Dynamic SfM

Detecting Scene ***Changes*** from Image Pairs

Tuanfeng Y. Wang, University College London
Pushmeet Kohli, Microsoft Research Cambridge
Niloy J. Mitra, University College London

Problem Statement

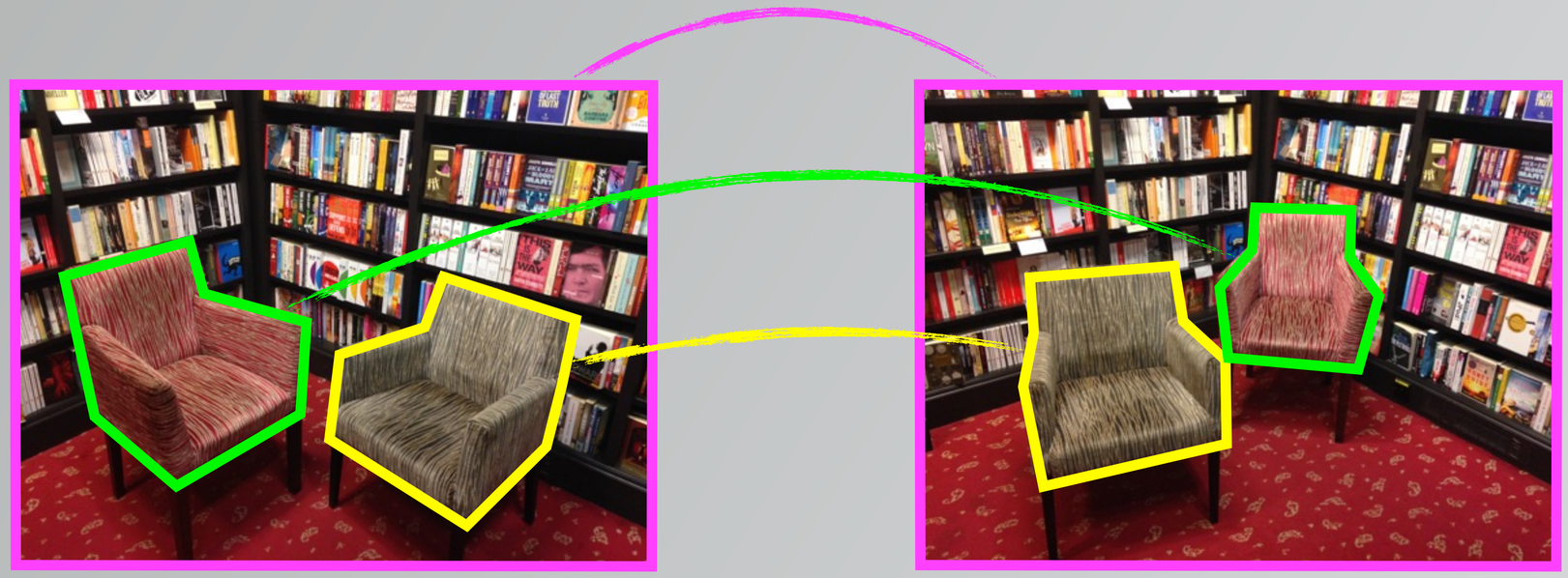
- Estimate Structure and Motion from a pair of uncalibrated images

Input： two RGB images with the scene changed

Problem Statement

- Estimate Structure and Motion from a pair of uncalibrated images

Input： two RGB images with the scene changed



Problem Statement

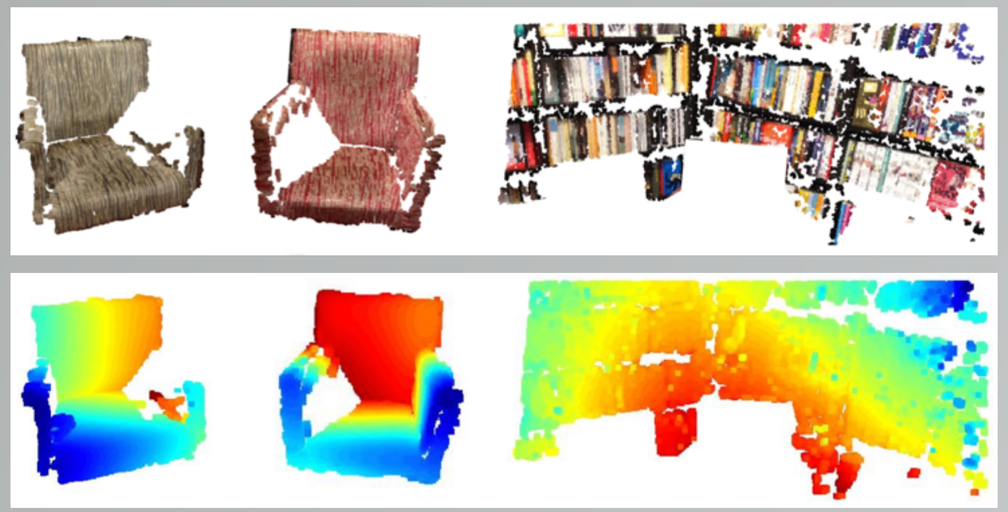
- Estimate Structure and Motion from a pair of uncalibrated images

Output： structure reconstruction & motions estimation & camera parameters calibration

Problem Statement

- Estimate Structure and Motion from a pair of uncalibrated images

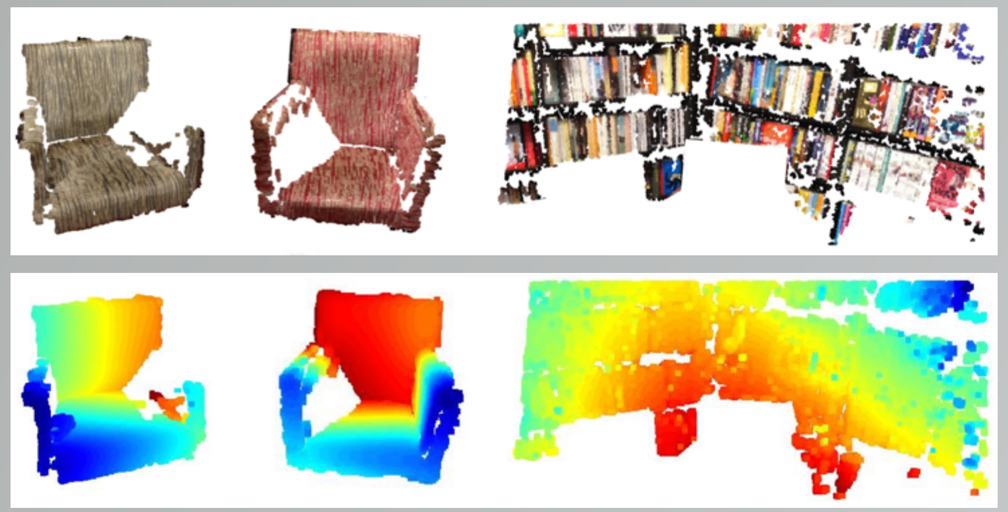
Output：structure reconstruction & motions estimation & camera parameters calibration



Problem Statement

- Estimate Structure and Motion from a pair of uncalibrated images

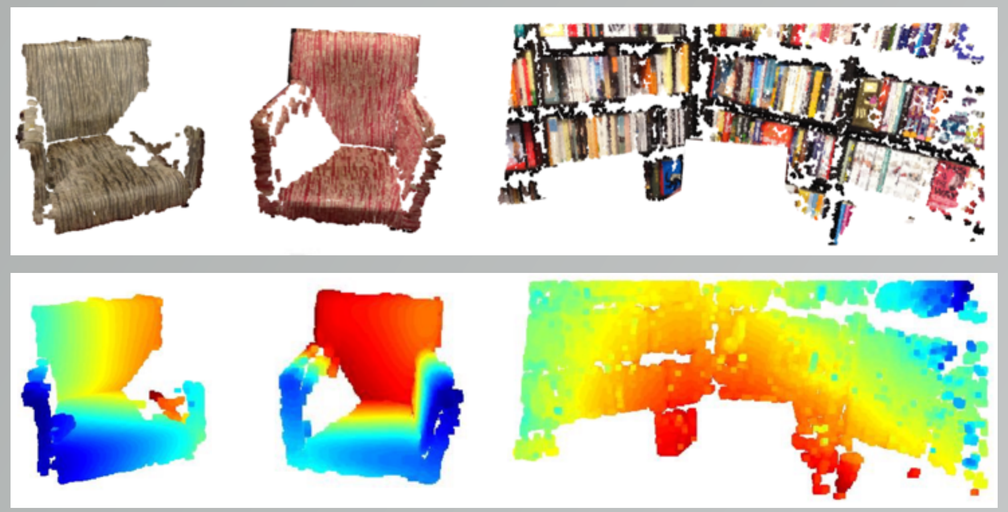
Output：structure reconstruction & motions estimation & camera parameters calibration



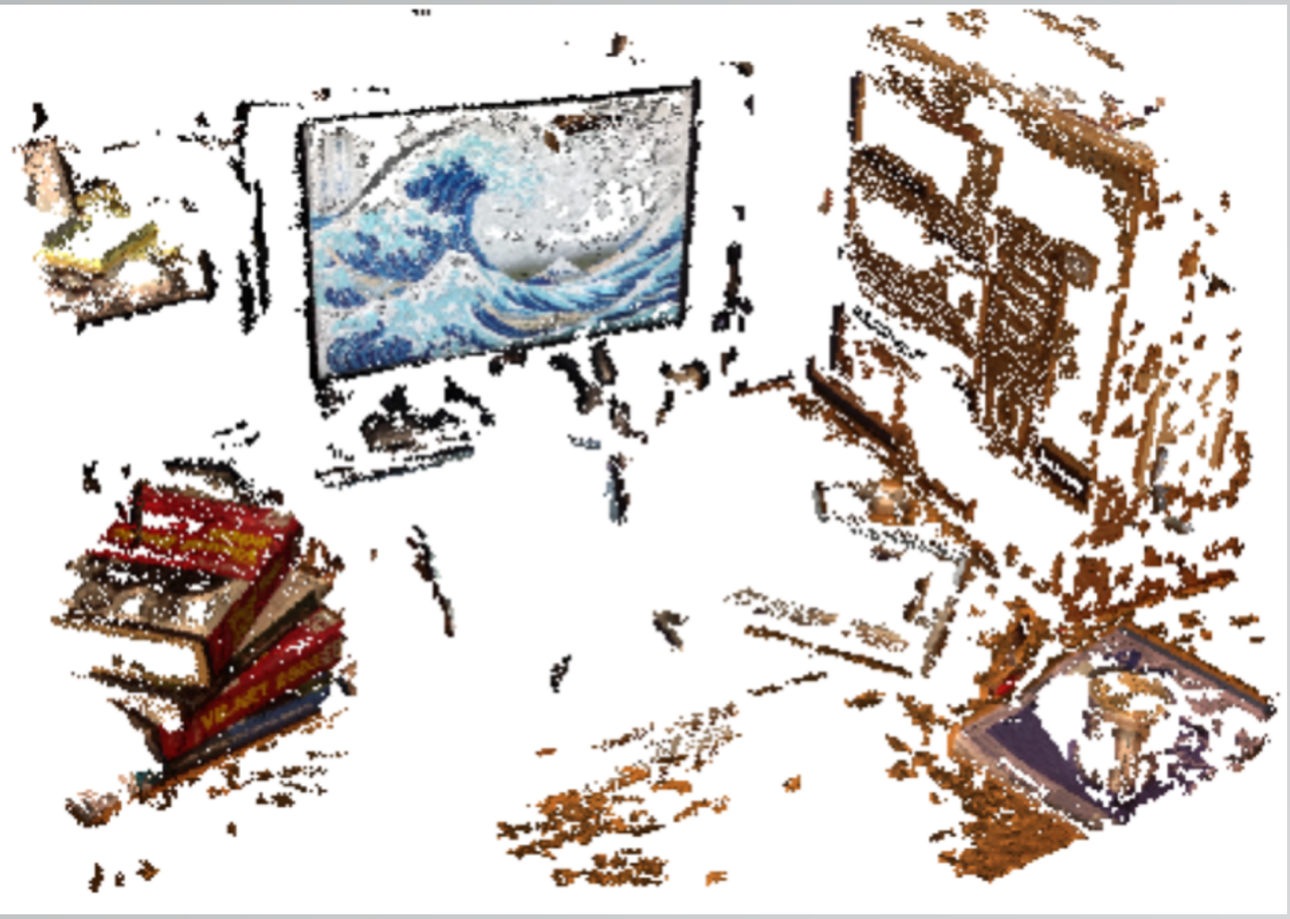
Problem Statement

- Estimate Structure and Motion from a pair of uncalibrated images

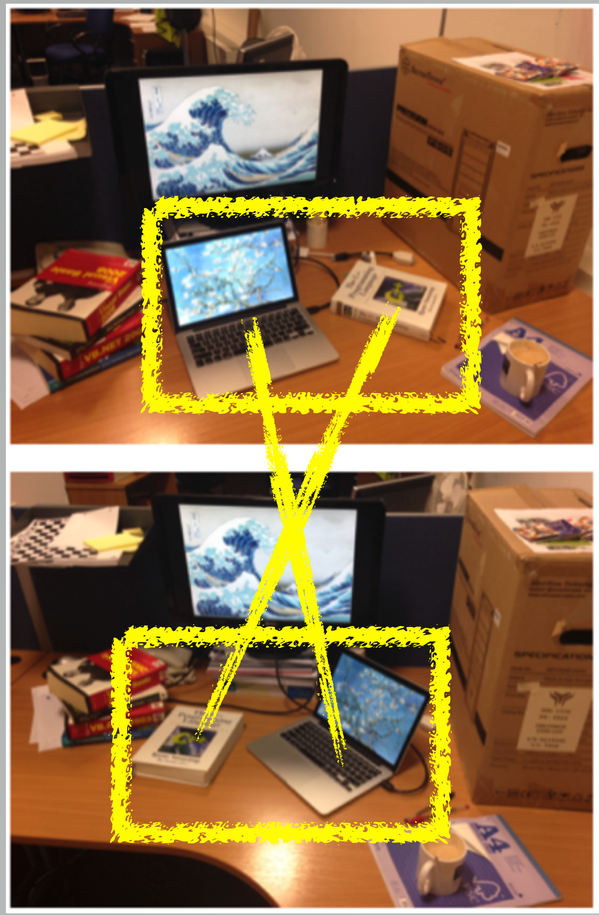
Output： structure reconstruction & motions estimation & camera parameters calibration

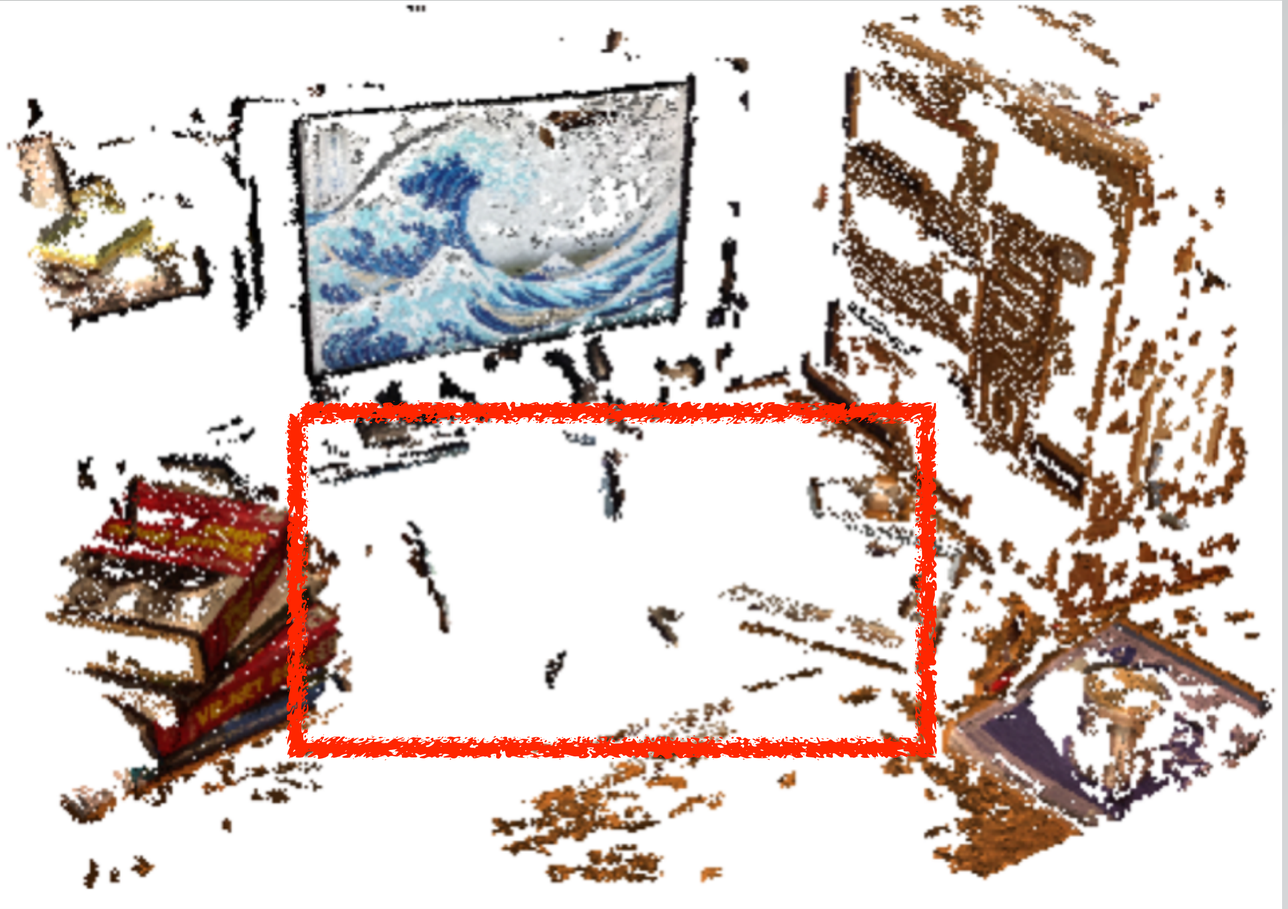


Direct Solution



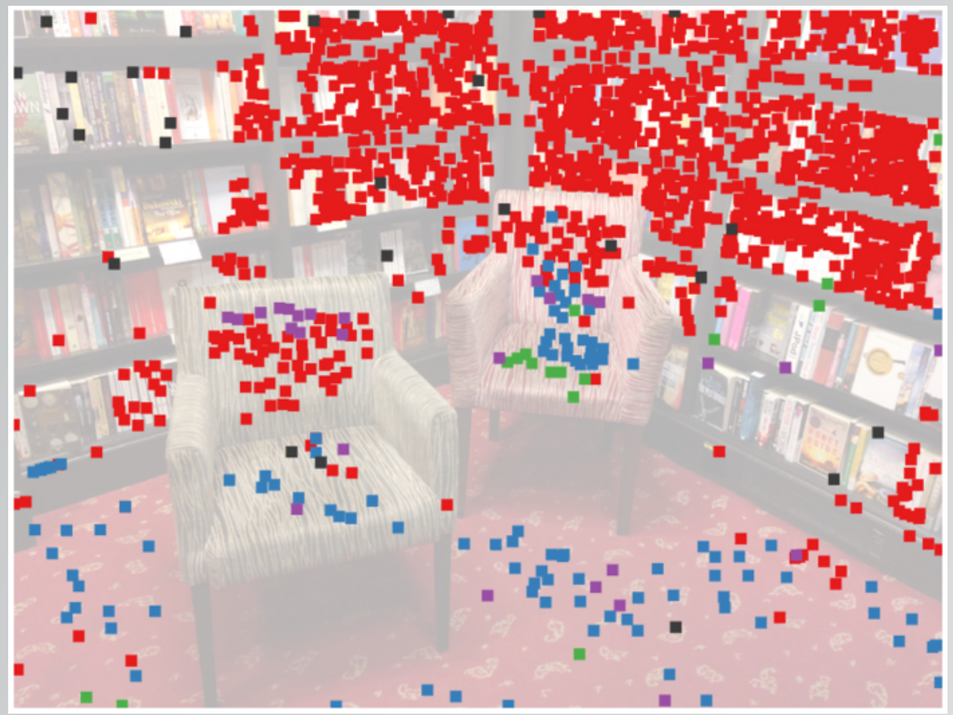
Direct Solution





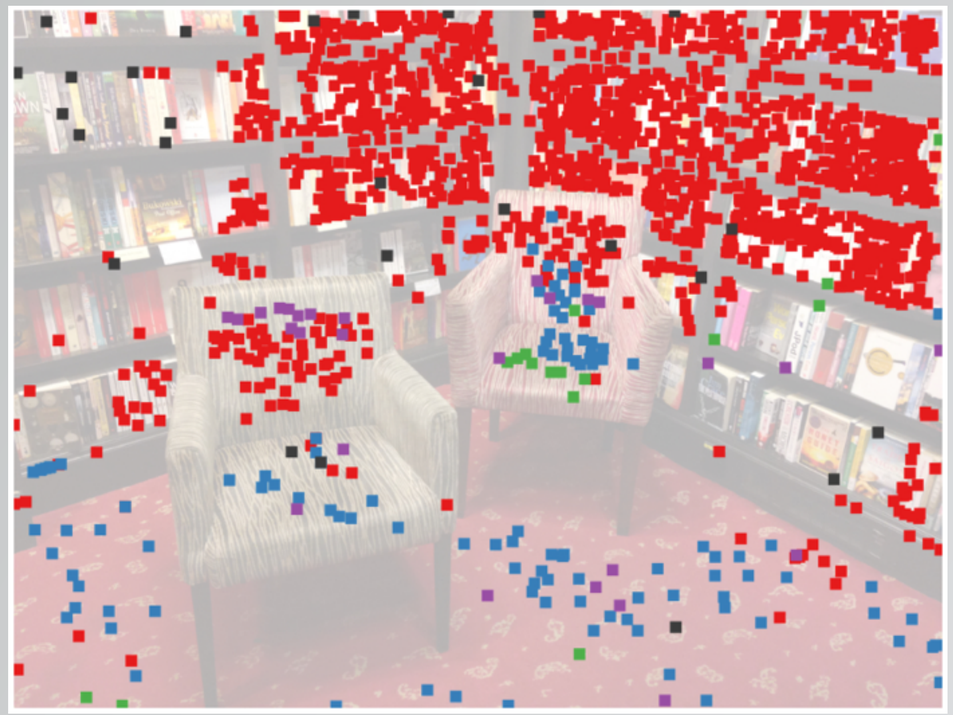
Combined Solution?

- Motion Segmentation
  - GPCA [VH04, YRM06, RTVM10]
  - Video based: [OSVG10], [JPS14], [RRGA12]
  - **PEARL** [IB12, DOIB12] *(state of the art)*



Combined Solution?

- Motion Segmentation
  - GPCA [VH04, YRM06, RTVM10]
  - Video based: [OSVG10], [JPS14], [RRGA12]
  - **PEARL** [IB12, DOIB12] *(state of the art)*
- Structure Estimation
  - **Structure from Motion** (SfM) [Wu11]
- Dense Reconstruction
  - **CMVS/PMVS** [FP10]
  - CMPMVS [JP11]
  - MVE [SFG14]
  - …



Challenges

Challenges

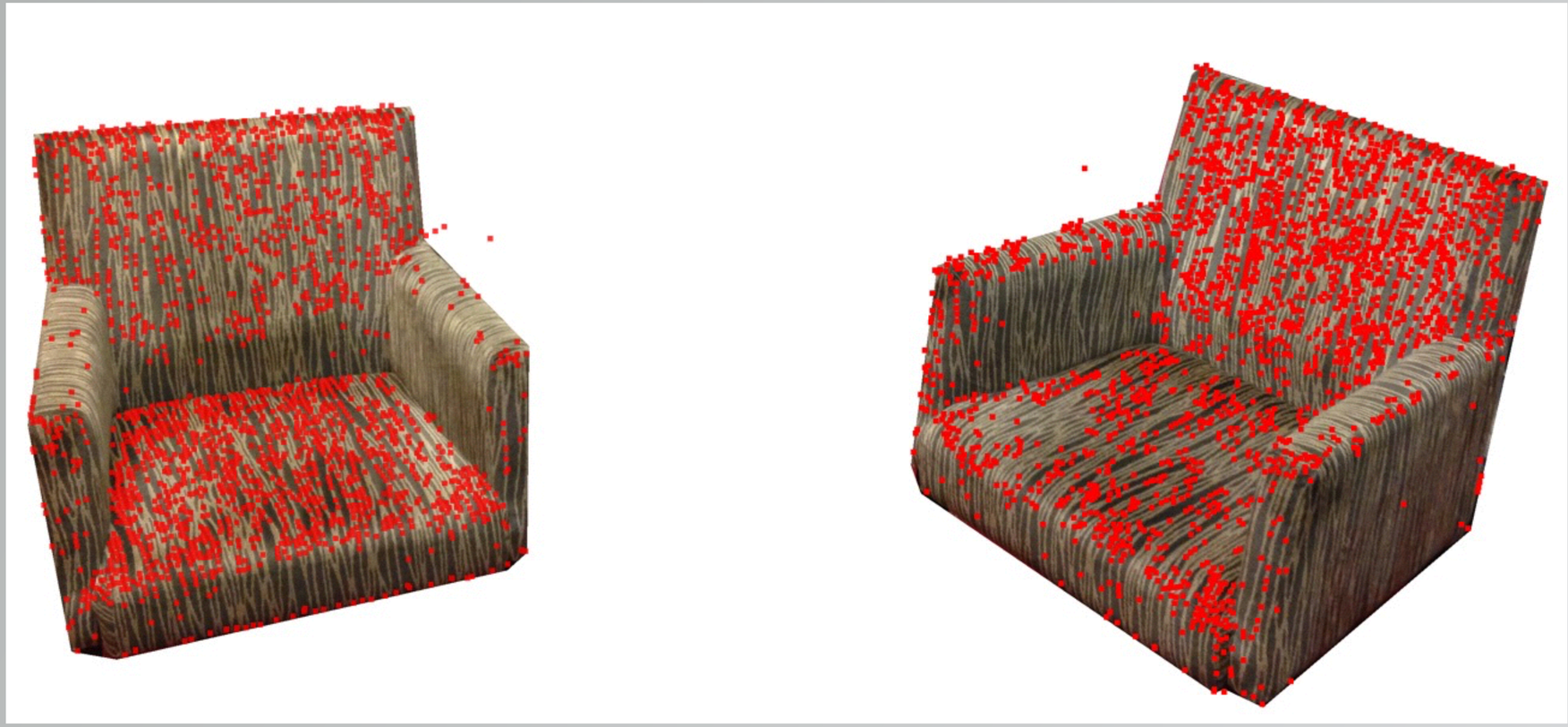
- Obtaining good quality point correspondences from only two images

Challenges

- Obtaining good quality point correspondences from only two images

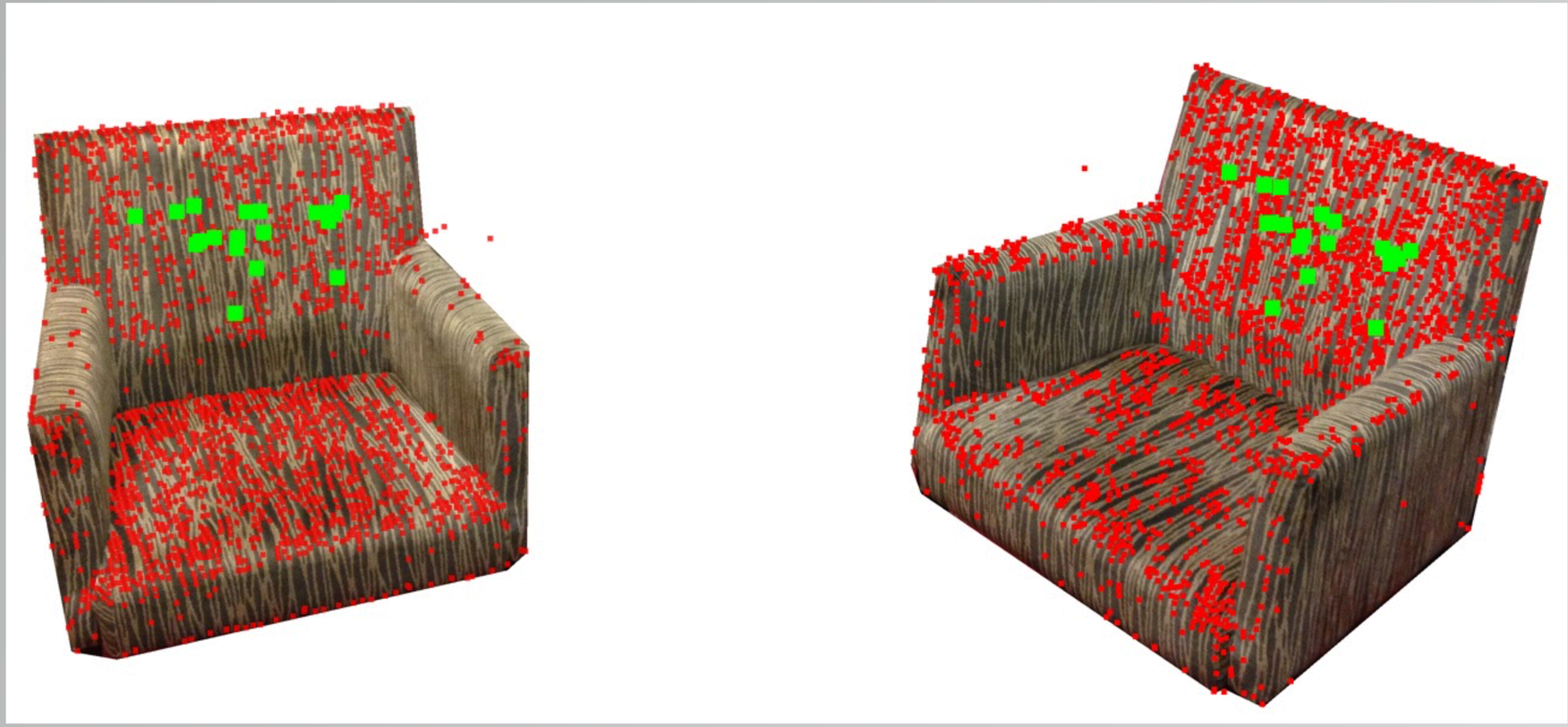
Challenges

- Obtaining good quality point correspondences from only two images



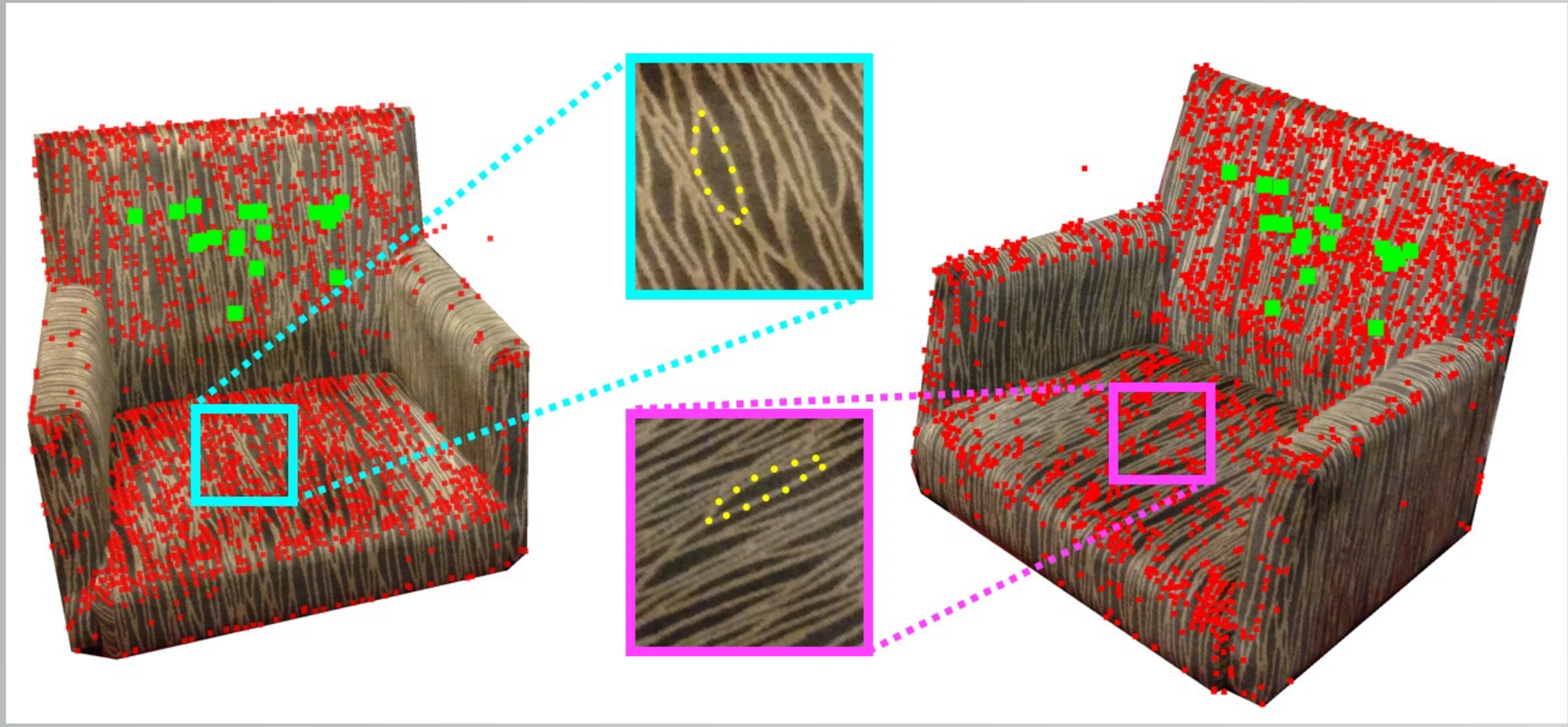
Challenges

- Obtaining good quality point correspondences from only two images



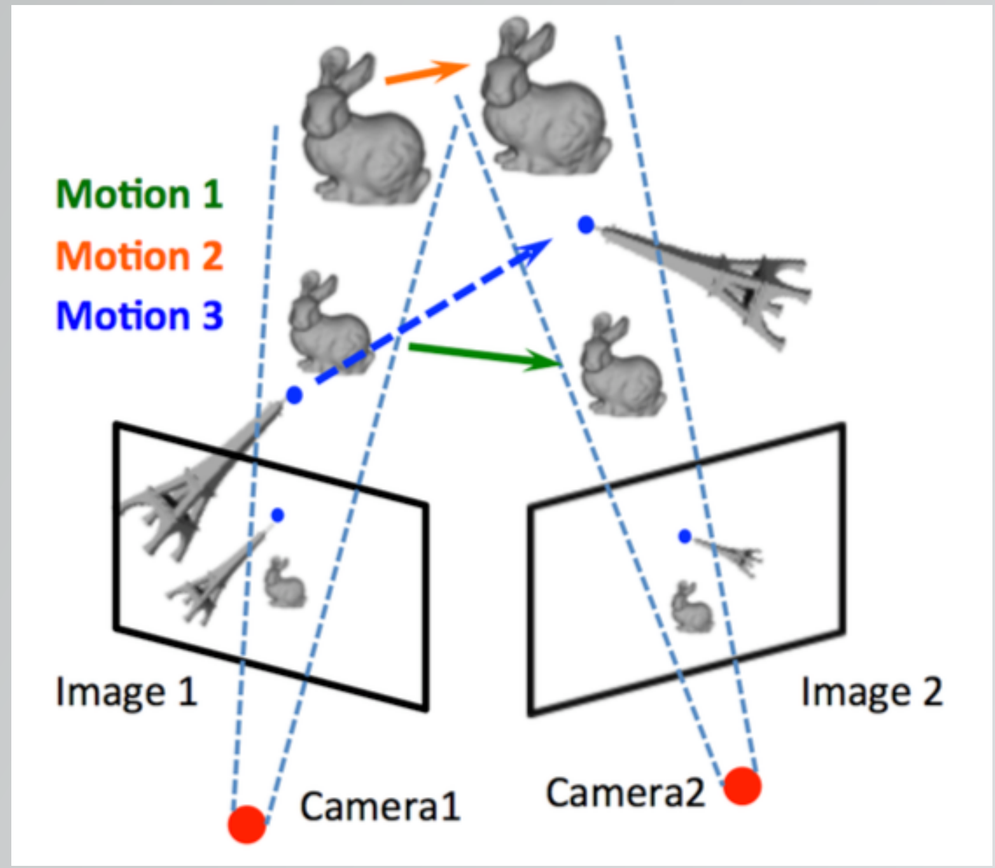
Challenges

- Obtaining good quality point correspondences from only two images



Challenges

- Obtaining good quality point correspondences from only two images
- Simultaneously estimate camera parameters, motions, groupings and structure via a continuous variable optimization
  - Initialize the optimization
  - two-view ambiguity



Challenges

- Obtaining good quality point correspondences from only two images
- Simultaneously estimate camera parameters, motions, groupings and structure via a continuous variable optimization
  - Initialize the optimization
  - two-view ambiguity
- Generate denser 3D point cloud

Dynamic SfM

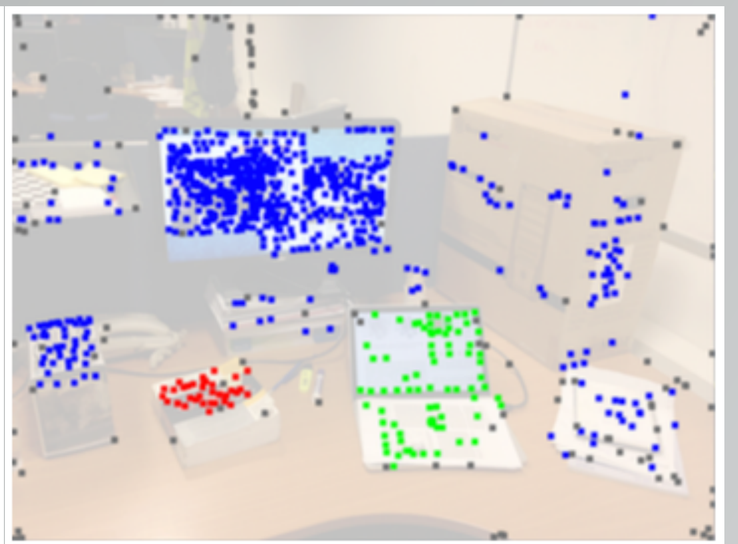
Input

Dynamic SfM

Input

Pre-boosting
generate correspondence

Dynamic SfM

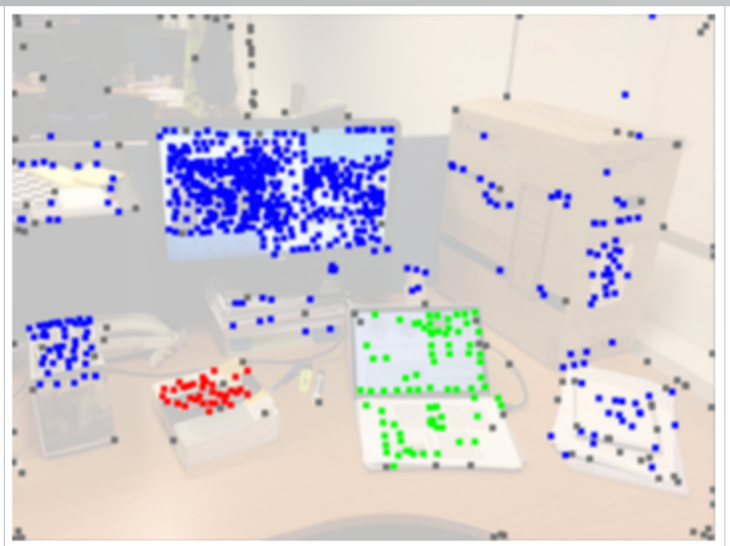


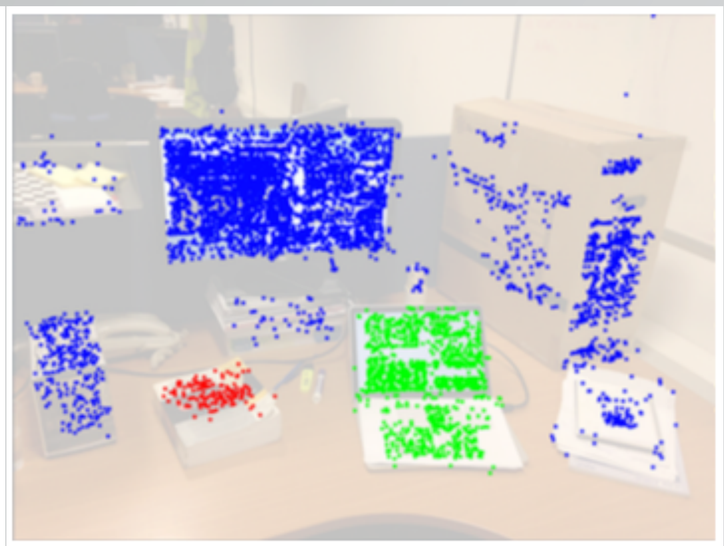
Input

continuous optimization
correspondence grouping

Pre-boosting
generate correspondence

Dynamic SfM





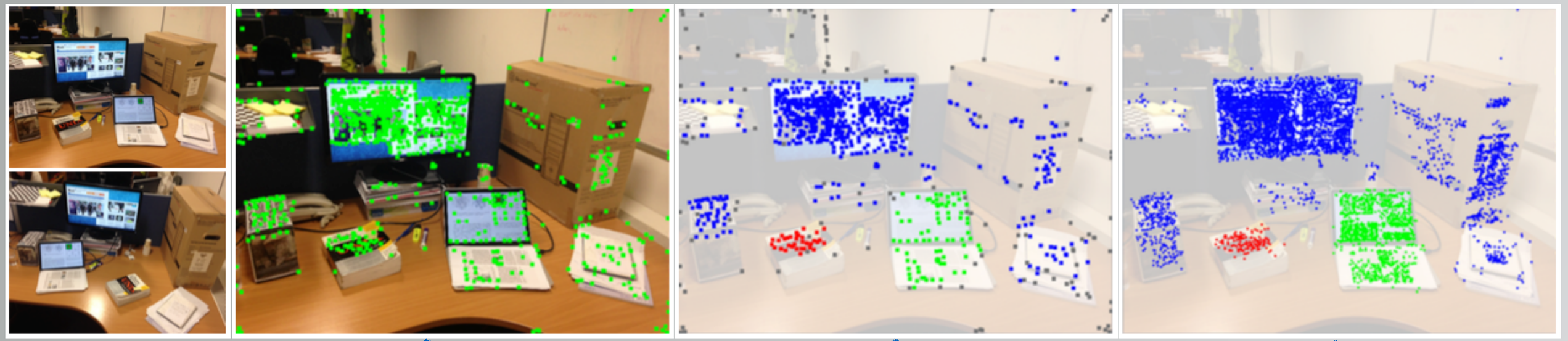
Input

continuous optimization
correspondence grouping

Pre-boosting
generate correspondence

Post-boosting
generate denser point cloud

Dynamic SfM



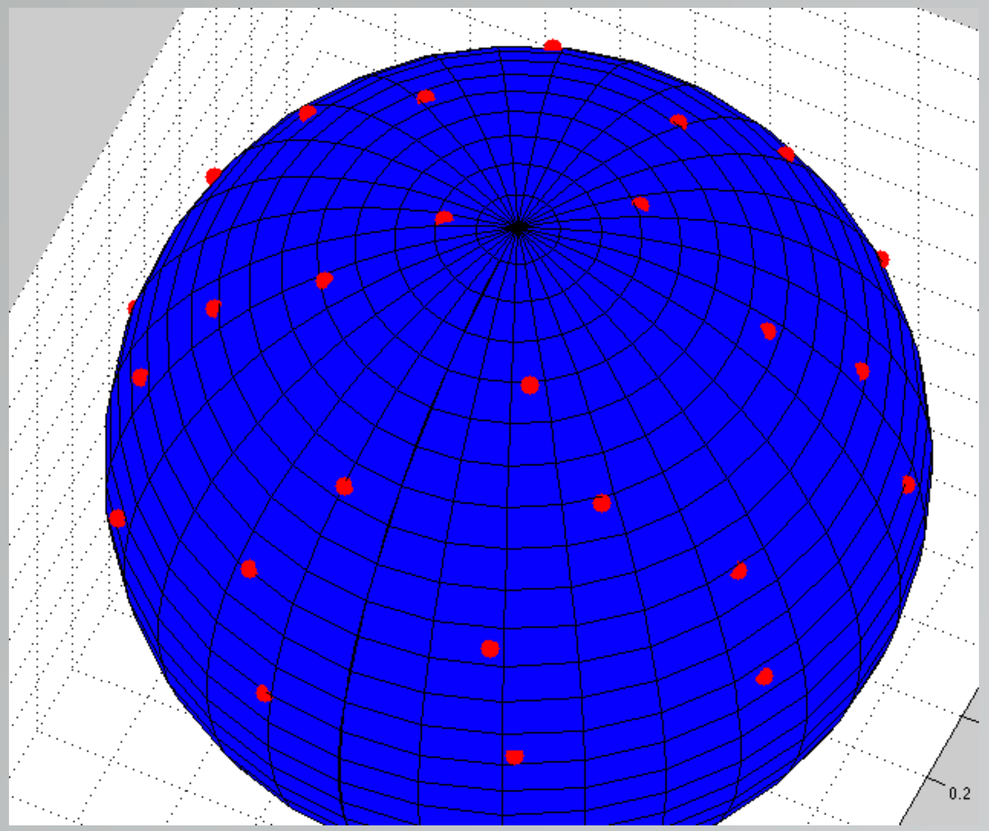
Input

continuous optimization
correspondence grouping

Pre-boosting
generate correspondence

Post-boosting
generate denser point cloud

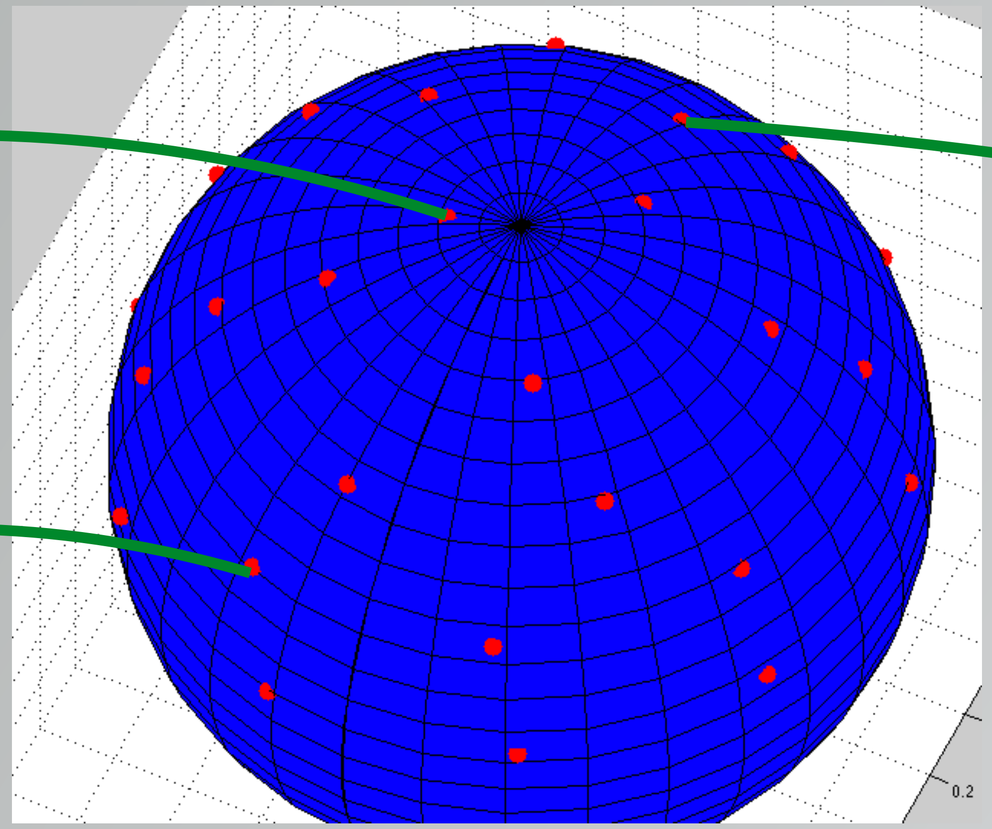
Pre-boosting



$$u_i = \arcsin\left(1 - \frac{2i-1}{2S}\right) \qquad v_i = u_i\sqrt{2\pi S}$$

$$\mathbf{n}_i = [\cos(u_i)\cos(v_i);\ \cos(u_i)\sin(v_i);\ \sin(u_i)]$$

Pre-boosting



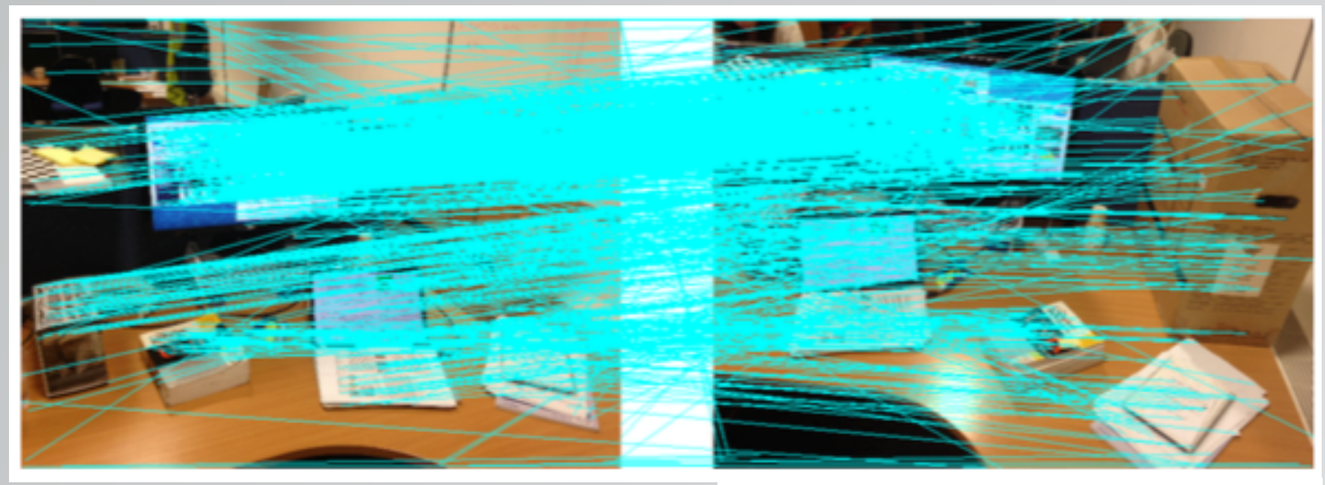
$$u_i = \arcsin\left(1 - \frac{2i-1}{2S}\right) \qquad v_i = u_i\sqrt{2\pi S}$$

$$\mathbf{n}_i = [\cos(u_i)\cos(v_i);\ \cos(u_i)\sin(v_i);\ \sin(u_i)]$$

Pre-boosting

Pre-boosting

Original match



After Pre-boosting

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Motion candidates
Camera parameters

3D positions

Group label

$$E(\{L_i\}, \{\mathbf{d}_k\}, \{\alpha^k\}) :=$$

$$E_{data} + E_{complexity} + E_{outlier} + E_{consistency}$$

Date term	lower reprojection error
Complexity term	penalizes having too many separate groups to describe the motions
Outlier term	avoid trivial solution of all outliers
Consistency term	regularizes false positive matches cause by two-view ambiguity

Continuous Optimization

Continuous Label

$$\alpha_i^k \in [0, 1]$$

$$\sum_{i=1}^{M} \alpha_i^k \leq 1 \quad \forall k$$

Continuous Optimization

Continuous Label

$$\alpha_i^k \in [0,1]$$

$$\sum_{i=1}^{M} \alpha_i^k \leq 1 \quad \forall k$$

$$E_{data}(\{L_i\},\{\mathbf{d}_k\},\{\alpha^k\}) = \sum_{k=1}^{N}\sum_{i=1}^{M} \alpha_i^k \| \mathbf{d}_k \|_{L_i}$$

$$E_{complexity}(\alpha) = \omega_1 \cdot \sum_{i=1}^{M} \left(\sum_{k=1}^{N} \alpha_i^k \right)^{\beta}$$

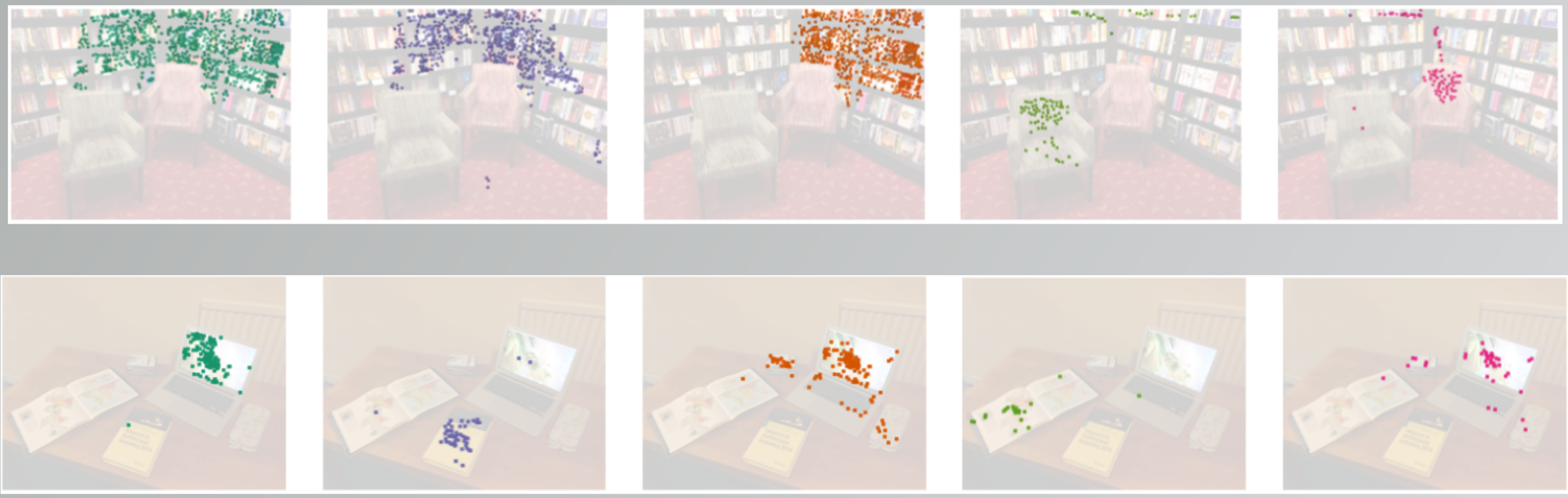
$$E_{outlier}(\alpha) = \omega_2 \cdot \left(\sum_{k=1}^{N} \delta_k \left(1 - \sum_{i=1}^{M} \alpha_i^k \right) \right)$$

$$E_{consistency}(\alpha) = \omega_3 \cdot \sum_{(p,q)\in DN} \theta_{p,q} \| \alpha^p - \alpha^q \|$$

Optimization initialization
with **Domain problem**

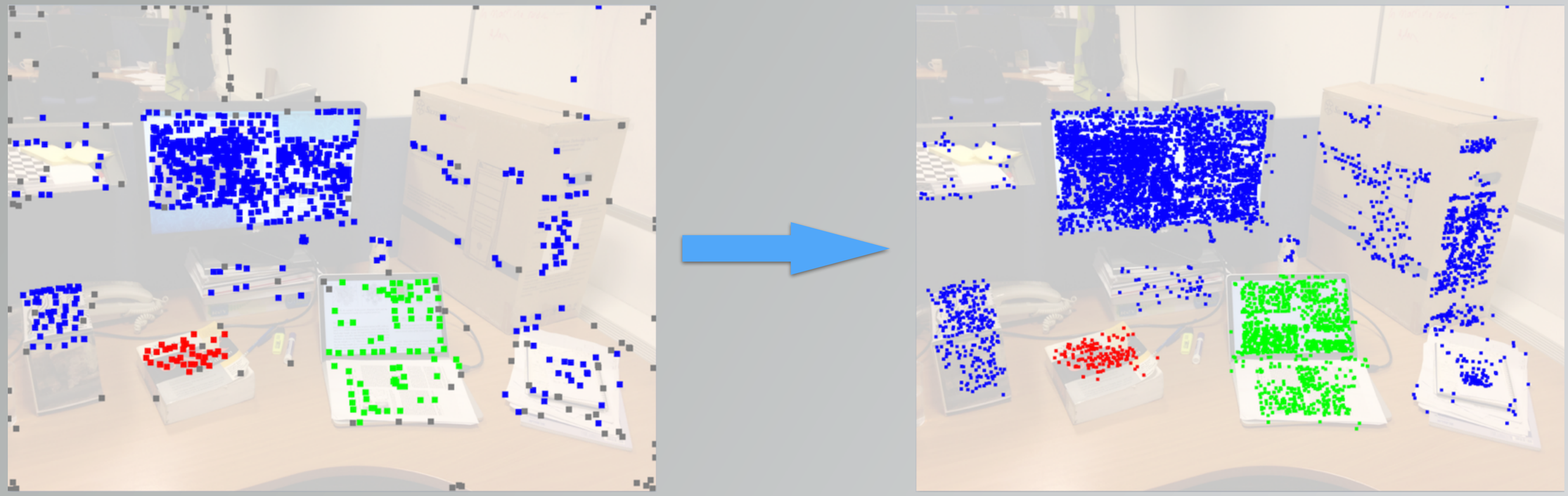
Optimization initialization
with **Domain problem**

RANSAC ⟶ ***Reweighted*** **RANSAC**



Post-boosting

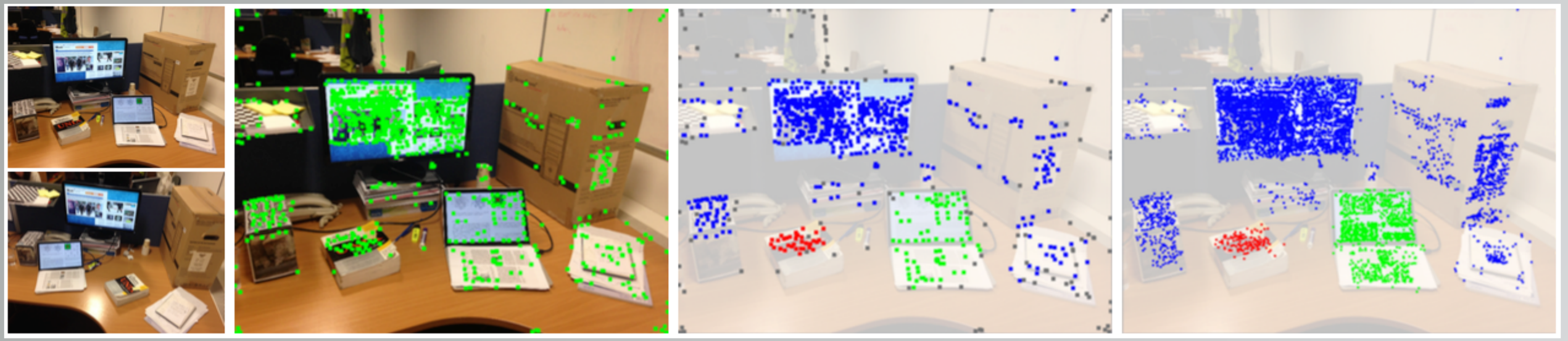
Patch based correspondence establishing

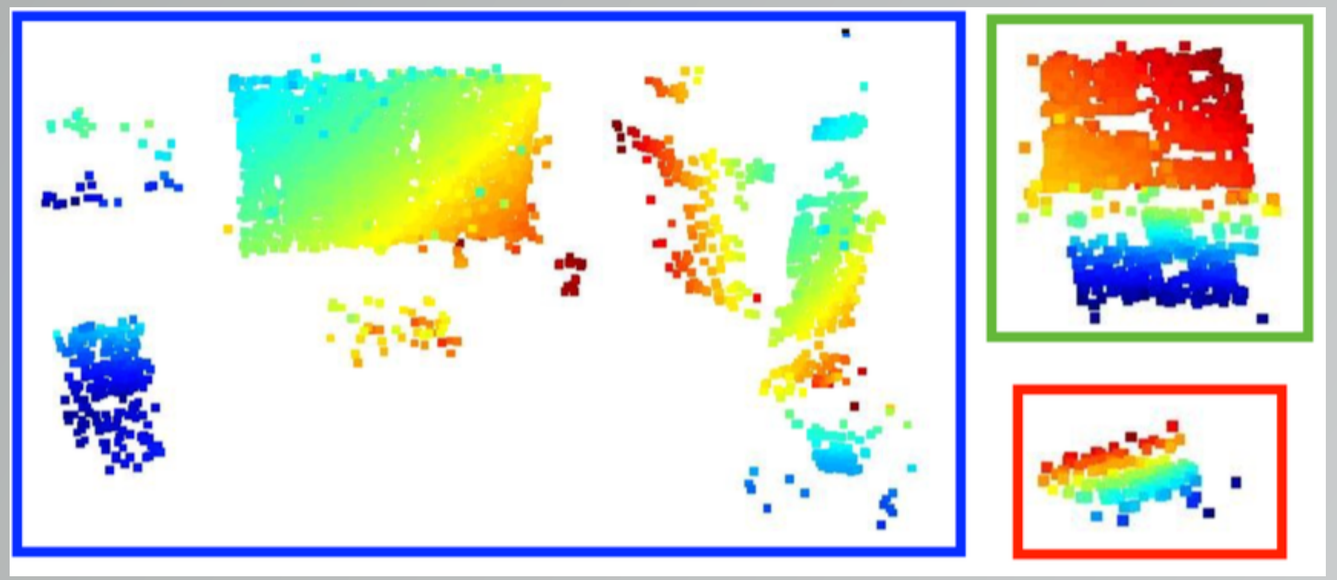


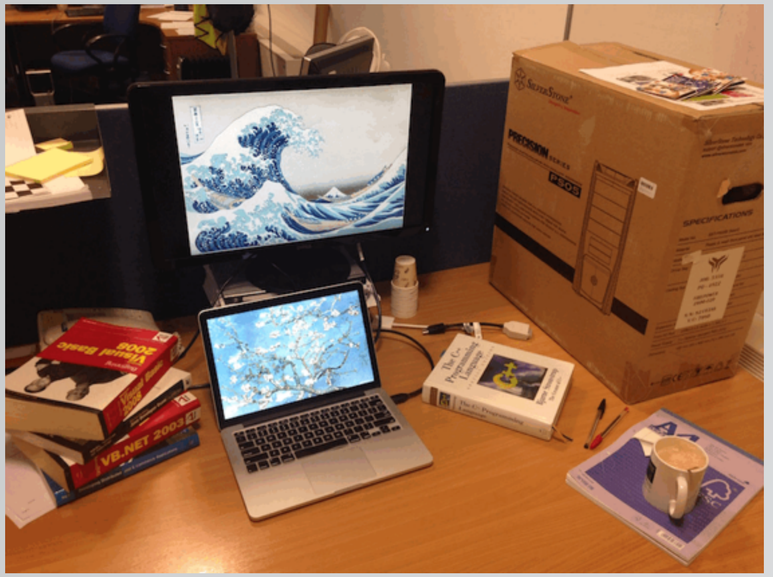
Post-boosting

Patch based correspondence establishing

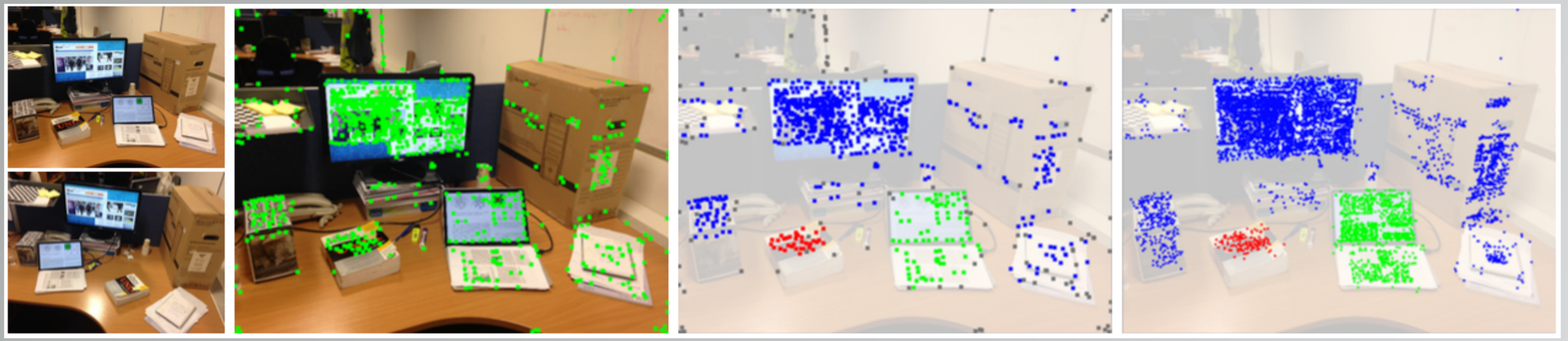
Dynamic SfM

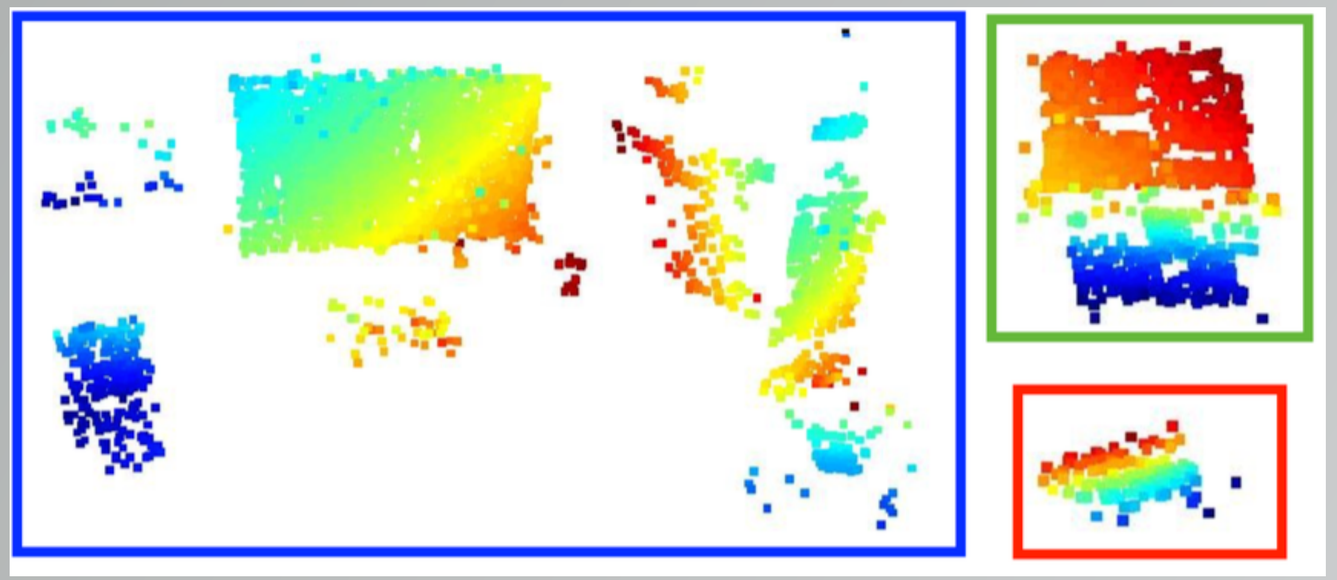


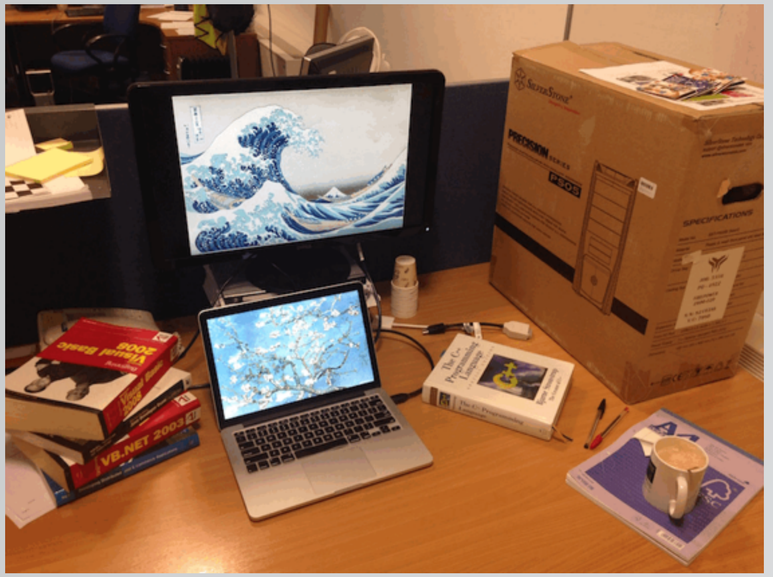




Dynamic SfM







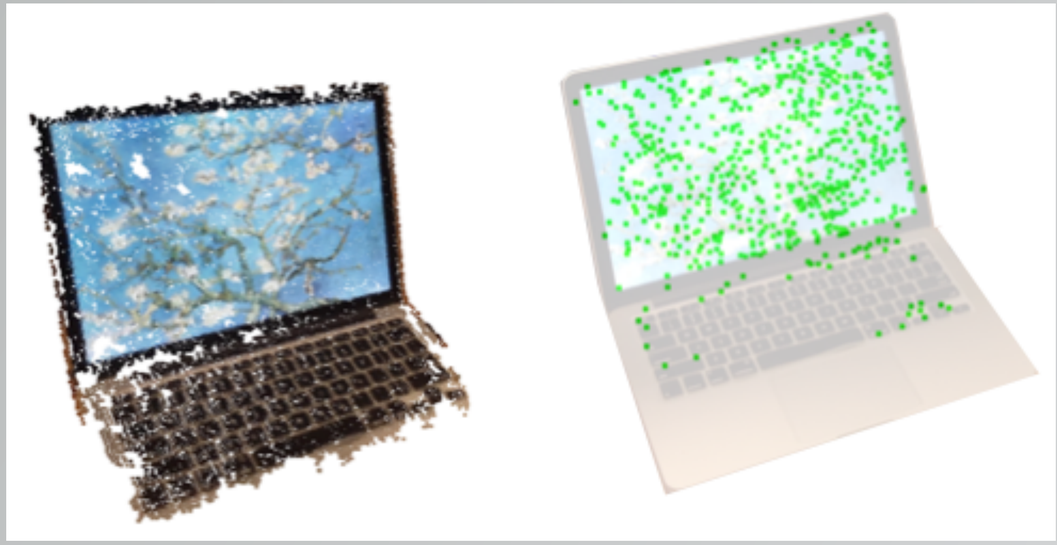
Dense Reconstruction

Dense Reconstruction

Dense Reconstruction

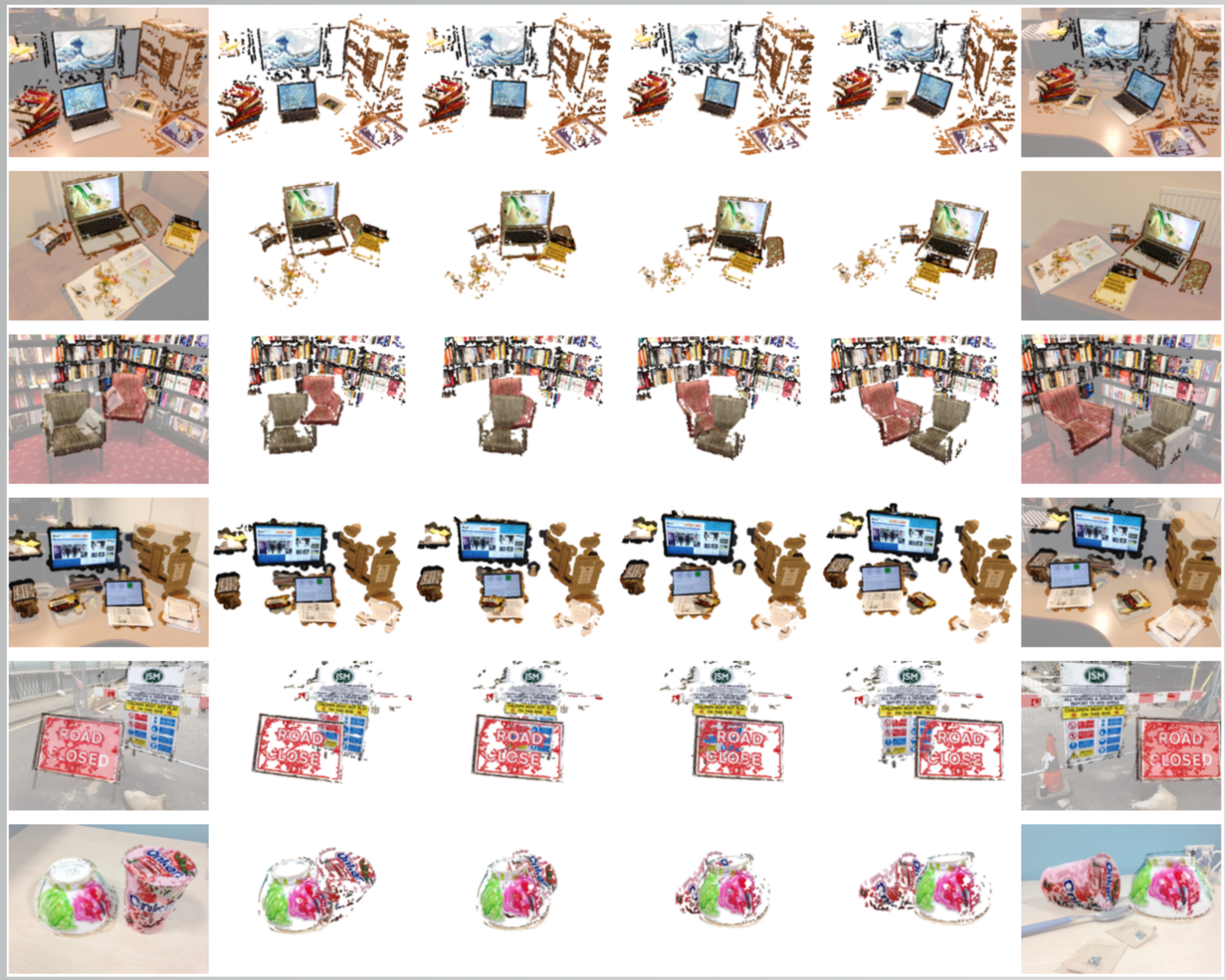
Dense Reconstruction

Dense Reconstruction



Application:

Motion Interpolation



Application:

Motion Interpolation

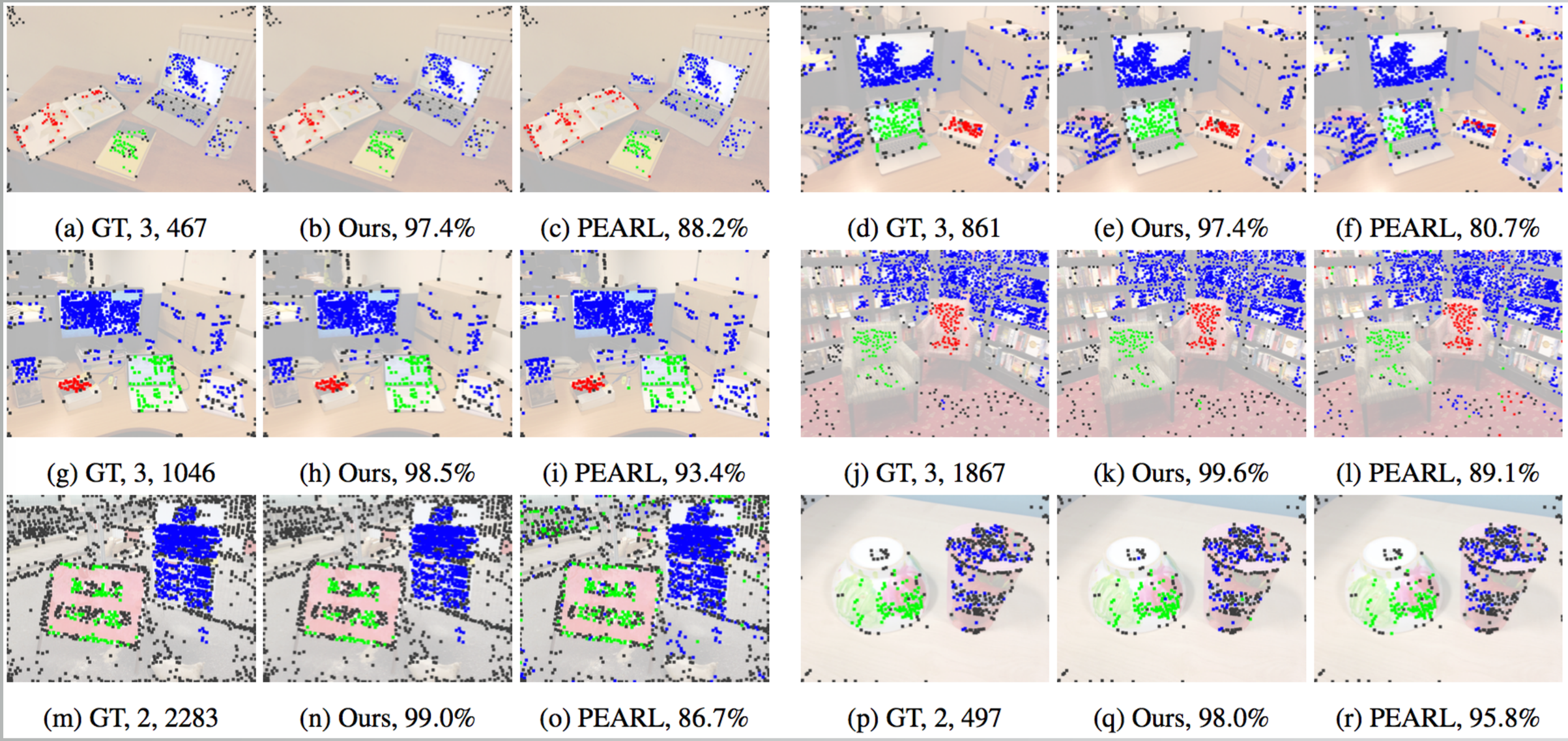
demo video available
on project page

Application:

Motion Interpolation

demo video available
on project page

Comparison with PEARL



(a) GT, 3, 467 (b) Ours, 97.4% (c) PEARL, 88.2%

(d) GT, 3, 861 (e) Ours, 97.4% (f) PEARL, 80.7%

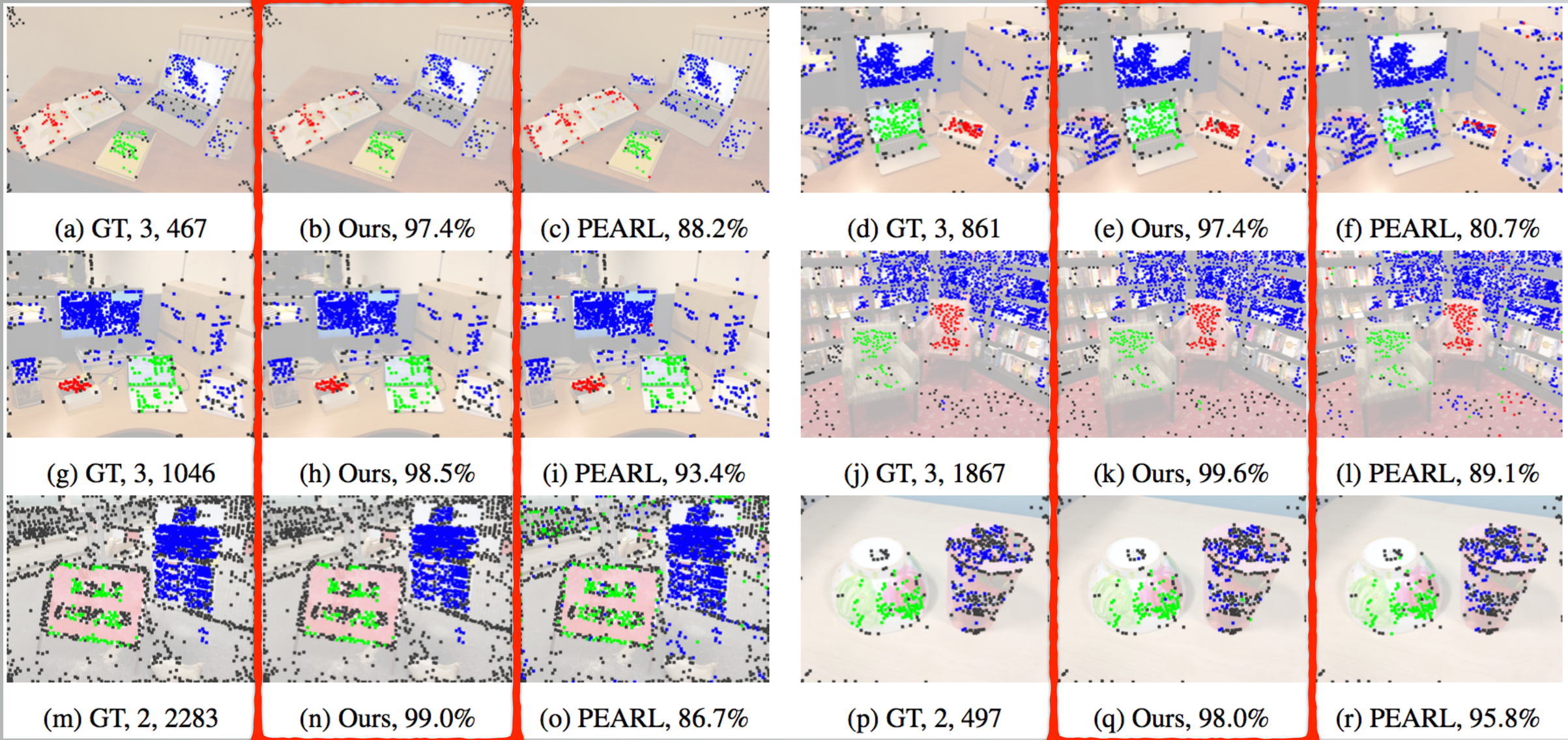
(g) GT, 3, 1046 (h) Ours, 98.5% (i) PEARL, 93.4%

(j) GT, 3, 1867 (k) Ours, 99.6% (l) PEARL, 89.1%

(m) GT, 2, 2283 (n) Ours, 99.0% (o) PEARL, 86.7%

(p) GT, 2, 497 (q) Ours, 98.0% (r) PEARL, 95.8%

Comparison with PEARL



(a) GT, 3, 467 (b) Ours, 97.4% (c) PEARL, 88.2%

(d) GT, 3, 861 (e) Ours, 97.4% (f) PEARL, 80.7%

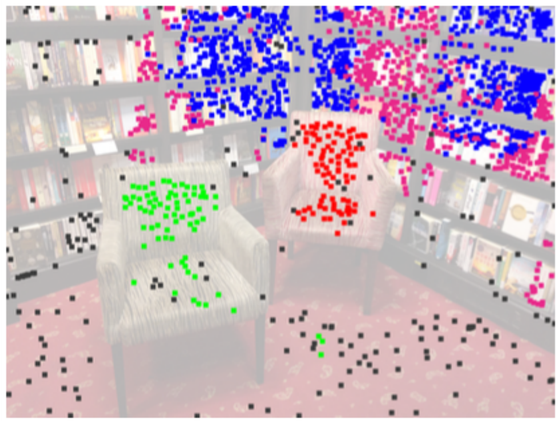
(g) GT, 3, 1046 (h) Ours, 98.5% (i) PEARL, 93.4%

(j) GT, 3, 1867 (k) Ours, 99.6% (l) PEARL, 89.1%

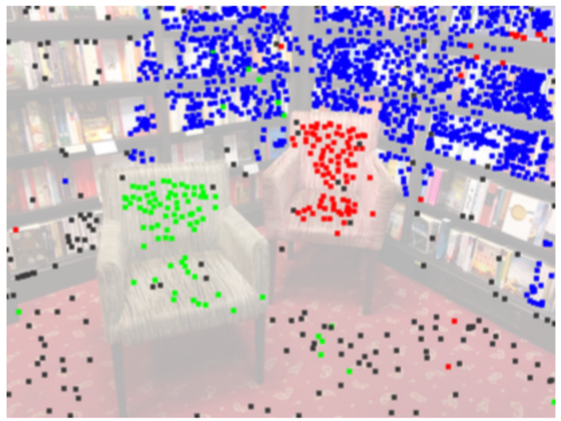
(m) GT, 2, 2283 (n) Ours, 99.0% (o) PEARL, 86.7%

(p) GT, 2, 497 (q) Ours, 98.0% (r) PEARL, 95.8%

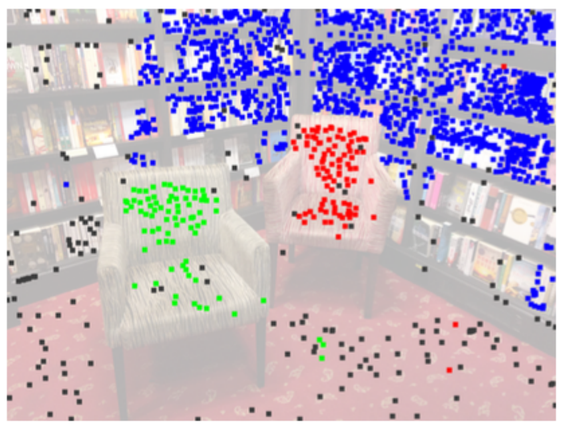
Effect of Terms



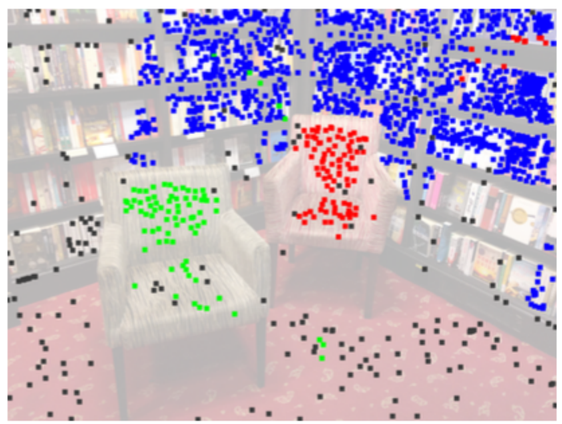
(a) no complexity term, 4 groups are generated.



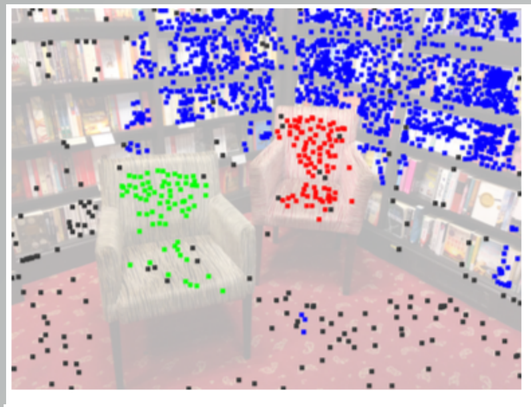
(b) uniform outlier penalty without consistency, correctness 97.4%



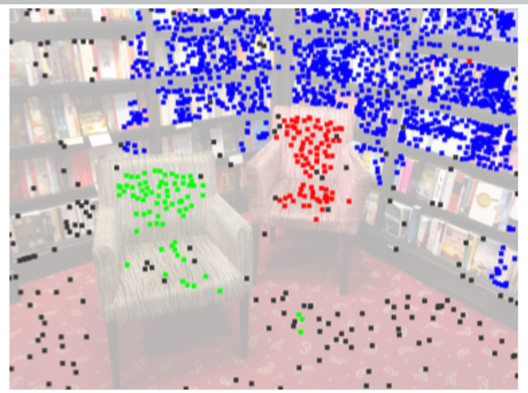
(c) uniform outlier penalty, correctness ratio 98.8%



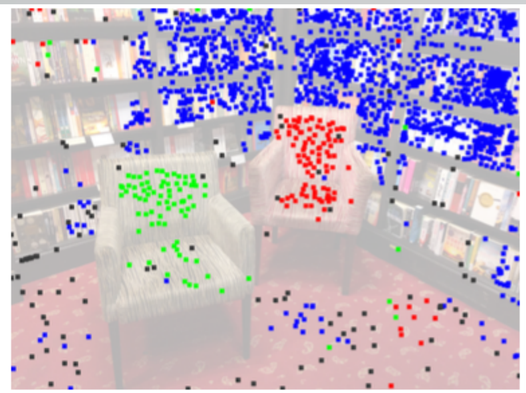
(d) no consistency term, correctness ratio 98.3%



GT

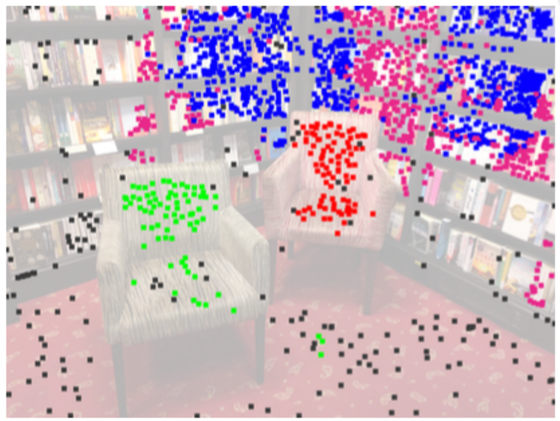


Ours , 99.6%

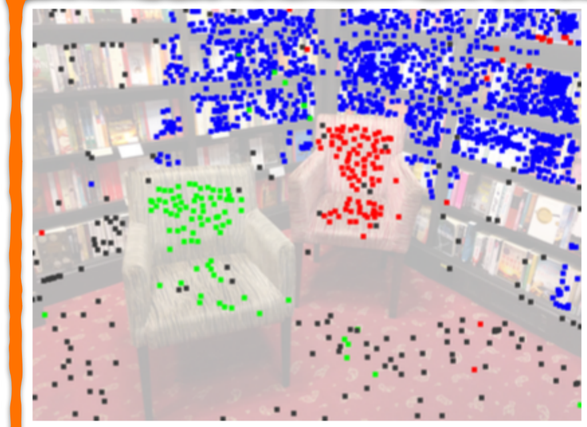


PEARL , 89.1%

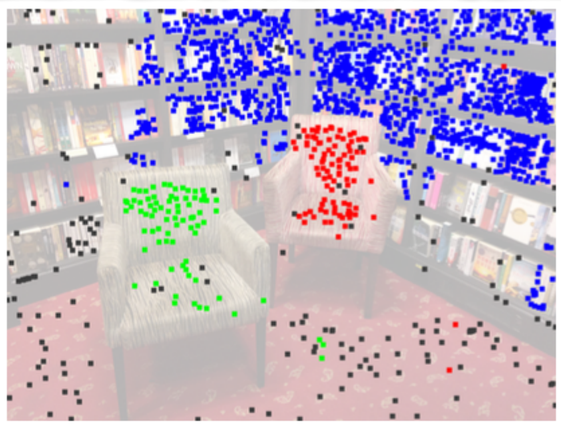
Effect of Terms



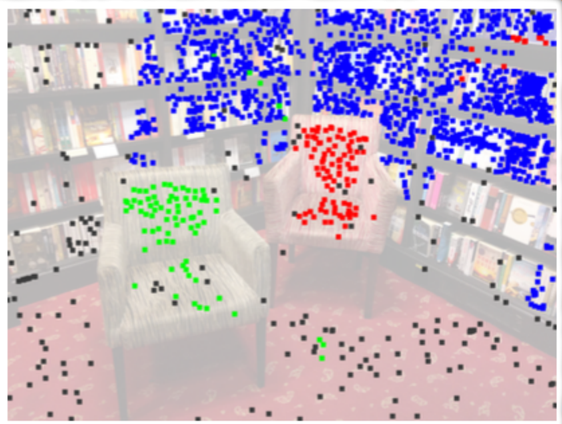
(a) no complexity term, 4 groups are generated.



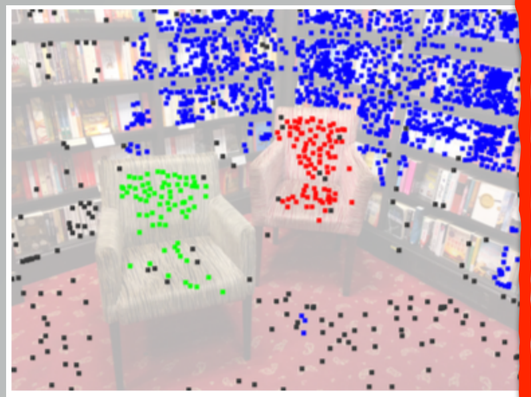
(b) uniform outlier penalty without consistency, correctness 97.4%



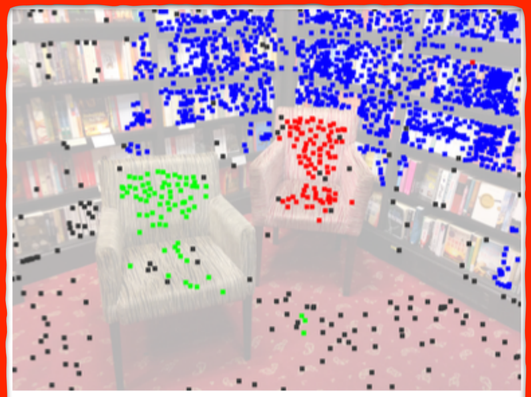
(c) uniform outlier penalty, correctness ratio 98.8%



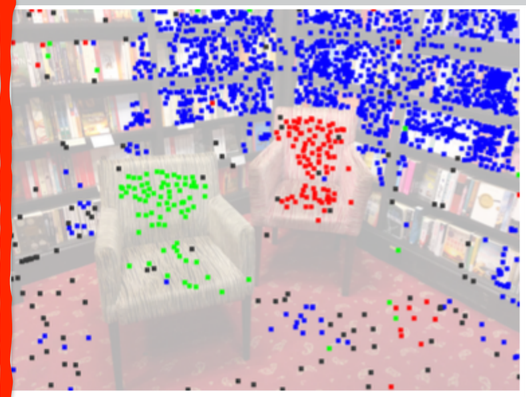
(d) no consistency term, correctness ratio 98.3%



GT



Ours , 99.6%



PEARL , 89.1%

Limitations

Limitations

- lack of texture

Limitations

- lack of texture

- small change of perspective

Conclusion

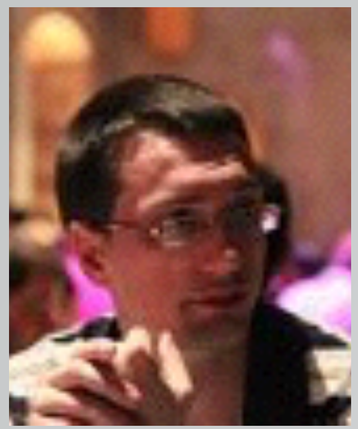
- **Dynamic SfM**
  - simultaneously recover both structure and motion of each part from a pair of input images
  - pre-boosting and post-boosting of correspondence

Conclusion

- **Dynamic SfM**
  - simultaneously recover both structure and motion of each part from a pair of input images
  - pre-boosting and post-boosting of correspondence
- **Future work**
  - prior knowledge of the structure
  - towards real time performance
  - multi-view input

Acknowledgement

- Invaluable comments, support and discussions from
  Aron Monszpart and **Duygu Ceylan** and **Christopher Russell**.



- This work was supported by ERC Starting Grant SmartGeometry (StG-2013335373) and Microsoft PhD scholarship

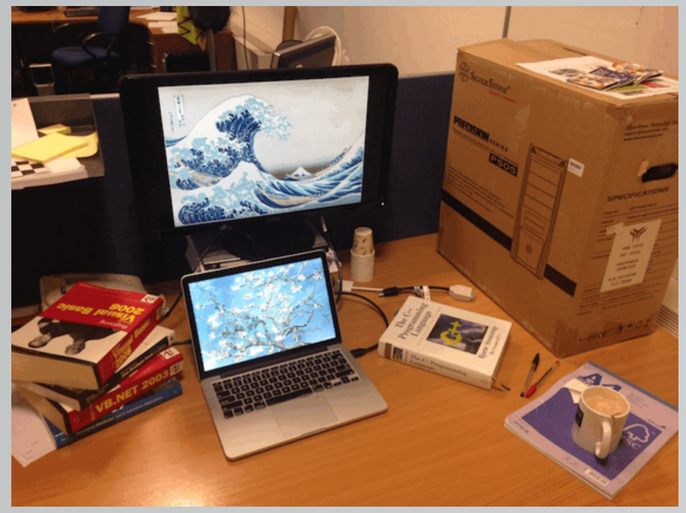
Microsoft® Research

erc

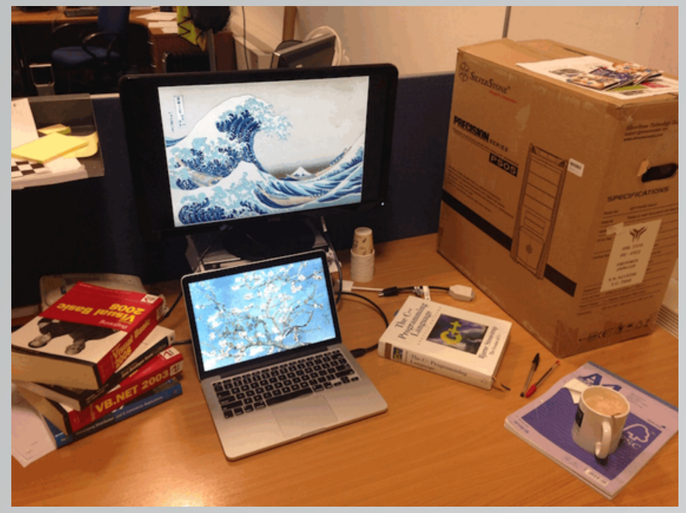
Thank you for your attention

- webpage (code+data)

http://geometry.cs.ucl.ac.uk/projects/2015/dynamicSfM/



Thank you for your attention

- webpage (code+data)

http://geometry.cs.ucl.ac.uk/projects/2015/dynamicSfM/