Modified Maximum Spanning Tree Clustering for Large-Scale Image Retrieval

Cheng Chang, Chundi Liu, Guangwei Yu, Maksims Volkovs
Layer 6 AI

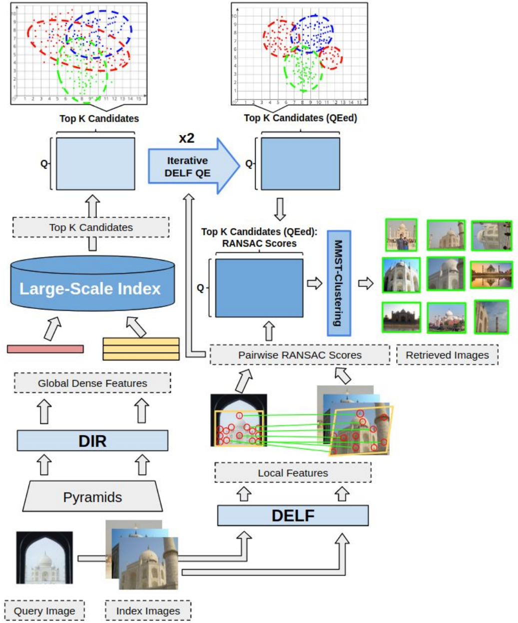
LAYER 6 AI

Motivation

Challenges:

- Retrieve all the images depicting the same landmark **regardless of visual similarity**.
- Ranking is necessary for the MAP metric.

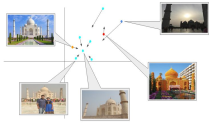
Architecture



Iterative DELF Query Expansion

Images with the same landmarks move closer to the approximated centers of the landmark clusters iteratively.

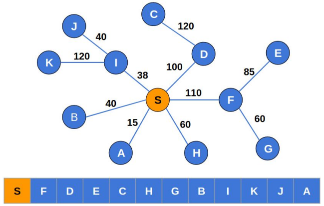
- Apply geometric verification using local descriptors.
- Test images can be used as additional data.

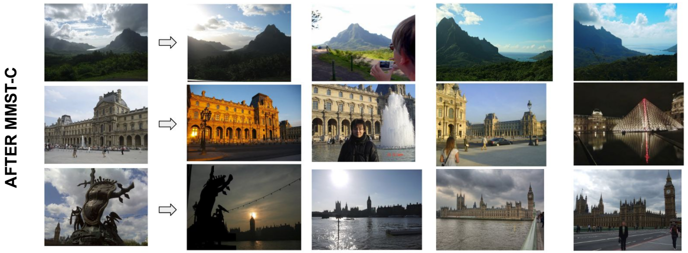


Iter*	VMAP*	VMD*
0	66.43%	1.02
1	76.68%	0.60
2	77.89%	0.45
3	75.84%	0.39
4	74.28%	0.36

Modified Maximum Spanning Tree Clustering

- As long as there is a bridging image, visually dissimilar images can be connected.
- Ranking is accomplished since we are adding the most confident image at a time.





Results

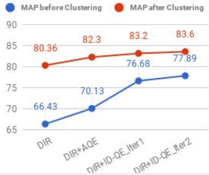
Google Landmark Retrieval Challenge:

- Largest public dataset for image retrieval.
- 15K unique landmarks, 1M training images, 1M index images, 100K test images.
- Images have various sizes and high resolutions; 329GB in total.

Highlights:

- No fine-tuning.
- Single model without ensemble.

Validation:



Experiment Results:

Method	PubMAP
DIR	42.3%
DIR + AQE	47.9%
DIR + ID-QE	55.7%
DIR + ID-QE + MMST-C*	62.7%

Competition Results:

Team	PvtMAP
1. CVSSP & Visual Atoms	62.7%
2. Layer 6 AI	**60.8%**
3. SevenSpace	59.8%
4. Naver Labs Europe	58.6%
5. VPP	58.3%

Conclusion

- **Iterative DELF QE** constrains the global feature space.
- **Clustering** resolves the challenge of visually dissimilar images through transition of bridging images.
- **Modified maximum spanning tree** clustering ranks the candidates across the connected images.

- **Scalability**: fast approximate update for both new images in index and in test.
- **Flexibility**: limit depth to constraint visual similarity.

References

[1] Noh H et al. Large-Scale Image Retrieval with Attentive Deep Local Features. Proc. ICCV 2017.
[2] Gordo A et al. Deep Image Retrieval: Learning Global Representations for Image Search. ECCV 2016.
[3] Gordo A et al. End-to-end Learning of Deep Visual Representations for Image Retrieval. IJCV 2017.