

Multi-Object Tracker Decision Worksheet

Four questions, four trackers, three topologies — pick the right MOT stack in ninety seconds.

AI for Video Engineering · Lesson 2.8

Four-question decision tree

1. Is the camera moving?

- ☐ Yes (phone, drone, dashcam) → BoT-SORT (CMC).
- ☐ No (fixed CCTV) → continue to question 2.

2. Non-linear / acrobatic motion?

- ☐ Yes (dance, gymnastics, animals) → OC-SORT.
- ☐ No (pedestrians, vehicles) → continue.

3. Many streams / modest hardware?

- ☐ Yes (10+ streams / server) → ByteTrack.
- ☐ No → continue to question 4.

4. Long-occlusion identity matters?

- ☐ Yes (sports, broadcast, POI) → BoT-SORT-ReID.
- ☐ No (counting, dwell) → ByteTrack.

Metric cheat sheet

- ☐ MOTA — detection-biased; legacy single number.
- ☐ IDF1 — association-focused; detection-blind.
- ☐ HOTA — balanced det + assoc; the 2026 default.

Four pitfalls — pre-launch checklist

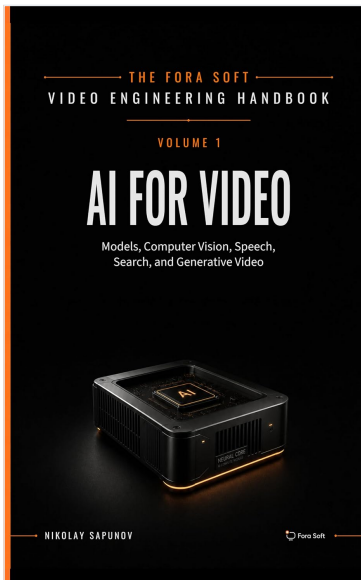
- ☐ Detector recall on real data above 90 percent.
- ☐ DeepSORT is not the default tracker.
- ☐ track_buffer / match_thresh tuned on real footage.
- ☐ Cross-camera IDs handled by separate Re-ID model.

Pipeline stages

- ☐ 1 · Detection — YOLO, RT-DETR (recall > 90 percent).
- ☐ 2 · Motion prediction — Kalman or observation-centric.
- ☐ 3 · Association — IoU or IoU + appearance.
- ☐ 4 · Lifecycle — birth / death / re-id thresholds.

Deployment topology — default tracker

Topology	Tracker	Cost owner	Latency
Browser	ByteTrack (ONNX Web)	User device	~16 ms
Edge (many)	ByteTrack on Jetson	Hardware	~30 ms
Edge (acc.)	BoT-SORT-ReID	Hardware	~50 ms
Server	BoT-SORT-ReID on CPU	Cloud / stream	~100 ms



GO DEEPER · VOLUME 6 OF 7

AI for Video: A Complete Guide to Intelligence in Video Products

From Models and Inference to Computer Vision, Speech, Search, and Generative Video

By Nikolay Sapunov, CEO at Fora Soft · 1,444 pages · Kindle & paperback on Amazon

[Read it on Amazon →](#)
amazon.com/dp/B0HH6TKPZ1

This worksheet picks a tracker, a topology and a metric in four questions. Volume 6 covers the computer vision behind it in depth: detection, tracking and re-identification, pose and segmentation, deployment topologies from browser to server, and the ML minimum you need to judge results.

WHAT YOU WILL FIND INSIDE

- Detection models and recall on real footage
- Tracking, re-identification and occlusion
- Deployment from browser and edge to server
- Evaluating vision models with the right metric

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI

Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF

Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery

Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing

Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs

Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video

Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Choosing a tracker for cameras, drones or sports video?

We will test ByteTrack, OC-SORT and BoT-SORT on your footage and recommend a stack with HOTA and cost per stream. 250+ video products shipped since 2005.

[Book a 30-min call →](#)
[WhatsApp →](#)
[Email us →](#)