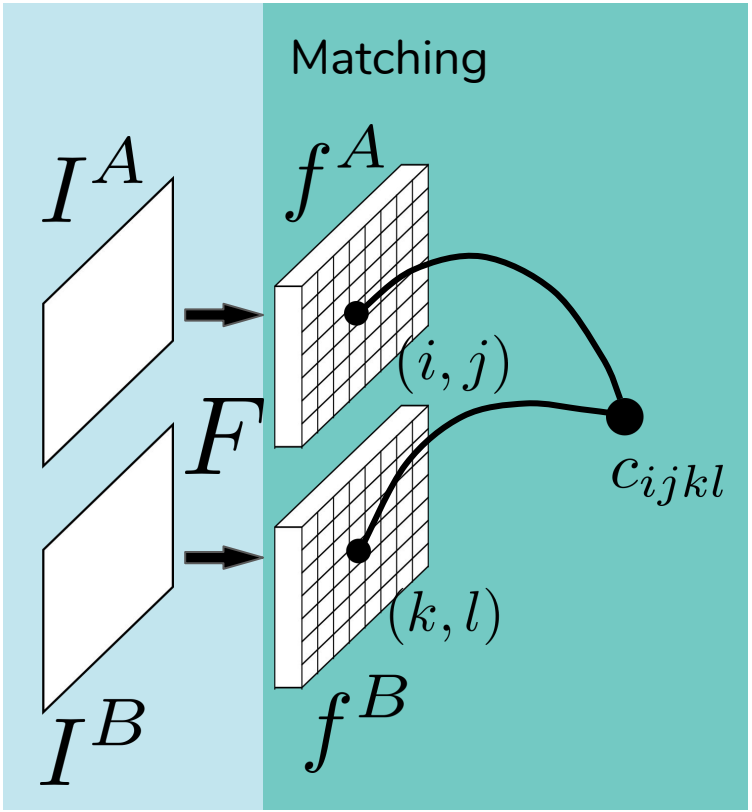
F: feature extraction CNN

f^A, f^B: dense feature maps

c_{ijkl}: 4D correlation tensor

$$c_{ijkl} = \frac{\langle f^A_{ij}, f^B_{kl} \rangle}{\|f^A_{ij}\|_2 \|f^B_{kl}\|_2}$$

Proposed method



F: feature extraction CNN

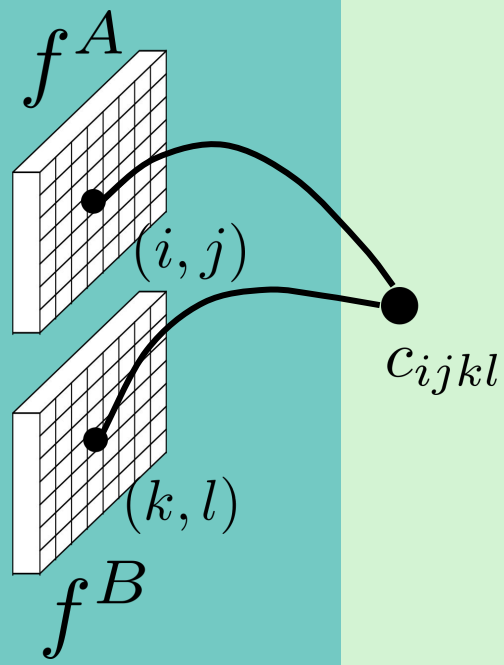
f^A, f^B: dense feature maps

c_{ijkl}: 4D correlation tensor

Number of matches: N^2
Correct matches: ~N

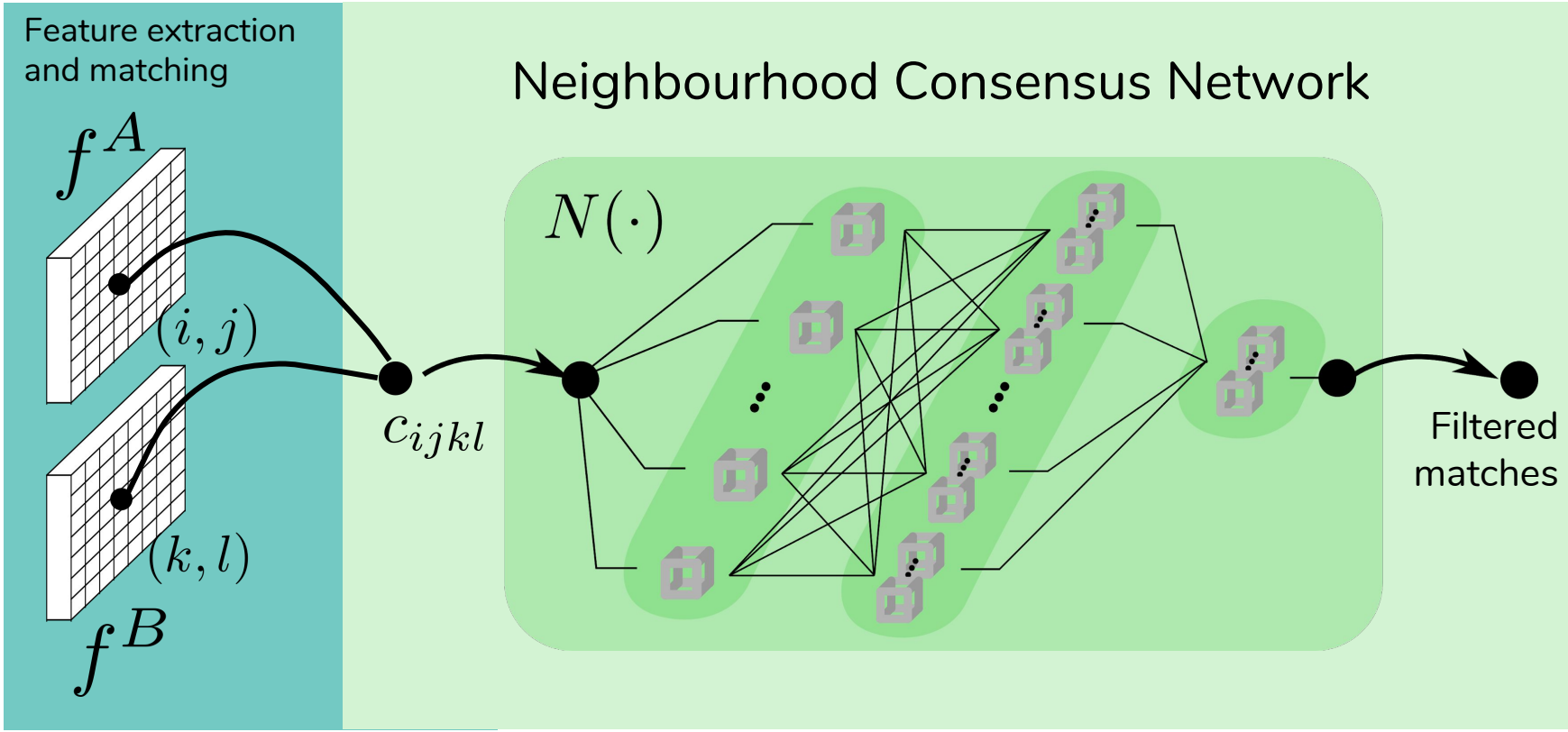
Proposed method

Feature extraction and matching

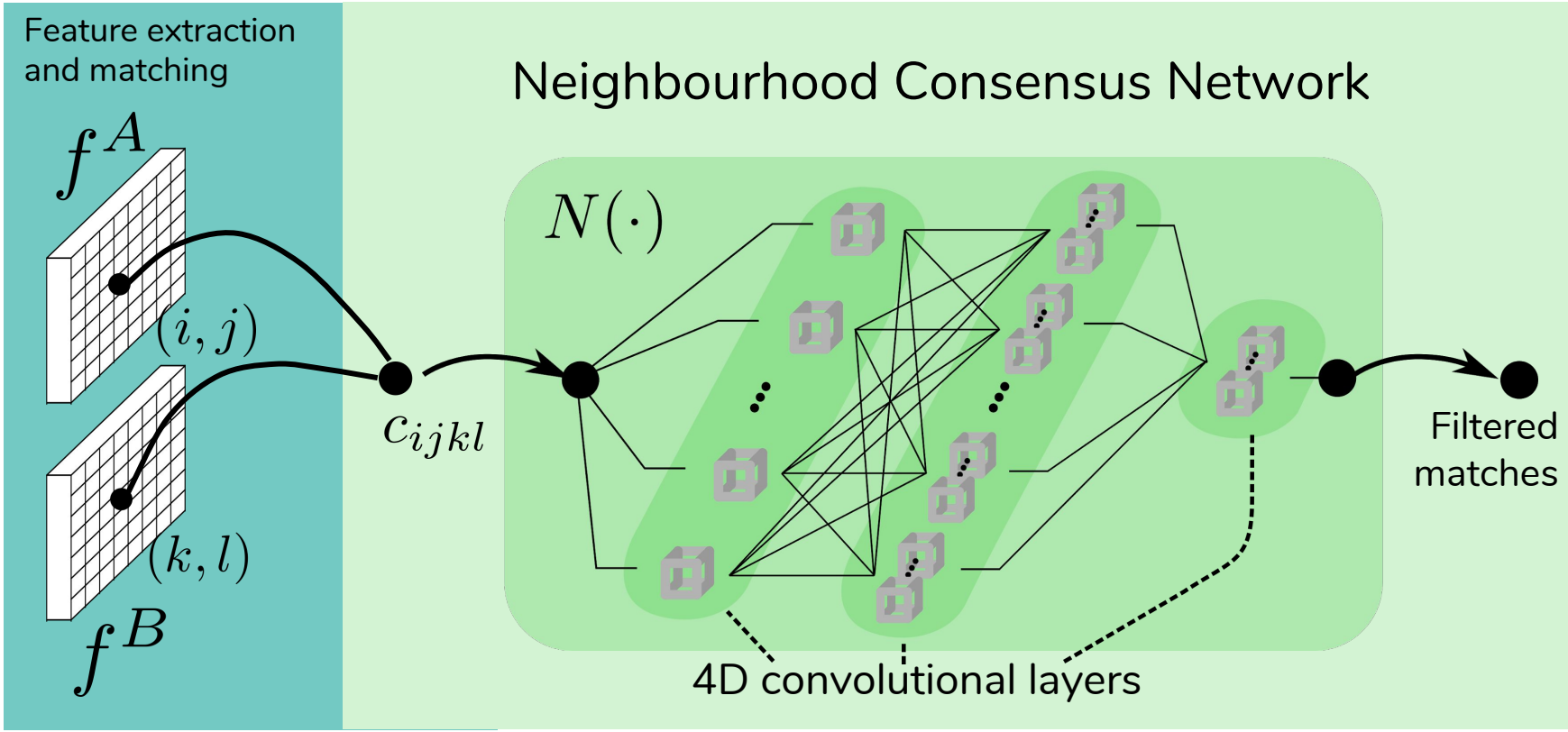


Neighbourhood Consensus Network

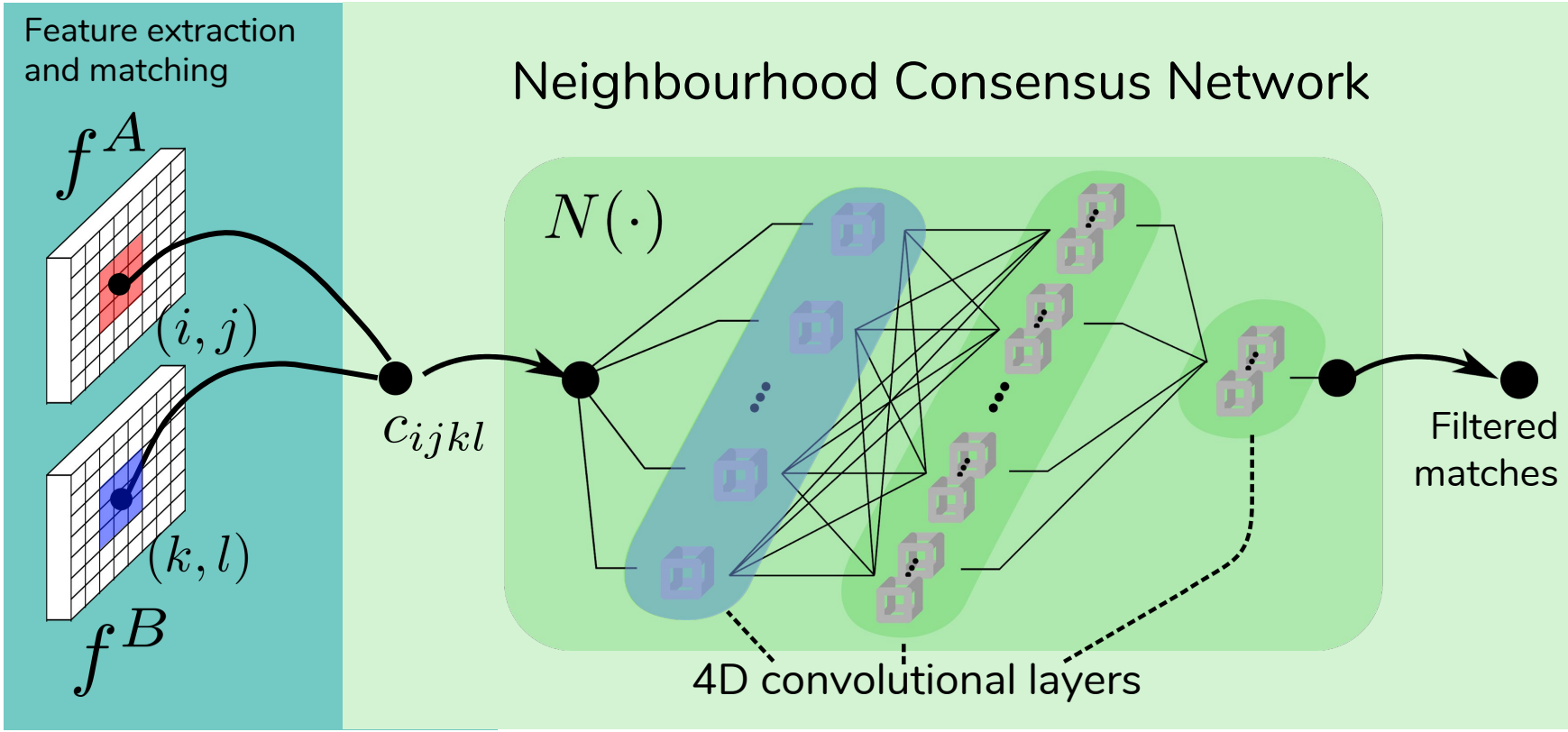
Proposed method



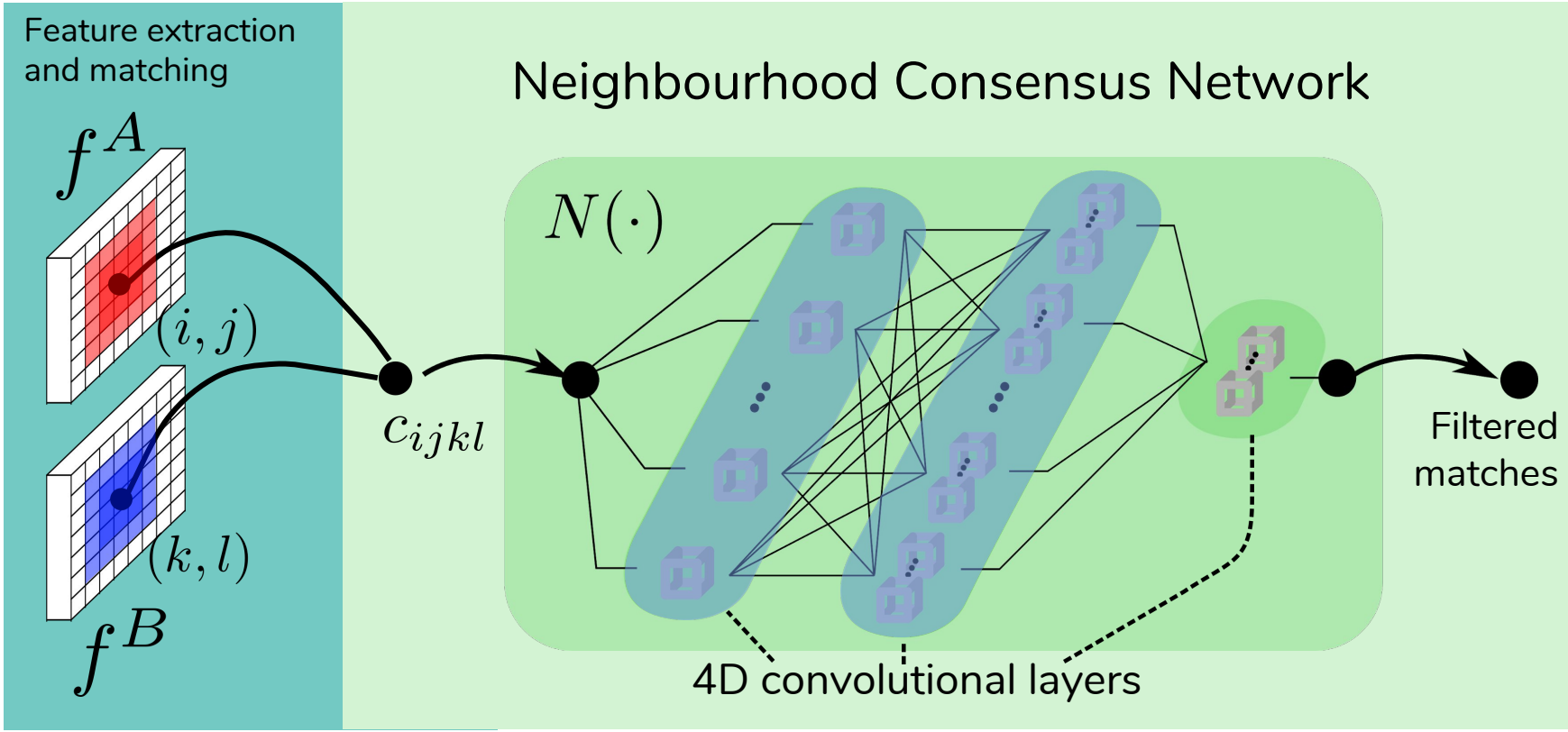
Proposed method



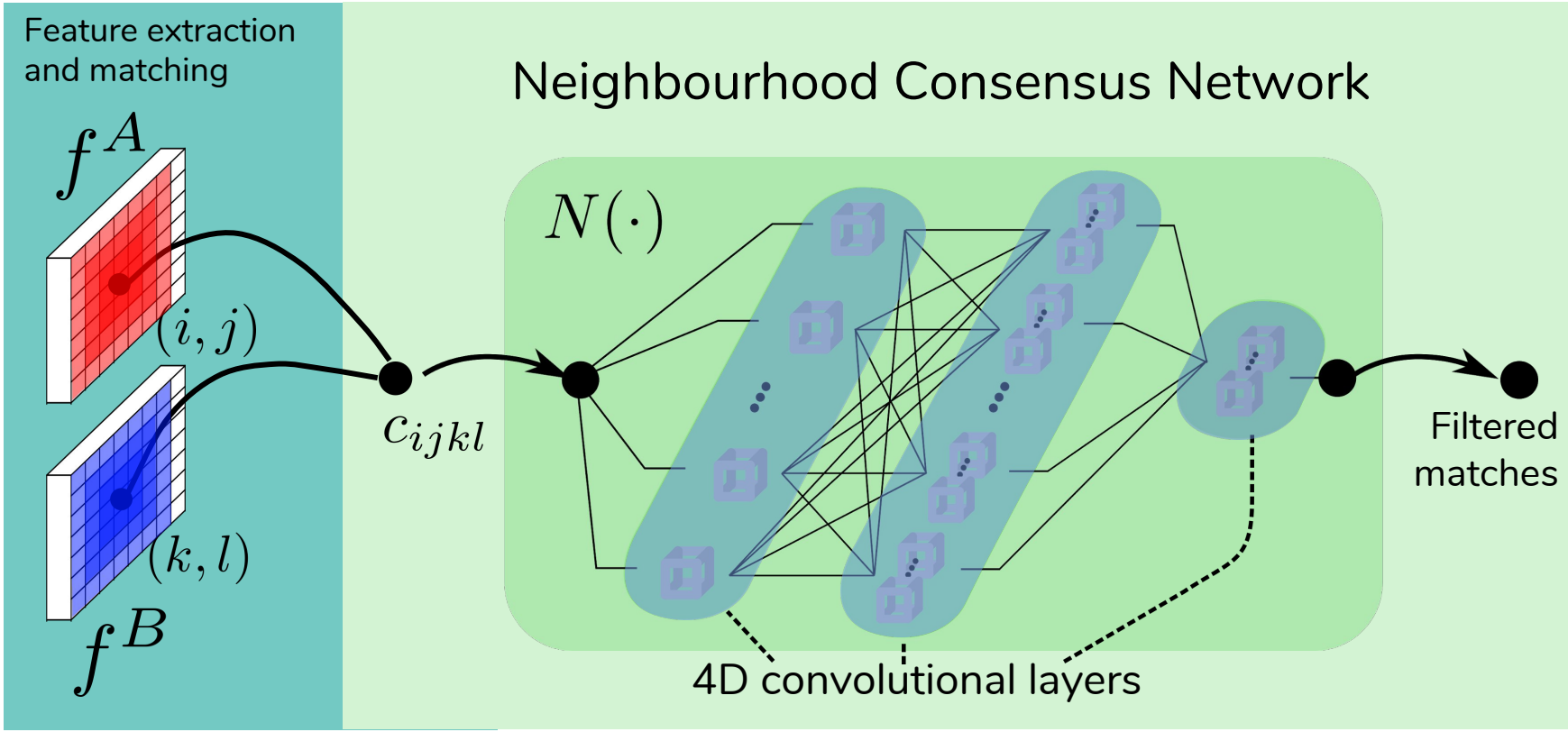
Proposed method



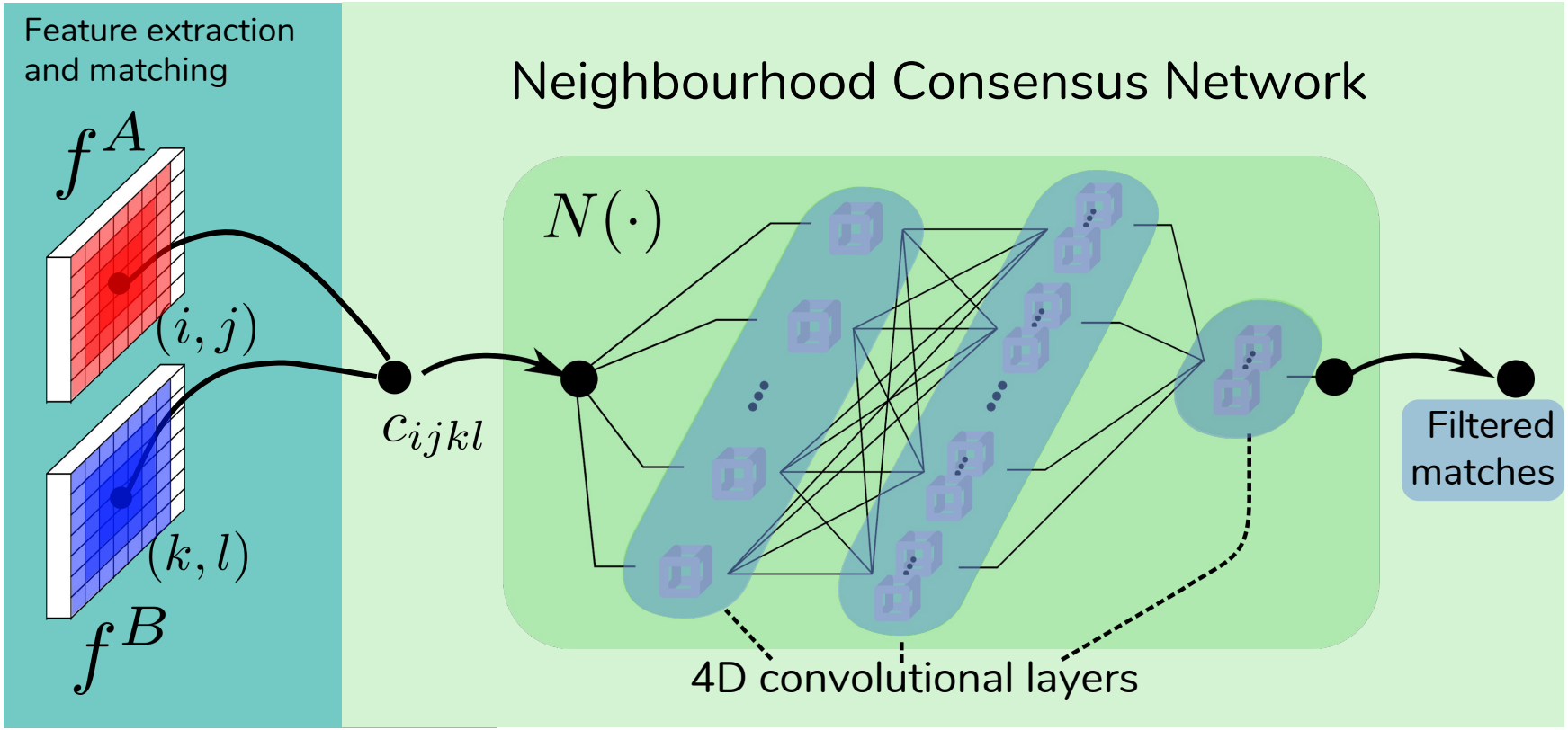
Proposed method



Proposed method



Proposed method



Weakly supervised training

Weakly supervised training

Only image-level supervision is required

Weakly supervised training

Only image-level supervision is required

Weakly supervised training

Only image-level supervision is required

Positive pair

Weakly supervised training

Only image-level supervision is required

Negative pair

✗

Weakly supervised training

Training objective:

Weakly supervised training

Training objective:



Positive pair

Weakly supervised training

Training objective:

↑ Maximize overall matching score

Weakly supervised training

Training objective:

Negative pair

↑ Maximize overall matching score

Weakly supervised training

Training objective:

↓ Minimize overall matching score

↑ Maximize overall matching score

Category-level matching (PF-Pascal)

Ham et al. PAMI 2017

Category-level matching (PF-Pascal)

- Task: *match similar semantic parts*

Ham et al. PAMI 2017

Category-level matching (PF-Pascal)

- Task: *match similar semantic parts*

Ham et al. PAMI 2017

Category-level matching (PF-Pascal)

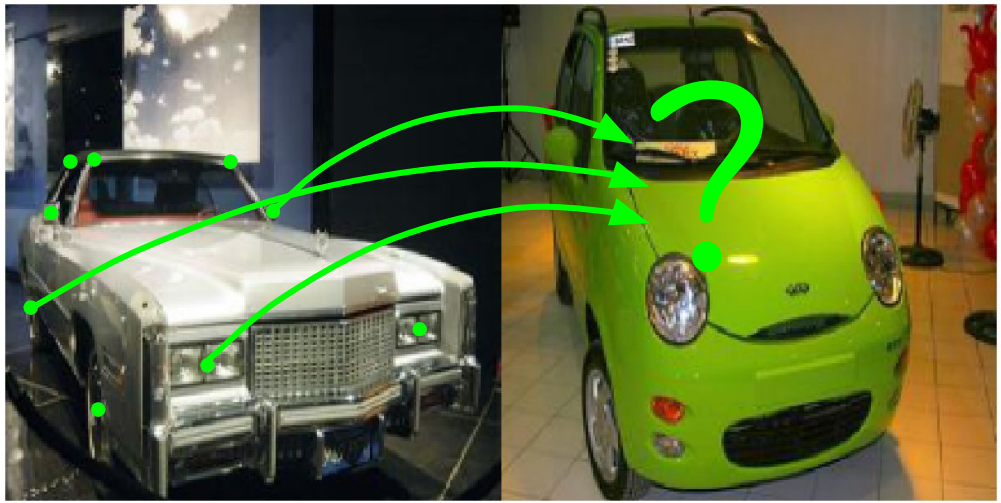
- Task: *match similar semantic parts*

annotated object parts

Ham et al. PAMI 2017

Category-level matching (PF-Pascal)

- Task: *match similar semantic parts*



annotated object parts

Ham et al. PAMI 2017

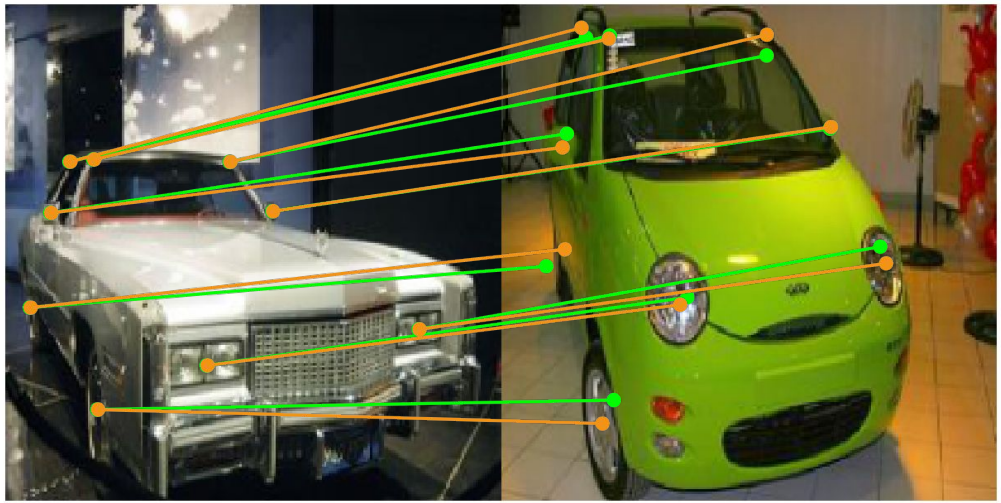
Category-level matching (PF-Pascal)

- Task: *match similar semantic parts*

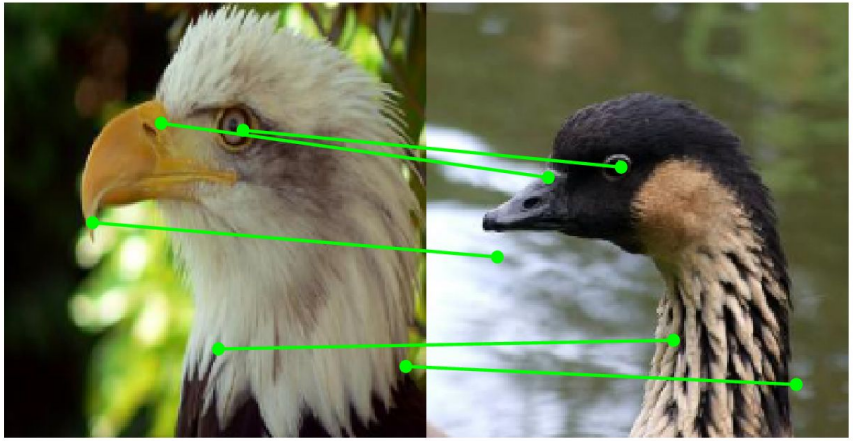
NCNet results

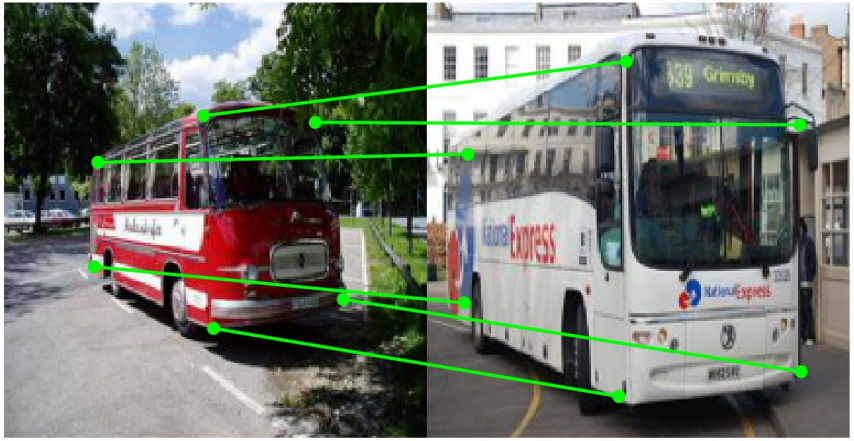
Category-level matching (PF-Pascal)

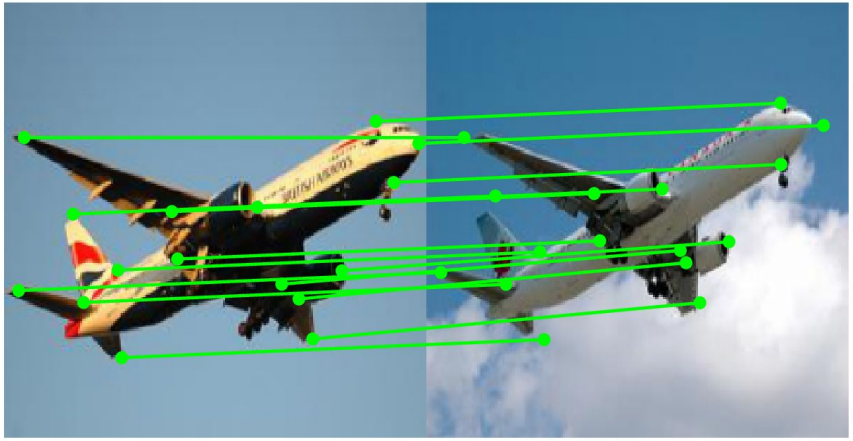
- Task: *match similar semantic parts*

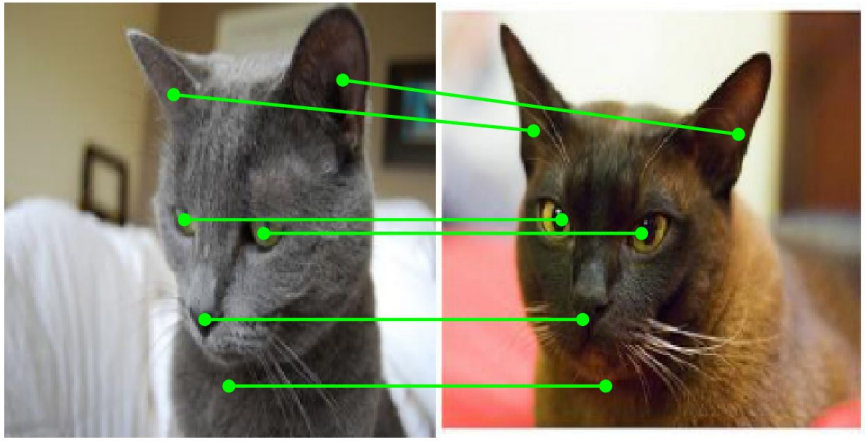


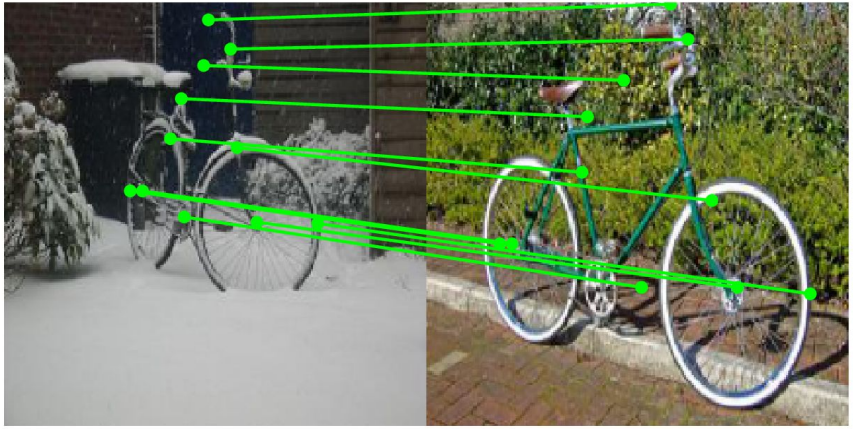
NCNet results

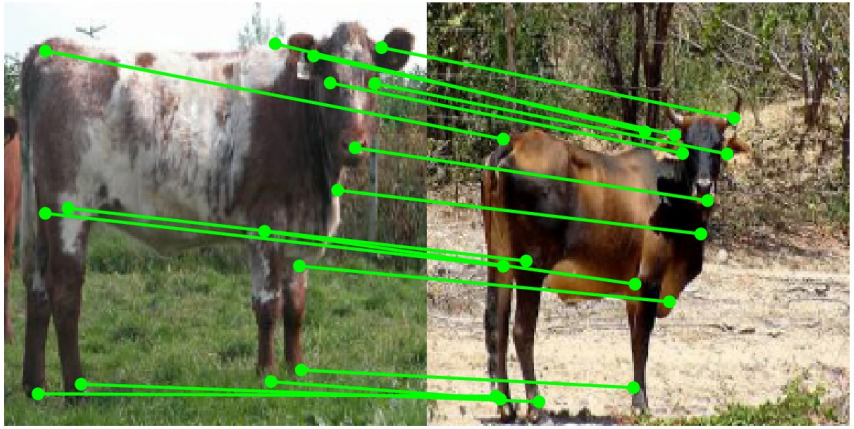


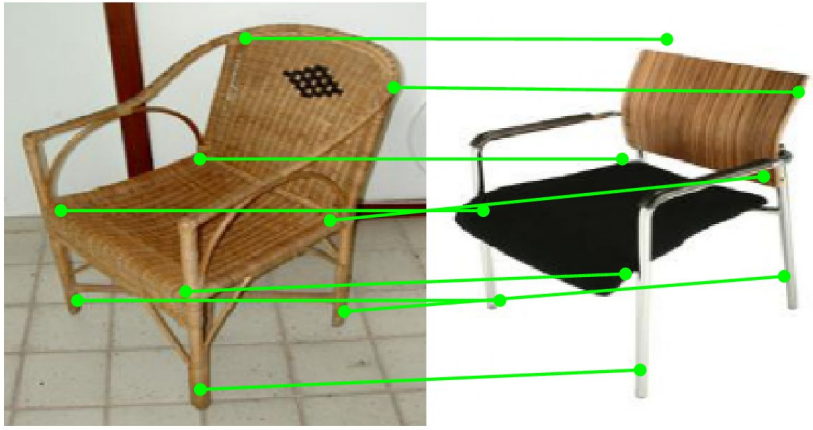

339 Grimsby
National Express

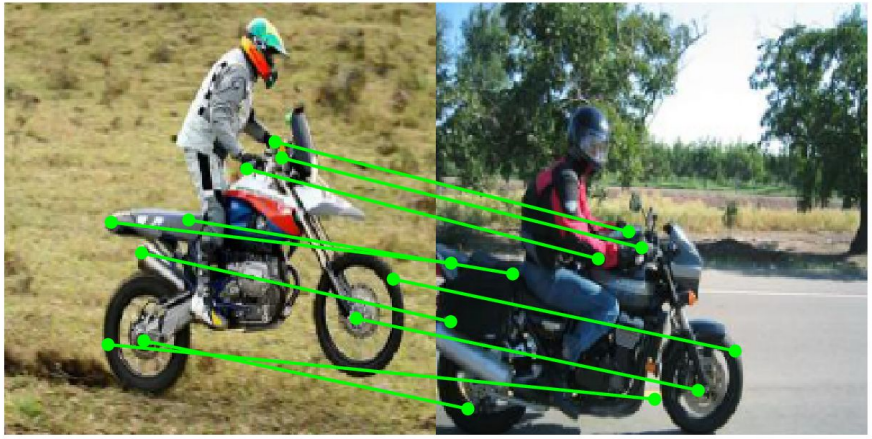












Instance-level matching (InLoc)

Taira et al. CVPR 2018

Instance-level matching (InLoc)

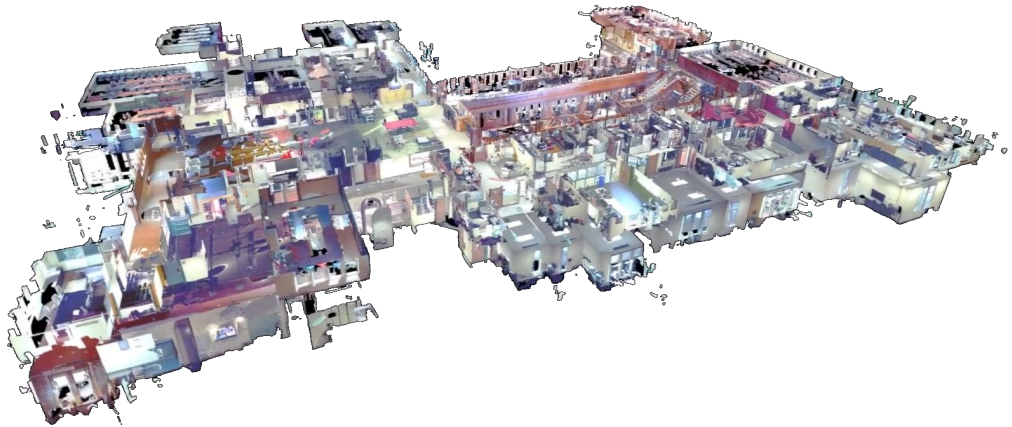
- Task: *recover query camera pose*

Taira et al. CVPR 2018

Instance-level matching (InLoc)

- Task: *recover query camera pose*

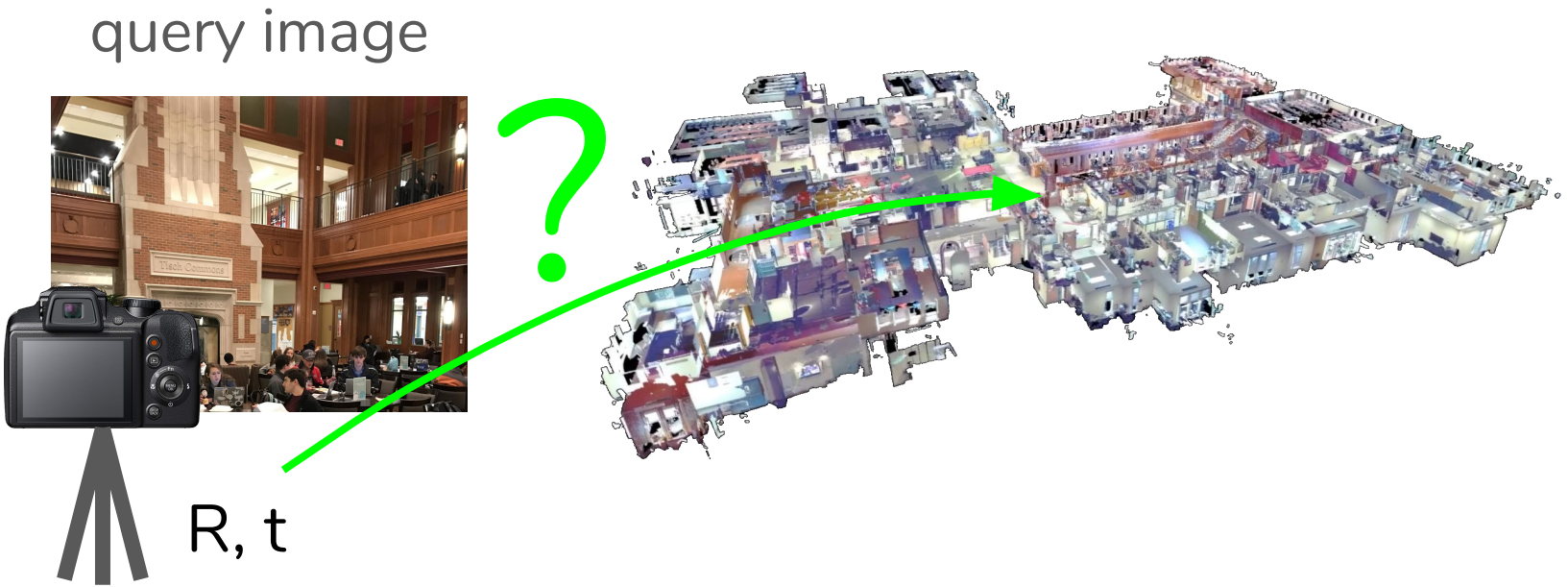
query image



Taira et al. CVPR 2018

Instance-level matching (InLoc)

- Task: *recover query camera pose*



Taira et al. CVPR 2018

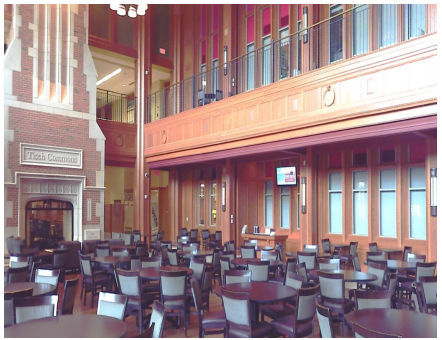
Instance-level matching (InLoc)

- Task: *recover query camera pose*

query image

image retrieval (NetVLAD)

RGBD database image



Taira et al. CVPR 2018
Arandjelović et al. CVPR 2016

Instance-level matching (InLoc)

- Task: *recover query camera pose*

query image

RGBD database image

pixel-wise correspondences

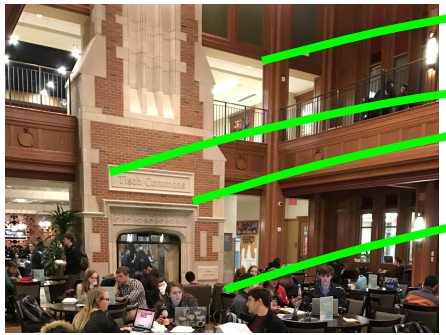
Taira et al. CVPR 2018
Arandjelović et al. CVPR 2016

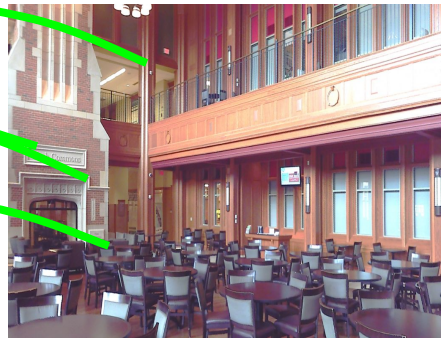
Instance-level matching (InLoc)

- Task: *recover query camera pose*

query image

RGBD database image





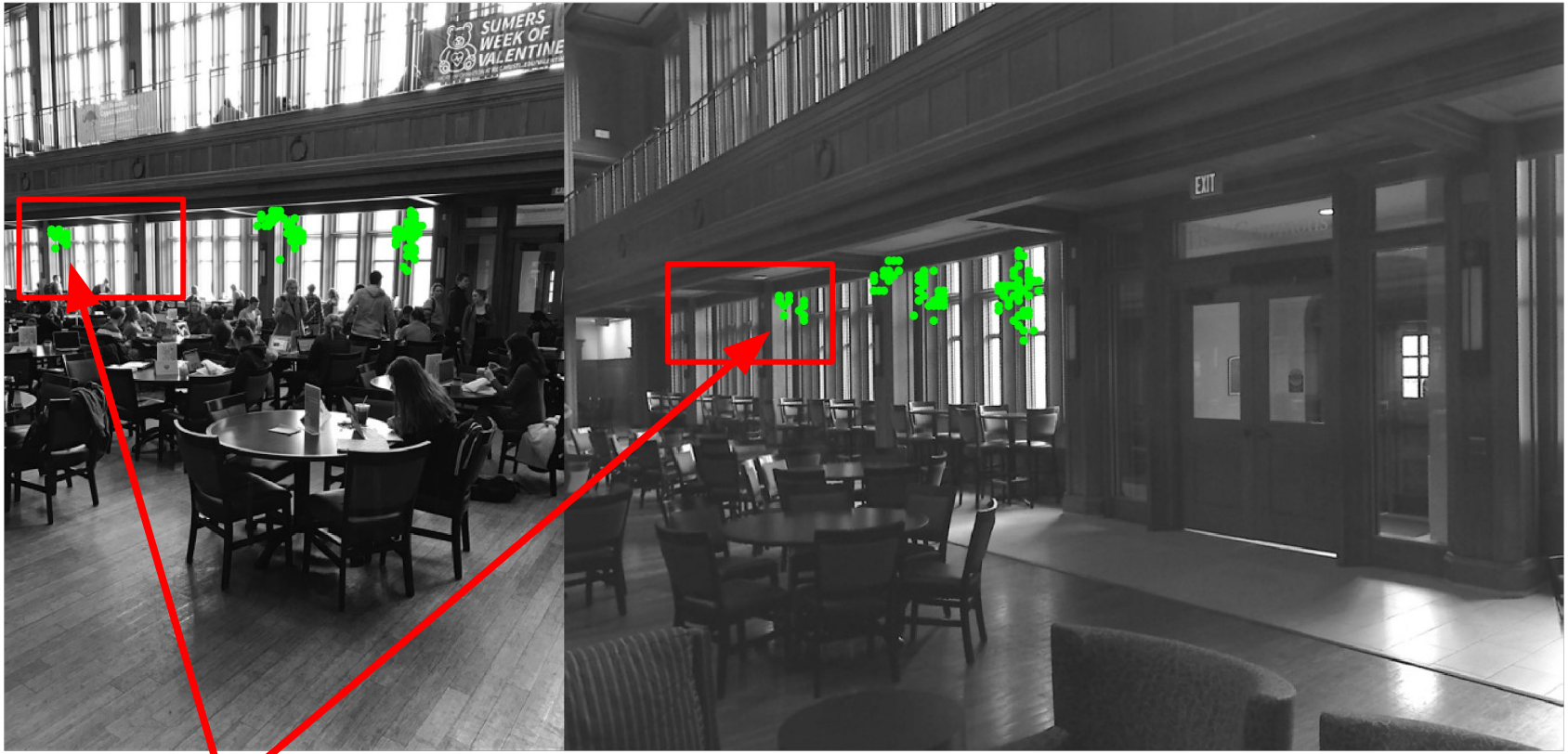
PnP

R, t

Taira et al. CVPR 2018
Arandjelović et al. CVPR 2016

baseline: InLoc

[Taira et al. CVPR'18]



mismatch

baseline: InLoc
[Taira et al. CVPR'18]

mismatch

baseline: InLoc
[Taira et al. CVPR'18]

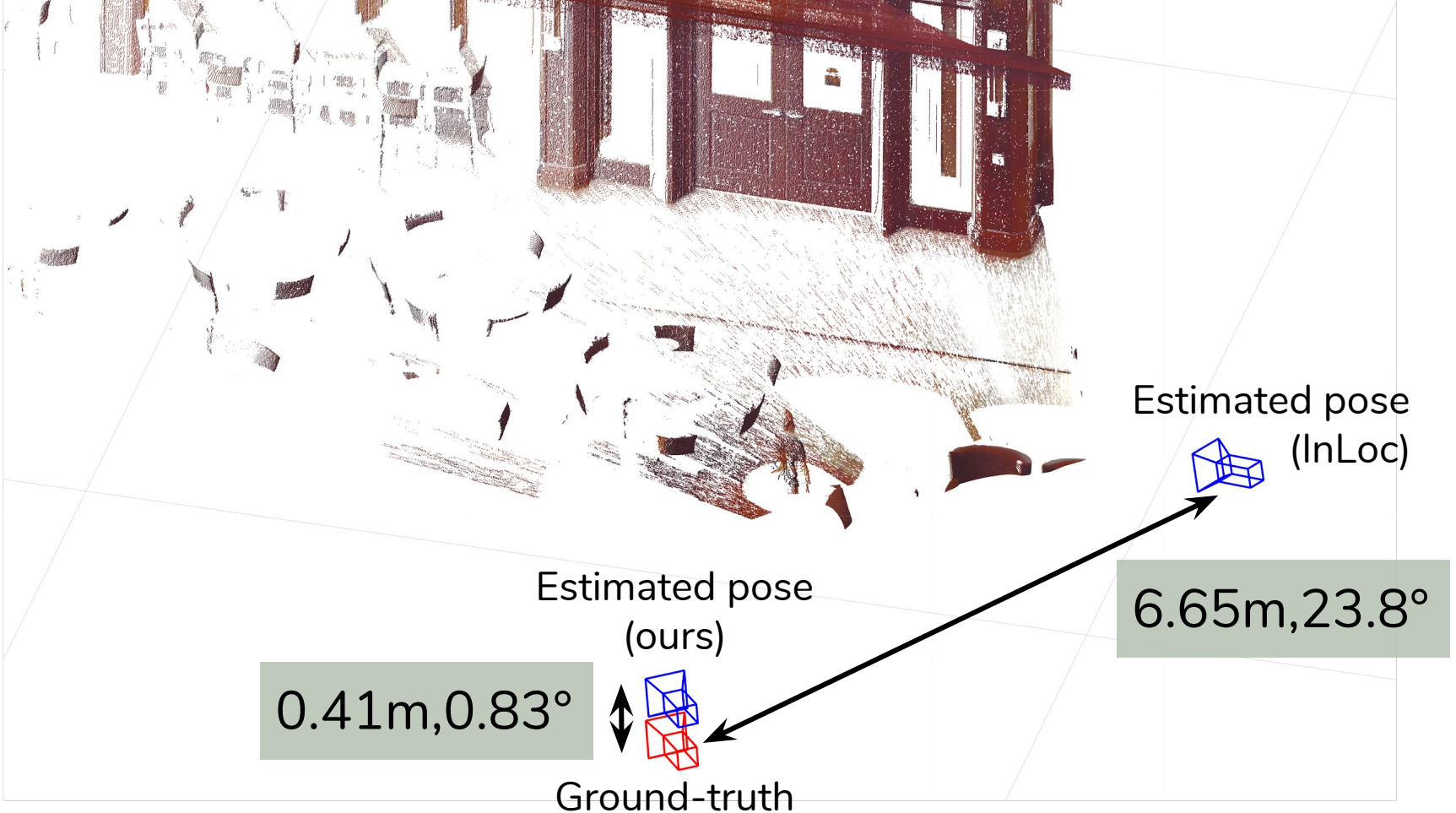
InLoc+NCNet

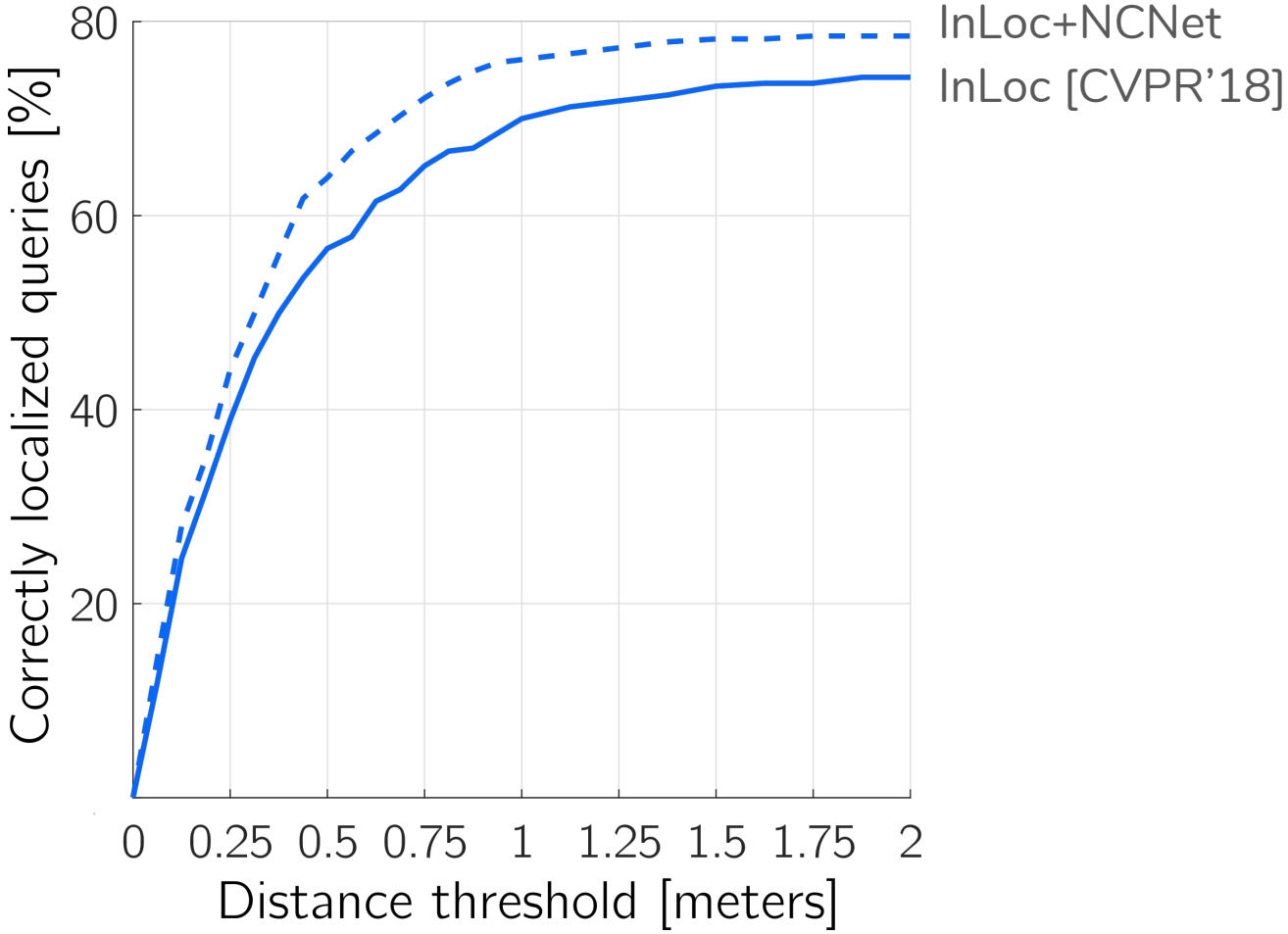
InLoc+NCNet

Estimated pose (InLoc)

Estimated pose (ours)

Ground-truth pose


Estimated pose
(InLoc)
6.65m,23.8°
Estimated pose
(ours)
0.41m,0.83°
Ground-truth
pose


InLoc+NCNet
InLoc [CVPR'18]
Correctly localized queries [%]
80
60
40
20
0
0.25
0.5
0.75
1
1.25
1.5
1.75
2
Distance threshold [meters]

Thank you!

Code and trained models at:
https://www.di.ens.fr/willow/research/ncnet/

Poster AB #118 @ Wed Session A