Visual Computing Institute
Computer Vision
Prof. Dr. Bastian Leibe

RWTH AACHEN UNIVERSITY

Fixing WTFs - Detecting Image Matches caused by Watermarks, Timestamps, and Frames in Internet Photos

Tobias Weyand, Chih-Yun Tsai, Bastian Leibe   RWTH Aachen University, Germany

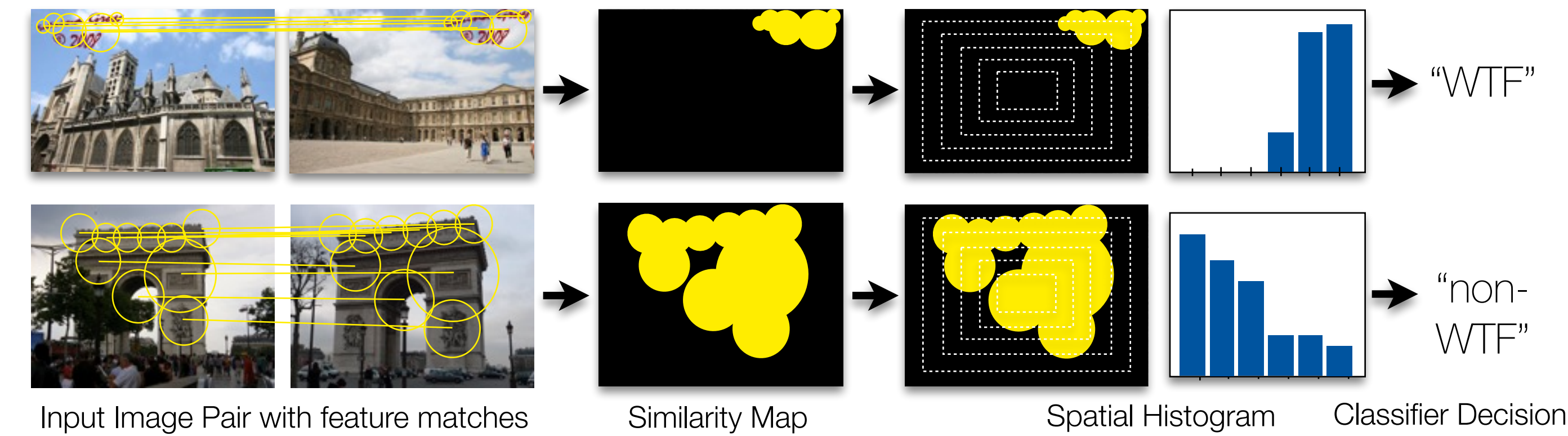Code and dataset available! tiny.cc/wtf

Google

Motivation

- Photos from photo sharing websites often contain Watermarks, Timestamps, or Frames.
- These WTFs can cause false-positive matches betwen photos showing different objects.
- Such WTF matches cause harm in many computer vision applications like image retrieval, image clustering and large-scale structure-from-motion.
- We propose a simple and fast method to fix WTFs by detecting them during matching.

Problems caused by WTFs

Invalid matches

False-positive image retrieval results

Query   Results

Image clusters containing multiple objects   Pseudo-clusters

Method

Key assumptions: WTFs have similar appearance and occur in certain image positions.

Input Image Pair with feature matches   Similarity Map   Spatial Histogram   Classifier Decision

Two-way Matching

Extract similarity maps both ways and combine the classifier decisions.
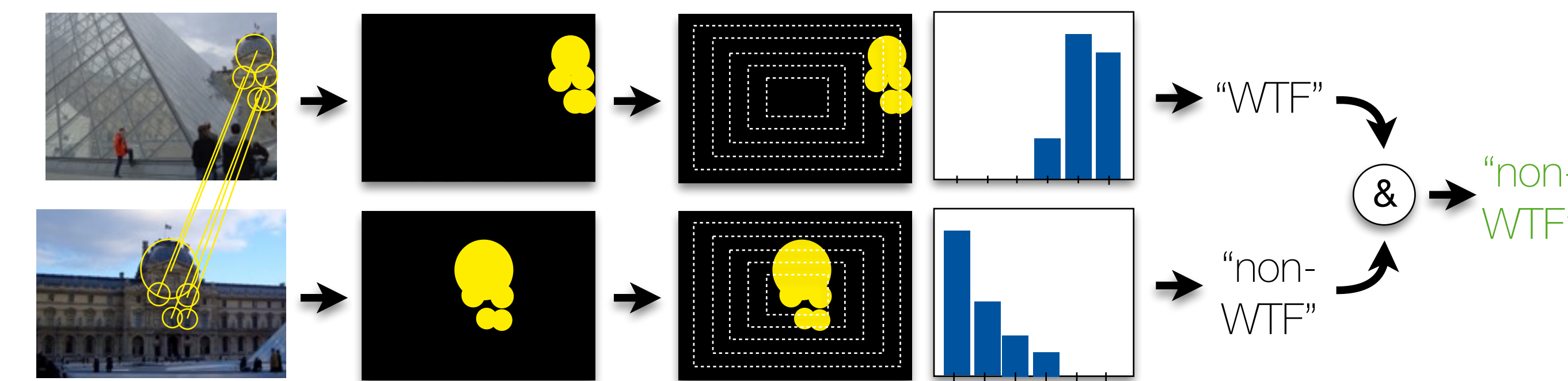
$\Rightarrow$ Avoids false-positive detections and doubles the amount of training data

Feature Design

dist. to border   dist. to center
cake   dartboard

- We tested four histogram shapes. Best: dist. to center, Worst: cake.
- Using the size of the matching region as an additional feature increased detection performance.
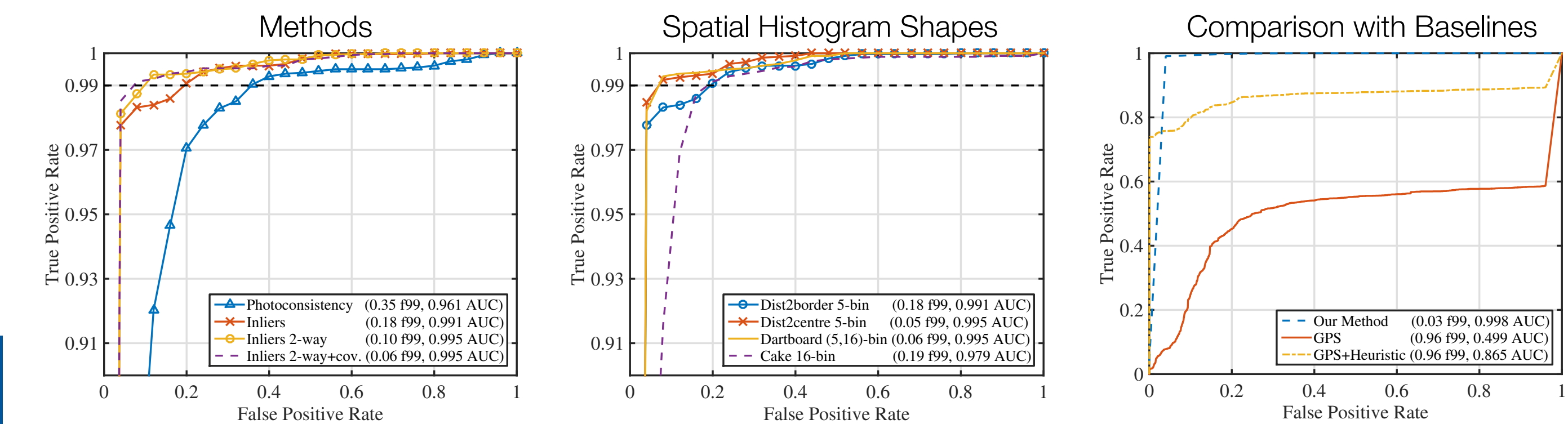$\Rightarrow$ Distinguishing features of WTFs: Distance to image center and size.

Alternative: Photoconsistency-based Similarity Maps

Input Image Pair   Warp using Homography   Compute Photoconsistency   Apply Threshold   Morphological Opening
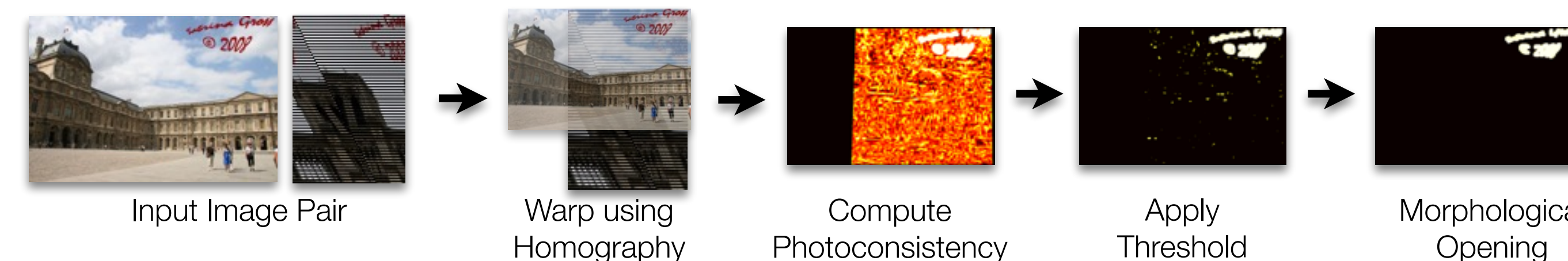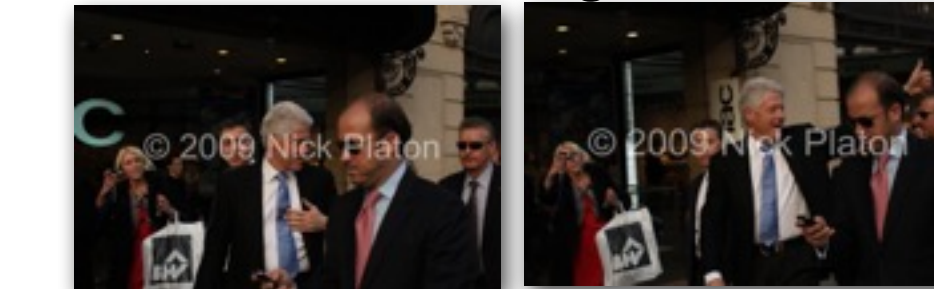
$\Rightarrow$ Worse precision due to misdetections in uniform image regions and 10x slower.
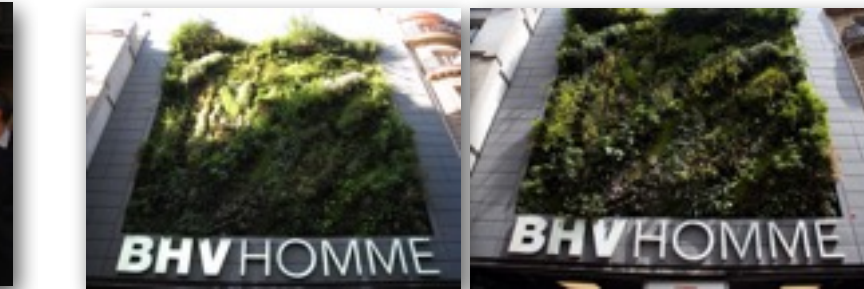
Results

Dataset: 36,240 image pairs from Flickr and Panoramio, 10% WTFs, 90% non-WTFs
Evaluation: Binary classification with 5-fold cross-validation. Classifier: AdaBoost
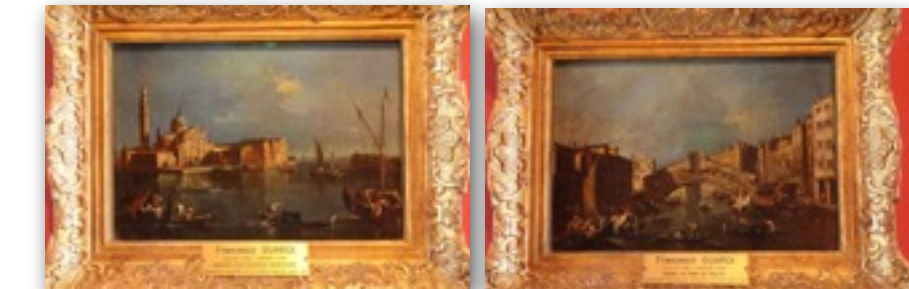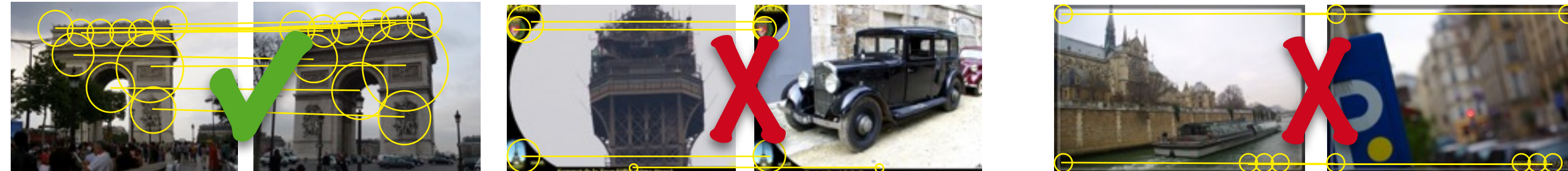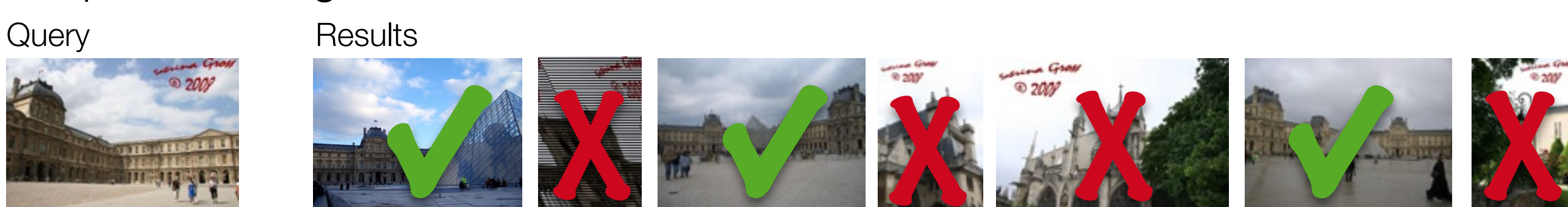Performance Measures: AUC (Area under the ROC curve), f99 (fpr at 99% tpr)

Methods

- Photoconsistency (0.35 f99, 0.961 AUC)
- Inliers (0.18 f99, 0.991 AUC)
- Inliers 2-way (0.06 f99, 0.995 AUC)
- Inliers 2-way+cov. (0.06 f99, 0.995 AUC)

Spatial Histogram Shapes

- Dist2border 5-bin (0.18 f99, 0.991 AUC)
- Dist2centre 5-bin (0.06 f99, 0.995 AUC)
- Dartboard (5,16)-bin (0.06 f99, 0.995 AUC)
- Cake 16-bin (0.19 f99, 0.979 AUC)

Comparison with Baselines

- Our Method (0.03 f99, 0.998 AUC)
- GPS (0.96 f99, 0.499 AUC)
- GPS+Heuristic (0.96 f99, 0.865 AUC)

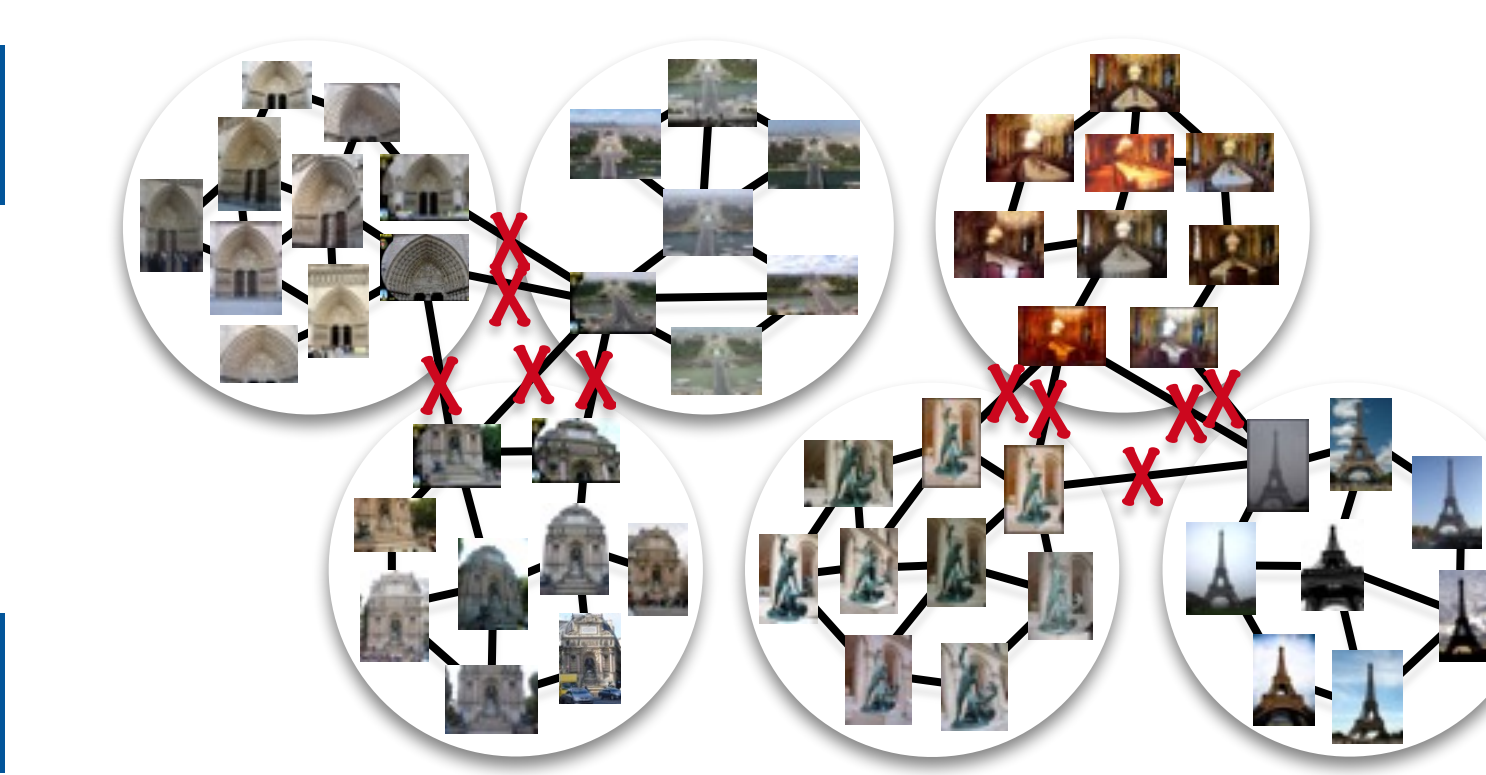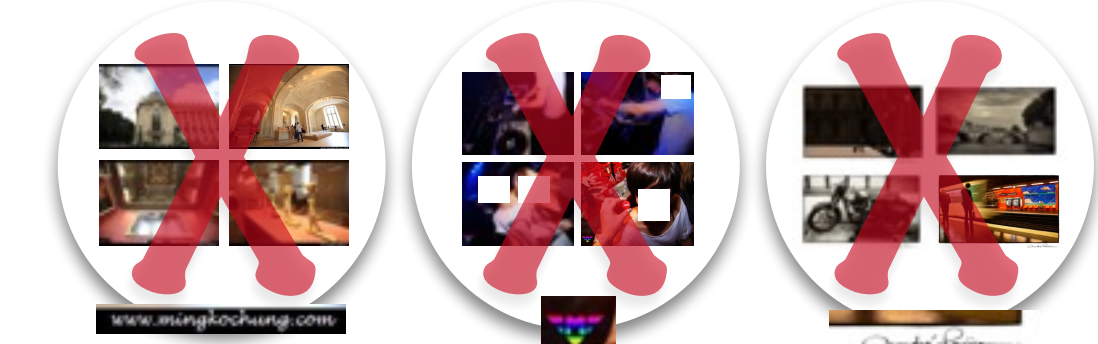False-Negative   False-Positive   False-Positive?

Application to Clustering

Setup: Iconoid Shift clustering [Weyand11ICCV] of Paris500k [Weyand09RMLE] and Oxford105k [Philbin07CVPR].
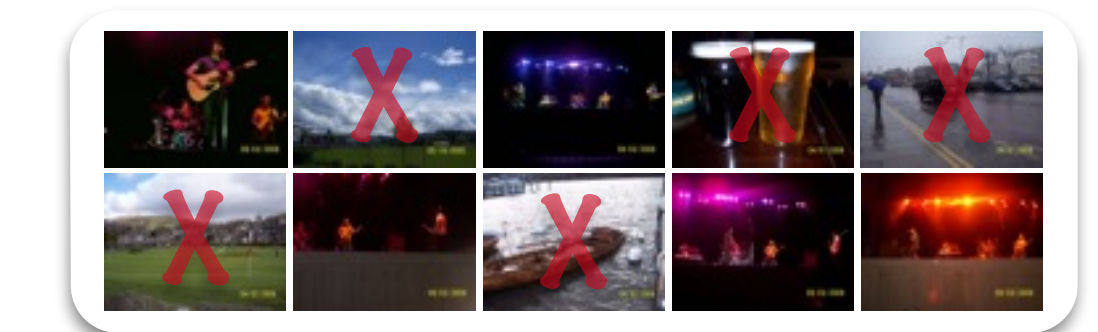
Clusters with multiple objects were split.   Pseudo-clusters were removed.

Polluted clusters were cleaned.

Conclusion

- WTFs in Internet photos cause false-positive matches that harm many applications.
- We detect WTF matches based on the image positions of matching local features.
- Our method achieves high accuracy and fixes many problems in clustering.
- The code and dataset are publicly available at: tiny.cc/wtf