



VISUAALISTEN TRANSFORMERIEIN LIIKEPERUSTAINEN TILANNEÄLY

Professori Laura Ruotsalainen,
tietojenkäsittelytiede, Helsingin yliopisto
Suomen tekoälykeskus FCAI



PROJEKTIN TIEDOT

VISUAALISTEN TRANSFORMERIEN AVULLA
LÄMPÖKAMERAN KUVISTA LIIKEPERUSTASEEN
TILANNEÄLYYN

*Thermal Visual Transformer for Motion Induced
Situational Intelligence (THERMORFER)*

Toteuttaja: Helsingin yliopisto, tietojenkäsittelytieteen osasto

MATINE-rahoitus 154 628€

Hankkeen kesto 2 vuotta, käynnissä 2. tutkimusvuosi



Taistelijan sijainti, ympäristössä havaittujen
kohteiden tunnistaminen ja etäisyyden määrittäminen



TUTKIMUKSEN PÄÄMÄÄRÄ JA TUTKIMUSSUUNNITELMA 1



ViT-menetelmään perustuva
semanttinen hahmontunnistus

semanttisen kartoituksen, eli tyypin
selvityksen, perusteella voidaan saada
likimääräinen tieto kohteen koosta



Inertiasensori- ja konenäköfuusio
syväneuroverkoilla

havainnoijan paikka ja liike



Kahteen edelliseen perustuva
kohteen etäisyyden määrittäminen

kohdetta seuraamalla ajan kuluessa voidaan
havainnointiin lisätä sen liiketieto => kohteen
etäisyys ja nopeus sekä tarkennettu arvio
sen koosta



MERKITYS PUOLUSTUSVOIMILLE

- Droonin tai taistelijan sijainti
- Hahmon tunnistus
- Hahmon etäisyys omaan sijaintiin
- Hahmon nopeus

Drooni

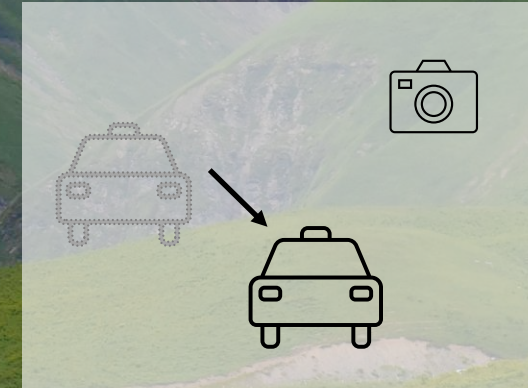


Kuvasekvenssi → liike → etäisyys

Mikä hahmo on?



Jalankulkija





TUTKIMUKSEN PÄÄMÄÄRÄ JA TUTKIMUSSUUNNITELMA 2

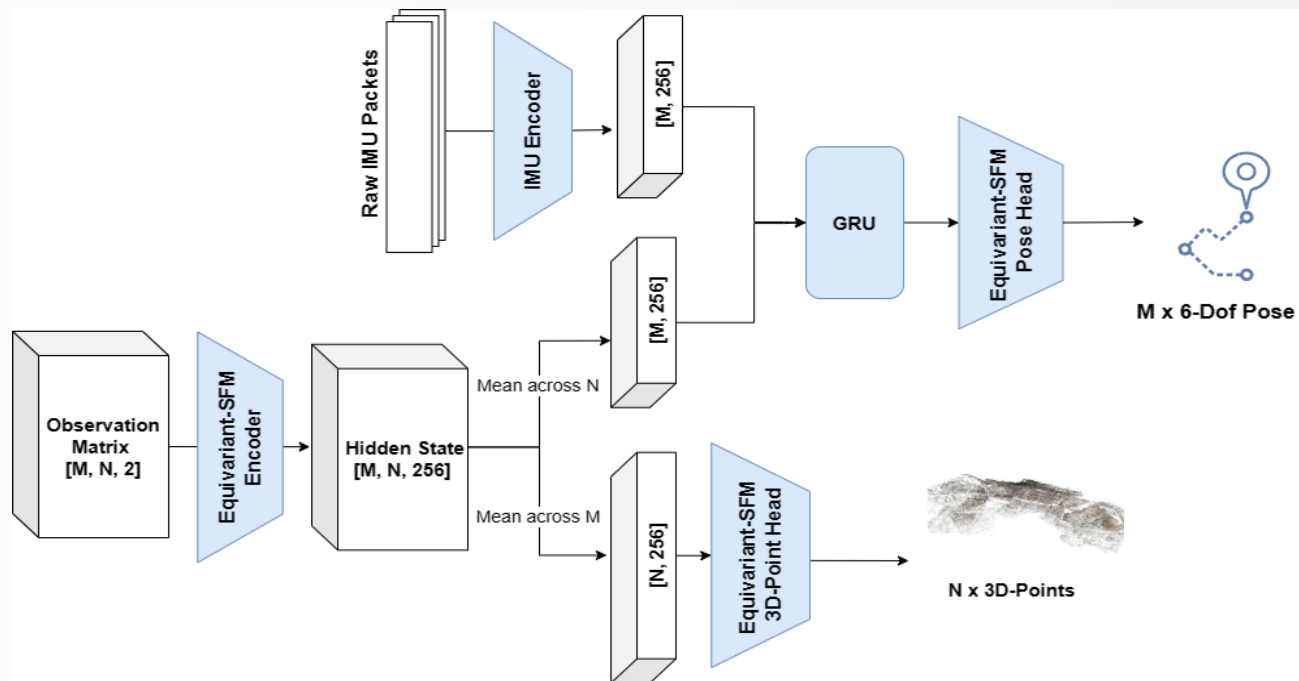
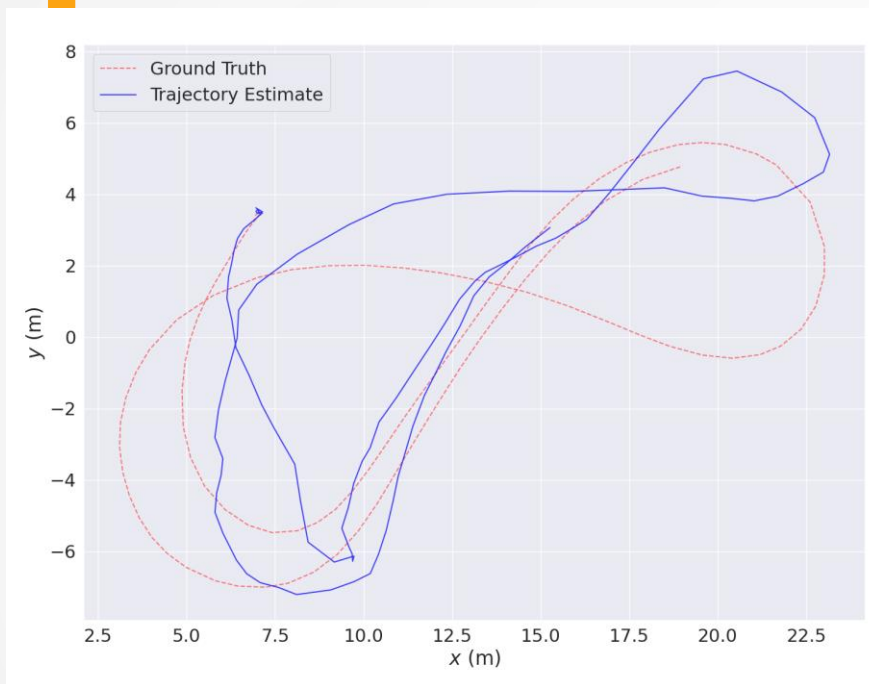
- Sijainnin määrittäminen kameran ja inertiasensorin mittausten fuusiolla
 - Lämpökamera Infraray E6+ V3.0, Hinta 3k€, koko 188x65x64 mm
 - Analog Devices Inc:n suhteellisen edullinen ja pieni taktisen tason IMU
- Alustoina droni sekä jalan kulkeva taistelija



Konenäön tehtäviin kehitetään menetelmiä perustuen viimeisimpiin syväoppimisen arkkitehtuureihin => **Vision Transformers (ViT)**



TRANSFORMER: SIJAINNIN MÄÄRITYS

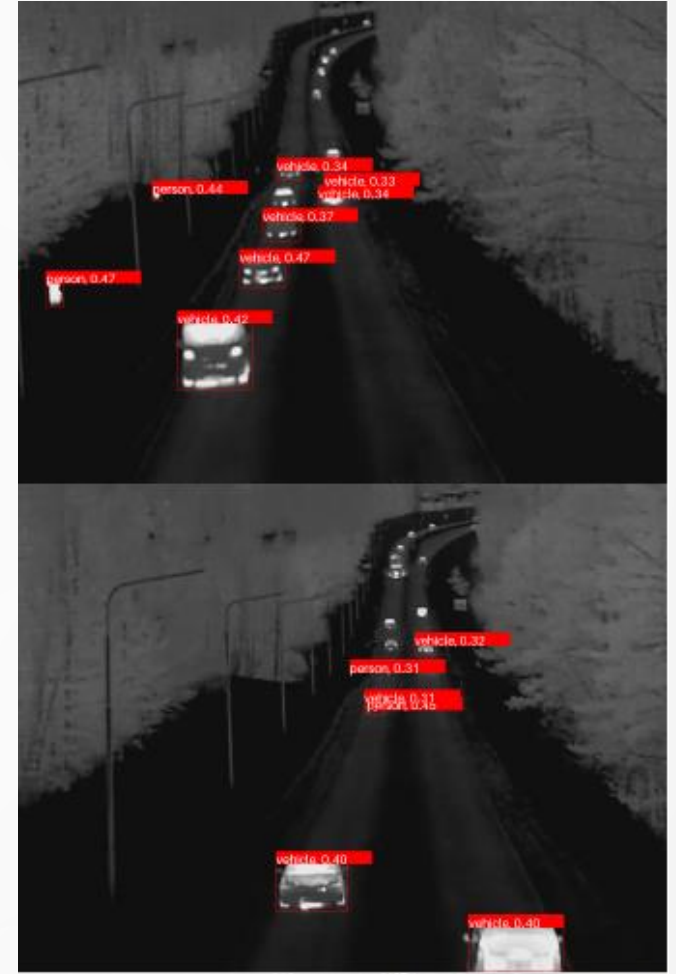


- Toistaiseksi kaikki transformer-pohjaiset sijaintia määrittävät menetelmät perustuvat optiseen virtaukseen (optical flow)
- Menetelmät eivät toimi kameran äkillisillä suurilla liikkeillä => menetelmämme toimii drooni-datalla



TRANSFORMER: HAHMON TUNNISTUS

- Transformer-pohjainen hahmontunnistus toimii myös lämpökuvilla
- Kuvat normalisoitu lämpötilan vaikutuksen häivyttämiseksi
- Lämpökuvat huonolaatuisia => tarvitaan paljon menetelmän opettamiseen
- Suuri tarve itseohjautuville menetelmille



R. Gao and L. Wang. MeMOTR: Long-term memory-augmented transformer for multi-object tracking (2023), ICCV.



ETÄISYYDEN MÄÄRITYS



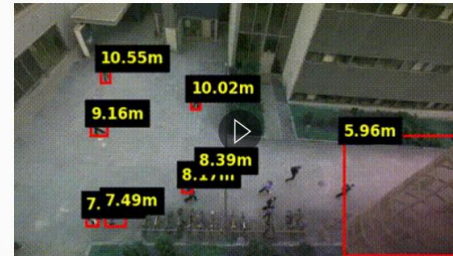
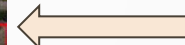
Etäisyys paikallaan olevaan kohteeseen



3D Seuranta



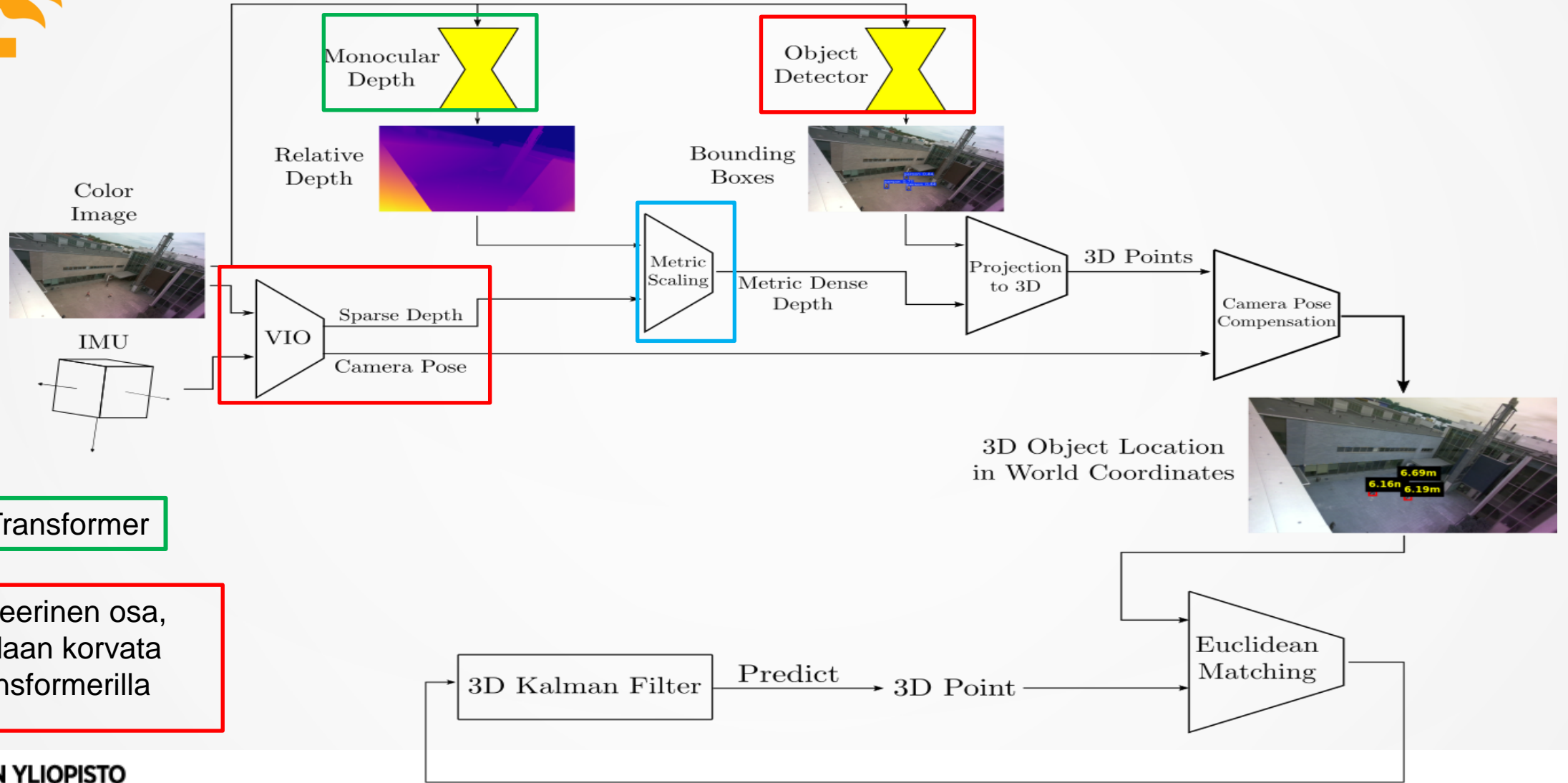
Suhteellinen etäisyys jokaiseen kuvapisteeeseen



- Etäisyyden määrittäminen yksilöllisellä kameralla yksinään on mahdotonta
- Liikkeen avulla etäisyys paikallaan oleviin kohteisiin
- Syväoppimisen avulla suhteellinen etäisyyskartta
- Suhteellinen kartta + paikallaan olevien kohteiden etäisyys => rajoitukset etäisyyden laskentaan



KONEOPPIMISEN ARKKITEHTUURI



Transformer

Geneerinen osa,
voidaan korvata
transformerilla



DATAN KERÄYS

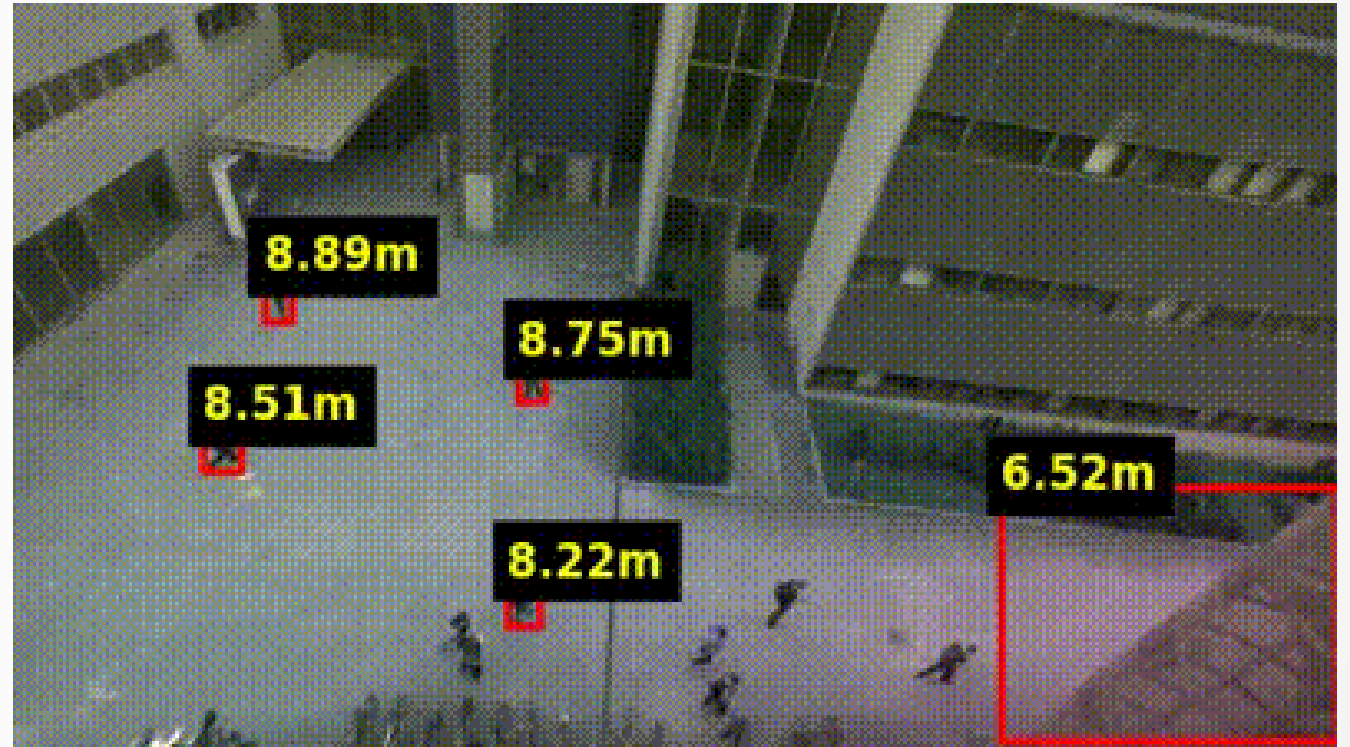
- Dataa kerättiin yliopiston katoilta => ”drooni”
- Oikea etäisyys lidarilla
- Tässä testissä värikuvia





ALUSTAVAT TULOKSET

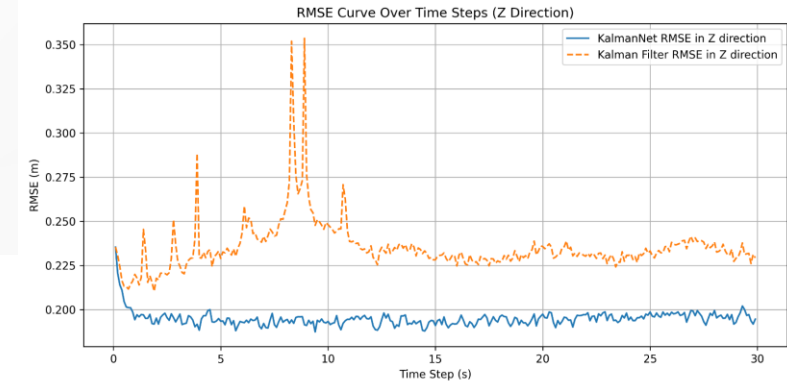
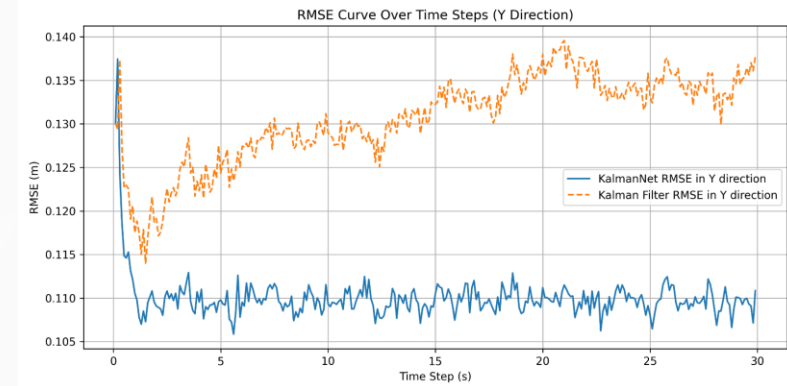
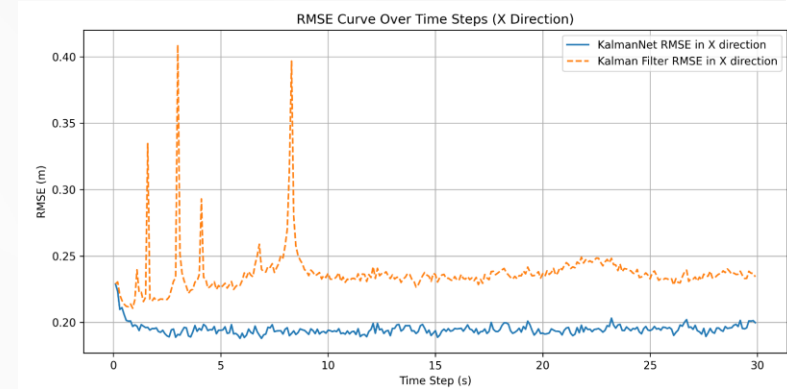
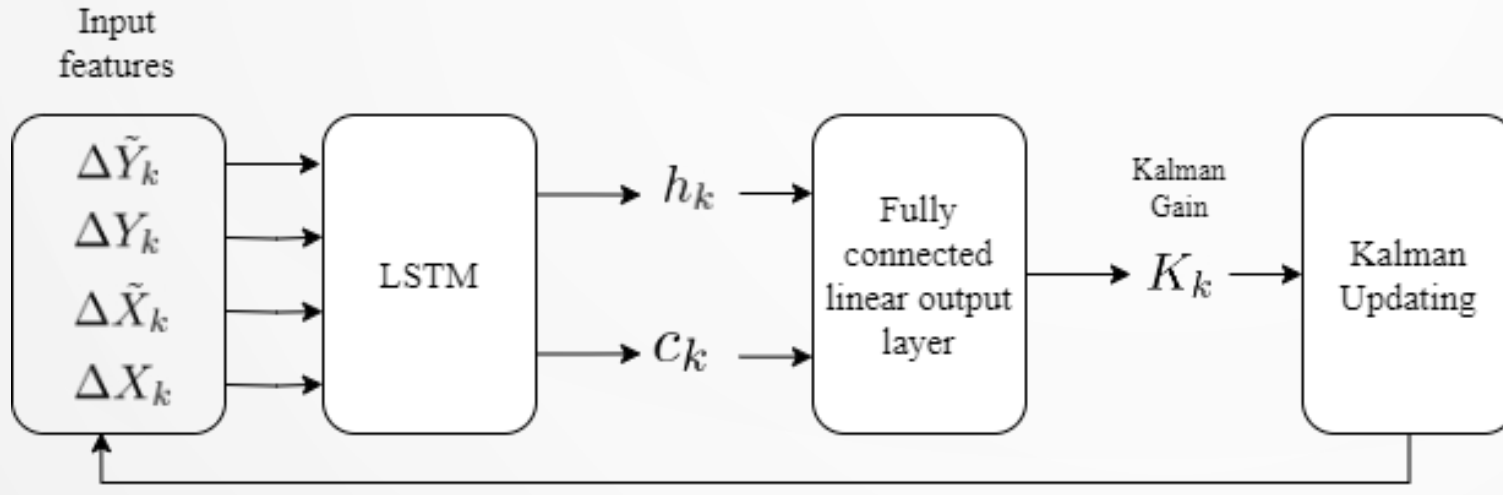
- Menetelmä laskee oikean etäisyyden
- Seuraa kohdetta melko tarkasti
- Löytää kohteen uudelleen





KALMAN-SUODIN JA SYVÄOPPIMINEN

- Mittaus- ja prosessivarianssien määrittäminen on vaikeaa => tarkkuus kärsii
- Tutkimuksessa käytetään neuroverkkoja parametrien määrittämiseen
- Uutta: useiden mittausten käyttö samanaikaisesti





JULKAISUT

- Pajula, Joswig, Morrison, Sokolova, Ruotsalainen (2023). Attention-based Visual-Inertial Fusion, IEEE Journal of Indoor and Seamless Positioning.
- Joswig, Morrison, Sokolova, Ruotsalainen (2024). Plane Prior for RGB-D based Visual Odometry. Intl. Conference on Indoor Positioning and Indoor Navigation (IPIN).
- Joswig, Pelvo, Ruotsalainen (2025). Elevated Tracking: 3D Object Detection, to be submitted.
- Gao, Joswig, Ruotsalainen (2025). Multi-source KalmanNet , to be submitted.
- Pajula, Pelvo, Ojala, Ruotsalainen (2024). Thermal Visual Transformer for Motion Induced Situational Intelligence. Kirjallisuuskatsaus, Matinen sarja.
- Pajula (2024). End-to-end Visual-Inertial Fusion for Ego-Motion Estimation. MSc thesis, University of Helsinki.
- Pelvo (2024). Transformer-based Object Detection and Tracking in Thermal Images MSc thesis, University of Helsinki.

