Introduction
- **Tags** are short textual annotations used to describe photos in order to provide meaningful information about them
- The success of **flickr** and **facebook** proves that users are willing to provide tags through manual annotations
- **Tagging** a lot of photos manually is a **time-consuming task**
- We propose an **interactive online platform** capable of performing semi-automatic image annotation and tag recommendation for an extensive online database of images containing various object classes
- We use **object-based tagging**, as the most salient regions in images usually correspond to specific objects
- The system performs **tag recommendation** and **tag propagation**:

Method
- **Feature extraction**: Fast-Hessian detector to detect salient regions, **SURF** – sparse local features robust to arbitrary changes in viewpoints
- **Vocabulary tree**: hierarchical k-means clustering to group the features according to their similarity
- **TF-IDF weighting scheme** (the importance of a visual word is higher if it is contained in only a few images):
  \[ d_i = m_i \cdot w_i = \frac{N_i}{\sum_j N_{ij}} \cdot \log \left( \frac{N}{N_i} \right) \]
  
  \( N \): # of images
  \( N_i \): # of images which have features in the subtree, if the i-th node is considered as a root
  \( N_{ij} \): # of occurrences of a visual word i within an image j
  \( \sum_j N_{ij} \): # of occurrences of all features within an image j
- **Image matching**: select a reduced set of candidate images which are most likely to contain the target object
  \[ s_j = |q - d_j|^2 = 2 \cdot \sum_i q_i \cdot d_{ij} \cdot \frac{q_i \cdot d_{ij}}{d_i} < T_i \]
  
  \( q \): weighting vector for a query image
  \( d_j \): weighting vector for an image j
- **Object duplicate detection**: detect and localize the target object – general Hough transform
- 3-dimensional histogram: each feature votes (IDF weight) for the position (center) and the scale of a bounding-box within the query image

Dataset
- **3200 images**: 8 classes of objects, 20 objects in each class, 20 sample images for each object

Results
- **PR curve**
- **F-measure**

Future direction
- Integration of the system within **social network environment** by allowing different users to use the system

Contact
Ivan Ivanov, e-mail: ivan.ivanov@epfl.ch