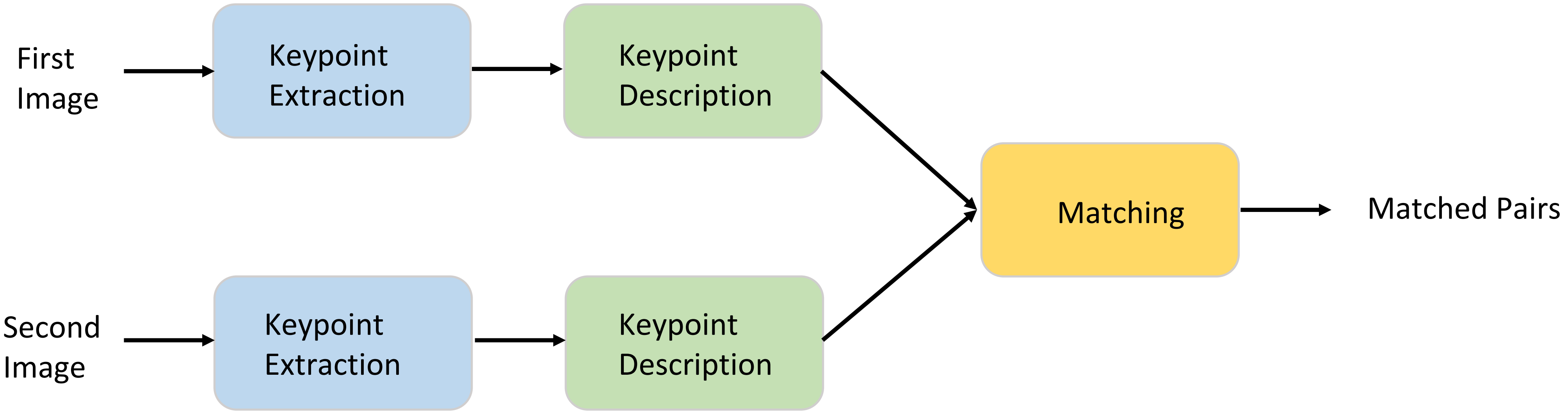
MULTI-SCALE KEYPOINT MATCHING

Sina Lotfian, Hassan Foroosh
University of Central Florida
Computer Science Department

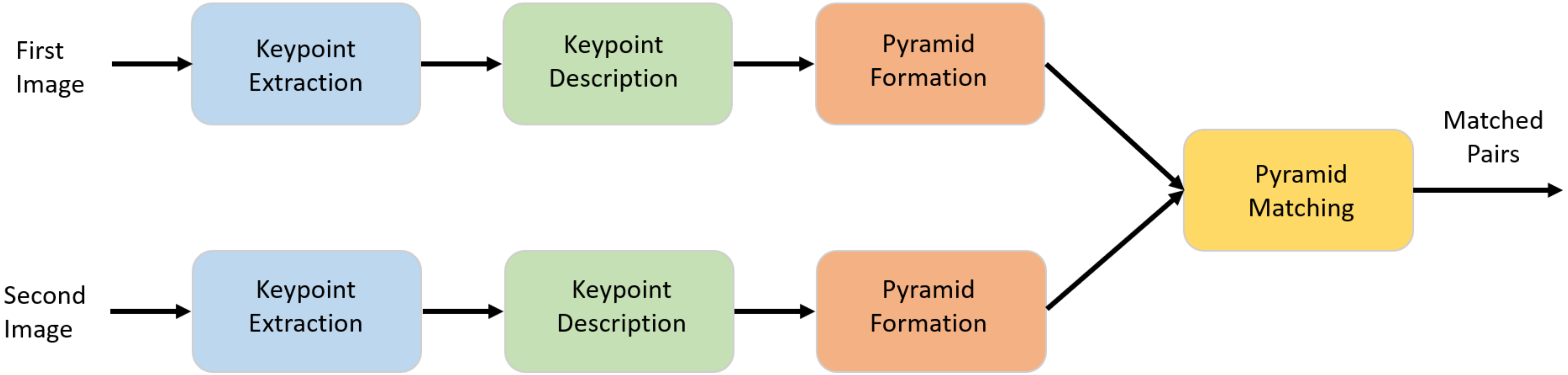
Computational Imaging Lab. (CIL)
Department of Computer Science

Single-Scale Matching



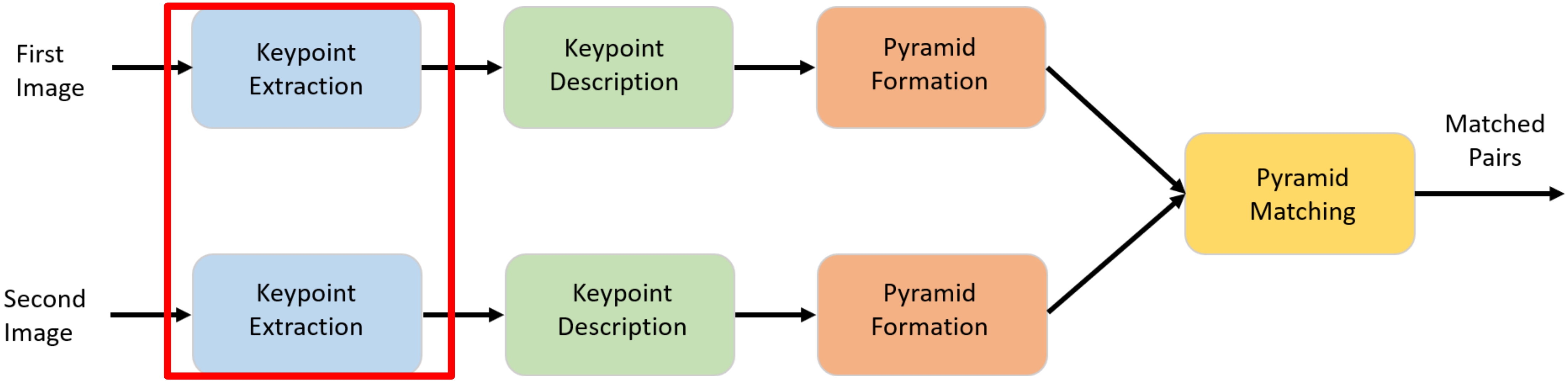
- There is not enough information for matching in a single scale.
- Humans use clues from different scales when matching points.
- How can we use data from multiple scales with least computational penalty?

Multi-Scale Matching



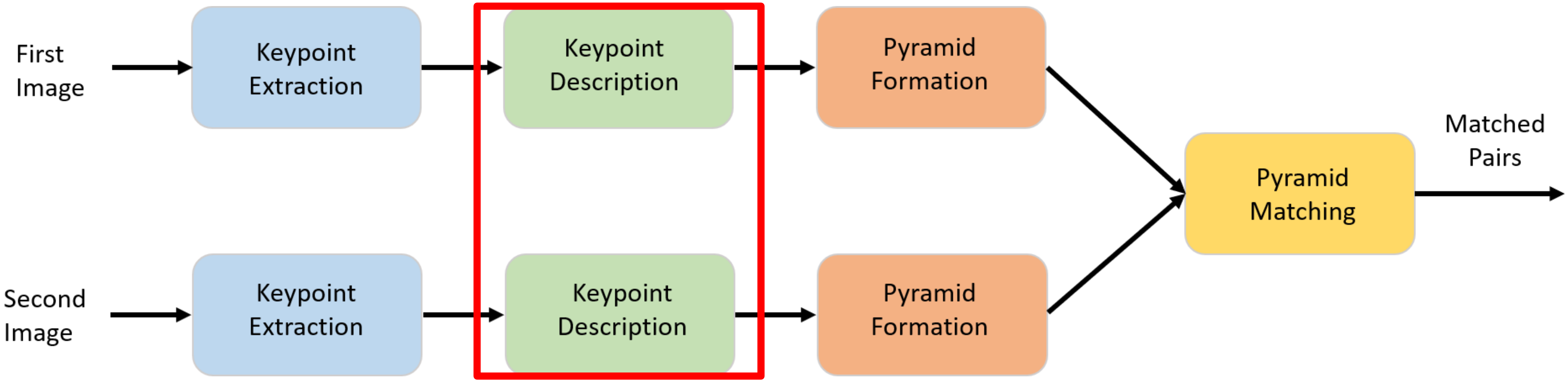
- Exploiting data from high and low scales
- Biologically plausible
- High scales contain Holistic information
- Lower Scale contain More localized data
- Iterative pruning

Multi-Scale Matching Pipeline



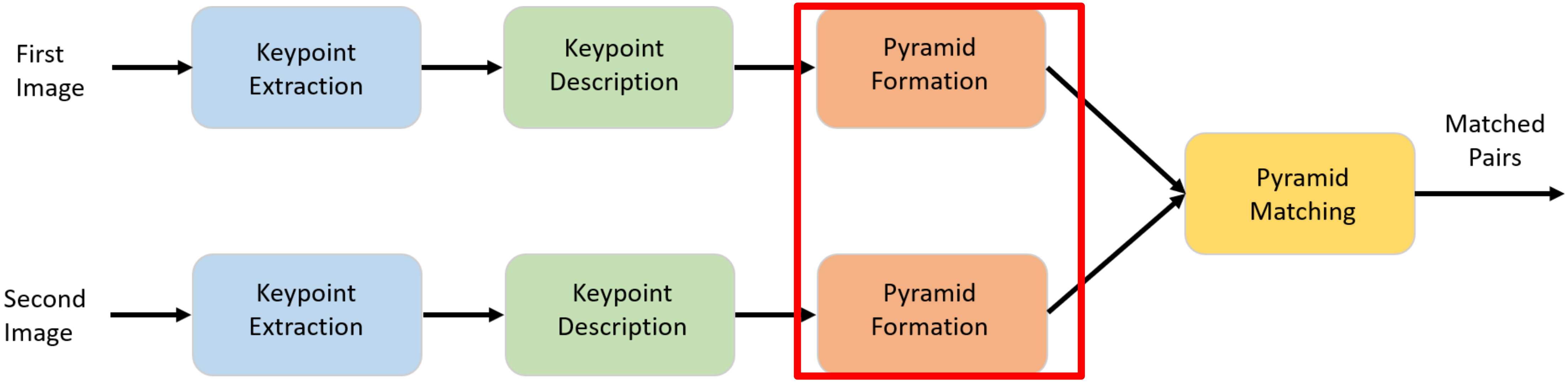
- Keypoints are extracted using Difference of Gaussian(DoG).

Multi-Scale Matching Pipeline



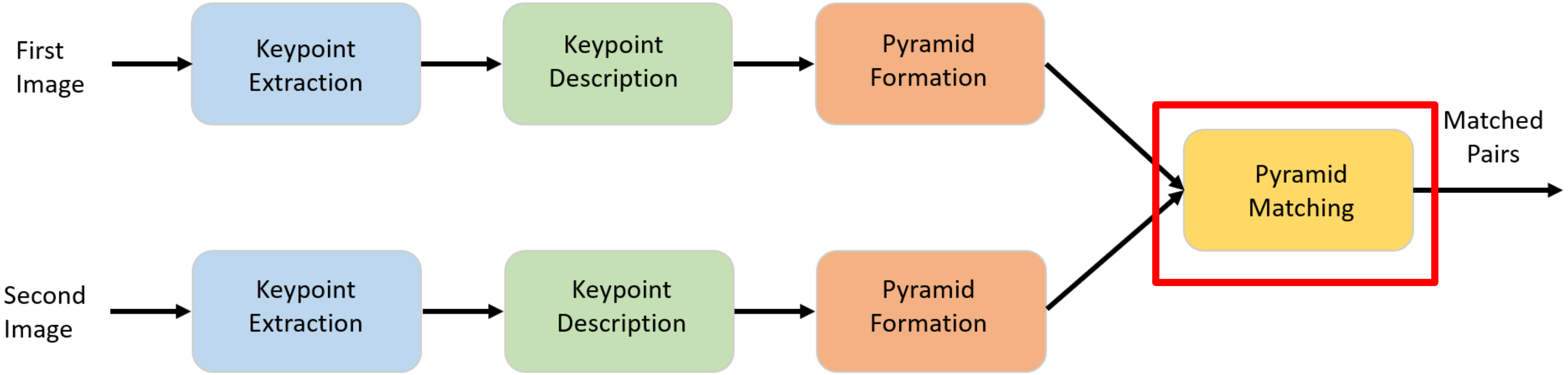
- Each keypoint is mapped into a feature space
- Descriptors can be generated from any method
  - Hand-Engineered features
  - Learned features
- This is done for N scales for each keypoint

Multi-Scale Matching Pipeline



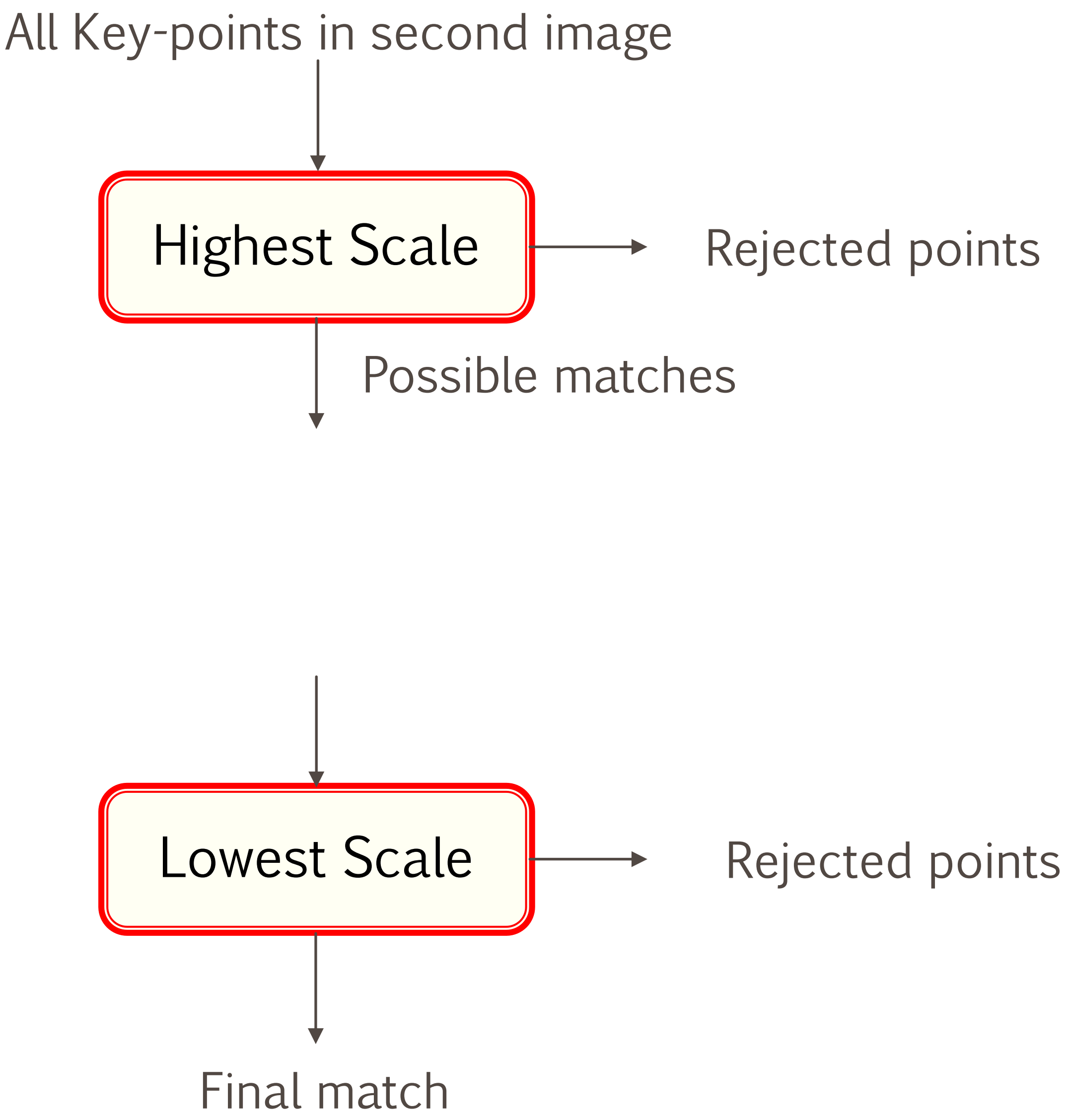
- Different scales are concatenated to form a pyramid
- Higher Scales are on Top → Lower Resolution
- Lower Scales Bottom → Higher Resolution

Multi-Scale Matching Pipeline



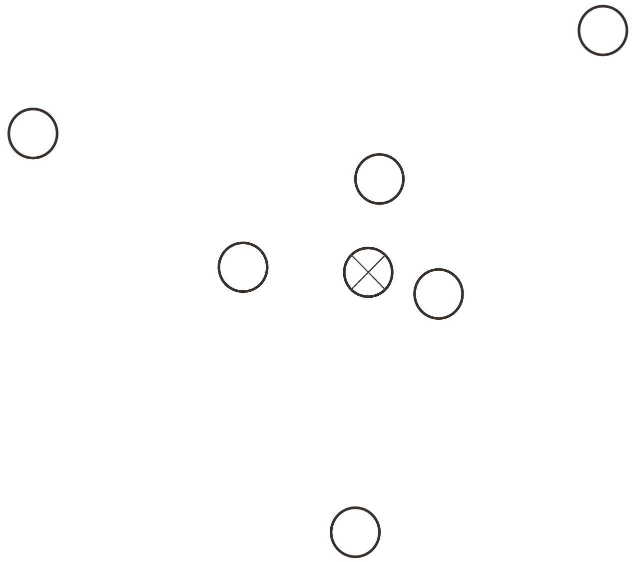
Top-Bottom Matching

- The goal is to choose the best match among points in the second image for a point in the first image.
- Start by considering all points in the second image.
- Use higher scales for early rejection
- Repeat until one point is remaining



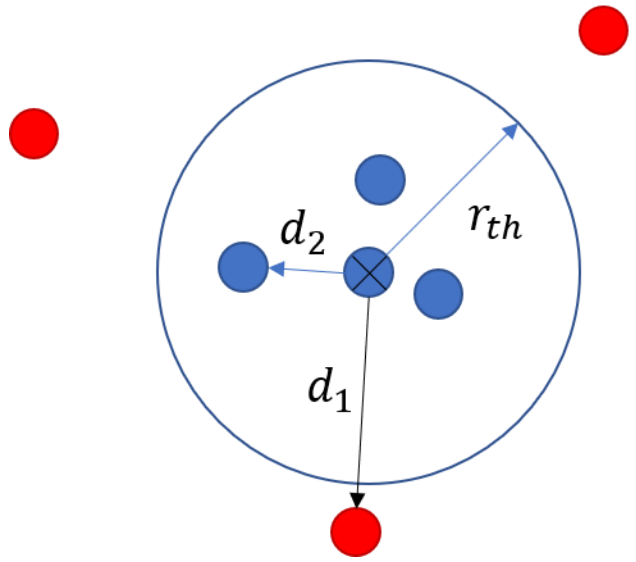
Maximum Margin Nearest Neighbor

- Each point has several possible match
- Which point should be rejected?



Maximum Margin Nearest Neighbor

- Threshold on distance
- Ant point inside the circle is a possible match
- Any point outside the circle is rejected
- Margin of confidence : Difference between distance of nearest rejected point d_1 and most distance accepted point d_2



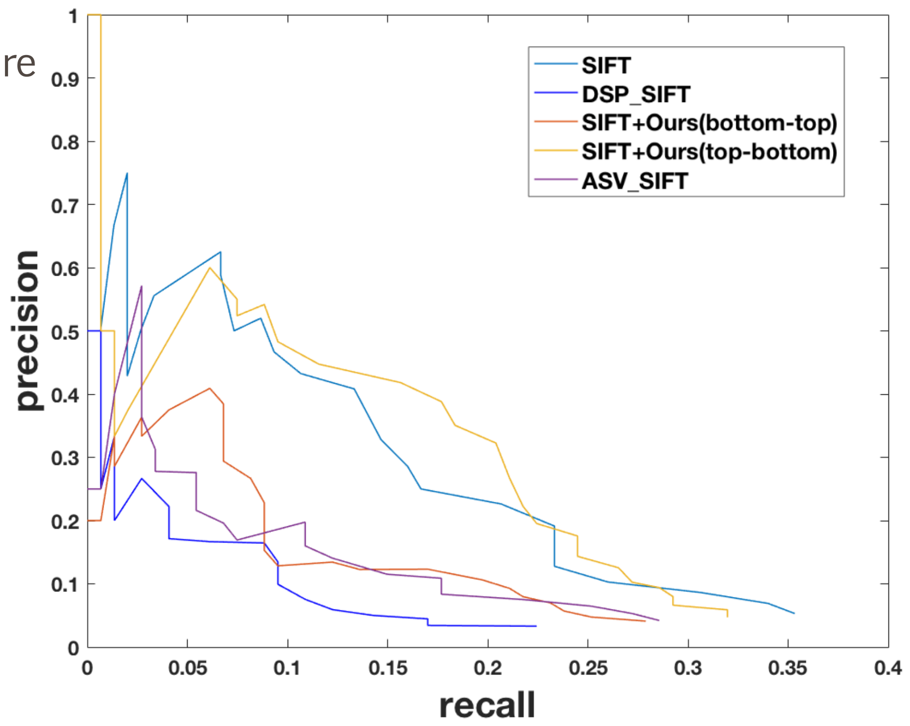
Results

- Mean average Precision

- We choose SIFT[1] as hand-crafted and SOSNet[2] as Learned feature
- ASV-SIFT and DSP-SIFT as competing multi-scale approaches

MAP for different datasets

Method	MAP		
	Oxford	Webcam	HPatches
SIFT [1]	58.82	16.29	42.28
SOSNet [2]	63.54	18.94	46.39
Root SIFT [3]	60.11	18.29	44.05
Raw Patch	30.67	3.97	21.13
LIOP [4]	40.54	1.79	33.87
DSP-SIFT [5]	60.43	22.28	45.17
ASV-SIFT [6]	60.94	23.16	45.53
SIFT + ours(top-bottom)	61.02	25.64	46.87
SIFT + ours(bottom-top)	60.41	25.10	47.04
SOSNet [2] + ours(top-bottom)	68.03	27.05	**51.31**
SOSNet [2] + ours(bottom-top)	**68.82**	**27.37**	50.63



Precision-Recall plot for Oxford dataset (View-Point change)

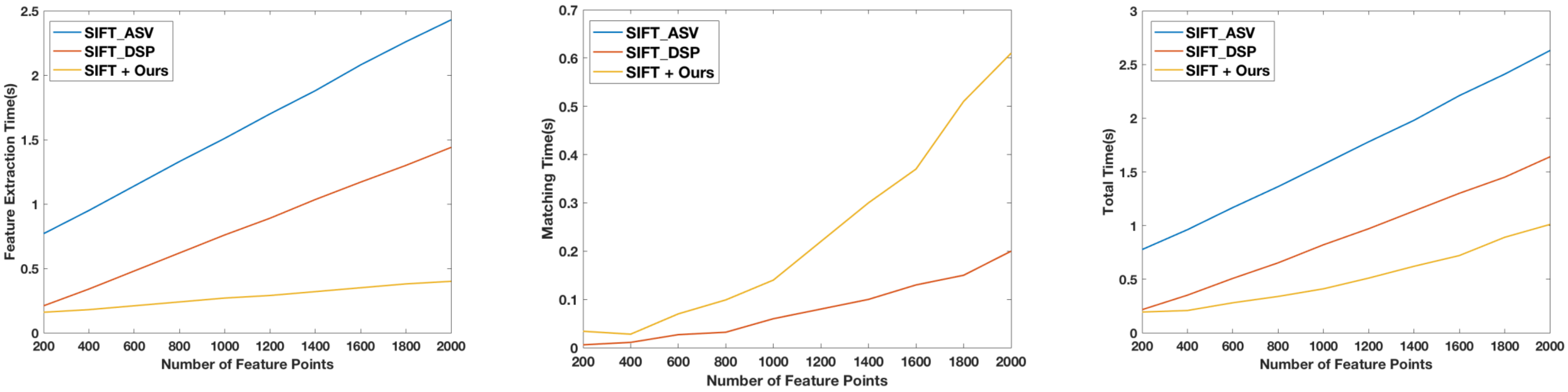
Result

1. Feature extraction time:
   - Grows linearly with number of keypoints
   - Grow linearly with number of sampled scales

2. Matching time :
   - Grows quadratically with number of keypoints
   - For the proposed method increases linearly with number of sampled scales

Results

- Time Analysis:



References

- [1] Lowe, David G. ”Distinctive image features from scale-invariant keypoints.” International journal of computer vision 60.2 (2004): 91-110.
- [2]Tian, Yurun, et al. ”SOSNet: Second order similarity regularization for local descriptor learning.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.
- [3] Arandjelovi´c, Relja, and Andrew Zisserman. ”Three things everyone should know to improve object retrieval.” 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012.
- [4] Miksik, Ondrej, and Krystian Mikolajczyk. ”Evaluation of local detectors and descriptors for fast feature matching.” Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). IEEE, 2012.
- [5] Dong, Jingming, and Stefano Soatto. ”Domain-size pooling in local descriptors: DSP-SIFT.” Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
- [6] Yang, Tsun-Yi, Yen-Yu Lin, and Yung-Yu Chuang. ”Accumulated stability voting: A robust descriptor from descriptors of multiple scales.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016