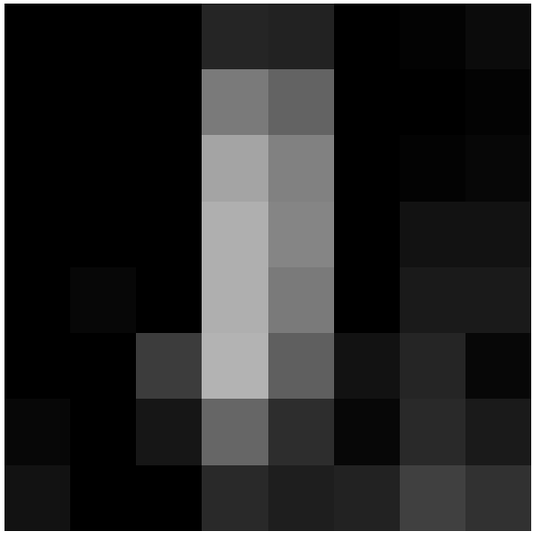
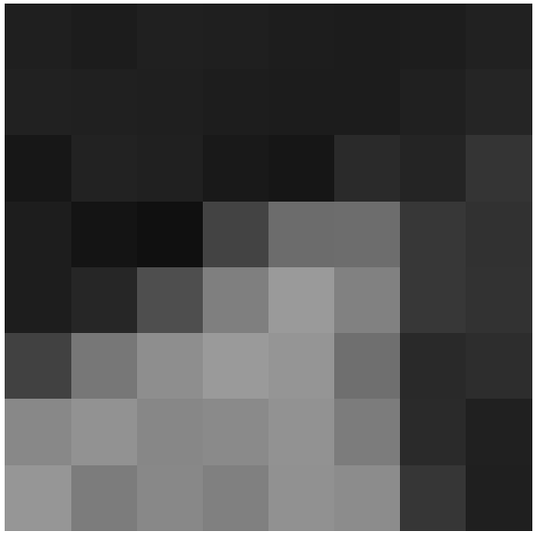
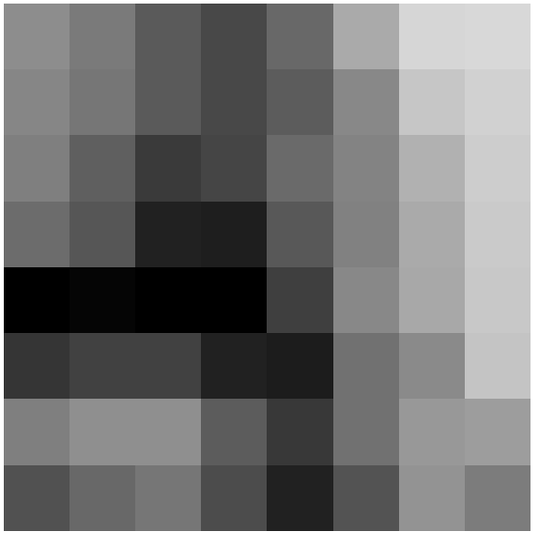
Tracking the Untrackable: How to Track When Your Object Is Featureless

Karel Lebeda[1], Jiří Matas[1], Richard Bowden[2]

[1]Center for Machine Perception, Czech Technical University
http://cmp.felk.cvut.cz/
[2]Centre for Vision, Speech and Signal Processing, University of Surrey
http://surrey.ac.uk/cvssp/

Conventional Tracking

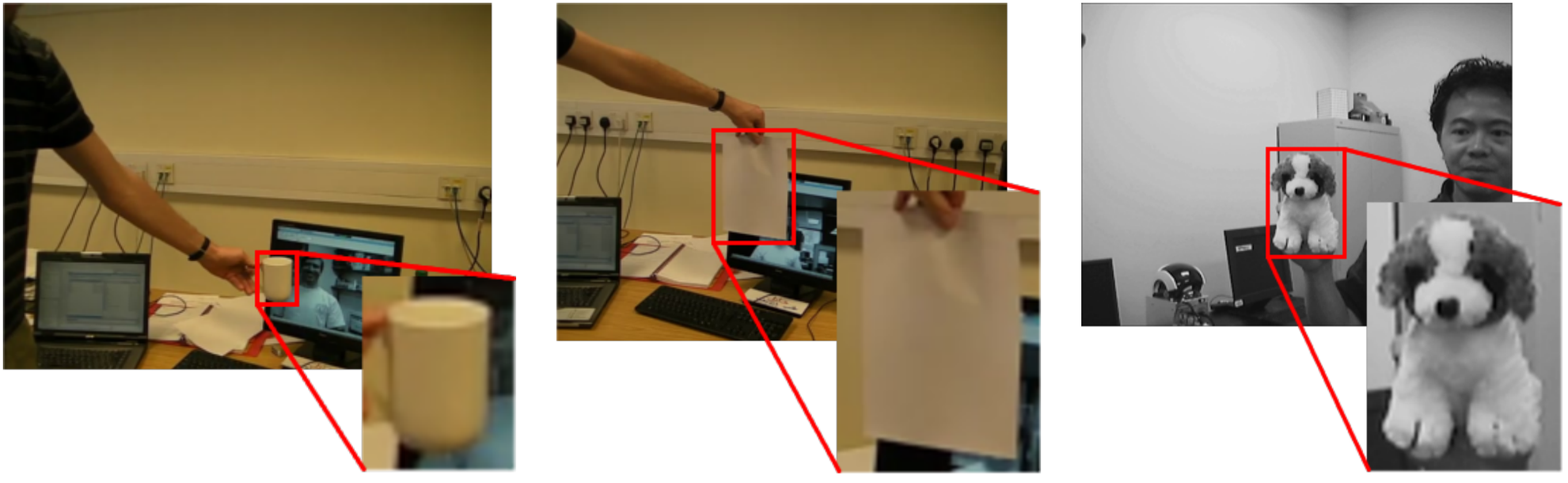




- well-localized features (blobs, corners,...)
- distinguishable from their neighbors

Moravec, H., Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover, Tech Report CMU-RI-TR-3,

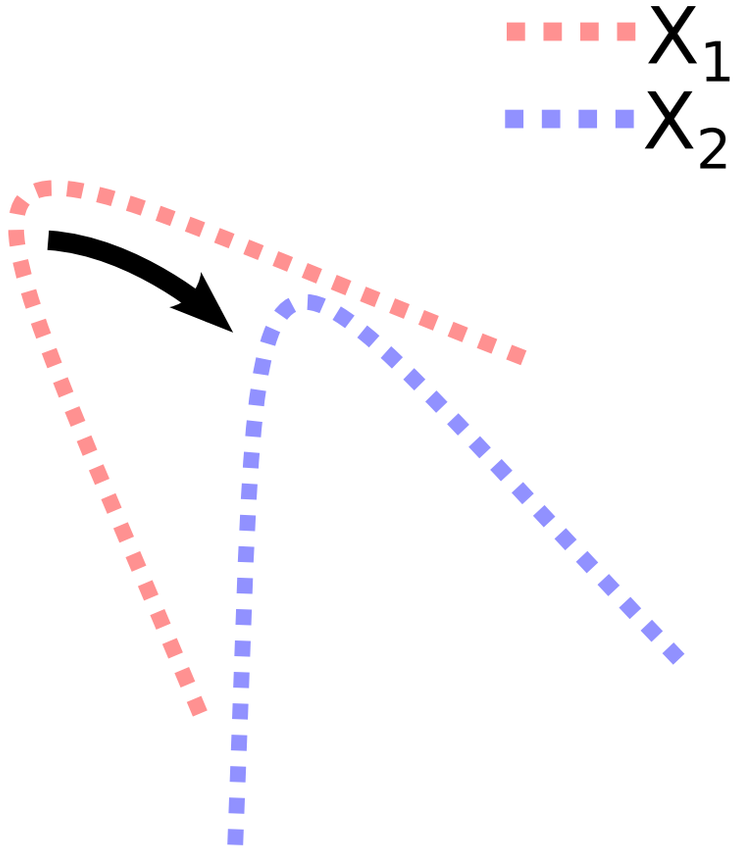
Carnegie-Mellon University, Robotics Institute, 1980.

Low-textured Objects?

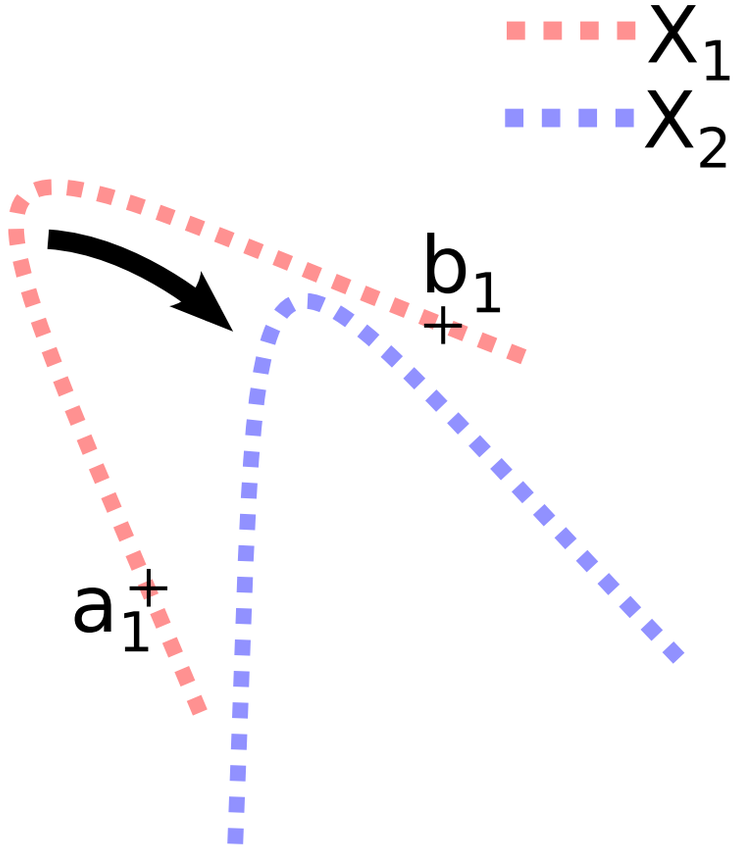


- some real-world objects have little of such features
- or these features lie on the object boundaries and are affected by the background

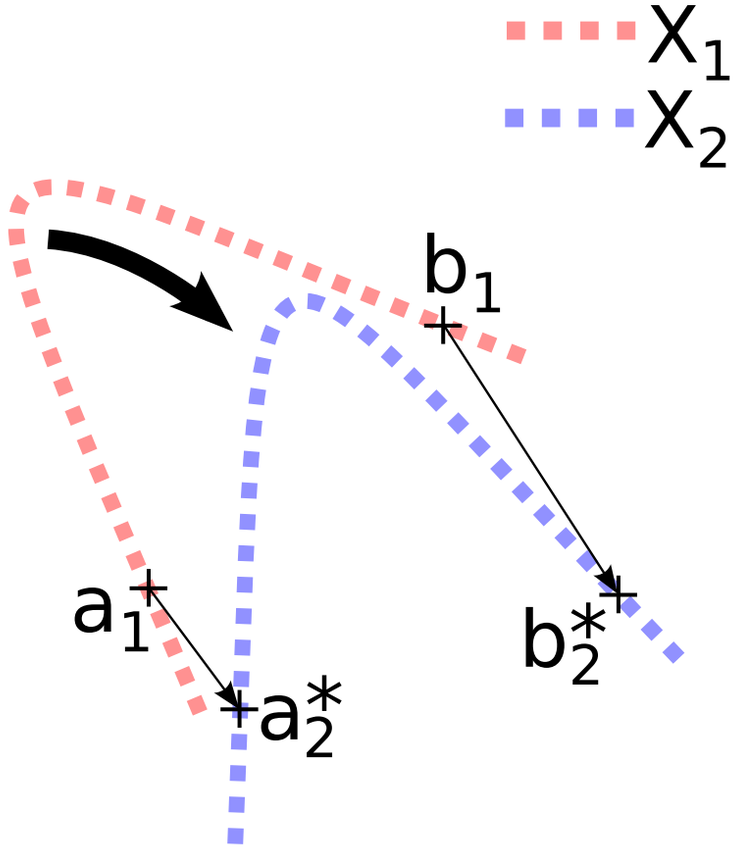
The Aperture Problem



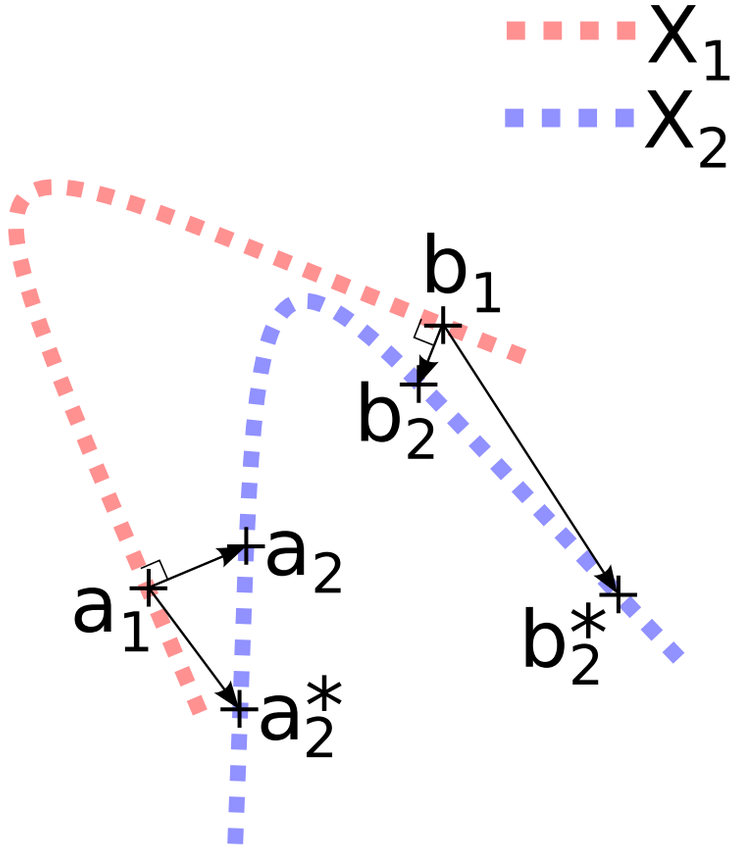
The Aperture Problem



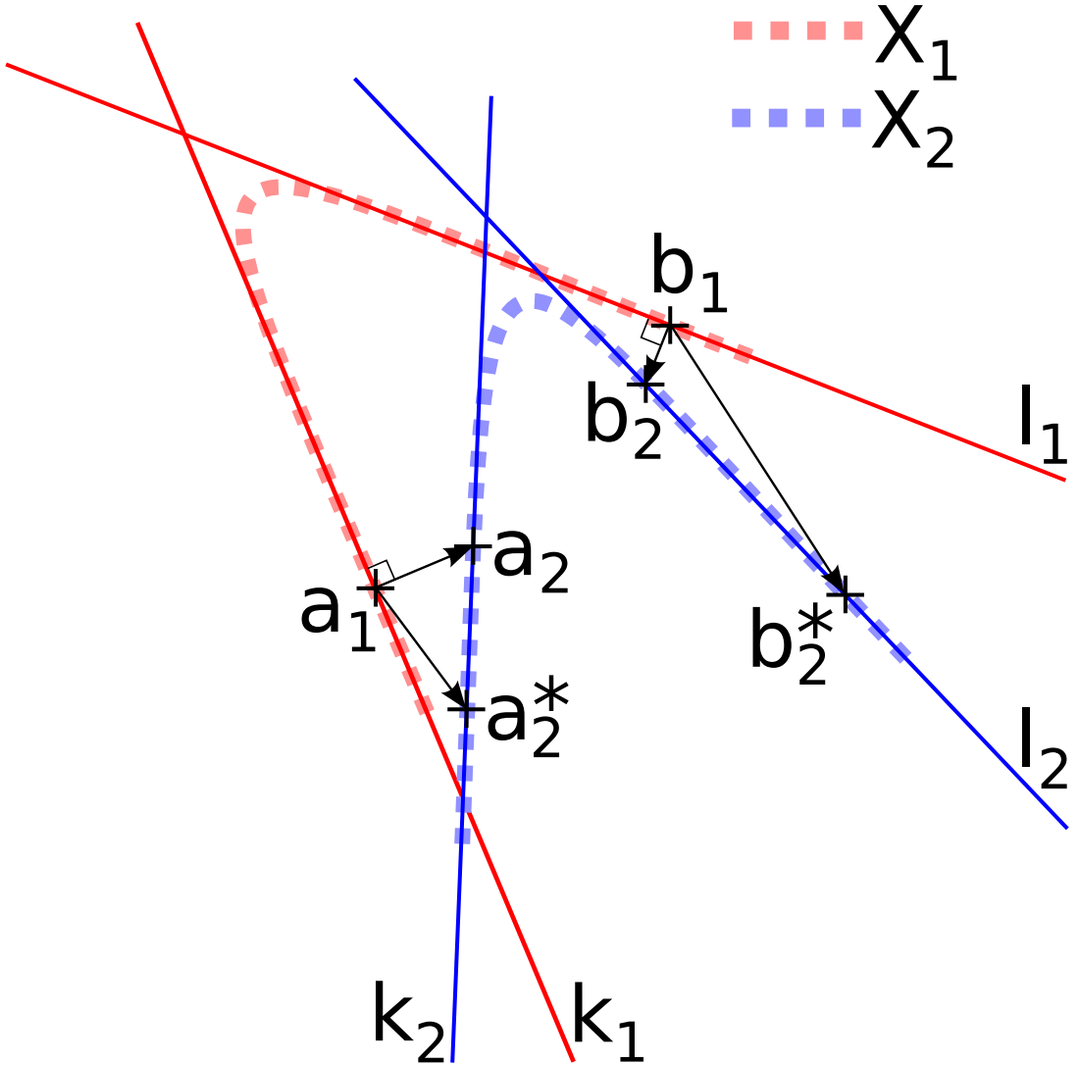
The Aperture Problem



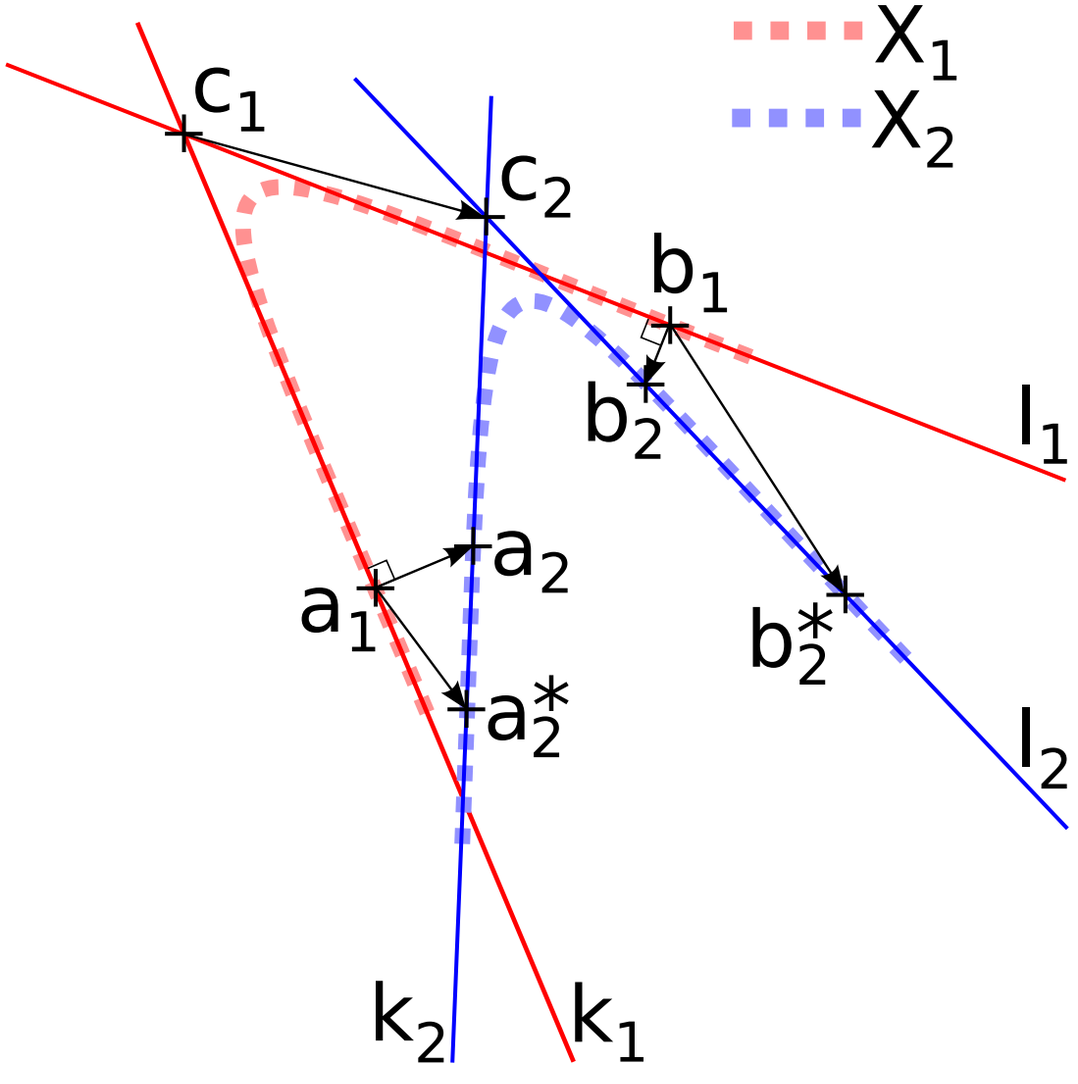
The Aperture Problem



The Aperture Problem



The Aperture Problem



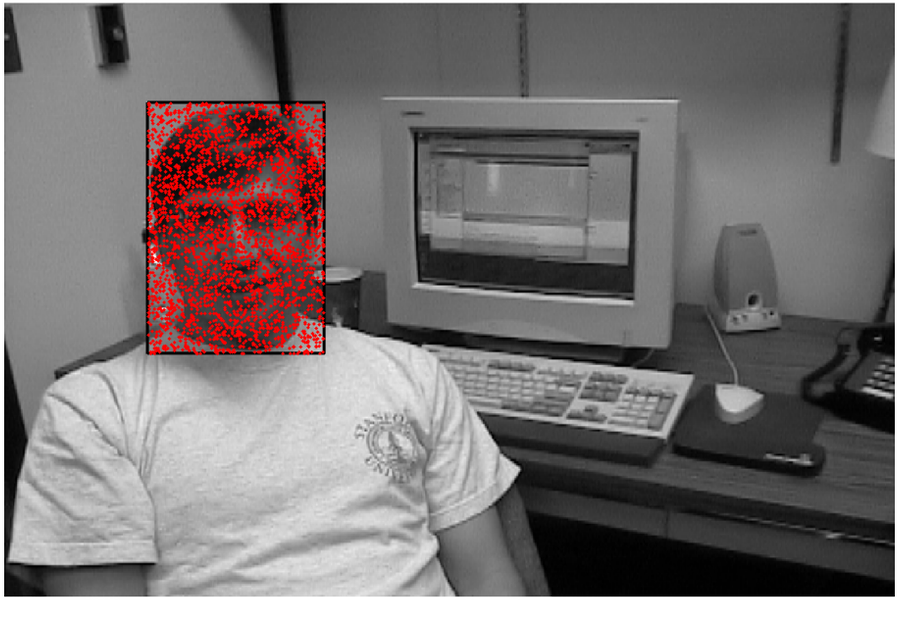
Points and Lines

Points

- edge points (edgels)
- randomly generated, then converged to the edges

Lines

- tangents to the edges



Points and Lines

Points

- edge points (edgels)
- randomly generated, then converged to the edges

Lines

- tangents to the edges

Points and Lines

Points

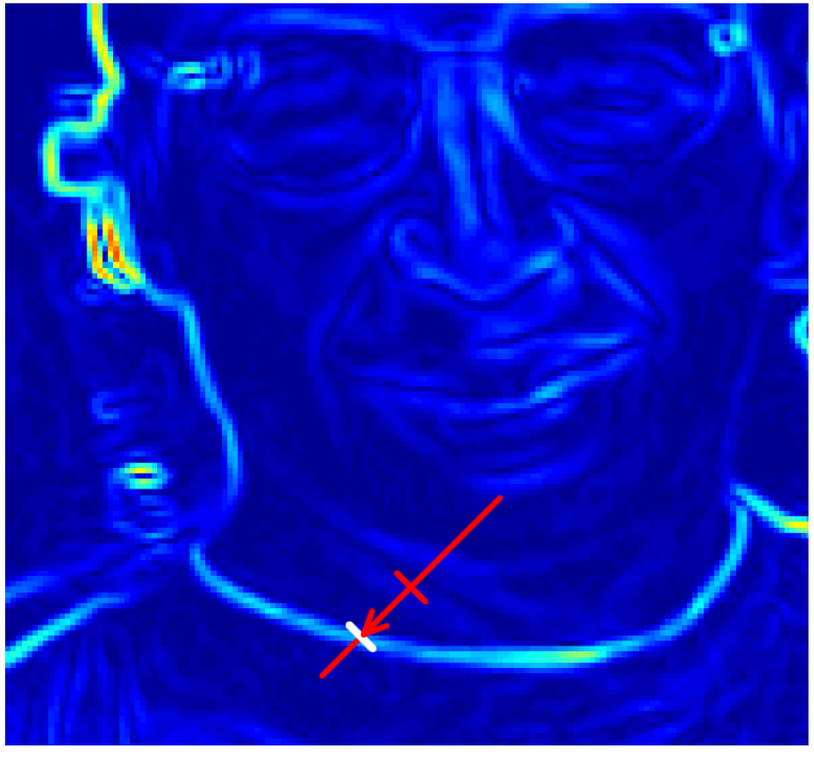
- edge points (edgels)
- randomly generated, then converged to the edges

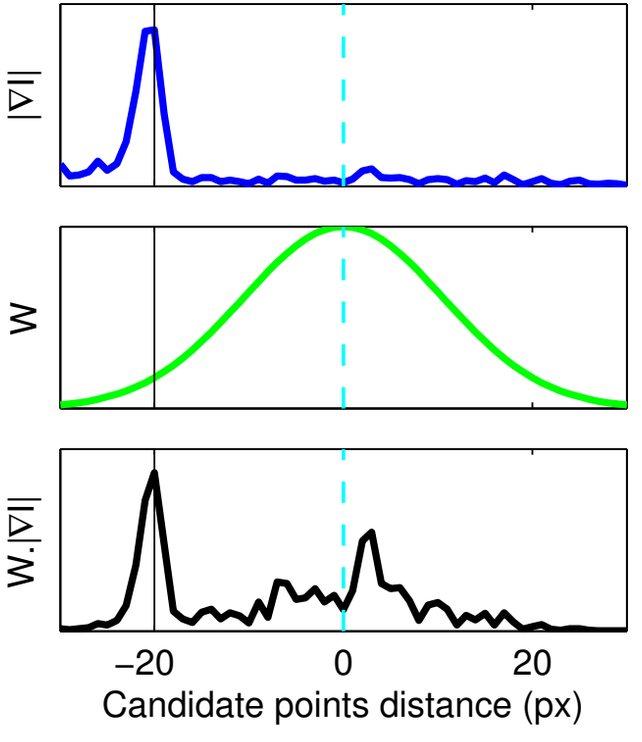
Lines

- tangents to the edges

Selecting Tangent Points

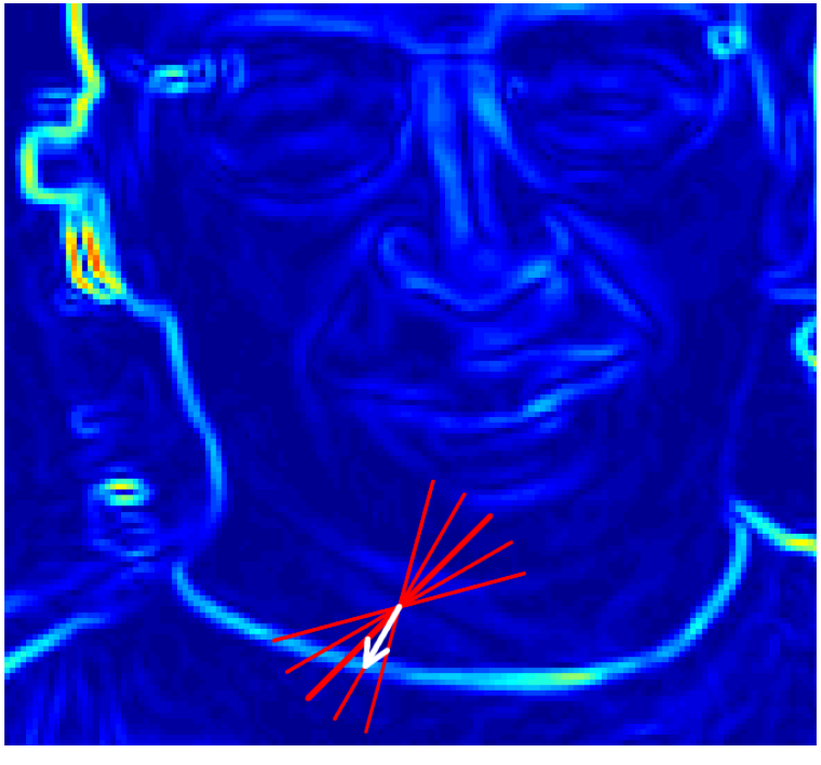
- used to find strong edges
- weak edges only when there are no strong ones near
- 1D search window in gradient direction
- weights: magnitude of gradient and distance





Matching Tangents

- used to obtain correspondences
- candidates are tangent points at local maxima of gradients in multiple directions; correspondence verification criterions are: gradient angle, appearance similarity and distance.



Line Correspondences

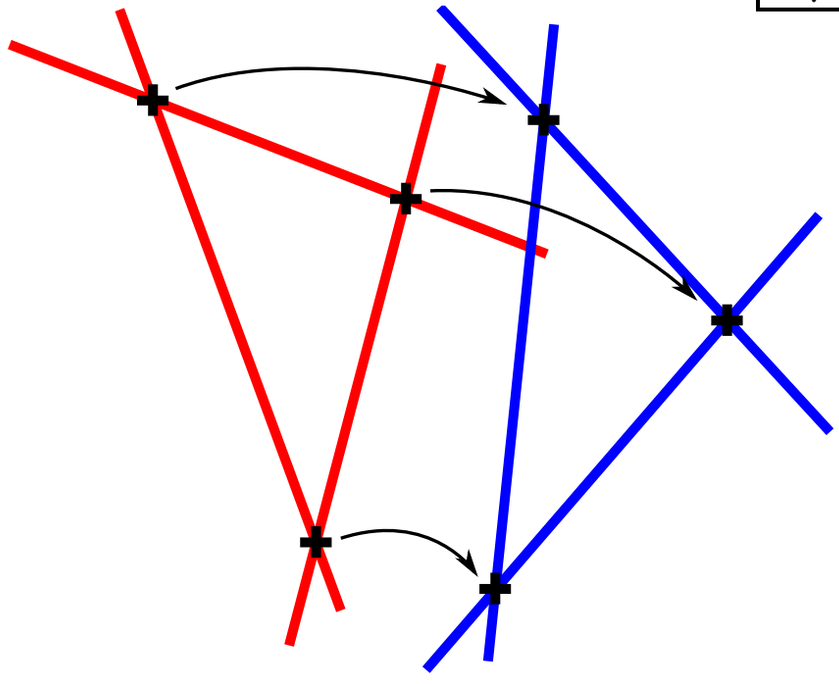
Estimation of transformation

- 2 constraints from one correspondence
- however, 2 correspondences are not enough (scale ambiguity)
- at least 3 correspondences for a similarity transformation (4 DoF), either directly or by intersection correspondences

RANSAC – RANdom SAmple Consensus

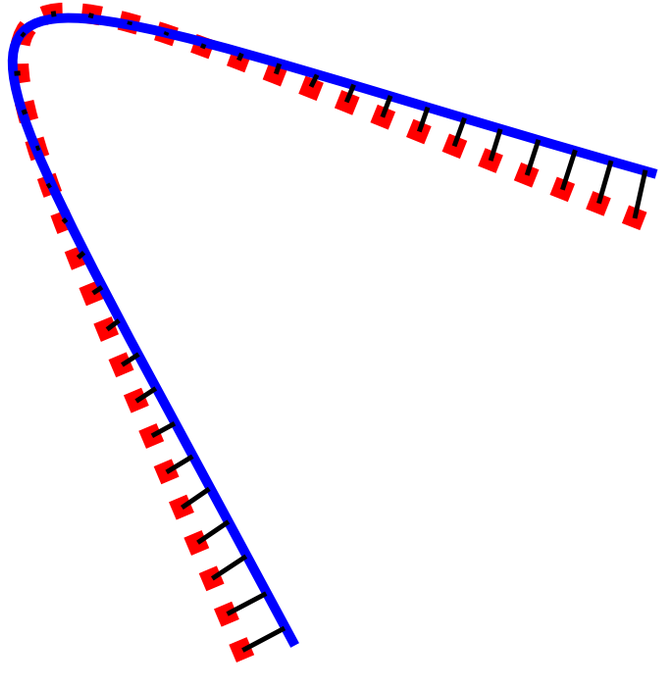
Random Sample

- minimal sample:
  3 line correspondences
- transformation by point
  correspondences of intersections



Consensus

- inlier ratio insufficient
  indicator of quality
- used together with
  image evidence of points
  (Chamfer distance)



Online Learning Points Reliability

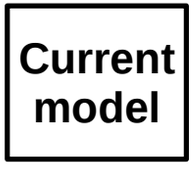
- remembering positions of edge points whose tangents are often transformed to edges
  $\Rightarrow$ field of observed point qualities
- gives an estimate of object structure
- fitting to points decreases drift

Model

- Encapsulates whole state of frame-to-frame tracking:
  - geometric information:
    position of tracker, points, lines,
  - field of observed point qualities.
- Initial model: user-given tracked bounding box, generated points.
- Online update at every frame.

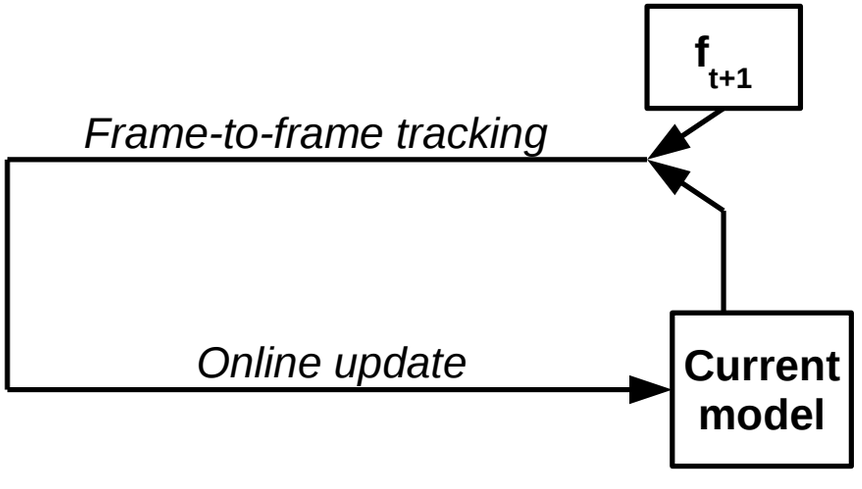
Using Multiple Models

Current model = state of the tracker.



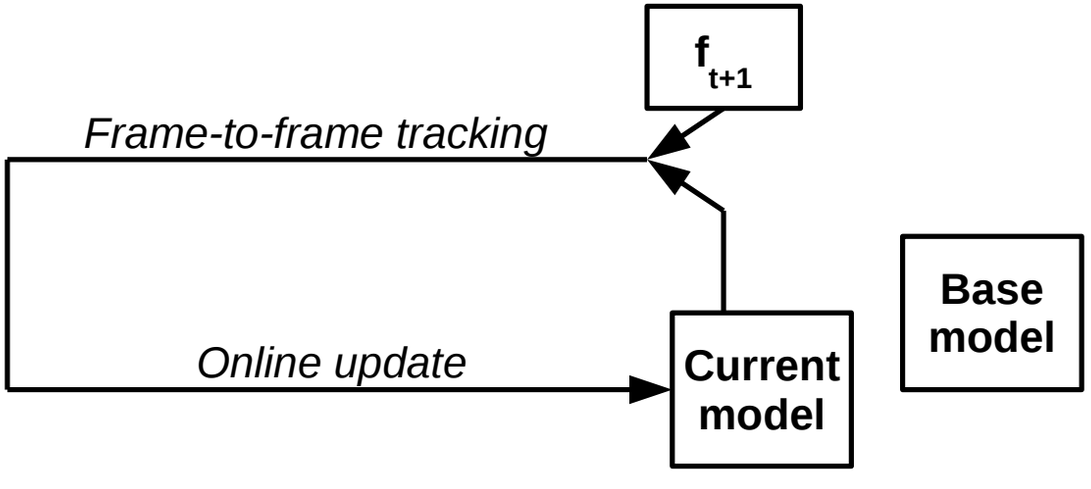
Using Multiple Models

At every frame, current model is updated. RANSAC score computed from image evidence and inlier ratio is used as a measurement of tracking quality.



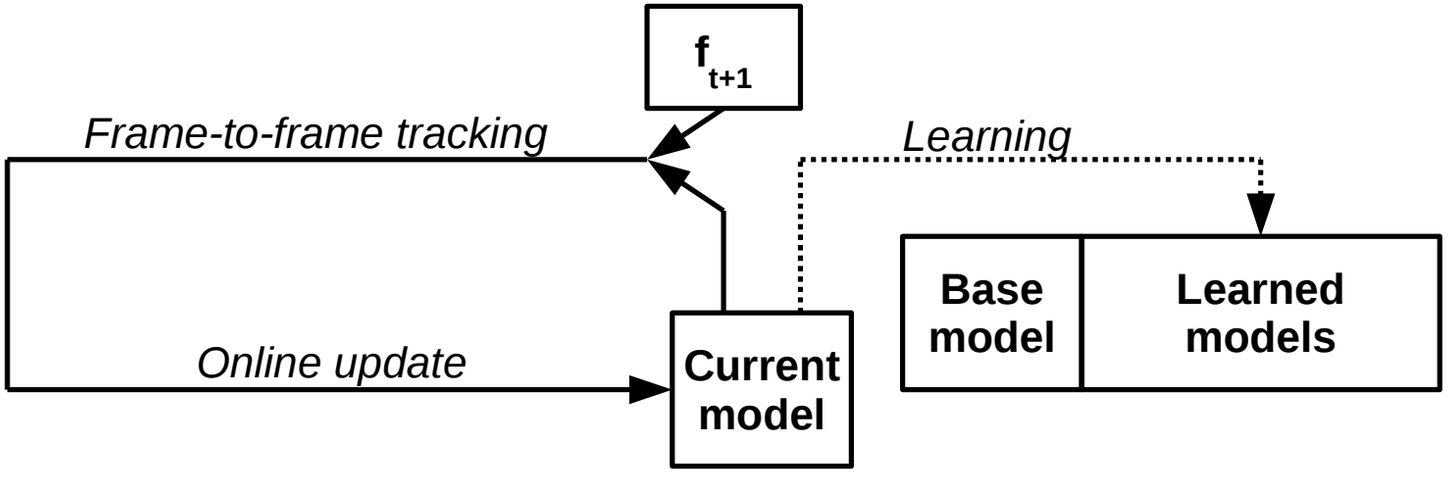
Using Multiple Models

Base model = initial state at the first frame.



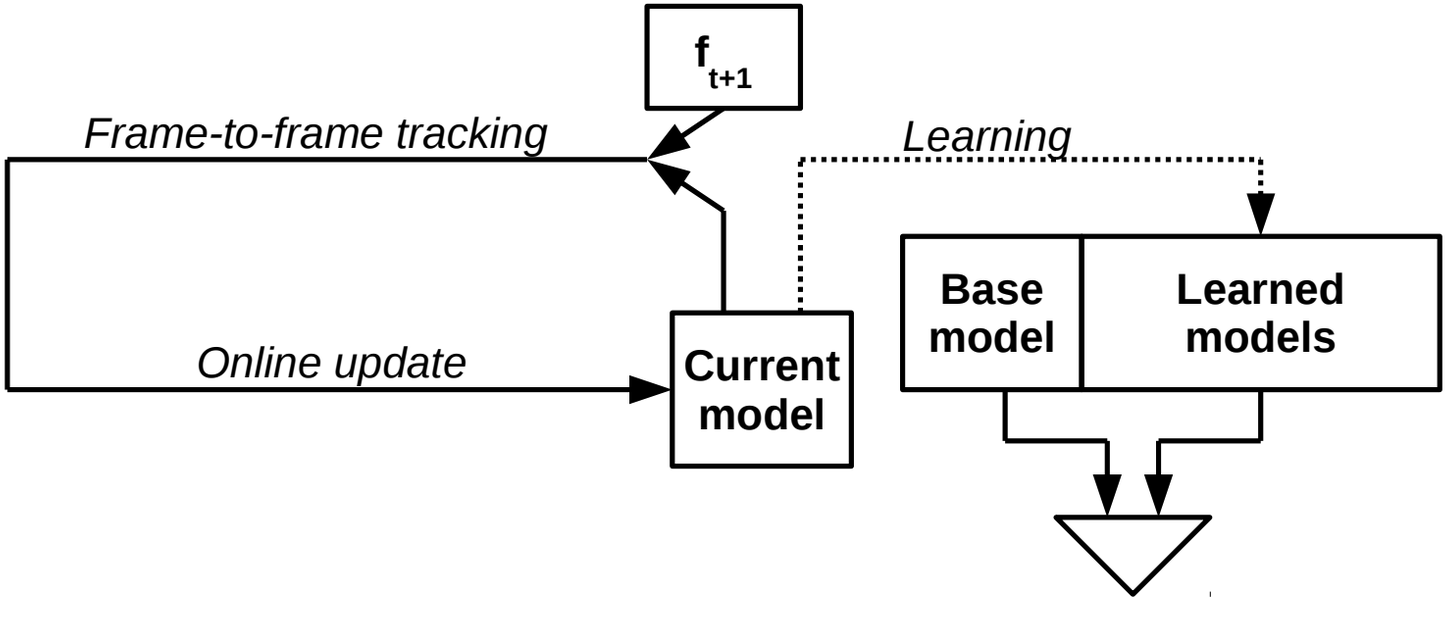
Using Multiple Models

Models with high confidence are stored for later use.



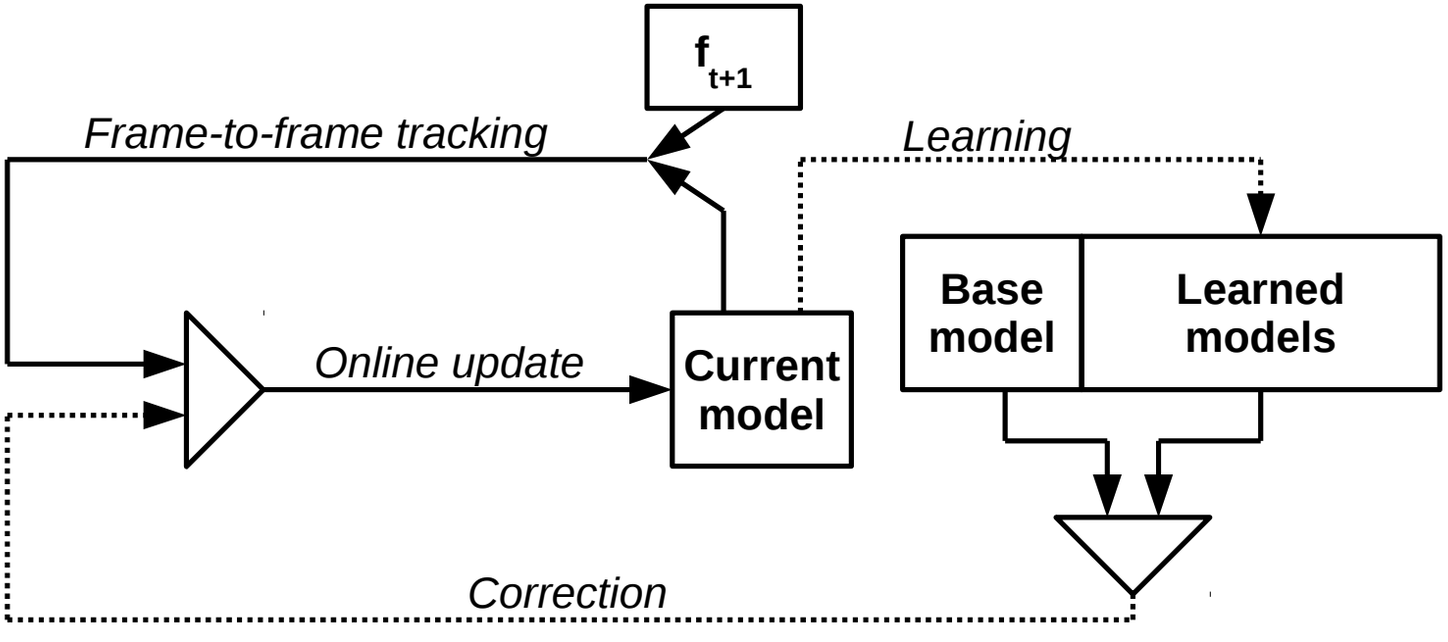
Using Multiple Models

On sudden drop of current model score, the base model and possibly other learned models are tried for a correction.

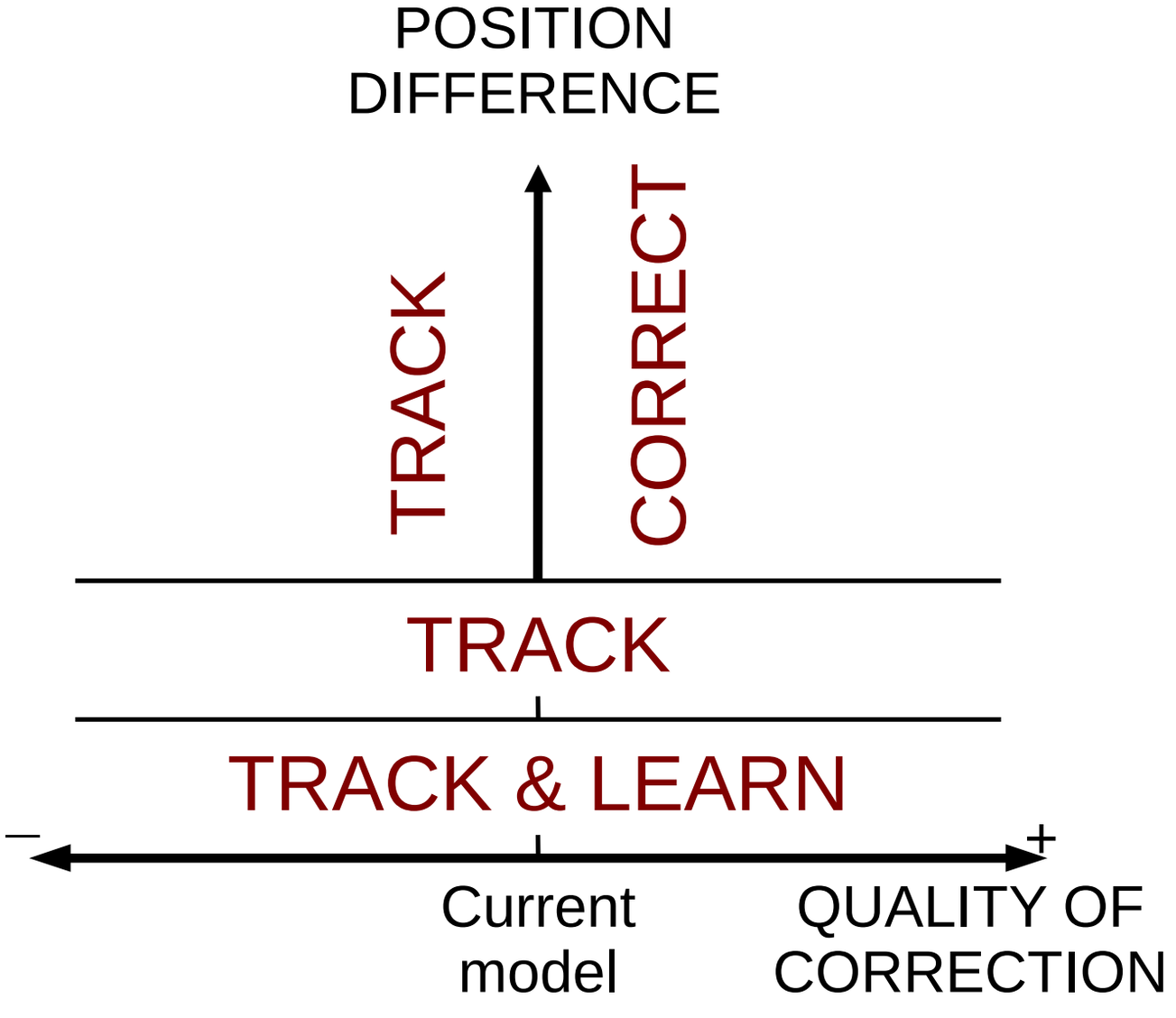


Using Multiple Models

The best of correction models is then compared with current model and tracker pose is recovered if necessary.



Possible Events In Correction of the Tracker

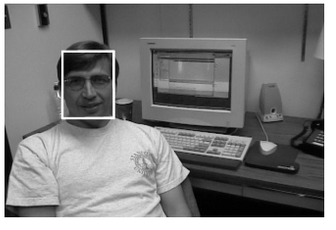


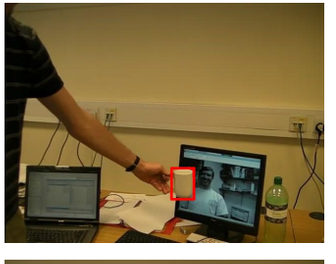
Experimental Evaluation

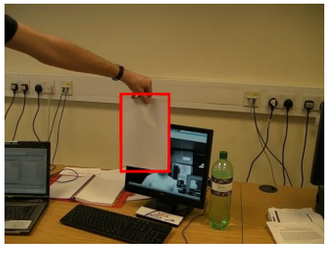
Compared with FoT (LK, Vojíř, '10), LGT (Čehovin, '11), TLD (Kálal, '11), results are competitive or superior.

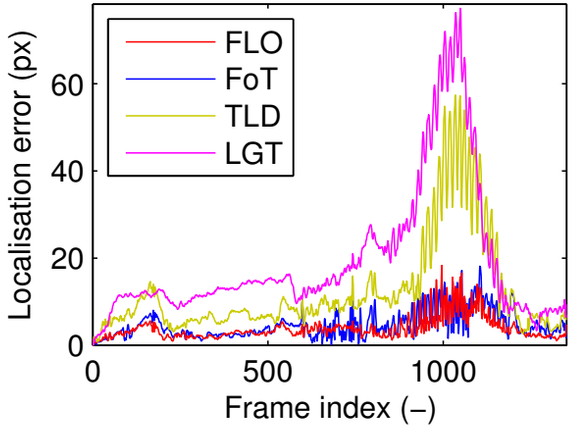
Measured:

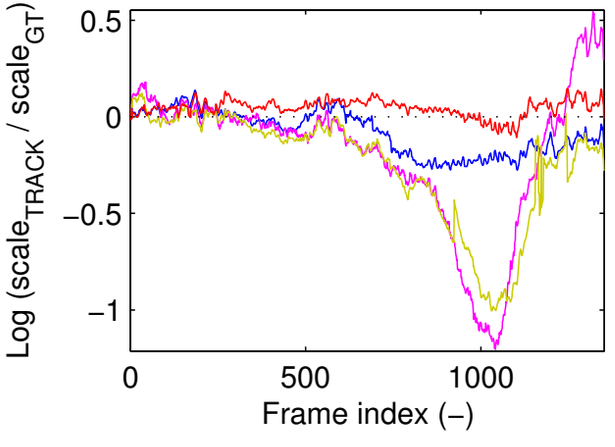
- localization error (distance of center)
- scale error (ratio of size to GT size)





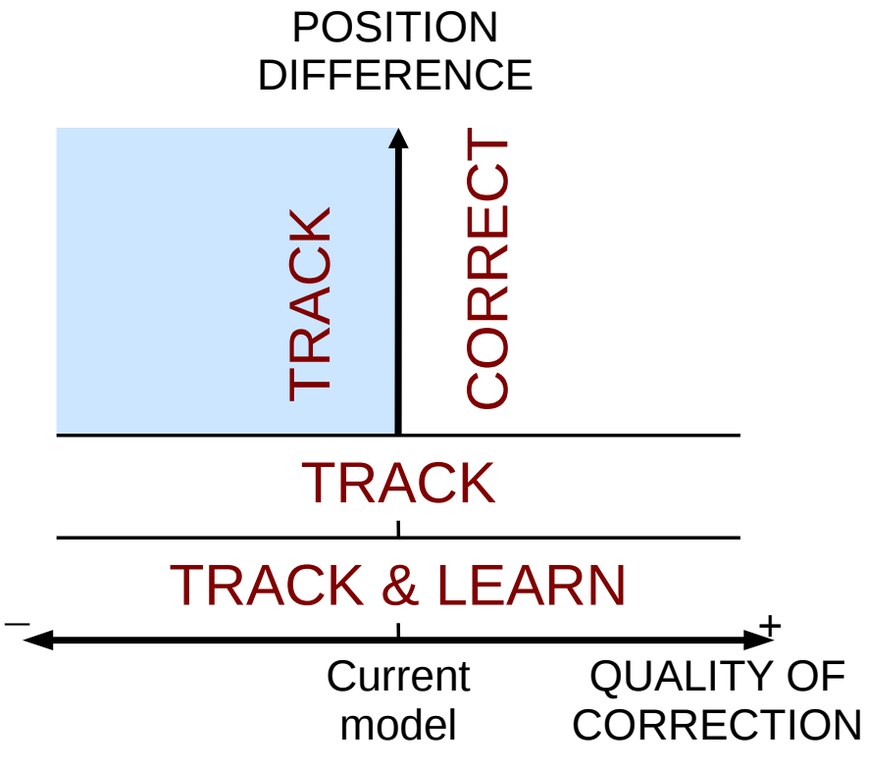


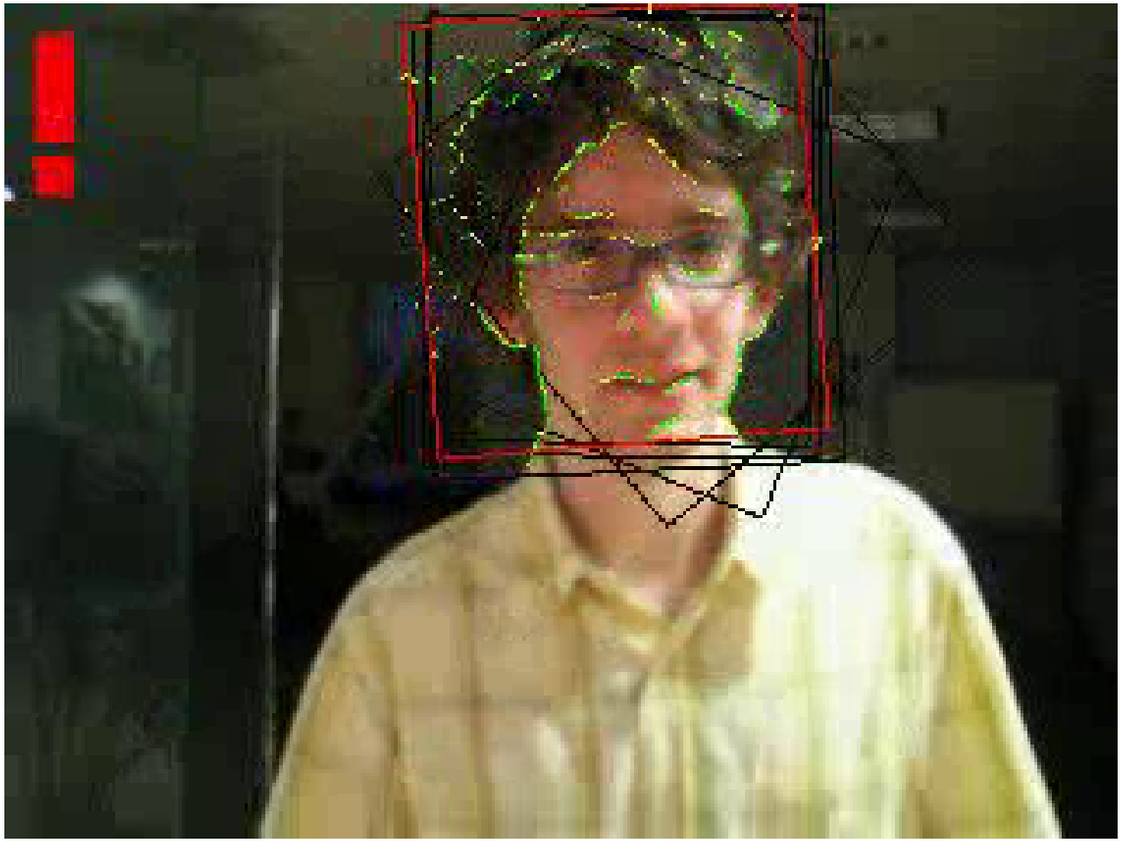




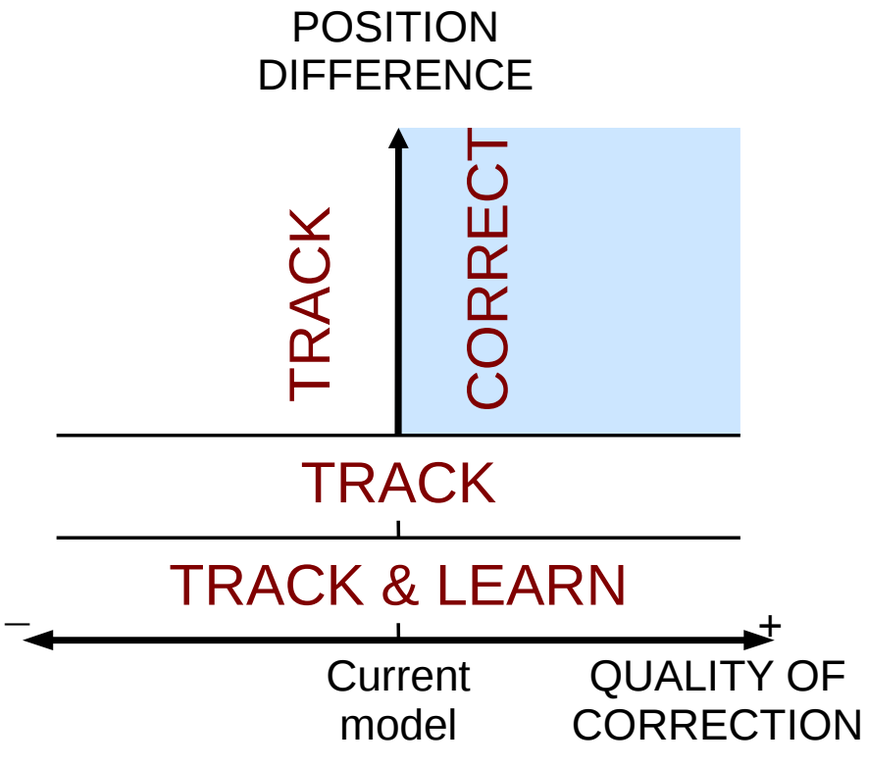
Matas, J. and Vojíř, T.: Robustifying the flock of trackers. Proc. of CVWW 2011.
Cehovin, L. *et al.*: An adaptive coupled-layer visual model for robust visual tracking. Proc. of ICCV 2011.
Kálal, Z. *et al.*: P-N learning: Bootstrapping binary classifiers by structural constraints. Proc. of CVPR 2010.

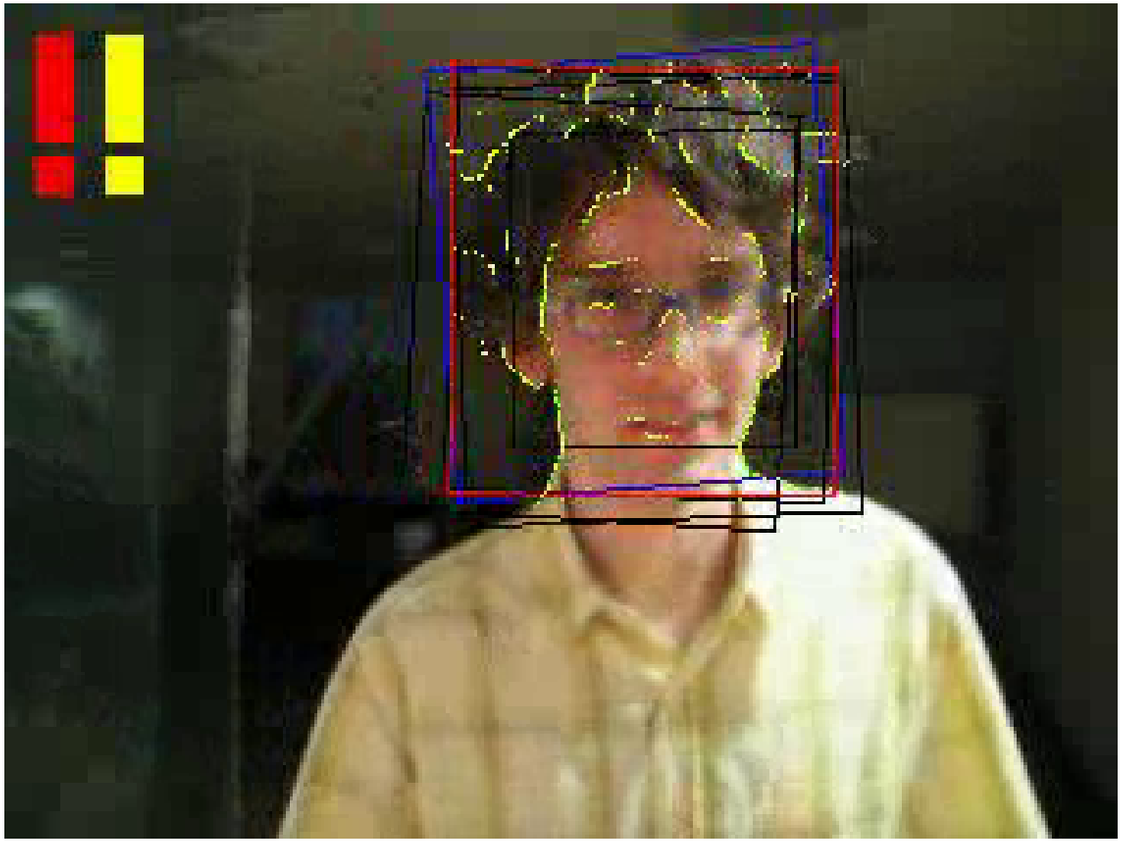
Video Example



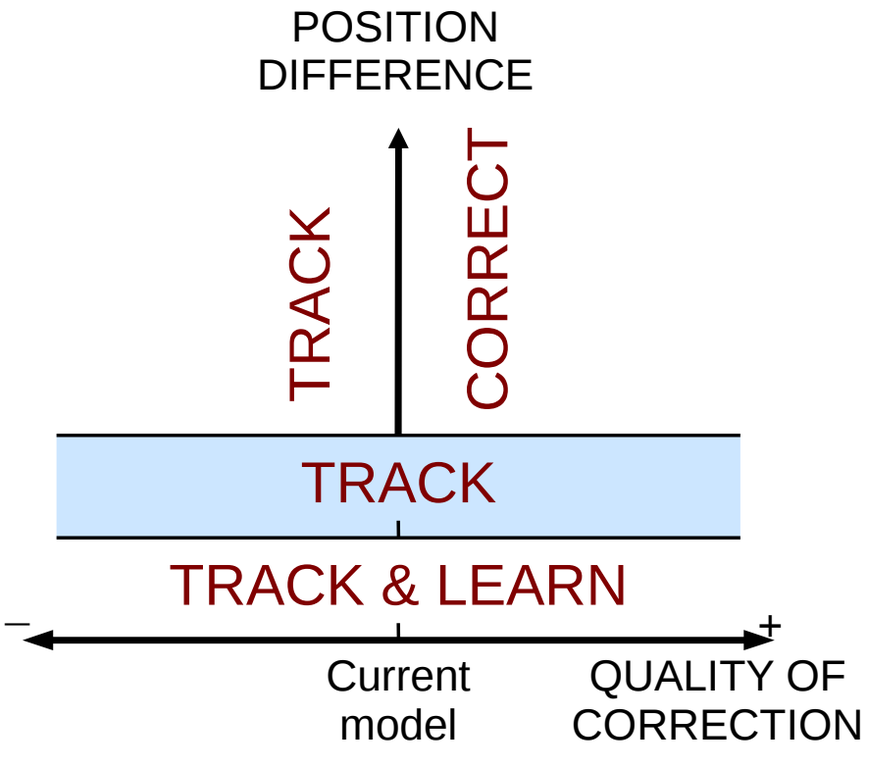


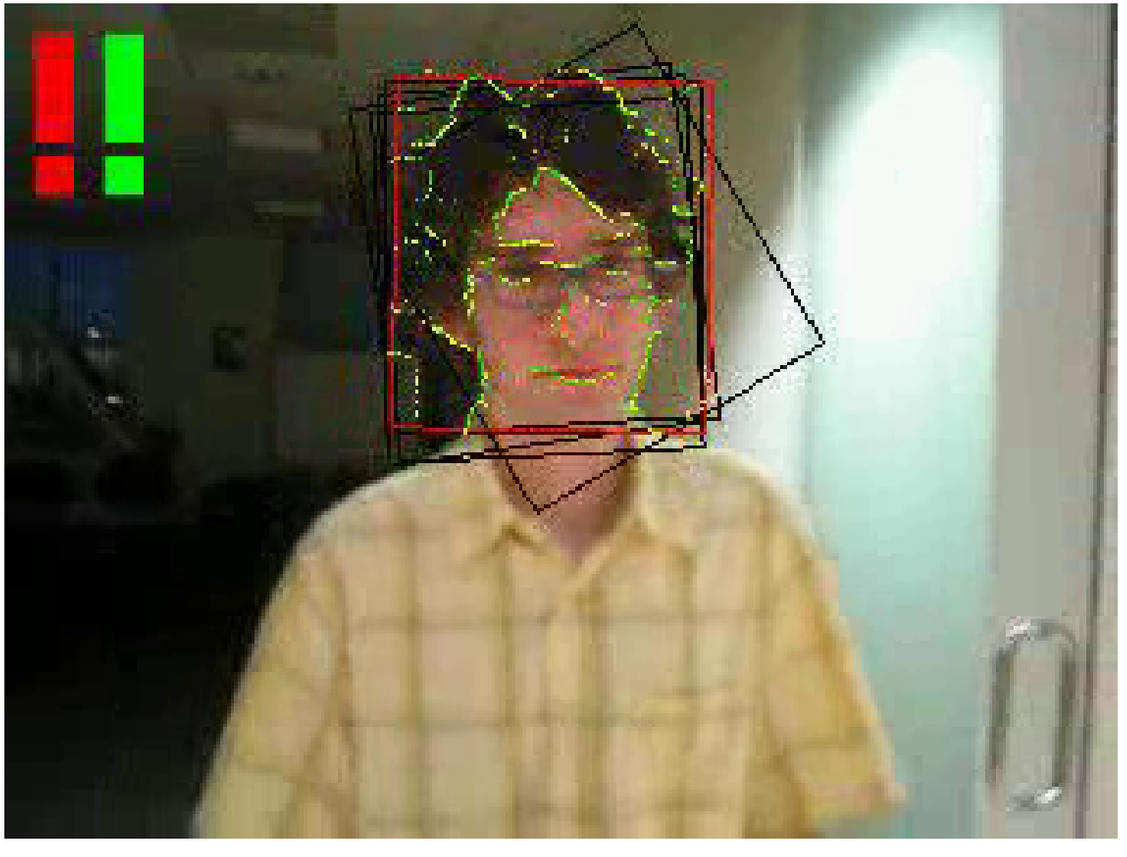
Video Example



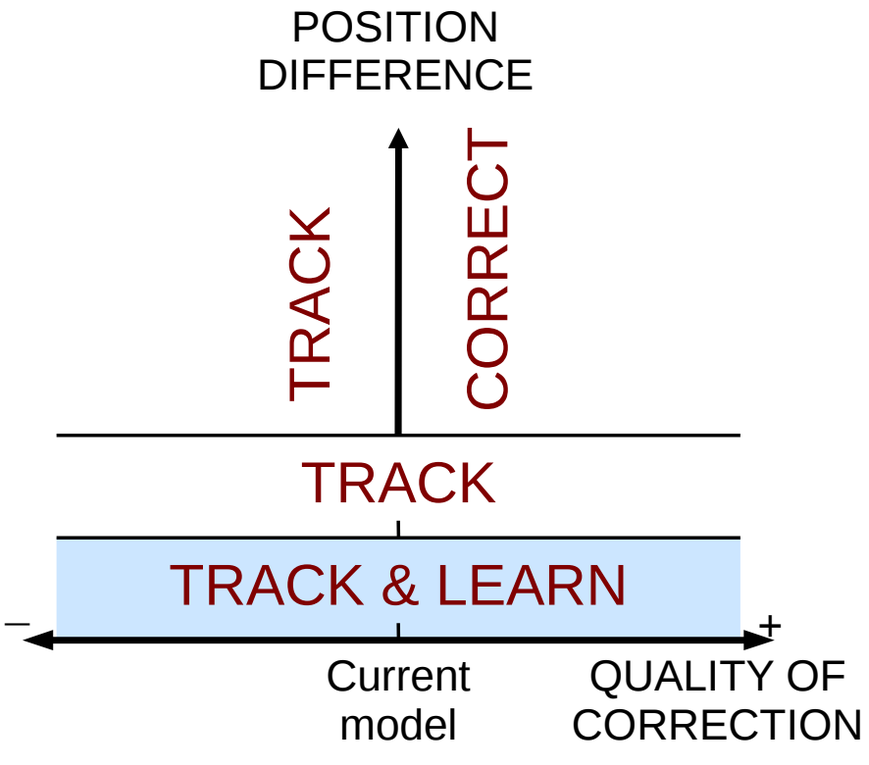


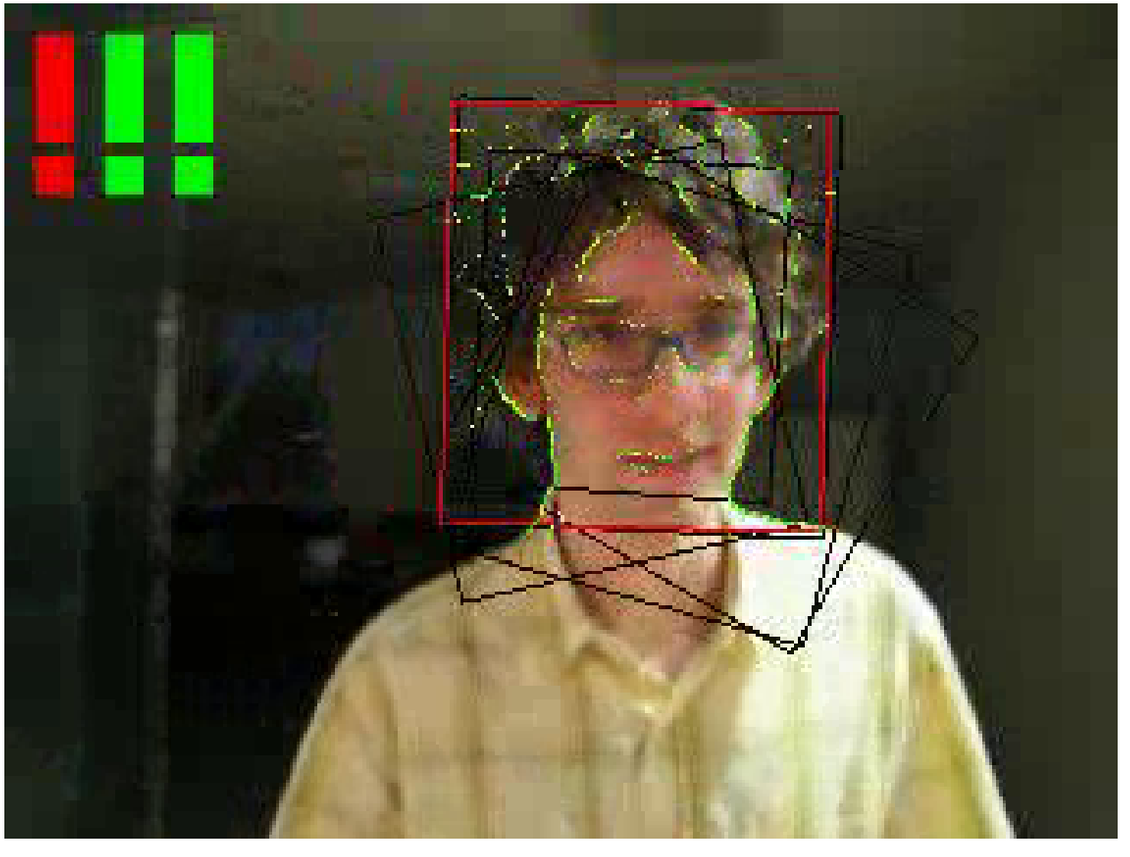
Video Example





Video Example





Video Example

David.avi

Conclusions

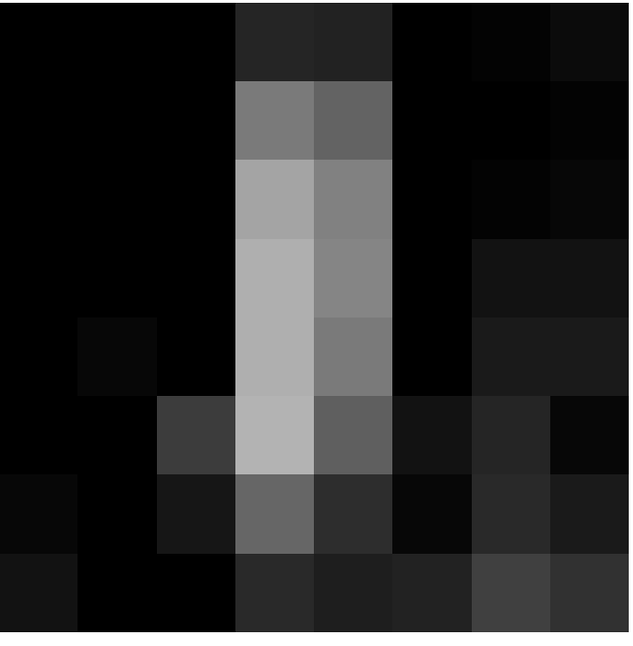
- designed, implemented and tested the FLO tracker using a novel idea of tracking tangent lines
- competitive on scenes with the sufficient number of distinguishable features
- superior on scenes without such features
- disadvantage: still short-term tracking (no redetection, once lost = forever lost)

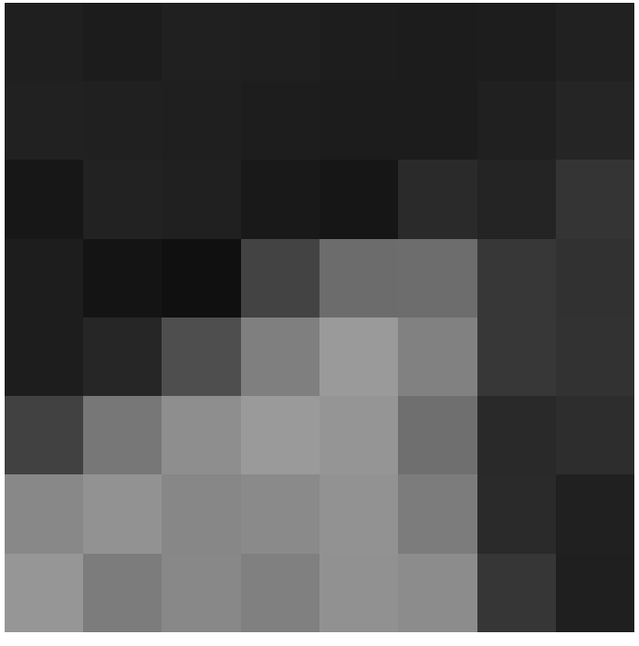
Thanks for attention! Any questions?

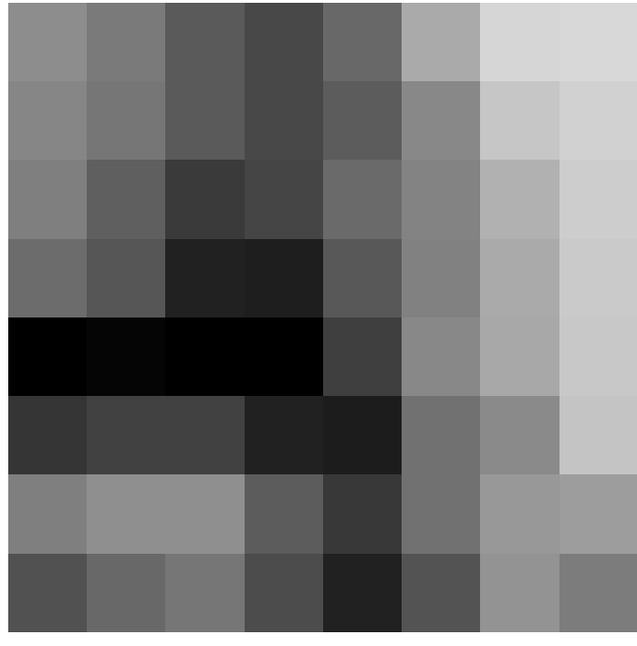
karel@lebeda.sk

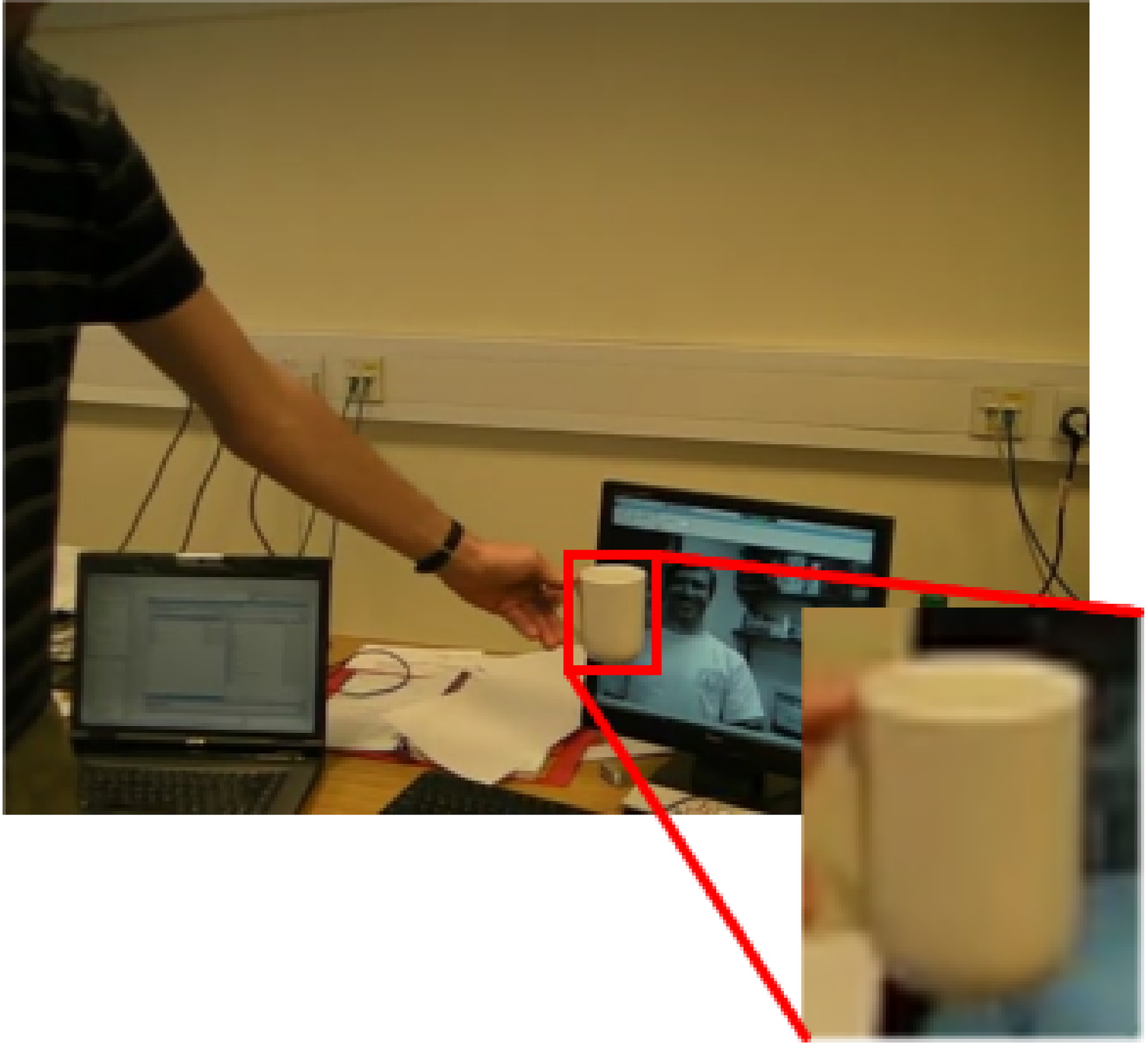

m p

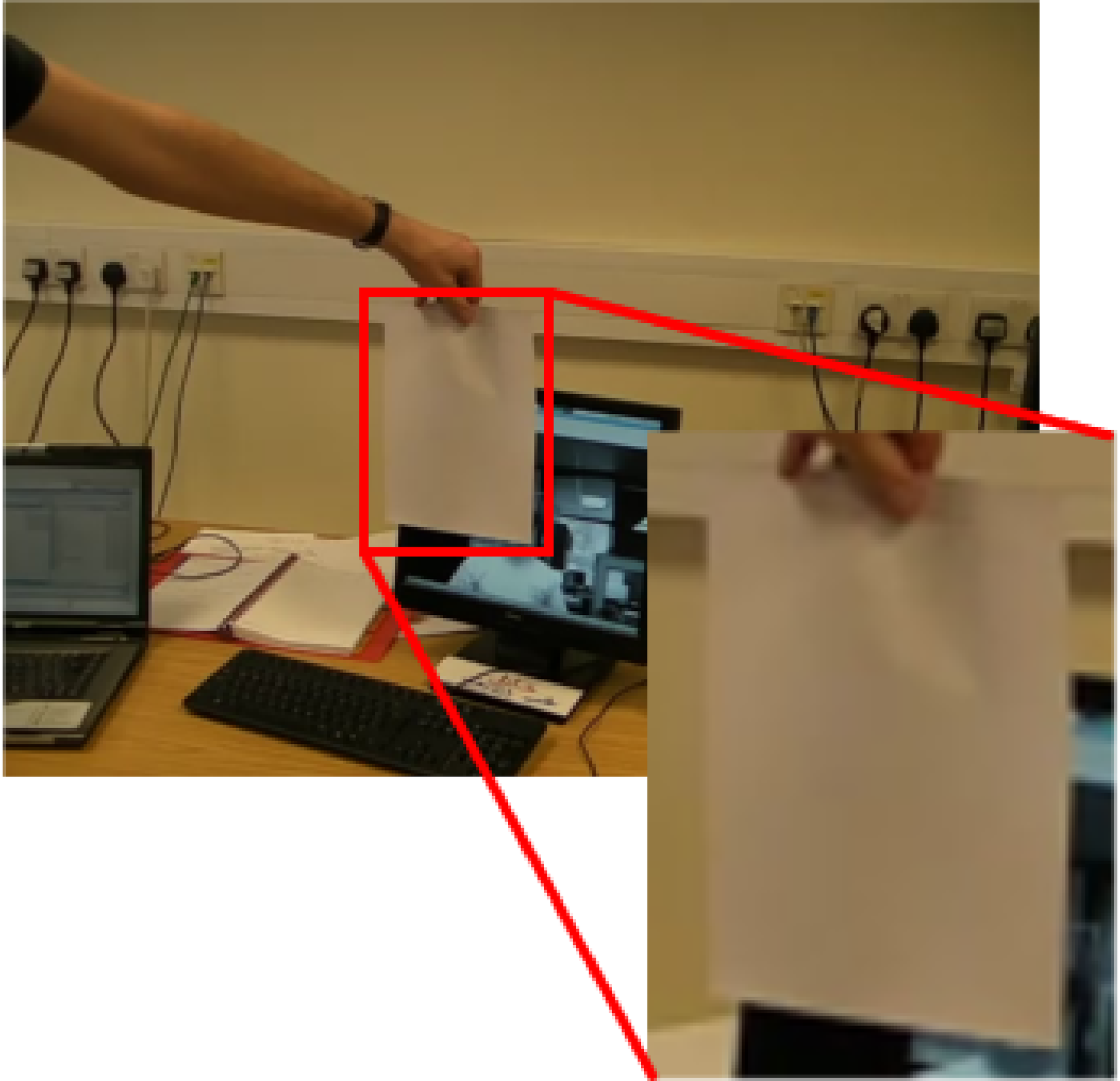
UNIVERSITY OF
SURREY

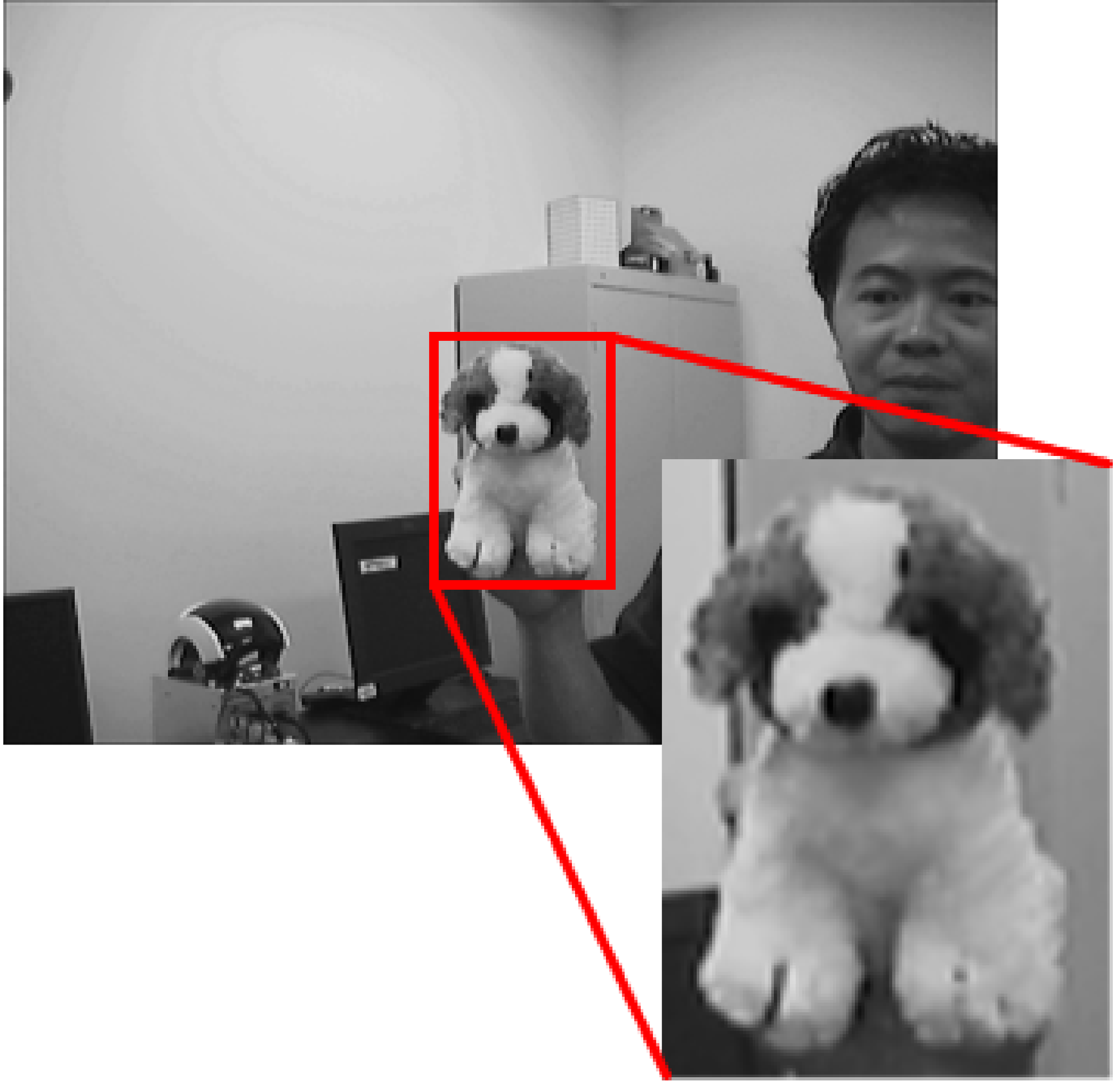


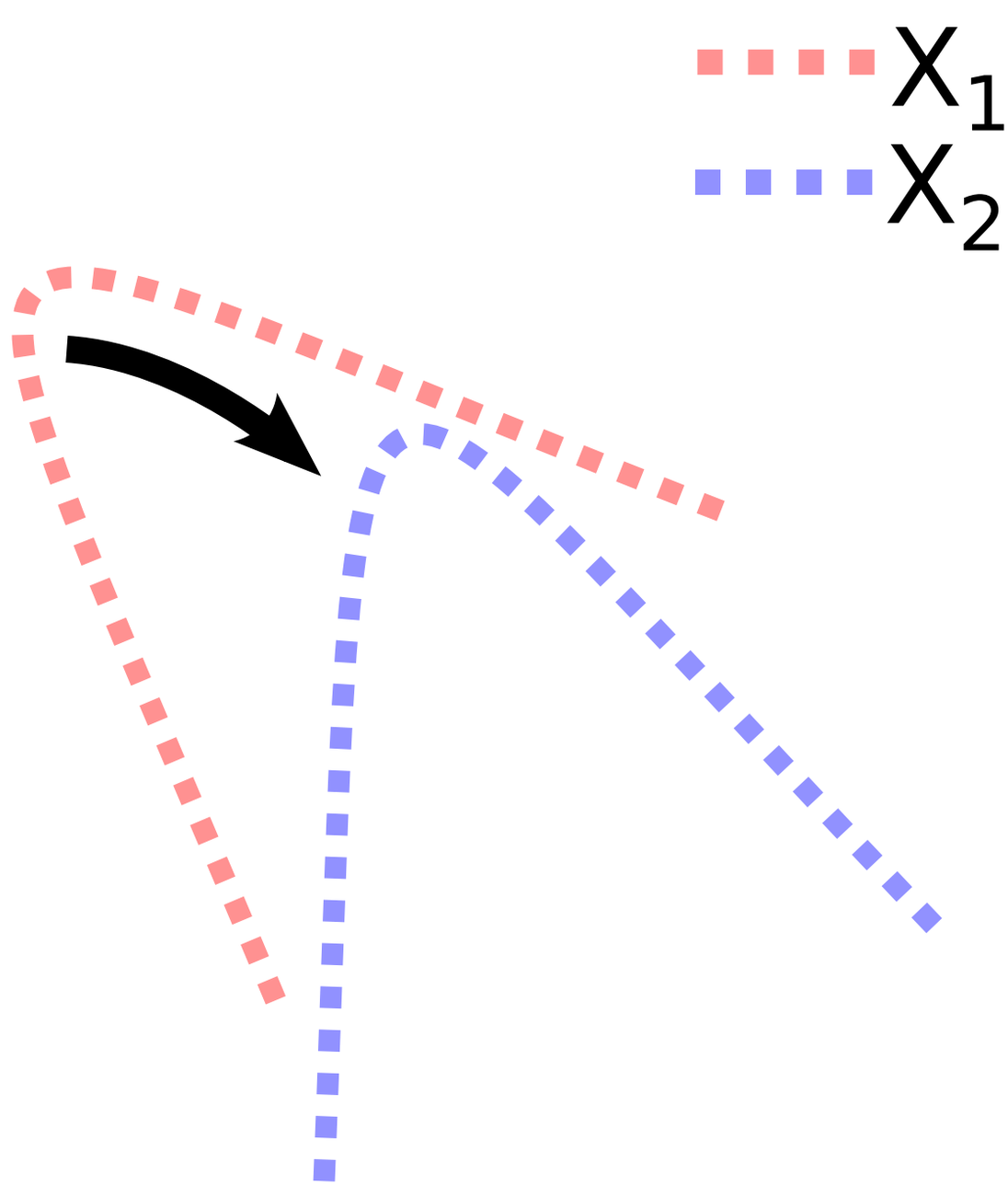


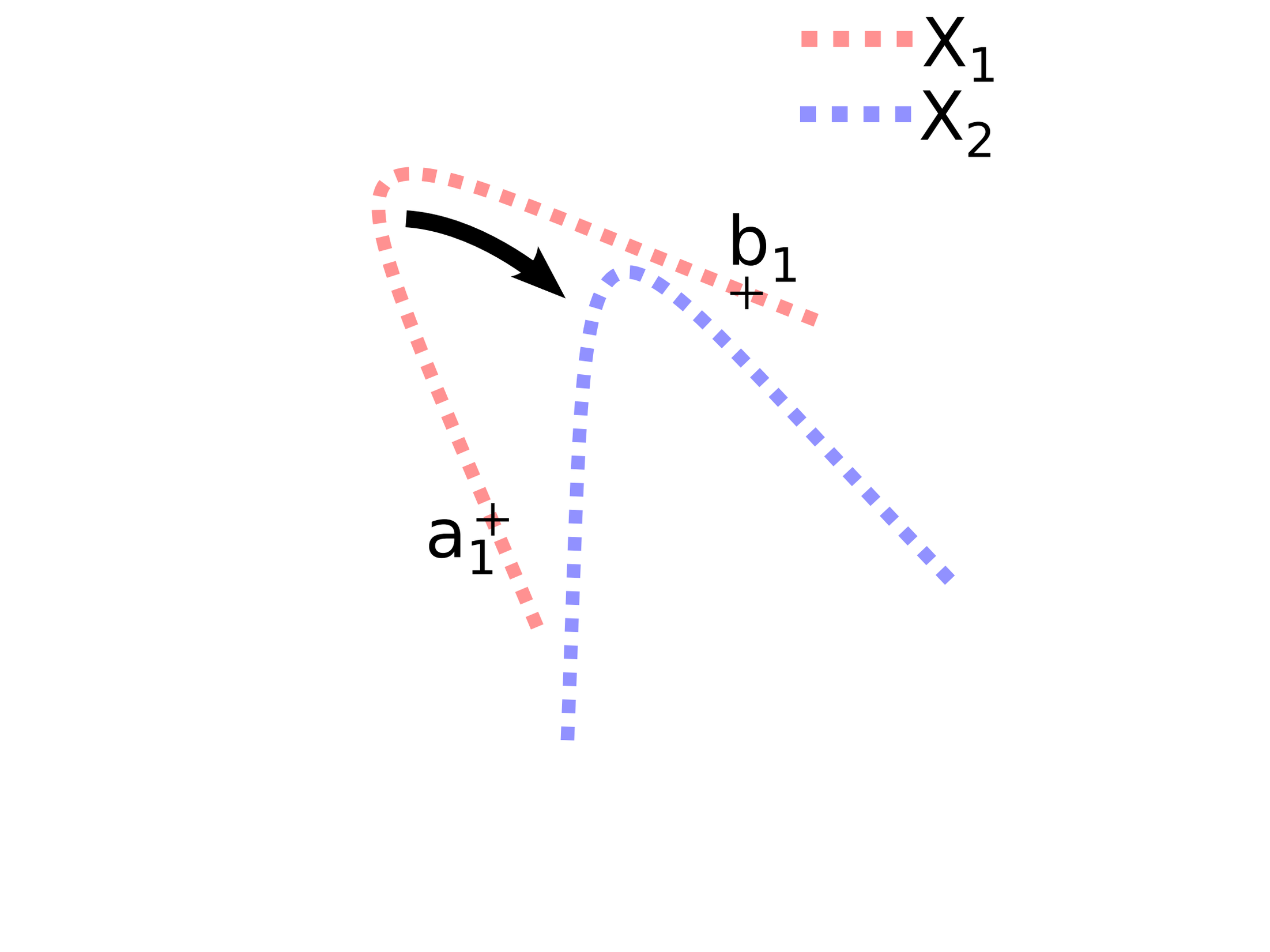


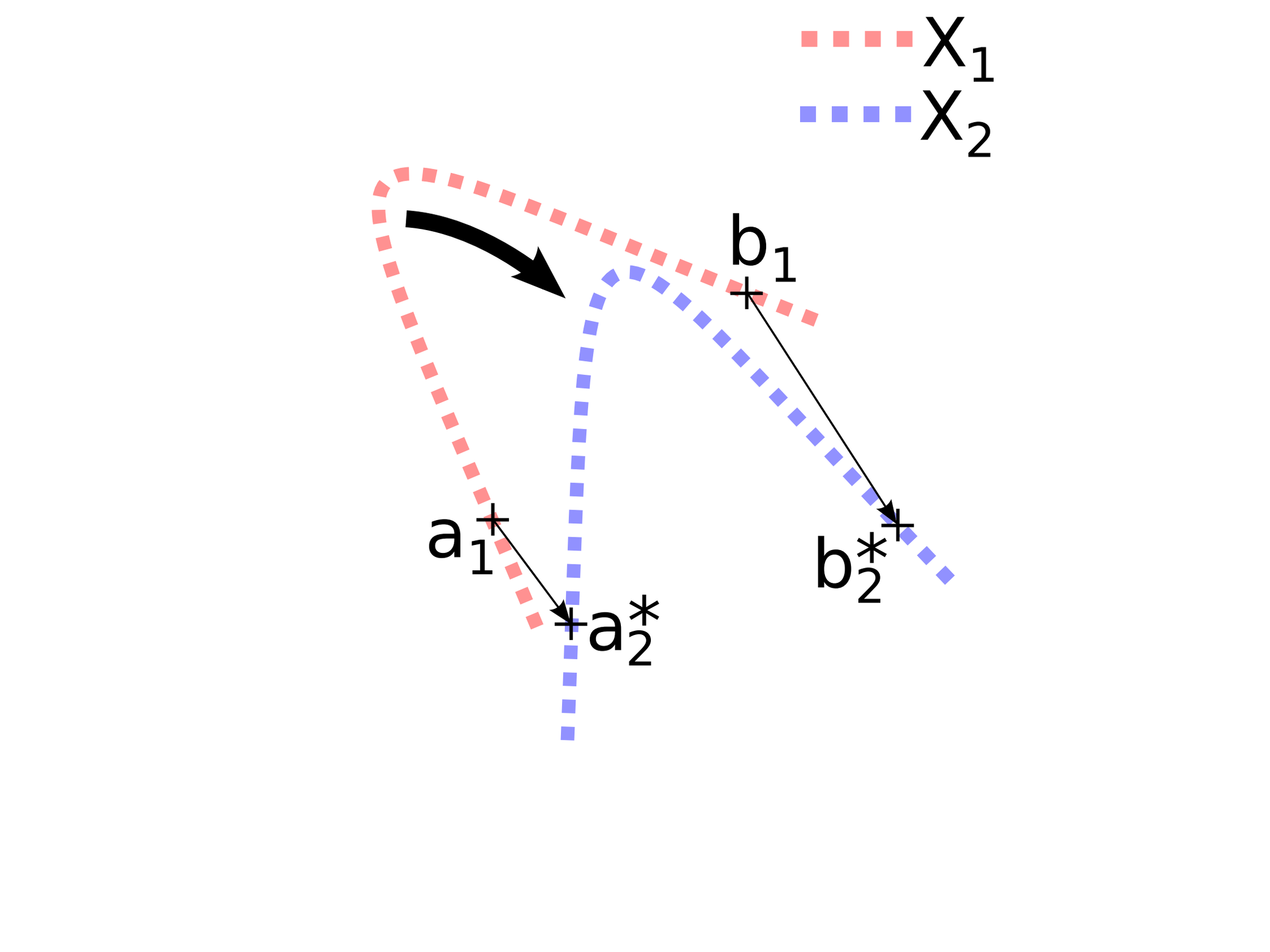


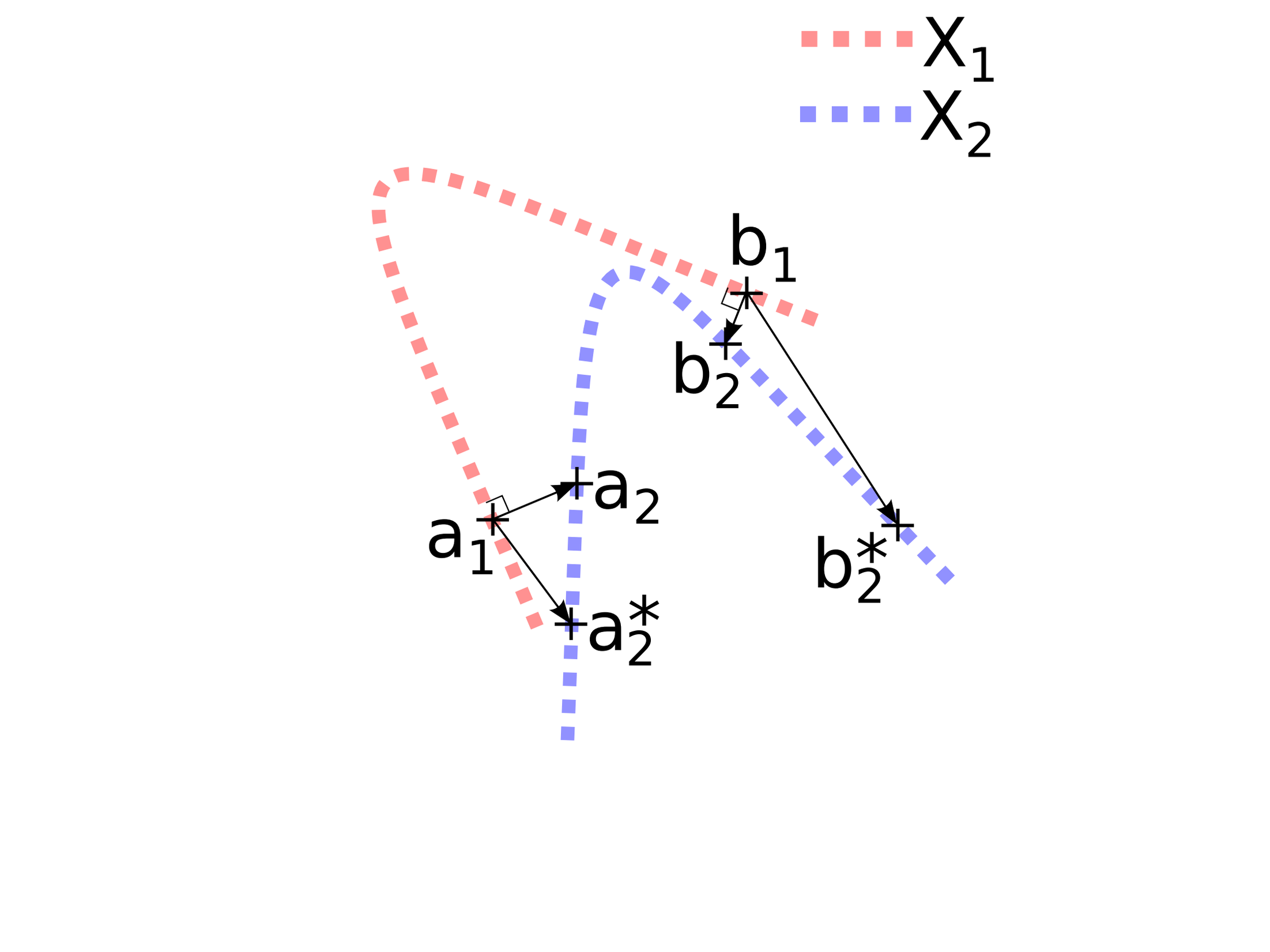


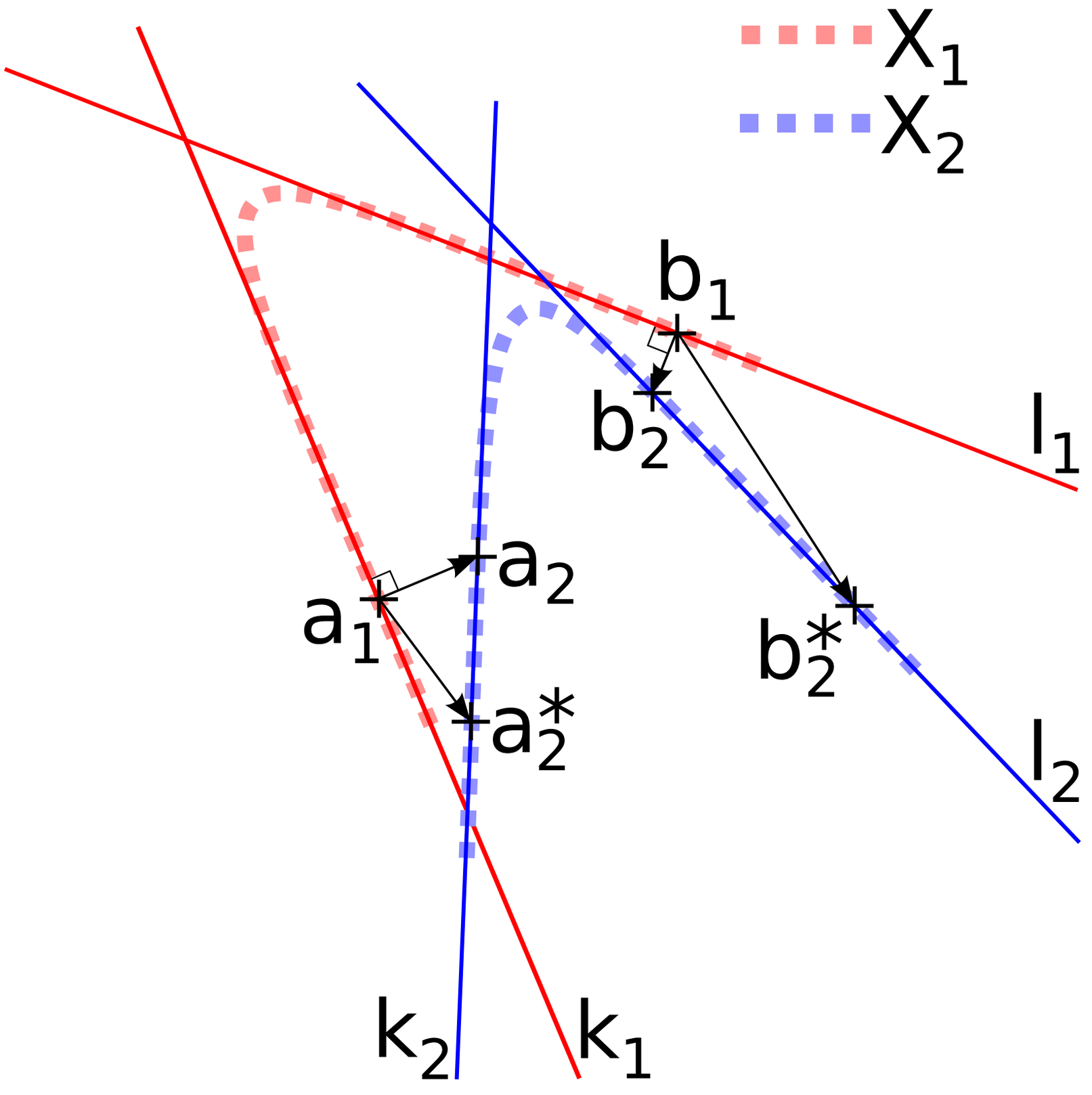


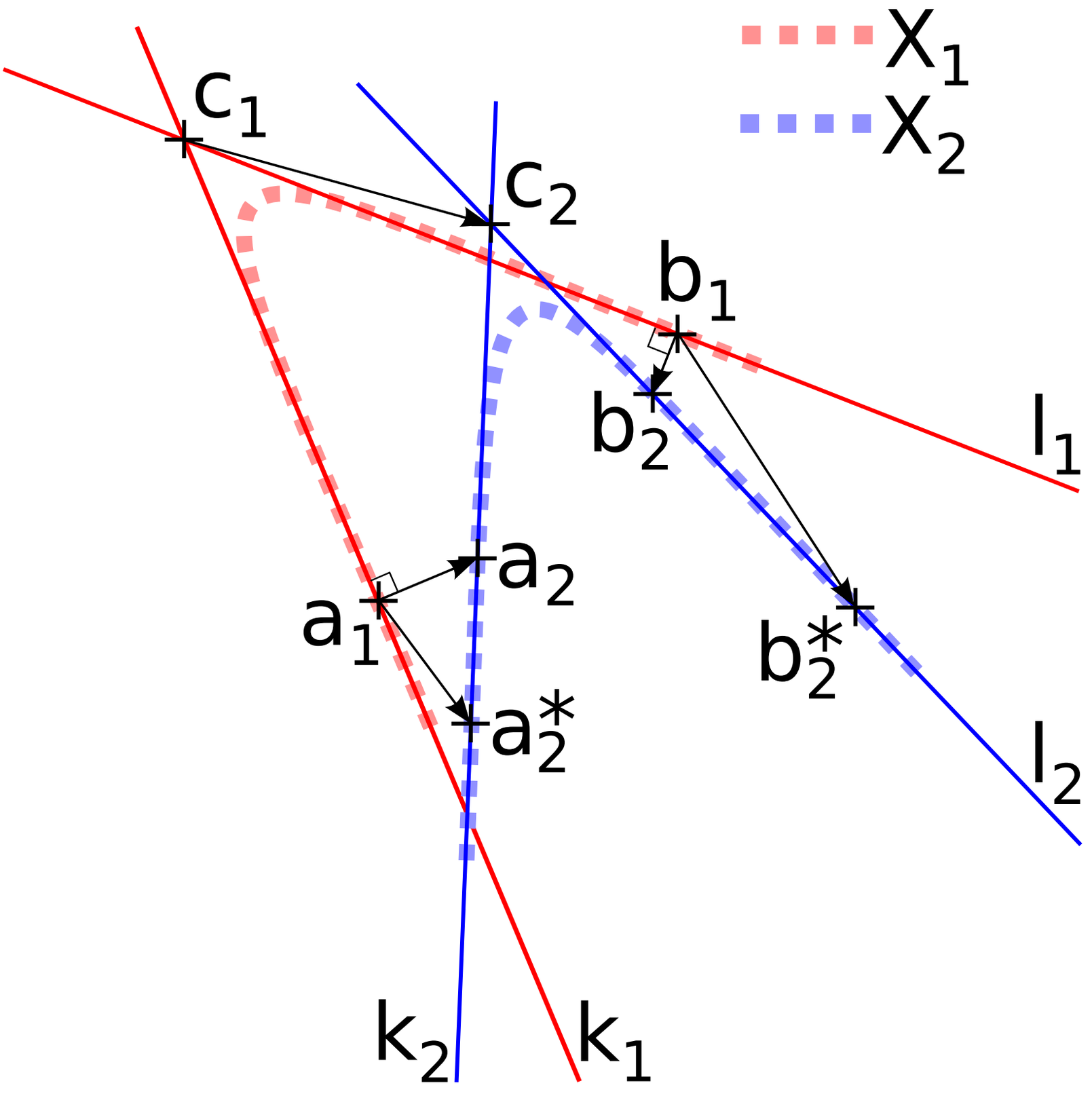

X_1
X_2

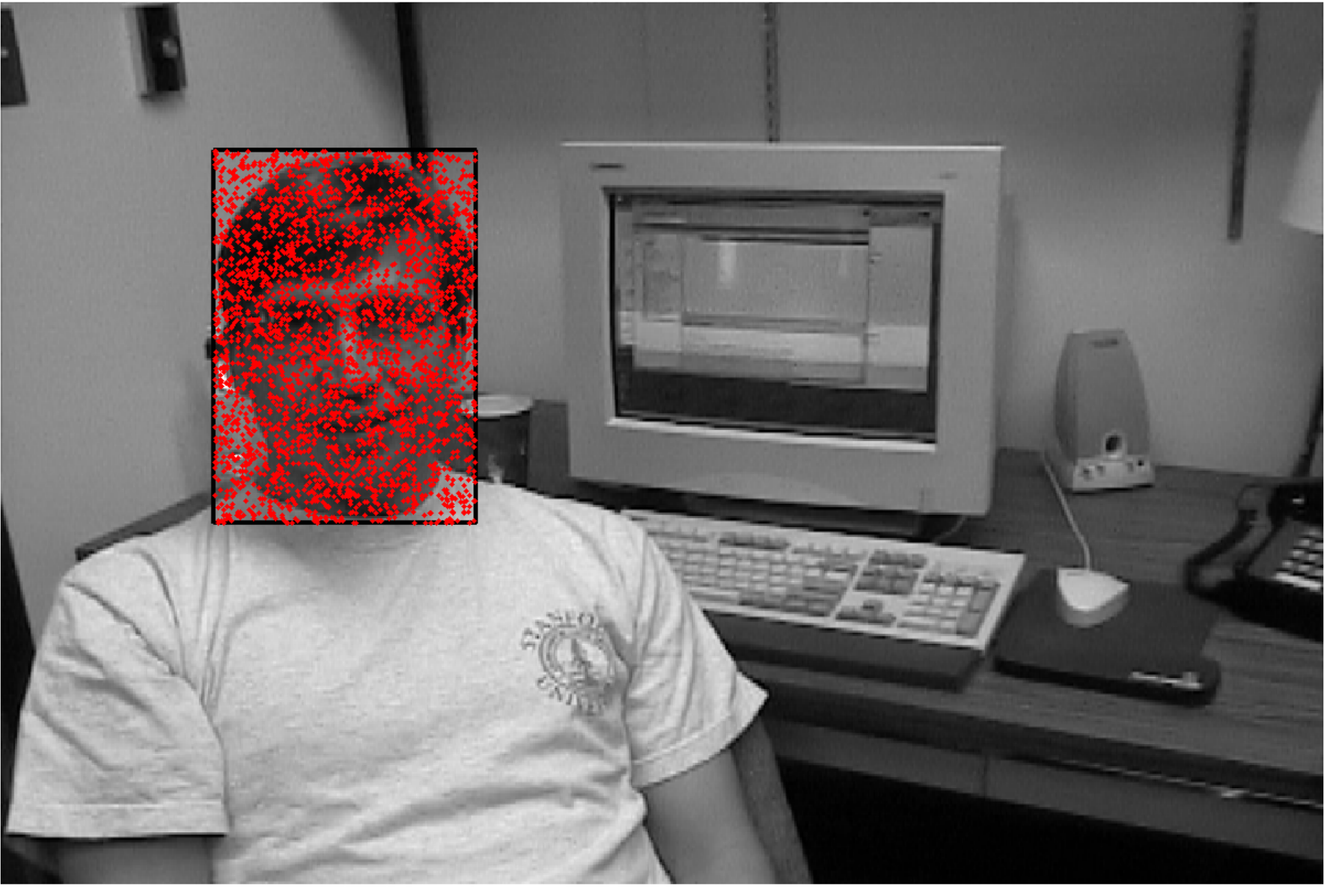

X_1
X_2
b_1^+
a_1^+

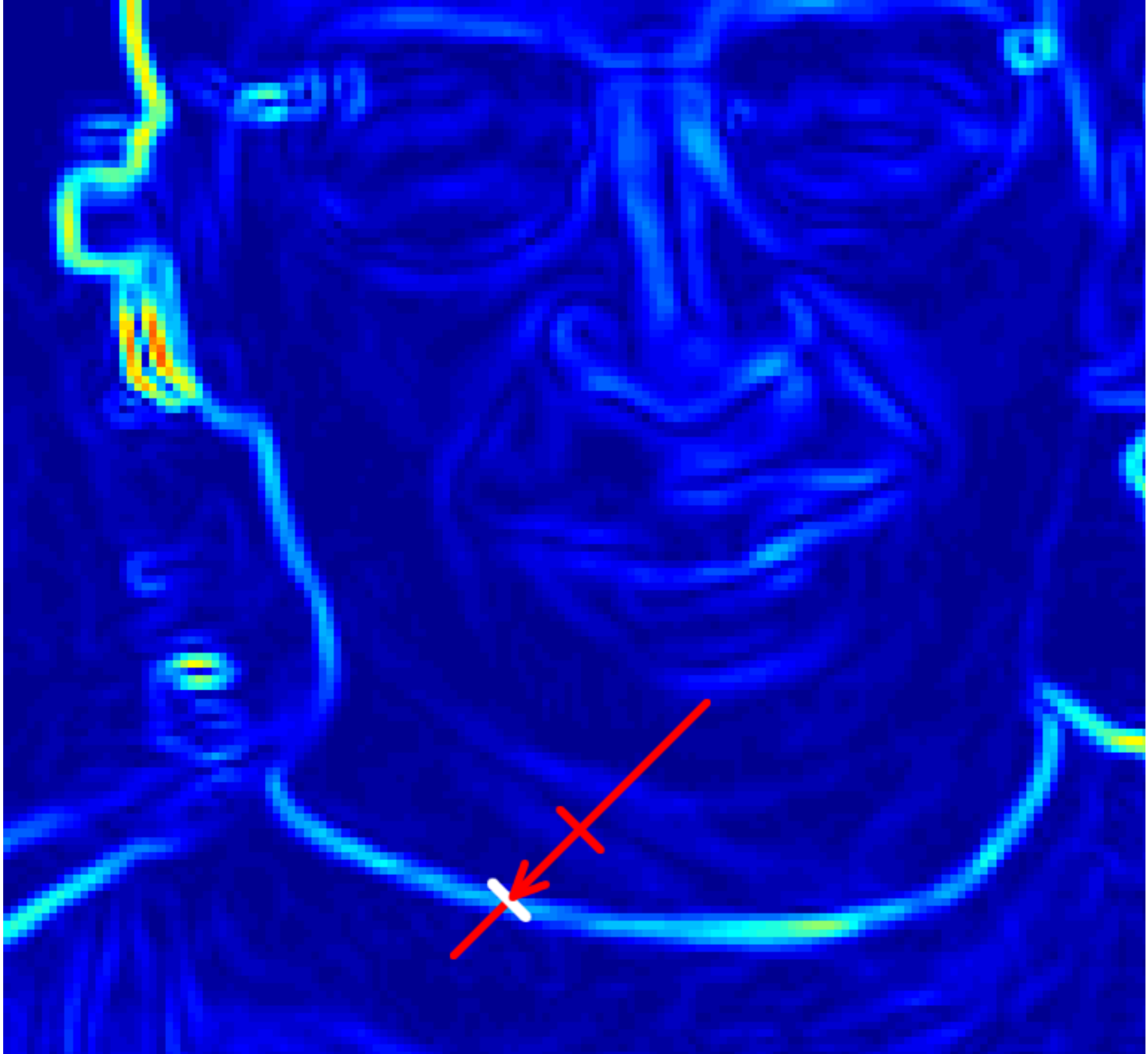

X_1
X_2
b_1
a_1
a_2^*
b_2^*

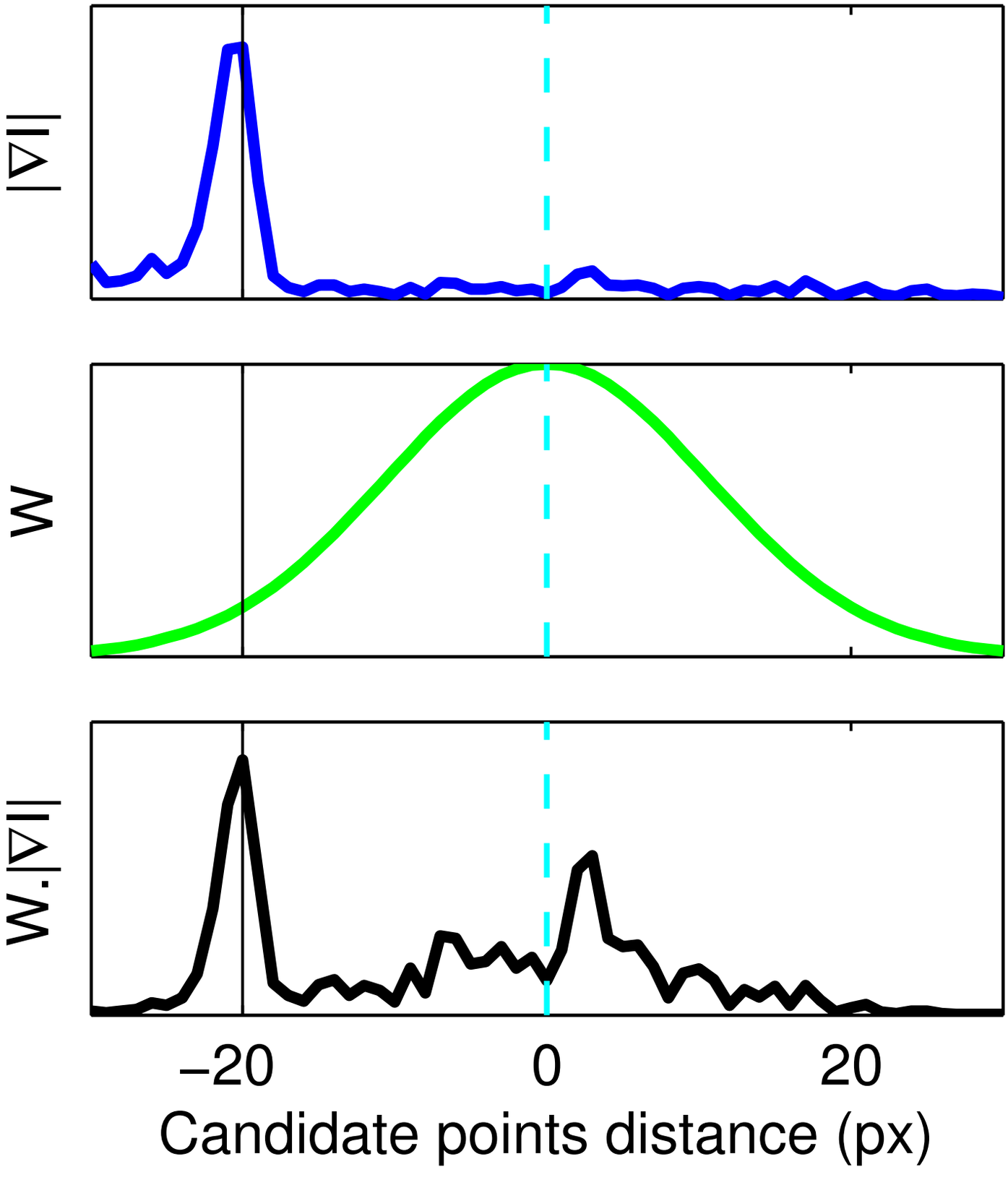

X_1
X_2
b_1
b_2
a_2
a_1
b_2^*
a_2^*

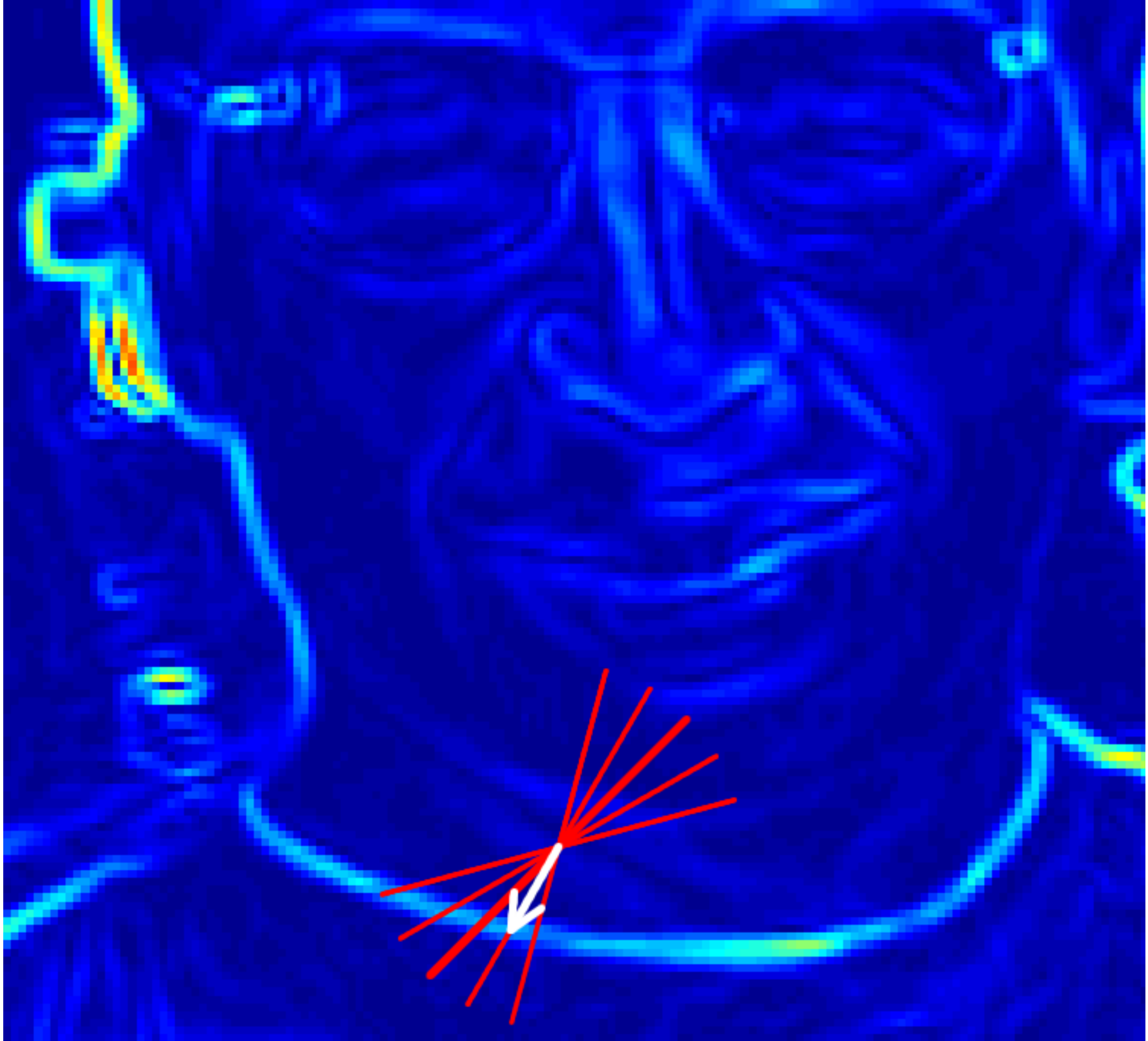

X1
X2
b1
b2
l1
a2
a1
b2*
a2*
l2
k2
k1

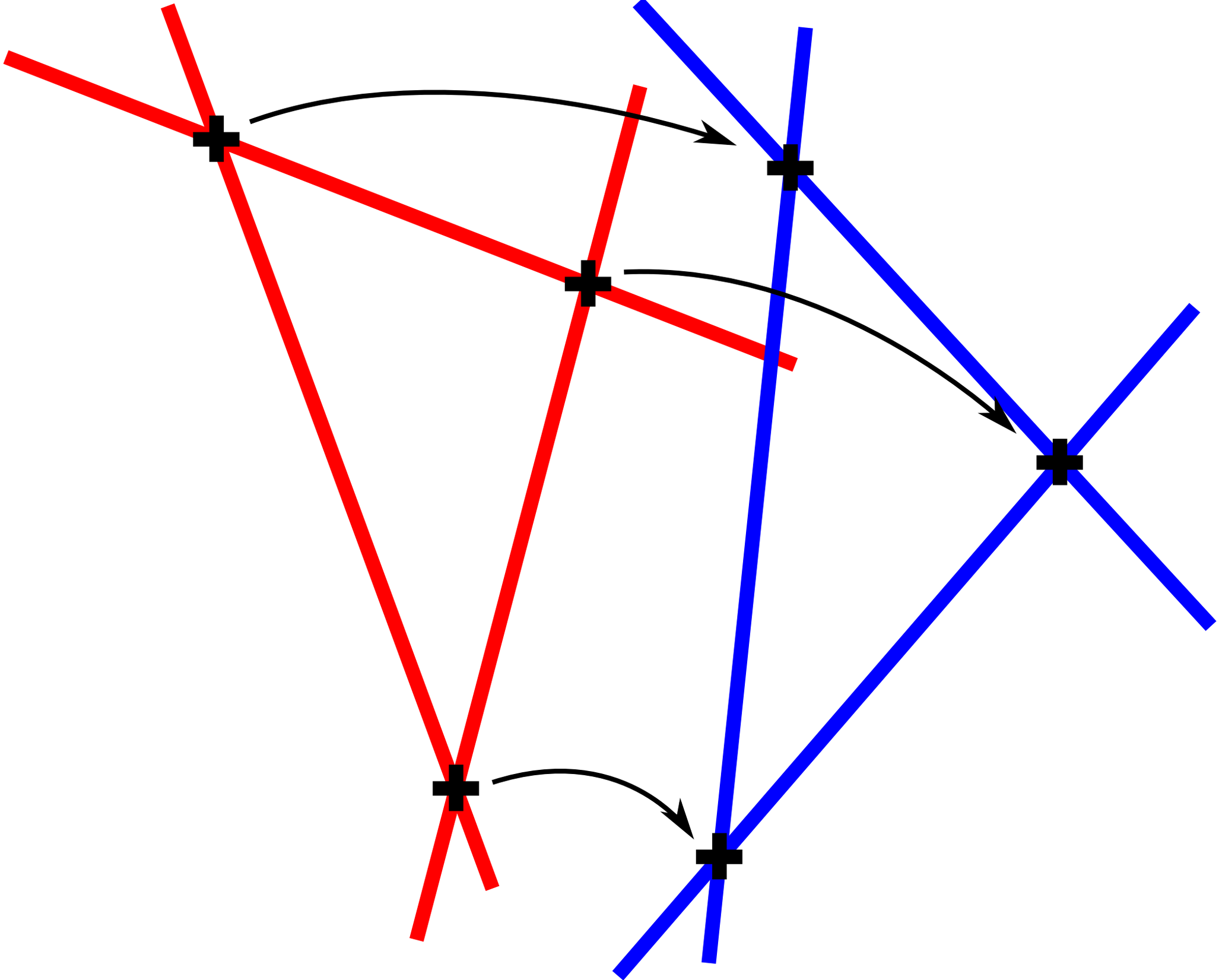

X1
X2
c1
c2
b1
b2
l1
a2
a1
b2*
a2*
l2
k2
k1

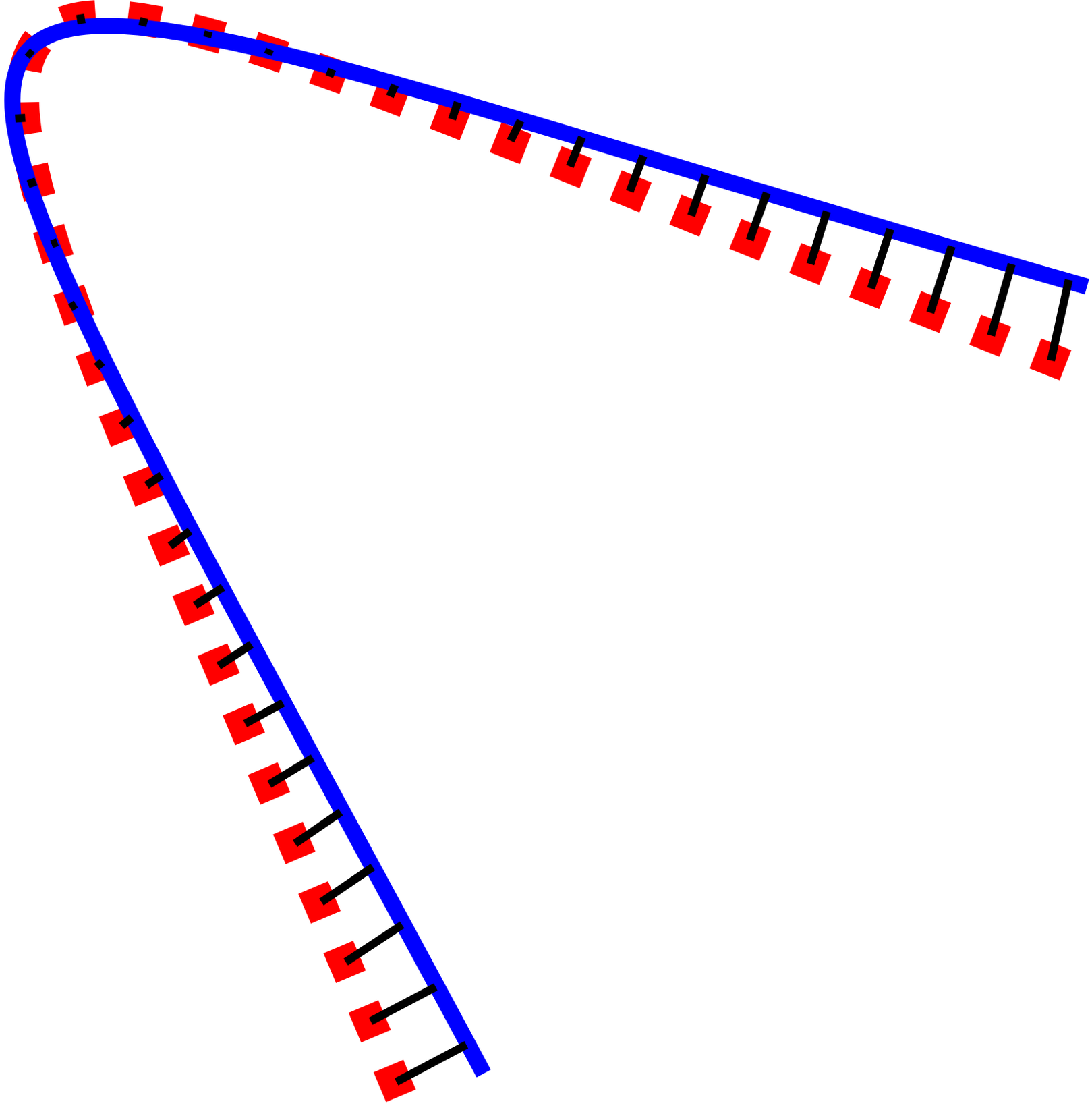


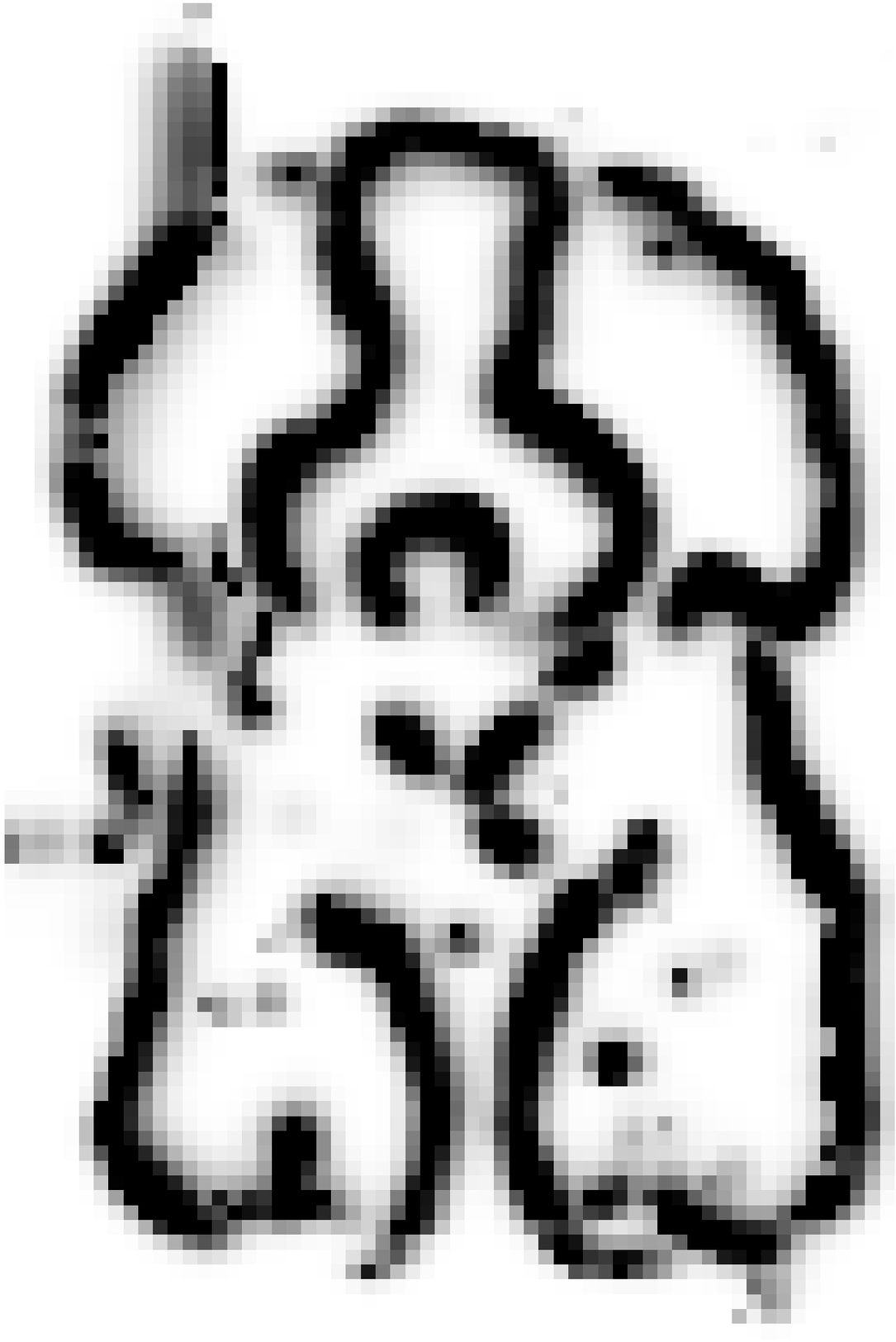


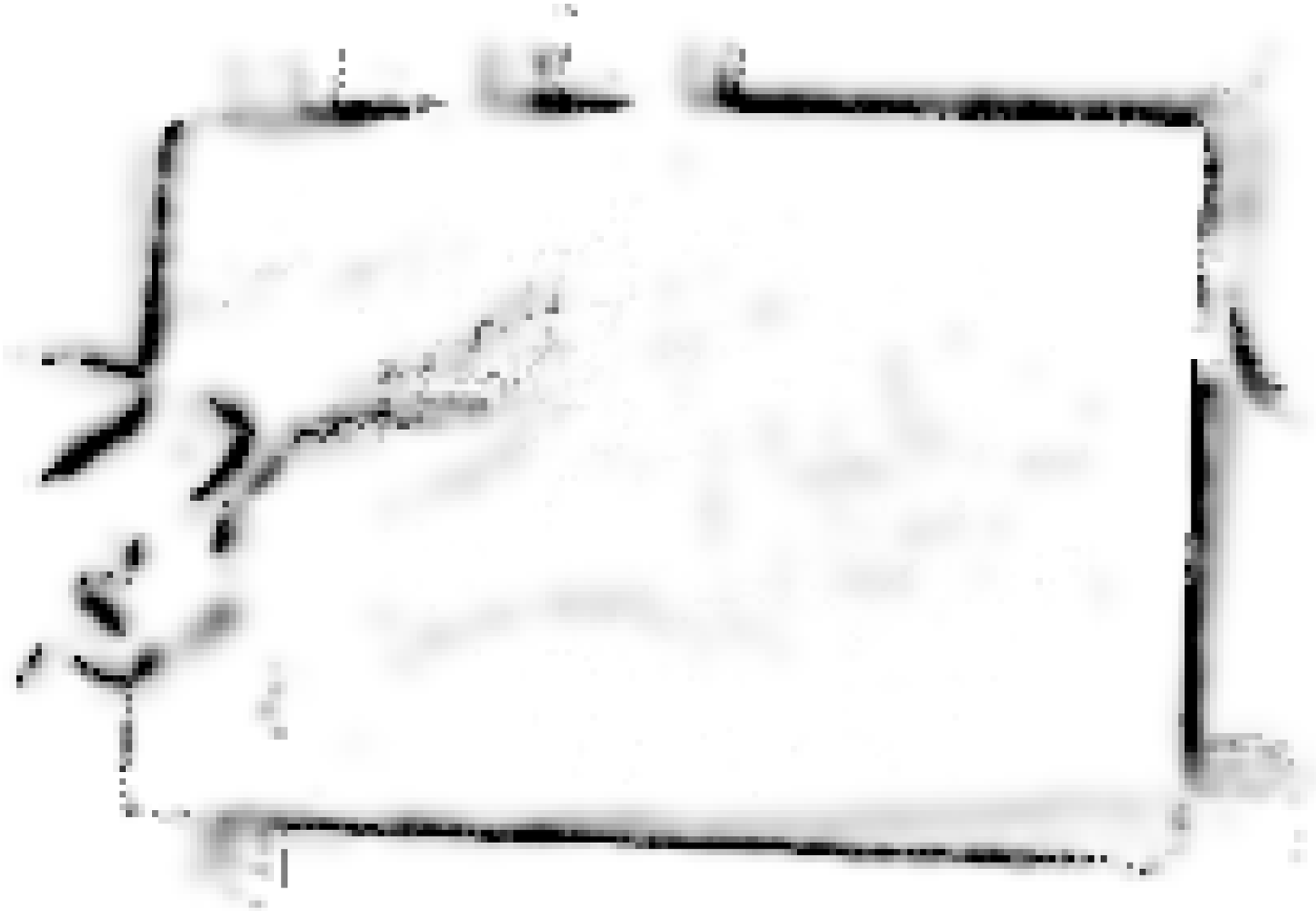

|∇I|
W
W.|∇I|
−20
0
20
Candidate points distance (px)

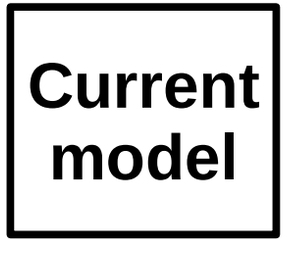


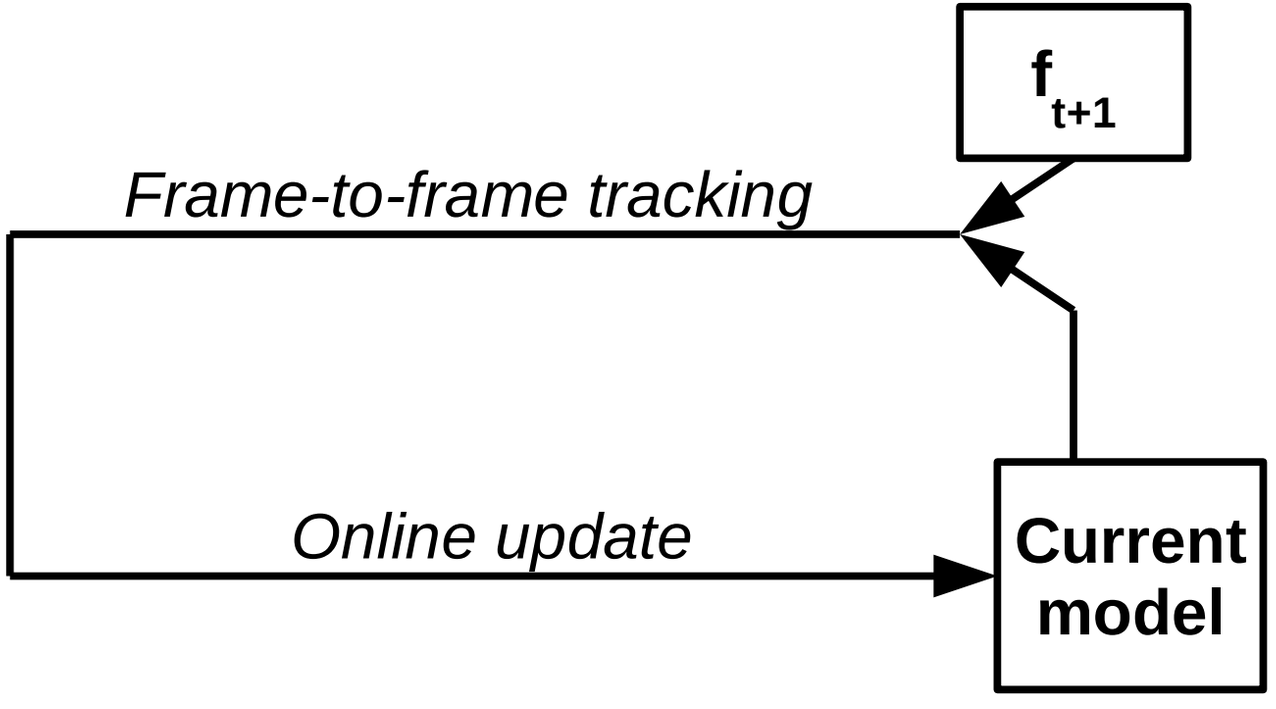


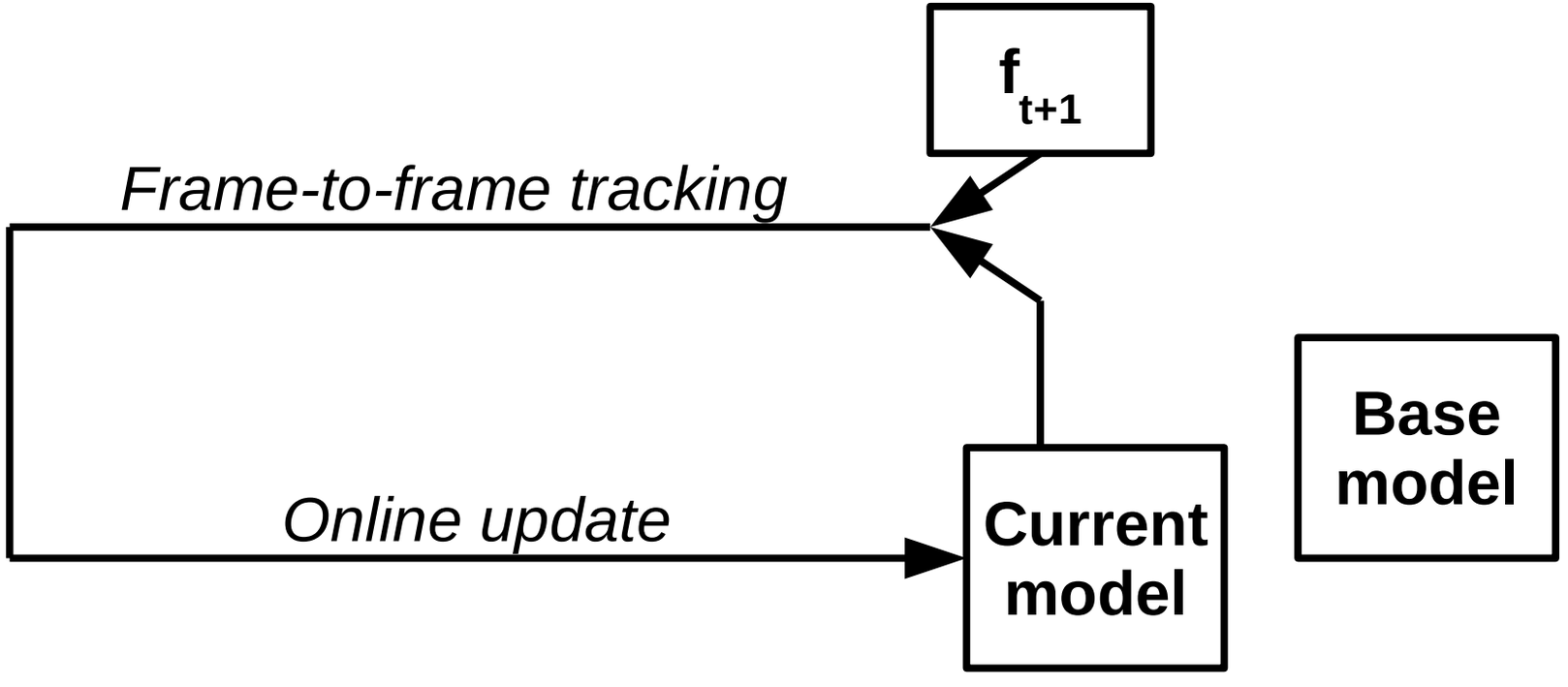


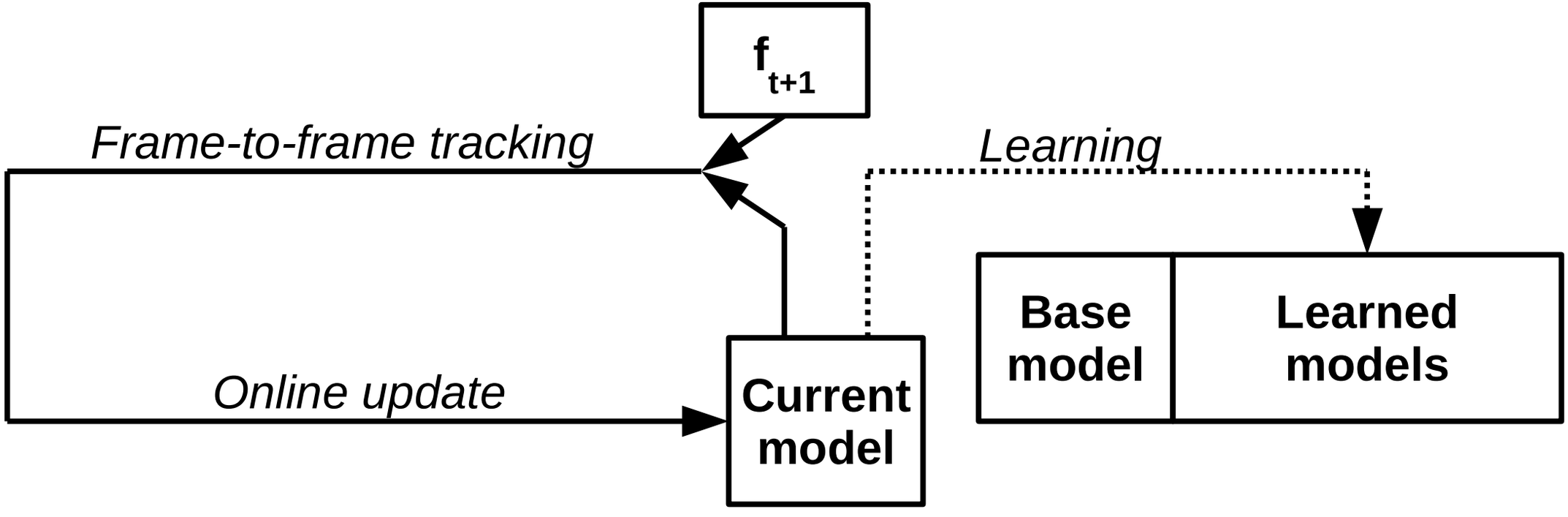


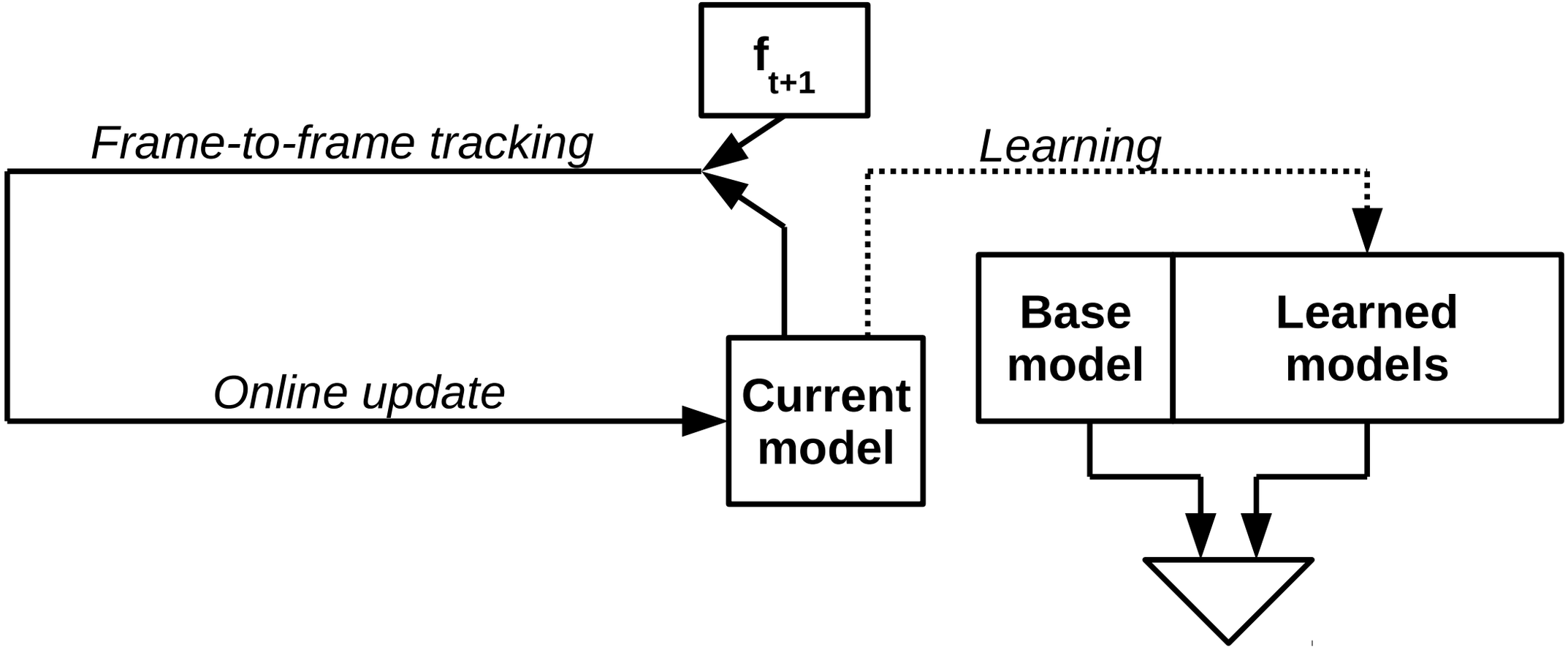


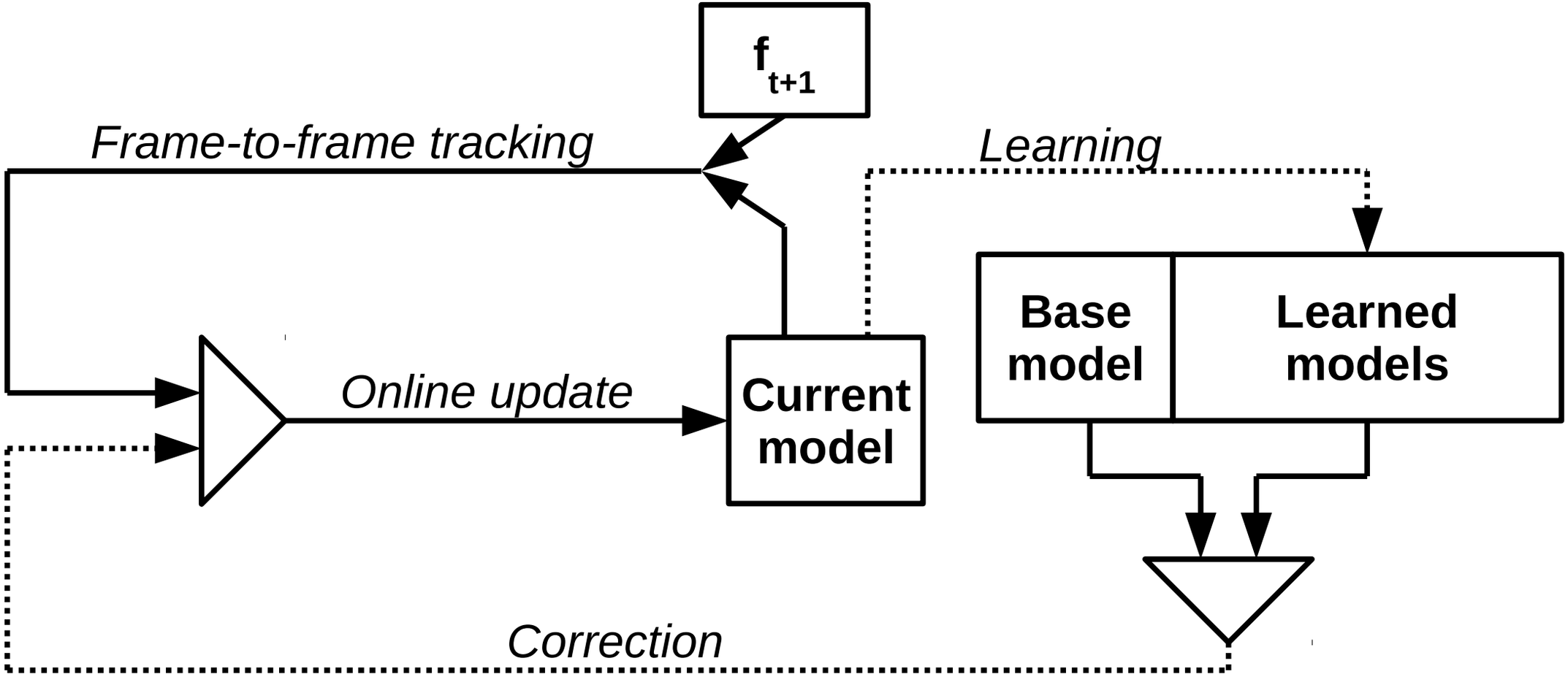

Current
model

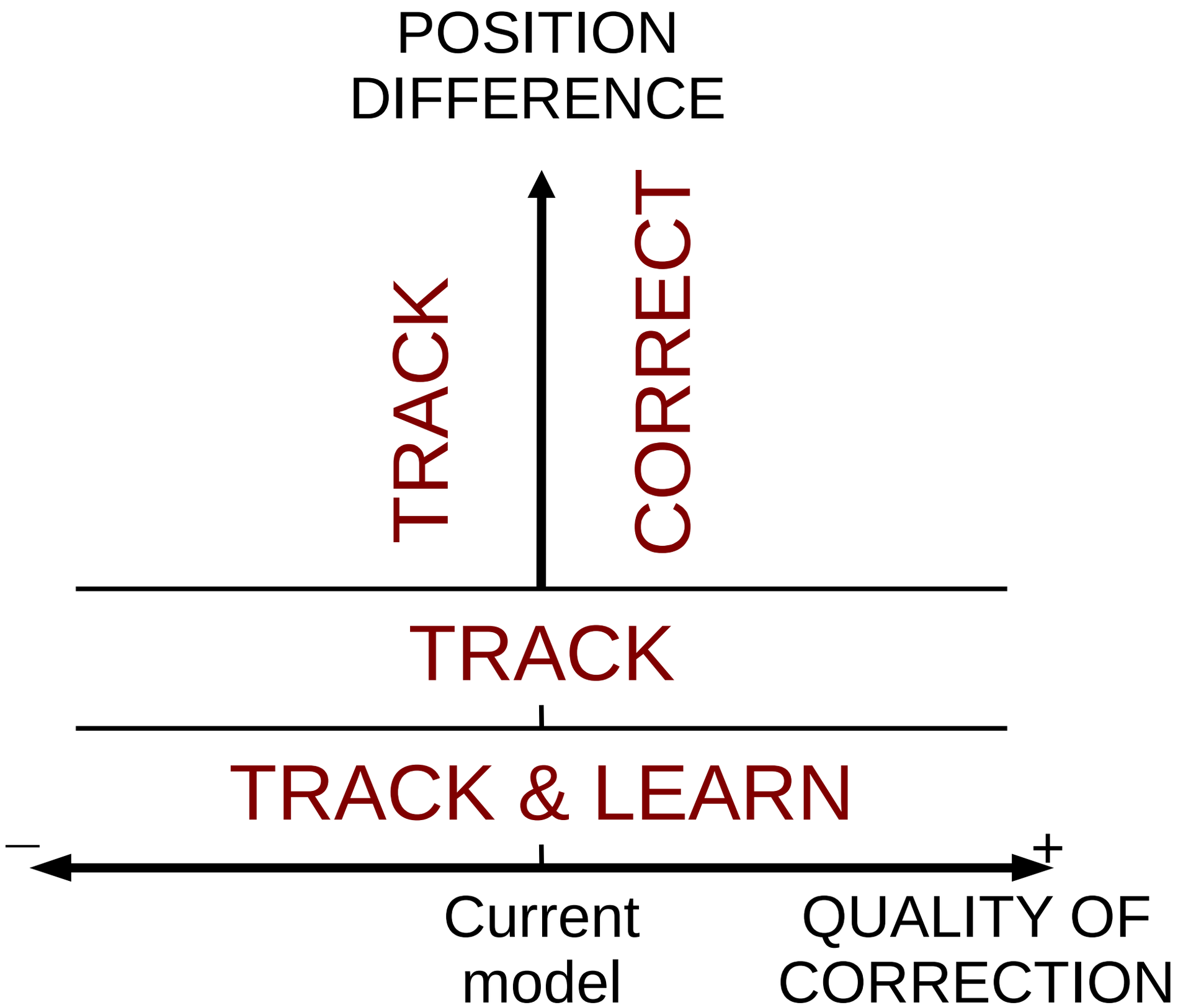

f_{t+1}
Frame-to-frame tracking
Online update
Current model

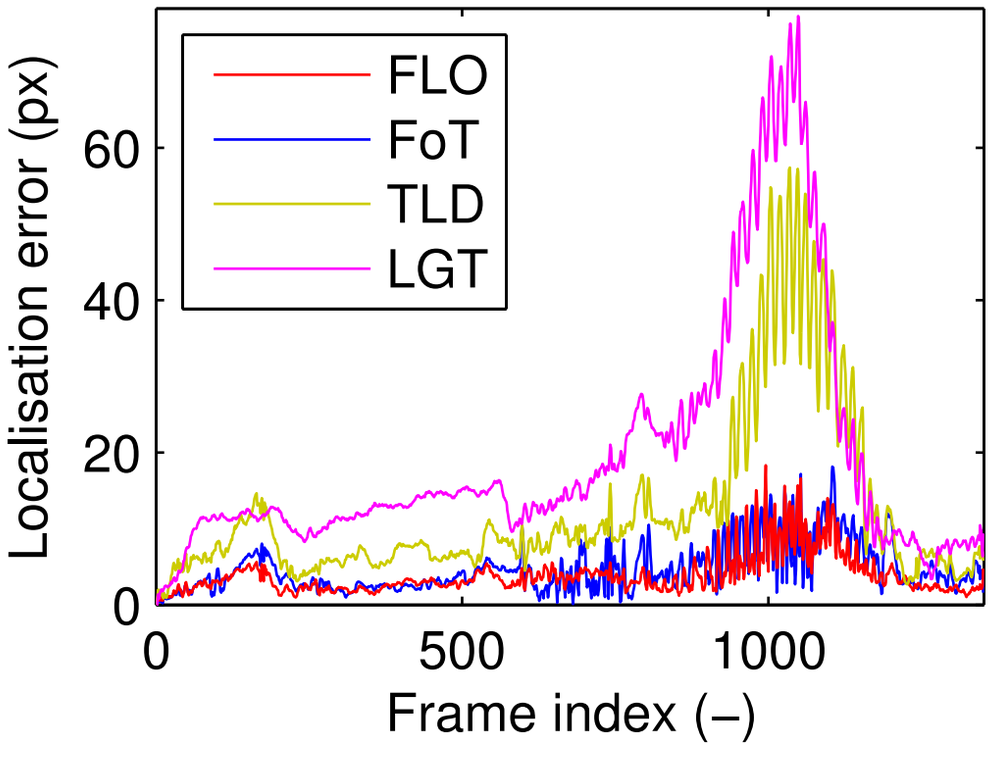

f_{t+1}
Frame-to-frame tracking
Online update
Current model
Base model

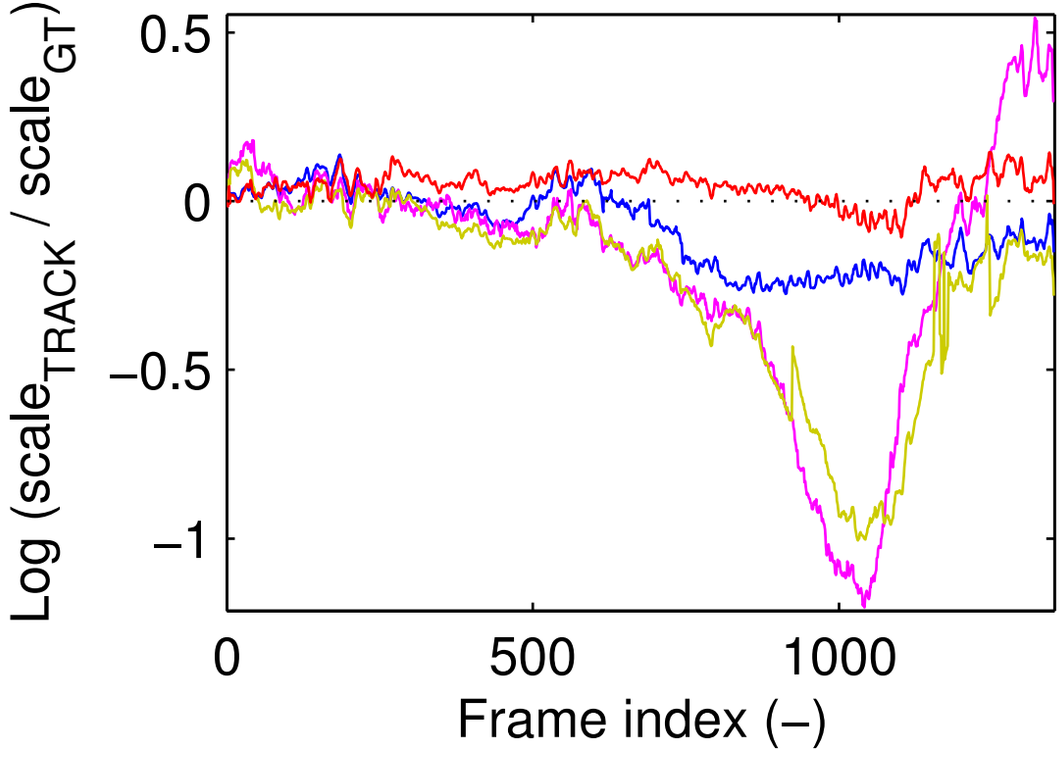

f_{t+1}
Frame-to-frame tracking
Learning
Base model
Learned models
Online update
Current model

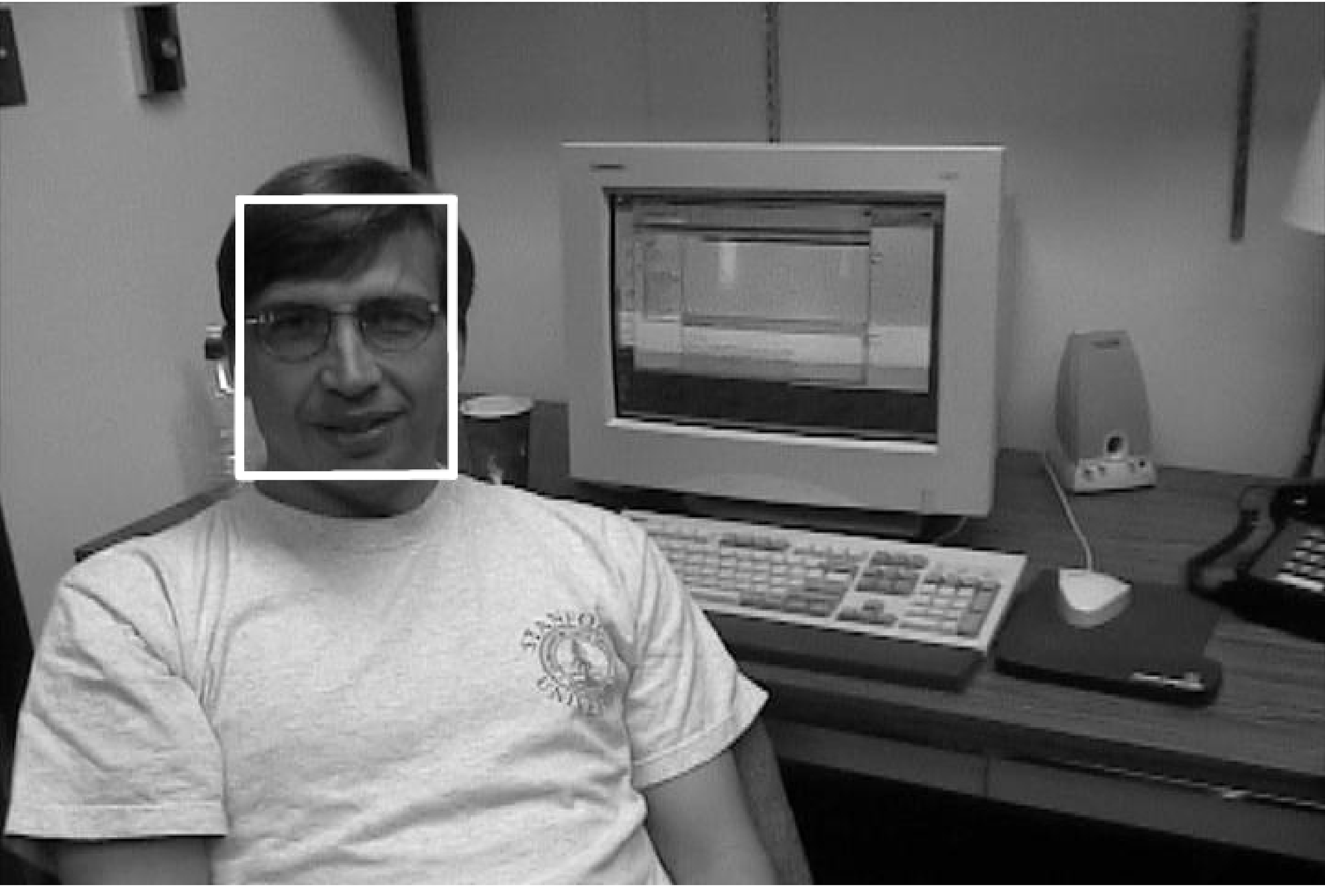

f_{t+1}
Frame-to-frame tracking
Learning
Base model
Learned models
Online update
Current model

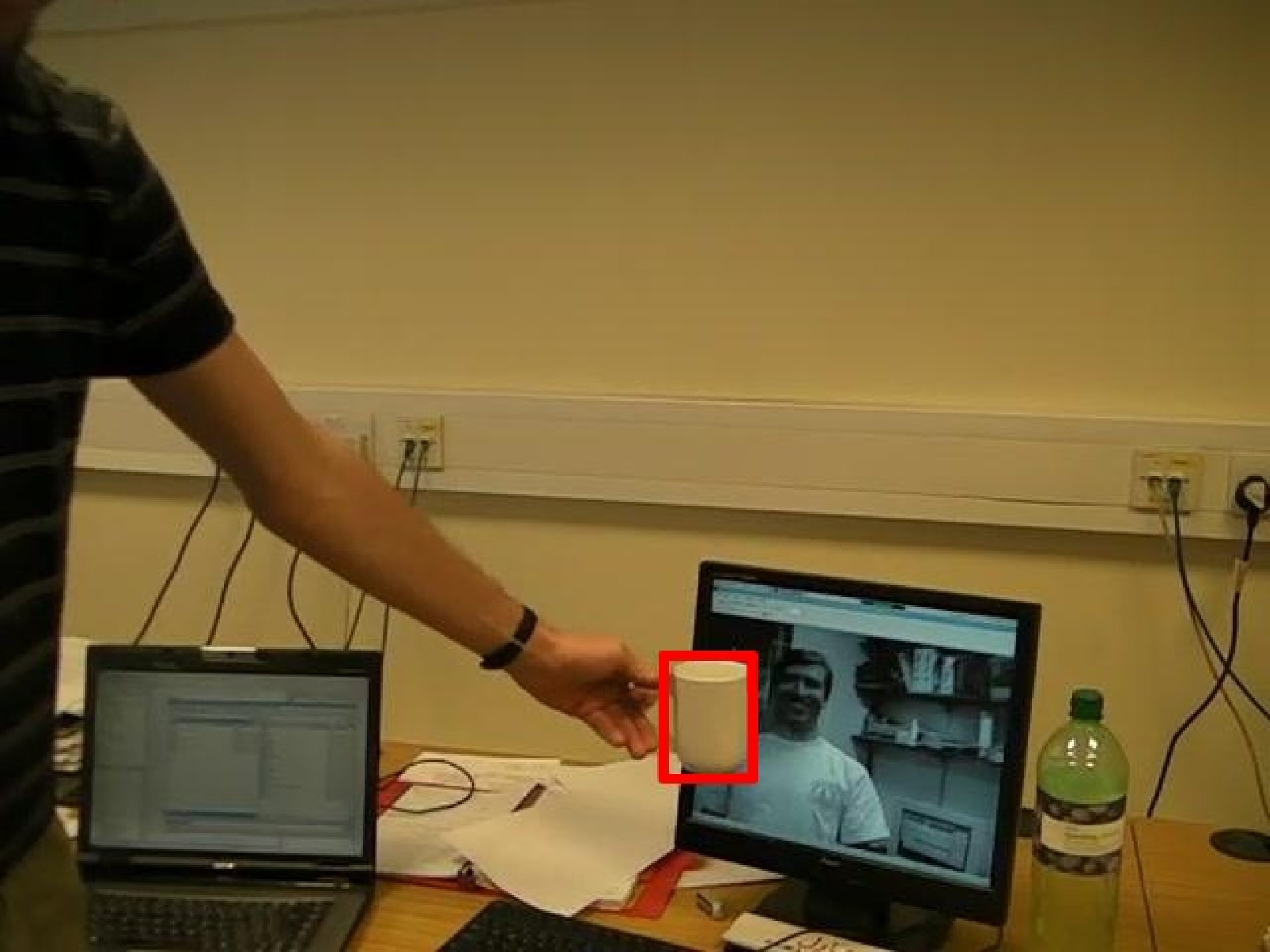

f_{t+1}
Frame-to-frame tracking
Learning
Base model
Learned models
Online update
Current model
Correction

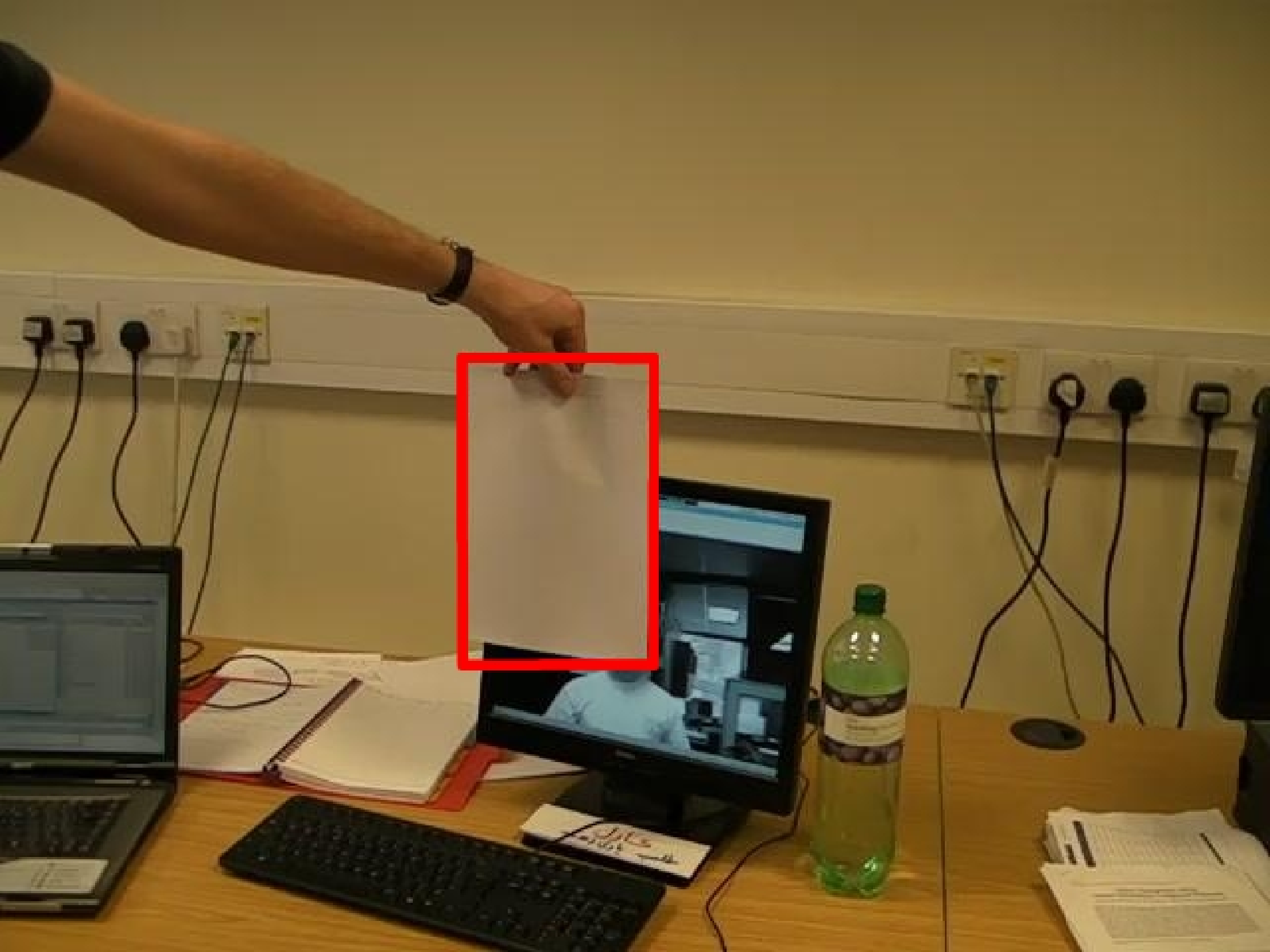

POSITION
DIFFERENCE
TRACK
CORRECT
TRACK
TRACK & LEARN
–
+
Current
model
QUALITY OF
CORRECTION

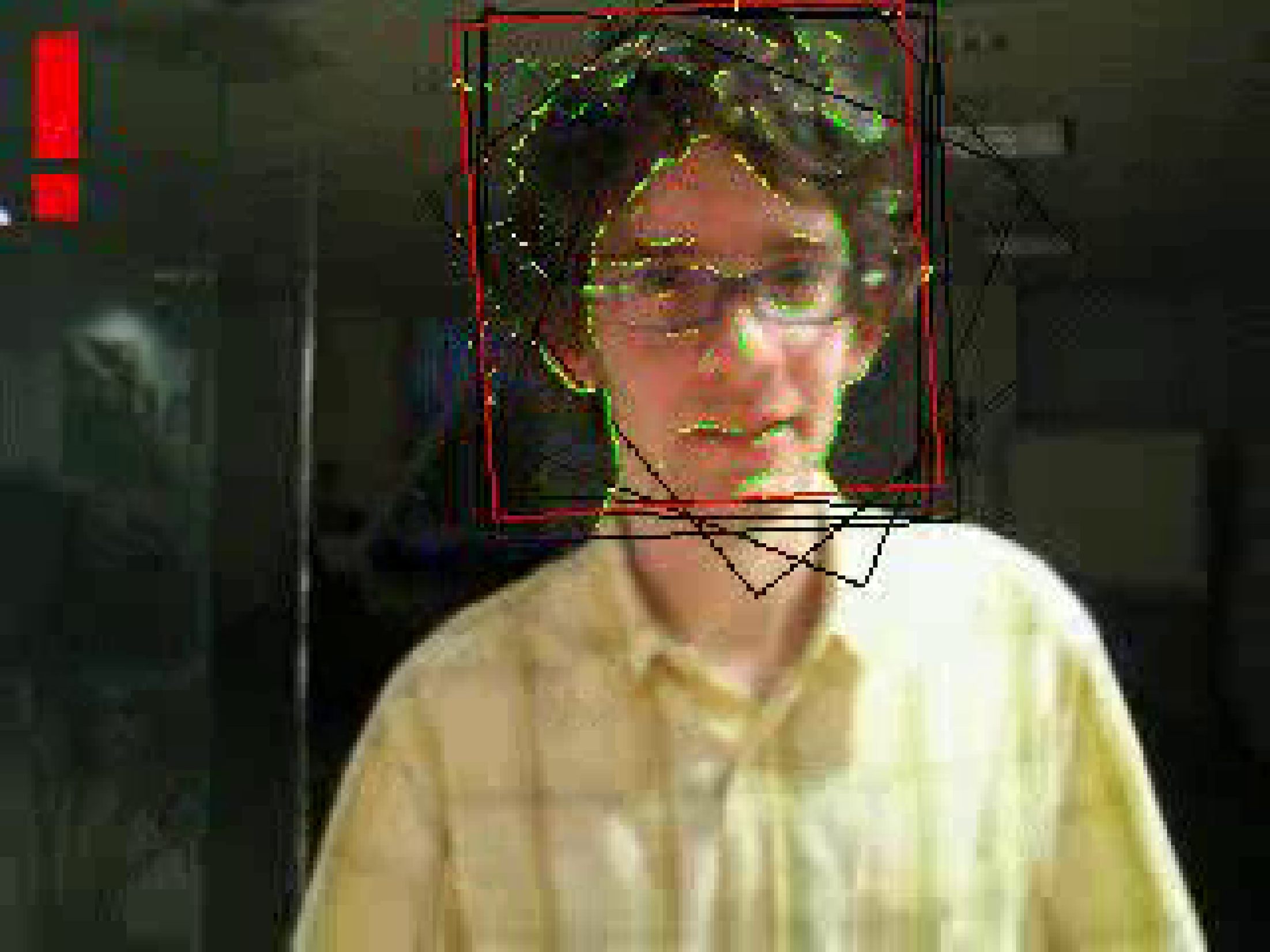

FLO
FoT
TLD
LGT
Localisation error (px)
0
20
40
60
0
500
1000
Frame index (–)

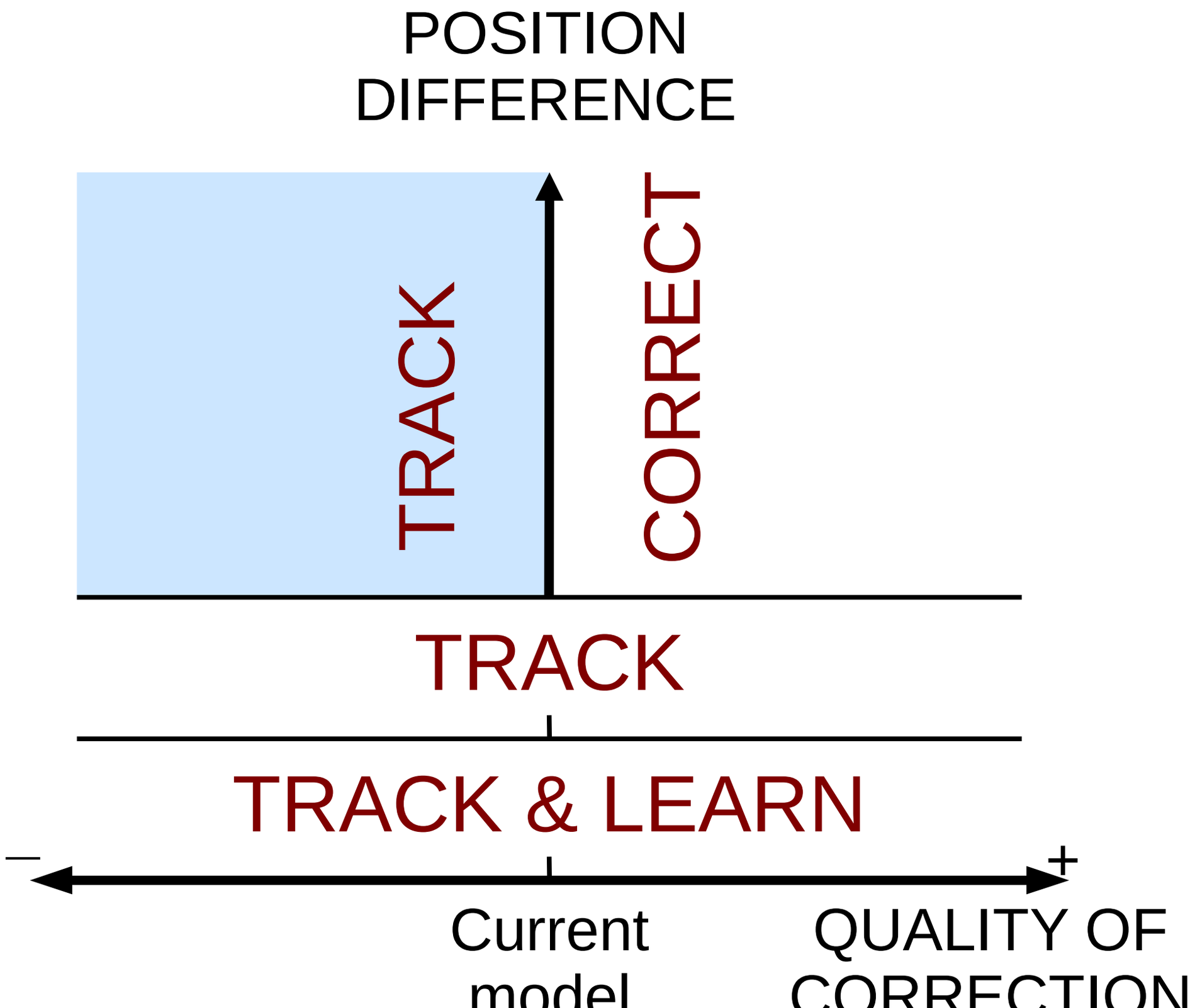

Log (scale$_{TRACK}$ / scale$_{GT}$)
0.5
0
–0.5
–1
0
500
1000
Frame index (–)

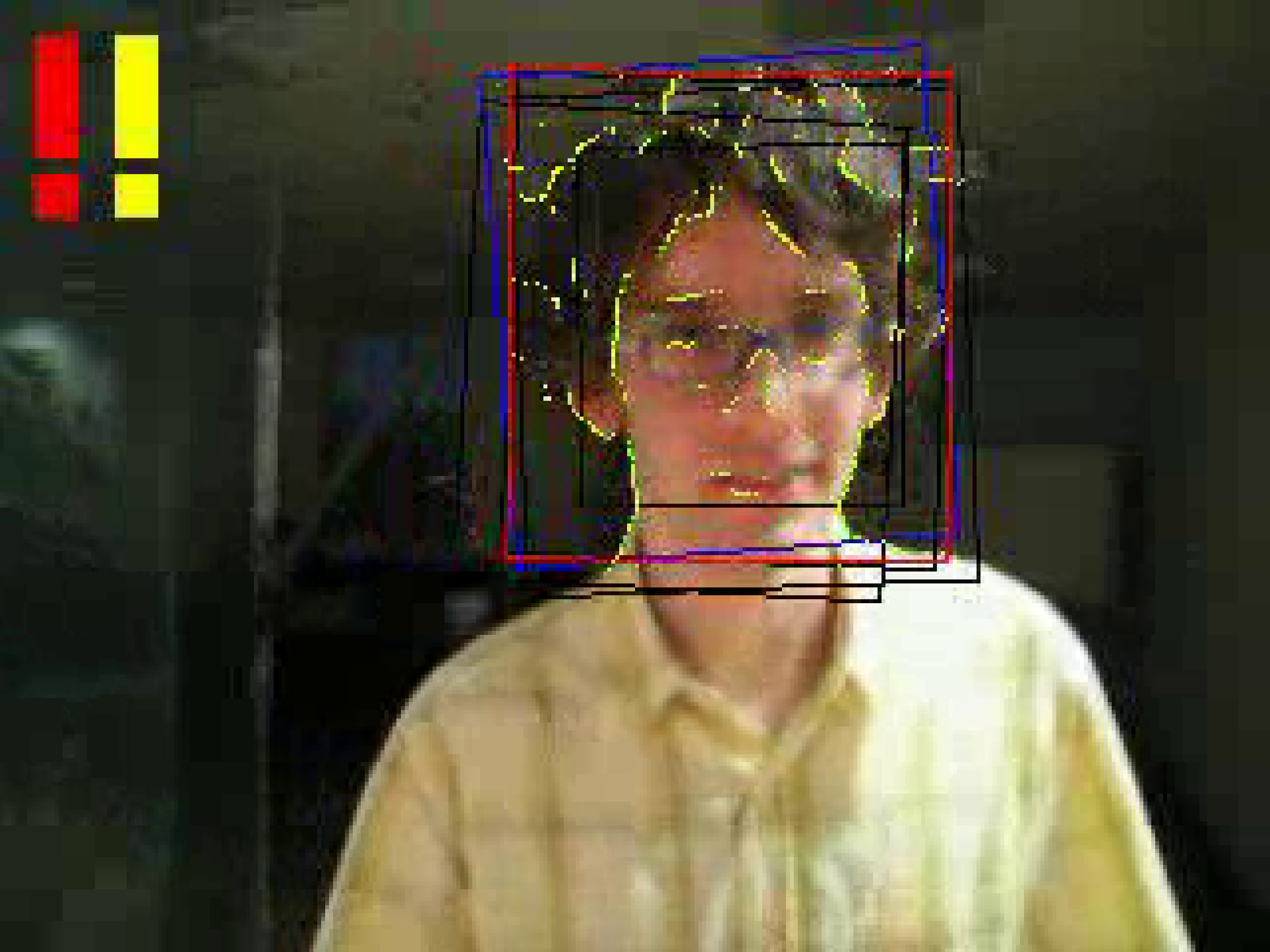


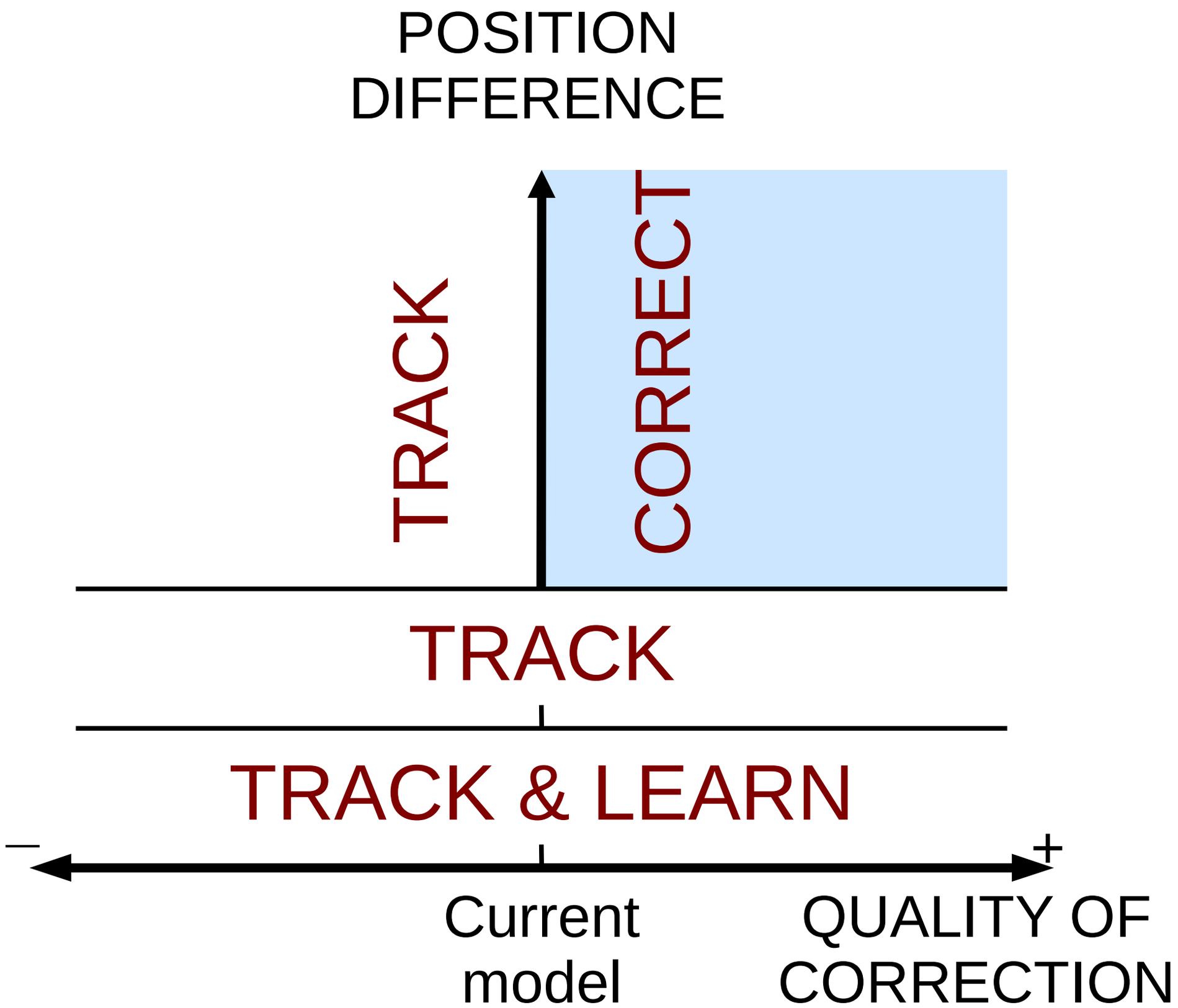


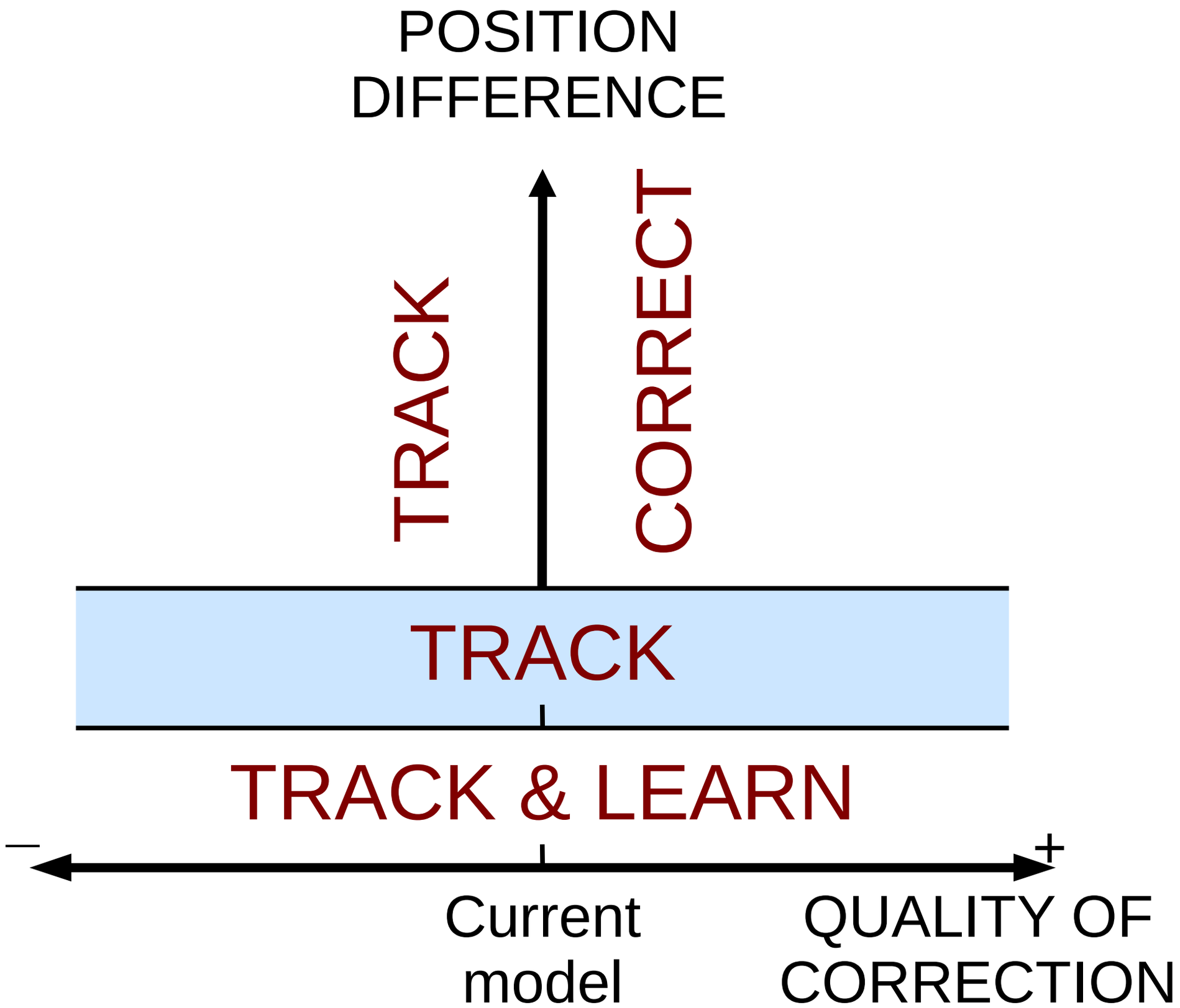


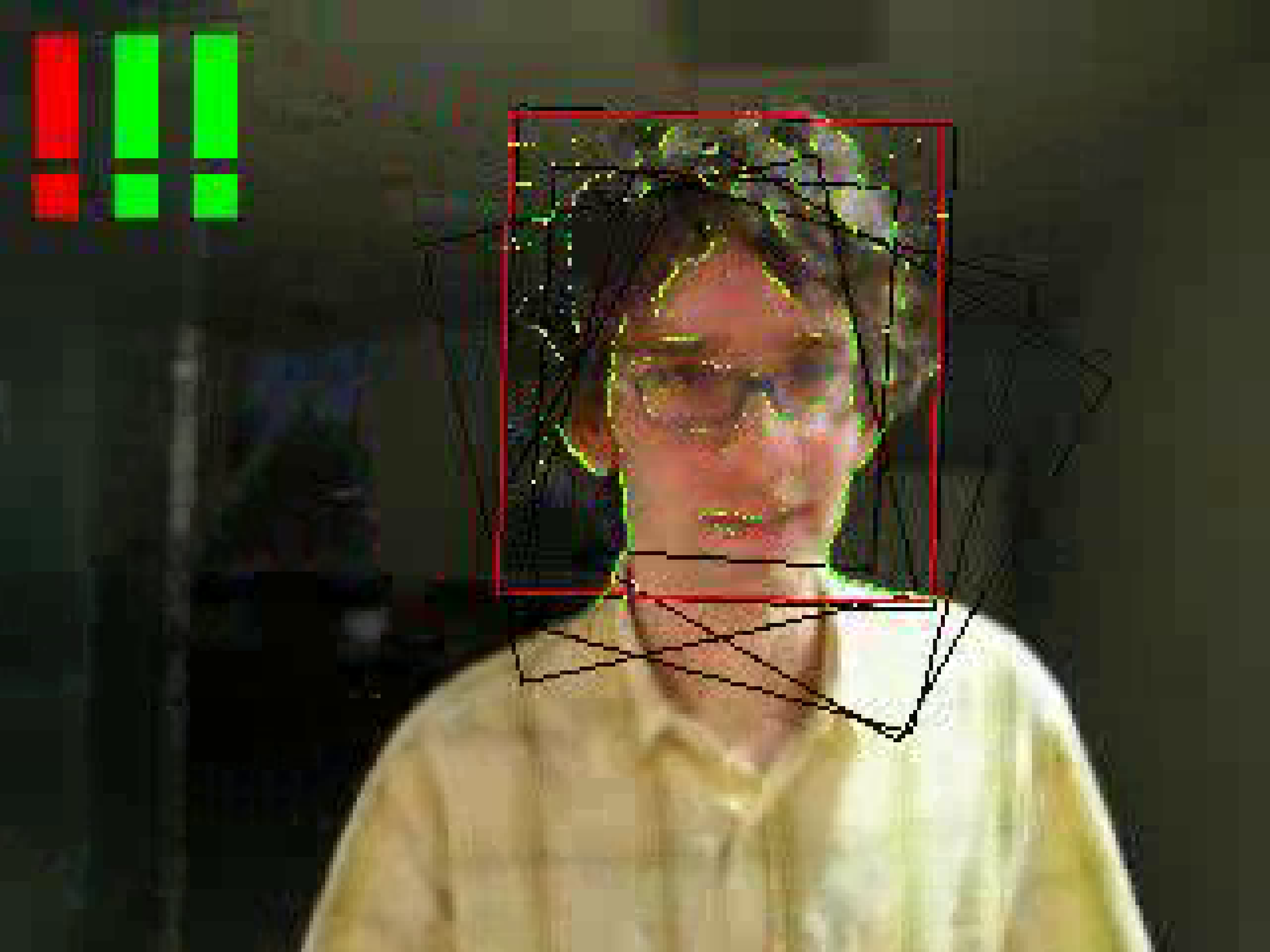


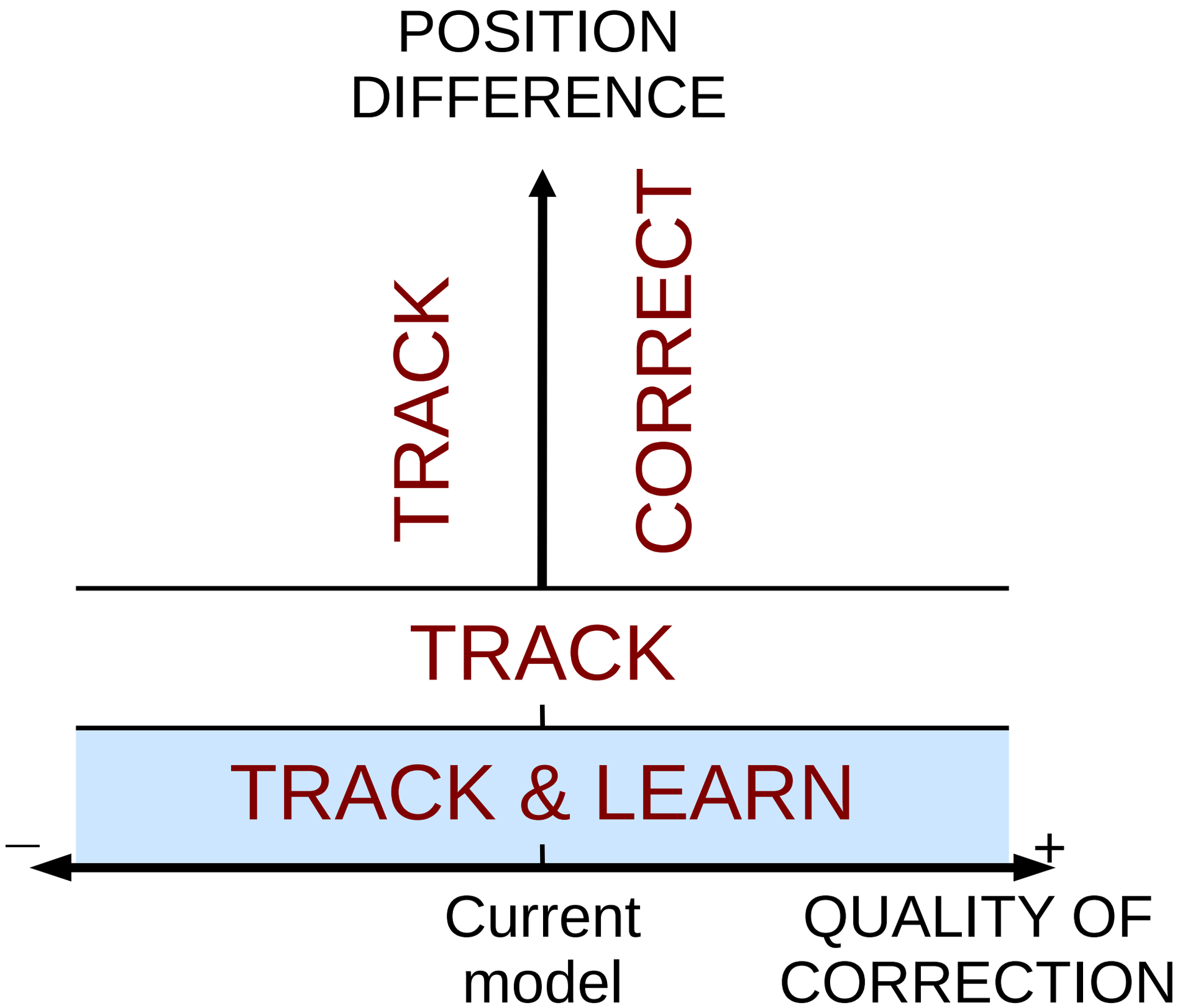

POSITION
DIFFERENCE
TRACK
CORRECT
TRACK
TRACK & LEARN
–
+
Current
model
QUALITY OF
CORRECTION




POSITION
DIFFERENCE
TRACK
CORRECT
TRACK
TRACK & LEARN
–
+
Current
model
QUALITY OF
CORRECTION


POSITION
DIFFERENCE
TRACK
CORRECT
TRACK
TRACK & LEARN
–
+
Current
model
QUALITY OF
CORRECTION




POSITION
DIFFERENCE
TRACK
CORRECT
TRACK
TRACK & LEARN
–
+
Current
model
QUALITY OF
CORRECTION