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

Motivation

- Photos from photo sharing websites often contain Watermarks, Timestamps, or Frames.
- These WTFs can cause false-positive matches between photos showing different objects.
- Such WTFs can 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.

Methods

- Key assumptions: WTFs have similar appearance and occur in certain image positions.
- We detect WTF matches based on the image positions of matching local features.

Feature Design

- 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.
- Distinguishing features of WTFs: Distance to image center and size.

Application to Clustering

Setup: Iconoid Shift clustering of Paris500k and Oxford105k.
- Clusters with multiple objects were split.
- Pseudo-clusters were removed.
- Polluted clusters were cleaned.

Results

- Dataset: 36,240 image pairs from Flickr and Panoramio, 10% WTFs, 90% non-WTFs

Performance Measures:

- True Positive Rate
- False Positive Rate
- Inliers 2-way+cov. (0.99, 0.995 AUC)
- Inliers 2-way (0.10 f99, 0.995 AUC)
- Inliers (0.18 f99, 0.991 AUC)
- Photoconsistency (0.35 f99, 0.961 AUC)

Comparison with Baselines

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

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