Advanced Social Media Analysis

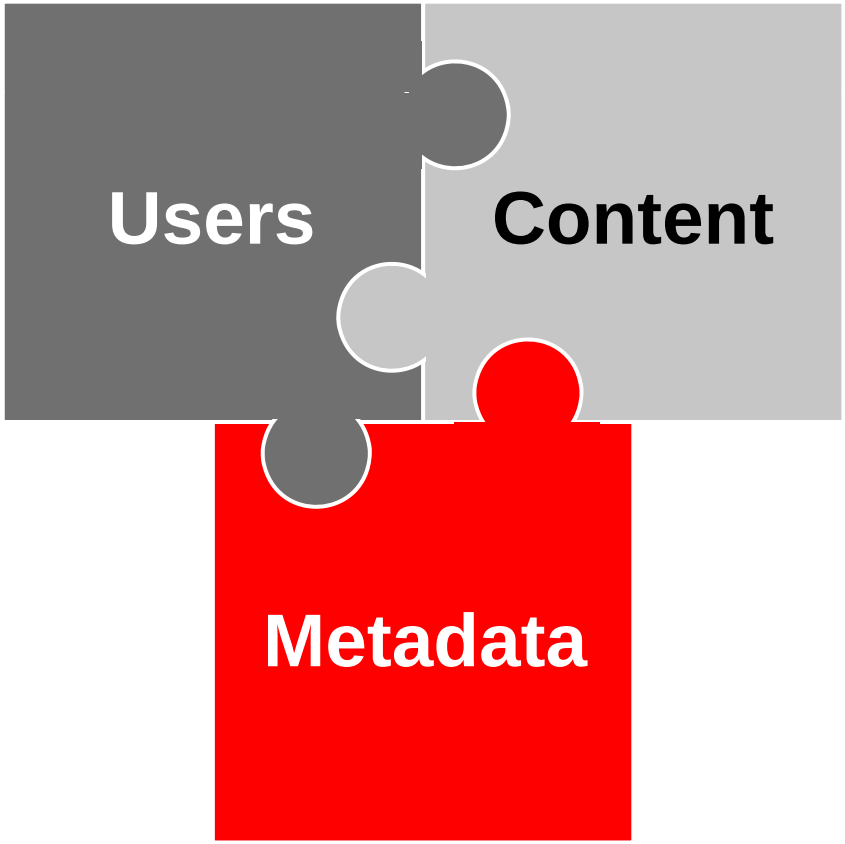
– PhD thesis –

Author:
Ivan Ivanov, PhD

Multimedia Signal Processing Group
Swiss Federal Institute of Technology

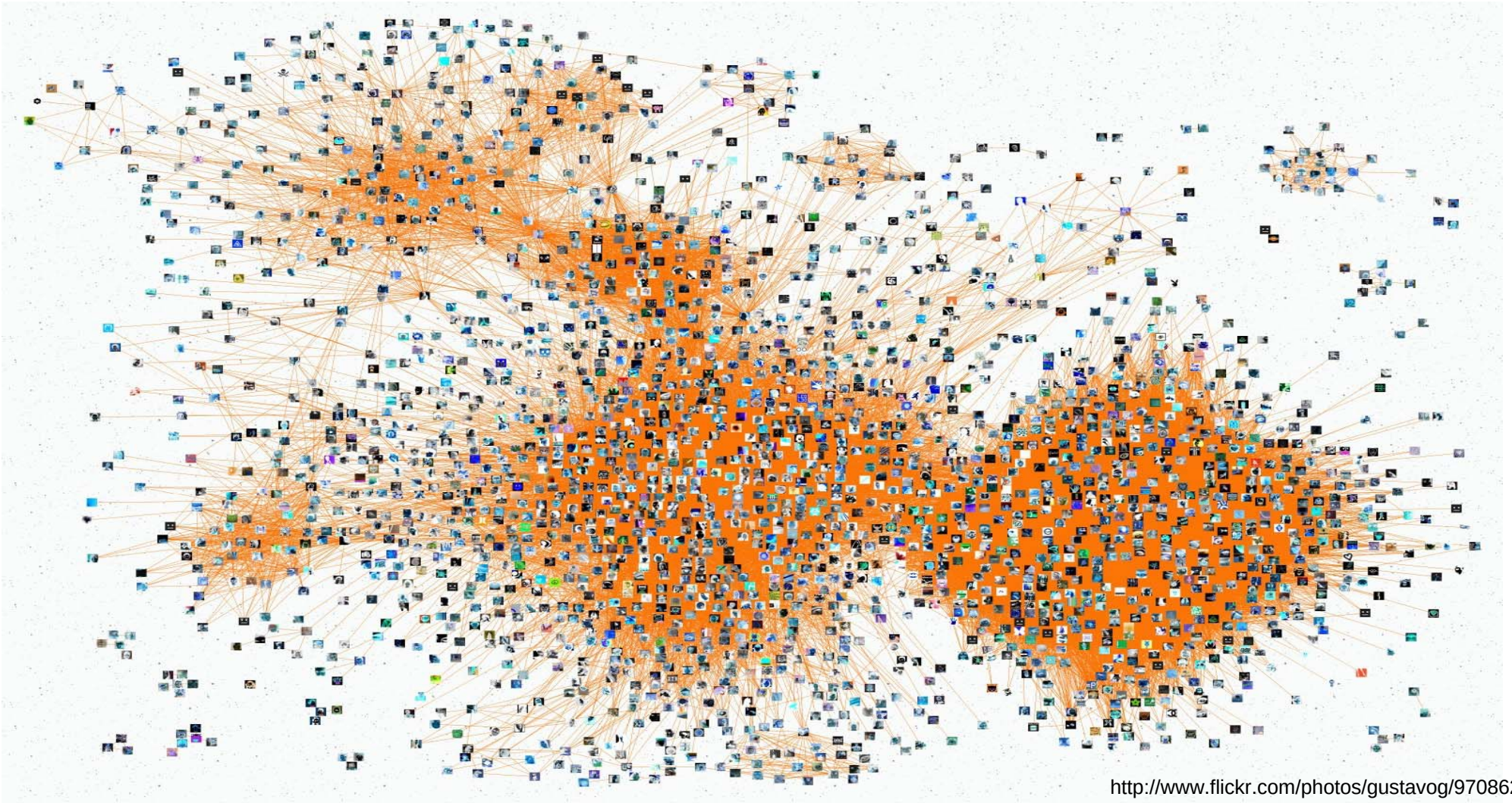
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Object-based Tag Propagation for Semi-Automatic Annotation of Images



Motivation

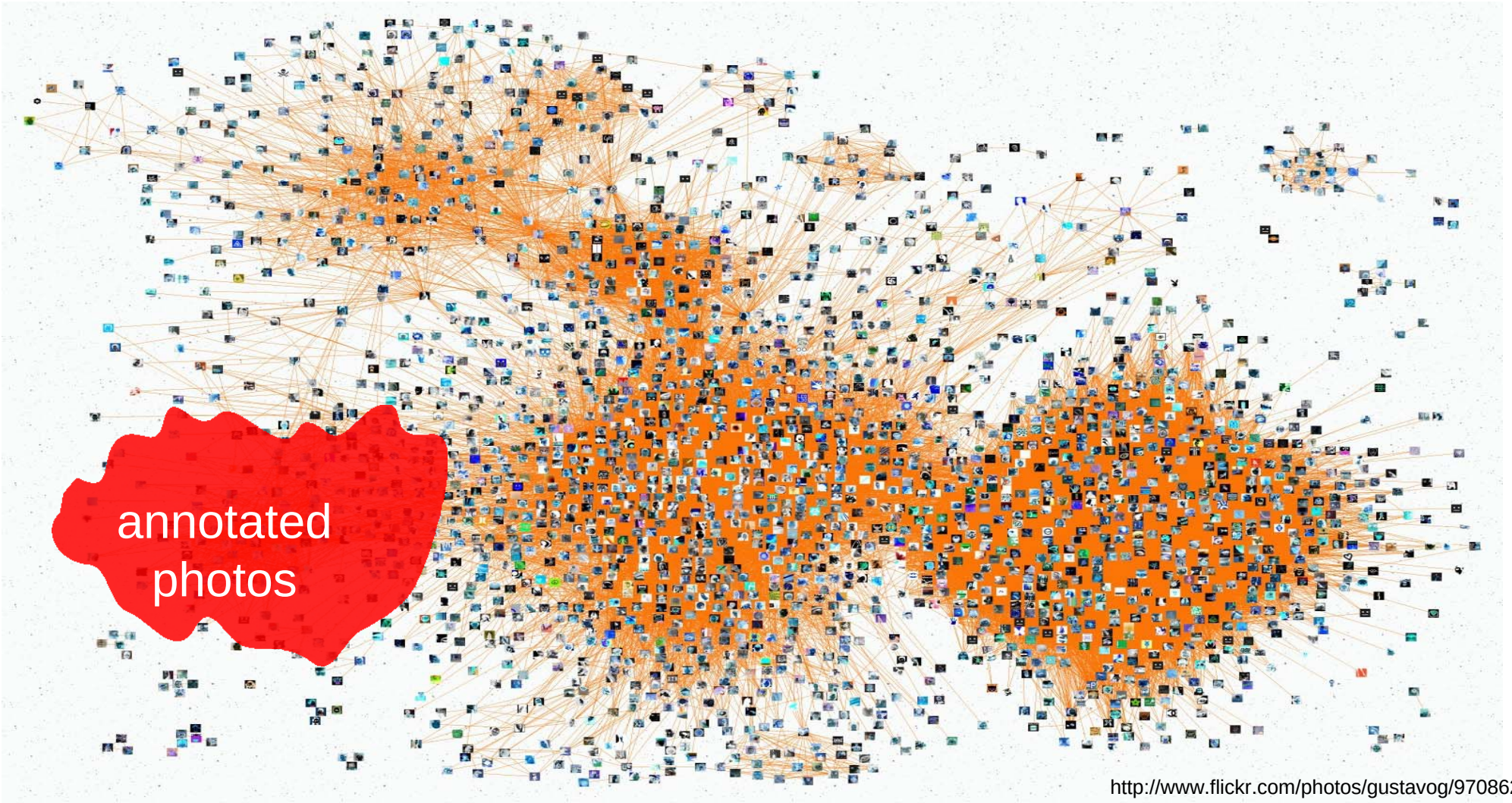
- Ease the process of annotating large photo collections
  - Content-based similarity to propose tags



http://www.flickr.com/photos/gustavog/9708628

Motivation

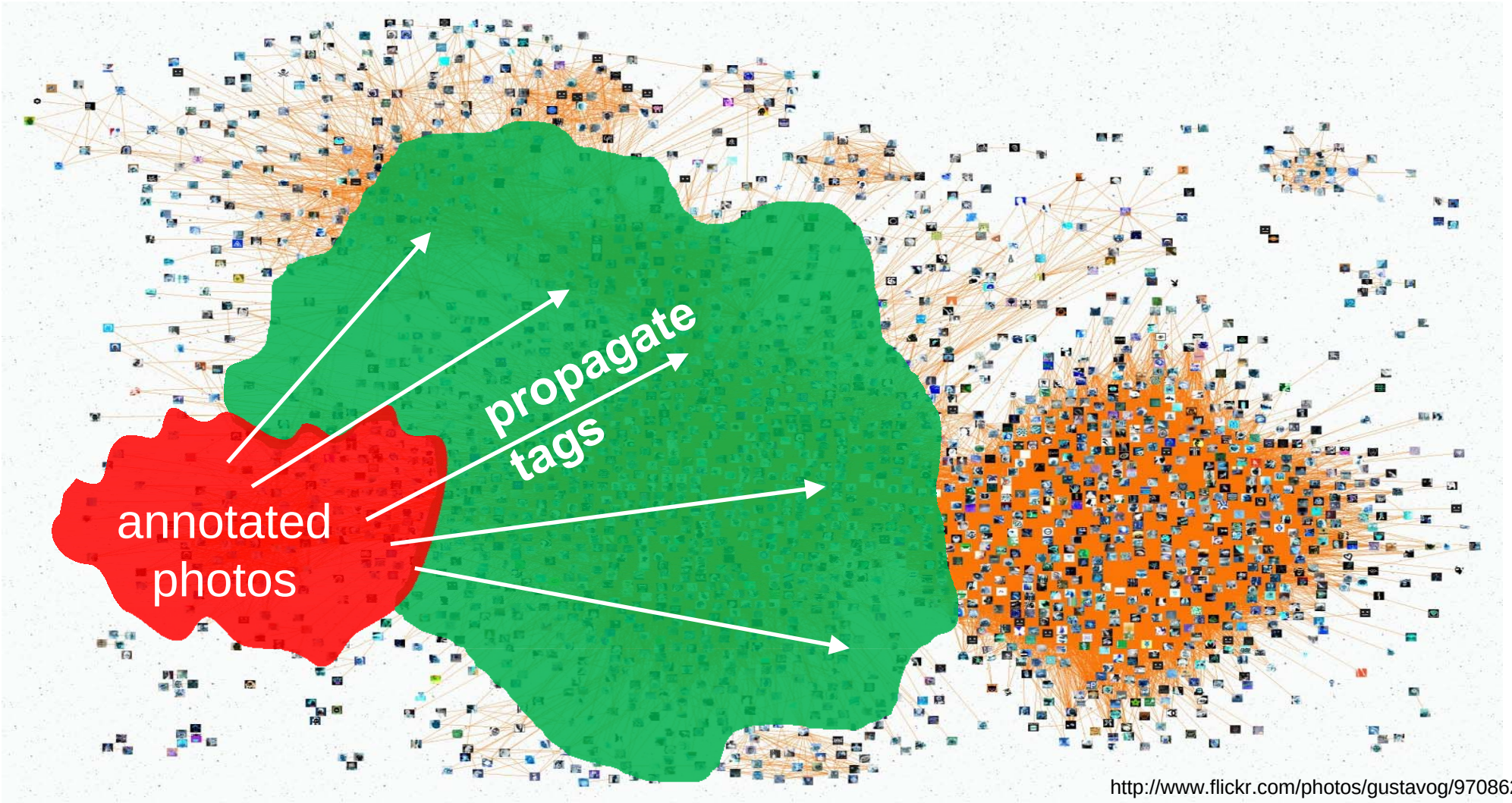
- Ease the process of annotating large photo collections
  - Content-based similarity to propose tags



http://www.flickr.com/photos/gustavog/9708628

Motivation

- Ease the process of annotating large photo collections
  - Content-based similarity to propose tags

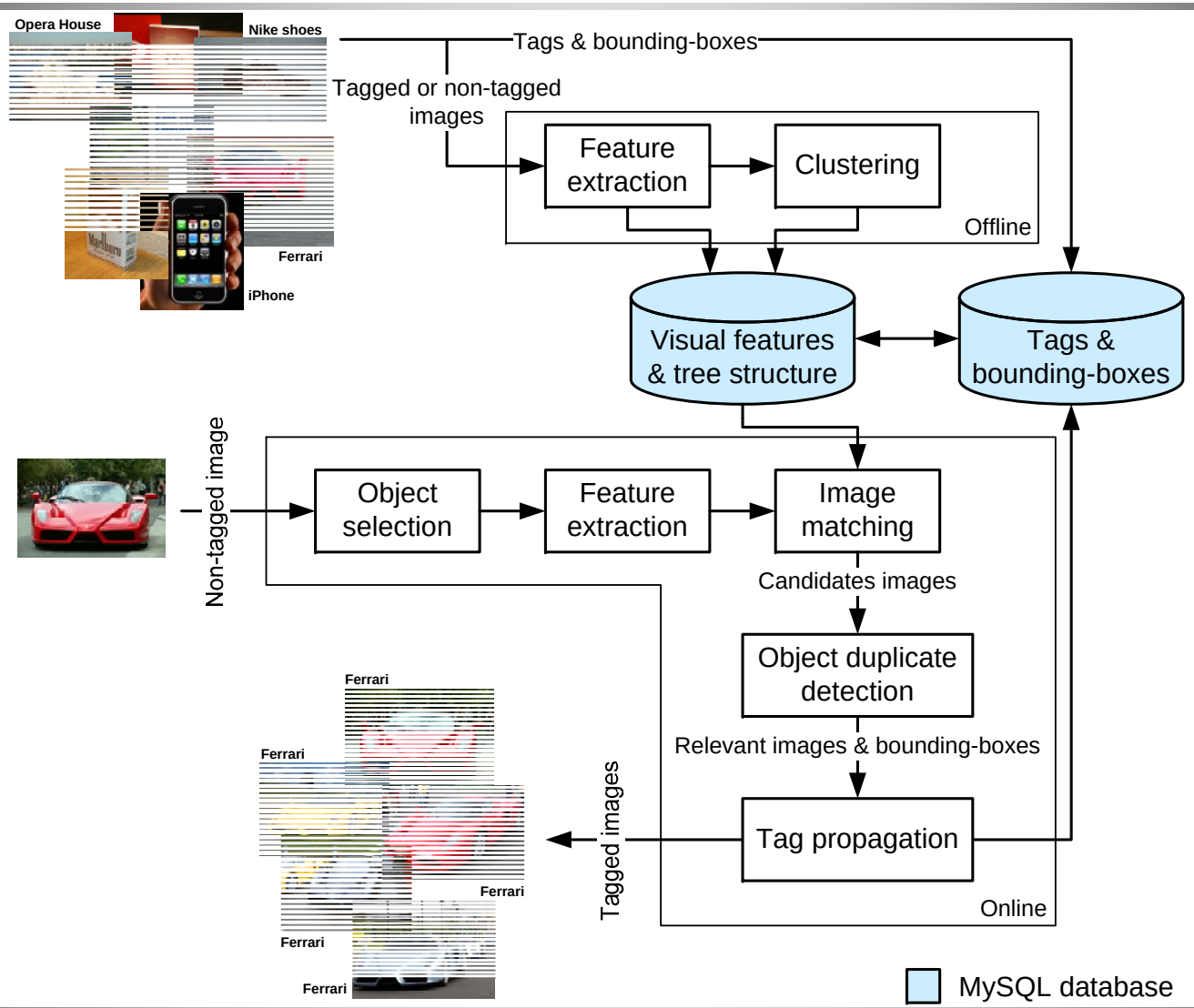


http://www.flickr.com/photos/gustavog/9708628

Tagging + Visual content analysis: State-of-the-art

- Games with a purpose
  - ESP game, TagCaptcha
- Specific applications:
  - People, Locations, Objects, Events
- Features
  - Textual: tags
  - Visual: color, texture, shape, edges, SIFT, SURF, HOG
  - Visual & textual features
- Commercial applications
  - Google Goggles, Kooaba, TinEye
- Our approach
  - Generality
  - Scalability
  - Object-based tagging
  - Interactivity

System overview

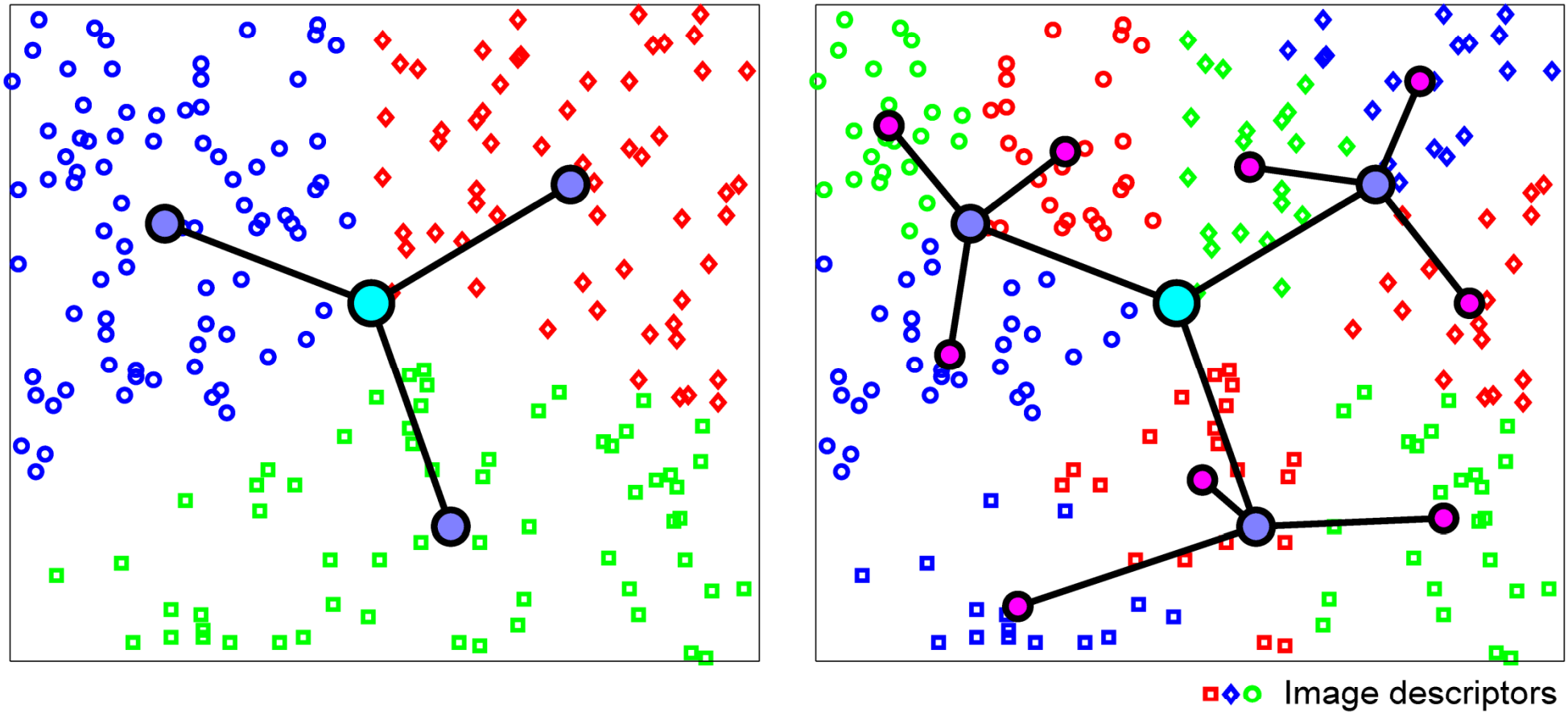


- Feature extraction

- Feature extraction
  - Fast-Hessian detector

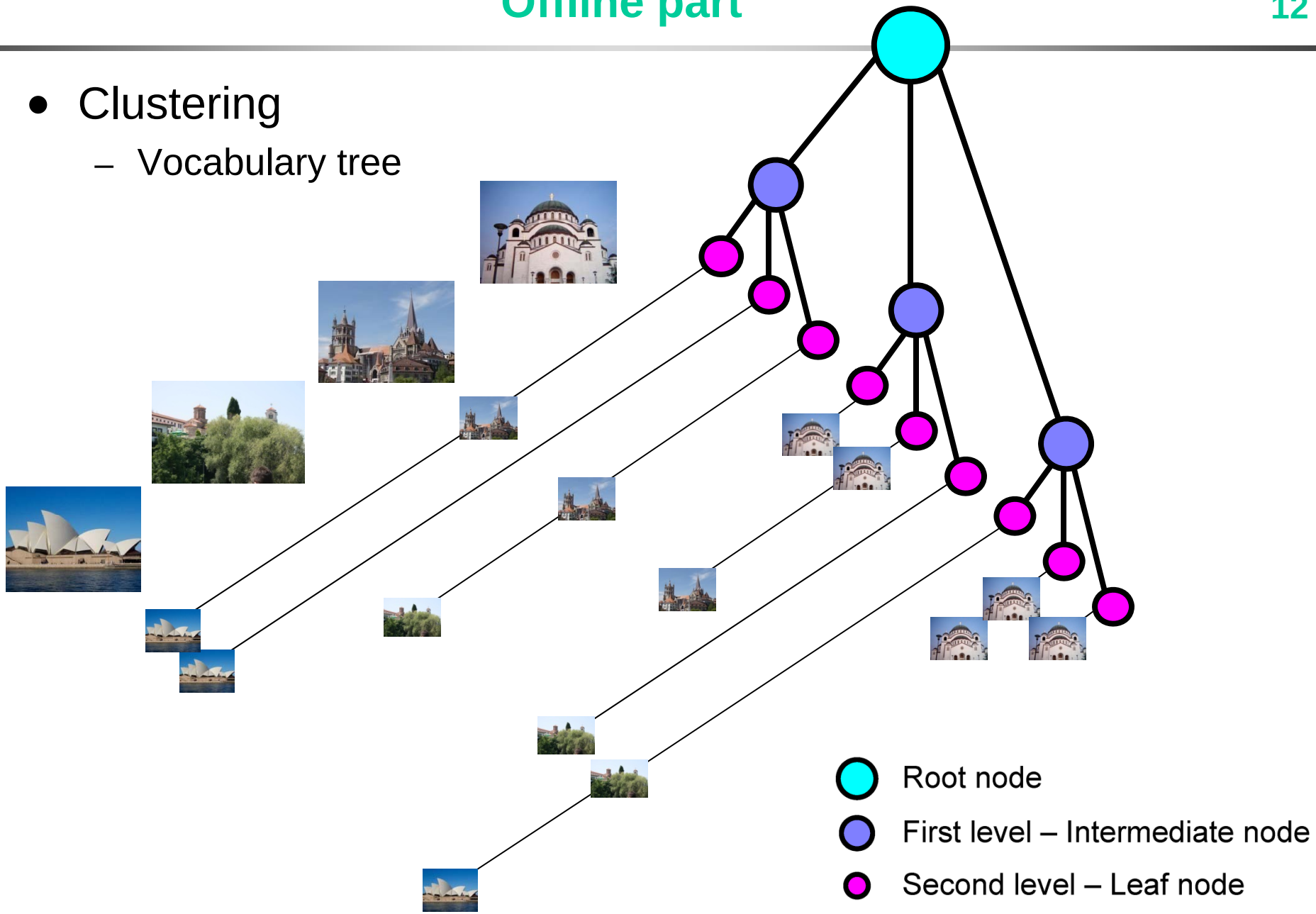
- Feature extraction
  - SURF descriptor

- Clustering
  - Hierarchical k-means



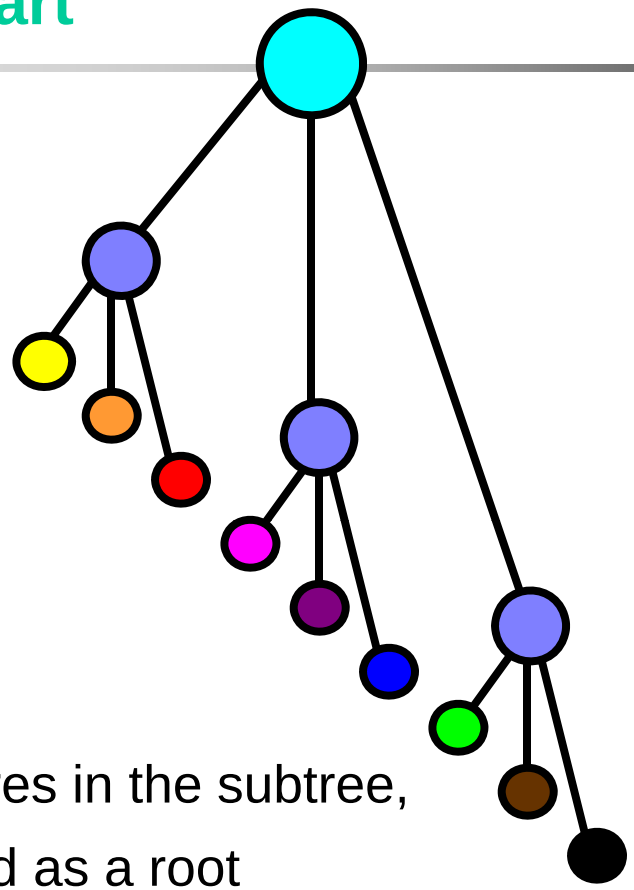
Offline part

- Clustering
  - Vocabulary tree



Offline part

- Clustering
  - TF-IDF weighting scheme



$$d_{ij} = m_{ij} \cdot w_i = \frac{N_{ij}}{\sum_k N_{kj}} \cdot log\left(\frac{N}{N_i}\right)$$

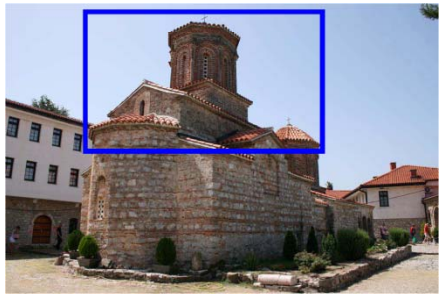
N : # images

N_i : # images which have features in the subtree, if the i-th node is considered as a root

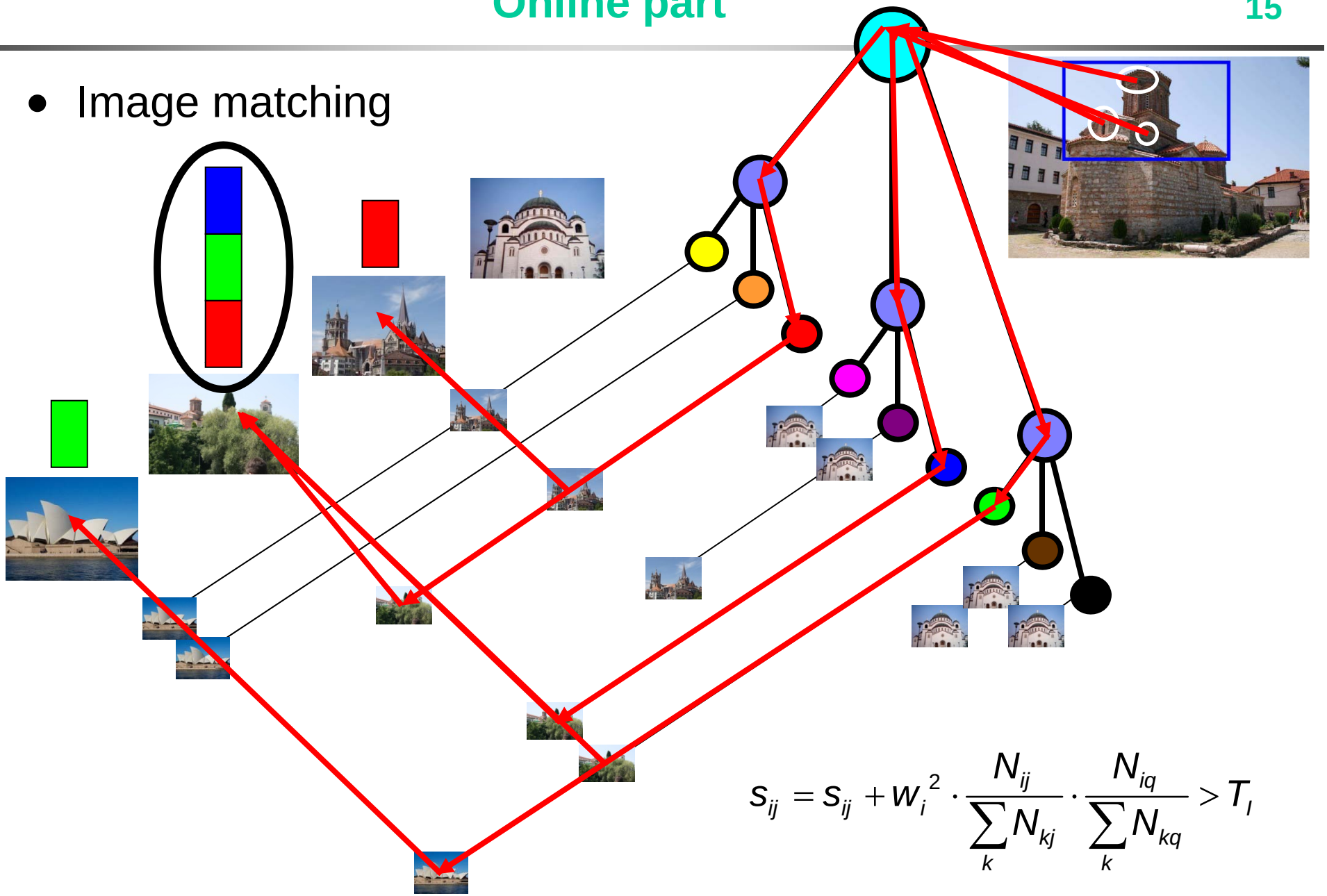
N_{ij} : # occurrences of a visual word i within an image j

$\sum_k N_{kj}$: # occurrences of all features within an image j

- Object selection
  - Bounding-box & tags

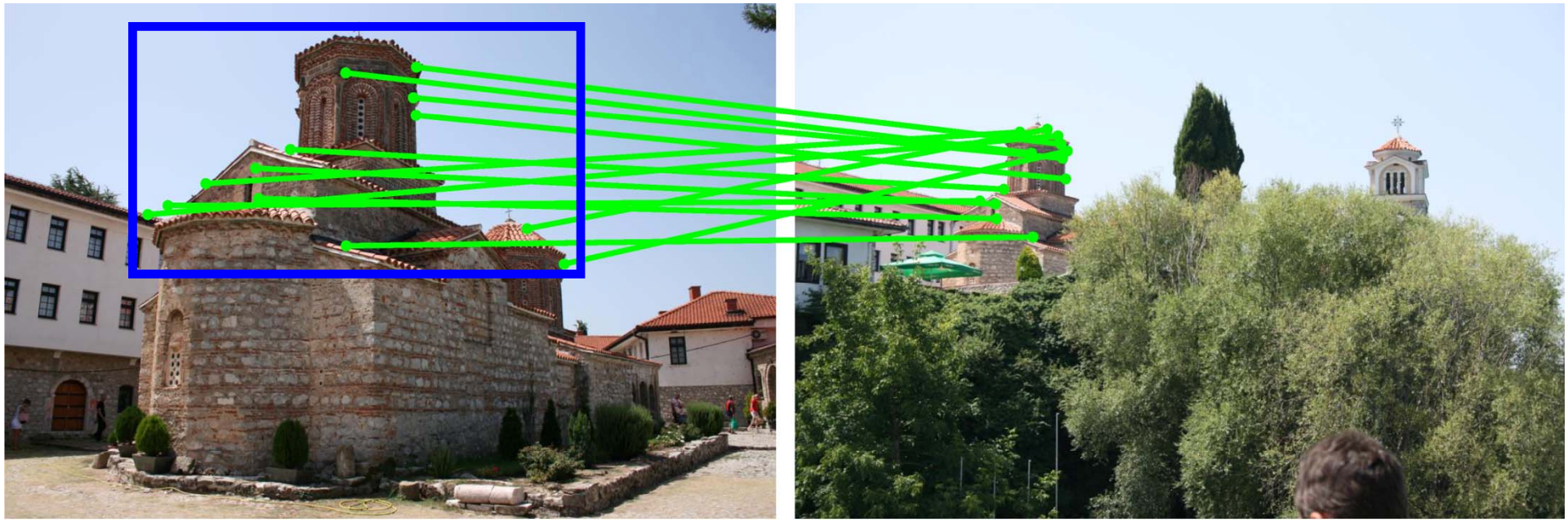


- Image matching

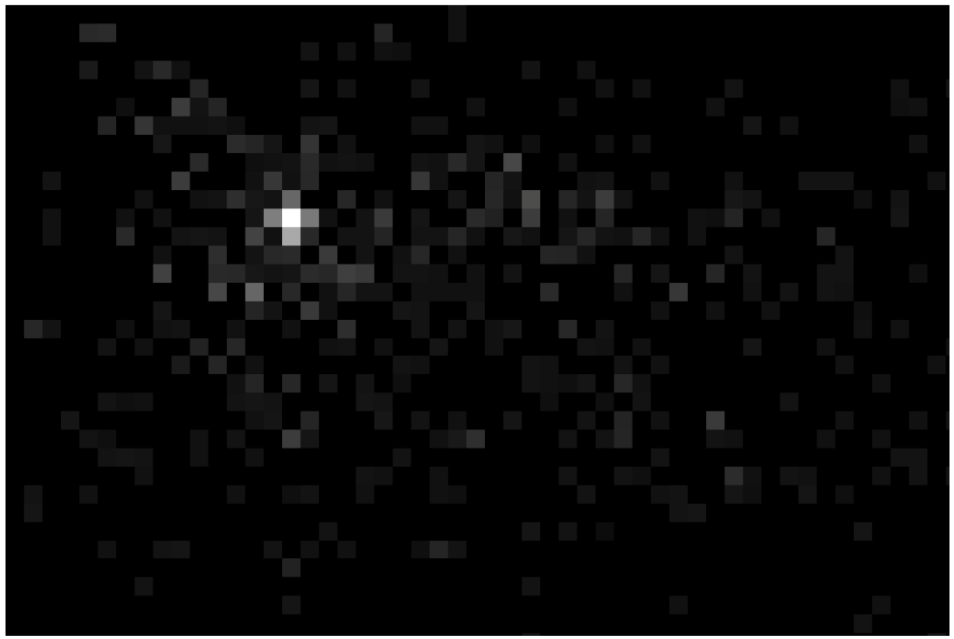


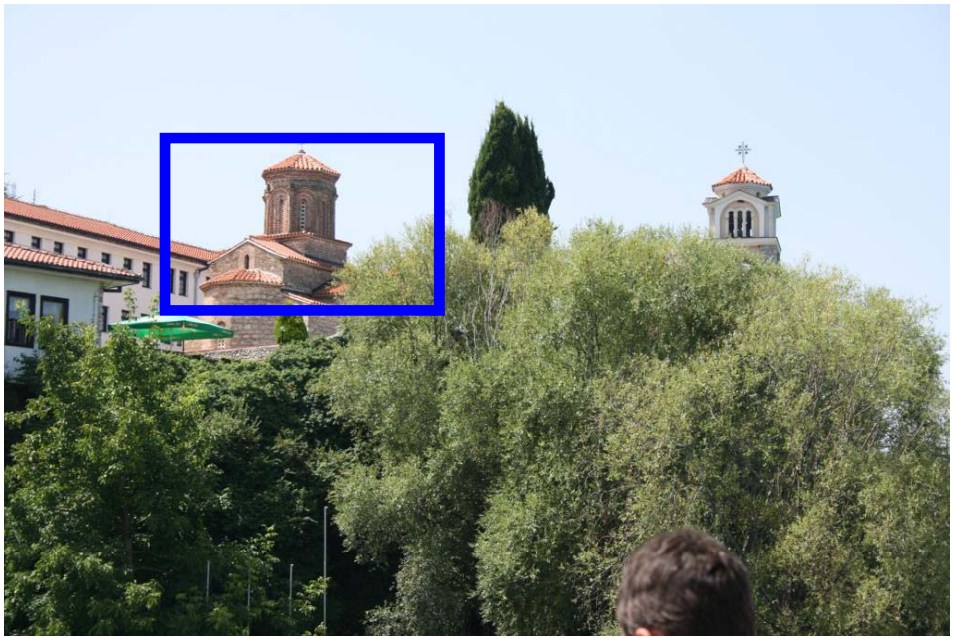
$$s_{ij} = s_{ij} + w_i^2 \cdot \frac{N_{ij}}{\sum_k N_{kj}} \cdot \frac{N_{iq}}{\sum_k N_{kq}} > T_l$$

- Object duplicate detection
  - Feature matching < T_F



- Object duplicate detection
  - Generalized Hough transform



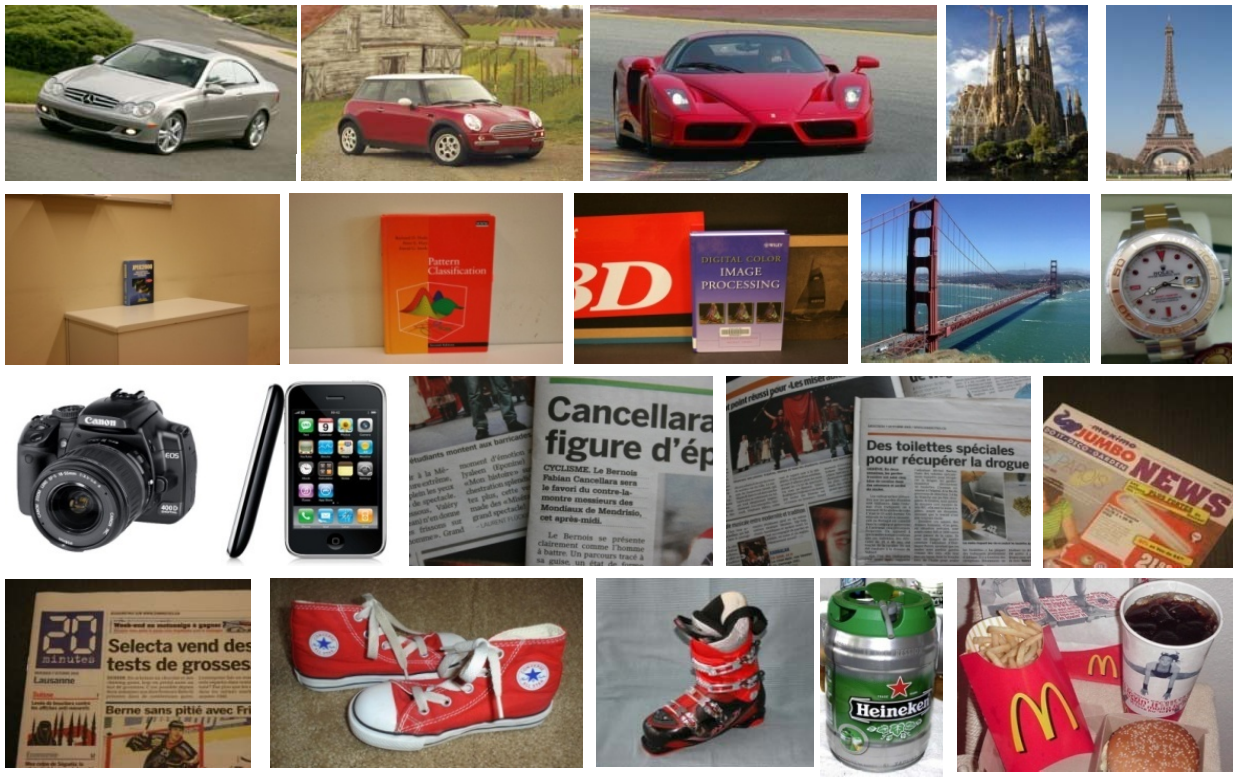


Online part

- Tag propagation

Experiments

- Database
  - Controlled database: 3200 images + ground truth
    8 classes of objects, 20 objects, 20 sample images



Experiments

- Database
  - Distractor database: MIRFLICKR-1M = 1+ million images

Experiments

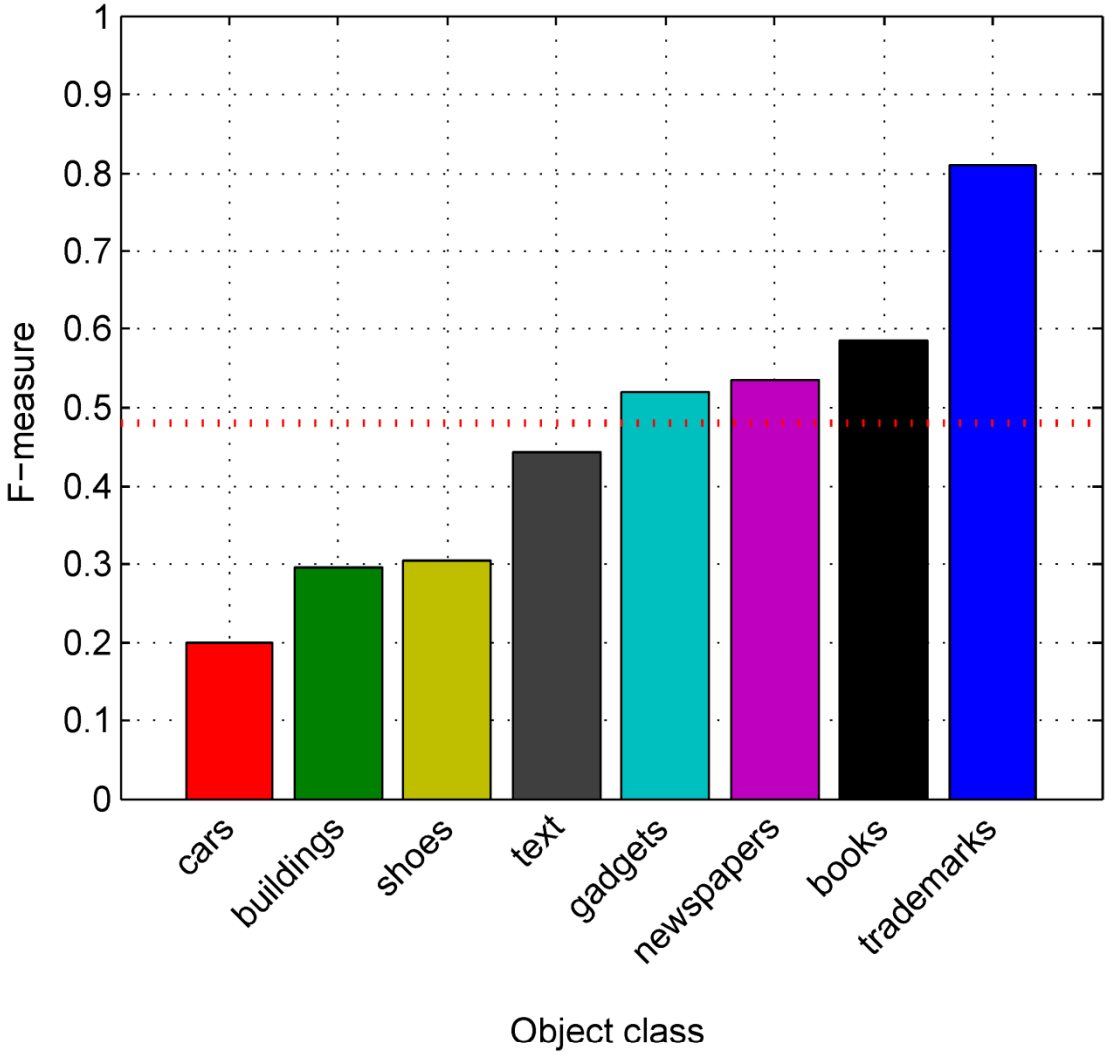
- Evaluation
  - First scenario: only controlled database
    Detection problem

$$P = \frac{TP}{TP + FP} \qquad R = \frac{TP}{TP + FN} \qquad F = \frac{2 \cdot P \cdot R}{P + R}$$

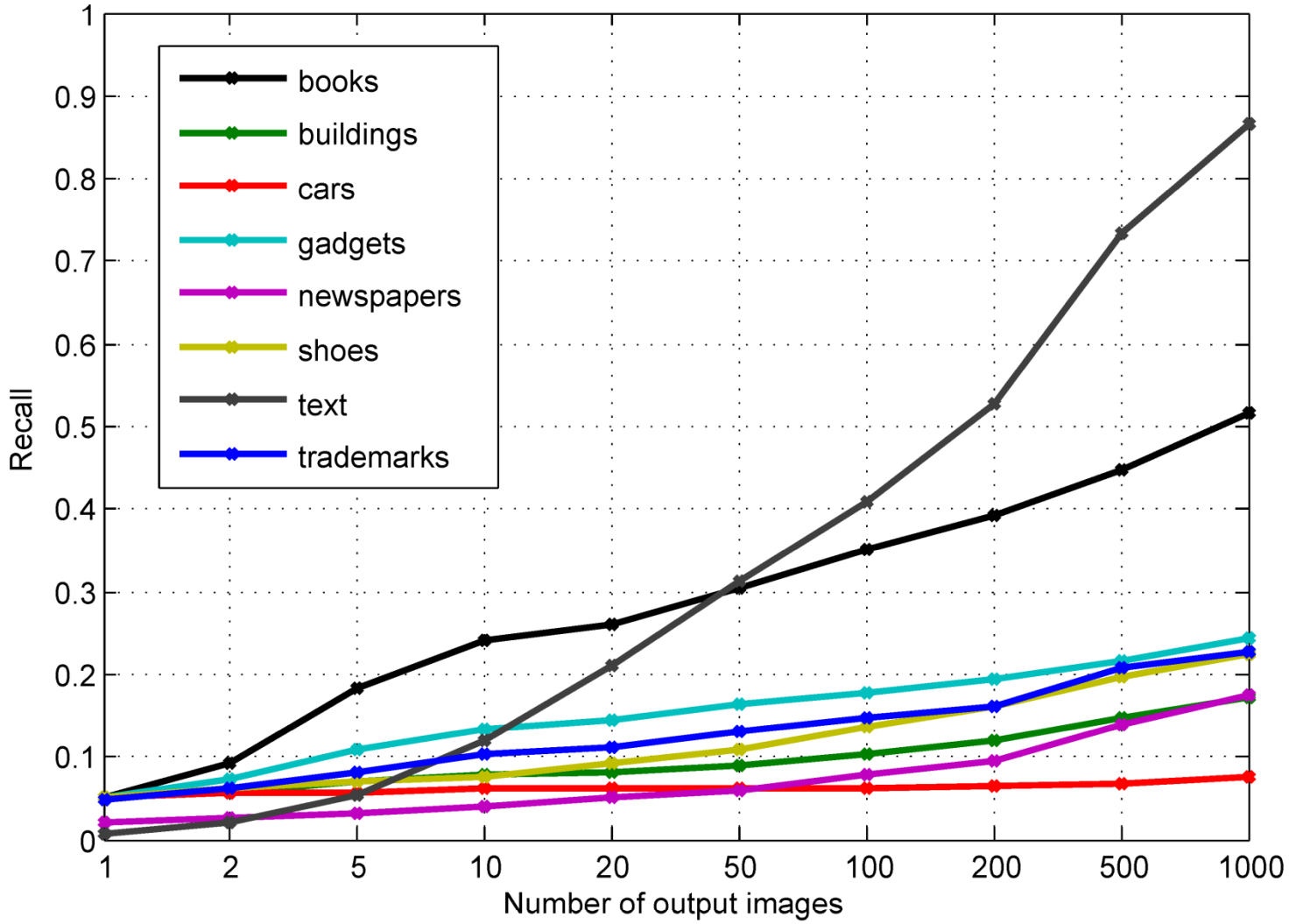
  - Second scenario: controlled + distractor databases
    Ranking problem

$$RankFactor\,(N,q) = \frac{\sum_{i=1}^{N} \omega(o_i,q) \cdot \frac{1}{i}}{\sum_{i=1}^{N} \frac{1}{i}}$$

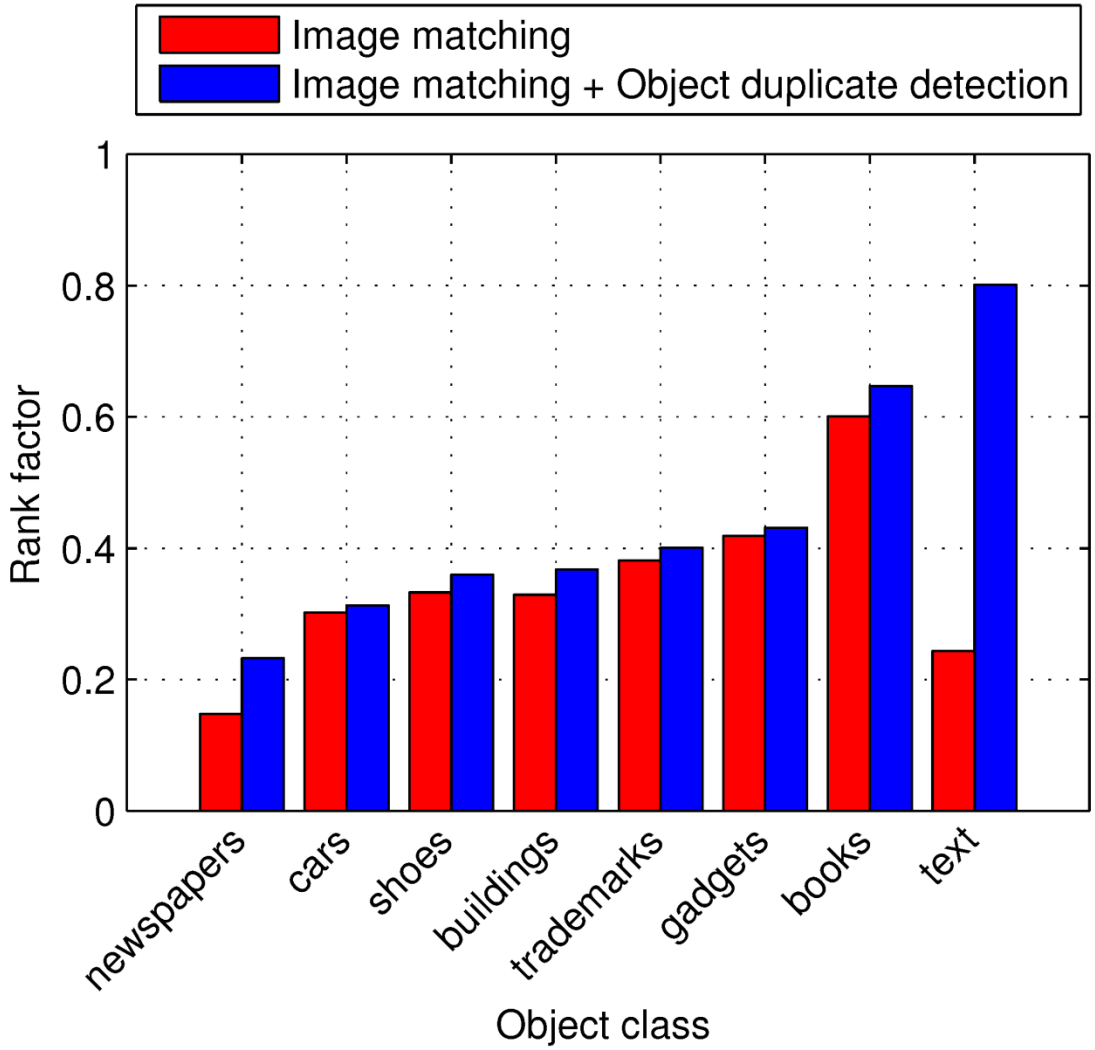
Results – Controlled database



Results – Controlled + Distractor databases

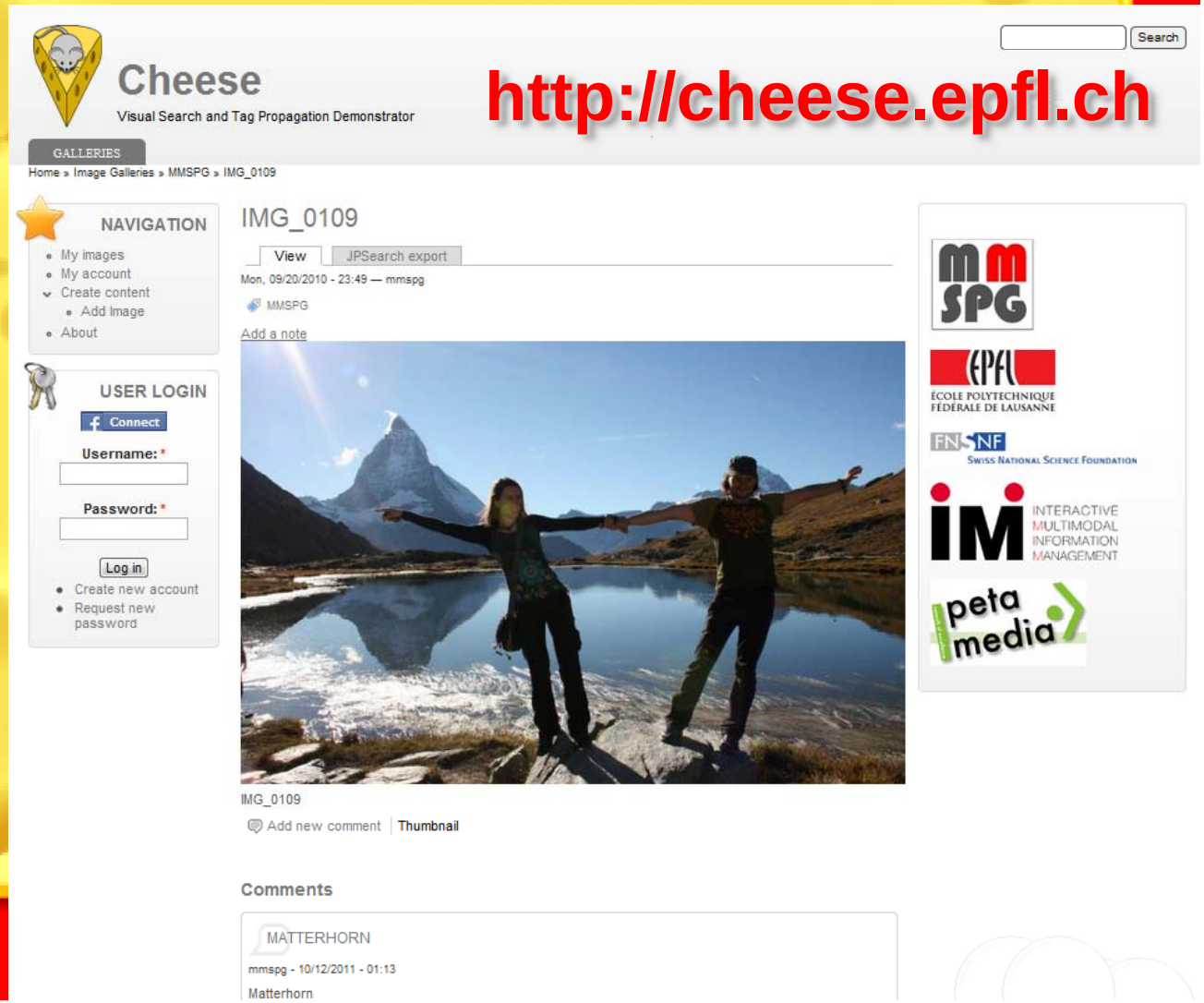


Results – Controlled + Distractor databases



Cheese

- **Interactive**
- **User-friendly**
- **Collecting** and **annotating** image database
- **1+ million images**
- Applications:
  Visual search
  Tag propagation



Conclusion

- An efficient system for **semi-automatic object tagging** in images
- The system performs **image matching and object duplicate detection** in the whole database and returns the search results with images containing similar objects
- The detection **works reliably for salient objects** such as trademarks, books, newspapers, and gadgets

- Future work
  - Support for other classes of objects
  - Tag recommendation