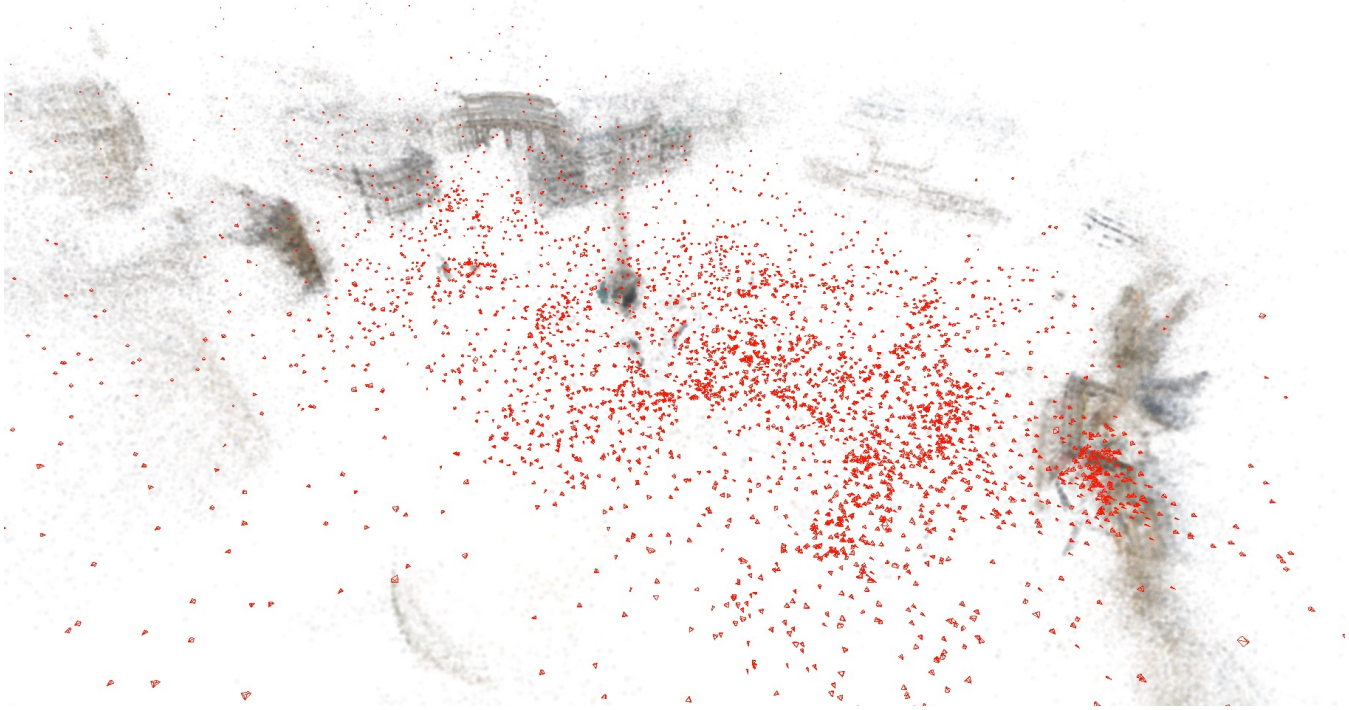
GraphMatch: Efficient Large-Scale Graph Construction for Structure from Motion



Qiaodong Cui[1], Victor Fragoso[2], Chris Sweeney[3], Pradeep Sen[1]

[1]University of California, Santa Barbara

[2]West Virginia University

[3]University of Washington

SfM background

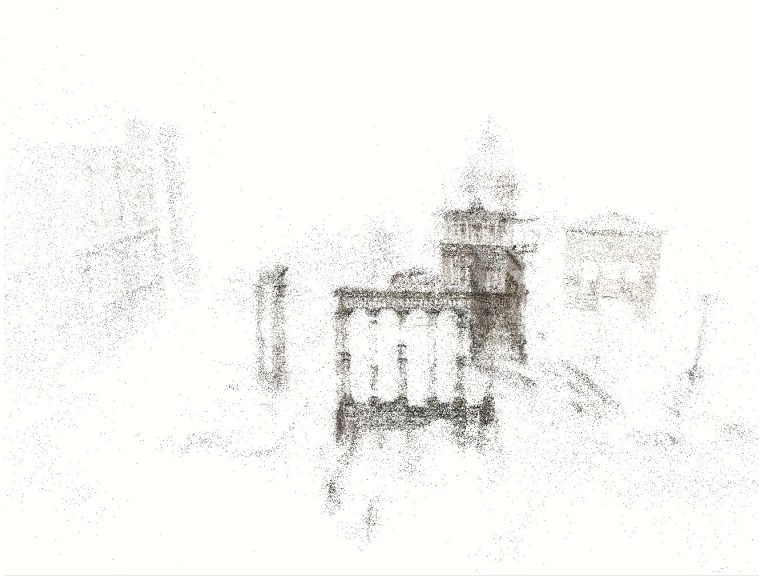
Images



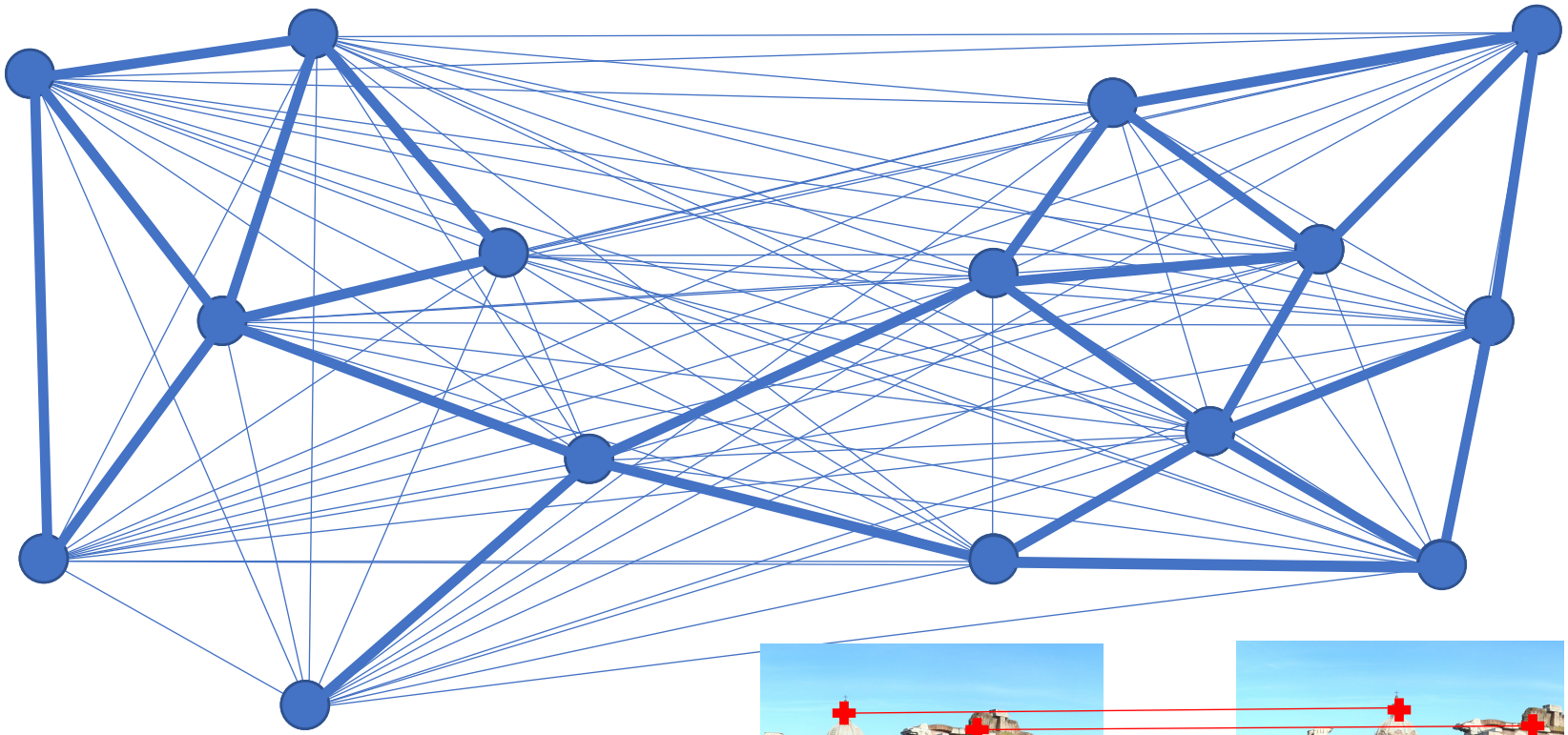


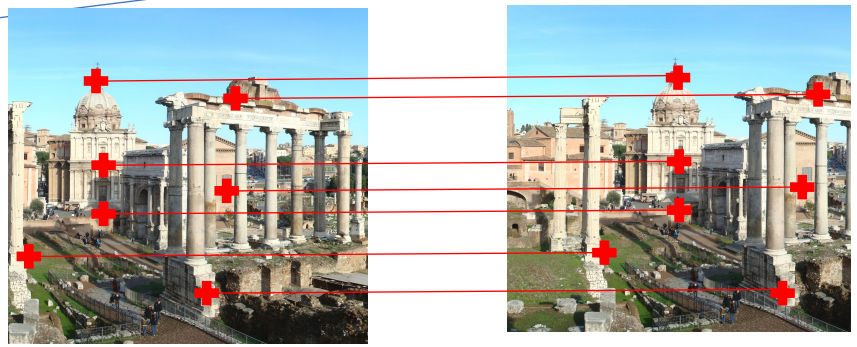
...

Reconstruction

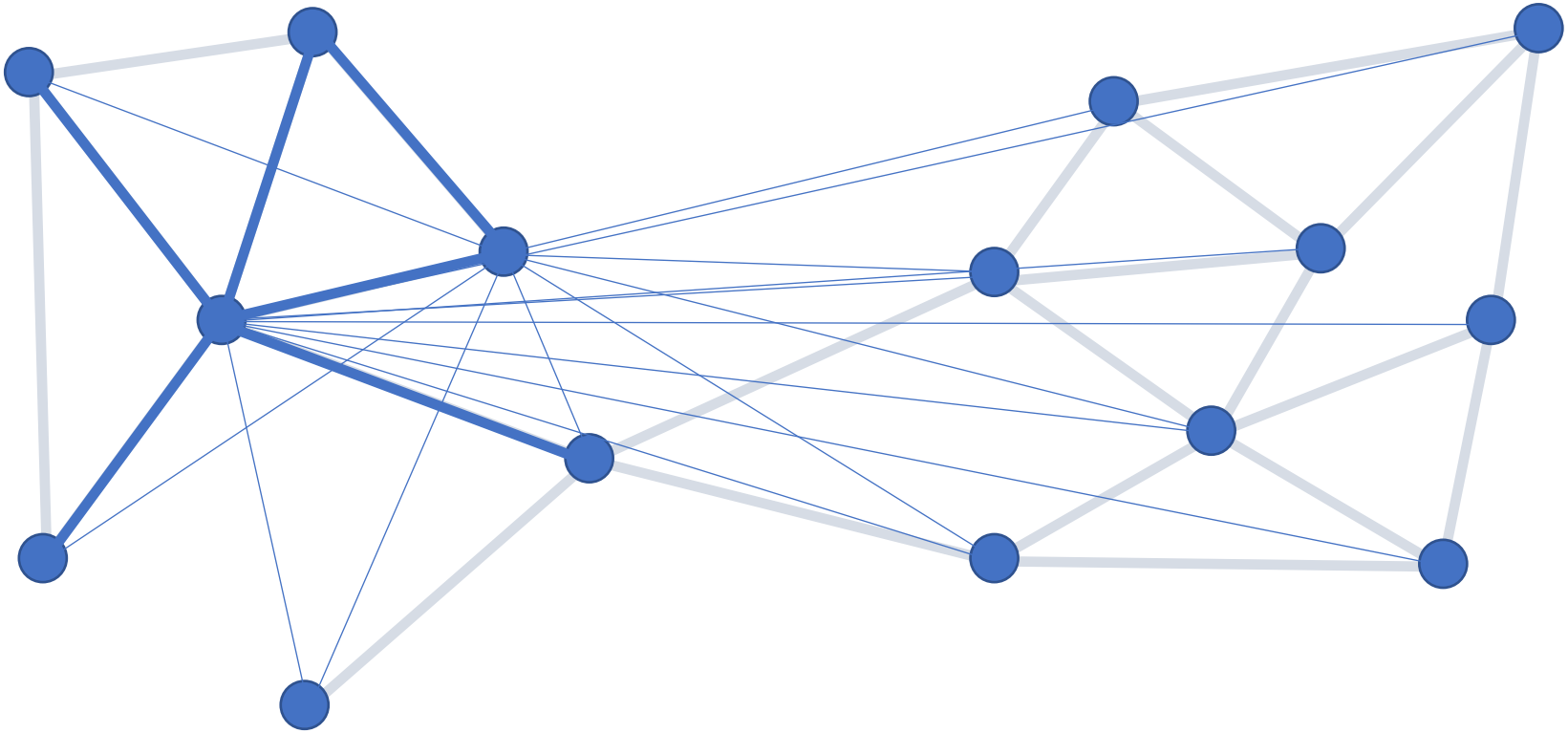


Matching graph





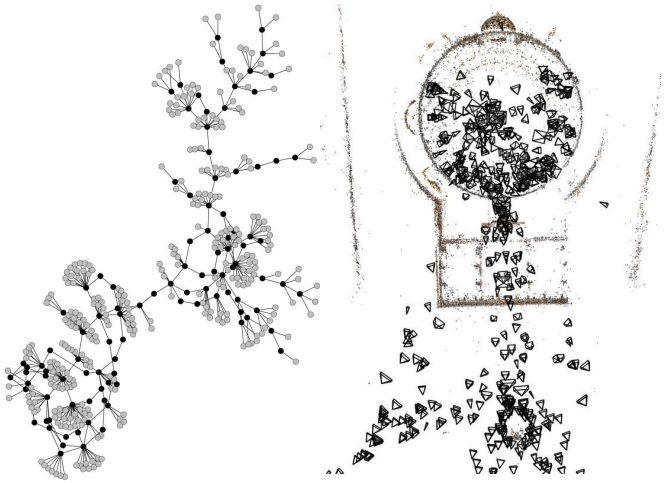
Brute Force



Building Rome in a day

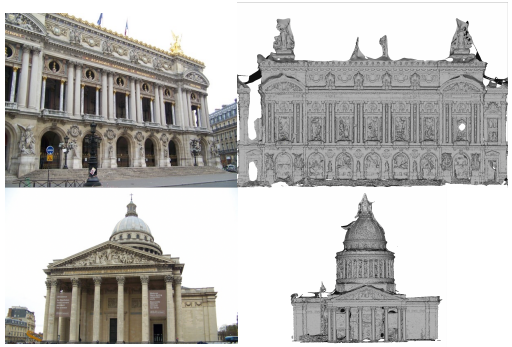
[Agarwal et al. 2009]

Skeletal graphs



[Snavely et al. 2008]

Global fusion

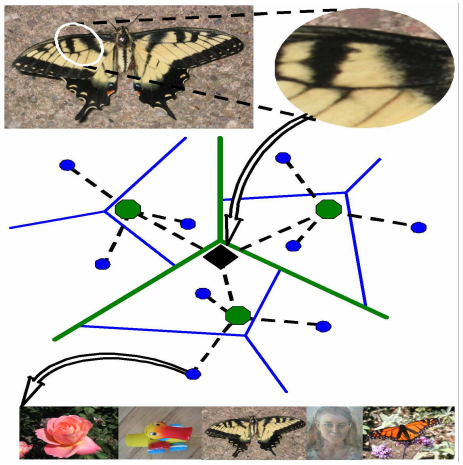


[Moulon et al. 2013]

SfM Revisited

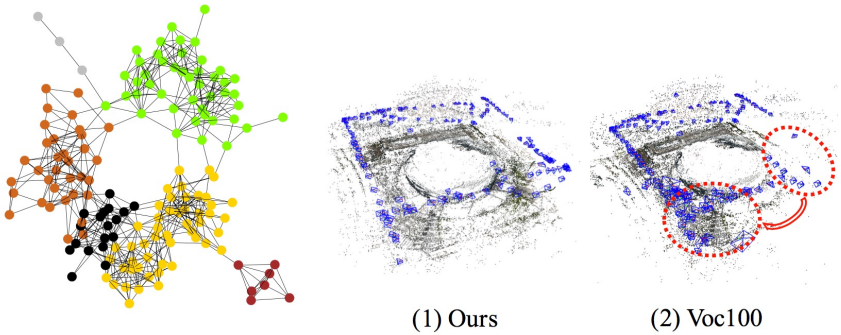
[Schönberger et al. 2016]

Vocabulary tree



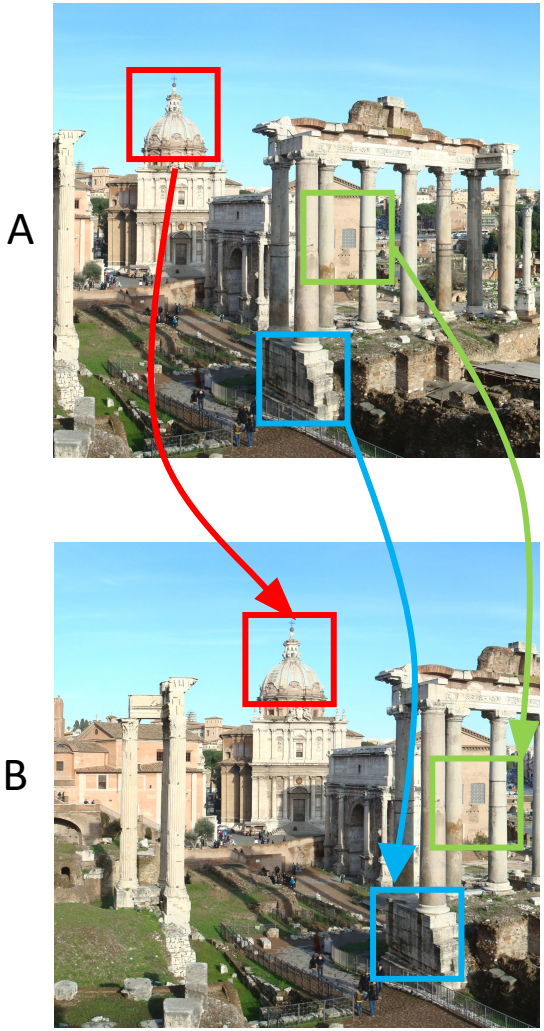
[Nister et al. 2006]

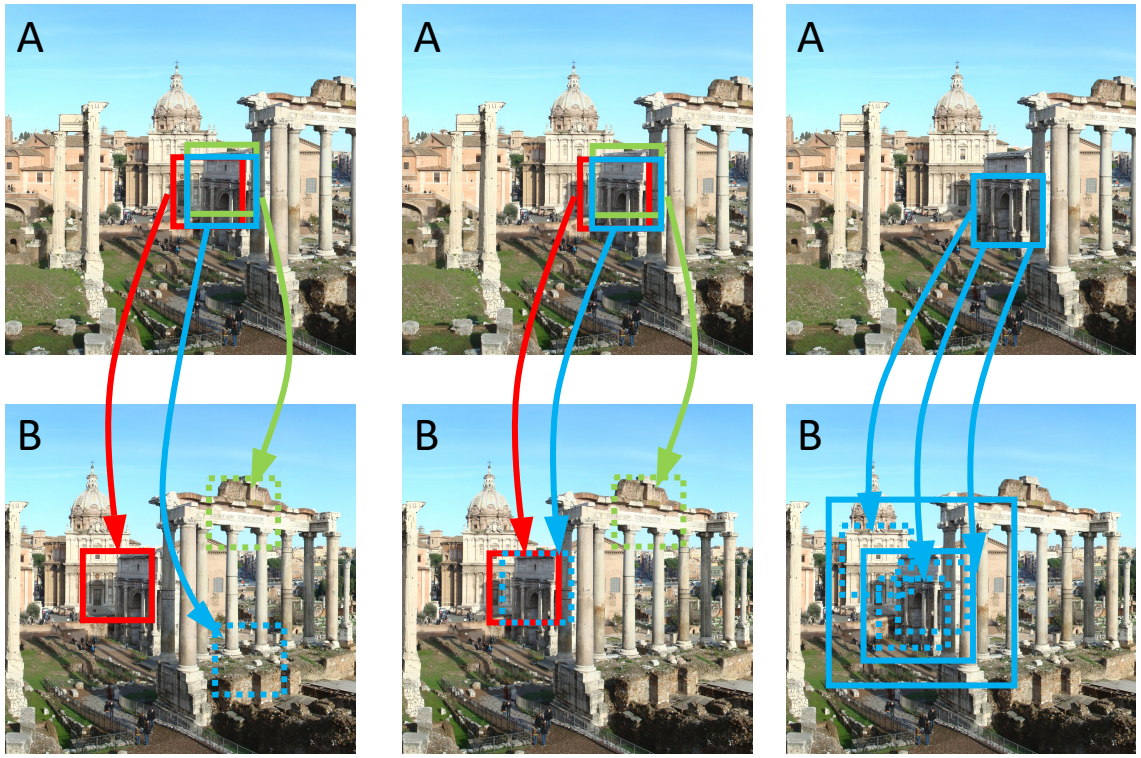
Graph based Match



[Shen et al. 2016]

PatchMatch





(a) Initialization (b) Propagation (c) Search

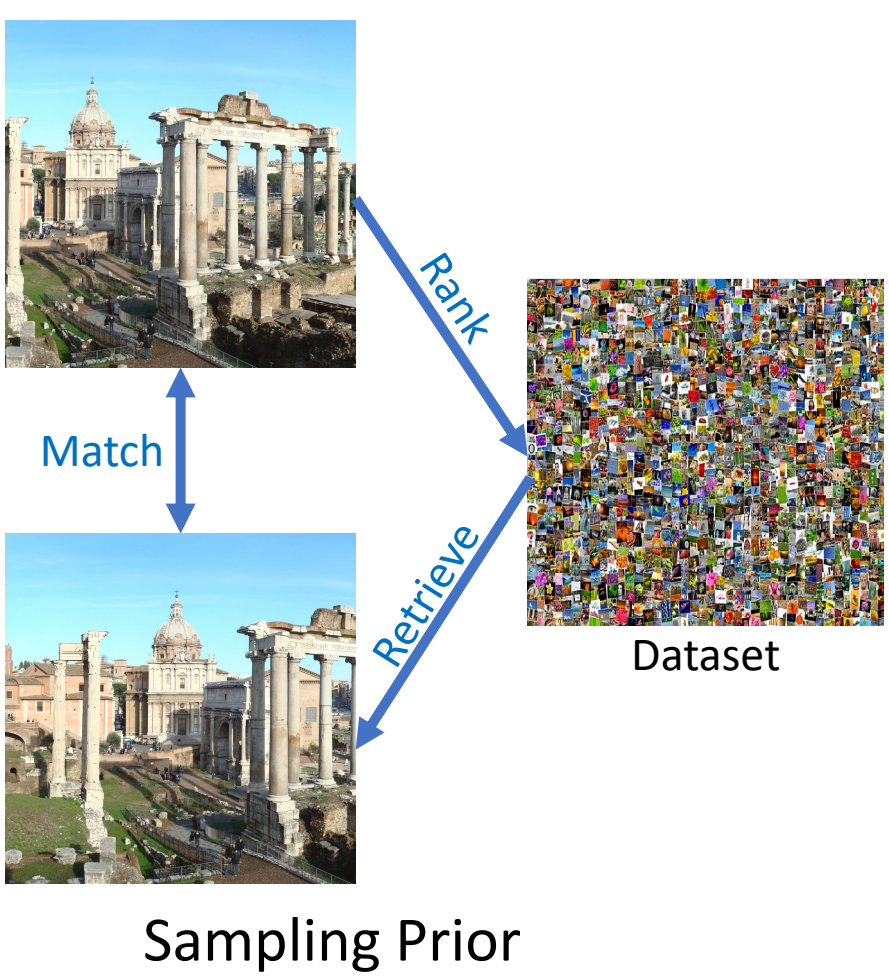
GraphMatch

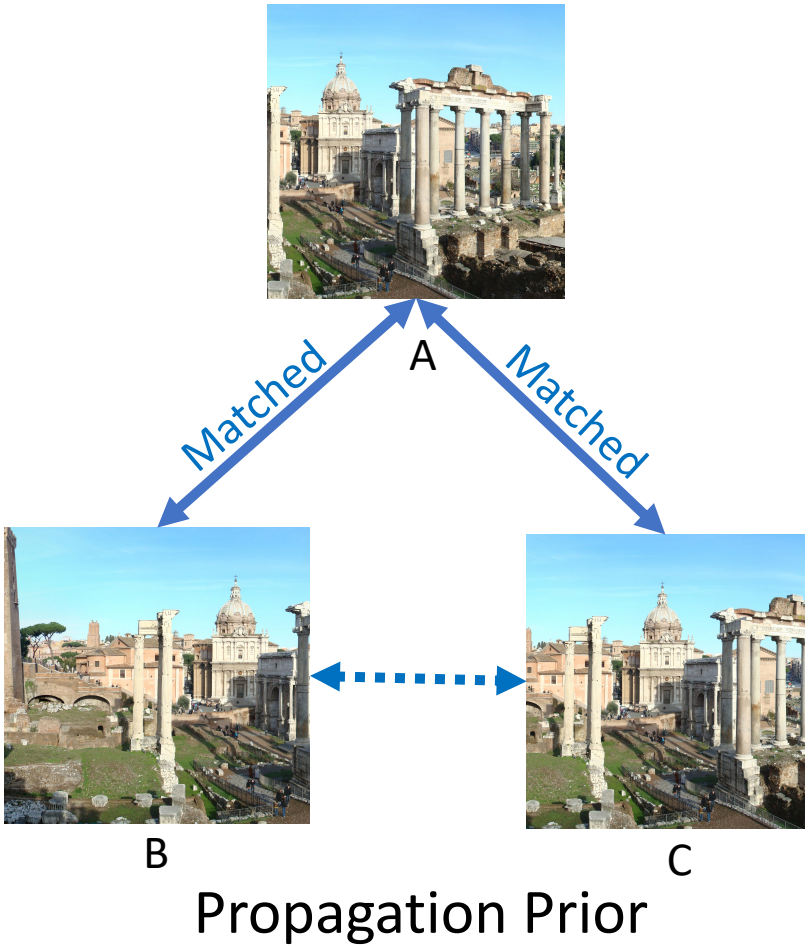
- Sampling step to identify new edges
- Propagation step to density matching graph
- Iterative “sample-and-propagate” scheme

GraphMatch

- Goal: Find most good matches as quickly as possible
- Good priors help predict potential matches
  - Sampling prior
  - Propagation prior

Sampling and Propagation priors

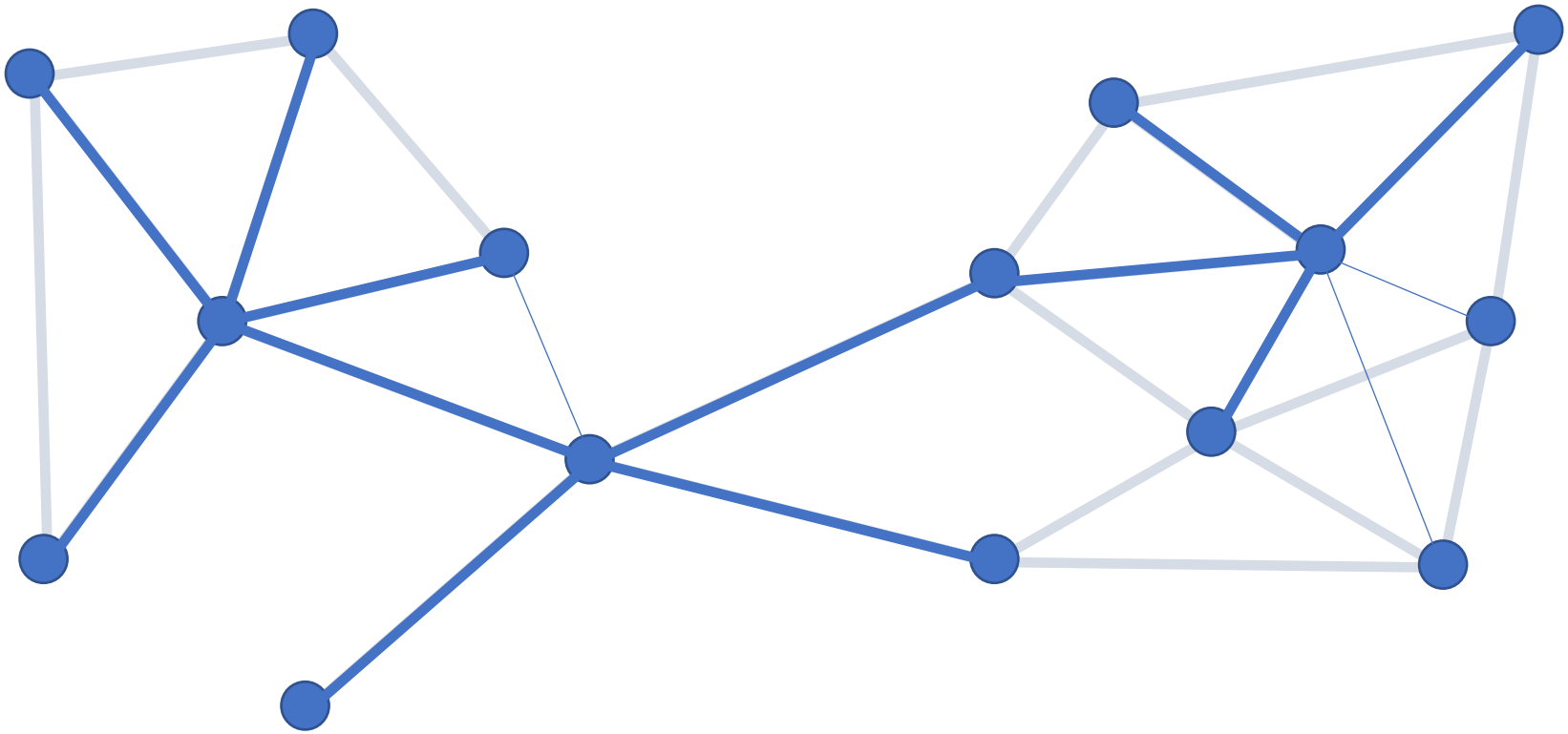




GraphMatch pipeline

1. Extract SIFT features and Fisher vector from images.
2. Compute fisher distance for all pairs of image.
3. While the algorithm has not converged do
    1. Sampling step
    2. Propagation step
4. Runs the reconstruction algorithm

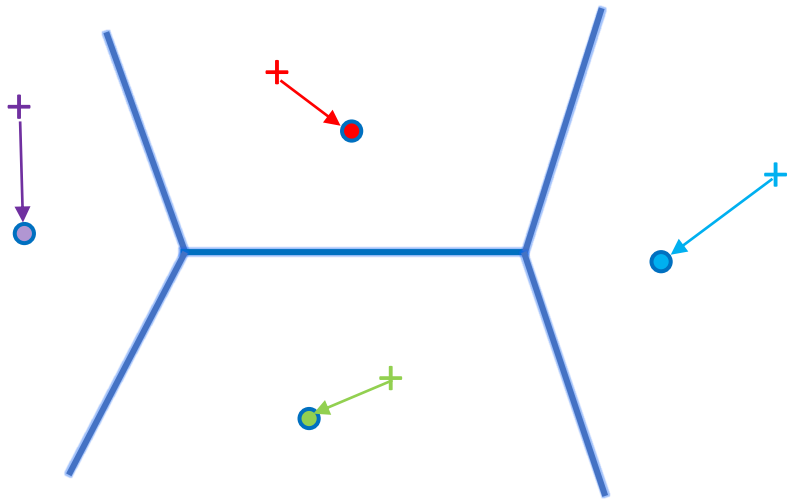
Sampling Step



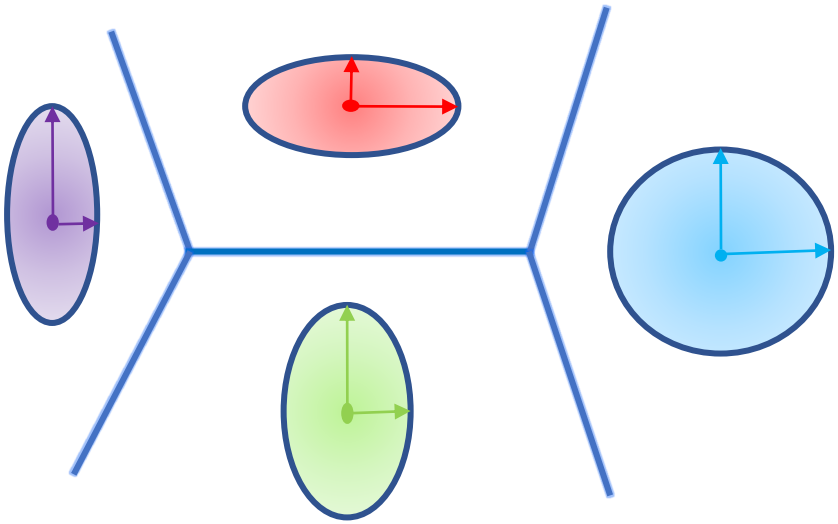
Sampling priors

- Vocabulary Tree
- VLAD vectors
- Fisher vectors

VLAD vectors

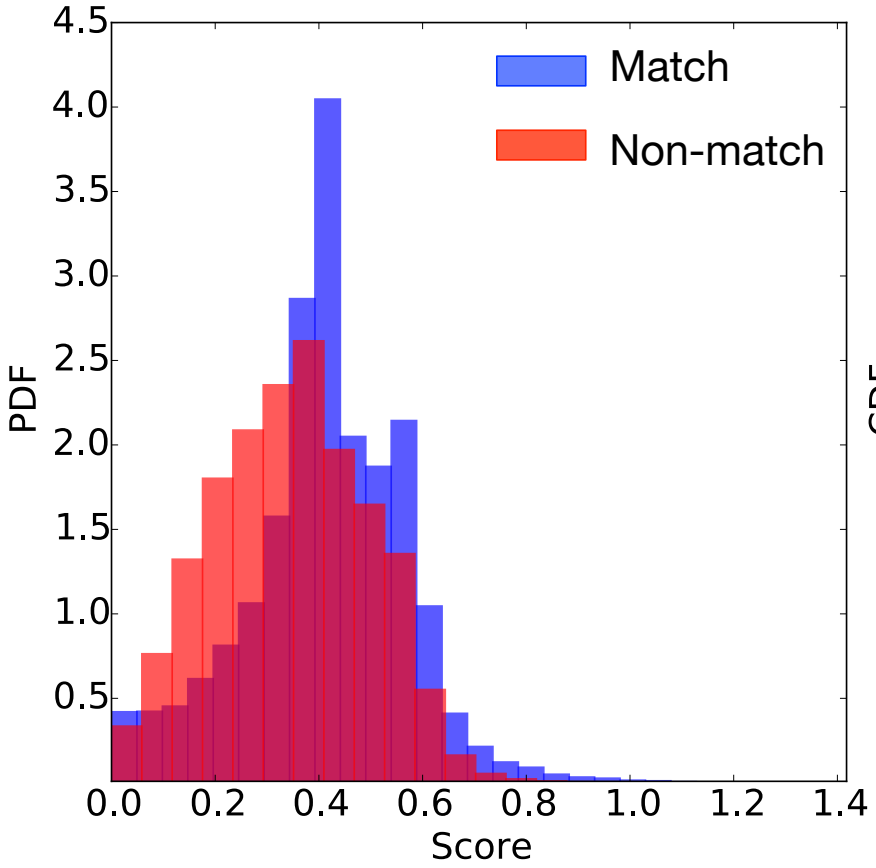


Fisher vectors

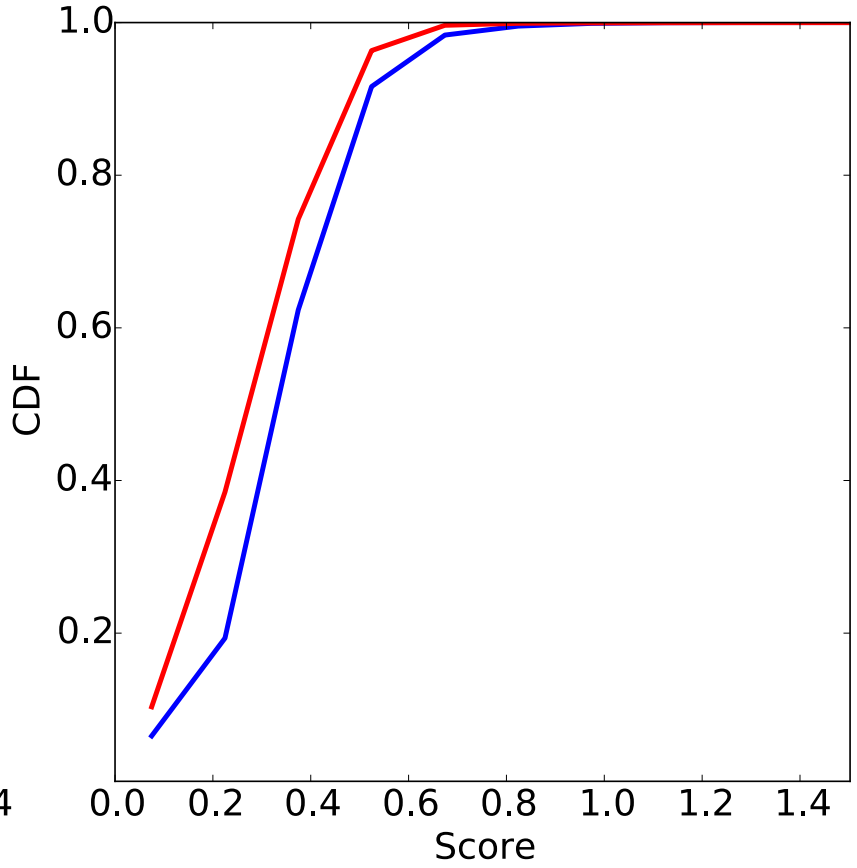


GraphMatch: Sampling

Voc. Tree scores are limited at predicting matching pairs.



(a) Vocabulary Trees Scores PDF

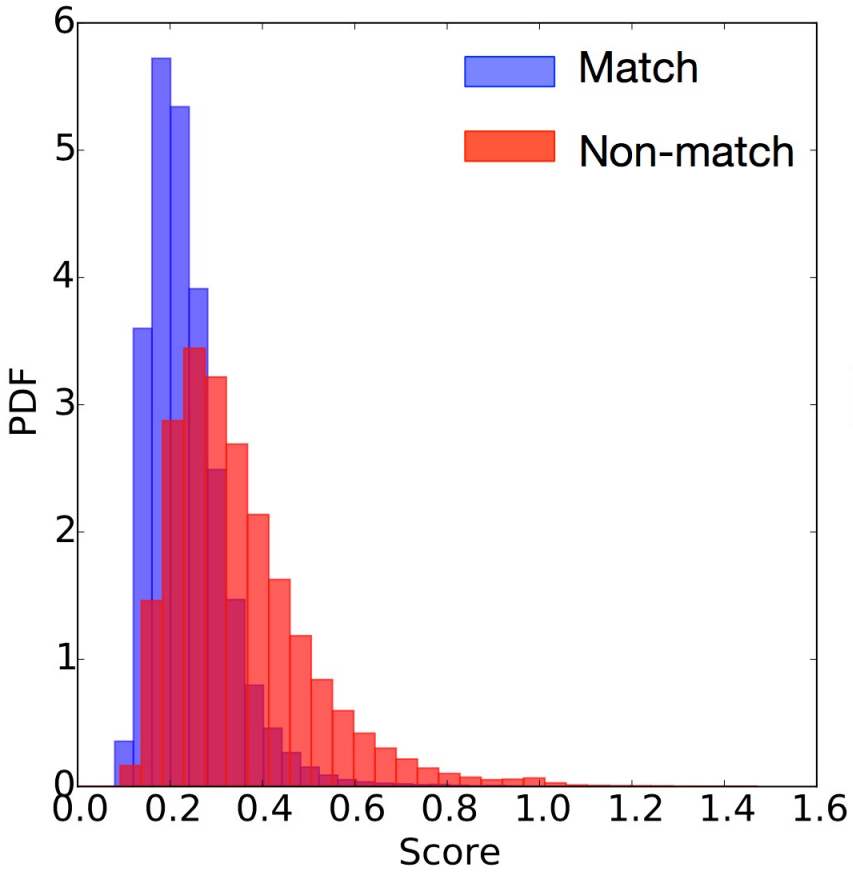


(b) Vocabulary Trees Scores CDF

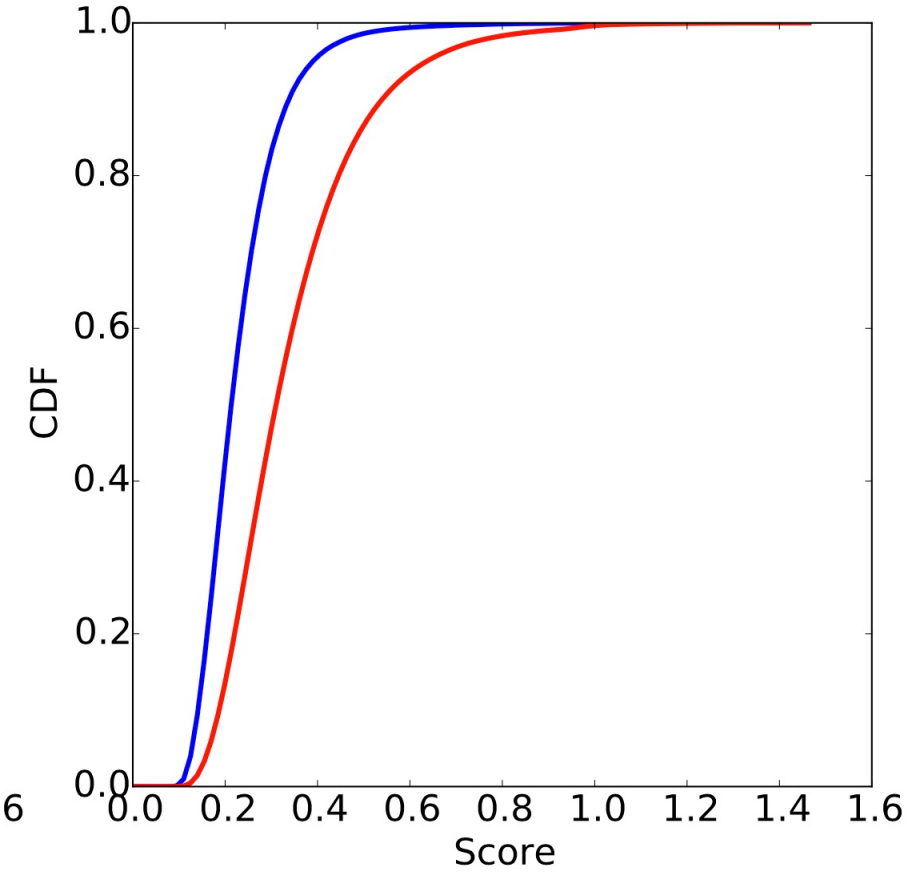
UCSB WVU

GraphMatch: Sampling

VLAD vectors are better



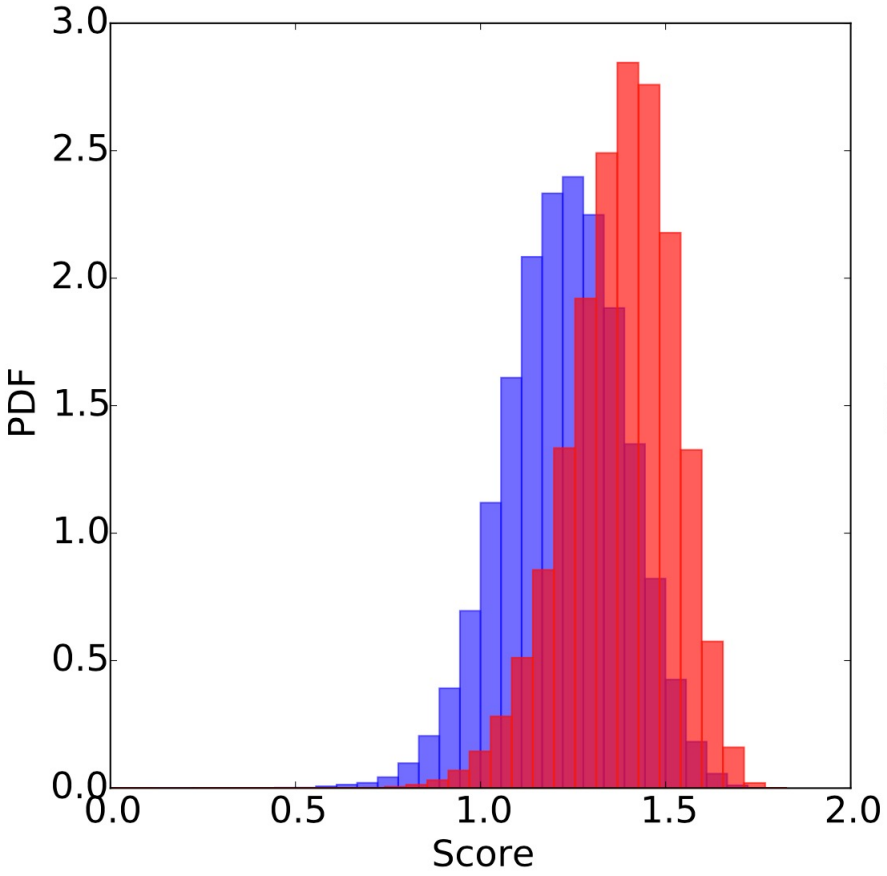
(a) VLAD PDF



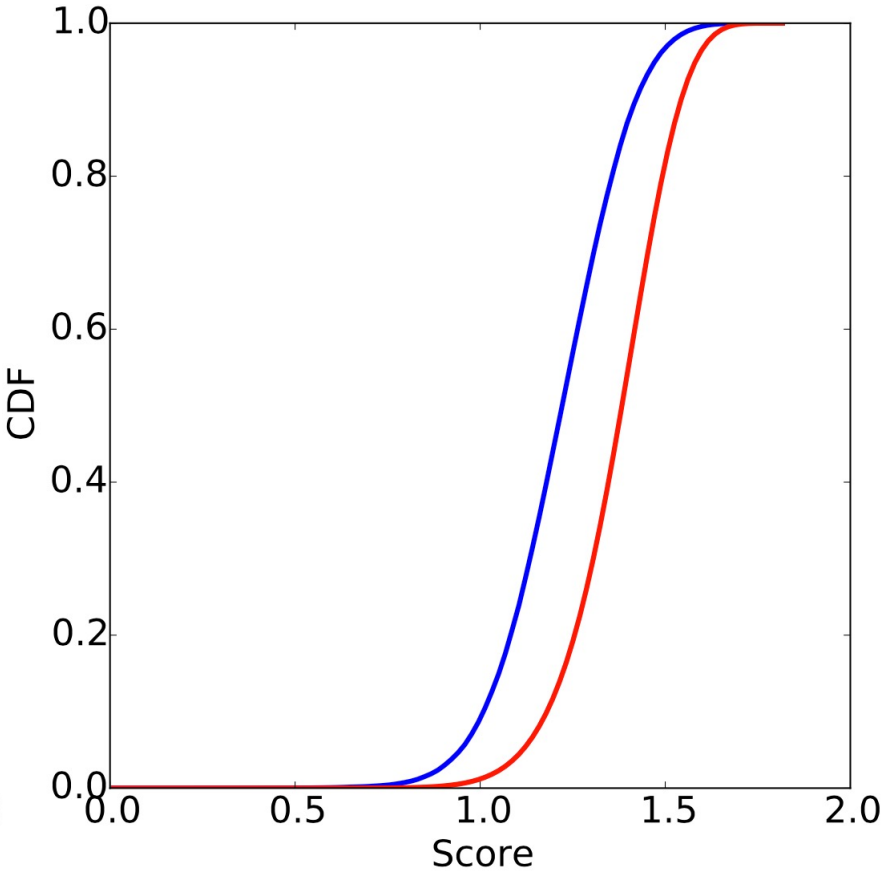
(b) VLAD CDF

GraphMatch: Sampling

Fisher vectors are best

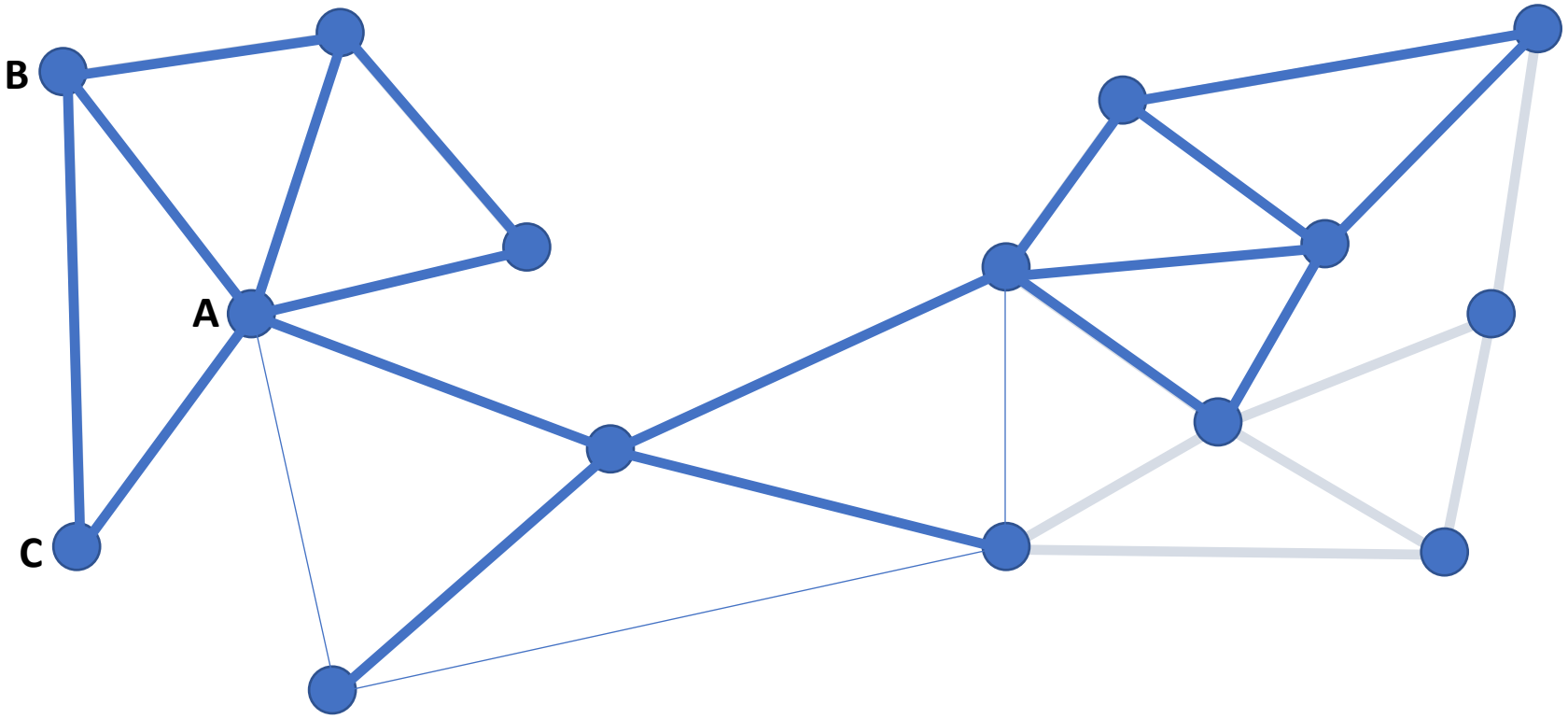


(c) Fisher PDF

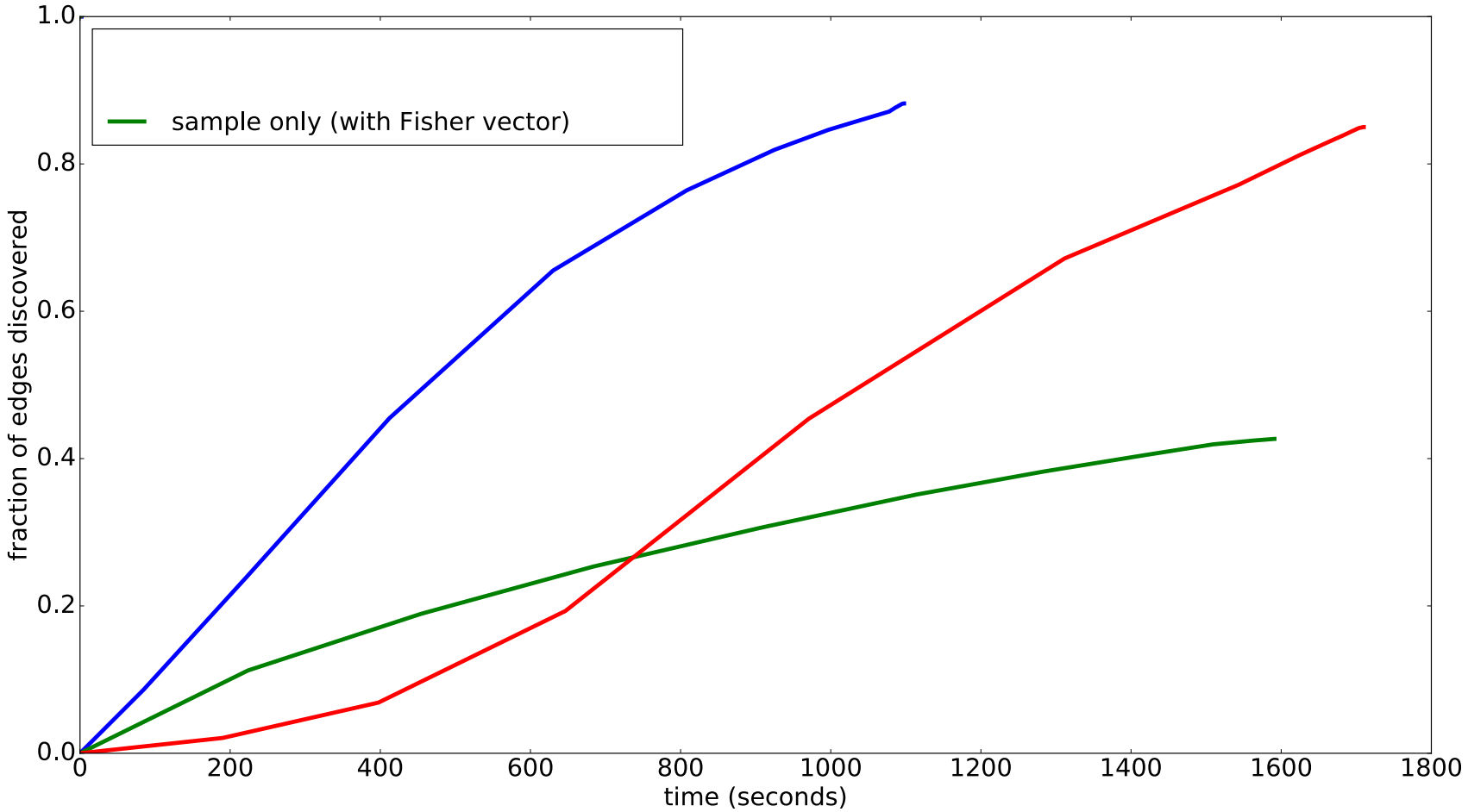


(d) Fisher CDF

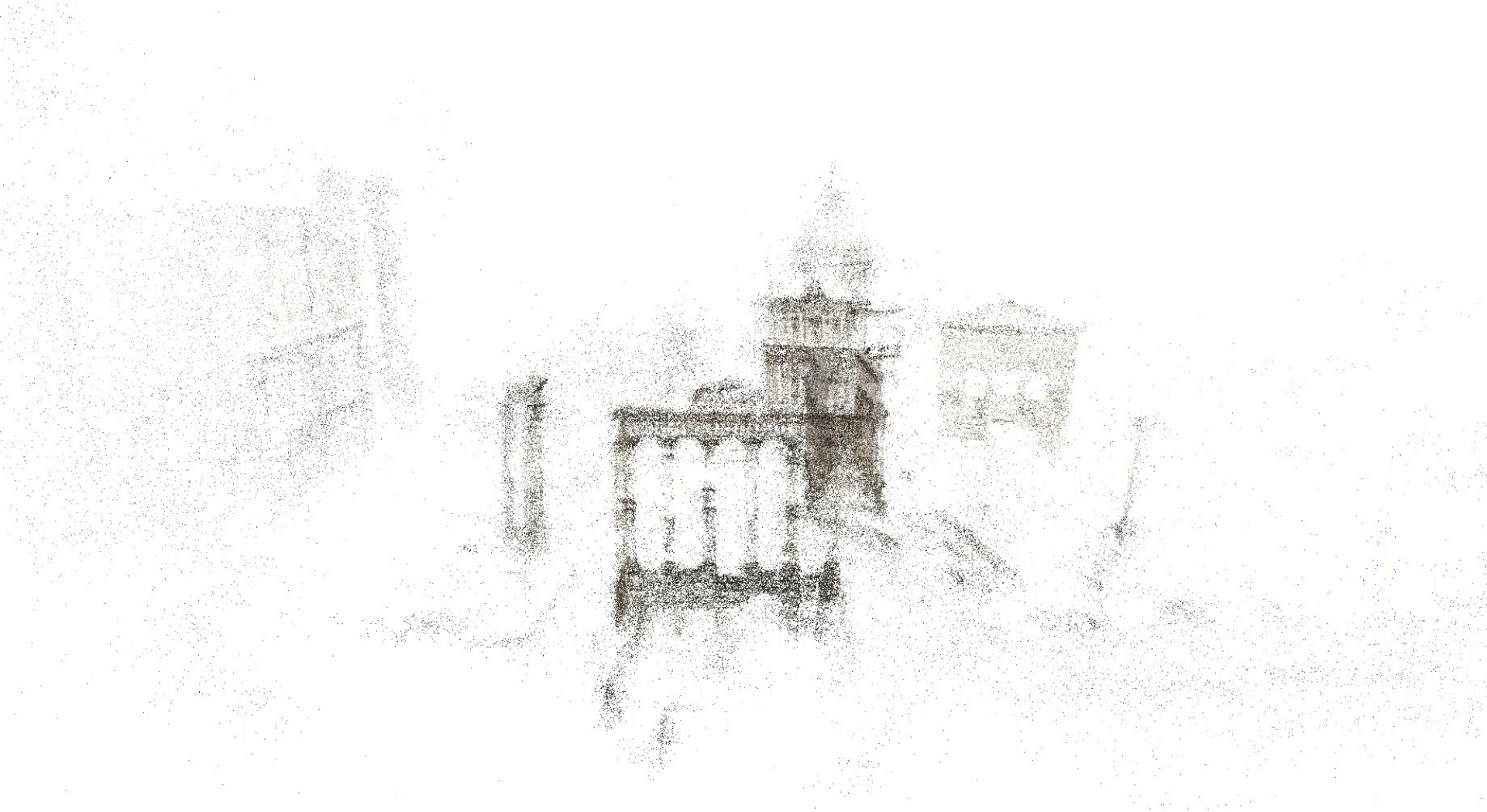
Propagation Priors



Interleaving

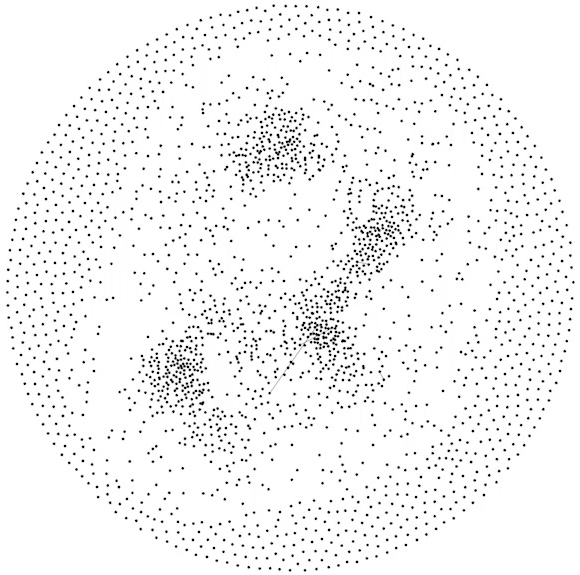


Reconstruction of Roman Forum Scene

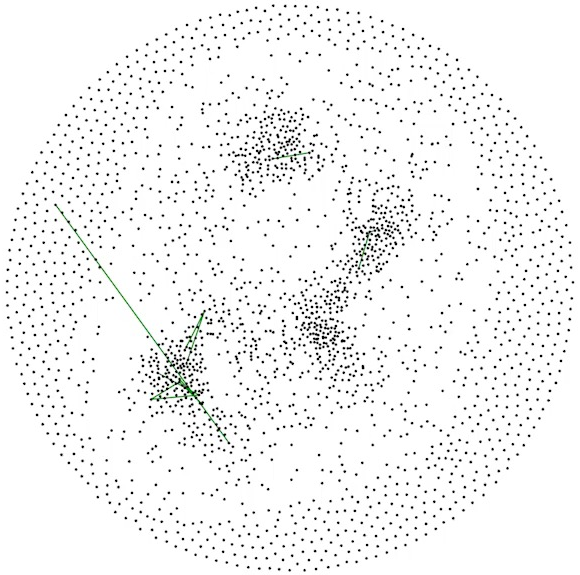


Matching graph for "Roman Forum" Scene

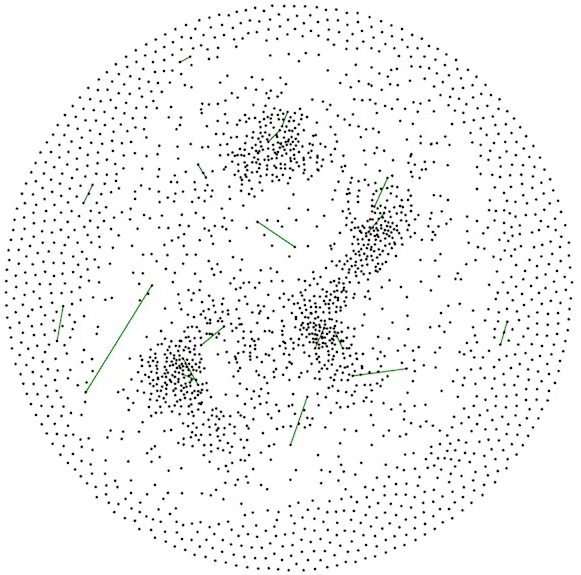
Voc. Trees



BRIAD [Agarwal et al. 2009]



Our approach



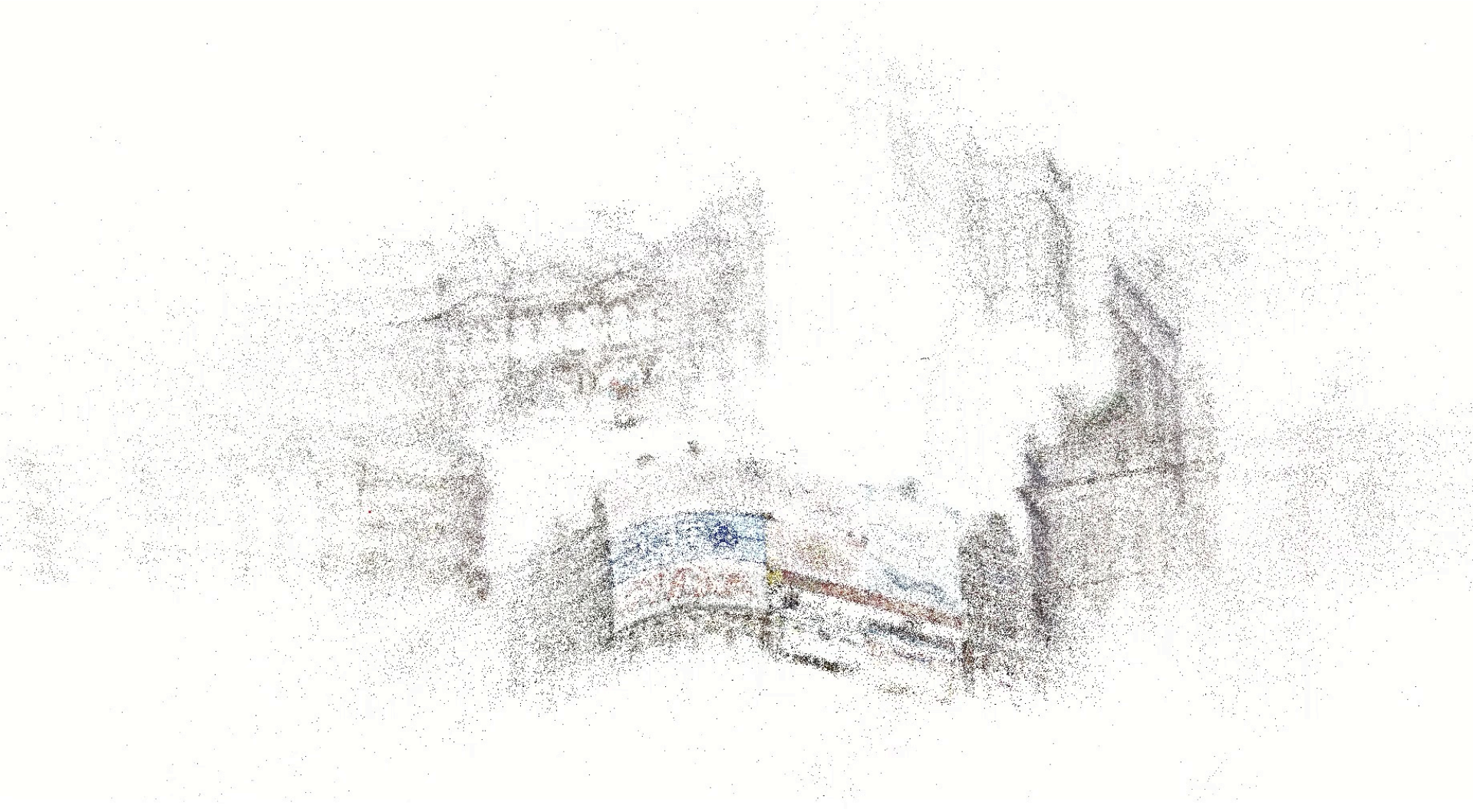
Iteration: 1

green: sampling edges
red: propagation edges

Vienna Cathedral

Baseline: 134.3 hrs (est.)
GraphMatch: 6.7 hrs

Piccadilly

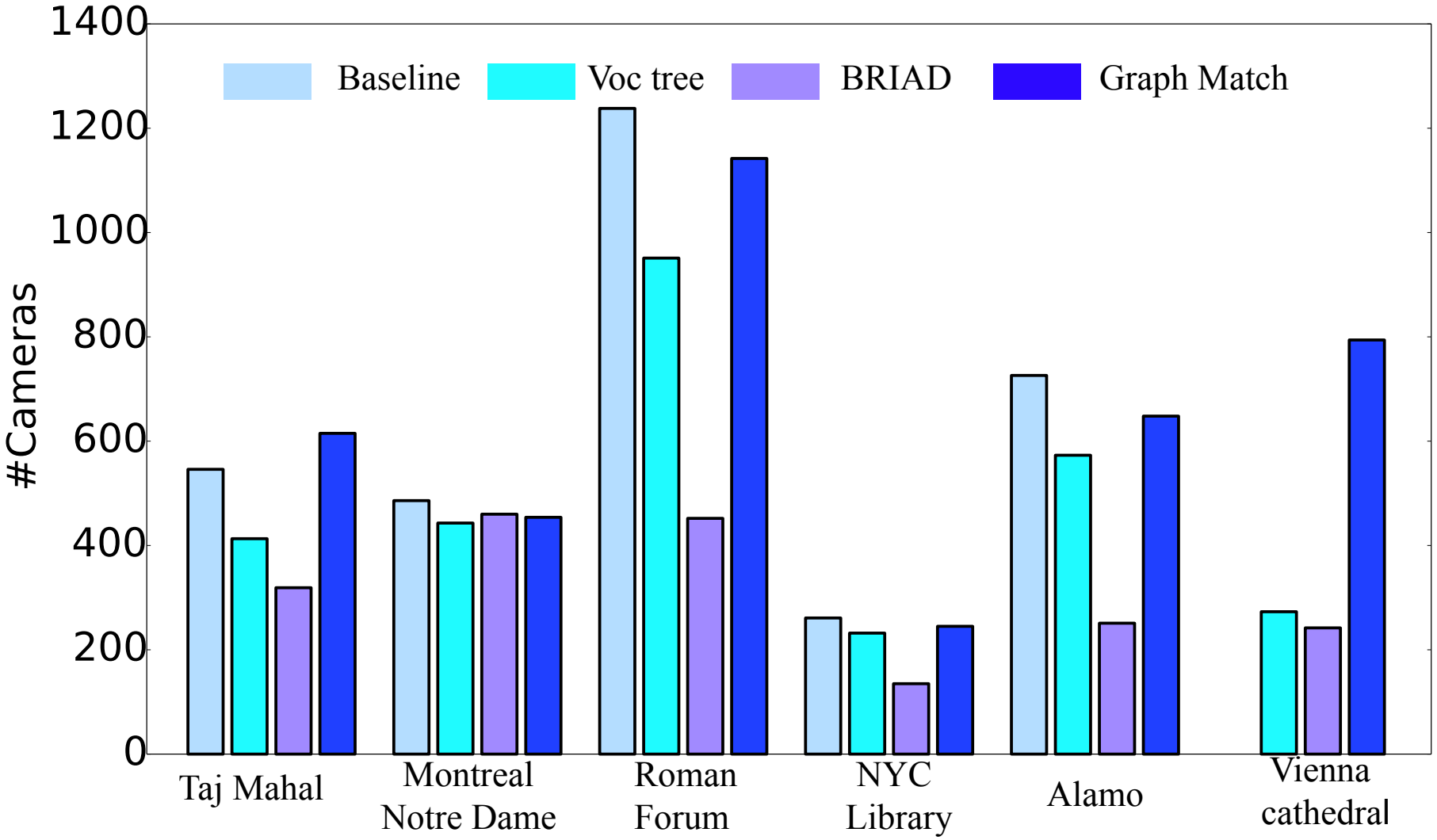


Baseline: 163.6 hrs (est.)
GraphMatch: 9.7 hrs

Trafalgar Square

Baseline: 835.4 hrs (est.)
GraphMatch: 16.9 hrs

Cameras reconstructed for different datasets



UCSB

Timing

DataSet	# recon cameras	Preprocess	Match	Recon	Total
VIENNA CATHEDRAL	794	3.37 min	367.58 min	44.28 min	450.60 min
PICADILLY	1863	3.61 min	512.53 min	98.34 min	655.38 min
Trafalgar	4057	7.93 min	1014.37 min	292.54 min	1380.68 min

Contributions

- Extension of PatchMatch to image matching.
  - “Sample-and-Propagate” Strategy
- Better priors (fisher vector)
- Efficiently finding more image matches
  - Achieving more reconstructed cameras
  - Maintaining equivalent speed-ups with Voc. tree

Limitations and Future Work

- Optimal parameter tuning.
- Image representation for sampling stage
  - PAIGE

Acknowledgements

- NSF grants IIS-1657179, IIS-1342931, and IIS-1321168
- Atieh Taheri for preliminary experiments.

Thank you