Final Project Presentation Guidelines

Chen Wang
CS231A
03/11/2022

Logistics

- Time:
  - 03/14/2022, 11:30am – 1:00pm
- Place:
  - 5 parallel Zoom sessions, attend the one corresponding to your assigned project mentor
- Format:
  - 6 minutes (5 minutes talk + 1 minute QA)
- Important info:
  - Each student should attend the whole session and watch other teams' presentations as well as presenting yourself.
  - If you have a team with multiple members, you should divide up the presentation amongst yourselves.
  - If you are unable to attend the live Zoom session due to a time zone issue or work commitments, or if you need to keep your project private, you can send us a recording of your presentation instead - please fill out this survey to indicate you need to do this.

Grading

We will be grading for completeness and clarity more so than the quality of your results. The rubric will be split into the following categories:

- 20% for problem statement, motivation, and background
- 30% for technical approach
- 30% for sufficient and informative quantitative and qualitative results
- 10% for visual style
- 10% for addressing questions raised during Q&A

Presentation Contents

- Tvsfpiq$Hijmrmxmsr$erh$Qsxmzexmsr
    - [lexmwxli$tvsfpiq$}syevixv}mrkxswspziC$Ls{$mwmxvipexihxsxli$gsyvwi$qexivmepC$[lex$mw$}syv$ksepC$[lex$mwxliglepprirkimrxlmw$tvsfpiqC
- Tvizmsyw$[svo$,Stxmsrep-$
    - Ls{hstvizmsyw$tistpi$wspzi$xlmw$tvsfpiqC$[lex$evi$xlimv$pmqmxexmsrwC
- Xiglrmgep$Hixempw
    - Lmklpmklx$}syv$qemr$xiglrmgep$gsrxvmfyxmsrw2$7$qmryxiwmwxss$pmxxpi$jsv$hixempih$qexl2
- I|tivmqirxw
    - I|tivmqirxep$wixyt2$Uyerxmxexmzi$viwypxw2$Uyepmxexmzi$viwypxw2$Sxliv$i|tigxih$viwypxw2
- Gsrgpywmsr$,Stxmsrep-$

Tips for the Presentation

1. Make a Storyline.

2. Highlight Your Contributions and Efforts.

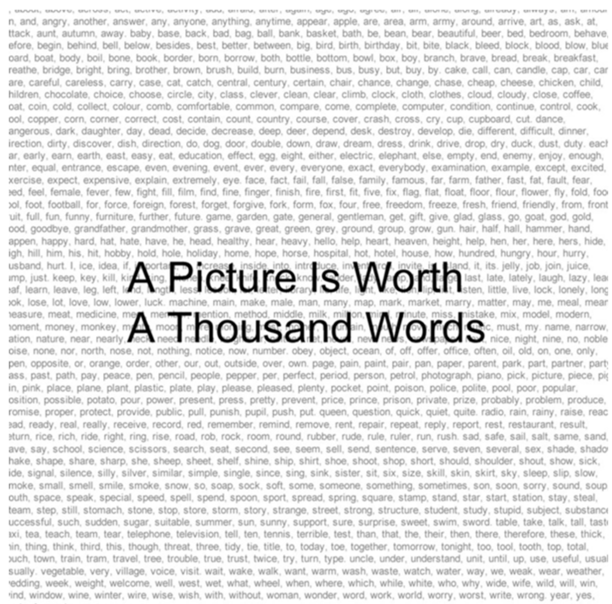
Your presentation is an advertisement of your project.

People will read your report later for details.

3. A picture is worth a thousand words.

After all, we are doing a computer vision course…

Animated figure is even better.



4. Less is More.

If you are not talk about a figure/text, remove it from your slides.

Only make 3~6 slides.

5. Practice.

Rehearse in front of your partners/friends.

Measure your time.

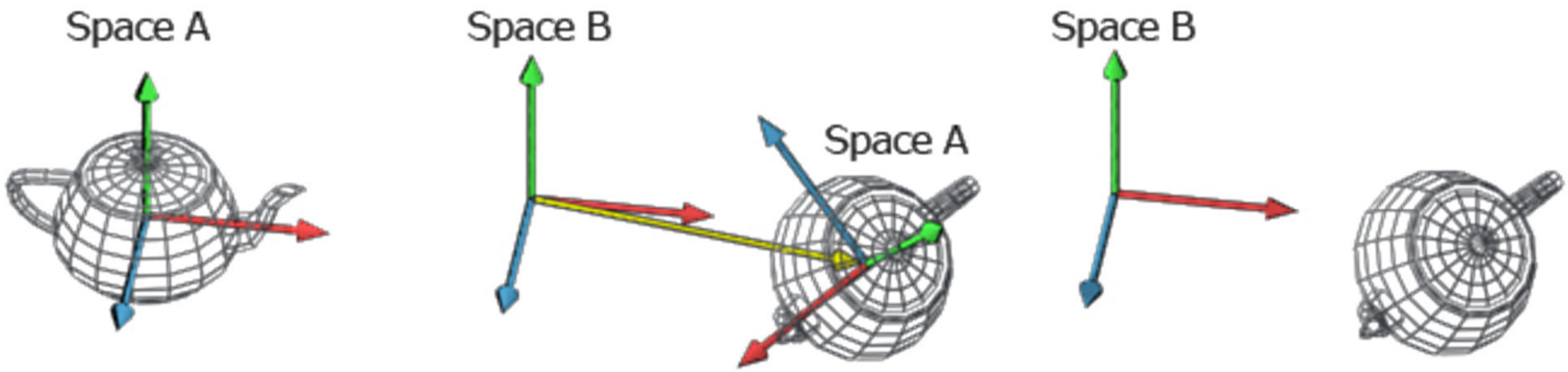
Record your voice.

DenseFusion: 6D Object Pose Estimation by Iterative Dense Fusion

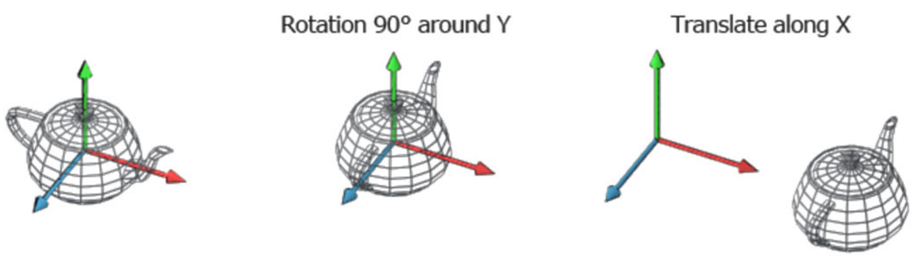
Chen Wang, Danfei Xu, Yuke Zhu, Roberto Martín-Martín
Cewu Lu, Li Fei-Fei, Silvio Savarese

Stanford
ARTIFICIAL
INTELLIGENCE

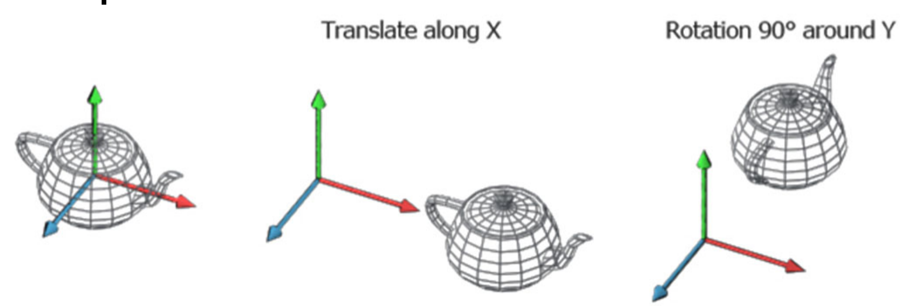
Representation of 6-DoF object pose: Rotation + Translation



example 1

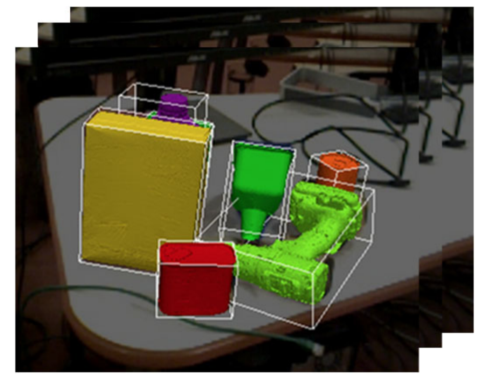


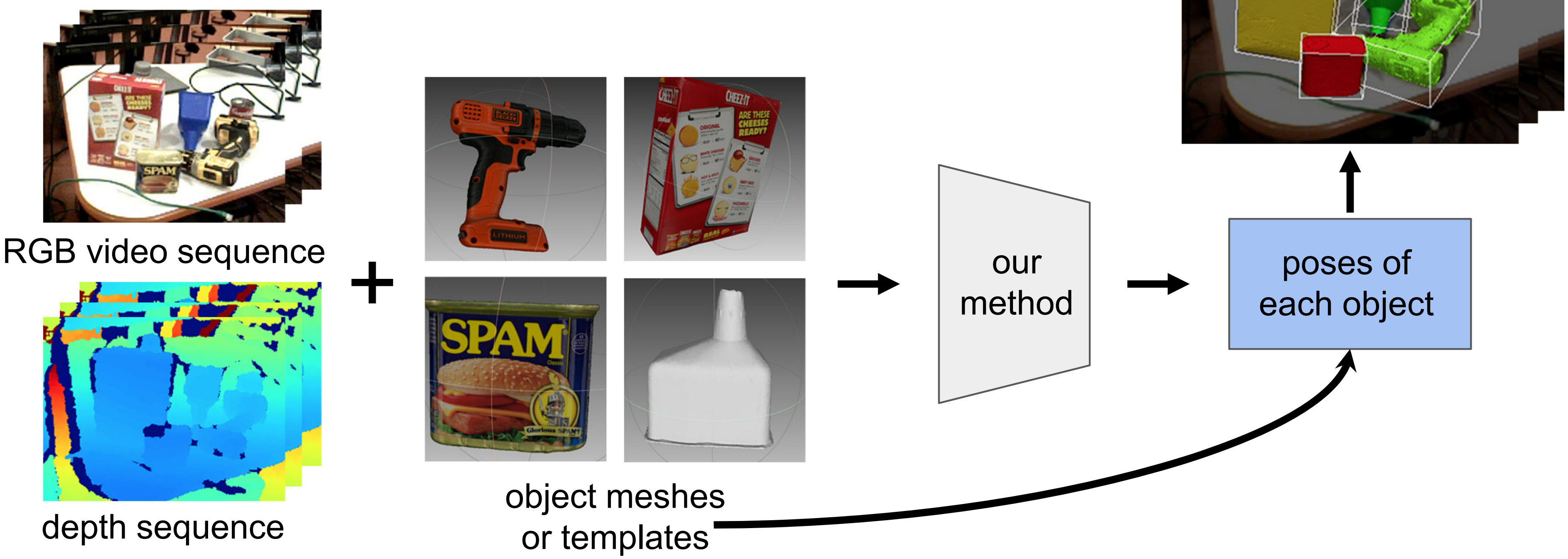
example 2



Problem Setting

- Goal: Uncover the Rotation and Translation of each object from a sequence of video frames
- Requirement:
  - depth image
  - the mesh models of the objects or their templates

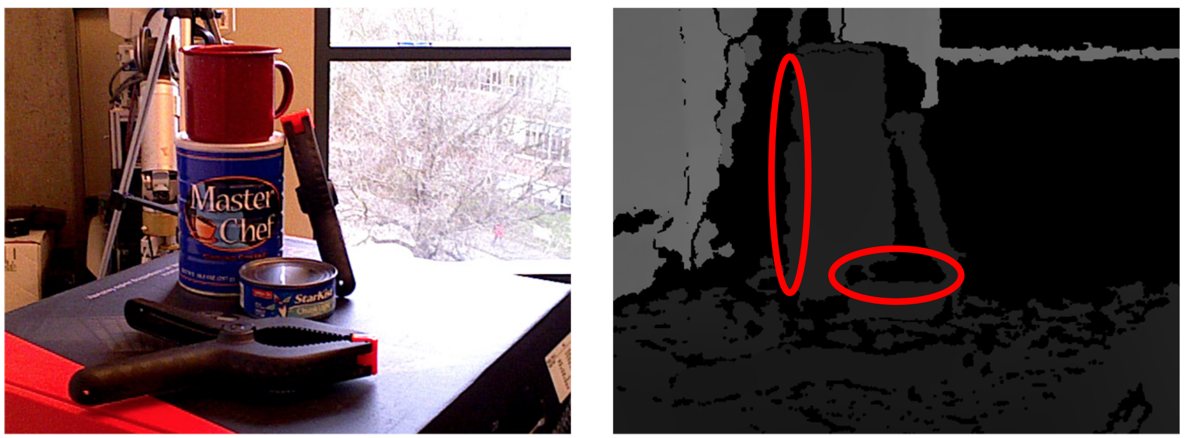




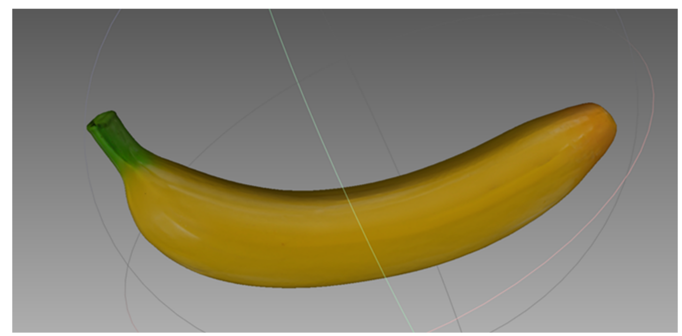
Challenges

- occlusion
- illumination changes of RGB
- noise of the depth image
- textureless object

occlusion

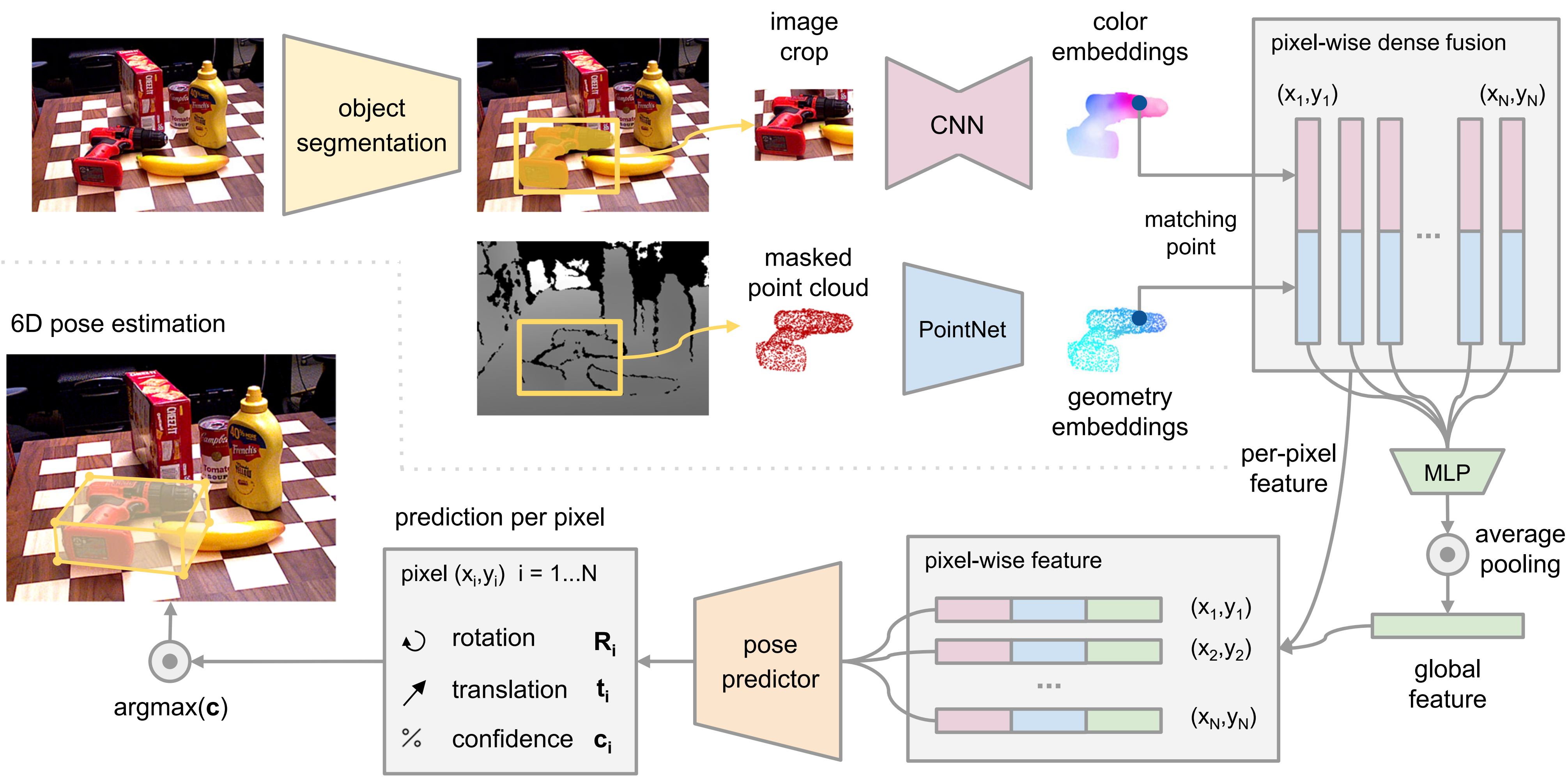


depth noise

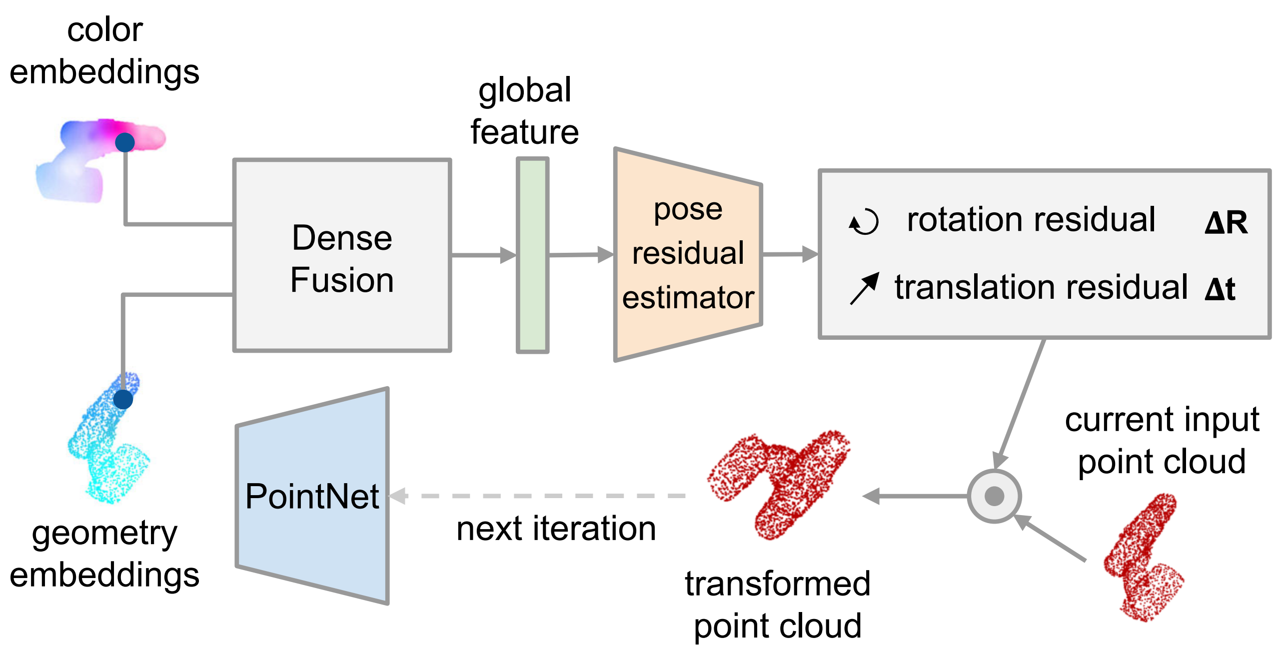


textureless

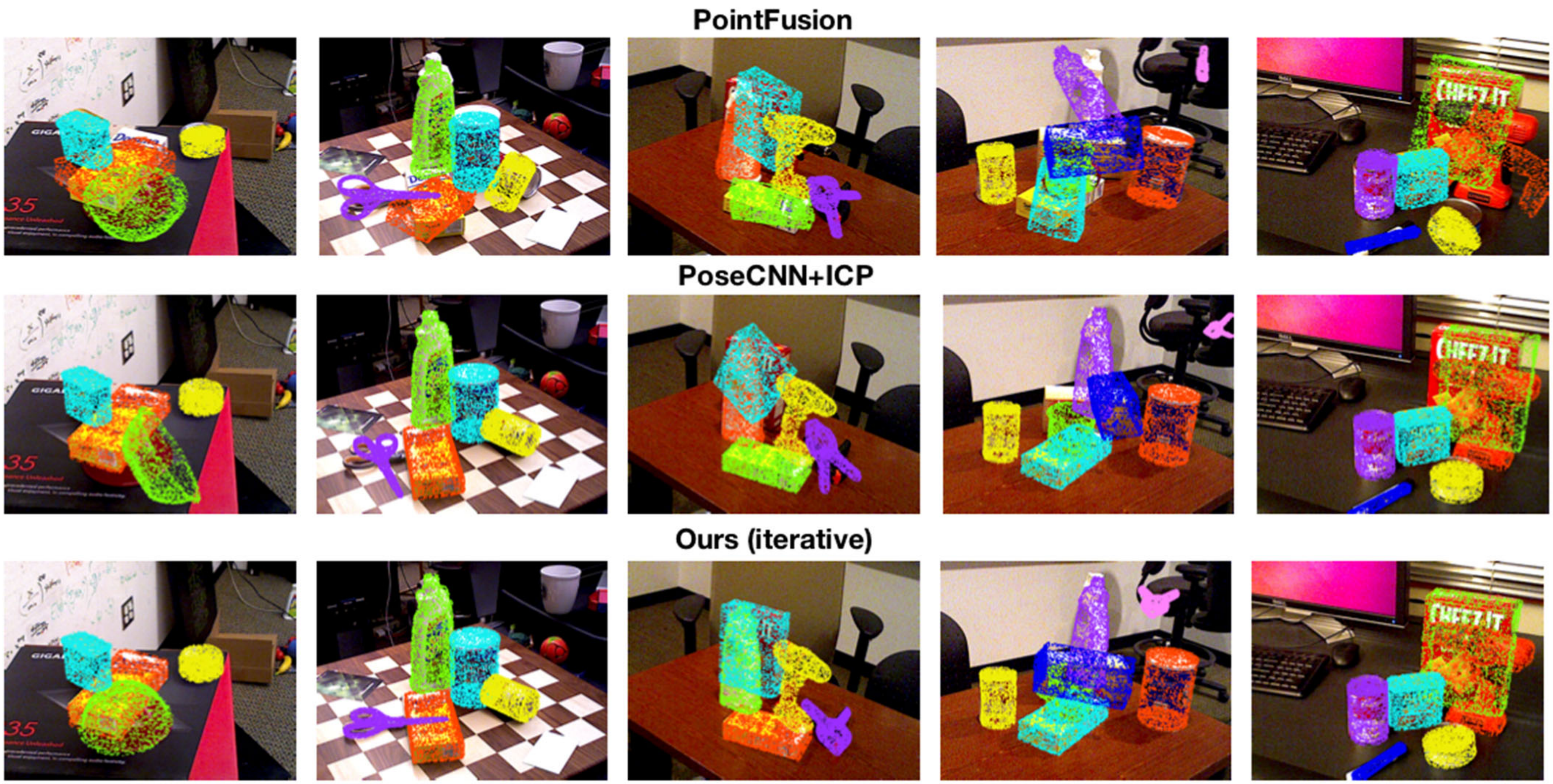
1. End-to-end training of the pose estimator



2. Iterative pose refinement



Compare to the baselines: (Qualitative results)

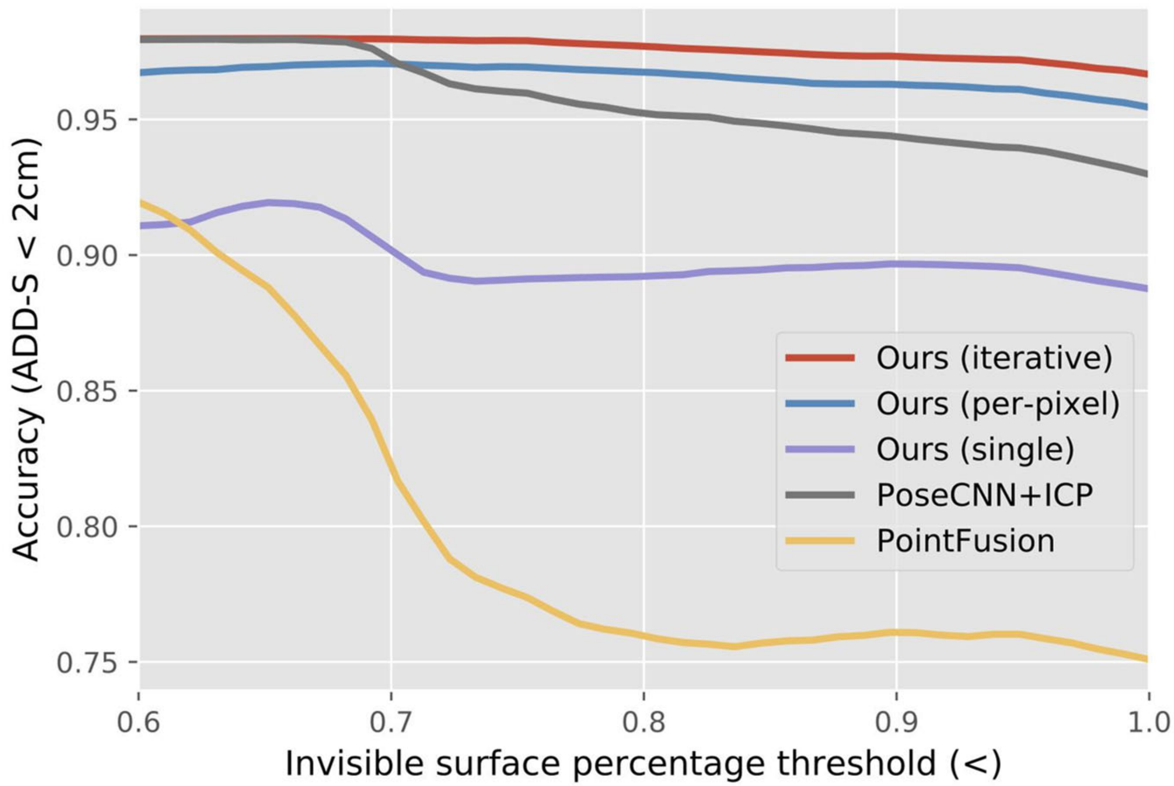


Compare to the baselines: (Quantitative results)

	PointFusion		PoseCNN+ICP		Ours (single)		Ours (per-pixel)		Ours (iterative)	
	AUC	<2cm	AUC	<2cm	AUC	<2cm	AUC	<2cm	AUC	<2cm
002_master_chef_can	90.9	99.8	95.8	100.0	93.9	100.0	95.2	100.0	**96.4**	100.0
003_cracker_box	80.5	62.6	92.7	91.6	90.8	98.4	92.5	99.3	**95.5**	**99.5**
004_sugar_box	90.4	95.4	**98.2**	100.0	94.4	99.2	95.1	100.0	97.5	100.0
005_tomato_soup_can	91.9	96.9	94.5	96.9	92.9	96.7	93.7	96.9	**94.6**	96.9
006_mustard_bottle	88.5	84.0	**98.6**	100.0	91.2	97.8	95.9	100.0	97.2	100.0
007_tuna_fish_can	93.8	99.8	**97.1**	100.0	94.9	100.0	94.9	100.0	96.6	100.0
008_pudding_box	87.5	96.7	**97.9**	100.0	88.3	97.2	94.7	100.0	96.5	100.0
009_gelatin_box	95.0	100.0	**98.8**	100.0	95.4	100.0	95.8	100.0	98.1	100.0
010_potted_meat_can	86.4	88.5	**92.7**	**93.6**	87.3	91.4	90.1	93.1	91.3	93.1
011_banana	84.7	70.5	**97.1**	99.7	84.6	62.0	91.5	93.9	96.6	**100.0**
019_pitcher_base	85.5	79.8	**97.8**	100.0	86.9	80.9	94.6	100.0	97.1	100.0
021_bleach_cleanser	81.0	65.0	**96.9**	99.4	91.6	98.2	94.3	99.8	95.8	**100.0**
024_bowl	75.7	24.1	81.0	54.9	83.4	55.4	86.6	69.5	**88.2**	**98.8**
025_mug	94.2	99.8	95.0	99.8	90.3	94.7	95.5	**100.0**	**97.1**	**100.0**
035_power_drill	71.5	22.8	**98.2**	**99.6**	83.1	64.2	92.4	97.1	96.0	98.7
036_wood_block	68.1	18.2	87.6	80.2	81.7	76.0	85.5	93.4	**89.7**	**94.6**
037_scissors	76.7	35.9	91.7	95.6	83.6	75.1	96.4	**100.0**	**95.2**	**100.0**
040_large_marker	87.9	80.4	97.2	99.7	91.2	88.6	94.7	99.2	**97.5**	**100.0**
051_large_clamp	65.9	50.0	**75.2**	74.9	70.5	77.1	71.6	78.5	72.9	**79.2**
052_extra_large_clamp	60.4	20.1	64.4	48.8	66.4	50.2	69.0	69.5	**69.8**	**76.3**
061_foam_brick	91.8	100.0	**97.2**	100.0	92.1	100.0	92.4	100.0	92.5	100.0
MEAN	83.9	74.1	93.0	93.2	88.2	87.9	91.2	95.3	**93.1**	**96.8**

DenseFusion achieved the **state-of-the-art performance** without time consuming ICP refinement. Our inference speed if **200x faster** than previous STOA method.

Occlusion experiments:



DenseFusion is the **most robust method** against occlusion situations.

Conclusion

- A principled way to combine color and depth information from the RGB-D input.
- Augmenting the information of each 3D point with 2D information from an embedding space learned for the task and use this new color-depth space to estimate the 6D pose.
- Integrating an iterative refinement procedure within the neural network architecture, removing the dependency of previous methods of a post-processing ICP step.