Image Retrieval 2.0

Filip Radenović

Center for Machine Perception
Czech Technical University in Prague

Outline

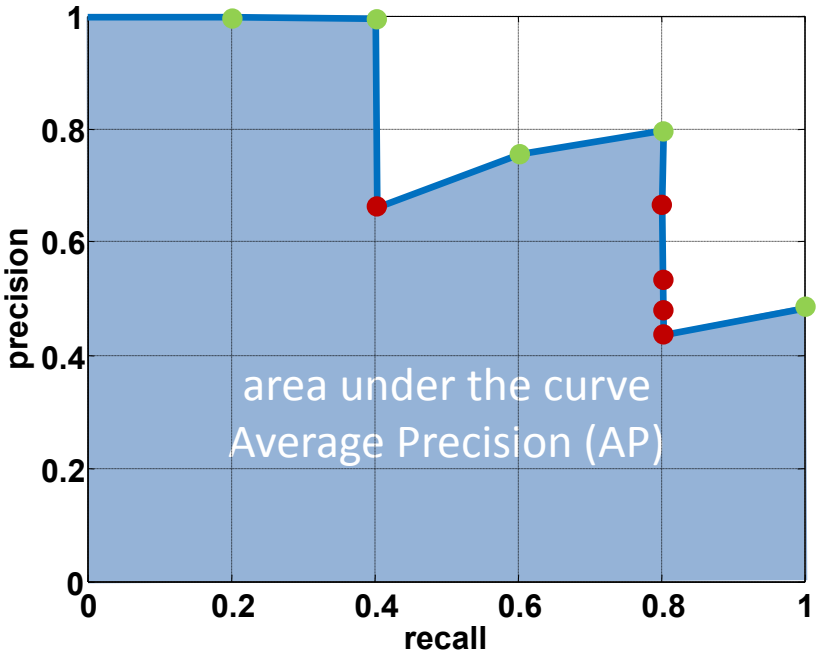
- 1.0: Standard image retrieval problems
  - Visually most similar
  - All visually similar
- 2.0: Beyond similarity retrieval
  - New (unseen) information
  - What/where is this?
  - What is interesting here?
  - Where should I look?
- 2.1: Image retrieval for 3D reconstruction

Standard Image Retrieval Evaluation

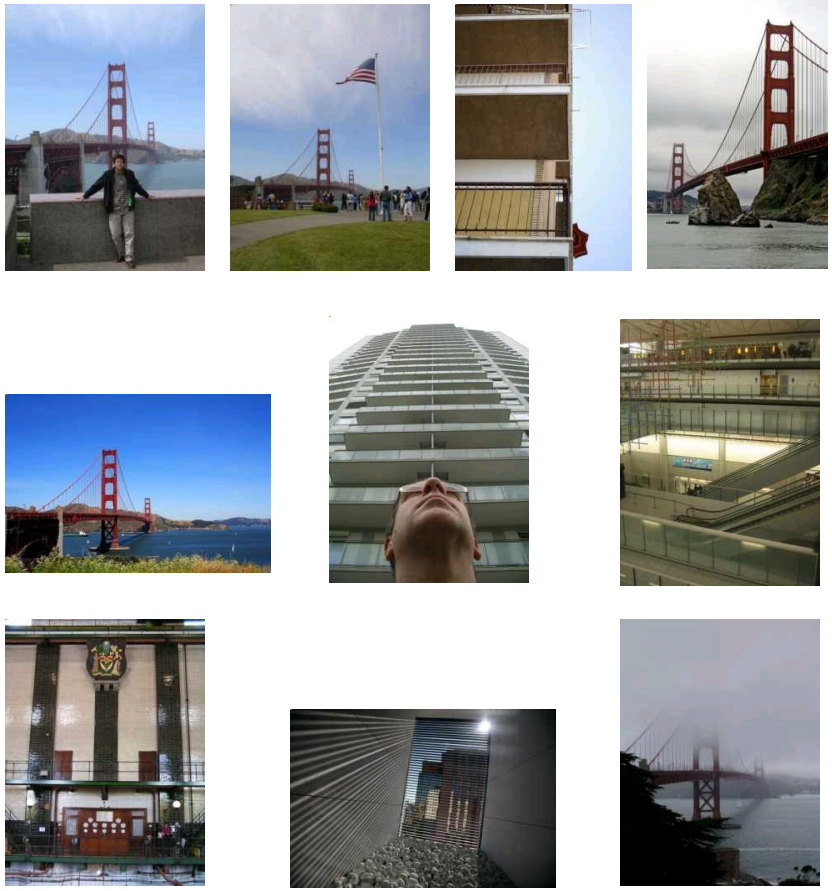
Query

Database size: 10 images
Relevant (total): 5 images

precision = #relevant / #returned
recall = #relevant / #total relevant



Results (ordered):

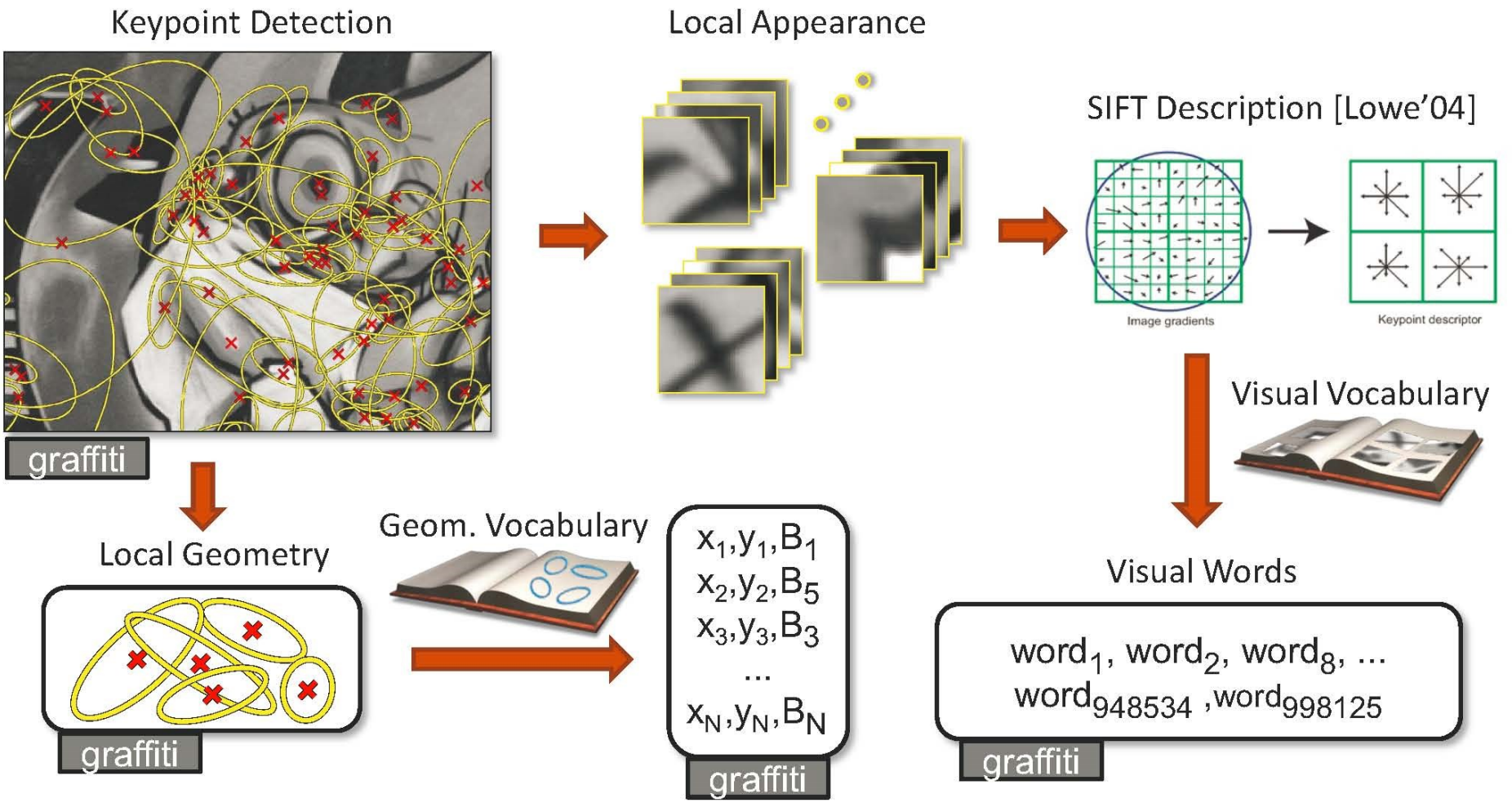


Is this what we want?

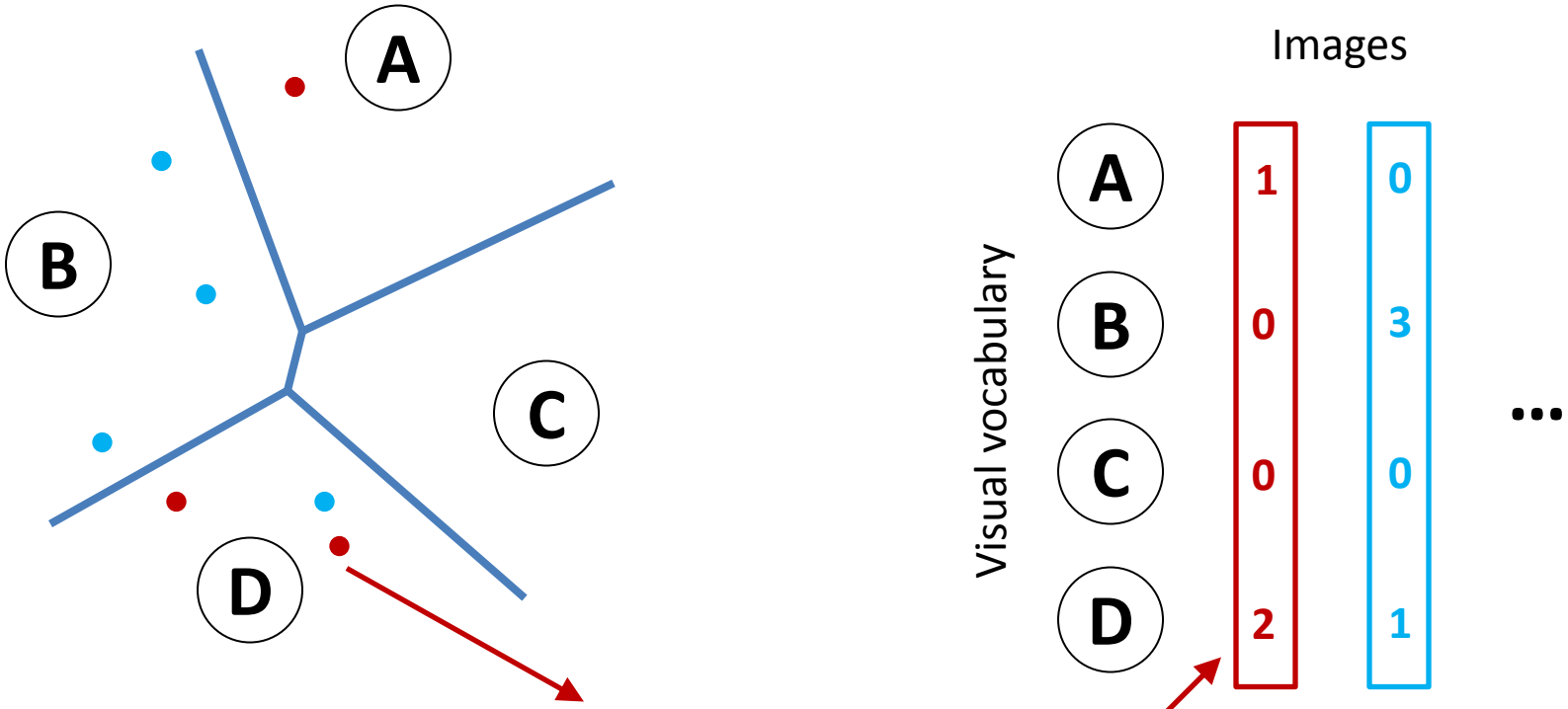
- Visually most similar
  - Results identical to query for large datasets
- All visually similar
  - Output of varying length
  - Ground truth hard to obtain
  - Users will never take a look at more than few tens of near-duplicate images!!!

1.0: Bag of Words (BoW) Image Retrieval

Bag of Words: Off-line Stage



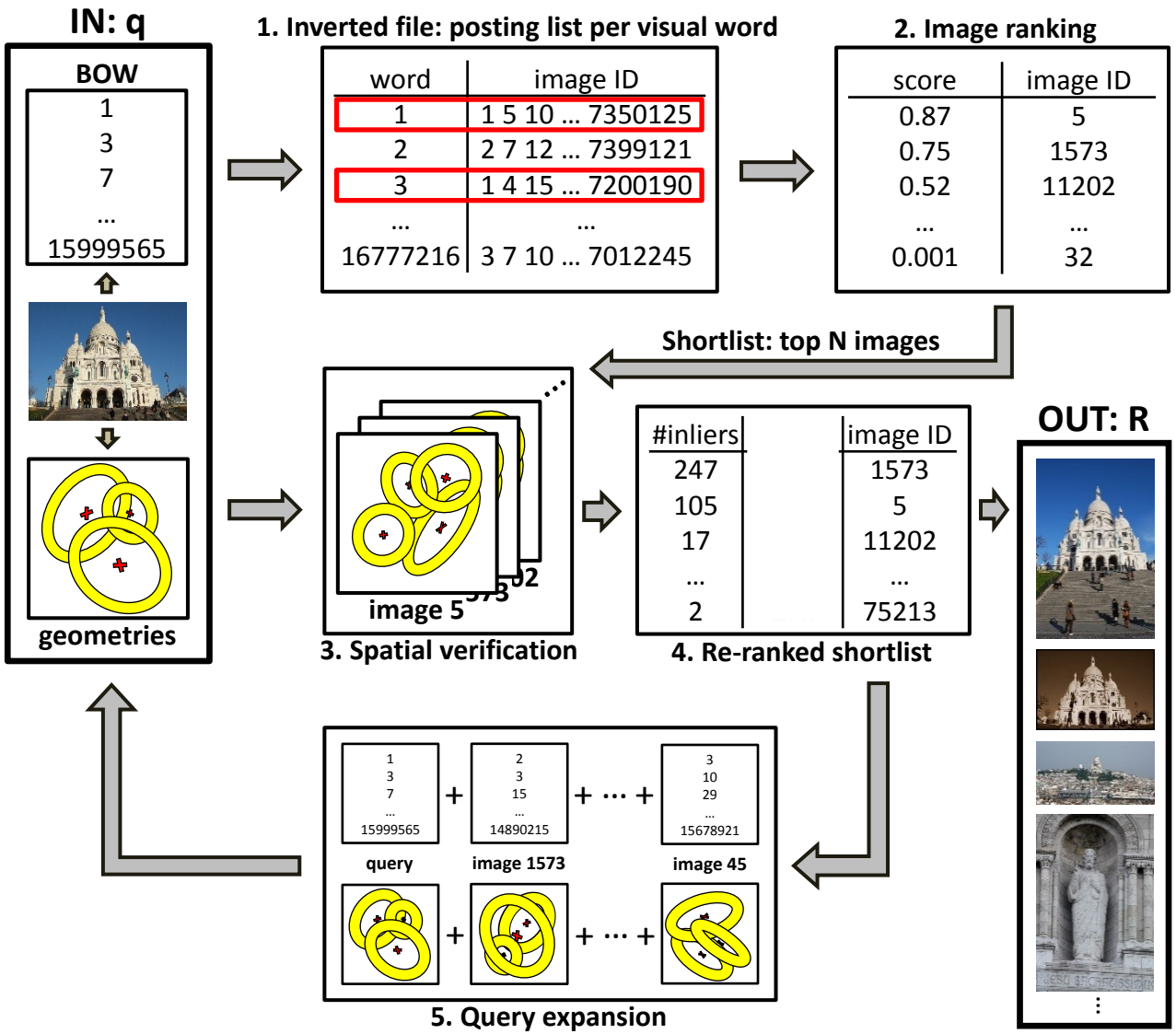
Bag of Words Image Representation



Term-frequency (tf) – visual word D is twice in the image

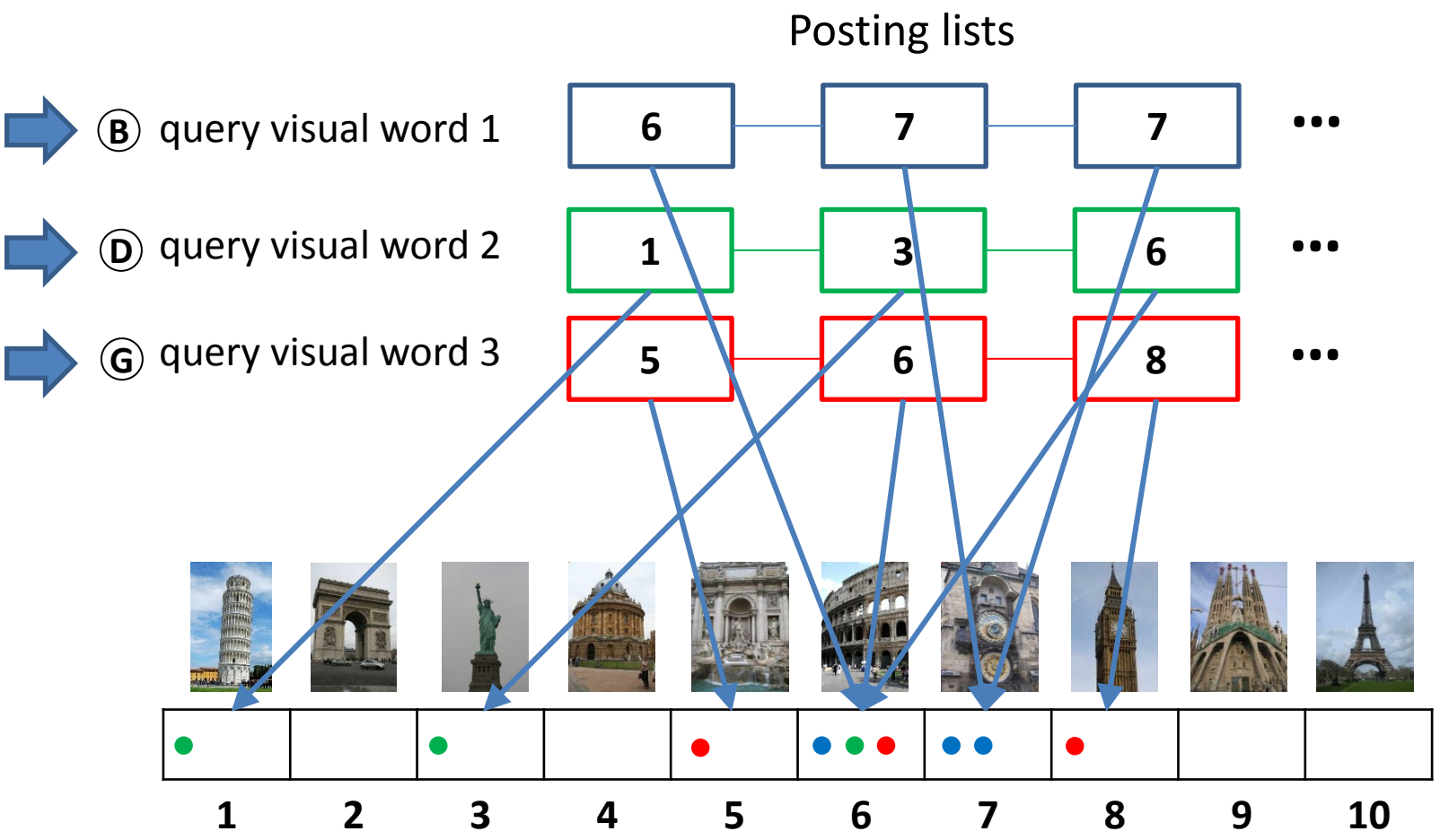
Images are represented by sparse vector / histogram of visual words present in them

Bag of Words : On-line Stage



Bag of Words Scoring

$$\text{score} = \frac{\mathbf{q}^\top \mathbf{x}}{||\mathbf{x}||}$$



Geometric Re-ranking

Re-rank top ranked images (removing false positives)

- RANSAC

NOTE: Standard BoW score ranking performed without geometric information

IMPORTANT: Geometric verification crucial for query expansion

Sivic, Zisserman: Video Google, ICCV 2003

Philbin, Chum, Isard, Sivic, Zisserman: Object retrieval with large vocabularies and fast spatial matching, CVPR'07

Query Expansion

Results

Query image

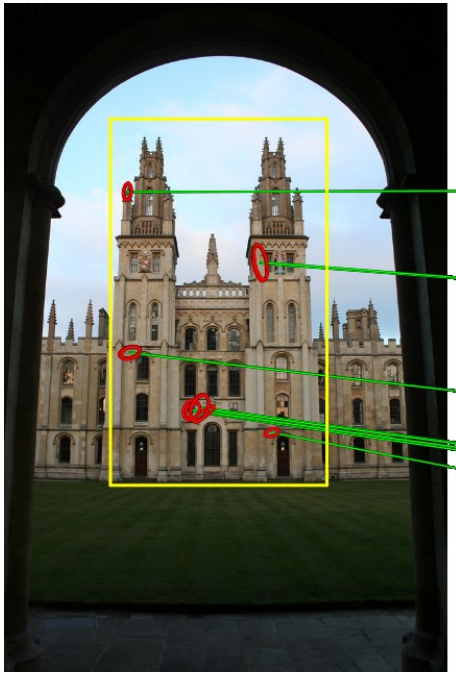
Spatial verification

New query

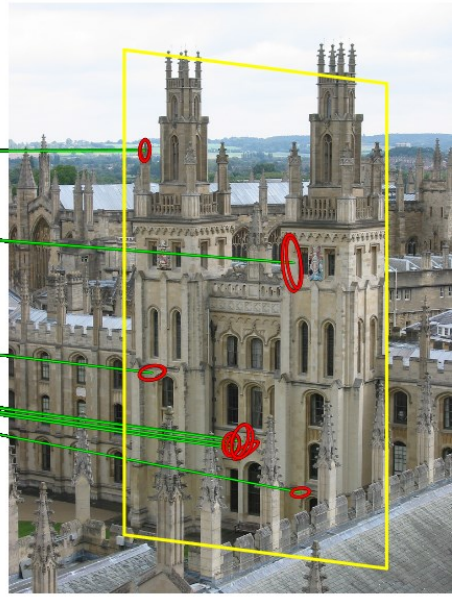
New results

Chum, Philbin, Sivic, Isard, Zisserman: Total Recall…, ICCV 2007

Query Expansion: Step by Step



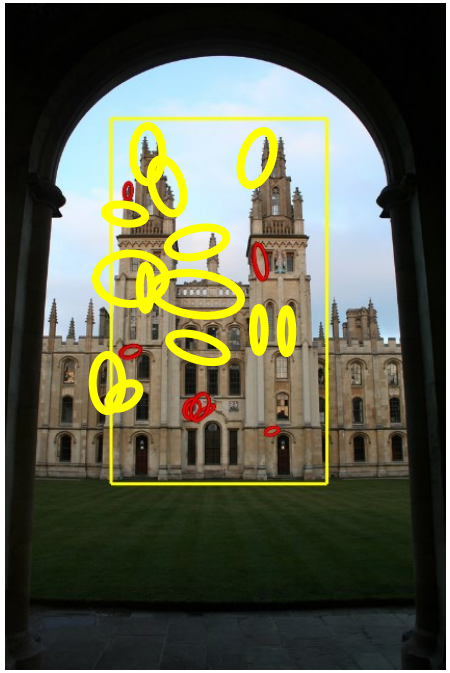
Query Image

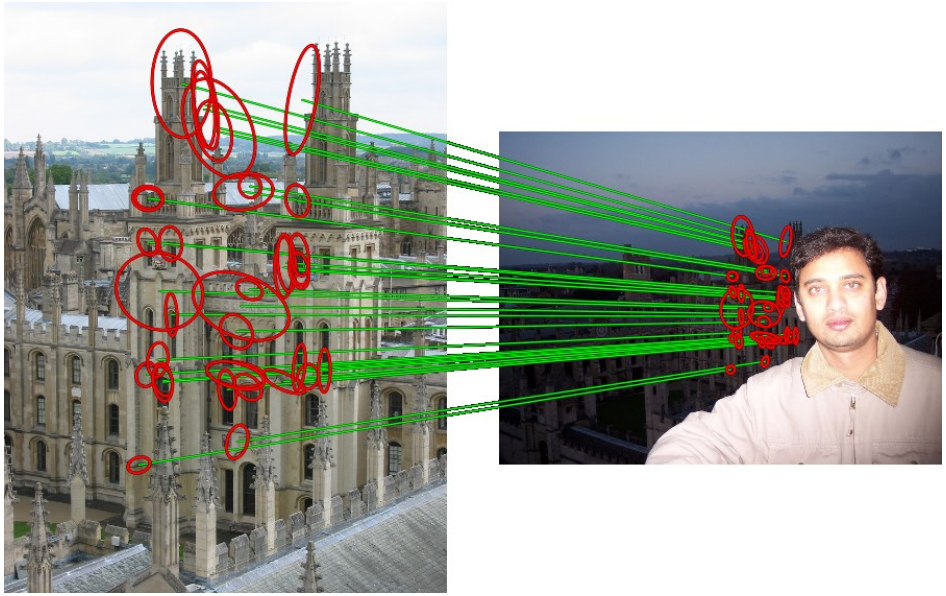


Retrieved image

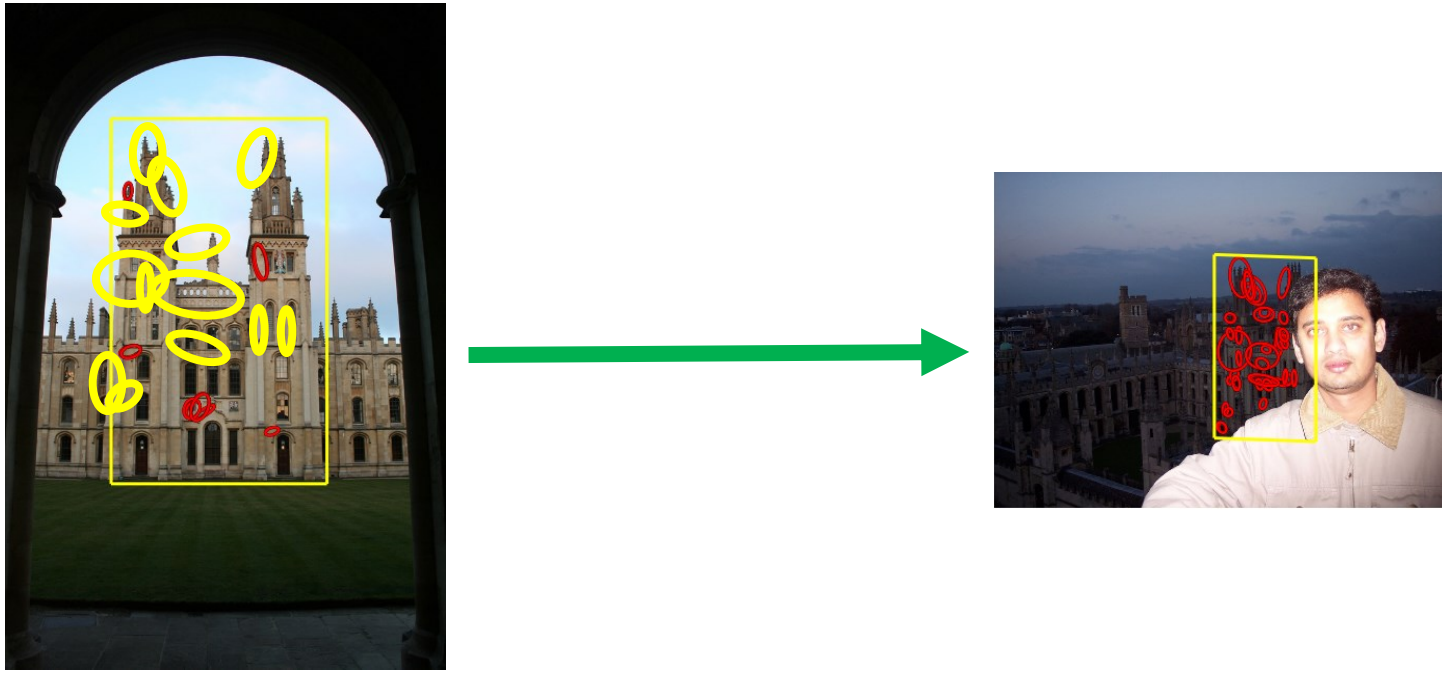
Originally not retrieved

Query Expansion: Step by Step





Query Expansion: Step by Step



2.0: Beyond Similarity Retrieval

Other Retrieval Problems

What is this?

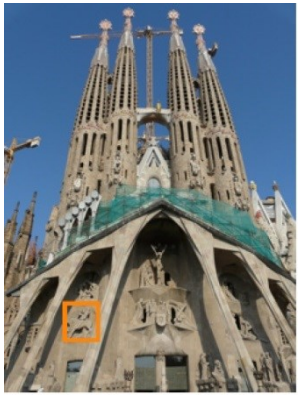
… and what is that?

Let's **zoom-in**!

Different Retrieval Problems

Top: visually most similar

Bottom: zoom-in



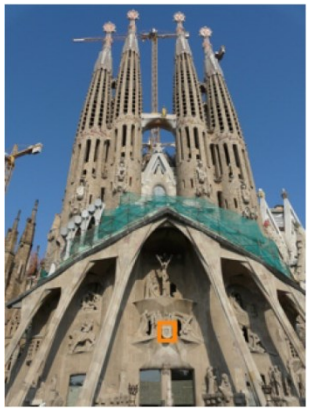
Query 1



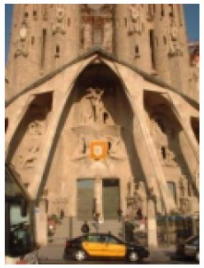






Query 2







Mikulik, Chum, Matas: **Image Retrieval for Online Browsing in Large Image Collections**, SISAP 2013.

Standard Retrieval and Details


query

rank:


1


2


32


64


65

EASY


query

rank:


1

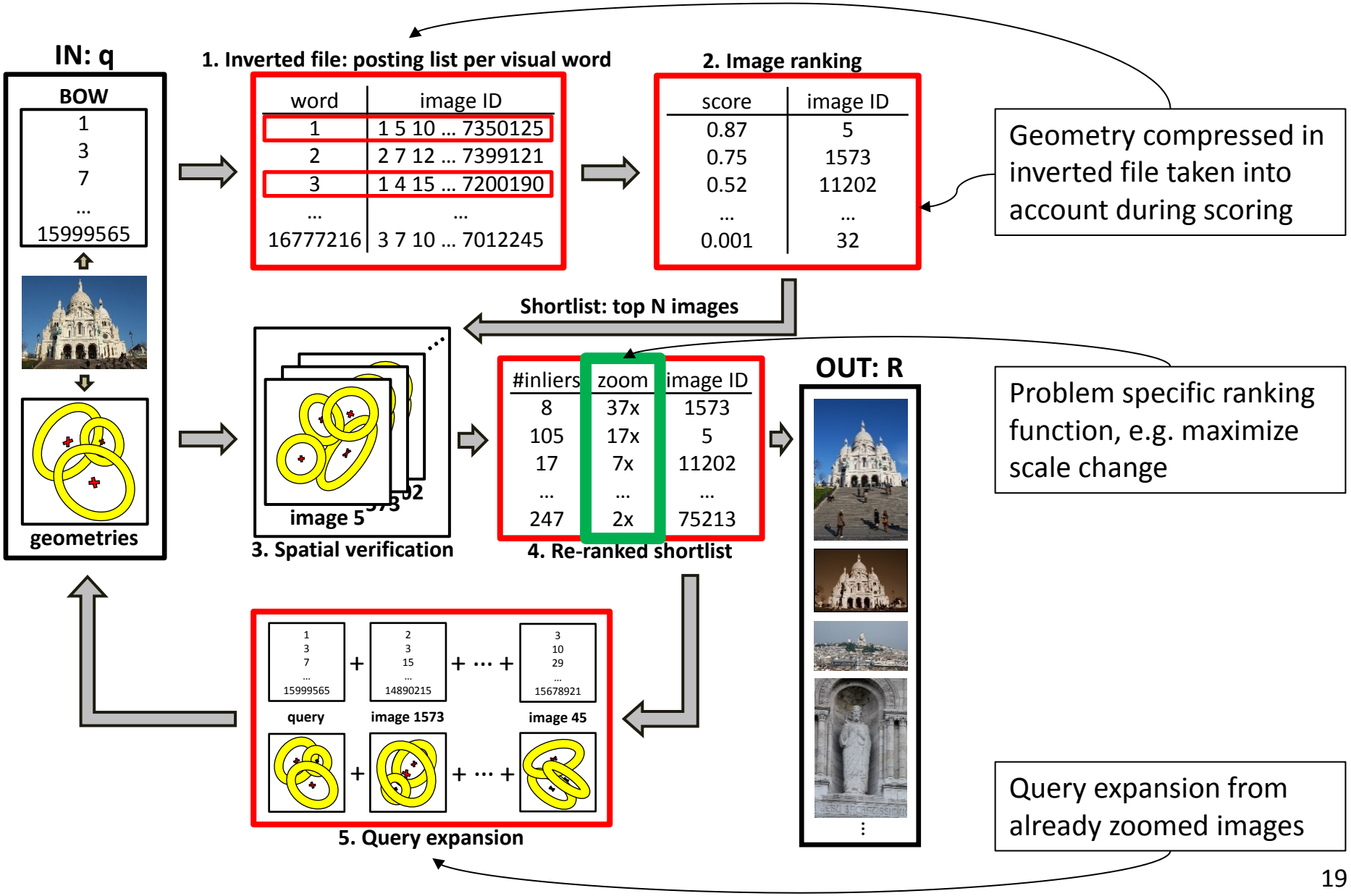

2048


16384

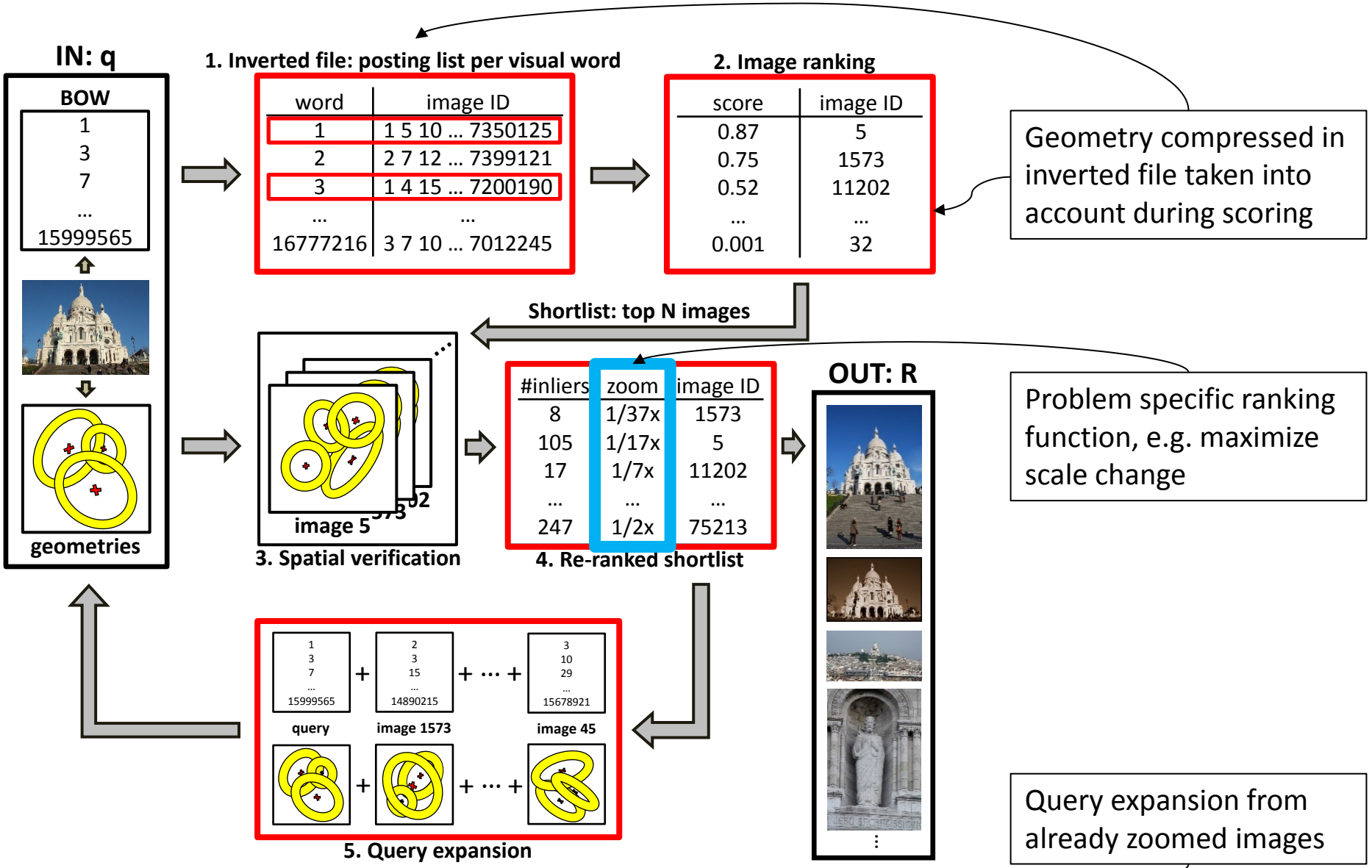

81368

DIFFICULT

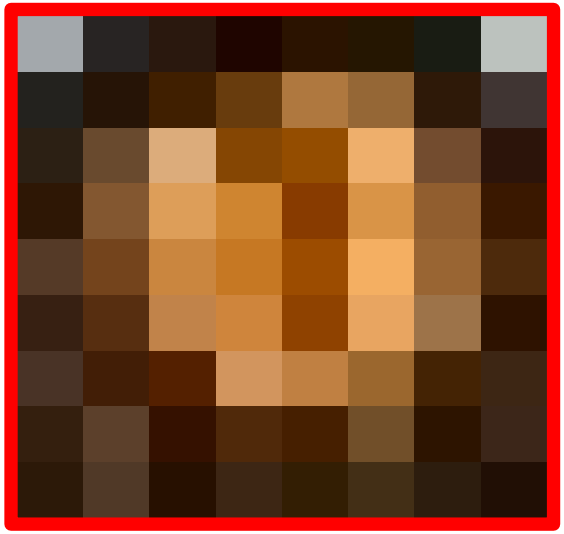
Zoom-in: On-line Stage



Zoom-out: On-line Stage

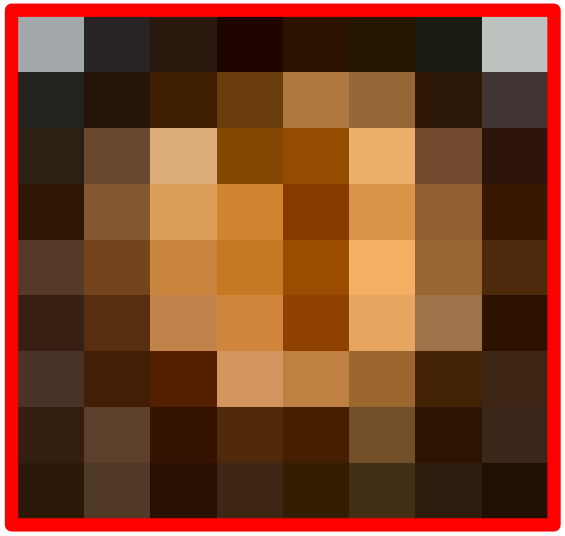


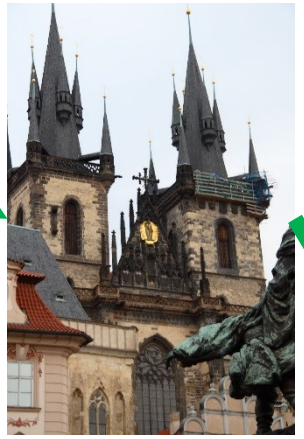
Zoom-in: Example



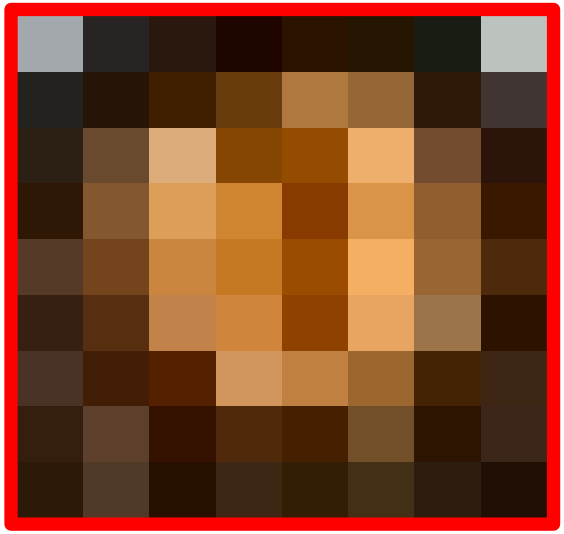
?

Zoom-in: Query Expansion

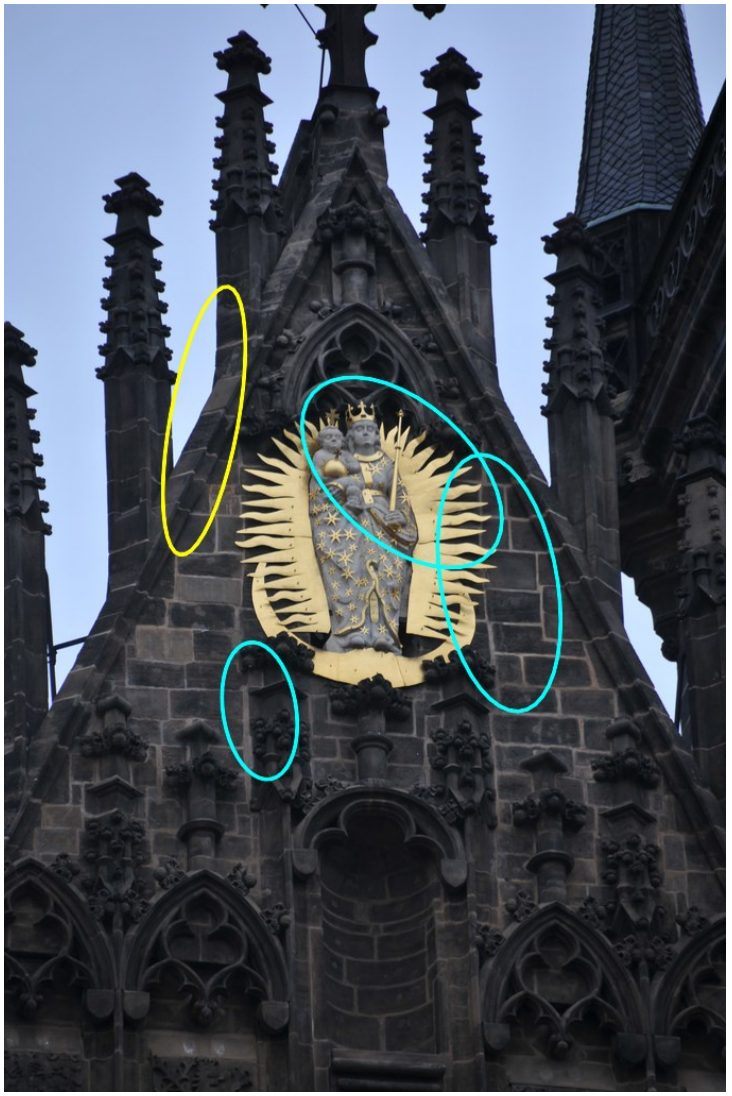




Zoom-in: Example

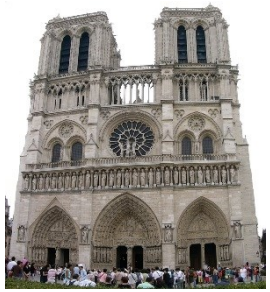


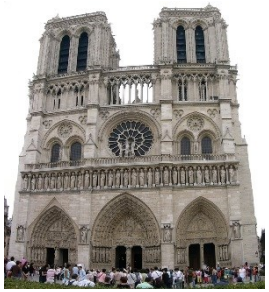
Zoom-in: Query Expansion



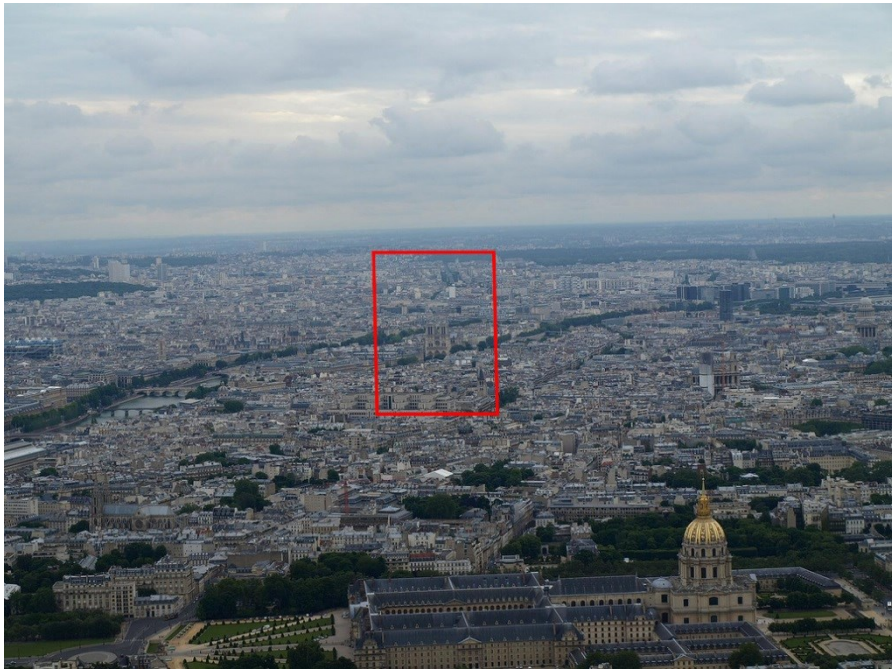
Zoom-in: Query Expansion

Zoom-out: Iterate

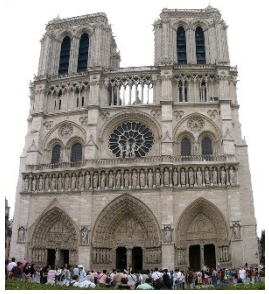


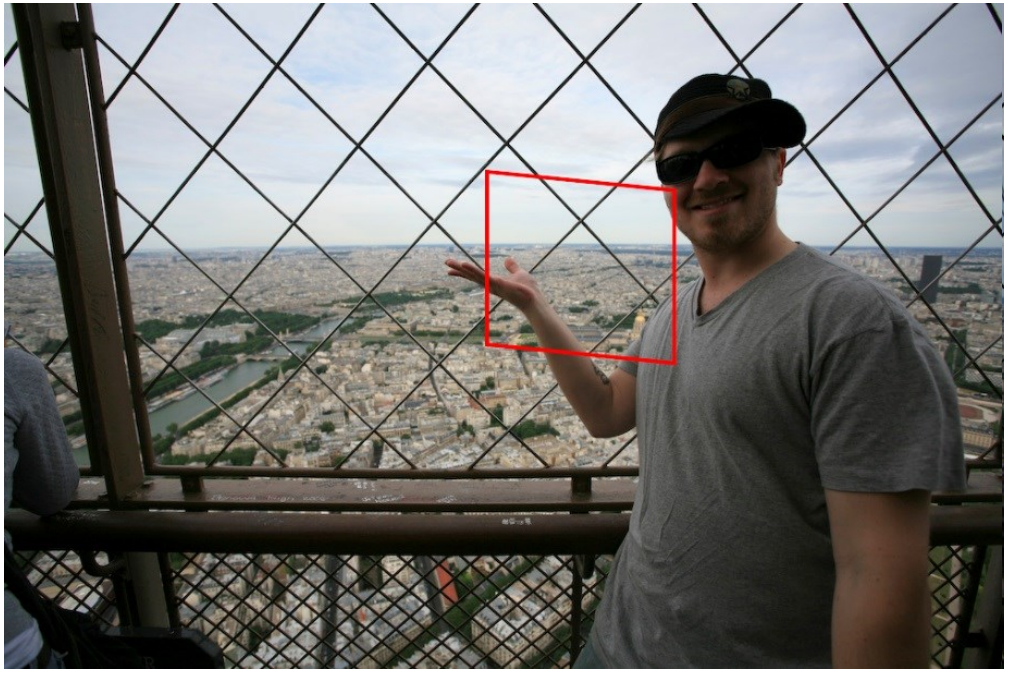


Zoom-out: Iterate



Zoom-out: Iterate



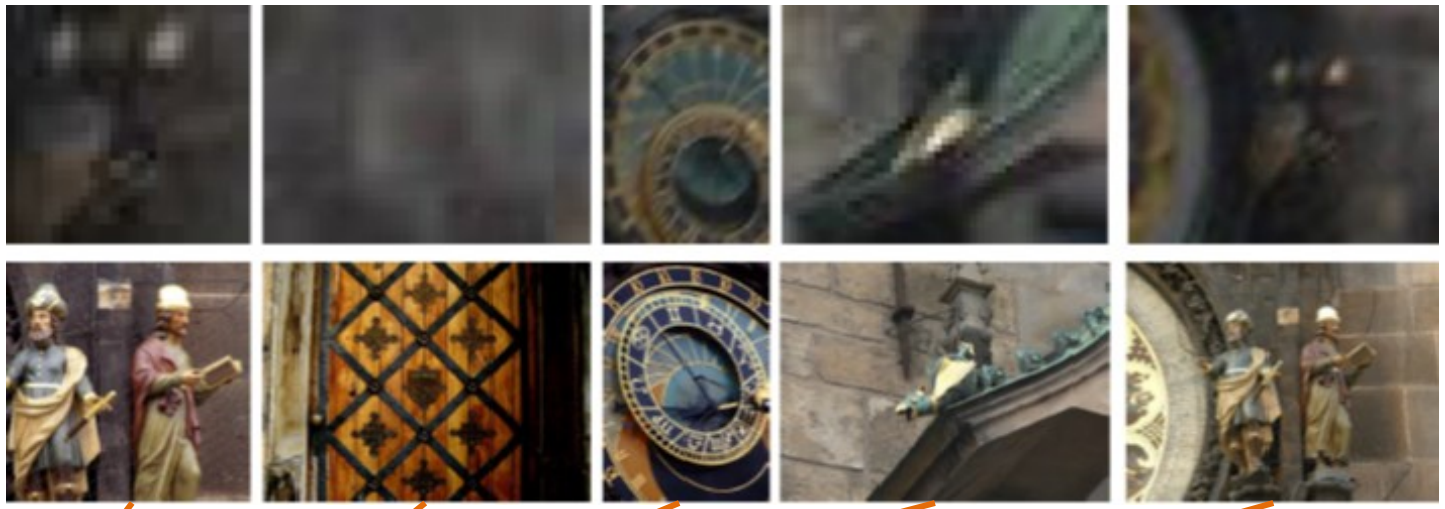


What is interesting here?

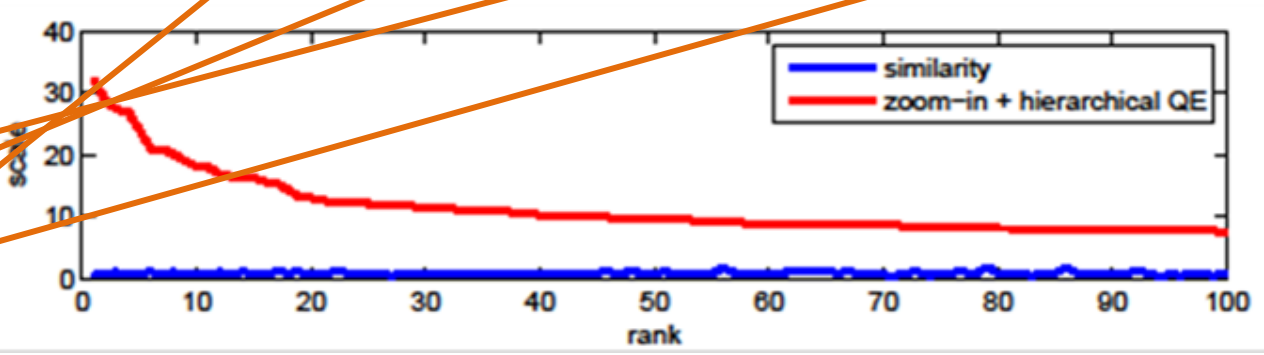
What should you not miss?

Highest Resolution Transform

Given a query and a dataset, for every pixel in the query image:
Find the database image with the maximum resolution depicting the pixel



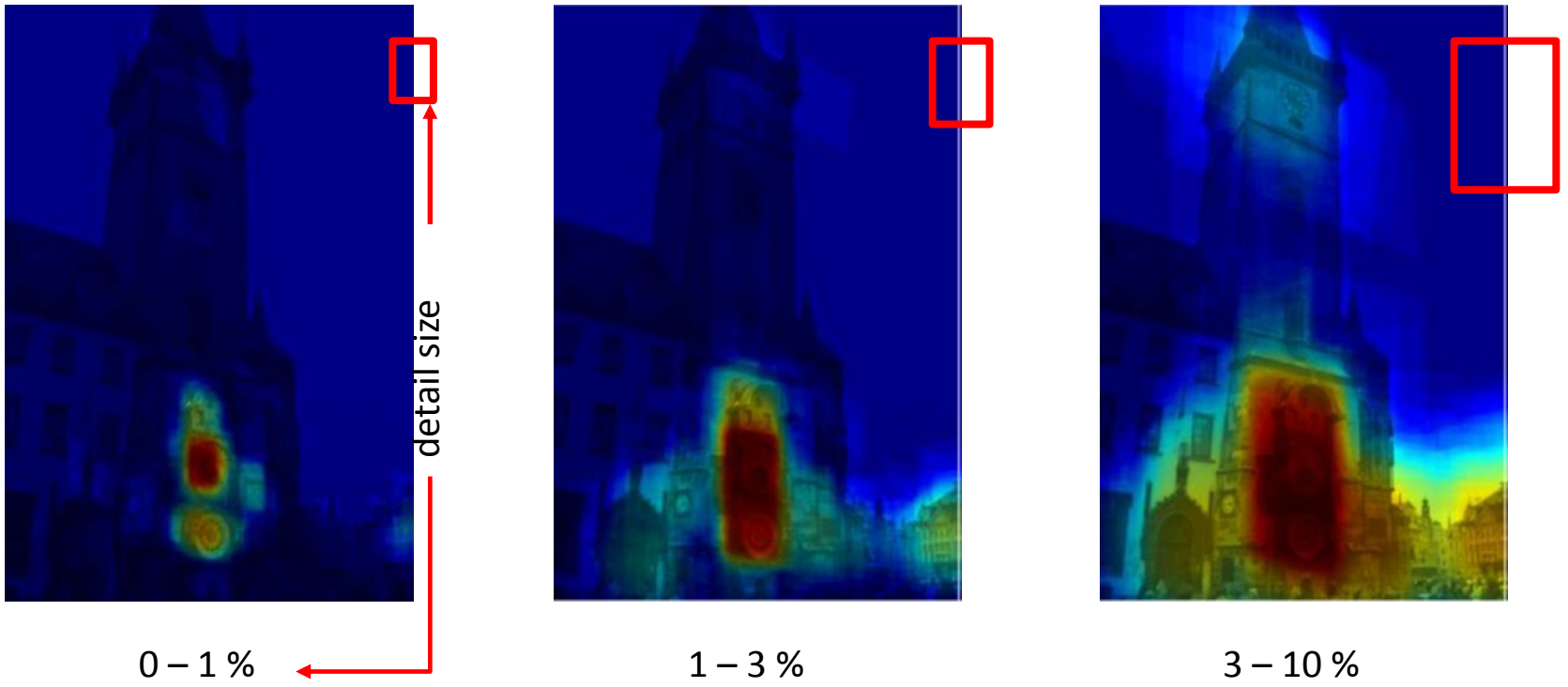
37.3x 27.0x 22.8x 21.9x 21.6x



Mikulík, Radenović, Chum, Matas: **Efficient Image Detail Mining,** ACCV 2014

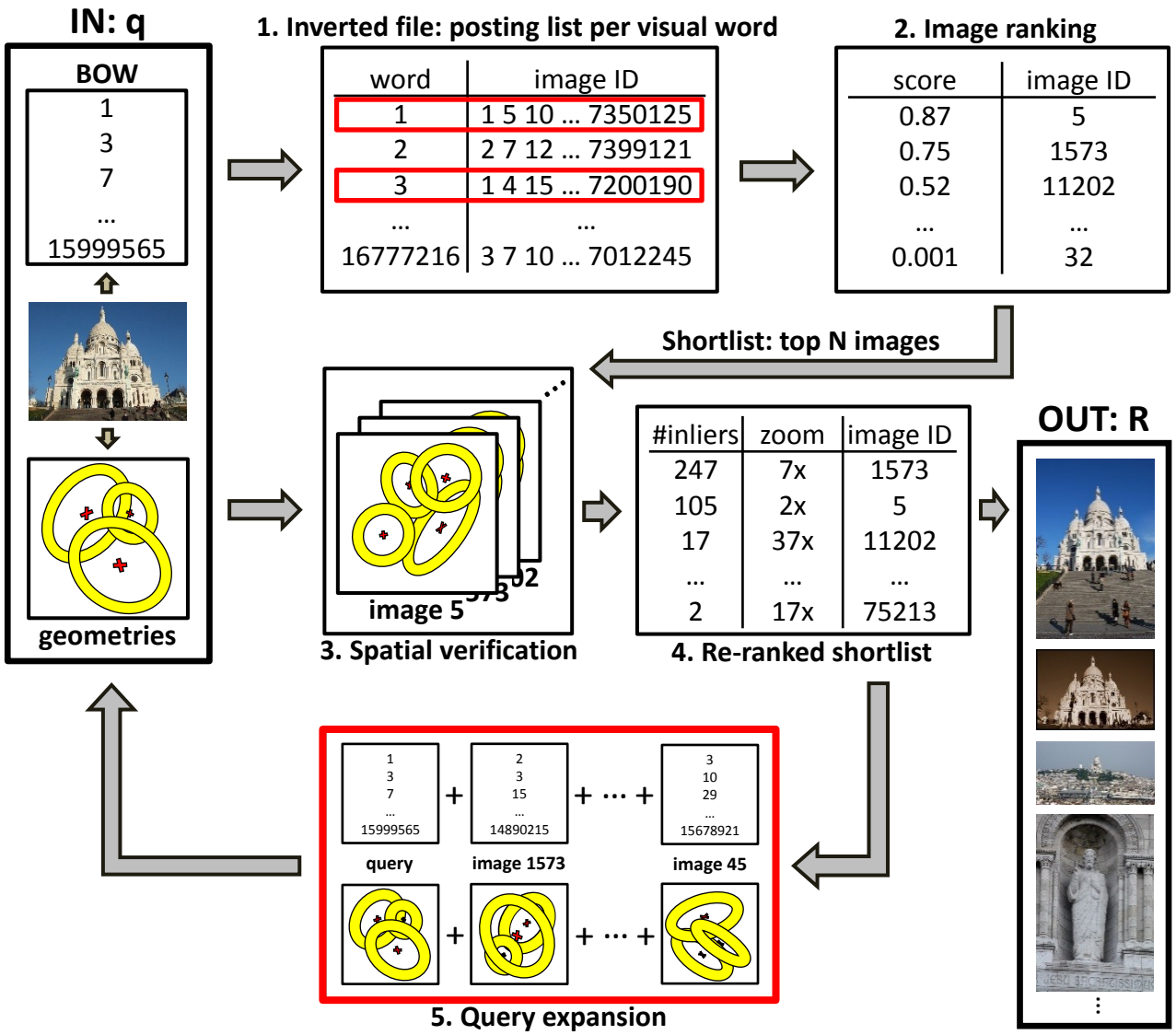
What most people find interesting?

Most commonly photographed parts
Given a query and a dataset, for every pixel in the query image:
Find the frequency with which it is photographed in detail

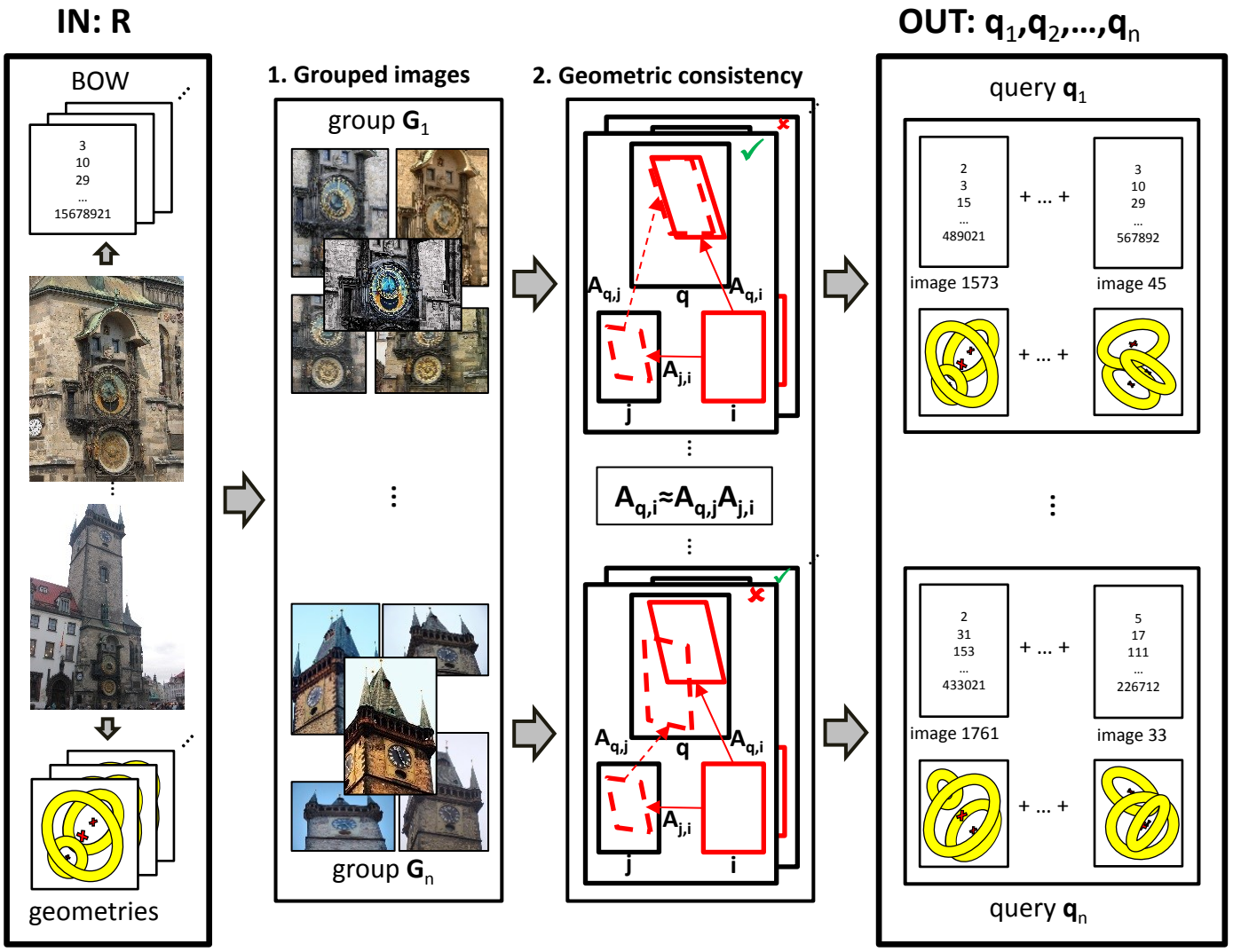


Mikulík, Radenović, Chum, Matas: **Efficient Image Detail Mining,** ACCV 2014

All Details: On-line Stage



All Details: Hierarchical Query Expansion



2.1: Image Retrieval for 3D Reconstruction

Structure-from-Motion 3D Reconstruction

- Few thousand images

  Exhaustive matching of all image pairs

  [Snavely, Seitz, Szeliski: **Photo tourism**, SIGGRAPH 2006]

  + High level of details reconstructed

  - Unfeasible for larger photo collections

- Few million images

  Matching images through standard image retrieval

  [Heinly, Schonberger, Dunn, Frahm: **Reconstructing the World in Six Days**, CVPR 2015]

  + Efficient and scalable image matching

  - Details not reconstructed

Retrieval for 3D Reconstruction

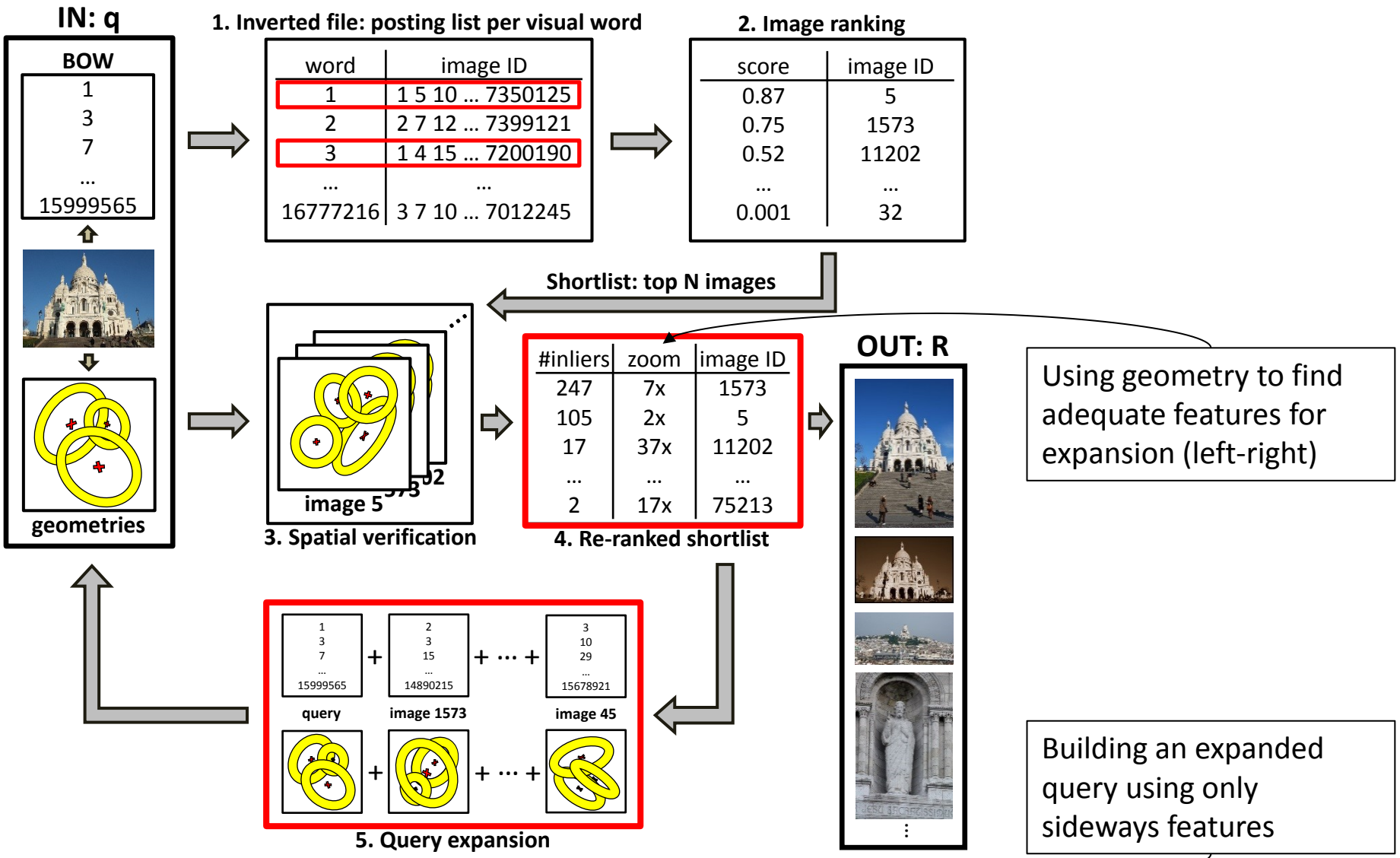
- Visually most similar search
  - Many near duplicates
  - Details lost
- Zoom-in and details search
  - Details retrieved
  - Transition images to match the details
- Zoom-out search
  - Viewpoint change
  - More context
- Sideways crawl
  - Significant viewpoint change
  - More context

Schoenberger, Radenović, Chum, Frahm: **From Single Image Query to Detailed 3D Reconstruction,** CVPR 2015

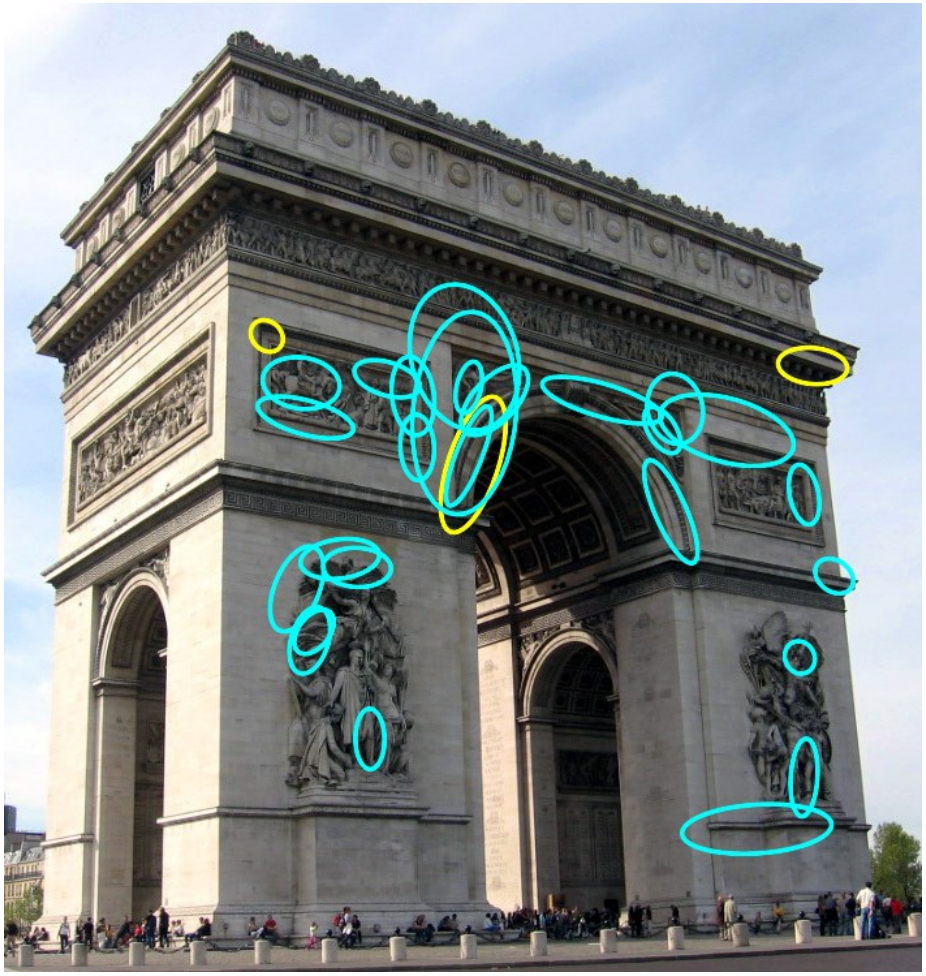
Sideways image crawl

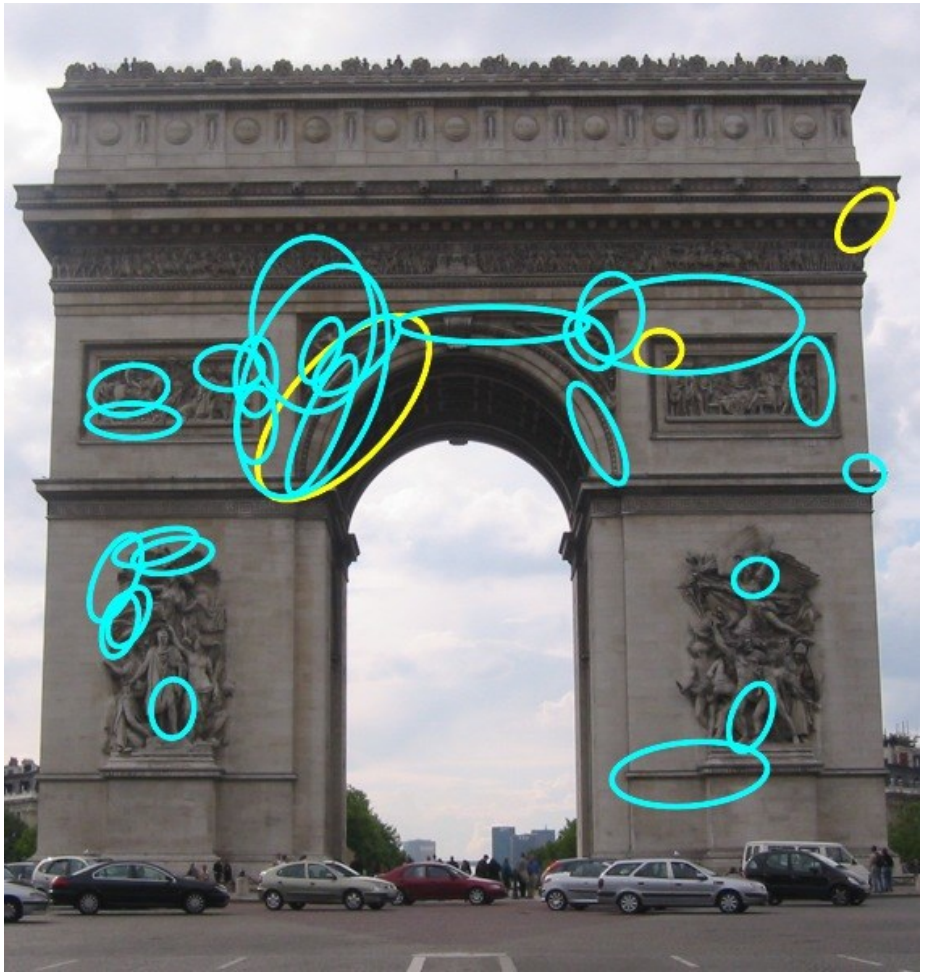
Schoenberger, Radenović, Chum, Frahm: **From Single Image Query to Detailed 3D Reconstruction,** CVPR 2015

Sideways crawl: On-line Stage

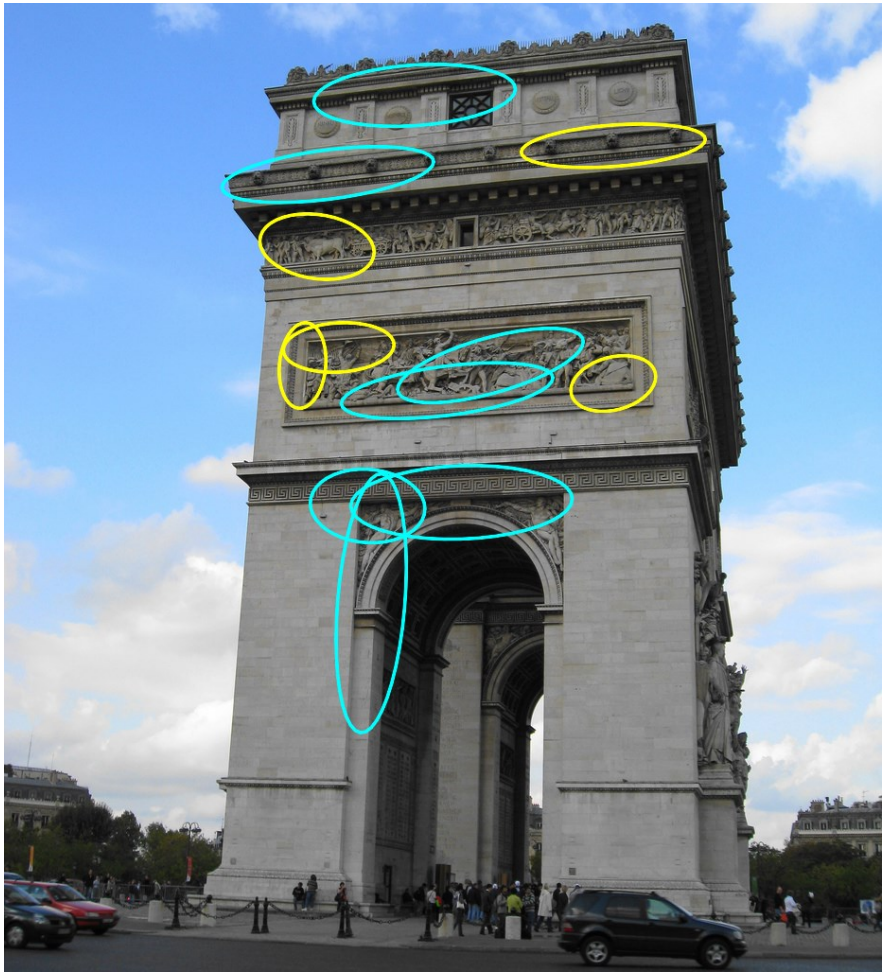


Sideways Left: Step by Step





Sideways Left: Step by Step



Retrieval for 3D Reconstruction

See our video at:

https://youtu.be/DIv1aGKqSIk

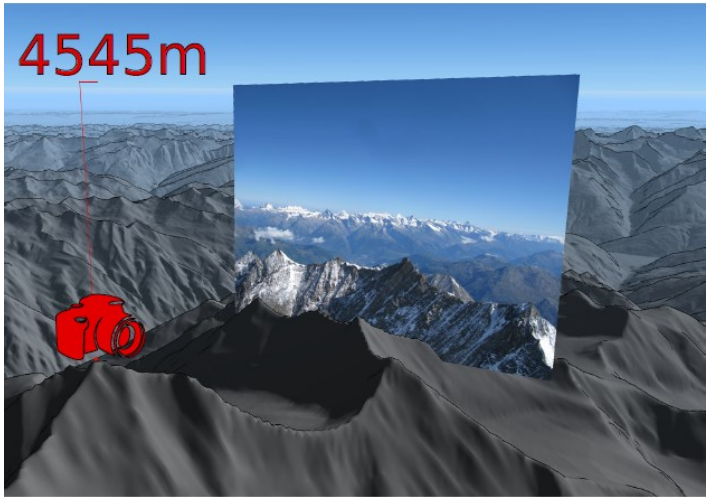
VIDEO

Schoenberger, Radenović, Chum, Frahm: **From Single Image Query to Detailed 3D Reconstruction,** CVPR 2015

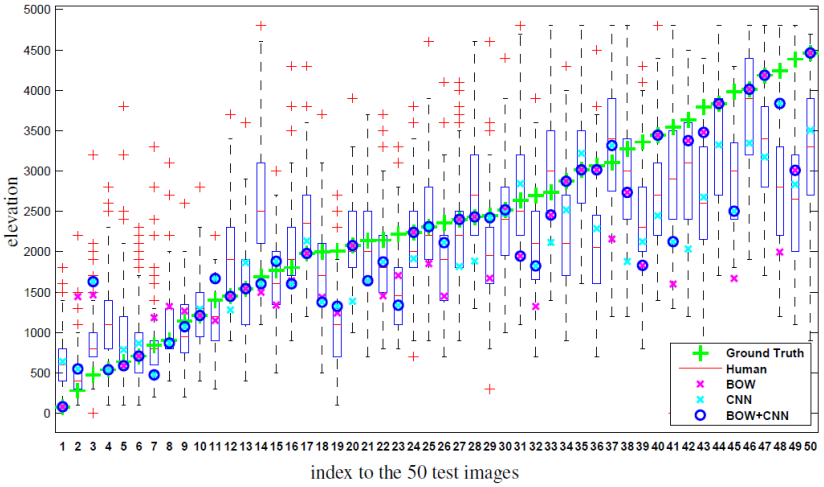
Localization: Most Similar Retrieval

Application: Camera Elevation Estimation

- Automatic elevation estimation from image content
- Location recognition in Alps
- Inferring height from a training dataset by using recognized location



Method	test dataset (13148 images)	user experiment set (50 images)
Baseline	801.49; *786.42*	1383.64; *1154.43*
Human	-	**879.95**
CNN	537.11	709.10
BOW	601.63	757.76
mVocab	610.36	811.00
BOW+mVocab	564.14	646.89
BOW+CNN	**500.44**	**531.05**

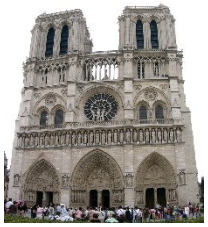


Čadík, Vašíček, Hradiš, Radenović, Chum: **Camera Elevation Estimation from a Single Mountain Landscape Photograph,** BMVC 2015

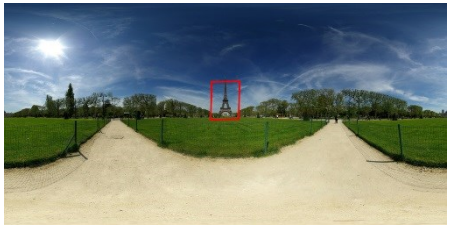
Summary

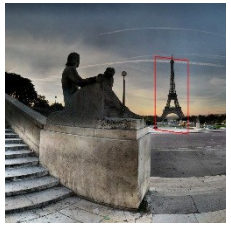



Visually most similar




Zoom-in / details



Zoom-out




Sideways right

Thank you!

Questions?