Long-Term Tracking Through Failure Cases

K. Lebeda[1], S. Hadfield[1], J. Matas[2], R. Bowden[1]

{k.lebeda, s.hadfield, r.bowden}@surrey.ac.uk, matas@cmp.felk.cvut.cz

[1]Centre for Vision, Speech and Signal Processing, University of Surrey
[2]Center for Machine Perception, Czech Technical University in Prague

LT-FLO track

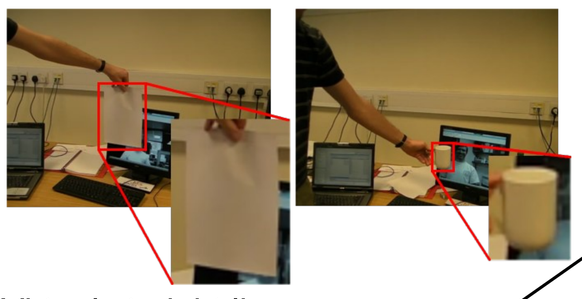
Long-Term FeatureLess Object tracker

MOTIVATION

Conventional Tracking

uses features (blobs, corners,...) that are distinct (Harris condition) and provide point-to-point correspondence.

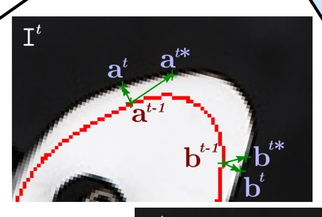
But what if there's not enough of these? Or if the majority lie on the object contour ⇒ influenced by the background.



Virtual straight lines – tangents – are present, but they have the Aperture Problem!

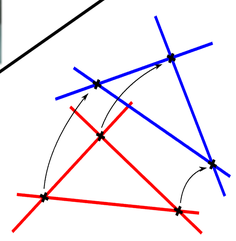
APPROACH

Correspondence of Edge Tangents

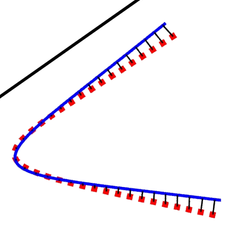


– edge point $\mathbf{a}^{t*}$, the true correspondence to $\mathbf{a}^{t-1}$ is difficult to find
– instead, edge point $\mathbf{a}^t$ is found by 1D search
– however, tangents at $\mathbf{a}^{t*}$ and $\mathbf{a}^t$ are the same ($\mathbf{k}^t$)
– note that $\mathbf{k}^t$ corresponds to $\mathbf{k}^{t-1}$

⇒ correspondence of intersection points **c** gives the correct transformation
– 1D perpendicular search; similarity of gradient angle, position, appearance



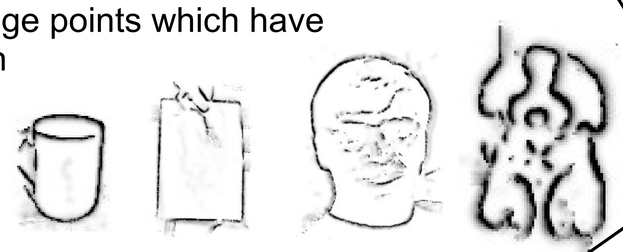
Frame-to-frame Transformation



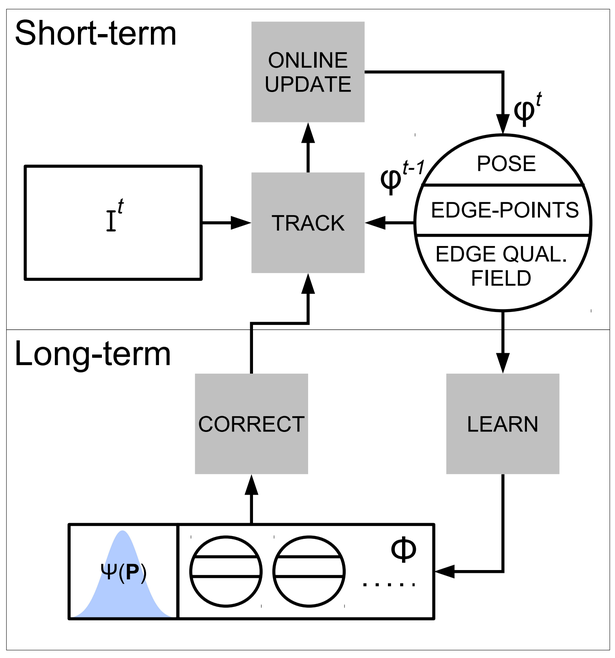
– similarity (T+R+S); estimated by LO-RANSAC
– sample: line triplet, using intersection points
– maximizing inlier count combined with image *evidence* (modified Chamfer distance, gives the fit of points to edges in the second image)

Online learning of reliable (tangent) points

– store a quality field of edge points which have predicted correct motion
– fitting points to the map decreases drift



Frame-to-Frame Tracking



Long-Term Tracking

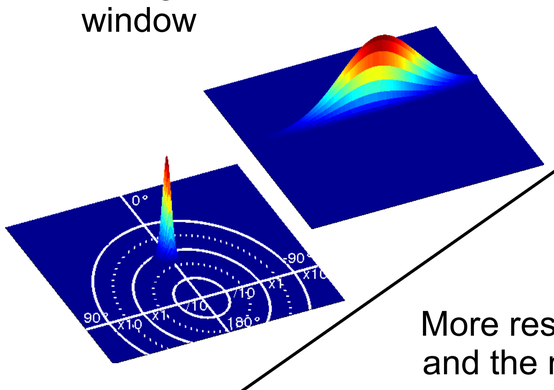
STATE =
ESTIMATED POSE
+ SET OF EDGE POINTS
+ LEARNED QUALITY FIELD

Corrections by Bank of States

– state of short-term tracker is updated online
– several previous states are stored in a *bank* of states
– when confidence in the current solution is low, a *correction* is performed, using the states from the bank
– tracking is restarted from the previous frame, the best solution is retained
– states with high confidence are stored for later corrections

Global Corrections

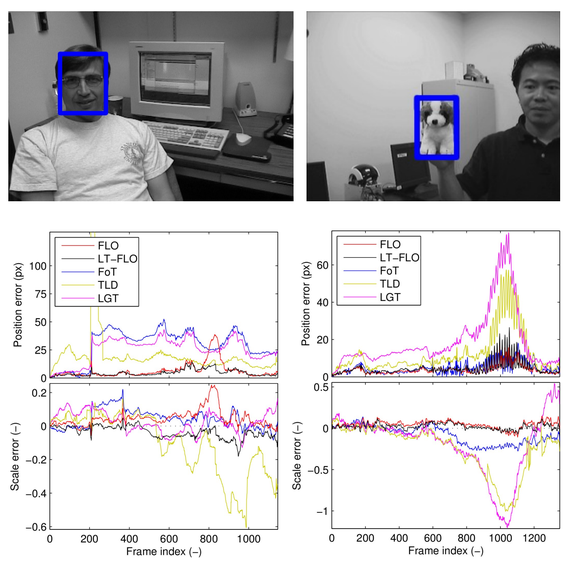
– executed after full occlusions
– analogous to local corrections
– instead of the last pose, random ones are drawn from observed probability distribution
– semi-global, more efficient than a sliding window



More results and the new sequences are available at:
http://cvssp.org/Personal/KarelLebeda/

RESULTS

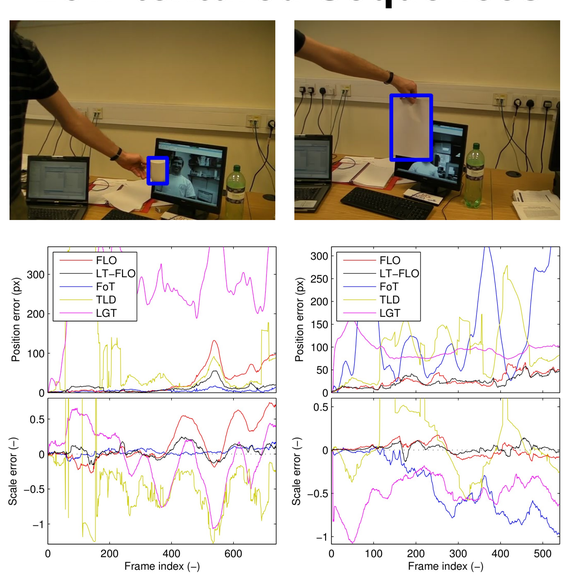
Standard Sequences



Long-Term Sequences

– succesful redetections, very low drift
– achieved mean F-measure 0.49, compared to 0.36 of TLD

Low-textured Sequences



Conclusions

– edge-based tracker, low reliance on texture and brightness
– robust to drift, redetects after full occlusions
– on high-textured scenes comparable to SOTA
– on low-textured scenes superior
– successfully tracks a sequence* of almost 30,000 frames!

This work was presented at the ICCV workshop on Visual Object Tracking Challenge, Sydney, Australia on 2nd December 2013. The authors were supported by the following projects: GACR P103/12/G084 and EPSRC EP/I011811/1.