



Download notebook paper:
http://bit.ly/ARTEMIS_INS

Authors

Andrei Bursuc
Titus Zaharia
Olivier Martinot
Françoise Prêteux

ARTEMIS Department
Institut TELECOM
TELECOM SudParis
UMR CNRS 8145 MAP5
9, rue Charles Fourier
91011 Evry Cedex
France

Partners

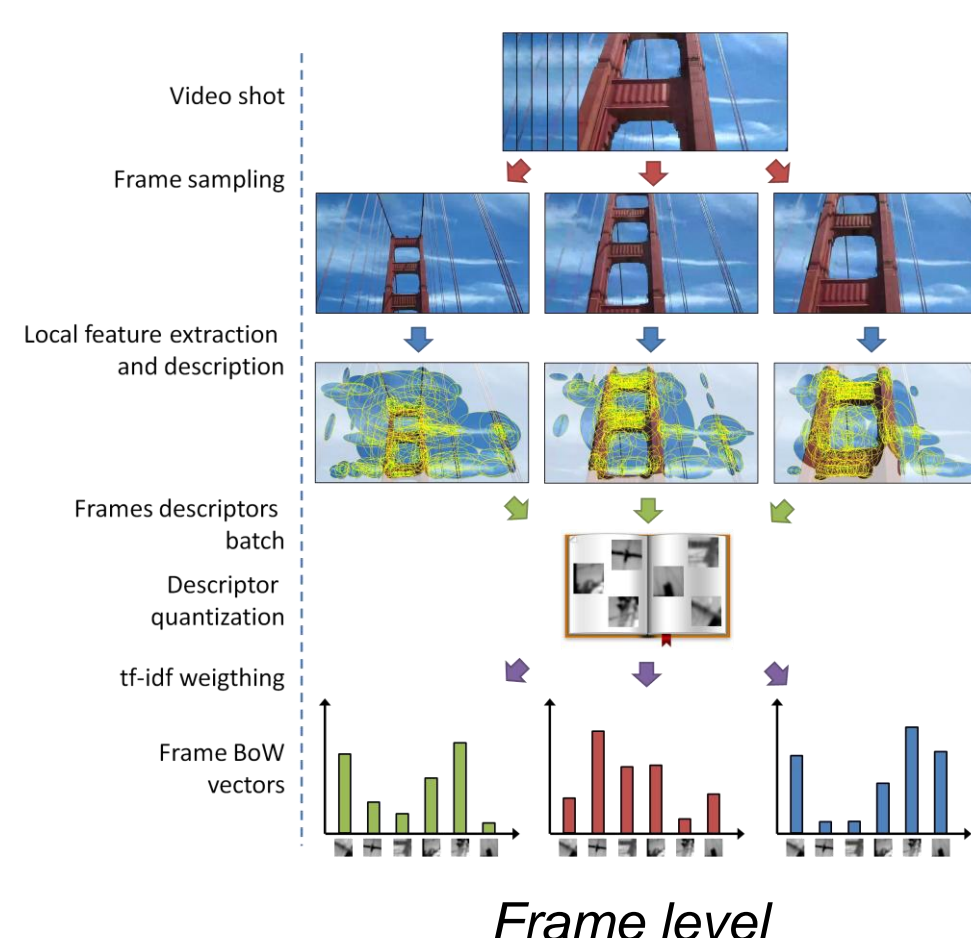
Alcatel-Lucent



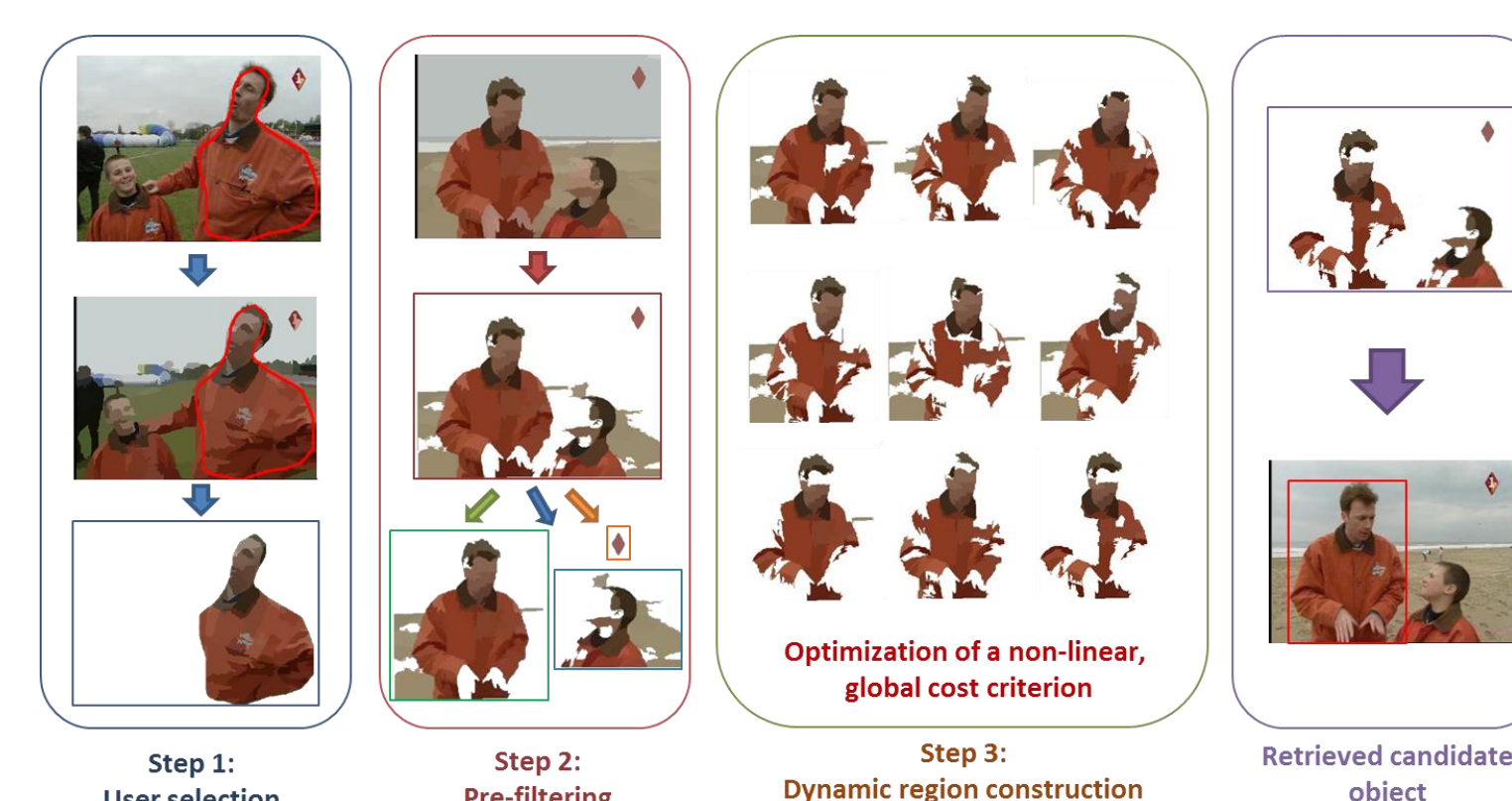
Approach overview

- Bag-of-Words + reranking using oversegmented regions and color descriptors
- Uniform frame sampling: 1 frame per second (683,433 keyframes)
- Keyframes resized to a maximum of 110592 pixels
- Description:
 - Hessian-Affine covariant regions + RootSIFT descriptors (245,575,000 regions)
 - MPEG-7 Color Structure Descriptor
 - MeanShift segments + extended MPEG-7 Dominant Color Descriptor (160 segments/keyframe)
- BoW quantization
 - Random sampling of 10% descriptors from each keyframe
 - Clusterization in a visual vocabulary of size 1,000,000 words

BoW video clip description:

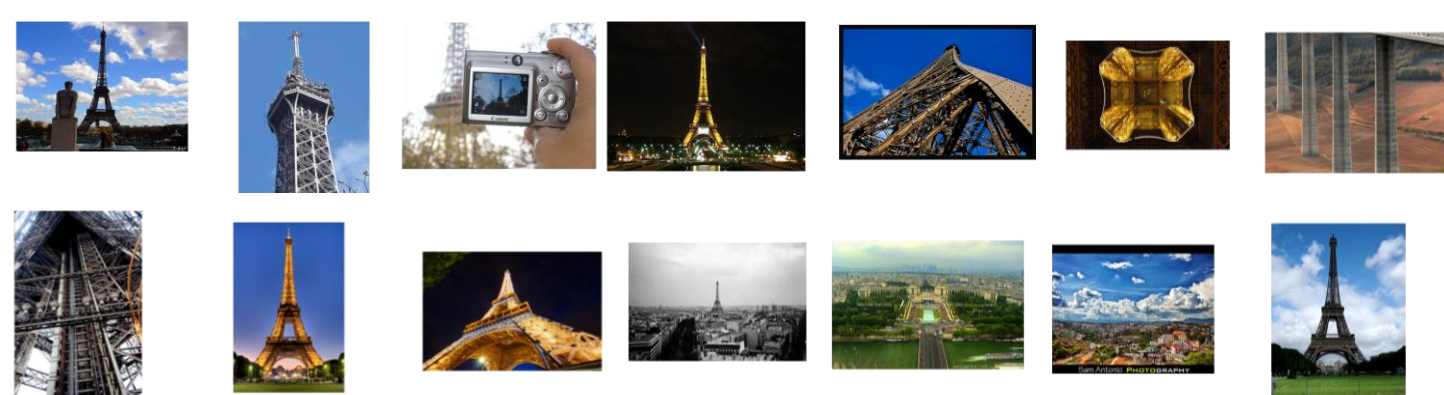


Region based image description:

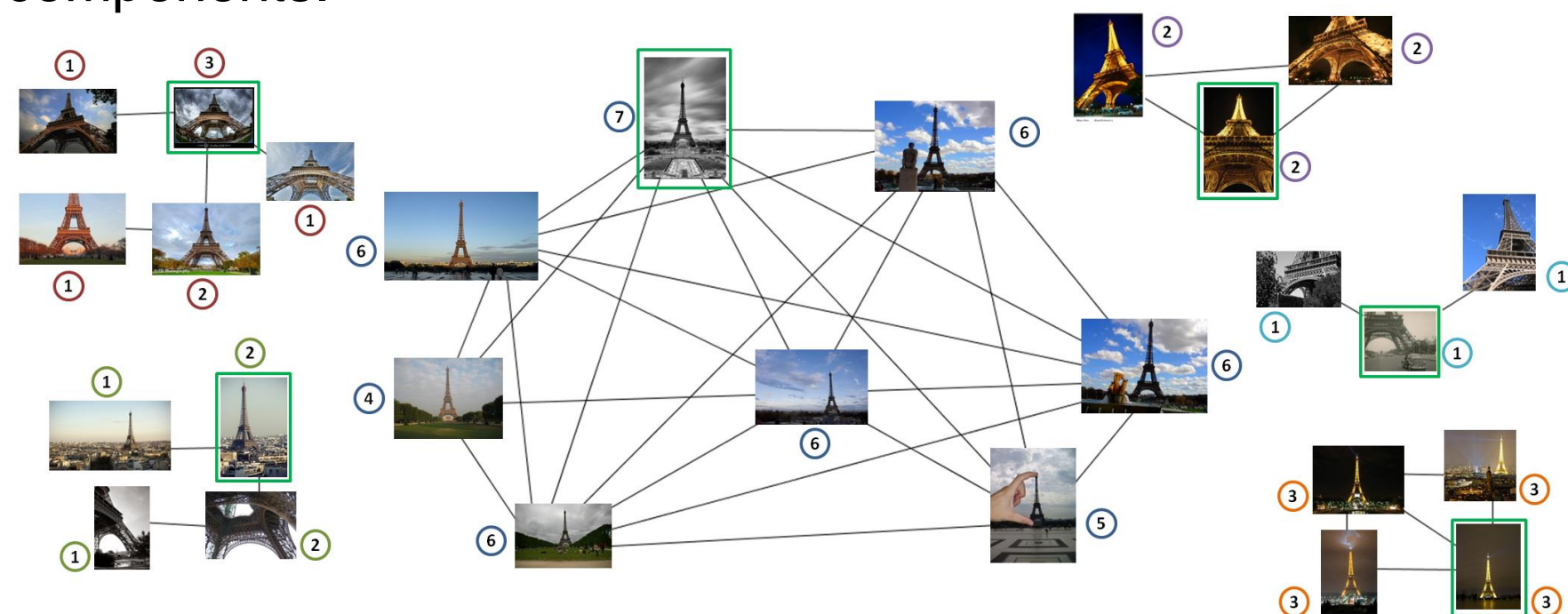


Multi-modal query expansion

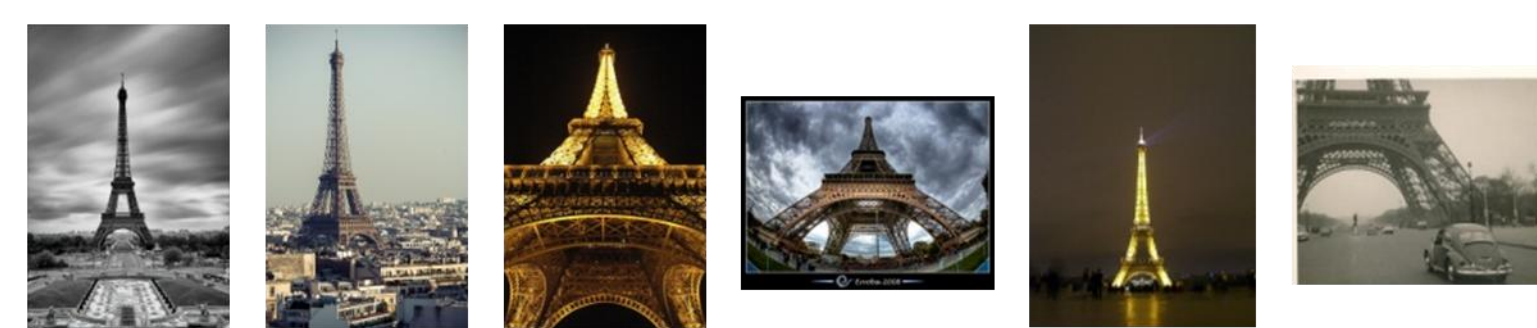
1. Download images from Flickr using textual descriptions (25-100).



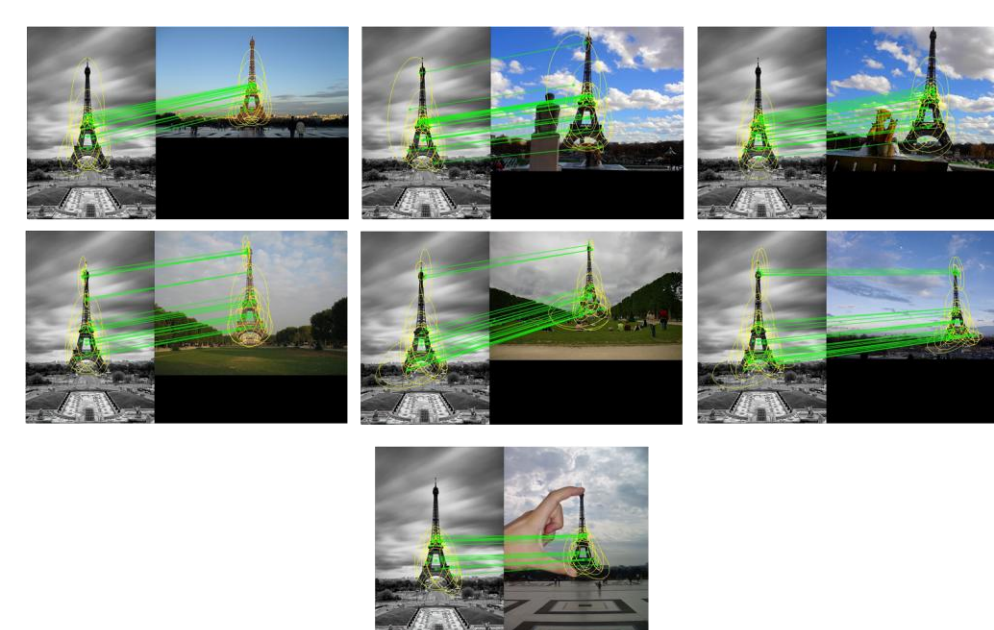
2. Discard negative images by matching the image set and projecting it into a query graph. Similar instances are identified in connected components.



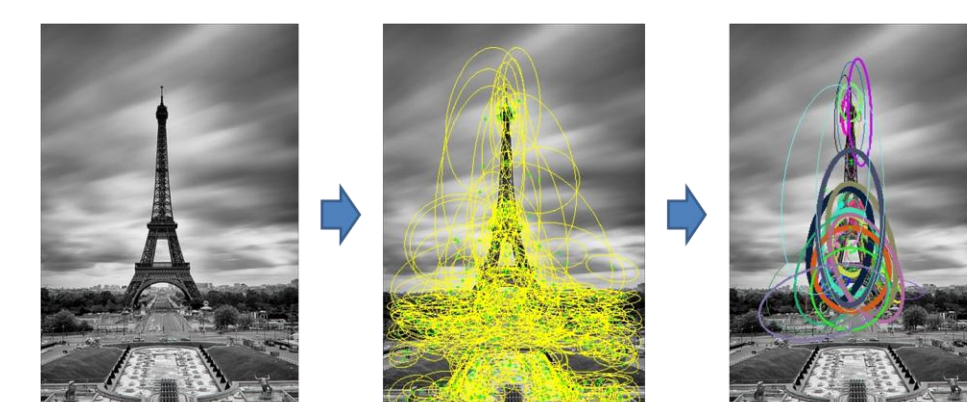
3. Identify representative images as images with highest degree from each connected component.



4. Compute new descriptors from representative images and their matches.

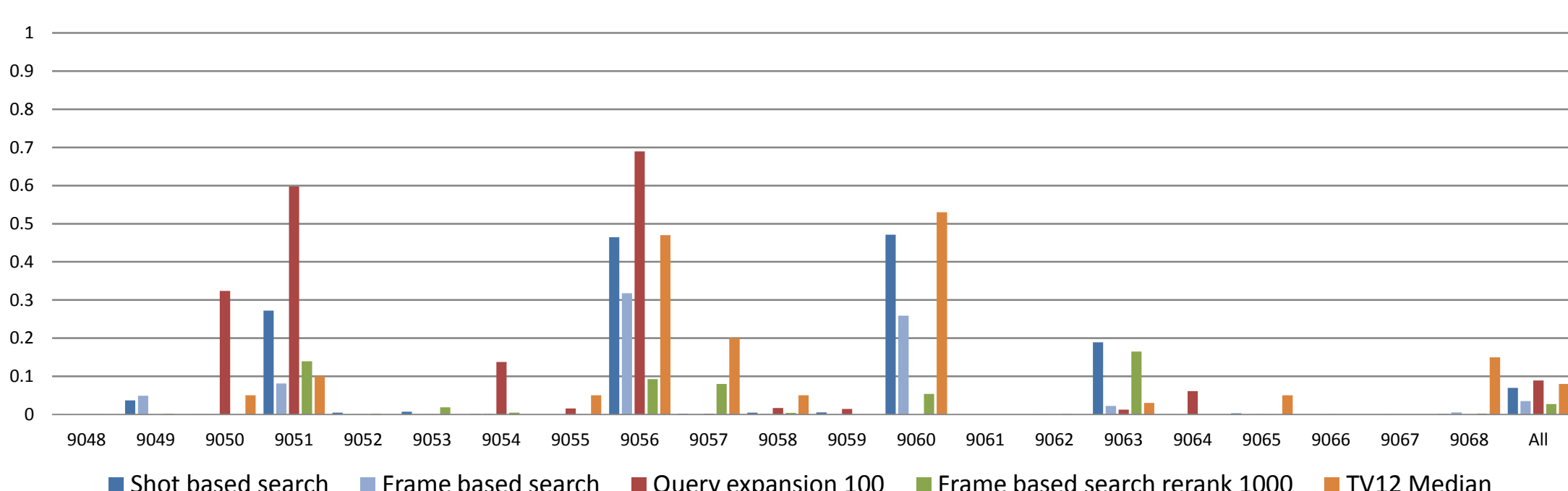


Representative image and its geometrically verified matches



Enriched representative image with regions weighted proportionally with the number of matches

Results



Conclusions

- Shot level descriptors yielded best results.
- Flickr-based expanded queries improved the performances.
- Some expanded queries were negative due to textual description ambiguity.
- For multiple example queries, positive results were discarded in the result aggregation process.
- Color based reranking improved results only for topics 9057, 9063, 9068.