

UNIVERZITET U BEOGRADU  
ELEKTROTEHNIČKI FAKULTET

Miloš S. Pavlović

**PRAĆENJE POKRETNIH OBJEKATA U  
SEKVENCI KRATKOTALASNIH  
INFRACRVENIH SLIKA KORELACIONIM  
METODAMA I TEHNIKAMA ROBUSNE  
KALMANOVE FILTRACIJE**

doktorska disertacija

Beograd, 2024.

UNIVERSITY OF BELGRADE  
SCHOOL OF ELECTRICAL ENGINEERING

Miloš S. Pavlović

**MOVING OBJECT TRACKING IN  
SHORT-WAVE INFRARED VIDEO SEQUENCE  
BASED ON CORRELATION FILTERS AND  
ROBUST KALMAN FILTERING TECHNIQUES**

Doctoral Dissertation

Belgrade, 2024.



**Mentor:**

dr Branko Kovačević, profesor emeritus  
Univerzitet u Beogradu, Elektrotehnički fakultet

**Članovi komisije:**

dr Željko Đurović, redovni profesor  
Univerzitet u Beogradu, Elektrotehnički fakultet

dr Aleksandra Krstić, vanredni profesor  
Univerzitet u Beogradu, Elektrotehnički fakultet

dr Zoran Banjac, viši naučni saradnik (komentor)  
Institut Vlatacom

dr Jelena Čertić, vanredni profesor  
Univerzitet u Beogradu, Elektrotehnički fakultet

dr Ilija Popadić, viši naučni saradnik  
Institut Vlatacom

Datum odbrane

---

# Zahvalnica

Zahvaljujem se svom mentoru prof. dr Branku Kovačeviću na pomoći pri izboru teme doktorske disertacije, za pruženu slobodu tokom istraživačkog rada, kao i za konstantnu podršku tokom rada na disertaciji.

Na velikom trudu, strpljenju i važnim savetima tokom sprovedenog istraživanja i pisanja disertacije zahvaljujem se svom komentoru iz Instituta Vlatacom, dr Zoranu Banjcu.

Veliku zahvalnost dugujem rukovodstvu Instituta Vlatacom na podršci svake vrste tokom izrade disertacije. Za inspirativno radno okruženje, konstruktivne diskusije i bodrenje tokom istraživanja zahvaljujem se koleginicama Jelici Radomirović i Tamari Parojčić i kolegi Lazaru Lazareviću, koji je značajno pomogao i u kreiranju skupa sekvenci kratkotalasnih infracrvenih slika za eksperimentalni rad. Na stručnoj pomoći i korisnim sugestijama zahvaljujem se dr Dragani Perić.

Na velikoj motivaciji i nadahnuću, kao i pomoći da disertacija dobije svoj konačni oblik zahvaljujem se Marini Tomašević.

Na kraju, zahvaljujem se svojim roditeljima Zorici i Saši na bezrezervnoj podršci, požrtvovanju i ohrabrenju tokom celokupnog perioda školovanja.

**Naslov teze:** Praćenje pokretnih objekata u sekvenci kratkotalasnih infracrvenih slika korelacionim metodama i tehnikama robusne Kalmanove filtracije

**Rezime:**

Primena kamera koje rade u kratkotalasnom infracrvenom spektralnom opsegu u video sistemima može obezbediti komplementarne informacije u izazovnim uslovima propagacije u kojima su termalne i kamere koje rade u vidljivom opsegu neefikasne, a jedna od najvažnijih komponenti svih inteligentnih video sistema jeste sposobnost praćenja pokretnih objekata.

U disertaciji je analiziran problem praćenja jednog pokretnog objekta u sekvenci kratkotalasnih infracrvenih slika. Cilj je postići dugotrajno praćenje pokretnog objekta u realnom vremenu u različitim izazovnim situacijama na sceni. U fokusu disertacije su metode za praćenje zasnovane na korelacionim filtrima, prvenstveno zbog svojih dobrih performansi u regularnim uslovima scene, ali i njihove računске efikasnosti. Praćenje pokretnih objekata korelacionim filtrima se zasniva na detekciji vršne vrednosti u korelacionom odzivu dobijenog korelacijom između uspostavljenog modela korelacionog filtra i prostora pretrage na svakoj slici sekvence. Ako se tokom praćenja između sukcesivnih slika sekvence objekat pomeri u odnosu na poziciju na prethodnoj slici sekvence, generisani korelacioni odziv treba da ima isti iznos pomeraja lokacije vršne vrednosti. Međutim, korelacioni filtri nisu dovoljno robusni da se nose sa izazovima poput promene veličine i orijentacije objekta, promene u osvetljenju scene, a poseban izazov je prisustvo različitih tipova okluzija. Dodatni izazov predstavlja nedostatak informacija o boji i generalno nizak kontrast kratkotalasne infracrvene slike.

U okviru disertacije je formulisan nov sistem za praćenje pokretnih objekata koji primenom Kalmanovog filtra kao estimatora stanja objekta, modula za detekciju okluzija i neregularnosti u praćenju, modula za estimaciju veličine objekta, a zatim adaptivnim ažuriranjem kontekstno regularizovanog korelacionog filtra postiže značajno poboljšanje performansi praćenja u navedenim izazovnim situacijama u odnosu na sistem oslonjen samo na korelacioni filter. Posebno, za prevazilaženje okluzija prepoznatih kao najizazovniji problem u video praćenju, u disertaciji se predlaže primena Kalmanovog filtra kao prediktora stanja objekta i metode za proširenje oblasti pretrage. Proširivanje oblasti pretrage pomaže u ponovnoj identifikaciji objekta nakon okluzije, ali i dovodi do povremene pojave grešaka u podacima merenja koje mogu dovesti do gubitka objekta i prestanka praćenja. Ove greške se mogu tretirati kao loši podaci ili autlajeri. Pojava autlajera u mernim podacima za posledicu ima loše performanse optimalnog Kalmanovog filtra, obzirom da time nije zadovoljena Gausova raspodela šuma merenja što je jedna od osnovnih pretpostavki na kojima se optimalna tehnika estimacije zasniva.

U disertaciji je predstavljen dizajn tehnike robusne Kalmanove filtracije, otporne na prisustvo autlajera koji kontaminiraju Gausovu raspodelu merenja. Prvo, klasa M-robustifikovanih algoritama dinamičke stohastičke aproksimacije je izvedena minimiziranjem u svakom koraku vremenski promenljivog M-robustnog indeksa performansi, koji je opšti za porodicu razmatranih algoritama. Matrica pojačanja estimatora se izračunava u svakom koraku minimizacijom dodatnog kriterijuma aproksimativne minimalne varijanse, koristeći metod statističke linearizacije. Kombinovanjem predloženog M-robustnog estimatora sa jednokoračnom optimalnom predikcijom, u smislu minimalne srednje kvadratne greške, izvedena je nova statistički linearizovana M-robustifikovana tehnika Kalmanove filtracije. Dve jednostavne praktične verzije predloženog M-robustifikovanog estimatora

stanja su izvedene aproksimacijom optimalnog srednje kvadratnog koeficijenta statističke linearizacije sa fiksnim i vremenski promenljivim faktorima.

Međutim, povećanje robusnosti može dovesti do smanjenja efikasnosti estimatora. Kako bi se napravio balans između efikasnosti estimatora i robusnosti u odnosu na autlajere, u disertaciji je predložen nov pristup gde se funkcija uticaja robusnog estimatora prilagođava u zavisnosti od situacije na sceni, pri čemu je uspostavljena relacija između praga saturacije Huberove funkcije uticaja i vršne vrednosti korelacionog odziva evaluiranog uzorka slike korelacionim filtrom. Dodatno unapređenje sistema za praćenje je postignuto i adaptacijom statistika šuma merenja u robusnom Kalmanovom filtru. Na ovaj način je obezbeđena sveobuhvatna adaptacija robusnog sistema za praćenje pokretnih objekata uslovima na sceni, kako u regularnim, tako i u scenarijima koji se smatraju veoma izazovnim za praćenje pokretnih objekata u sekvenci slika.

**Ključne reči:** praćenje pokretnih objekata, kratkotalasna infracrvena slika, korelacioni filtri, Kalmanov filter, robusna estimacija, adaptivne robusne metode.

**Naučna oblast:** Elektrotehnika i računarstvo

**Uža naučna oblast:** Upravljanje sistemima i obrada signala

**UDK broj:** 621.3

**Dissertation title:** Moving object tracking in short-wave infrared video sequence based on correlation filters and robust Kalman filtering techniques

**Abstract:**

The application of cameras operating in the short-wave infrared spectral range in video systems can provide complementary information in challenging propagation conditions where visible-light and thermal cameras are ineffective. Additionally, one of the most important components of intelligent video systems is the ability to track moving objects.

The main research topic of this dissertation is the tracking of a single moving object in short-wave infrared video sequence. The goal is to achieve long-term tracking of a moving object in real-time in various challenging situations on the scene. The focus is on the tracking methods based on correlation filters, primarily due to their good performance under regular scene conditions, but also due to their computational efficiency. The tracking of moving object using correlation filter is based on the detection of a peak in the correlation response obtained by correlation between the established correlation filter model and the search area in each frame. If the object is moved relative to the position in the previous frame of the video sequence, the generated correlation response should have the same displacement of the peak location. However, correlation filters are not robust enough to deal with challenges such as changes in object size and orientation, changes in scene illumination conditions, and a particular challenge is the presence of different types of occlusions. An additional challenge is the lack of color information and the low contrast of the short-wave infrared image.

Dissertation formulates a system for moving object tracking that applies the Kalman filter as an estimator of the object state, a module for occlusion and failure detection, a module for object size estimation, as well as adaptive updating the context-aware correlation filter and achieves a significant improvement in tracking performance in the specified challenging situations compared to a system based only on a correlation filter. In particular, to overcome occlusions, recognized as the most challenging problem in video tracking, a Kalman filter as a predictor and an extended search area were used, so that tracking can be continued even in the case of full occlusion. Although using an extended search area helps in better re-detection of the object after occlusion, it may also lead to occasional errors in the measurement data that can result in object loss. These errors can be treated as bad data or outliers. The appearance of outliers in the measurement data may cause significant deviations from the theoretically optimal performance of the Kalman filter, since the distribution of the measurement data which contain outliers does not satisfy the Gaussian distribution, which is one of the basic assumptions on which the optimal Kalman filter is based.

In this dissertation, the design of a robust Kalman filtering technique that is insensitive to outliers contaminating Gaussian distributed measurements, is presented. First, a class of M-robustified dynamic stochastic approximation algorithms is derived by minimizing at each step the time-varying M-robust performance index, which is general for the family of algorithms to be considered. The gain matrix of a particular estimator is calculated at each step by minimizing an additional criterion of the approximate minimum variance type, using the statistical linearization method. By combining the proposed M-robust estimator with one-step optimal prediction, in terms of minimum mean-square error, a new statistically linearized M-robustified Kalman filtering technique has been derived. Two simple practical versions of the proposed M-robustified state



estimator are derived by approximating the mean-square optimal coefficient of statistical linearization with fixed and time-varying factors.

However, strong robustness to outliers may decrease the estimator's efficiency under regular conditions. To make a balance between desired estimator efficiency and robustness to outliers, a new adaptive approach was proposed in the dissertation. Influence function of the robust estimator is adjusted depending on the situation on the scene, whereby a relationship between the saturation threshold of Huber's influence function and the peak of the correlation response map has been established. Additional improvement of the tracking system was achieved by adapting the noise statistics in the robust Kalman filter. In this way, a comprehensive adaptation of a robust moving object tracking system to scene conditions is provided, both in regular and challenging video tracking scenarios.

**Keywords:** moving object tracking, short-wave infrared imagery, correlation filter, Kalman filter, robust estimation, adaptive robust methods.

**Scientific field:** Electrical Engineering and Computer Science

**Scientific subfield:** System Control and Signal Processing

**UDC number:** 621.3

# Sadržaj

<b>1</b>	<b>UVOD</b>	<b>1</b>
1.1	ORGANIZACIJA DISERTACIJE	4
<b>2</b>	<b>KRATKOTALASNA INFRACRVENA SLIKA</b>	<b>7</b>
2.1	KARAKTERISTIKE SWIR KAMERA, PREDNOSTI I OGRANIČENJA	9
2.2	PRIMENA SWIR KAMERA ZA DETEKCIJU I PRAĆENJE OBJEKATA	13
2.3	SKUP SWIR VIDEO SEKVENCI ZA PRAĆENJE POKRETNIH OBJEKATA	15
<b>3</b>	<b>PRAĆENJE POKRETNIH OBJEKATA U SEKVENCI SLIKA</b>	<b>22</b>
3.1	KOMPONENTE ALGORITMA ZA PRAĆENJE OBJEKATA U SEKVENCI SLIKA	25
3.2	IZAZOVI U PRAĆENJU OBJEKATA U SEKVENCI SLIKA	26
3.2.1	Klater	27
3.2.2	Šum	28
3.2.3	Promena položaja i orijentacije	28
3.2.4	Promena veličine objekata	29
3.2.5	Promene u uslovima scene	29
3.2.6	Okluzije	30
3.3	METODE INICIJALIZACIJE PRAĆENJA OBJEKATA	31
3.3.1	Manuelna inicijalizacija	31
3.3.2	Automatska inicijalizacija	32
3.3.2.1	Detekcija pokreta	33
3.3.2.2	Detekcija objekata na SWIR slici	35
3.4	KLASIFIKACIJA PRISTUPA ZA PRAĆENJE OBJEKATA U SEKVENCI SLIKA	40
<b>4</b>	<b>KORELACIONE METODE ZA PRAĆENJE POKRETNIH OBJEKATA</b>	<b>44</b>
4.1	FUNDAMENTALNI KORELACIONI FILTRI	45
4.1.1	Sintetičke diskriminacione funkcije	45
4.1.2	Sintetičke diskriminacione funkcije minimalne varijanse	47
4.1.3	Filtar minimalne prosečne korelacione energije	48
4.1.4	Sintetičke diskriminacione funkcije optimalnog kompromisa	49
4.1.5	Sintetičke diskriminacione funkcije minimalne srednje kvadratne greške	50
4.2	SAVREMENI KORELACIONI FILTRI ZA PRAĆENJE POKRETNIH OBJEKATA	52
4.2.1	Filtar minimalne sume kvadrata grešaka	52
4.2.2	Kernelizovani korelacioni filtri	54
4.2.2.1	Ciklična pomeranja i cirkularne matrice	55
4.2.2.2	Pojednostavljenje linearne regresije za cirkularne matrice	56
4.2.2.3	Kernelizacija korelacionih filtara	57
4.2.2.4	Višestruki kanali i ažuriranje filtra	60
4.2.3	Opšta struktura algoritma za praćenje korelacionim filtrima	61
4.2.4	Tehnike za savladavanje izazova u praćenju pokretnih objekata	62
4.2.4.1	Estimacija veličine objekta	62
4.2.4.2	Detekcija okluzija	63
4.2.4.3	Regularizacija i ažuriranje korelacionih filtara	65

4.2.4.4	Prevažilaženje okluzija i ponovna identifikacija objekta .....	66
4.2.4.5	Fuzija prediktora i korelacionih filtara .....	67
4.3	DIZAJN SISTEMA ZA PRAĆENJE POKRETNIH OBJEKATA ZASNOVANOG NA KORELACIONIM FILTRIMA .....	68
4.3.1	Statistička analiza rezultata rada predloženog sistema za praćenje .....	76
<b>5</b>	<b>ROBUSNA ESTIMACIJA STANJA STOHAŠTIČKIH SISTEMA .....</b>	<b>79</b>
5.1	KALMANOV FILTAR .....	80
5.1.1	Osobine i ograničenja standardnog Kalmanovog filtra .....	81
5.2	ROBUSNA ESTIMACIJA STANJA DINAMIČKIH SISTEMA.....	83
5.3	METOD DINAMIČKE STOHAŠTIČKE APROKSIMACIJE .....	85
5.4	M-ROBUSNI KALMANOV FILTAR .....	87
5.4.1	Statistički linearizovan M-robusni Kalmanov filtar .....	92
5.4.2	Optimalni koeficijent statističke linearizacije.....	94
5.4.3	Funkcije uticaja.....	96
5.5	ADAPTIVNI M-ROBUSNI KALMANOV FILTAR .....	99
5.5.1	Adaptivna M-robusna estimacija statistika šuma merenja .....	100
5.5.2	Adaptivna M-robusna estimacija statistika šuma stanja .....	101
<b>6</b>	<b>ROBUSNI SISTEM ZA PRAĆENJE POKRETNIH OBJEKATA U SEKVENCI</b>	
	<b>KRATKOTALASNIH INFRACRVENIH SLIKA.....</b>	<b>103</b>
6.1	SISTEM ZA PRAĆENJE KORELACIONIM I STANDARDNIM KALMANOVIM FILTROM.....	105
6.2	ROBUSNI KALMANOV FILTAR.....	108
6.2.1	Analiza statistički linearizovanog M-robusnog Kalmanovog filtra.....	108
6.2.2	Sistem za praćenje korelacionim i M-robusnim Kalmanovim filtrom .....	112
6.2.3	Analiza funkcija uticaja.....	114
6.3	SISTEM ZA PRAĆENJE KORELACIONIM FILTROM I M-ROBUSNIM KALMANOVIM FILTROM SA ADAPTIVNOM FUNKCIJOM UTICAJA .....	117
6.4	SISTEM ZA PRAĆENJE KORELACIONIM I ADAPTIVNIM M-ROBUSNIM KALMANOVIM FILTROM .....	121
6.4.1	Komparativna analiza predloženog sistema i savremenih algoritama za praćenje pokretnih objekata u sekvenci SWIR slika .....	124
<b>7</b>	<b>ZAKLJUČAK.....</b>	<b>129</b>
	<b>LITERATURA .....</b>	<b>132</b>

# Lista skraćenica

AP	<i>Average Precision</i>
APCE	<i>Average Peak Correlation Energy</i>
CCD	<i>Charge Coupled Device</i>
CEE	<i>Cumulative Estimation Error</i>
CMOS	<i>Complementary Metal Oxide Semiconductor</i>
EKF	<i>Extended Kalman Filter</i>
FFT	<i>Fast Furier Transform</i>
FGV	<i>Funkcija Gustine Verovatnoće</i>
FIR	<i>Far Infrared</i>
FPS	<i>Frames Per Second</i>
HOG	<i>Histogram Orijentisanih Gradjenata</i>
IFFT	<i>Inverse Fast Furier Transform</i>
InGaAs	<i>Indijum Galijum Arsenid</i>
IOU	<i>Intersection Over Union</i>
KCF	<i>Kernelized Correlation Filter</i>
LWIR	<i>Long-Wave Infrared</i>
MACE	<i>Minimum Average Correlation Energy</i>
MAD	<i>Median Absolute Deviation</i>
ML	<i>Maximum Likelihood</i>
MODTRAN	<i>MODerate resolution atmospheric TRANsmission</i>
MOG	<i>Mixture Of Gaussians</i>
MOSSE	<i>Minimum Output Sum of Squared Error</i>
MWIR	<i>Mid-Wave Infrared</i>
NIR	<i>Near Infrared</i>
OPE	<i>One Pass Evaluation</i>
PSR	<i>Peak to Sidelobe Ratio</i>
PTZ	<i>Pan-Tilt-Zoom</i>
QWIP	<i>Quantum Well Infrared Photodetector</i>
SDF	<i>Synthetic Discriminant Function</i>
SVM	<i>Support Vector Machine</i>
SWIR	<i>Short-Wave Infrared</i>
TIR	<i>Thermal Infrared</i>

# 1 Uvod

Praćenje pokretnih objekata u sekvenci slika predstavlja proces estimacije trajektorije kretanja objekata na osnovu datog inicijalnog stanja objekta na prvoj slici sekvence [1]. Jedan je od fundamentalnih problema u kompjuterskoj viziji. Sposobnost da se precizno prate objekti u dinamičkim i nepredvidivim okruženjima predstavlja osnovu za razvoj inteligentnih sistema koji mogu interagovati sa svojim okruženjem na efikasan način. Spektar primena video praćenja objekata je širok i uključuje oblasti kao što su inteligentni video nadzor, praćenje saobraćaja, autonomna vožnja, robotika, prepoznavanje aktivnosti, sportska analitika, medicinska dijagnostika, interakcija čovek – računar, proširena realnost, kao i mnoge druge. Nezavisno od primene, primarni ciljevi razvoja svakog algoritma za praćenje jesu ostvarivanje visoke preciznosti praćenja u različitim okolnostima na sceni, mogućnost dugotrajnog praćenja, kao i postizanje brzine procesiranja dovoljne za rad u realnom vremenu. Početna komponenta u svakom sistemu za praćenje jeste senzor, koji se koristi za snimanje scene.

Najčešće korišćene kamere su one koje rade u vidljivom opsegu, odnosno kolor kamere. Stoga je i razvoj algoritama za video praćenje dominantno usmeren na praćenje objekata u sekvencama snimljenim kamerama koje rade u vidljivom opsegu [2] – [7]. Za uspešno praćenje objekata u uslovima slabog osvetljenja ili potpunog mraka, u sistemima za video praćenje se često primenjuju i termalne kamere. Iz tog razloga, razvijeni su i mnogi algoritmi za praćenje objekata u sekvencama termalnih slika [8] – [11]. Međutim, u uslovima snimanja scene na kojoj ima prisustva dima, izmaglice ili magle, isparenja, kao i jake kiše, kamere koje rade u vidljivom opsegu imaju tendenciju da generišu slike limitirane u detaljima i informativnim obeležjima, te na taj način direktno utiču na performanse celokupnog sistema za praćenje objekata. U ovim izazovnim uslovima scene, mogućnost efikasnog snimanja može se obezbediti korišćenjem kamere koje rade u kratkotalasnom infracrvenom opsegu (eng. *Short-Wave InfraRed* - SWIR) i obezbeđuju slike bogatije detaljima [12], [13]. Iako se svetlost u SWIR opsegu reflektuje od objekata na sličan način kao i u vidljivom opsegu, vidljivost je u navedenim izazovnim uslovima sredine poboljšana usled dužih talasnih dužina u SWIR opsegu, koje su manje pod uticajem efekata rasejanja [12]. S druge strane, kamere koje rade u termalnom infracrvenom opsegu mogu da obezbede snimanje u uslovima izmaglice, magle ili dima, ali u opštem slučaju mogu samo da detektuju prisustvo toplijeg objekta na hladnijoj pozadini. Primenom SWIR kamera može se postići identifikacija objekta i pružiti više informacija o objektu u navedenim izazovnim vremenskim uslovima. Pored toga, SWIR kamere imaju i mogućnost snimanja slika kroz staklo. Zbog svojih prednosti koje se odnose na smanjene efekte rasejanja, SWIR kamere sve više nalaze primenu u različitim sistemima video nadzora i analitike.

Fokus razvoja algoritama kompjuterske vizije za primenu u kratkotalasnom infracrvenom opsegu je dominantno usmeren ka modelima mašinskog i dubokog učenja za detekciju objekata na SWIR slikama [14] – [17]. U oblasti praćenja objekata u sekvencama SWIR slika u literaturi se uglavnom razmatra praćenje više objekata na osnovu detekcija, gde su detekcije generisane modelima

mašinskog ili dubokog učenja [18] – [22]. Nasuprot tome, predmet istraživanja ove doktorske disertacije je praćenje jednog pokretnog objekta u sekvenci SWIR slika snimljenih jednom kamerom.

Cilj disertacije je postizanje dugotrajnog praćenja u izazovnim scenarijima nezavisno od tipa objekta, uz mogućnost rada u realnom vremenu. Praćenje u pristupima oslonjenim na detekcije dobijene modelima neuralnih mreža podrazumeva da su klase objekata koji se mogu pratiti unapred definisane tokom obučavanja modela za detekciju objekata. Za svaki novi tip objekata koji se želi pratiti, mora se prvo obučiti model za detekciju, što predstavlja otežavajuću okolnost za ispunjenje cilja da se kreira sistem za praćenje nezavisan od tipa objekta. Pristupi za praćenje jednog objekta nezavisni od tipa objekta koji se prati obuhvataju tradicionalne algoritme i algoritme zasnovane na dubokom učenju. Metode za praćenje objekata zasnovane na dubokom učenju postižu veoma dobre performanse, ali sa potrebom za velikim skupovima podataka za obučavanje. Za rad u realnom vremenu, ove metode zahtevaju procesiranje na grafičkim procesorskim jedinicama, a pored toga imaju i nepoznato ponašanje u scenarijima za koje nisu obučavane. S druge strane, tradicionalne metode za praćenje pokretnih objekata, koje obuhvataju generativne i diskriminativne algoritme, mogu postići brzine izvršavanja dovoljne za rad u realnom vremenu i bez procesiranja na grafičkim procesorskim jedinicama. Kod generativnih algoritama praćenje se vrši pretragom za regionom slike koji najbolje odgovara referentnom modelu objekta, koristeći samo informacije o objektu. Kod diskriminativnih algoritama se pored informacija o objektu koriste i informacije o neposrednoj pozadini objekta. Performanse generativnih algoritama ograničene su dimenzijama prostora za reprezentaciju modela objekta, a u kompleksnijim scenarijima pokazuju manju sposobnost diskriminacije objekta od okoline. Zahvaljujući uvršćenju i pozadinskih informacija u model objekta, diskriminativni algoritmi imaju bolju sposobnost da obuhvate širi spektar promena u izgledu objekta. Među diskriminativnim algoritmima za praćenje, sa mogućnošću rada u realnom vremenu, posebno se izdvajaju diskriminativni korelacioni filtri [6], [7]. Analizom prirode SWIR slike, kao i osobina diskriminativnih algoritama, može se zaključiti da pristupi zasnovani na korelacionim filtrima pružaju dobru osnovu za razvoj efikasnog sistema za praćenje pokretnih objekata u sekvenci SWIR slika.

U regularnim uslovima, korelacioni filtri pokazuju veoma dobre performanse u praćenju objekata. Međutim, korelacioni filtri generalno nisu dovoljno robusni da se nose sa promenama veličine ili orijentacije objekta, promenama u osvetljenju scene, a poseban izazov predstavlja prisustvo različitih tipova okluzija.

Polazeći od ograničenja korelacionih filtara, u ovoj disertaciji se najpre predlaže dizajn sistema za praćenje sa modulima za estimaciju veličine objekta, detekciju okluzija, ponovnu identifikaciju objekta nakon perioda pod okluzijom, kao i adaptivno ažuriranje modela korelacionog filtra. Budući da su okluzije identifikovane kao najizazovniji problem, a posebno manevar praćenog objekta pod okluzijom, u sistem za praćenje je dalje implementiran Kalmanov filter kao prediktor stanja objekta i modul za pretragu objekta u proširenoj oblasti. Ovim bi se omogućilo prevazilaženje okluzija i ponovna detekcija objekta nakon perioda okluzije. Proširivanje oblasti za pretragu objekta pomaže u efikasnijoj ponovnoj detekciji praćenog objekta nakon okluzije, ali takođe dovodi do pojave grešaka u podacima estimacije pozicije objekta, koje mogu dovesti do gubitka objekta i prestanka praćenja. Greške u estimaciji pozicije objekta velikog intenziteta, koje dovode do prekida praćenja, u kontekstu estimacionih procedura mogu se tretirati kao loši podaci ili autlajeri (eng. *outliers*) [23]. Iz tog razloga, disertacija ima za cilj da istraži karakteristike autlajera prilikom praćenja objekata u sekvenci SWIR slika sistemom zasnovanim na korelacionim filtrima. Međutim, u javno dostupnim skupovima sekvenci za razvoj i evaluaciju algoritama za praćenje, kako u vidljivom opsegu [2] – [7],

[24] – [29], tako i u infracrvenom [8] – [11], nedostaju sekvence snimljene SWIR kamerom. Stoga je za potrebe istraživanja ove disertacije kreiran novi reprezentativni skup SWIR video sekvenci. Pored kretanja objekta u regularnim uslovima, skup obuhvata i sekvence sa izazovnim scenarijima za praćenje. Statistička analiza grešaka dizajniranog sistema koji obuhvata korelacioni filter sa proširenom zonom pretrage i Kalmanov filter, ukazuje na uobičajene greške koje degradiraju performanse praćenja, kao i na greške koje dovode do potpunog gubitka objekta.

U slučaju kada u mernim podacima postoje greške koje se po svojoj prirodi mogu tretirati kao autlajeri, standardni Kalmanov filter, kao prediktor u dizajniranom sistemu za praćenje, ne predstavlja optimalno rešenje. Da bi se dobile optimalne performanse Kalmanovog filtra, neophodno je obezbediti tačan apriori opis dinamike stanja sistema i statistike slučajnih merenja. U tom smislu, ako dinamika stanja sistema i pridružena merenja imaju izražene nelinearnosti, koje se ne mogu pravilno opisati linearizacijom kao i ako osnovne stohastičke sekvence nemaju Gausovsku (normalnu) raspodelu, performanse Kalmanovog filtra mogu biti značajno degradirane [30], [31]. Dakle, postoji interes za klasu estimacionih procedura koje ne moraju biti optimalne, prema određenoj statističkoj meri učinka, ali imaju ograničenu ukupnu grešku estimacije u praktičnim primenama.

Porodica procedura dinamičke stohastičke aproksimacije (eng. *dynamic stochastic approximation*) nudi prihvatljiv razuman izbor, obezbeđujući dobre rezultate koji su potvrđeni u različitim aplikacijama [32] – [36]. U tom smislu, svaki Kalmanov filter sa pogrešnom sekvencom pojačanja, usled odstupanja od teorijski optimalnih uslova u praksi, može se posmatrati kao algoritam dinamičke stohastičke aproksimacije. Konkretno, optimalni Kalmanov filter je osetljiv na prisustvo autlajera u mernoj sekvenci usled linearne zavisnosti estimacije stanja objekta od merenja, odnosno nije robusan na prisustvo autlajera u mernoj sekvenci. Stoga, postoji dodatni praktični interes za dizajniranje klase robusnih tehnika filtriranja, koje efikasno rade u prisustvu autlajera. U statističkoj literaturi je za potiskivanje dejstva autlajera predloženo više robusnih procedura [23], [37] – [39]. U inženjerskoj praksi se često primenjuje Huberov M-robusni estimator, jer aproksimira optimalni estimator maksimalne verodostojnosti (eng. *maximum likelihood* – ML), što ga čini prirodnim izborom koji se lako implementira u praksi [40]. U tom smislu, u literaturi su predložene različite kombinacije M-robusnog estimatora i optimalnog Kalmanovog filtra [41] – [45].

Uopšteno govoreći, svaka procedura estimacije je kombinacija kriterijuma koji treba minimizirati, modela varijabli stanja koje se estimiraju i algoritma za estimaciju nepoznatih stanja na osnovu merenja. U tom smislu, robusni estimatori u literaturi mogu se klasifikovati u dve grupe. Prva grupa je porodica nerekurzivnih (eng. *offline*) robusnih procedura sa paketnom obradom mernih podataka. Robusnost ovih estimatora postiže se po cenu povećanja računarskih zahteva, te nisu pogodni za primenu u realnom vremenu. Zbog toga, za rad u realnom vremenu, rekurzivne robusne procedure imaju prednost u odnosu na nerekurzivne. U tom smislu, drugu grupu robusnih metoda predstavlja porodica estimatora koji rekurzivno izračunavaju estimacije stanja sistema na osnovu raspoloživih merenja. U novijoj literaturi je predložen rekurzivni robusni estimator, koji predstavlja kombinaciju Huberovog M-robusnog estimatora sa specifičnim kvazi-linearnim oblikom dinamičke stohastičke aproksimacije, kako bi se redizajnirao korak korekcije u optimalnom Kalmanovom filtru [44], [45]. Pored toga, mnogi suboptimalni nelinearni estimatori stanja sistema su dizajnirani primenom Tejlorovog reda za opisivanje dinamike stanja nelinearnog sistema [46] – [48]. Druga često korišćena metoda je statistička aproksimacija, koja generalno daje bolju aproksimaciju nelinearnosti u estimacionoj proceduri od metode Tejlorovog reda [46] – [48]. Najjednostavniji oblik takve metode je poznat kao statistička linearizacija. Metod statističke linearizacije ima potencijalnu prednost za

projektovanje suboptimalnog nelinearnog filtra (estimatora stanja sistema) [46] – [48]. Stoga je predložen novi estimator stanja dinamičkog sistema koji je zasnovan na kombinaciji Huberovog M-robustnog estimatora i algoritma dinamičke stohastičke aproksimacije, uz primenu metode statističke linearizacije za aproksimaciju nelinearne robusne kriterijumske funkcije u M-robustnom estimatoru [49] – [51]. U tom smislu, Huberov M-robustni koncept se koristi za dizajniranje porodice M-robustnih procedura dinamičke stohastičke aproksimacije, minimiziranjem u svakom koraku vremenski promenljivog M-robustnog indeksa performansi, zasnovanog na Huberovoj M-robustnoj nelinearnoj kriterijumskoj funkciji. Da bi se postigla brza konvergencija, matrica pojačanja robusnog nelinearnog estimatora stanja sistema se izvodi u svakom koraku minimiziranjem aproksimativnog kriterijuma minimalne varijanse greške, po analogiji sa optimalnim Kalmanovim filtrom. Postavljeni problem nelinearne optimizacije je aproksimativno rešen primenom metode statističke linearizacije. Štaviše, aproksimacijom koeficijenta statističke linearizacije, optimalnog u srednje kvadratnom smislu, sa usrednjenim nagibom Huberove M-robustne funkcije uticaja, koja predstavlja prvi izvod osnovne robusne kriterijumske funkcije, izvedena je nova statistički linearizovana M-robustifikovana procedura dinamičke stohastičke aproksimacije. Dodatno, aproksimacijom usrednjenog nagiba Huberove funkcije uticaja trenutnim uzorkom, dobijena je adaptivna verzija predloženog robusnog rekursivnog estimatora stanja sistema. Polazeći od optimalne strukture Kalmanovog filtra, u kojoj su koraci predikcije i korekcije nezavisni, izvedeni robusni rekursivni estimator stanja sistema se koristi za redizajniranje koraka korekcije, dok korak predikcije stanja ostaje neizmenjen po analogiji sa optimalnim Kalmanovim filtrom. Na taj način, korak predikcije je jednokoračni optimalni prediktor stanja sistema u srednje kvadratnom smislu.

Implementacijom predložene M-robustifikovane verzije Kalmanovog filtra u sistem za praćenje korelacionim filtrom, u disertaciji se dalje predlaže nov pristup za praćenje objekata u sekvenci SWIR slika. Takođe, nova adaptivna verzija M-robustifikovanog Kalmanovog filtra je dizajnirana koristeći informacije iz korelacionog odziva za prilagođavanje praga saturacije Huberove M-robustne funkcije uticaja. Na taj način, robusni algoritam za praćenje postiže prihvatljiv kompromis između efikasnosti praćenja manevra objekta zasnovanog na estimaciji stanja sistema i robusnosti sistema za praćenje u smislu eliminacije uticaja autlajera na kvalitet praćenja.

## 1.1 Organizacija disertacije

Disertacija, pored uvoda, obuhvata još šest međusobno povezanih poglavlja u kojima su kroz postupnu analizu detaljno prikazani predmet istraživanja disertacije, kao i ostvareni rezultati.

U drugom poglavlju se nakon definicije elektromagnetnog spektra analizira priroda senzora koji rade u kratkotalasnom infracrvenom opsegu. Potom je izvršena karakterizacija slike sa SWIR kamere i analizirane su osobine SWIR kamera u razmatranim izazovnim uslovima sredine. Uz eksperimentalne testove u istim uslovima scene izvršena je komparativna analiza slika sa SWIR, termalne i kamere koja radi u vidljivom opsegu. Nakon potvrde efikasnosti SWIR kamera u karakterističnim uslovima scene, prikazan je detaljni razvoj algoritama kompjuterske vizije za detekciju i praćenje objekata na SWIR slici. Analizom javno dostupnih skupova sekvenci za razvoj i evaluaciju algoritama za praćenje, utvrđeno je da ne postoji skup reprezentativnih sekvenci koje su snimljene SWIR kamerom. Stoga, u poslednjem segmentu drugog poglavlja se predlaže nov



reprezentativni skup SWIR video sekvenci za razvoj i testiranje metoda za praćenje pokretnih objekata u SWIR opsegu.

Treće poglavlje definiše problem praćenja pokretnih objekata u sekvenci slika. Prikazane su osnovne komponente algoritama za praćenje, a zatim detaljno analizirani izazovi koji se javljaju u praćenju pokretnih objekata u sekvenci slika. Potom su razmatrane metode inicijalizacije praćenja objekata i predložene dve nove metode za automatsku inicijalizaciju praćenja. Prva se zasniva na detekciji pokreta u sekvenci slika, dok se druga oslanja na modele dubokog učenja za detekciju objekata definisane klase. Za obučavanje modela dubokog učenja za detekciju objekata na SWIR slici, u ovom poglavlju se predlaže novi automatizovani metod kros-spektralne anotacije skupova SWIR slika. Na kraju, dat je detaljni prikaz razvoja algoritama za praćenje pokretnih objekata u sekvenci slika.

Četvrto poglavlje se bavi korelacionim metodama za praćenje pokretnih objekata. Prikazan je razvoj od fundamentalnih do savremenih korelacionih filtara, koji se koriste u algoritmima za praćenje pokretnih objekata. U skladu sa izazovima u praćenju prikazanim u trećem poglavlju, izvršena je analiza tehnika adaptacije korelacionih filtara koje imaju za cilj prevazilaženje navedenih izazova. Analiza obuhvata sledeće aspekte: estimaciju veličine objekta, detekciju okluzija, regularizaciju i ažuriranje korelacionih filtara, prevazilaženje okluzija i fuziju prediktora sa korelacionim filtrima. Zatim je dat predlog dizajna sistema za praćenje pokretnih objekata zasnovanog na kontekstno regularizovanom korelacionom filtru i Kalmanovom filtru. Za tako dizajniran kombinovani sistem izvršena je statistička analiza rezultata rada i ustanovljeni su osnovni uzroci grešaka, posebno u estimaciji pozicije objekta. Kao posebno izazovan tip grešaka izdvojene su greške veoma velikog intenziteta, označene kao autlajeri.

U cilju kreiranja robusnog sistema za praćenje pokretnih objekata, u petom poglavlju se razmatra robusna estimacija stanja stohastičkih sistema. Najpre je prikazan standardni Kalmanov filtar kao linearni optimalni estimator stanja sistema u smislu minimalno moguće varijanse greške estimacije, a zatim su razmatrane osobine i ograničenja standardnog Kalmanovog filtra. Drugi segment ovog poglavlja se bavi tehnikama robusne estimacije stanja dinamičkih sistema. Zatim, polazeći od metode dinamičke stohastičke aproksimacije izložene u trećem segmentu ovog poglavlja, u četvrtom segmentu poglavlja se predlaže novi statistički linearizovani M-robusni tip Kalmanovog filtra i teorijski se izvodi optimalni koeficijent statističke linearizacije Huberove M-robusne nelinearne funkcije uticaja tipa zasićenja. Takođe, dat je prikaz različitih robusnih funkcija uticaja, koje se mogu primeniti u robustifikaciji Kalmanovog filtra. U poslednjem segmentu poglavlja prikazan je M-robusni Kalmanov filtar sa adaptivnim estimacijama statistika šumova stanja i merenja, paralelno sa robusnom estimacijom stanja sistema.

Šesto poglavlje predlaže robusni sistem za praćenje pokretnih objekata inkorporiranjem robusnog estimatora, predstavljenog u petom poglavlju, u dizajnirani sistem za praćenje zasnovan na korelacionom filtru, iz četvrtog poglavlja. Nakon opisa postupka inicijalizacije sistema, analizirani su rezultati video praćenja sistema koji primenjuje standardni Kalmanov filtar. Nakon potvrde efikasnosti M-robusnog Kalmanovog filtra na simuliranim podacima u drugom segmentu ovog poglavlja, dalje su analizirane performanse praćenja celokupnog sistema primenom robusnog Kalmanovog filtra, a zatim je razmatrana primena različitih funkcija uticaja u robusnom Kalmanovom filtru. U narednom segmentu je predložen nov metod za adaptivno podešavanje parametra Huberove funkcije uticaja na osnovu informacija iz korelacionog odziva. Dalje su prikazane i performanse praćenja robusnog sistema u kome se adaptivno ažuriraju statistike šuma merenja u robusnom

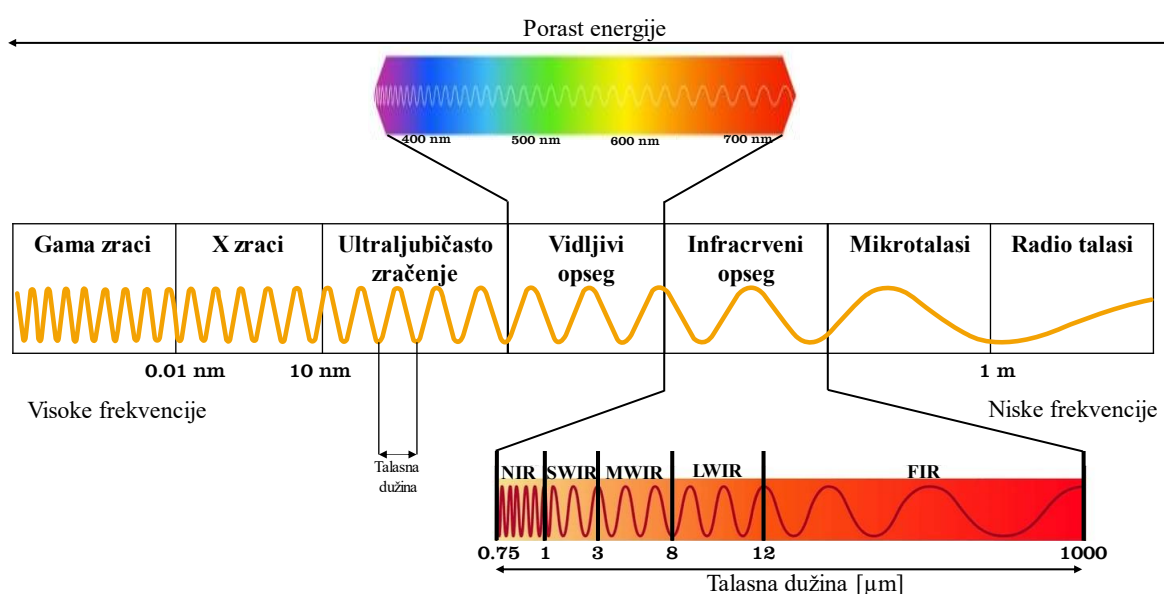
Kalmanovom filtru sa adaptivnom estimacijom parametara funkcije uticaja. Na kraju je izvršena komparativna analiza predloženog adaptivnog robusnog sistema za praćenje i poznatih efikasnih algoritama za praćenje pokretnih objekata iz referentne literature.

Sedmo poglavlje čine zaključak disertacije sa predloženim pravcima daljeg istraživanja.

Na kraju disertacije navedena je referentna literatura, koja obuhvata knjige i naučne radove objavljene u celini u časopisima i zbornicima sa referentnih naučnih skupova, korišćena tokom istraživanja i prilikom izrade ove doktorske disertacije.

## 2 Kratkotalasna infracrvena slika

Odabir pravog senzora za problem praćenja objekata je od velike važnosti. Izbor senzora može imati značajan uticaj na mogućnost efikasnog rada sistema u različitim situacijama i uslovima okruženja. Radi boljeg razumevanja kratkotalasnih infracrvenih senzora, najpre će biti objašnjen pojam elektromagnetnog spektra. Elektromagnetni spektar, ilustrovan na slici 2.1, obuhvata opsege zračenja poznate kao (navedene po redosledu povećanja talasne dužine): gama zraci, rendgenski (X) zraci, ultraljubičasto zračenje, vidljivo svetlo, infracrveno zračenje, mikrotalasi i radio talasi.

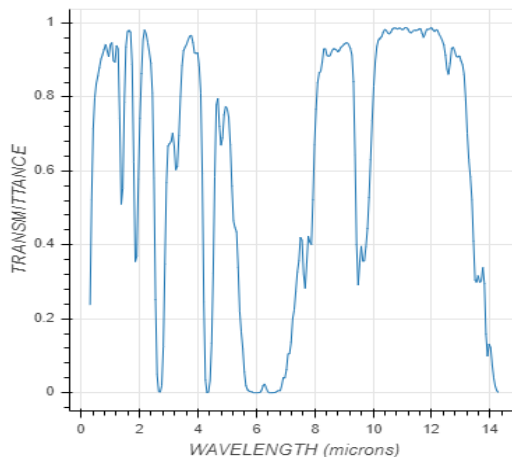


Slika 2.1 Elektromagnetni spektar sa izdvojenim infracrvenim spektrom.

Infracrveno zračenje se definiše kao deo spektra talasnih dužina dužih od onih koje je ljudsko oko u stanju da percipira, približno od oko  $0.7 \mu\text{m}$  što je granica crvene boje u vidljivom opsegu, ali kraćih od onih u mikrotalasnom opsegu, koji nominalno počinje od  $1000 \mu\text{m}$  talasne dužine. Infracrveni opseg se dalje može podeliti na podopsege, bazirane na svojim specifičnim svojstvima: bliski infracrveni opseg (eng. *Near Infrared* - NIR) talasnih dužina  $0.75\text{--}1 \mu\text{m}$ , kratkotalasni infracrveni opseg (eng. *Short-Wave Infrared* - SWIR) talasnih dužina  $1\text{--}3 \mu\text{m}$ , srednjetalasni infracrveni opseg (eng. *Mid-Wave Infrared* - MWIR) talasnih dužina  $3\text{--}8 \mu\text{m}$ , dugotalasni infracrveni opseg (eng. *Long-Wave Infrared* - LWIR) talasnih dužina  $8\text{--}12 \mu\text{m}$ , i daleki infracrveni opseg (eng. *Far Infrared* - FIR), talasnih dužina  $12\text{--}1000 \mu\text{m}$ . U literaturi se često MWIR i LWIR opsezi združeno nazivaju termalni infracrveni opseg (eng. *Thermal Infrared* - TIR).

Podela infracrvenog opsega zračenja najčešće je definisana kombinacijom transmisije u atmosferi i opsezima detekcije dostupnih senzora [12]. Dok se NIR spektar može meriti CMOS (eng. *Complementary Metal Oxide Semiconductor*) ili CCD (eng. *Charge Coupled Device*) tehnologijom, SWIR spektar zahteva drugačiji detektor [52]. Tako se SWIR opseg može definisati kao deo elektromagnetnog spektra od tačke gde reakcija silicijumskog detektora pada pa do početka MWIR

atmosferskog prozora prenosa. Iako je SWIR opseg izvan NIR opsega, SWIR senzori prvenstveno reaguju na reflektovanu elektromagnetnu energiju od objekata, a ne na emitovanu energiju, pa se nominalno ne koriste za merenje toplotne emisivnosti. Većina energije u SWIR opsegu se odbija ili apsorbira od strane objekata, slično svojstvima svetlosti u vidljivom i NIR opsegu. Na slici 2.2 je prikazan grafik transmisivnosti po talasnim dužinama dobijen MODTRAN (eng. *MODerate resolution atmospheric TRANsmission*) atmosferskim modelom prostiranja [53] za srednje geografske širine u letnjim uslovima uz domet senzora do 5 km i zenit od 45°.



**Slika 2.2** Atmosferska propustljivost za različite talasne dužine (MODTRAN model).

U protekle dve decenije, SWIR detektori su postali dostupniji i komercijalne kamere koje rade u SWIR opsegu se mogu naći na tržištu. Materijali na kojima se zasnivaju SWIR detektori su: germanijum (Ge), indijum galijum arsenid (InGaAs), indijum antimonid (InSb), živin kadmijum telurid (HgCdTe) [12], kao i oni koji sadrže strukture super rešetke tipa II [54]. Tehnologija infracrvenih fotodetektora kvantne jame (eng. *Quantum Well Infrared Photodetector - QWIP*) nedavno je postigla značajan napredak u aplikacijama za LWIR slike [54], [55]. Međutim, zbog relativno niske kvantne efikasnosti, primena QWIP tehnologije u SWIR opsegu ne izgleda izvodljivo u aplikacijama gde je detektor izložen veoma niskom fluksu zračenja kao što su noćni uslovi snimanja pod prirodnim izvorima osvetljenja. Iako InGaAs tehnologija (na indijum-fosfor (InP) supstratu) nudi najbolji odnos performansi i cene, granična talasna dužina ovih senzora je ograničena na 1.7  $\mu\text{m}$  kako bi se izbegli degradirajući efekti neusklađenosti rešetke. Fotodetektori  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$  usklađenih rešetki dostupni su sa gustinom tamne struje na sobnoj temperaturi reda veličine  $\text{nA}/\text{cm}^2$  (što odgovara tammim strujama piksela reda veličine  $\text{fA}$ ) omogućavajući snimanje pod veoma niskim nivoima fluksa. Međutim, postoje mnoge aplikacije sa zahtevom da senzor pokrije prošireni opseg talasnih dužina sa graničnom talasnom dužinom od 2.5  $\mu\text{m}$  ili više, pri čemu postoji velika neusklađenost rešetke između InGaAs apsorbera i InP supstrata. Iako tehnologija HgCdTe nema ovo ograničenje, infrastruktura materijala II-VI i visokokvalitetni CdZnTe supstrati nisu tako široko dostupni kao III-V, što ovu tehnologiju čini još uvek skupom, posebno za komercijalne primene. S druge strane, tehnologija super rešetke tipa II još uvek nije ponudila očekivane performanse detektora na nivou kvantne efikasnosti i tamne struje.

U poređenju sa drugim infracrvenim sensorima ili termalnim sensorima, troškovi SWIR senzora su obično veći, manje su rezolucije i veće težine [52]. Međutim, jedinstvena svojstva SWIR senzora mogu pružiti vredne informacije tamo gde su senzori koji rade u vidljivom opsegu, NIR ili termalni senzori neefikasni.

## 2.1 Karakteristike SWIR kamera, prednosti i ograničenja

Cilj sistema za video praćenje pokretnih objekata jeste da obezbedi mogućnost efikasnog rada nezavisno od doba dana ili noći, kao i u svim vremenskim uslovima. Kada je scena koja se snima dobro osvetljena (prirodnim ili veštačkim izvorom osvetljenja) i kada su vremenski uslovi povoljni, kamere koje rade u vidljivom opsegu, prvenstveno kolor kamere, daju najinformativniju i ljudskom oku najprijemčiviju sliku. Za rad u noćnim uslovima i uslovima slabog osvetljenja, infracrvene kamere koje rade u MWIR ili LWIR opsegu predstavljaju neophodnu komponentu sistema za video praćenje. Ove kamere snimaju temperaturnu raspodelu na sceni, u mogućnosti su da detektuju temperaturnu razliku između objekta i pozadine i na taj način obezbeđuju sliku dobrog kontrasta. Ovo je posebno važno za detekciju i praćenje objekata na velikim rastojanjima u noćnim uslovima. SWIR kamere su komplementarne MWIR, LWIR i kolor kamerama u određenim uslovima loše vidljivosti na sceni poput situacija gde postoji prisustvo dima ili isparenja, ali i u vremenskim uslovima kao što su izmaglica i magla [56].

Iako se svetlost u SWIR opsegu reflektuje od objekata na sličan način kao i u vidljivom i NIR opsegu, vidljivost je u navedenim izazovnim uslovima sredine poboljšana usled dužih talasnih dužina u SWIR opsegu koje su manje pod uticajem Rejljevog efekta rasejanja [52]. Dok male čestice (izmaglica ili čestice dima) dovode do rasipanja vidljive svetlosti, efekti rasipanja u SWIR opsegu su manji. Interagovanje svetlosti sa atomima i molekulima koji su značajno manji od talasne dužine svetlosti utiče da se svetlosni zraci rasipaju (odnosom suprotno proporcionalnom talasnoj dužini, na 4. stepen) i predstavlja fenomen poznat kao Rejljevo rasejanje, što onemogućava kratkim talasnim dužinama prostiranje do velikih udaljenosti. Kako svetlost nailazi na sferne čestice čije su dimenzije blizu ili veće od njene talasne dužine, ona slabi stopom sporijom od odnosa  $\frac{1}{\lambda}$ , gde  $\lambda$  označava talasnu dužinu. Sposobnost snimanja kroz izmaglicu je ključna prednost za poboljšanu vidljivost na velikim udaljenostima u aplikacijama za nadzor u poređenju sa kamerama koje rade u vidljivom opsegu. U slučaju šumskih požara ili požara naftnih postrojenja, sposobnost SWIR-a da obezbedi mogućnost efikasnog snimanja kroz čestice dima ima veliki značaj. U prisustvu velikih čestica, koje su 10 puta veće od talasne dužine, sve talasne dužine su podjednako oslabljene [12].

Meteorološke studije prostiranja kroz izmaglicu i maglu, koje potiču još iz pedesetih godina dvadesetog veka [57], dokazale su teorijska očekivanja u pogledu prenosa elektromagnetne energije koja pripada vidljivom i infracrvenom opsegu. U referentnoj metodologiji za proračun spektrofotometrijske apsorpcije koristi se optička gustina po kilometru, veličina koja zavisi od radijusa i broja kapi sadržanih u jedinici zapremine atmosfere. Za izmaglicu ova vrednost je manja od 2, a za maglu ova vrednost raste i do 30. Opšti zaključak je da u slučaju izmaglice talasna dužina za maksimalnu gustinu varira između 0.4 i 0.55  $\mu\text{m}$ , a gustina se brzo smanjuje sa povećanjem talasne dužine. U ovom scenariju, prenos u SWIR opsegu je mnogo veći nego u vidljivom opsegu. U slučaju magle, rezultati su pokazali da prenos u SWIR opsegu nije bio veći u odnosu na prenos u vidljivom opsegu.

U literaturi se navode i druge prednosti SWIR opsega u odnosu na vidljivi i NIR opseg [13]. Komparativna analiza prikazana je u tabeli 2.1. Identifikovano je nekoliko potencijalnih upotreba gde SWIR ima prednosti u odnosu na vidljivi opseg i NIR, od kojih su, u slučaju sistema za video nadzor, najvažniji: prostiranje kroz izmaglicu na moru, prostiranje u uslovima šumskih i naftnih požara,

atmosferska transmisija, kontrast pomorskih i kopnenih objekata, turbulencija i identifikacija na velikim dometima. Iako SWIR kamere najčešće imaju manju rezoluciju, one su uporedive sa mnogim trenutno dostupnim termalnim infracrvenim kamerama.

**Tabela 2.1** Poređenje vidljivog, NIR i SWIR opsega u dnevnim uslovima.

	Vidljivi opseg	NIR	SWIR
Prostiranje kroz izmaglicu na moru	loše	umereno	dobro
Prostiranje kroz maglu	loše	loše	umereno
Atmosferska transmisija	loše	umereno	dobro
Prostiranje kroz oblake	loše	loše	loše
Prostiranje u uslovima šumskih požara	loše	umereno	dobro
Vidljivost laserskih zraka	umereno	umereno	dobro
Detekcija/identifikacija kamuflaže	umereno	umereno	umereno
Kontrast urbane i ruralne pozadine	dobro	dobro	umereno
Kontrast pomorskih i kopnenih ciljeva	umereno	dobro	dobro
Detekcija kože	umereno	umereno	umereno
Spektralna diskriminacija	umereno	umereno	dobro
Turbulencija (blur i izobličenje slike)	umereno	umereno	dobro



a)



b)

**Slika 2.3** Scena raskrsnice u uslovima magle na rastojanju od 1 km: a) vidljivi opseg, b) SWIR.



a)



b)

**Slika 2.4** Zgrada u uslovima magle na rastojanju od 4.3 km: a) vidljivi opseg, b) SWIR.

Eksperimentalnim testovima sprovedenim korišćenjem multisenzorskog sistema VMSIS (*Vlatacom Multi Sensor Imaging System*) [58] potvrđene su izložene prednosti SWIR kamere u odnosu na kamere koje rade u vidljivom opsegu. Na slikama 2.3 i 2.4, predstavljene su scene u uslovima magle snimljene kamerom u vidljivom opsegu i SWIR kamerom. Na datim slikama jasno

se mogu videti prednosti korišćenja SWIR kamera u uslovima magle. Na slikama 2.5 – 2.7 je dat uporedni prikaz scena u uslovima jake kiše snimljenih kamerama u vidljivom opsegu, termalnom kamerom (MWIR opseg) i SWIR kamerom. I u ovom slučaju se zapaža da slike snimljene SWIR kamerom mogu pružiti najviše detalja o objektima na sceni. Pored toga što same padavine smanjuju vidljivost objekata na sceni, kada se kapljice vode nađu na površini objektiva kamere, one posledično utiču na zamućenje slike. Iz datih eksperimentalnih primera se može videti da SWIR kamere pokazuju rezistentnost na opisanu pojavu.



a) b) c)  
**Slika 2.5** Zgrada u uslovima kiše: (a) vidljivi opseg, (b) MWIR, (c) SWIR.



a) b) c)  
**Slika 2.6** Scena ulice sa pokretnim objektima u uslovima kiše: a) vidljivi opseg, b) MWIR, c) SWIR.

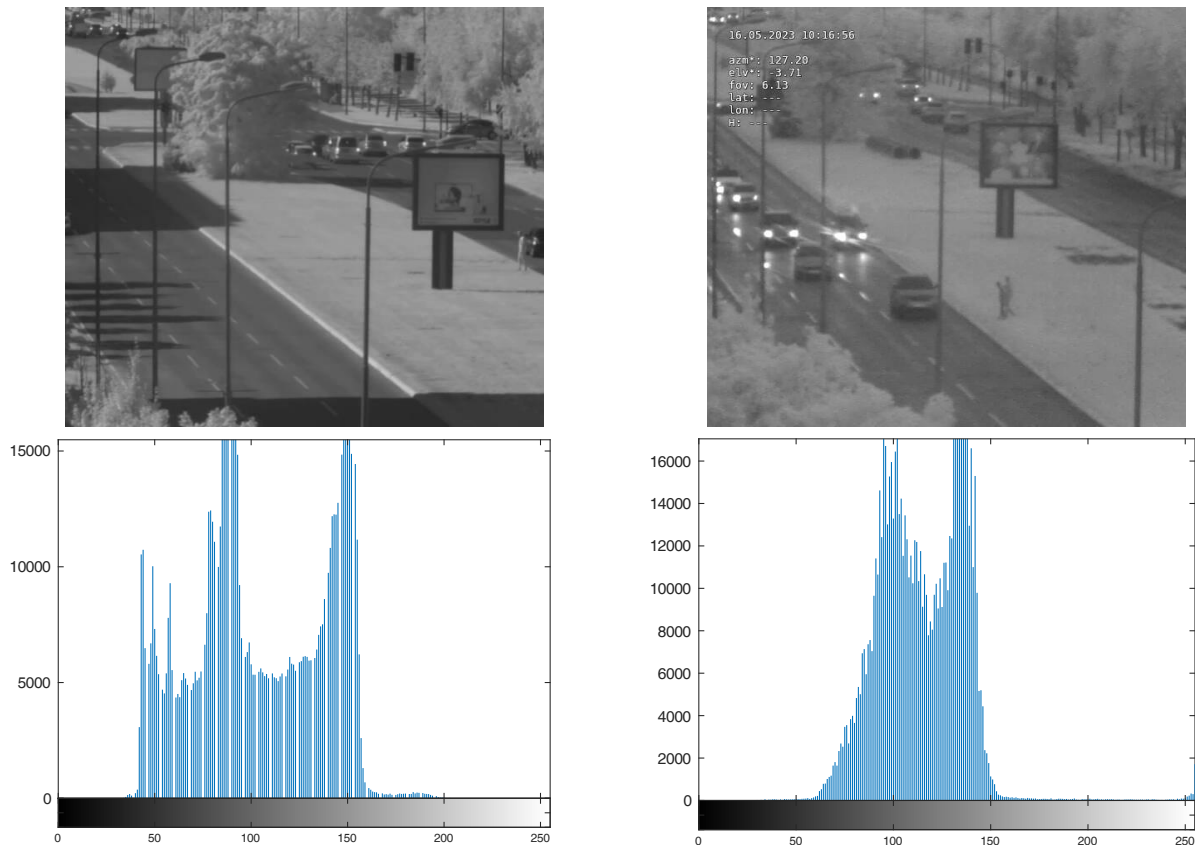


a) b) c)  
**Slika 2.7** Scena raskrsnice sa pokretnim objektima u uslovima kiše: a) vidljivi opseg, b) MWIR, c) SWIR.

Dalja analiza se može izvršiti posmatranjem iste scene snimljene SWIR kamerom u različitim vremenskim uslovima i uslovima osvetljenja. Na slici 2.8 levo je prikazana scena snimljena u dobrim vremenskim uslovima sunčanog dana što ukazuje da je scena dobro osvetljena. Na slici desno je prikazana ista scena snimljena tmurnog dana u uslovima kiše i izmaglice, uz slabu osvetljenost scene. Može se videti da obe slike sadrže dovoljno informacija o objektima, što ukazuje da navedeni izazovni uslovi ne utiču značajno na promenu sadržaja slike. Ako posmatramo histograme ovih slika prikazane na slici 2.8 ispod slika scena, vidi se da histogram slike snimljene u dobrim uslovima ne pokriva ceo dinamički opseg, odnosno da su SWIR slike snimljene i u dobrim uslovima scene nižeg



kontrasta. Niski kontrast na slici može biti veoma izazovan problem u praćenju objekata što otežava razlikovanje objekta od pozadine. Ako se pogleda histogram slike snimljene u lošim vremenskim uslovima, vidi se da je širina histograma neznatno uža u odnosu na histogram slike snimljene u dobrim uslovima. Pored loših vremenskih uslova i smanjenog osvetljenja scene, uticaj na izgled histograma ima i to što na sceni nisu prisutni potpuno isti objekti kao oni snimljeni na prvoj slici u dobrim uslovima.



**Slika 2.8** Ista scena snimljena SWIR kamerom u dobrim vremenskim uslovima i pri dobrom osvetljenju (levo) i snimljena u uslovima kiše i izmaglice pri slabom osvetljenju (desno). Ispod slika scena prikazani su histogrami ovih slika.

U zavisnosti od potrebe za dodatnim izvorom zračenja, senzori se mogu podeliti na pasivne i aktivne. Pasivni senzori ne zahtevaju dodatne izvore zračenja za formiranje slike scene, dok je za aktivne taj izvor neophodan. Aktivni senzori iz tog razloga mogu biti otkriveni, što je naročito važno u vojnim primenama. U tom pogledu, pasivni senzori su u prednosti. Međutim, za problem identifikacije objekata, posebno onih na velikim rastojanjima od senzora, konvencionalni pasivni infracrveni senzori koji rade u MWIR i LWIR opsegu imaju ograničenja [56], [13].

Za identifikaciju na velikim rastojanjima, najadekvatnije karakteristike imaju reflektivne NIR i SWIR kamere [59]. Pošto je njihova talasna dužina mala, veličina piksela može biti manja, a rezolucija veća od MWIR i LWIR kamera. U poređenju sa vidljivim opsegom, ovi senzori su manje podložni negativnim atmosferskim efektima. Iako je rezolucija SWIR kamera povećana, zahteva za radom i u dnevnim i u noćnim uslovima, kao i u svim vremenskim prilikama, konvencionalne pasivne SWIR kamere ne mogu da zadovolje. Da bi se poboljšale performanse pasivnih SWIR kamere u uslovima slabog osvetljenja ili noćnim uslovima, potrebno je aktivno snimanje, odnosno korišćenje laserskog izvora za osvetljenje. To je poznata tehnika koja se koristi za noćno snimanje ili za



poboljšanje efikasnosti snimanja u okruženjima sa rasipanjem. Uklanjanje efekata povratnog rasipanja dovodi do značajnog povećanja dometa snimanja [59]. Za aktivno snimanje najčešće se koristi impulsni laser talasnih dužina oko 1550 nm, jer je bezbedan za oči (eng. *eye safe laser*), a takođe odgovara atmosferskom prozoru koji ima minimalno slabljenje. Princip aktivnog snimanja može se primeniti korišćenjem dva koncepta. Prvi koncept koristi tipičan impulsni laser i naziva se *gated snimanje*, dok je drugi koncept onaj u kome se koristi kontinualni laser za osvetljenje [60] – [62].

Komparativna analiza reflektivnih aktivnih i pasivnih NIR/SWIR senzora, kao i emisivnih MWIR/LWIR senzora u različitim uslovima i prema različitim kriterijumima, data je u tabeli 2.2 [59].

**Tabela 2.2** Poređenje mogućnosti reflektivnih i emisivnih senzora.

	Reflektivni NIR/SWIR		Emisivni
	Aktivni	Pasivni	MWIR/LWIR
Snimanje u potpunom mraku	✓	✗	✓
Merenje razdaljine	✓	✗	✗
Snimanje kroz staklo	✓	✓	✗
Identifikacija	✓	✓	✗
Scena niskog termalnog kontrasta	dug domet	kratak domet	limitirano
Otpornost na parazitske izvore svetla	visok nivo	limitirano	-
Rad u uslovima kiše, snega, magle	uspešno	umereno	umereno

Pored primena u aplikacijama za nadzor, zahvaljujući dobrom prostiranju kroz različite uslove sredine, posebno u uslovima velike vlažnosti, izmaglice, magle i dima, SWIR kamere nalaze primenu i u mnogim drugim oblastima. SWIR senzori mogu da obezbede vidljivost kroz mnoge korišćene materijale za pakovanje koji su u SWIR opsegu providni, dok su u vidljivom spektru neprozirni. U proizvodnji poljoprivrednih proizvoda koriste se za uočavanje nedostataka i procenu kvaliteta voća i povrća, za sortiranje po stepenu vlažnosti ili za procenu sadržaja masti. Primenu nalaze i u farmaceutskoj industriji za procenu kvaliteta farmaceutskih proizvoda, ali i u medicini zbog mogućnosti prostiranja elektromagnetskog zračenja u ovom opsegu kroz tkiva, pa se SWIR kamere mogu koristiti za detekciju vena. Za inspekciju različitih proizvoda i materijala, SWIR kamere se zahvaljujući svojim osobinama koriste u nizu aplikacija, uključujući inspekciju elektronskih ploča, inspekciju prenosa poluprovodnika, inspekciju solarnih ćelija, automobilsku inspekciju, kao i za kontrolu kvaliteta procesa [12], [13]. Međutim, od interesa za ovu disertaciju su prvenstveno sistemi za nadzor, kao i aplikacije i algoritmi razvijeni za detekciju i praćenje objekata u SWIR video sekvenci.

## 2.2 Primena SWIR kamera za detekciju i praćenje objekata

Razvoj algoritama kompjuterske vizije za obradu slika i video sekvenci u SWIR opsegu pokrenula je upotreba SWIR kamera koje mogu pružiti komplementarne informacije u situacijama i aplikacijama u kojima su termalne kamere i kamere koje rade u vidljivom opsegu neefikasne. Algoritmi kompjuterske vizije za detekciju i praćenje objekata razvijeni za primenu na SWIR slici uključuju metode koje koriste tradicionalne pristupe izdvajanja obeležja sa slike, obučavanje modela mašinskog učenja koji se dalje koriste za detekciju, ali i metode koje se oslanjaju na duboko učenje i obučavaju se po principu od kraja do kraja (eng. *end-to-end*).

U [63], autori su predstavili *WVU Outdoor SWIR Gait (WOSG)* skup podataka kreiran SWIR kamerom za evaluaciju performansi algoritama za prepoznavanje hoda u SWIR video sekvencama. Evaluirani algoritmi ne zahtevaju veliki skup podataka, te ovaj skup podataka sadrži 155 subjekata sa informacijama o hodu prikupljenim u nekontrolisanom, spoljašnjem okruženju. Tri algoritma za prepoznavanje hoda na kreiranom SWIR skupu podataka su evaluirana, sa posebnim naglaskom na dizajn algoritama za segmentaciju i prepoznavanje pojedinaca u SWIR domenu zbog značajnih operativnih prednosti u odnosu na slike u vidljivom opsegu. Autori rada [64] su predstavili aktivni SWIR video sistem za detekciju, praćenje i identifikaciju ljudi na udaljenostima od nekoliko stotina metara. Predstavljen sistem ima mogućnost generisanja slika na kojima je moguće prepoznavati lica, u dnevnim i noćnim uslovima, sa uspešnim prepoznavanjem na udaljenostima do 350 metara. Detekcija pešaka na SWIR slikama u različitim uslovima vidljivosti, pri vedrom nebu, u uslovima izmaglice i magle je analizirana u [14]. Tradicionalni pristup zasnovan na histogramu orijentisanih gradijenata (HOG) i klasifikatoru nosećih vektora (eng. *Support Vector Machine - SVM*) oslonjenog na modele deformabilnih delova, i obučen na PASCAL skupu slika u opsegu vidljive svetlosti, primenjen je za detekciju pešaka. Isti pristup, koristeći SVM klasifikator i HOG obeležja, primenjen je u [15], gde je SVM klasifikator obučen na obeležjima dobijenim sa slika u vidljivom opsegu INRIA skupa, a testiranja su sprovedena u uslovima dobre vidljivosti. U [16] je primenjen aktivni SWIR sistem za taktičko snimanje za dnevno/noćni nadzor na velikim rastojanjima. Izvršena je analiza detekcije i praćenja osobe tokom hodanja, kao i automatskog prepoznavanja lica u dnevnim i noćnim uslovima na različitim udaljenostima od kamere u opsegu od 100 m do 350 m. U [18] se razmatra praćenje vozila direktno u domenu kompresivnih merenja, pri čemu su korišćene SWIR sekvence degradiranog kvaliteta slike sa nedostajućim podacima, odnosno prisustvom mrtvih piksela. Primenjena je mreža *ResNet-18* (neuralna mreža sa 18 konvolucionih slojeva) kako bi se izbegao problem saturacije performansi neuralne mreže, kao i degradacije prilikom obučavanja dubljih slojeva. Obučen je model za dve klase, vozilo i pozadina, koristeći oko 150.000 slika vozila i oko 500.000 slika pozadine. Autori su pokazali da metod za praćenje zasnovan na dubokom učenju ima bolje performanse od konvencionalnih algoritama za praćenje i predložili nova unapređenja i pravce budućih istraživanja. Dalje, autori su poboljšali performanse praćenja automobila u [19] – [21] uključujući *YOLO* model verzije 3 za praćenje objekata. Klasifikacija se i ovde vrši primenom *ResNet* mreže zbog bolje klasifikacije u odnosu na *YOLO* model u domenu kompresivnih merenja za SWIR video sekvence. Međutim, praćenje objekata u sekvenci degradiranih SWIR slika nije predmet istraživanja ove disertacije. U [17] je razvijen sistem kamera za nadzor koji se sastoji od SWIR, termalne i hiperspektralne kamere sa mogućnošću detekcije objekata. Nezavisno od detektora, razvijena je procedura za detekciju objekata integrišući tehniku detekcije pokreta (na principu oduzimanja pozadine) i duboke konvolucione neuralne mreže. Na osnovu detektora objekata (oduzimanje pozadine, *Faster R-CNN* i *YOLOv2*), mogući detektovani objekti na termalnim i SWIR slikama su zatim projektovani u lokalni koordinatni sistem primenom transformacione matrice, a rezultujuće koordinate su zatim projektovane na ravan slike hiperspektralne kamere koristeći inverznu transformacionu matricu. Za projektovane objekte može se zatim direktno vršiti fuzija spektralnih informacija iz hiperspektralnog opsega. Finalne detekcije objekata se izdvajaju nakon njihove spektralne verifikacije i prepoznavanja, na osnovu asocijacije podataka za koju se eksploatišu njihova temporalna spektralna obeležja. U [65] je predstavljena autonomna platforma koja sadrži četiri SWIR kamere sa zadatkom prikupljanja podataka i testiranja algoritama za autonomna vozila. Platforma predstavlja sistem kompjuterske vizije, pa je pored SWIR kamera opremljena i sa deset

kamera male fokalne dužine, dve kamere velike fokalne dužine i četiri lidara. Kao takva, predstavljena platforma ima mogućnost prikupljanja video sekvenci, podataka u 3D-u, kao i testiranja algoritama za trodimenzionalnu rekonstrukciju, semantičku segmentaciju i klasifikaciju prepreka. U [22] je predložen algoritam za praćenje više objekata u SWIR video sekvenci sa malim brojem frejmova u sekundi za primenu u aplikacijama nadzora u pametnim gradovima. Predloženi algoritam za praćenje je iz klase algoritama praćenja na osnovu detekcija. Za detekciju objekata iz klasa ljudi, automobila, motocikala, autobusa i kamiona su primenjeni modeli *YOLOv5* i *YOLOv7* obučeni na namenski kreiranom skupu od oko 7300 SWIR slika. Za praćenje objekata, primenjena je dvokoračna metoda za pridruživanje novih detekcija uspostavljenim tragovima. Metrika za pridruživanje u prvom koraku predstavlja Euklidsko rastojanje između centara okružujućih okvira detekcija i tragova, a u drugom koraku je primenjena IOU (eng. *Intersection Over Union*) metrika između okružujućih okvira detekcija i tragova. Samo pridruživanje detekcija tragovima se vrši primenom Mađarskog algoritma.

Analizom dostupne literature može se uočiti da su istraživanja u domenu praćenja objekata u SWIR video sekvencama uglavnom fokusirana na praćenje više objekata, odnosno uočava se nedostatak pristupa koji se bave praćenjem jednog objekta. Fokus je prvenstveno na razvoju modela mašinskog i dubokog učenja za detekciju objekata na SWIR slikama. Oslanjajući se na obučene modele za detekciju objekata, predloženi algoritmi se bave praćenjem više objekata u paradigmi praćenja na osnovu detekcija (eng. *Tracking by Detection*). Stoga, praćenje u navedenim pristupima podrazumeva unapred definisane klase objekata koji se mogu pratiti, a koje se moraju odrediti prilikom obučavanja modela za detekciju objekata. Dodatno, može se uočiti nedostatak pristupa koji se usko bave praćenjem objekata u realnom vremenu. S druge strane, predmet istraživanja u ovoj disertaciji je praćenje jednog pokretnog objekta u SWIR video sekvenci, nezavisno od tipa objekta, uz mogućnost dugotrajnog praćenja u realnom vremenu.

## 2.3 Skup SWIR video sekvenci za praćenje pokretnih objekata

Za razvoj sistema za video praćenje neophodno je imati odgovarajući skup video sekvenci. Analizirani su referentni skupovi podataka za ocenu performansi algoritama za praćenje objekta: OTB100 [24], TC128 [25], NUS-PRO [26], UAV123 [27], OxUvA [28], TrackingNet [29], LaSOT [66], GOT-10K [67], LTIR [68], PTB-TIR [10], LSOTB-TIR [11], kao i svi *VOT challenges* skupovi podataka od VOT2015 [5] do VOT2022 [2]. Iako analizirani skupovi podataka sadrže veliki broj različitih sekvenci koje su snimljene kamerama koje rade u vidljivom opsegu, tako i onim u infracrvenom opsegu, može se uočiti nedostatak javno dostupnih skupova video sekvenci u kratkotalasnom infracrvenom opsegu kreiranih za problem praćenja pokretnih objekata. Stoga, za razvoj algoritma za praćenje objekata u SWIR opsegu, analizu algoritma, kao i za finalno testiranje u izazovnim scenarijima potrebno je kreirati novi skup SWIR video sekvenci.

Skup sekvenci koji se kreira treba da ispuni nekoliko uslova. Prvo, za pravilnu statističku analizu algoritma za praćenje, skup sekvenci treba da sadrži dovoljno veliki broj frejmova. Zatim, trebalo bi da sadrži pokretni objekat od interesa u različitim scenarijima. Ovi scenariji uključuju promene u dinamici kretanja objekta, promene u orijentaciji, promene veličine objekta tokom praćenja, i što je najizazovnije, prisustvo okluzija (delimičnih i potpunih, kratkotrajnih i dugotrajnih, pokretnih i statičnih), što sve predstavlja uzroke grešaka u praćenju. Izazovi koji se javljaju u video

praćenju pokretnih objekata će detaljnije biti razmatrani u sekciji 3.1. Takođe, za pravilnu analizu, svi scenariji treba da budu snimljeni istom kamerom i u istim uslovima.

Skup sekvenci kreiran za potrebe istraživanja u okviru ove disertacije je veoma izazovan i sadrži sekvence koje ilustruju realistične urbane scene. Treba da pruži dobru osnovu za razvoj kvalitetnog algoritma za praćenje pokretnih objekata u SWIR opsegu u različitim scenarijima koji se mogu naći u aplikacijama za nadzor. U referentnim skupovima podataka za video praćenje najčešći objekti od interesa za praćenje su pešaci, pa su pešaci odabrani kao objekti od interesa za istraživanje i statističku analizu algoritma za praćenje i u ovoj disertaciji. Pešaci predstavljaju tipične objekte sa relativno sporom i brzom promenom dinamike kretanja, što je posebno važno u analizi algoritma za praćenje. Tipičan frejm iz skupa sekvenci sa označenim objektom od interesa je prikazan na slici 2.9.



**Slika 2.9** Frejm iz skupa SWIR sekvenci sa okružujućim okvirom oko objekta od interesa za praćenje.

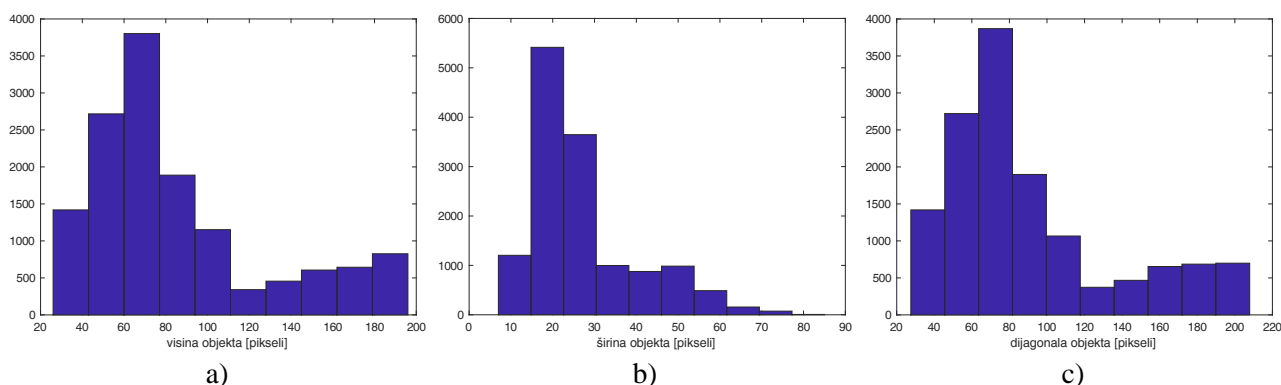
Kreirano je 16 video sekvenci, gde je pešak koji se kreće objekat od interesa za praćenje. Kreirani skup podataka sadrži ukupno 13852 frejma, što se može smatrati dovoljnim brojem za detaljnu statističku analizu rada algoritma za praćenje. Objekat od interesa na svakom frejmu je manuelno anotiran odgovarajućim okružujućim okvirom. Centar okružujućeg okvira predstavlja poziciju praćenog objekta i zajedno sa širinom i visinom okružujućeg okvira za dati frejm predstavlja referentne podatke (eng. *ground-truth*) za dalje eksperimente. Video sekvence snimljene su SWIR kamerom, rezolucije  $576 \times 504$  piksela i brzinom od 25 frejmova u sekundi (eng. *frames per second - FPS*). Korišćena kamera je implementirana u tehnologiji isprepletenog prikaza (eng. *interlaced*), gde se naizmenično iscrtavaju parne i neparne linije slike, što predstavlja dodatni izazov za algoritam praćenja prilikom ekstrakcije obeležja iz slike. Posebno kod objekata u pokretu, na slici dobijenoj u ovoj tehnologiji, javljaju se iskrzane ivice jer se na slici istovremeno nalaze parni redovi prethodnog i neparni redovi posmatranog (trenutnog) frejma sekvence.

U kreiranom skupu sekvenci, prvu grupu predstavljaju sekvence snimljene kamerom u fiksnoj položaju. U okviru ove grupe snimljeno je 10 video sekvenci sa različitim kombinacijama situacija koje se mogu smatrati izazovnim za video praćenje. Detaljne informacije o svim scenarijima i prisutnim izazovima praćenja su date u tabeli 2.3. Izdvojeni frejmovi iz ove grupe sekvenci su prikazani na slikama 2.11 – 2.20. Na svakoj slici prvi frejm s leve strane predstavlja prvi frejm u video sekvenci, na kome se vrši inicijalizacija praćenja. Drugu grupu sekvenci predstavljaju sekvence snimljene kamerom koja je tokom snimanja bila u pokretu na pan-tilt pozicioneru. Ova grupa video

sekvenci je prikazana na slikama 2.21 – 2.26. I ova grupa sekvenci obuhvata različite kombinacije izazova koji se javljaju u praćenju objekata. Kako okluzije predstavljaju jedan od najvećih izazova u praćenju, tabela 2.3 za svaku sekvencu daje opis okluzija koje se u njoj javljaju. Dodatno, pešak se tokom snimanja video sekvenci kretao na udaljenosti od 70 m do 200 m od kamere. Površina praćenog objekta u ravni slike u celokupnom skupu sekvenci se kreće od 196 do 15315 piksela. Minimalna dijagonala okružujućeg okvira u skupu iznosi 27, a maksimalna 208 piksela. Na slici 2.10 su dati histogrami koji prikazuju raspodele dimenzija objekta u kreiranom skupu SWIR video sekvenci, gde se uočava da su dimenzije objekta na slici veoma male, a što je karakteristično za primenu sistema za praćenje objekata koji se nalaze na velikim razdaljinama od kamere.

**Tabela 2.3** Karakteristike kreiranog skupa SWIR video sekvenci za praćenje pokretnog objekta.

Sekvenca	Broj frejmova	Okluzije						Klater	Promena dinamike	Promena veličine	Kretanje kamere
		Dugotrajne	Kratkotrajne	Delimične	Potpune	Statične	Pokretne				
1.	825		✓	✓	✓	✓		✓	✓		
2.	396		✓	✓	✓	✓		✓			
3.	522	✓	✓	✓	✓	✓		✓			
4.	420		✓	✓		✓	✓				
5.	225		✓	✓	✓	✓	✓				
6.	199	✓	✓	✓	✓	✓	✓				
7.	1224	✓	✓	✓	✓	✓	✓	✓	✓		
8.	1392	✓			✓	✓		✓	✓		
9.	1122	✓	✓	✓	✓	✓					
10.	1724	✓			✓	✓	✓	✓	✓		
11.	392		✓		✓	✓				✓	
12.	594		✓		✓	✓		✓		✓	
13.	670	✓	✓	✓	✓	✓	✓		✓	✓	
14.	675						✓	✓	✓	✓	
15.	728	✓	✓	✓	✓	✓	✓			✓	
16.	2744	✓	✓	✓	✓	✓			✓	✓	



**Slika 2.10** Histogrami dimenzija okružujućih okvira objekta u skupu SWIR video sekvenci:  
 a) visina b) širina c) dijagonala.



**Slika 2.11** Sekvenca 1. Kretanje pešaka u scenariju kratkotrajnih delimičnih i potpunih statičnih okluzija sa promenom orijentacije i veličine objekta.



**Slika 2.12** Sekvenca 2. Kretanje pešaka u scenariju kratkotrajnih delimičnih i potpunih statičnih okluzija sa promenom orijentacije i smera kretanja.



**Slika 2.13** Sekvenca 3. Kretanje pešaka uz promenu dinamike kretanja u scenariju sa jednom dugotrajnom i više kratkotrajnih statičnih okluzija.



**Slika 2.14** Sekvenca 4. Kretanje pešaka u scenariju klatera (blizina sličnog objekta objektu od interesa), sa jednom kratkotrajnom delimičnom okluzijom i podrhtavanjem kamere uzrokovano poremećajem.



**Slika 2.15** Sekvenca 5. Kretanje pešaka u scenariju sa delimičnim i potpunim statičnim okluzijama i delimičnim pokretnim okluzijama.





**Slika 2.16** Sekvenca 6. Kretanje pešaka u scenariju sa jednom kratkotrajnom potpunom statičnom okluzijom, jednom dugotrajnom potpunom pokretnom i više delimičnih pokretnih okluzija, uz prisustvo klatera.



**Slika 2.17** Sekvenca 7. Kretanje pešaka u scenariju sa kratkotrajnim delimičnim i potpunim statičnim okluzijama, kao i kratkotrajnim delimičnim i dugotrajnim potpunim pokretnim okluzijama.



**Slika 2.18** Sekvenca 8. Kretanje pešaka u scenariju sa dugotrajnom potpunom statičnom okluzijom, uz izrazitu promenu dinamike kretanja i promene uslova osvetljenja scene u kojoj se kreće pešak.



**Slika 2.19** Sekvenca 9. Kretanje pešaka u scenariju sa kratkotrajnom delimičnom i dugotrajnom potpunom statičnom okluzijom, pri veoma niskom kontrastu između objekta i pozadine.



**Slika 2.20** Sekvenca 10. Kretanje pešaka u scenariju sa kratkotrajnom delimičnom i dugotrajnom potpunom statičnom okluzijom, pri veoma niskom kontrastu između objekta i pozadine, uz prisustvo klatera.



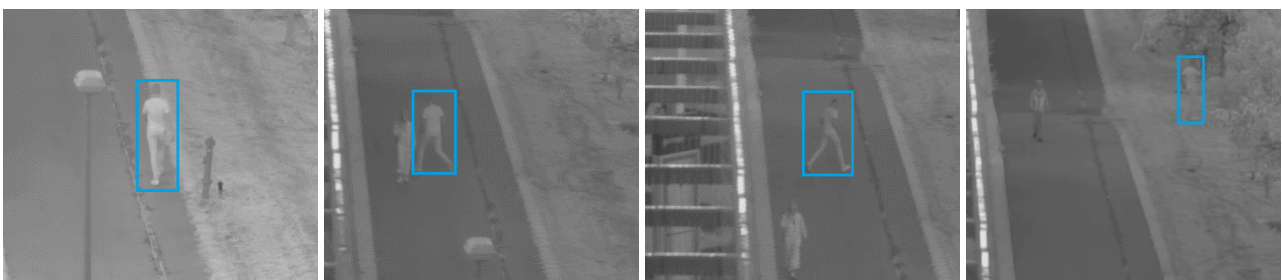
**Slika 2.21** Sekvenca 11. Kretanje pešaka u scenariju sa kratkotrajnim delimičnim i potpunim statičnim okluzijama. Sekvenca je snimljena kamerom u pokretu.



**Slika 2.22** Sekvenca 12. Kretanje pešaka u scenariju sa kratkotrajnim delimičnim i potpunim statičnim okluzijama uz promenu orijentacije i dinamike kretanja. Sekvenca je snimljena kamerom u pokretu.



**Slika 2.23** Sekvenca 13. Kretanje pešaka u scenariju sa jednom dugotrajnom statičnom okluzijom i više kratkotrajnih delimičnih statičnih okluzija. Sekvenca je snimljena kamerom u pokretu.



**Slika 2.24** Sekvenca 14. Kretanje pešaka u scenariju sa klaterom, promenom orijentacije i dinamike kretanja, kao i veličine objekta usled udaljavanja od kamere. Sekvenca je snimljena kamerom u pokretu.



**Slika 2.25** Sekvenca 15. Kretanje pešaka u scenariju sa dugotrajnom potpunom statičnom okuzijom i delimičnim kratkotrajnim pokretnim i statičnim okluzijama. Sekvenca je snimljena kamerom u pokretu.





**Slika 2.26** Sekvenca 16. Kretanje pešaka u scenariju sa statičnim delimičnim i potpunim okluzijama uz smanjenje veličine objekta tokom kretanja. Sekvenca je snimljena kamerom u pokretu.

### 3 Praćenje pokretnih objekata u sekvenci slika

Praćenje pokretnih objekata u sekvenci slika, odnosno video praćenje, predstavlja proces lociranja pokretnih objekata i procene njihove trajektorije tokom vremena [1]. Pozicija objekta koji se prati određuje se uspostavljanjem relacije između odgovarajućih piksela na trenutnom frejmu video sekvence i modela objekta. Model objekta predstavlja definiciju objekta koji se prati u video sekvenci, odnosno način na koji algoritam vidi objekat od interesa. Ovu relaciju nije jednostavno uspostaviti, štaviše, predstavlja jedan od najvećih izazova u video praćenju.

U zavisnosti od broja objekata koji se istovremeno prate u video sekvenci, mogu se razlikovati praćenje jednog objekta i praćenje više objekata [1]. Praćenje više objekata je poslednjih godina doživelo veliki razvoj, prvenstveno razvojem dubokih neuralnih mreža koje predstavljaju osnovu na koju se algoritmi za praćenje više objekata oslanjaju [22], [69] – [71]. Praćenje jednog objekta se može posmatrati kao podskup praćenja više objekata. Gledano na taj način, može se pomisliti da je praćenje jednog objekta pojednostavljenje problema praćenja više objekata. Međutim, praćenje jednog objekta pruža mogućnosti koje su znatno ograničene kod praćenja više objekata. Prvenstveno se to ogleda u dužini praćenja objekata. Prilikom praćenja jednog objekta moguće je koristiti PTZ (eng. *Pan-Tilt-Zoom*) [1] kamere i tako obezbediti značajno duži period praćenja objekta. Takođe, razlike u pristupu dizajna algoritma za praćenje jednog objekta mogu se videti u manjoj kompleksnosti samog algoritma, što za posledicu ima mogućnost izvršavanja algoritma u realnom vremenu na platformama koje imaju manju procesorsku moć. Predmet od interesa u ovoj disertaciji jeste praćenje jednog pokretnog objekta u video sekvenci.

Praćenje jednog objekta u video sekvenci može se formulisati kao estimacija vremenske serije:

$$\mathbf{x} = \{x_k\}; \quad k = 1, 2, 3, \dots \quad (3.1)$$

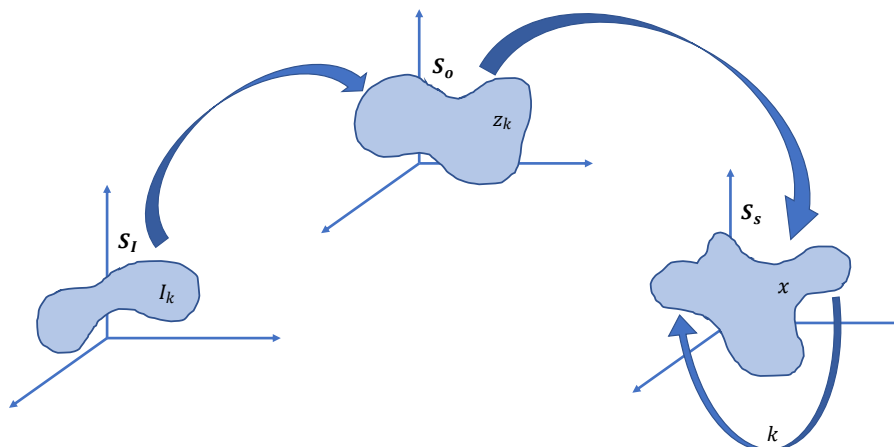
u skupu diskretnih vremenskih instanci indeksiranih sa  $k$ , na osnovu informacija prikupljenih iz skupa:

$$\mathbf{I} = \{I_k\}; \quad k = 1, 2, 3, \dots \quad (3.2)$$

gde je  $I_k$  frejm video sekvence u trenutku  $k$ , definisan u prostoru svih mogućih slika  $S_I$  [1]. Vektor  $x_k$  predstavlja stanje objekta koji se prati u trenutku  $k$  i definisan je u prostoru stanja  $S_S$ . Dimenzionalnost, kao i tip informacija koje će biti sadržane u vektoru stanja objekta  $x_k$  zavise od dizajna algoritma za praćenje, ali i krajnje primene samog algoritma. Definisana vremenska serija  $\mathbf{x}$  u (3.1) zapravo predstavlja trajektoriju objekta koji se prati, u prostoru stanja  $S_S$ .

Kao prelazni prostor između prostora slika  $S_I$  i prostora stanja  $S_S$ , definiše se prostor obeležja  $S_O$ . Tako se slike  $I_k$  iz video sekvence mapiraju u promenljive  $z_k$  u prostoru obeležja  $S_O$ . Ovaj prostor obično ima manju dimenzionalnost u odnosu na ulazni prostor  $S_I$ , ali sadrži informacije o objektu koje su relevantne za zadatak praćenja. Postupak kojim se vrši transformisanje ulaznog prostora slika

$S_I$  u prostor obeležja  $S_o$  predstavlja operaciju za ekstrakciju obeležja [1]. Na osnovu ekstrahovanih obeležja, algoritam za praćenje informacije dobijene sa slike koristi za rekurzivnu procenu stanja objekta. Dalje, strategija za lokalizaciju objekta definiše kako se mogu koristiti dobijena obeležja za što bolju procenu stanja objekta  $x_k$ . Tok informacija između opisanih prostora stanja važnih za praćenje objekta u video sekvenci prikazan je na slici 3.1.



**Slika 3.1** Tok informacija između prostora stanja u video praćenju.

Informacije koje su sadržane u vektoru stanja  $x_k$  mogu se grupisati u tri kategorije: informacije o lokaciji i obliku objekta, informacije o izgledu objekta i informacije o temporalnim varijacijama oblika i izgleda objekta [1].

1. Informacije o lokaciji i obliku objekta zavise od tipa objekta koji je odabran za praćenje, kao i od količine informacija koje se mogu dobiti iz frejmova video sekvence. U sukcesivnim frejmovima video sekvence, lokacija i oblik objekta se dobijaju kao rezultat izvršavanja algoritma i logike koju sam algoritam poseduje.
2. Informacije o izgledu objekta, odnosno način kodiranja informacija o izgledu ima funkciju u modelovanju varijacija izgleda objekta tokom vremena. Modelovanje varijacija izgleda objekta može da obezbedi dugotrajnije praćenje objekta čiji se izgled tokom vremena usled različitih faktora menja u odnosu na izgled koji je imao u trenutku inicijalizacije praćenja.
3. Parametri informacija o temporalnim varijacijama oblika i izgleda objekta predstavljaju izvod parametara prethodne dve kategorije. Najčešće su opcioni.

Sadržaj vektora stanja  $x_k$  treba da bude takav da dobro modeluje objekat od interesa za praćenje. Elementi vektora stanja, pored pozicije objekta na slici, što najčešće i predstavlja izlaz samog algoritma za praćenje, mogu sadržati i informacije vezane za veličinu objekta, dinamiku kretanja, rotaciju, ali i izgled i specifičnosti vezane za posebne objekte od interesa. Odatle se vidi da svi elementi vektora stanja ne moraju predstavljati izlaz algoritma za praćenje. Iako dodavanje većeg broja informacija o objektu u vektor stanja može poboljšati performanse praćenja, mora se voditi računa i o složenosti estimatora stanja. Povećanje dimenzionalnosti vektora stanja objekta povećava i složenost estimatora, pa se mora napraviti kompromis između čuvanja dovoljnog broja informacija koje poboljšavaju performanse praćenja i dimenzionalnosti vektora stanja. U slučaju ograničenih hardverskih resursa za izvršavanje algoritma, poželjno je da dimenzionalnost vektora  $x_k$  bude što manja kako bi se postiglo izvršavanje algoritma u realnom vremenu.

Za praćenje objekta u ravni slike, vektor stanja  $x_k$  minimalno može sadržati dva elementa. Ova dva elementa predstavljaju centar objekta koji se prati, odnosno horizontalnu ( $x_c$ ) i vertikalnu ( $y_c$ ) koordinatu u ravni slike. Vektor stanja definisan koordinatama centra objekta izražen je kao:

$$x_k = (x_c, y_c) \quad (3.3)$$

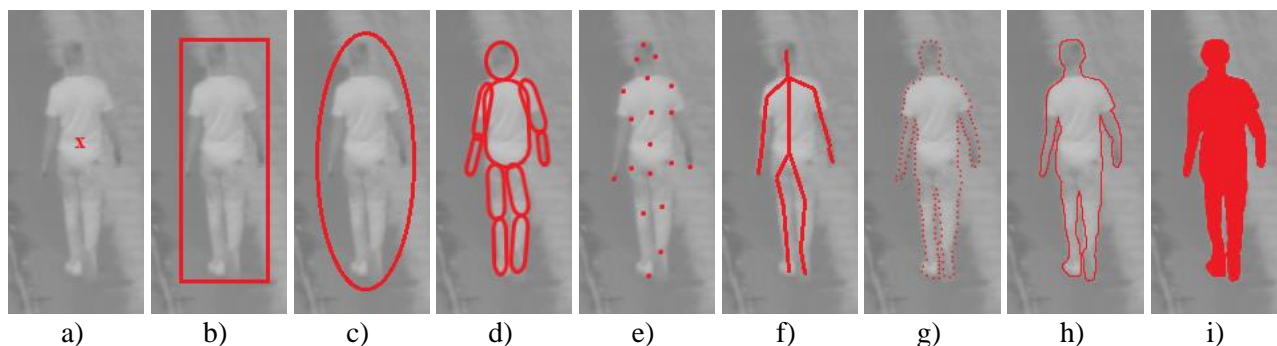
Međutim, za praćenje objekta u video sekvenci obično nije dovoljno poznavati samo položaj centra objekta na slici [1]. Zahtev je najčešće da se oko objekta koji se prati estimira i okružujući okvir [1] – [5] koji definiše dimenzije objekta na slici. U tom slučaju se pored pozicije centra objekta mora voditi evidencija i o širini ( $w$ ) i visini ( $h$ ) objekta u ravni slike. Tada vektor stanja  $x_k$  dobija oblik:

$$x_k = (x_c, y_c, w, h) \quad (3.4)$$

Štaviše, u vektor stanja objekta se može uvesti i parametar koji predstavlja ugao rotacije ( $\theta$ ) praćenog objekta u ravni slike, pa vektor stanja  $x_k$  dalje može dobiti oblik:

$$x_k = (x_c, y_c, w, h, \theta) \quad (3.5)$$

U savremenim algoritmima se sreću i složeniji modeli za reprezentaciju objekta koji se prati. Mogu se koristiti skupovi karakterističnih tačaka na objektu, sekvence tačaka na konturi objekta, cela kontura objekta, specifična obeležja objekta ili silueta objekta od interesa. Na slici 3.2 su prikazani različiti načini predstavljanja objekata u video praćenju.



**Slika 3.2** Predstavljanje objekta: a) centroid, b) pravougaonik, c) elipsa, d) delovi objekta, e) ključne tačke objekta, f) skelet objekta, g) ključne tačke konture objekta, h) kompletna kontura objekta, i) silueta objekta.

Model izabran da predstavlja oblik objekta može ograničavati tip kretanja ili deformacije kojima se objekat može podvrgnuti [72]. Tako se za objekat koji je predstavljen svojim centrom može koristiti samo model translacije. U slučaju kada se za objekat koristi reprezentacija geometrijskog oblika poput elipse, prikladni mogu biti parametarski modeli kretanja kao što su affine ili projektivne transformacije. Ovi prikazi mogu aproksimirati kretanje rigidnih objekata na sceni. Za objekte čiji oblik nije rigidan, kontura objekta ili cela silueta predstavljaju najdeskriptivniji prikaz, pa se mogu koristiti kako parametarski, tako i neparametarski modeli za specifikaciju njihovog kretanja.

Primarni cilj razvoja svakog algoritma za praćenje jeste postizanje visoke preciznosti praćenja u različitim okolnostima na sceni, mogućnost dugotrajnog praćenja, kao i dovoljna brzina izvršavanja za rad u realnom vremenu. Na račun različitih matematičkih teorijskih postavki, vremenom su proistekle mnoge metode za video praćenje. Principi detekcije, regresije, segmentacije i algoritmi optimizacije su osnova različitih metoda za video praćenje. U narednim sekcijama ovog poglavlja će biti opisane glavne komponente koje u opštem slučaju predstavljaju osnovu algoritama za video

praćenje. Takođe, navedeni su izazovi koje treba prevazići kako bi algoritam za video praćenje uspešno pratio objekte od interesa. Kako bi se stekao sveobuhvatan uvid u algoritme za video praćenje, izloženi su različiti pristupi dizajniranja algoritama za praćenje, kao i napredak u njihovom razvoju.

### 3.1 Komponente algoritma za praćenje objekata u sekvenci slika

U algoritmima za praćenje pokretnih objekata u sekvenci slika može se izdvojiti pet glavnih komponenti, odnosno modula. To su: modul za izdvajanje relevantnih informacija sa slike, modul za pravilnu reprezentaciju objekta, modul za lokalizaciju objekta (modul za propagaciju stanja objekta tokom vremena), zatim modul za rukovođenje strategijom praćenja i modul za izdvajanje meta-podataka [1]. U nastavku ove sekcije opisani su pojedinačni moduli algoritama za video praćenje.

1. Prvi modul na putu toka signala u video praćenju jeste modul za izdvajanje relevantnih informacija sa dela slike na kome se nalazi objekat od interesa za praćenje. Karakteristike koje su značajne za detekciju objekata od interesa jesu one koje izdvajaju objekat iz pozadinske scene, kao i one koje objekat od interesa razdvajaju od drugih objekata. Funkcionalnost ovog modula se zasniva na karakteristikama koje se mogu podeliti u tri glavne kategorije: karakteristike niskog nivoa (boja, gradijenti, pokreti), srednjeg nivoa (ivice, uglovi, regioni) i karakteristike visokog nivoa (detekcije objekata generisane modelima dubokog učenja).
2. Naredni modul ima zadatak da obezbedi pravilnu reprezentaciju objekta. Ova reprezentacija predstavlja kodiranje izgleda i oblika objekta za praćenje. Definišu se karakteristike objekta, odnosno gradi se model objekta koji će algoritam koristiti za praćenje. Model za određeni objekat se može definisati na različite načine: *a priori*, može biti slika objekta ili se može naučiti iz obučavajućeg skupa uzoraka. Tako se za predstavu oblika objekta mogu koristiti pravougaoni okvir, elipsa, deformabilna kontura ili neka od predstava prikazana na slici 3.2. Informacije o izgledu mogu biti kodirane kao funkcija gustine verovatnoće određenih karakteristika izgleda izračunatih unutar oblasti objekta na slici. Takođe, model objekta treba i da obezbedi kompromis između tačnog opisa objekta i fleksibilnog prilagođenja promenama koje se dešavaju na objektu tokom praćenja. Dovoljno tačnim opisom objekta se izbegavaju problemi sa klaterom (više detalja je dato u sekciji 3.2.1), odnosno izbegava se praćenje objekata koji prema izgledu mogu biti jako slični objektu od interesa. S druge strane, fleksibilnost omogućava da se praćenje objekta održi neprekidnim u situacijama kada objekat menja svoju veličinu, ali i izgled usled različitih uticaja, što je detaljnije opisano u sekcijama 3.2.2 – 3.2.5.
3. Kao treći, definiše se modul za lokalizaciju objekta, odnosno propagaciju stanja objekta tokom vremena. Na osnovu informacija dobijenih iz prethodnih modula za izdvajanje obeležja i uspostavljenog modela objekta, koristi se iterativna procedura kako bi se odredila trajektorija kretanja objekta. Na taj način se ažurira stanje objekta i povezuju se instance istog objekta

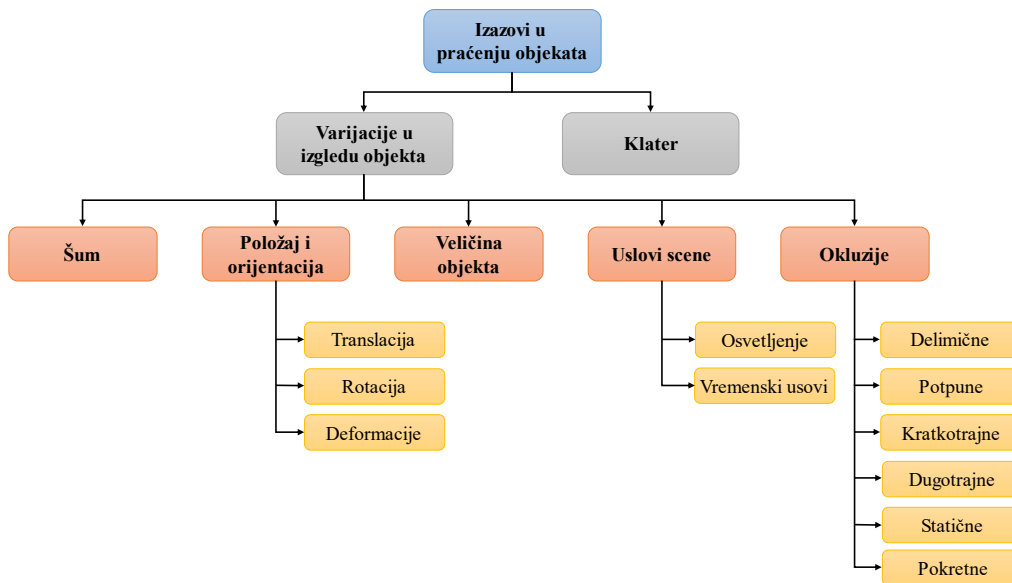
tokom vremena. Ovaj modul treba da obezbedi i kompenzaciju različitih efekata koji se mogu javiti tokom praćenja, a jedan od najizazovnijih predstavlja prevazilaženje okluzija.

4. Kreiranje i brisanje tragova objekata koji se prate se obavlja u modulu koji rukovodi strategijom praćenja. Ovaj modul treba da obezbedi inicijalizaciju praćenja kada se objekat od interesa detektuje na sceni, ali i da izvrši prekid praćenja kada taj objekat više nije uočljiv na sceni. Do inicijalizacije praćenja dolazi kada se objekat od interesa za praćenje pojavi na sceni, odnosno uđe u vidno polje kamere, što se odražava pojavom objekta na ivicama slike. Takođe, karakteristične situacije u kojima dolazi do inicijalizacije praćenja jesu scenariji kada se objekat nađe u definisanim regionima od interesa na slici, kada se objekat pojavi na slici nakon što je bio zaklonjen nekim drugim objektom ili kada objekat koji je u vidnom polju kamere sa velike distance prilazi samoj kameri, pa projekcija tog objekta na ravan slike dostigne dovoljno veliku površinu da se može prepoznati kao objekat od interesa za praćenje. Nasuprot situacijama u kojima dolazi do inicijalizacije praćenja, do prekida praćenja najčešće dolazi kada objekat koji se prati napusti vidno polje kamere, odnosno ponovo se nađe van ivica slike. Situacija kada je objekat predugo zaklonjen drugim objektom na sceni takođe može dovesti do prekida praćenja. Do prekida praćenja dolazi i kada se objekat unutar vidnog polja kamere isuviše udalji od kamere i više ne tumači kao objekat za praćenje. Štaviše, može se usvojiti i kriterijum koji govori o performansama praćenja. Stoga, kada performansa praćenja po definisanom kriterijumu padne ispod definisanog praga, doći će do prekida praćenja.
5. Kao poslednji modul definiše se modul za izdvajanje meta-podataka [1]. U zavisnosti od specifične aplikacije u kojoj se primenjuje video praćenje, u okviru ovog modula se izdvajaju željeni podaci određenog tipa i namene. Ovi podaci se koriste za opisivanje situacije u video sekvenci, razumevanje scene i prepoznavanje ponašanja pojedinih objekata na sceni. Takođe, dobijeni meta-podaci mogu pomoći u povećanju produktivnosti u aplikacijama gde se očekuje manuelni rad korisnika aplikacije koji je neophodan za izvršenje određenog zadatka.

## 3.2 Izazovi u praćenju objekata u sekvenci slika

U cilju obezbeđivanja dugotrajnog praćenja objekta važno je upoznati se sa izazovima koji se javljaju u video praćenju. Poznavanje ovih izazova može pomoći prilikom dizajniranja algoritma, posebno modula za pravilnu reprezentaciju objekta.

Osnovni zahtevi koje algoritam za video praćenje treba da zadovolji jesu mogućnost razlikovanja objekta od interesa za praćenje od pozadine, ali i mogućnost prilagođenja promenama koje se dešavaju na samom objektu usled različitih uticaja. Stoga, primarni izazovi koji se javljaju u video praćenju jesu sličnost predstave objekta koji se prati sa pozadinom i okolnim objektima (problem klatera (eng. *clutter*)), kao i promene u izgledu objekta tokom vremena. Na slici 3.3 je dat dijagram najvažnijih izazova koji se sreću u video praćenju. U nastavku su detaljnije opisani pojedinačni izazovi prikazani dijagramom na slici 3.3.



Slika 3.3 Izazovi u video praćenju.

### 3.2.1 Klater

Sličnost objekta od interesa za praćenje sa okolnim objektima na sceni predstavlja problem koji se naziva klater. Ukoliko se objekat od interesa nađe u okolini njemu sličnih objekata, može doći do zamene objekta koji se prati, pa algoritam nakon takve situacije može nastaviti da prati drugi objekat sličan objektu od interesa. Za prevenciju ovog problema, najvažnije je pravilno definisanje modela izgleda objekta koji će napraviti distinkciju između objekta od interesa i okolnih objekata i pozadine. Model izgleda objekta gradi se na osnovu vrednosti piksela dela slike na kom je inicijalno definisan objekat od interesa, na osnovu boje ili na osnovu specifičnih standardnih ili dubokih obeležja ekstrahovanih sa tog dela slike. Stoga, kako bi se što bolje opisao model izgleda objekta koji neće dovesti do zamene objekta u slučaju prisustva klatera, moguće je kombinovati više različitih informacija i obeležja objekta. Ovo s druge strane dovodi do povećanja kompleksnosti modela, a posledično i do sporijeg izvršavanja algoritma. Zato prilikom dizajna algoritma za praćenje treba postići kompromis između kompleksnosti modela kako bi se omogućio rad u realnom vremenu i dovoljne deskriptivnosti kako bi se prevazišao problem klatera. Na slici 3.4 je dat primer klatera, odnosno situacije gde se objekat od interesa za praćenje nalazi u okolini njemu sličnih objekata.



Slika 3.4 Problem klatera. Objekat koji se prati se nalazi u okolini njemu sličnih objekata.

### 3.2.2 Šum

Gotovo neizbežna pojava kojoj je izložen svaki signal koji se dobija sa senzora jeste prisutnost šuma. Šum na slici, odnosno u video sekvenci, zavisi prvenstveno od tehnologije u kojoj je senzor izrađen. Kao i kod ostalih tipova infracrvenih senzora, kod SWIR senzora je česta pojava mrtvih piksela – slika 3.5 a), što se može posmatrati i kao impulsni ili “so” šum. Takođe, u konkretnom slučaju, skup video sekvenci za analizu u ovoj disertaciji kreiran je kamerom koja je implementirana u tehnologiji isprepletenog prikaza, što kod objekata u pokretu dovodi do pojave iskrzanih ivica. Takođe, šum može biti uzrokovan i samom prirodom scene koja se snima. Štaviše, podrhtavanje kamere usled različitih poremećaja može dovesti do zamućenja na slici, što se takođe može posmatrati kao prisustvo šuma – slika 3.5 b). Šum prisutan na slici degradira kvalitet slike, a samim tim utiče na gubitak informacija iz slike. Može uticati na izgled objekta koji se prati, što posledično dovodi do smanjenja performansi algoritma za praćenje. Stoga je poželjno da model objekta bude invarijantan na prisustvo šuma. Kako bi se predupredio uticaj šuma na model objekta i performanse praćenja, u sistemima za video praćenje se velika pažnja posvećuje pretprocesiranju video sekvence, algoritmima za filtriranje šuma, popravci kvaliteta slike i video stabilizaciji [73] – [76].



Slika 3.5 Prisutnost šuma na SWIR slici: a) usled mrtvih piksela, b) usled podrhtavanja kamere.

### 3.2.3 Promena položaja i orijentacije

Prilikom praćenja pokretnih objekata, normalno je da objekat menja svoj položaj na sceni. Međutim, tokom kretanja, pored translatornog pomeranja, objekat može menjati i svoju orijentaciju. Promenom orijentacije, izgled objekta može biti potpuno drugačiji u odnosu na izgled koji je objekat imao prilikom inicijalizacije praćenja i uspostavljanja modela objekta. Takođe, sličan problem se javlja i kada objekat ima promene u položaju svojih delova. Ilustrativni primer je dat na slici 3.6. Kako u ovim situacijama ne bi došlo do prekida praćenja, model objekta treba da ima mogućnost prilagođavanja promenama koje se dešavaju na objektu tokom vremena. Ovo je zahtev oprečan zahtevu za prevazilaženje problema klatera opisanog u sekciji 3.2.1, pa se i u ovom slučaju prilikom dizajniranja algoritma za praćenje mora postići kompromis.





**Slika 3.6** Promena orijentacije objekta tokom praćenja.

### 3.2.4 Promena veličine objekata

Na performanse praćenja veliki uticaj ima i promena veličine objekta tokom vremena. Promena relativne veličine objekta u ravni slike je uzrokovana kretanjem objekta na sceni, udaljavanjem ili približavanjem kameri u odnosu na položaj koji je objekat imao u trenutku inicijalizacije praćenja. Ako model objekta ima fiksnu veličinu, izdvojene karakteristike će biti nepotpune ako se veličina objekta poveća. S druge strane, ako se veličina objekta smanji, promenljive pozadinske karakteristike će biti uvedene u model izgleda objekta i tako ga narušiti. Primer promene veličine objekta tokom praćenja je dat na slici 3.7.



**Slika 3.7** Promena veličine objekta tokom praćenja usled udaljavanja od kamere.

### 3.2.5 Promene u uslovima scene

U aplikacijama za rad na otvorenom veliki izazov predstavljaju promene u uslovima scene. Ovde se prvenstveno misli na promene u ambijentalnom osvetljenju i promene u vremenskim uslovima. Promene u osvetljenju na otvorenom prostoru se jako često dešavaju, a kako SWIR kamere snimaju reflektovanu svetlost, slično kao u vidljivom opsegu, ove promene mogu imati značajan uticaj na kvalitet praćenja. Kao promena u uslovima osvetljenja scene se može videti i situacija kada objekat tokom kretanja prelazi iz dela osvetljene scene u deo koji predstavlja senku nekog drugog

objekta – slika 3.8. Znajući da se promene u uslovima osvetljenja dešavaju često, obeležja objekta koja se izdvajaju i opisuju objekat treba da budu invarijantna na promene u osvetljenju.



**Slika 3.8** Promena u uslovima osvetljenja na sceni usled ulaska objekta koji se prati u senku drugog objekta.

### 3.2.6 Okluzije

Kao najveći izazov u video praćenju mogu se izdvojiti okluzije. Okluzija predstavlja situaciju u kojoj objekat koji se prati nije vidljiv na slici, iako se objekat nalazi na sceni u vidnom polju kamere. Izostanak vidljivosti ključnih delova objekata koji su važni za detekciju i praćenje takođe predstavlja okluziju. Stoga, mogu se razlikovati delimične i potpune okluzije. Delimične okluzije se dešavaju kada samo pojedini delovi objekta koji se prati nisu vidljivi, dok potpuna okluzija predstavlja izostanak celog objekta. Uzroci okluzija mogu biti različiti. Javljaju se kada objekat koji se prati bude zaklonjen strukturom pozadinske scene ili nekim drugim objektom koji se kreće na sceni. Odatle možemo razlikovati statične i pokretne okluzije. Takođe, okluzija se dešava i kada jedan deo objekta zakloni neki drugi deo istog objekta, što predstavlja samookluziju. Ovakva situacija se najčešće dešava kod praćenja zglobnih objekata. Poseban je izazov kada se okluzija dogodi usled zaklanjanja objekta od interesa drugim objektom iste klase. Navedeni scenario predstavlja kombinaciju problema klatera i okluzija. Ilustrativni primeri situacija u kojima se dešavaju okluzije su prikazani na slici 3.9.



a)

b)



c)



d)

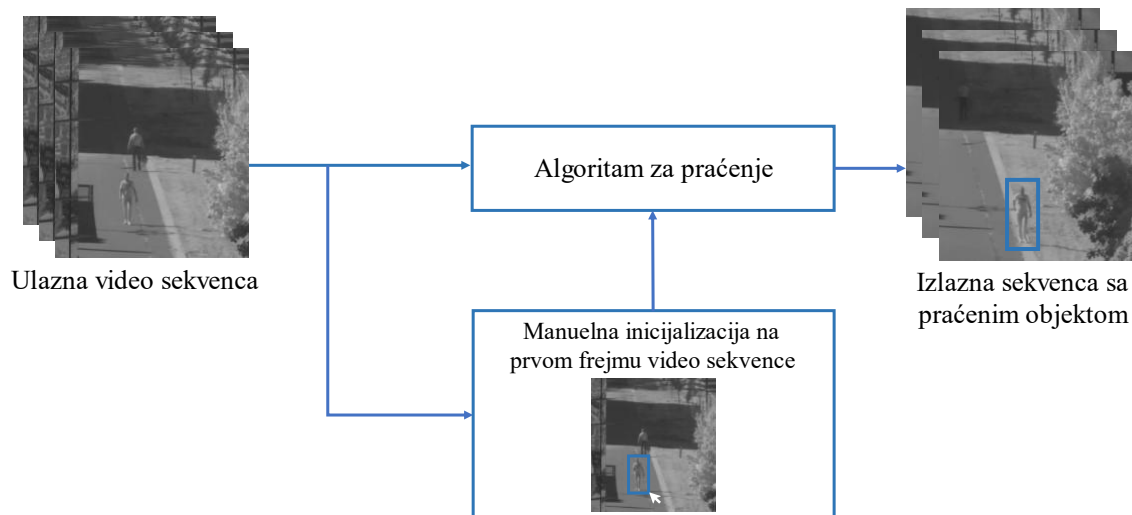
**Slika 3.9** Problem okluzija. Objekat koji se prati a) je zaklonjen statičnom okluzijom b); Objekat koji se prati c) biva delimično zaklonjen pokretnom okluzijom d).

### 3.3 Metode inicijalizacije praćenja objekata

U zavisnosti od potrebe za interakcijom sa aplikacijom za video praćenje, a posebno prilikom inicijalizacije algoritma za praćenje, mogu se razlikovati manuelna i automatska inicijalizacija. Sam proces inicijalizacije praćenja je nezavisan od algoritma koji vrši praćenje objekta za koji je inicijalizovan i ne utiče na dizajn algoritma za praćenje. Međutim, način na koji je algoritam inicijalizovan može imati uticaja na performanse praćenja. To prvenstveno zavisi od toga koliko su precizno definisane granice objekta koji treba pratiti.

#### 3.3.1 Manuelna inicijalizacija

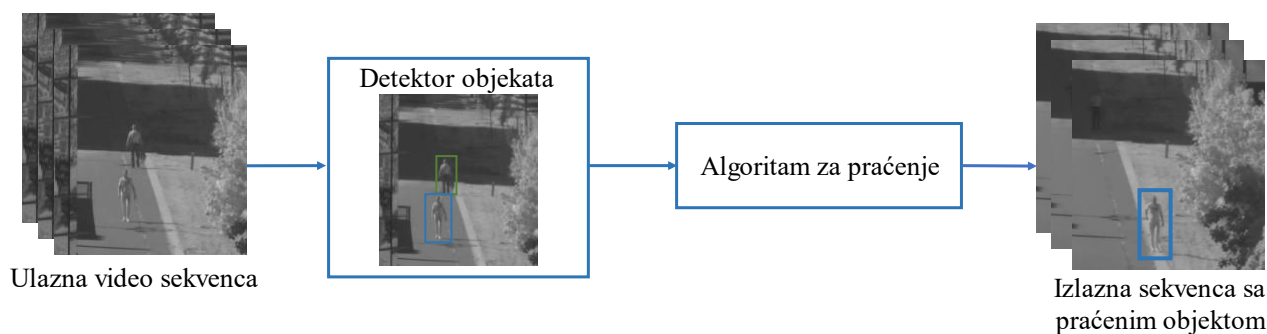
Kada se inicijalizacija praćenja vrši manuelno govori se o interaktivnom ili poluautomatskom pristupu u video praćenju [1]. Predstavlja prelazni stupanj ka potpuno automatskom video praćenju. U ovom slučaju se očekuje da informaciju o položaju i dimenzijama objekta za praćenje na slici obezbedi sam korisnik aplikacije. Prednost ovog tipa inicijalizacije jeste što korisnik manuelno može veoma precizno da odredi granice objekta. Međutim, veoma veliki izazov može biti u slučaju potrebe za preciznom inicijalizacijom praćenja objekata koji su u pokretu, a posebno ako se veoma brzo kreću kroz vidno polje kamere. Ostale situacije u kojima korisnik može imati potrebu za interakcijom sa aplikacijom jesu u toku procesa ocene rada algoritma ili u slučaju eventualnih korekcija praćenja. Šematski prikaz procedure manuelne inicijalizacije video praćenja dat je na slici 3.10.



Slika 3.10 Manuelna inicijalizacija praćenja objekta.

### 3.3.2 Automatska inicijalizacija

Automatska inicijalizacija praćenja podrazumeva potpuno odsustvo aktivnosti od strane korisnika. Omogućena je postojanjem nekog tipa detektora objekata. Iako predstavlja nezavisan segment od algoritma praćenja, kada se detektoru objekata prepusti autonomija pri inicijalizaciji praćenja, on zapravo postaje sastavni deo algoritma za praćenje. Pored same inicijalizacije, informacije o objektu dobijene iz detektora se mogu iskoristiti za poboljšanje estimacije stanja objekta tokom vremena. Ovaj pristup se široko koristi u algoritmima za praćenje više objekata, u takozvanoj paradigmi praćenja na osnovu detekcija gde se algoritam za praćenje u velikoj meri oslanja na detektor objekata od interesa [22], [69] – [71]. Kao što je već naznačeno, inicijalizacija praćenja objekata koji su u pokretu, u trenutku inicijalizacije, može biti veoma izazovna ako se inicijalizacija vrši manuelno. Primenom detektora objekata u pokretu, ovaj problem se može prevazići. U slučaju da je od interesa automatska inicijalizacija praćenja objekata određene klase, potrebni su modeli (najčešće dubokog učenja) posebno obučeni za detekciju objekata željene klase. Tako detektovani objekti se zatim prosleđuju algoritmu za praćenje koji ih dalje prati tokom vremena. Šematski prikaz procedure automatske inicijalizacije video praćenja dat je na slici 3.11.



Slika 3.11 Automatska inicijalizacija praćenja primenom detektora objekata od interesa.

### 3.3.2.1 Detekcija pokreta

Inicijalizacija praćenja objekata koji se veoma brzo kreću kroz vidno polje kamere može biti zadatak sa kojim i veoma iskusni operateri aplikacija video praćenja mogu imati poteškoća. To rezultira neuspešnim aktiviranjem procesa video praćenja ili praćenjem nedovoljno precizno označenog objekta i samim tim lošim performansama praćenja. Da bi se prevazišao problem inicijalizacije praćenja pokretnih objekata i postigao određeni stepen automatizacije ove procedure, potrebno je implementirati algoritam za detekciju pokretnih objekata i inicijalizaciju praćenja. Štaviše, algoritam za detekciju pokretnih objekata može biti koristan za otkrivanje objekata koji ulaze na scenu koja se snima i time predstavljati suštinski korak u analizi scene.

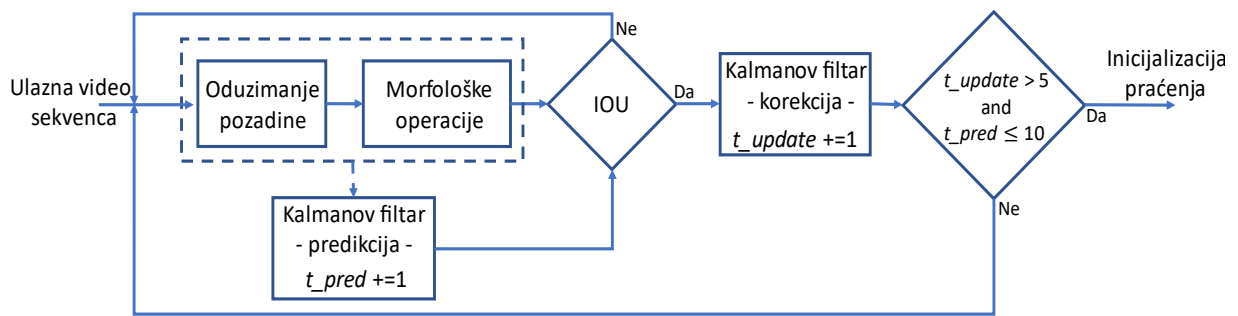
Protetkih godina razvijeni su mnogi pristupi za detekciju pokreta u video sekvenci [77]. Neki od glavnih pravaca u detekciji pokreta su zasnovani na oduzimanju pozadine [78], [79], optičkom toku (eng. *optical flow*) [80] i pristupima zasnovanim na dubokom učenju [81].

Kako je cilj da se blagovremeno otkrije pokretni objekat i aktivira praćenje tog objekta, metoda detekcije pokreta treba da ima malo vreme procesiranja. Iz tog razloga, za mogućnost rada u realnom vremenu, jedna od metoda na kojoj može biti bazirana detekcija pokreta u algoritmu automatske inicijalizacije praćenja jeste oduzimanje pozadine odnosno segmentacija pozadine i objekata u prednjem planu zasnovana na adaptivnom miksu Gausovih komponenti (*Mixture of Gaussians - MOG*) [79].

Rezultat primene *MOG* algoritma za oduzimanje pozadine je binarna slika u kojoj su pokretni objekti predstavljeni belim pikselima dok je pozadina predstavljena crnim. Beli pikseli na ovoj slici su rezultat kretanja objekata na sceni, ali i uticaja šuma i podrhtavanja kamere. Takođe, čest je slučaj da jedan pokretan objekat bude predstavljen sa nekoliko razdvojenih klastera belih piksela. Da bi se prevazišli problemi lažnih pokreta i detekcija pokreta istog objekta koje nisu unitarne, nad dobijenom binarnom slikom se primenjuju morfološke operacije. Postupak morfološkog otvaranja (erozija praćena dilatacijom) primenjuje se za uklanjanje malih grupa belih piksela koje predstavljaju lažne pokrete. Zatim se primenjuje morfološko zatvaranje (dilatacija praćena erozijom) za popunjavanje malih praznina između razdvojenih segmenata belih piksela koji predstavljaju kretanje istog objekta. Na ovaj način se u velikoj meri eliminiše problem lažnih detekcija malih pokreta, a postiže se i da jedan pokretni objekat bude predstavljen jednim regionom belih piksela. Nakon primenjenih morfoloških operacija, potrebno je definisati zatvorene konture belih piksela, a zatim pravougaone okružujuće okvire oko dobijenih zatvorenih kontura [82]. Za dobijene okružujuće pravougaone okvire oko detektovanih pokretnih objekata primenjuje se poređenje njihove površine sa definisanim pragom. U skupu pokretnih objekata treba zadržati samo one čija je površina veća od definisanog praga kako bi se eliminisali veoma mali objekti koje nije moguće pratiti, odnosno detektovani pokreti koji su zapravo posledica prisustva šuma na slici.

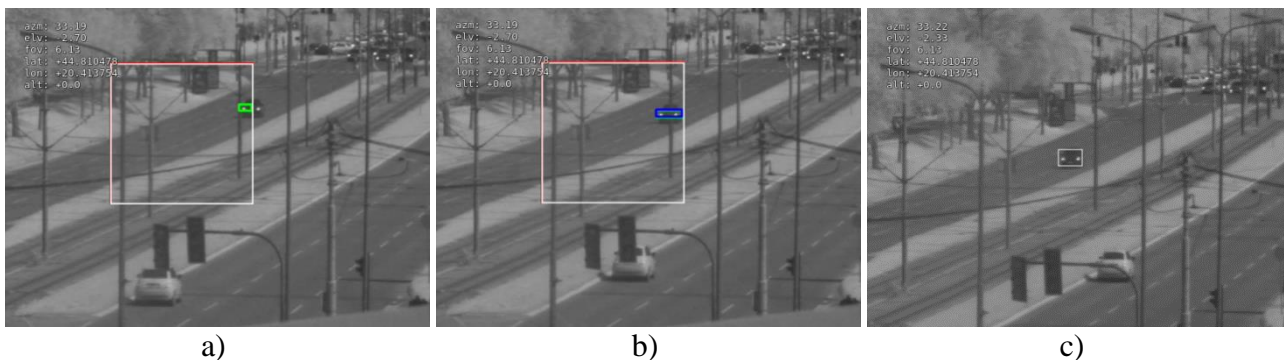
Kako bi se dodatno smanjila verovatnoća aktivacije praćenja lažne detekcije pokretnog objekta, primenjuje se takozvano privremeno praćenje detektovanog kretanja [82]. Za detekciju pokretnog objekta koji se pojavi na sceni se kreira instanca standardnog Kalmanovog filtra, zadužena za praćenje date detekcije kroz naredne frejmove video sekvence i inicijalizuje se brojač perioda postojanja objekta na sceni ( $t_{pred}$ ). Model u Kalmanovom filtru koji opisuje stanje objekta kao elemente u vektoru stanja sadrži poziciju i dimenzije objekta u ravni slike, kao i odgovarajuće brzine. Više detalja o standardnom Kalmanovom filtru se može naći u sekciji 5.1. Blok dijagram koji opisuje metod za inicijalizaciju praćenja na osnovu detekcije pokreta u video sekvenci dat je na slici 3.12.





**Slika 3.12** Blok dijagram metode za inicijalizaciju praćenja na osnovu detekcije pokreta.

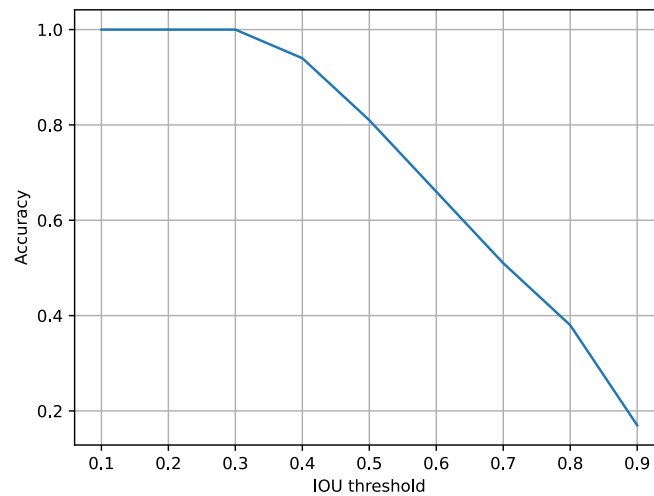
Vodeći se pristupom *tracking-by-detection* [69], okružujući okviri oko novootkrivenih pokreta u narednom frejmu video sekvence i predikcija pozicije i veličine objekta iz Kalmanovog filtra se zatim prosleđuju bloku za izračunavanje IOU metrike između dobijenih okružujućih okvira. IOU metrika zapravo određuje da li je novootkrivena detekcija pokreta na slici uzrokovana privremeno praćenim objektom. Ako preklapanje prema IOU metrici između detekcije pokreta i predikcije Kalmanovog filtra premašuje prag od 0.5, može se smatrati da je novootkriveni pokret zaista uzrokovao objekat koji se privremeno prati. Tada se stanje privremeno praćenog objekta u Kalmanovom filteru može ažurirati, a brojač ažuriranja ( $t\_update$ ) se inkrementira. Ako se stanje objekta ažurira u 5 frejmova tokom privremenog praćenja, algoritam za praćenje će se automatski inicijalizovati za detektovani pokretni objekat. Ako ovaj uslov nije ispunjen, procedura se nastavlja sve dok se ne otkrije pravi pokretni objekat ili se proces ponovo pokreće ako se praćenje ne inicijalizuje nakon prvih 10 frejmova privremenog praćenja [82]. Primer detekcije automobila u pokretu u SWIR video sekvenci i automatske inicijalizacije praćenja dat je na slici 3.13.



**Slika 3.13** a) Detekcija pokreta u regionu od interesa, b) privremeno praćenje detektovanog pokreta, c) inicijalizovano praćenje detektovanog pokretnog objekta.

Na slici 3.14 su prikazane performanse opisane metode za inicijalizaciju praćenja pokretnog objekta dobijene za 100 inicijalizacija praćenja različitih objekata iz kreiranog skupa SWIR video sekvenci. Performanse su izražene tačnošću za različite vrednosti IOU metrike između estimiranog okružujućeg okvira pokretnog objekta i stvarnog (*ground-truth*) okružujućeg okvira na frejmu koji predstavlja trenutak inicijalizacije praćenja. Tačnost je predstavljena kao procenat inicijalizacija praćenja za koje je IOU metrika između estimiranog i stvarnog okružujućeg okvira veća od datog praga. Rezultati sa slike 3.14 pokazuju da se pozicija pokretnog objekta veoma dobro estimira u trenutku inicijalizacije praćenja. Na tačnost najviše uticaja ima estimacija dimenzija okružujućeg okvira, što je posebno izazovno usled mogućeg prisustva delimičnih i potpunih okluzija u trenutku

inicijalizacije praćenja, kao i usled prisustva senke objekta koja ima istu dinamiku kretanja u video sekvenci kao i sam objekat.



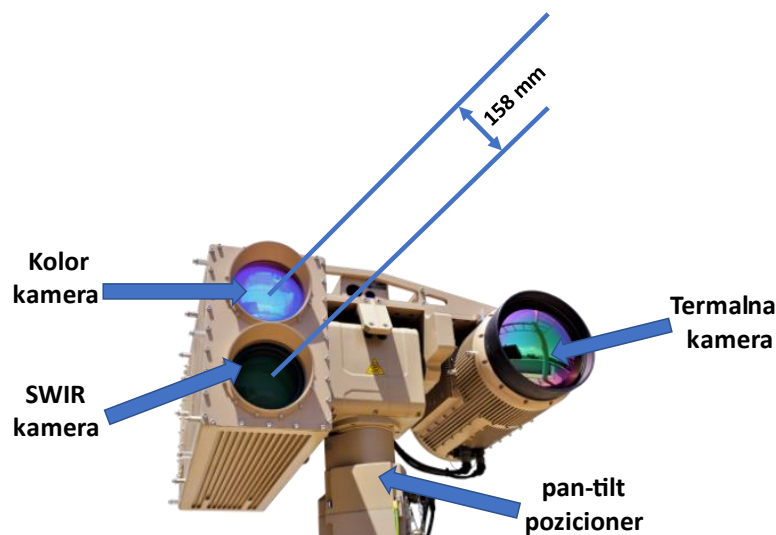
**Slika 3.14** Tačnost inicijalizacije praćenja objekta na osnovu detekcije pokreta za različite IOU vrednosti.

### 3.3.2.2 Detekcija objekata na SWIR slici

Za inicijalizaciju praćenja određenog tipa objekata, potreban je model koji je posebno obučen za detekciju željene klase objekata. Za detekciju objekata na slikama najčešće se koriste modeli dubokog učenja zbog svojih dobrih performansi. Međutim, u dostupnoj literaturi u oblasti detekcije objekata uočava se nedostatak modela dubokog učenja visokih performansi za detekciju objekata na SWIR slikama.

Osnovni preduslov za obučavanje modela dubokog učenja jeste posedovanje adekvatnog obučavajućeg skupa koji treba da obezbedi dovoljnu generalizaciju modela za detekciju objekata u različitim situacijama koje se mogu javiti u realnim aplikacijama [83], [84]. Kvalitetan obučavajući skup treba da sadrži veliki broj slika sa objektima od interesa, koje treba snimiti iz različitih uglova, na različitim udaljenostima od kamere, na različitim pozadinama, u različitim vremenskim uslovima i uslovima osvetljenja. Najčešći slučaj kreiranja obučavajućeg skupa podataka za određeni problem je kombinovanje postojećih javno dostupnih skupova podataka sa slikama snimljenim u uslovima i scenarijima koji najbolje odgovaraju svrsi aplikacije. Međutim, ne mogu se naći veliki, javno dostupni skupovi SWIR slika sa objektima iz urbanih scenarija kakvi su obuhvaćeni u kreiranoj bazi za praćenje opisanoj u sekciji 2.3. Stoga se i za potrebe obučavanja modela za detekciju objekata na SWIR slici mora kreirati novi skup podataka za obučavanje.

Manuelno labeliranje velikog skupa podataka je zamoran i dugotrajan posao. U cilju izbegavanja manualnog labeliranja velikog skupa podataka SWIR slika za obučavanje detektora objekata na SWIR slici, mogu se iskoristiti prednosti transfera učenja i multisenzorskih sistema koji poseduju SWIR i kolor kamere [85]. Ovaj postupak je poznat kao kros-spektralna anotacija [86], [87]. Cilj je detekciju objekata izvršiti na slici snimljenoj kolor kamerom dobro obučanim modelima, a zatim dobijene detekcije mapirati na SWIR sliku iste scene snimljene u istom trenutku. Sistem koji poseduje SWIR i kolor kameru i obezbeđuje paralelno snimanje iste scene obema kamerama prikazan je na slici 3.15.



**Slika 3.15** Sistem na pan-tilt pozicioneru koji obezbeđuje snimanje iste scene SWIR, kolor i termalnom kamerom [58].

U cilju obezbeđivanja slika iste scene obema kamerama, optičke ose kamera u multisenzorskom sistemu moraju biti poravnate. Ovo se postiže *boresight* kalibracionom procedurom [88] multisenzorskog sistema u elektrooptičkoj laboratoriji. Postupak je prikazan na slici 3.16.



**Slika 3.16** Kalibraciona procedura multisenzorskog sistema u elektrooptičkoj laboratoriji [87].

Iako je vidno polje kamera podesivo, vidna polja SWIR i kolor kamere se ne mogu podesiti na potpuno iste vrednosti. Jedna kamera može imati širi vertikalni, dok druga može imati širi horizontalni ugao vidnog polja. Međutim, mapiranje detekcija sa kolor kamere na sliku sa SWIR kamere moguće je samo u zoni preklapanja vidnih polja SWIR i kolor kamere. Stoga, da bi se kreirao skup podataka visokog kvaliteta, bez propuštenih detekcija, slike obe kamere moraju biti izrezane u zoni preklapanja vidnih polja [87].

Uprkos zadržavanju samo dela slike u zoni preklapanja, određena izobličenja između kolor i SWIR slika mogu biti i dalje prisutna. Kako bi se problem izobličenja prevazišao i detekcije objekata sa kolor slike precizno mapirale na SWIR sliku u zoni preklapanja, potrebno je uspostaviti geometrijsku relaciju između SWIR i kolor slike. Prvi korak u celokupnoj proceduri je izračunavanje



matrice transformacije između slika, odnosno matrice homografije, za šta su potrebni odgovarajući parovi tačaka na SWIR i kolor slici. Određivanje parova tačaka na SWIR i kolor slici nekom automatizovanom procedurom na osnovu obeležja može biti veoma nepouzđano usled različite prirode nastanka slika u spektralnim opsezima u kojima kamere rade. Stoga, matrica homografije se može odrediti na osnovu ručno odabranih odgovarajućih parova tačaka na obe slike (što je potrebno odraditi samo na prvom frejmu SWIR i kolor video sekvence).

Za ove potrebe, dobro se pokazala nereflektivna transformacija sličnosti (eng. *non-reflective similarity*) [89]. Ova transformacija se koristi u slučaju kada su oblici na odgovarajućim slikama nepromenjeni, ali je prisutna distorzija nekom kombinacijom translacije, rotacije i skaliranja. Transformacija ima četiri stepena slobode i potrebna su joj minimalno dva para odgovarajućih tačaka. Podržava rotaciju, translaciju i izotropno skaliranje [89]. Nereflektivna transformacija sličnosti izražava se matricom  $H$ :

$$H = \begin{bmatrix} h_1 & -h_2 \\ h_2 & h_1 \\ h_3 & h_4 \end{bmatrix} \quad (3.6)$$

Projekcija  $(\hat{x}, \hat{y})$  tačke  $(x, y)$  u ravni slike matricom  $H$  je data sa:

$$[\hat{x} \ \hat{y}] = [x \ y \ 1] H \quad (3.7)$$

Nakon nereflektivne transformacije sličnosti, prave linije ostaju prave, a paralelne ostaju paralelne.

Nakon što se izračuna matrica homografije, vrši se mapiranje detekcija sa kolor slike na transformisanu SWIR sliku. Pošto su za obučavanje modela za detekciju objekata potrebne slike u originalnom domenu (netransformisane), detekcije moraju biti konvertovane u originalni domen inverznom transformacijom. Prednost ovog pristupa je u tome što ne zahteva transformaciju celokupne kolor i SWIR slike, što je računski skup proces, već se samo koordinate detektovanih objekata na kolor slici konvertuju između referentnih sistema [87].

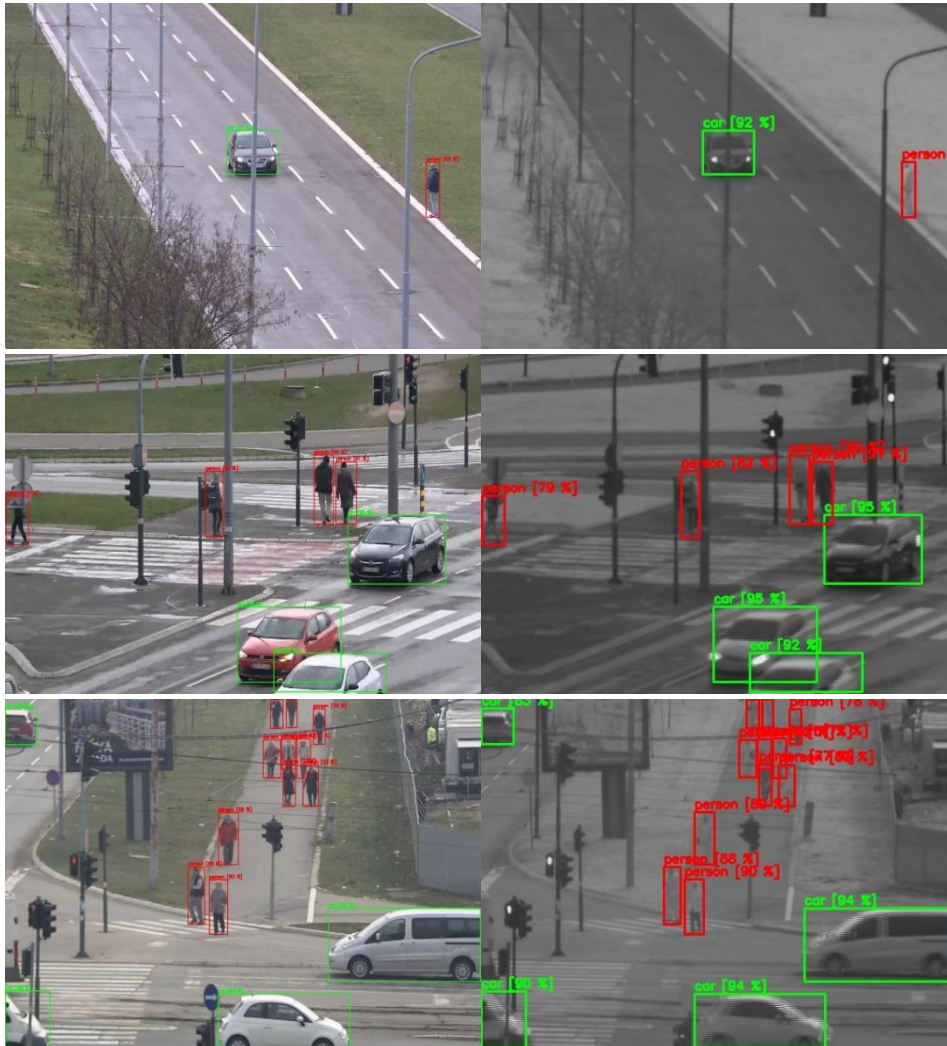
Na sceni koja se snima za kreiranje obučavajućeg skupa može postojati mnogo pokretnih objekata. Opisana homografska transformacija je nezavisna od pokretnih objekata na sceni. Iako je izvršena precizna transformacija između SWIR i kolor slike, ona ne garantuje tačno mapiranje detekcija pokretnih objekata ako kamere nisu pravilno sinhronizovane. SWIR kamera radi sa frekvencijom  $f_s$  (frejmova po sekundi), dok kolor kamera radi sa frekvencijom  $f_c$ . Da bi se izvršilo mapiranje detekcija pokretnih objekata između odgovarajućih parova slika u video sekvenci, potrebna je precizna sinhronizacija kamera. Pošto je korelacija između uzastopnih frejmova u video sekvenci veoma velika, ne mora se svaki frejm iz video sekvence koristiti za kreiranje skupa podataka. Detekcija objekata na kolor slici i mapiranje detekcija na SWIR sliku mogu se primeniti između parova slika definisanih sa:

$$N_s = N_c \frac{f_s}{f_c} \quad (3.8)$$

gde je  $N_s$  redni broj frejma u SWIR video sekvenci, dok je  $N_c$  redni broj frejma u video sekvenci kolor kamere.

Za kreiranje modela za detekciju objekata na SWIR slici iskorišćeno je nekoliko varijanti modela dubokog učenja YOLOX [90], [87]. Za detekciju objekata na kolor slici u okviru metode za kros-spektralnu anotaciju iskorišćen je model YOLOX-x [90] sa rezolucijom ulazne slike od 640×640

piksela, a prag za prihvatanje detekcija za mapiranje je postavljen na 0.3. SWIR slike izrezane u zoni preklapanja vidnih polja SWIR i kolor kamere ulaze u obučavajući skup SWIR slika. Skup SWIR slika kreiran predloženom metodom kros-spektralne anotacije je podeljen na obučavajući deo i deo za validaciju, gde je u obučavajući skup uvršćeno 10.023 slika, a u skup za validaciju 717. Kreirani skupovi SWIR slika sadrže slike urbanih scena, a kao objekti od interesa za detekciju su odabrani pešaci i automobili. Prikaz mapiranja detekcija sa kolor na SWIR slike u okviru kreiranja obučavajućeg skupa SWIR slika metodologijom kros-spektralne anotacije prikazan je na slici 3.17.



**Slika 3.17** Mapiranje detekcija objekata od interesa (pešaci i automobili) sa kolor na SWIR slike metodologijom kros-spektralne anotacije.

Za obučavanje YOLOX modela za detekciju objekata na SWIR slikama, radi pravilnog poređenja, praćena je postavka parametara predložena u [90]. Obučavanje je rađeno kroz 300 epoha, sa 5 *warmup* epoha, stohastičkim gradijentnim spustom sa početnom stopom učenja od 0.01, i kosinusnom adaptacijom stope učenja (*cosine learning rate schedule*), momentum 0.9 i *weight decay* 0.0005. Tokom obučavanja YOLOX-x modela korišćeni su *mix up* i *mosaic* augmentacija podataka sa opsegom skale [0.1, 2.0], dok je za obuku manjih modela YOLOX-Nano i YOLOX-Tiny izostavljena *mix up* augmentacija, a opseg skaliranja *mosaic* augmentacije postavljen je na [0.5, 1.5]. Veličina ulazne slike je 640×640. Obučavanje je izvršeno na računaru sa i7-8700K 3.70 GHz CPU-om, 32 GB RAM-a i NVIDIA GeForce RTX 2080 GPU-om.

Za testiranje obučениh YOLOX modela kreiran je skup od 715 SWIR slika i ručno labeliran za klase objekata pešaka i automobila. Prema pretraživanju referentne literature i poznatih baza podataka, javno dostupni skupovi labeliranih SWIR slika pešaka i automobila, na kojima se modeli za detekciju objekata mogu testirati i upoređivati, se ne mogu pronaći. U tabeli 3.1 je predstavljeno poređenje rezultata detekcije originalnih YOLOX modela (obučениh na COCO skupu kolor slika [91]) i YOLOX modela obučениh na automatski kreiranom skupu SWIR slika metodom kros-spektralne anotacije. Evaluacija modela je izvršena na kreiranom test skupu SWIR slika primenom biblioteke *COCO\_eval* [92]. Modeli su evaluirani i prema veličini objekata na slici. Objekti male veličine predstavljaju objekte do 32×32 (1024) piksela, srednji: od 32×32 (1024) do 96×96 (9216) piksela, a veliki objekti su oni čija je veličina veća od 96×96 (9216) piksela [92]. Performanse su predstavljene prosečnom preciznošću (*Average Precision* - AP) 0.5:0.95. Rezultati jasno ukazuju na značajno poboljšanje performansi modela obučenog na SWIR skupu slika. Radi korektnog poređenja, kako bi se prikazao samo uticaj automatski kreiranog skupa SWIR slika kros-spektralnom metodologijom na performanse modela, svi modeli su obučeni pod istim uslovima kao i originalni. Jedini slučaj smanjenih performansi se vidi u detekciji velikih objekata (veličine veće od 96 × 96 (9216) piksela) modelom YOLOX-x. Razlog za to se može videti u procentu velikih objekata u skupu podataka za obuku. Kreirani set za obučavanje predloženom metodologijom sadrži 58.605 objekata, od čega je 5.94% velikih, 59.95% srednjih i 34.11% malih. Sveukupni pogled na rezultate detekcije ukazuje da automatski generisani skup podataka kros-spektralnom metodologijom značajno doprinosi poboljšanju performansi detekcije objekata na SWIR slikama [87].

**Tabela 3.1** Performanse originalnog YOLOX modela i modela obučavanog na SWIR skupu slika kreiranog metodologijom kros-spektralne anotacije.

AP [%]		Sve dimenzije		Mali objekti		Objekti srednje veličine		Veliki objekti	
		Osnovni model	Obučen na SWIR-u	Osnovni model	Obučen na SWIR-u	Osnovni model	Obučen na SWIR-u	Osnovni model	Obučen na SWIR-u
YOLOX Nano	Svi objekti	19.4	<b>58.9</b>	8.9	<b>45.7</b>	25.0	<b>64.8</b>	35.1	<b>68.2</b>
	Pešaci	17.5	<b>53.7</b>	5.2	<b>36.6</b>	22.9	<b>60.2</b>	34.0	<b>58.9</b>
	Automobili	21.2	<b>64.1</b>	12.7	<b>54.8</b>	27.2	<b>69.4</b>	36.3	<b>77.4</b>
YOLOX Tiny	Svi objekti	28.4	<b>62.3</b>	18.5	<b>52.2</b>	33.8	<b>67.0</b>	50.1	<b>73.1</b>
	Pešaci	25.9	<b>57.6</b>	17.1	<b>45.1</b>	29.6	<b>62.3</b>	50.0	<b>68.5</b>
	Automobili	30.9	<b>67.1</b>	19.9	<b>59.3</b>	38.1	<b>71.8</b>	50.2	<b>77.6</b>
YOLOX x	Svi objekti	55.9	<b>66.2</b>	51.0	<b>57.3</b>	58.1	<b>70.3</b>	<b>86.5</b>	67.6
	Pešaci	48.6	<b>61.8</b>	43.6	<b>50.6</b>	50.0	<b>66.2</b>	<b>93.3</b>	58.6
	Automobili	63.2	<b>70.5</b>	58.5	<b>64.1</b>	65.7	<b>74.5</b>	<b>79.7</b>	76.7

Dobijeni detektor objekata na SWIR slikama se dalje efikasno može koristiti za inicijalizaciju praćenja određenog tipa objekata. Predložena metodologija kros-spektralne anotacije se može primeniti za kreiranje skupova podataka bilo koje klase objekata. Štaviše, predložena metodologija se može primeniti i za obučavanje bilo kog drugog modela dubokog učenja za detekciju objekata. Takođe, kombinacijom detektora objekata na SWIR slikama i predloženog algoritma za detekciju pokreta u video sekvenci može se dobiti metod za detekciju pokretnog objekta od interesa [93]. Na taj način se praćenje može automatski inicijalizovati samo za pokretni objekat definisane klase.

### 3.4 Klasifikacija pristupa za praćenje objekata u sekvenci slika

Algoritmi za praćenje objekata mogu se klasifikovati i analizirati na nekoliko načina. Prvo, u zavisnosti od broja kamera koje se koriste za praćenje objekata, razlikuju se monokularno i stereo praćenje. Monokularno praćenje podrazumeva korišćenje samo jedne kamere, pa se algoritmi za praćenje dizajniraju za rad sa informacijama u dvodimenzionalnom prostoru. S druge strane, u stereo praćenju se koriste dve ili više kamera postavljene na određenom rastojanju, što pruža trodimenzionalne (3D) informacije o sceni i objektima, uključujući dubinu scene i poziciju objekata u 3D prostoru. Informacije o dubini scene i poziciji objekata u 3D prostoru mogu biti veoma korisne i povećati performanse praćenja, ali se istovremeno i računaska obrada značajno usložnjava, što utiče na mogućnost rada algoritama u realnom vremenu. Od interesa za istraživanje u ovoj disertaciji je monokularno praćenje.

Prema vremenskoj skali izvršavanja, mogu se razlikovati onlajn (eng. *online*) i oflajn (eng. *offline*) algoritmi. Oflajn algoritmi se primenjuju za praćenje objekata u video sekvenci nakon što je snimanje završeno. S druge strane, onlajn praćenje podrazumeva praćenje objekata u realnom vremenu, dok se video sekvenca još uvek snima. Dok oflajn algoritmi imaju pristup svim frejmovima od početka do kraja video sekvence, onlajn algoritmi nemaju pristup budućim frejmovima, pa se mogu oslanjati samo na informacije sa trenutnog i prethodnih frejmova. Budući da se kod oflajn algoritama ne zahteva rad u realnom vremenu, oni mogu koristiti više vremena za kompleksniju obradu i detaljniju analizu, za razliku od onlajn algoritama gde je brzina i efikasnost od suštinskog značaja za mogućnost rada u realnom vremenu. Kod onlajn algoritama, pod radom u realnom vremenu se podrazumeva da se obrada trenutnog frejma, odnosno praćenje na trenutnom frejmu video sekvence, završi u vremenskom periodu pre pristizanja narednog frejma sa senzora. Iz ovih razlika proističu i različite primene algoritama za praćenje. Oflajn algoritmi nalaze primenu u filmskoj i video produkciji, u analizi medicinskih video sekvenci, forenzici, sportskoj analizi, dok se onlajn algoritmi koriste u aplikacijama gde je potrebna brza reakcija na dešavanja na sceni, kao što su video nadzor, autonomna vožnja, robotika, kao i mnoge druge. Predmet istraživanja ove disertacije je onlajn video praćenje za rad u realnom vremenu.

Na osnovu karakteristika koje se koriste u modelu praćenja kod monokularnih onlajn algoritama, mogu se razlikovati tradicionalni algoritmi koji koriste ručno definisane (eng. *hand-crafted*) karakteristike i algoritmi zasnovani na dubokim obeležjima. Pristupi zasnovani na ručno definisanim karakteristikama izdvajaju obeležja sa slika prema definisanom algoritmu baziranom na ekspertskom znanju. S druge strane, algoritmi zasnovani na dubokim obeležjima izdvajaju semantičke oznake iz slika korišćenjem najčešće konvolucionih neuralnih mreža. U poslednje vreme, u praćenju objekata u vidljivom opsegu, kao i u praćenju objekata u termalnim infracrvenim video sekvencama, metode zasnovane na dubokom učenju stekle su posebnu pažnju. Zbog mogućnosti hijerarhijskog učenja, pristupi zasnovani na dubokim obeležjima značajno nadmašuju algoritme koji su zasnovani na ručno definisanim karakteristikama [94]. Međutim, u literaturi je primena metoda dubokog učenja u SWIR domenu dominantno usmerena na detekciju objekata [17], [87]. Algoritmi zasnovani na dubokom učenju koji se odnose na praćenje objekata u SWIR video sekvencama primenjeni su u [22] za praćenje više objekata po principu *tracking-by-detection*, a u [18] – [21] su primenjene konvolucione neuralne mreže za praćenje više objekata u sekvenci degradiranih SWIR slika, sa značajnim stepenom kontaminacije slika šumom i prisustvom loših piksela. Međutim,

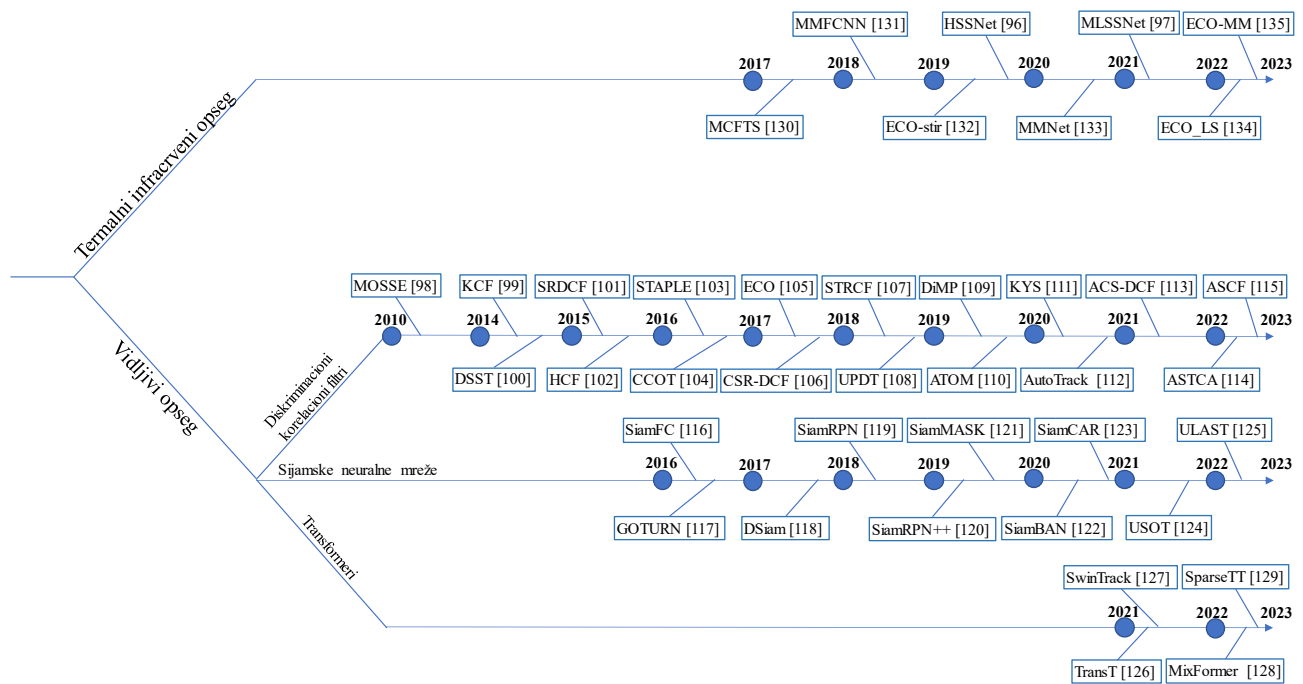
praćenje više objekata, a posebno praćenje objekata u sekvenci degradiranih SWIR slika je van okvira ove disertacije. Posmatrano u aplikacijama u vidljivom i termalnom infracrvenom opsegu, metode za praćenje objekata zasnovane na dubokom učenju postižu izuzetno dobre performanse, ali sa potrebom za velikim skupom podataka za obučavanje. Međutim, kao što je već komentarisano u sekciji 2.3, javno dostupni veliki skupovi SWIR slika kreirani za obučavanje modela za praćenje pokretnih objekata se ne mogu naći. Štaviše, za modele dubokog učenja se generalno može reći da imaju nepoznato ponašanje u scenarijima za koje nisu obučeni, što ukazuje na opšti nedostatak pristupa zasnovanih na dubokom učenju.

U zavisnosti od toga kako algoritmi razlikuju objekat od interesa za praćenje od okruženja, mogu se razlikovati generativni i diskriminativni algoritmi. Generativni algoritmi praćenje objekata tretiraju kao problem gde se traži region slike koji najbolje odgovara referentnom modelu objekta, koristeći samo informacije o objektu. Kod diskriminativnih algoritama cilj je da se napravi razlika između objekta i pozadine, koristeći informacije o objektu ali i njegovoj neposrednoj pozadini. Performanse generativnih algoritama ograničene su dimenzijama prostora za reprezentaciju modela objekta, a u kompleksnijim scenarijima pokazuju manju sposobnost diskriminacije objekta od okoline. S druge strane, diskriminativni algoritmi, koji pored informacija o objektu koriste i pozadinske informacije u bliskoj okolini objekta, imaju bolju sposobnost da obuhvate širok spektar promena u izgledu objekta. U novijim sistemima za praćenje, arhitekture poznate kao sijamske neuralne mreže (eng. *Siamese neural networks*) [95] se koriste za praćenje objekata koje je bazirano na podudaranju sličnosti. Proteklih godina predložen je veliki broj sijamskih mreža za praćenje objekata kako u vidljivom [95], tako i u termalnom infracrvenom opsegu [96], [97]. Osnovni razlog popularnosti pristupa zasnovanih na sijamskim mrežama je taj što su pokazale dobre performanse sa zadovoljavajućom računskom efikasnošću. Međutim, diskriminativne sposobnosti algoritama baziranih na sijamskim mrežama su prilično slabe [95], pošto se fokusiraju samo na učenje vizuelne reprezentacije ciljnog objekta za podudaranje sličnosti, ignorišući pozadinu i specifične informacije o objektu. Usled toga, pokazuju niske performanse u situacijama sa prisustvom okluzija, kao i prilikom deformacija scene.

Klasifikacija algoritama za praćenje u literaturi se može izvršiti i na osnovu posedovanja mehanizma predikcije. Modul za predikciju stanja objekta nije apsolutno neophodan. Algoritam se može konstruisati tako da se detekcija objekta radi na svakom frejmu. Kada su objekti detektovani, može se izvršiti asocijacija podataka kako bi se povezali tragovi praćenih objekata sa trenutno detektovanim objektima. Stoga, u ovim pristupima detekcija mora biti efikasna ako se želi postići rad algoritma u realnom vremenu. Međutim, prilikom praćenja objekta na trenutnom frejmu video sekvence, značajne informacije se mogu dobiti iz praćenja na prethodnim frejmovima. U najmanju ruku, iz prethodnih frejmova video sekvence poznate su procenjene pozicije objekta. Stoga, mnogi algoritmi koriste informacije o poziciji objekta kao početnu tačku na trenutnom frejmu, čime se značajno može smanjiti prostor za pretragu objekta i time povećati efikasnost algoritma. Na osnovu informacija o prethodnim pozicijama objekta može se estimirati brzina kretanja objekta i tako proceniti pozicija na trenutnom frejmu. U praćenju objekata u video sekvenci najčešće korišćeni mehanizmi za predikciju pozicije objekta su različite varijante Kalmanovog i čestičnog filtra. Prednost korišćenja modula za predikciju se posebno primećuje u situacijama kada je objekat zaklonjen usled okluzija, te se sa slike ne može dobiti nikakva informacija o objektu.

Izazovi koji se javljaju u video praćenju opisani u sekciji 3.2 konstantno podstiču razvoj novih algoritama za praćenje objekata. Osnovni cilj svih algoritama za praćenje jeste da na što bolji način

modeluju izgled i kretanje objekta i da na taj način postignu što precizniju lokalizaciju objekta na frejmovima video sekvence i estimiraju trajektoriju kretanja. S druge strane, želja je da se postigne što manja kompleksnost algoritma, kako bi se omogućio rad u realnom vremenu. Analizirajući literaturu o praćenju objekata u video sekvenci, mogu se uočiti tri glavna pristupa u razvoju algoritama. To su diskriminacioni korelacioni filtri [98] – [115], sijamske neuralne mreže [116] – [125], a zatim i pristupi bazirani na transformerima [126] – [129]. Pored praćenja u vidljivom opsegu, ovi pristupi su dominantni i u praćenju objekata u termalnim infracrvenim video sekvencama [96], [97], [130] – [135]. Na slici 3.18 je prikazan detaljan razvoj sa najvažnijim algoritmima iz svakog od navedenih pravaca.



**Slika 3.18** Razvoj algoritama za praćenje objekata u vidljivom i infracrvenom opsegu.

Iako se može primetiti da u dosadašnjoj literaturi nedostaju algoritmi posebno dizajnirani za praćenje objekata u SWIR opsegu, analiza navedenih pristupa razvijenih za vidljivi i termalni infracrveni opseg može poslužiti kao dobra osnova za razvoj algoritma za praćenje objekata u SWIR video sekvenci. Mada SWIR pokriva deo infracrvenog spektra, svetlost u SWIR domenu se pretežno reflektuje od objekata, kao u vidljivom opsegu. Štaviše, razlika SWIR slika u odnosu na termalne infracrvene slike vidi se i u njihovim različitim karakteristikama, obzirom da obrasci na termalnim slikama potiču iz razlika u materijalima od kojih su objekti napravljeni, kao i od njihovih različitih temperatura. Dodatno, promena emitovanog zračenja je proces koji je sporiji od promene reflektovanog zračenja. To znači da se promene u termalnoj slici dešavaju sporije od promena u SWIR slici, što rezultira i različitim karakteristikama šuma u SWIR i termalnim slikama. Stoga, od interesa su prvenstveno algoritmi razvijeni za praćenje u vidljivom opsegu. Kao što je već opisano u sekciji 2.3, u javno dostupnim skupovima video sekvenci za praćenje objekata se ne mogu naći video sekvence napravljene SWIR kamerom, a posebno veliki broj sekvenci kakav je potreban za razvoj algoritama za praćenje baziranih na dubokom učenju. Tu je takođe i generalna mana ovih metoda, izražena kroz nepoznato ponašanje u uslovima za koje nisu obučavane. Analizirajući prednosti i

nedostatke generativnih i diskriminativnih algoritama, diskriminativni algoritmi prilikom modelovanja objekta pored informacija o objektu uzimaju u obzir i informacije o pozadini, te imaju bolje sposobnosti da se lakše nose sa izazovima koji potiču od promene u izgledu objekta tokom praćenja. Uz sve navedeno, veoma važan zahtev jeste i mogućnost da algoritam može procesirati frejmove video sekvence u realnom vremenu, odnosno da kompleksnost algoritma bude što manja. Dakle, analizom prirode SWIR slike kao i literature o diskriminativnim algoritmima za praćenje sa mogućnošću rada u realnom vremenu, prvenstveno u vidljivom opsegu, zaključuje se da pristupi zasnovani na korelacionim filtrima pružaju najbolju osnovu za razvoj algoritma za praćenje pokretnih objekata u sekvenci SWIR slika.

## 4 Korelacione metode za praćenje pokretnih objekata

Metode za praćenje pokretnih objekata u video sekvencama zasnovane na korelacionim filtrima se u osnovi oslanjaju na pristup podudaranja šablona. Kao cilj se postavlja pronalaženje šablona, a koji se naziva korelacioni filtar, na takav način da korelacija ovog filtra i instance objekta kandidata treba da rezultira odgovarajućim korelacionim odzivom. Kvalitet ovako dobijenog korelacionog odziva se meri u odnosu na njegovu bliskost željenom odzivu, a prvenstveno u odnosu na bliskost pika korelacionog odziva piku željenog odziva, gde pik predstavlja maksimalnu vrednost korelacionog odziva. Ako se tokom praćenja pokretnog objekta u video sekvenci između sukcesivnih frejmova instanca objekta kandidata pomeri u odnosu na poziciju centra praćenog objekta, korelacioni filtar treba da generiše korelacioni odziv sa istim iznosom pomeraja lokacije pika.

Kada se korelacioni filtar jednom izračuna, primena ovog filtra za lokalizaciju objekta u frejmovima video sekvence uvek sledi isti postupak. Na frejmu  $x$  video sekvence, primenom naučenog korelacionog obrasca (filtra)  $h$ , korelacioni odziv  $g$  se može izračunati kao:

$$g = x \otimes h \quad (4.1)$$

gde  $\otimes$  predstavlja operaciju korelacije. Prelaskom u Furijeov domen, operacija korelacije se preslikava u operaciju množenja (element po element u slučaju množenja matrica (*Hadamard product*)), a na taj način se proračun korelacije značajno ubrzava. Ako se sa  $\mathcal{F}$  označi Furijeova, a sa  $\mathcal{F}^{-1}$  inverzna Furijeova transformacija, tada se odziv  $g$  iz (4.1) može izračunati kao:

$$g = \mathcal{F}^{-1}(\hat{x} \odot \hat{h}^*) \quad (4.2)$$

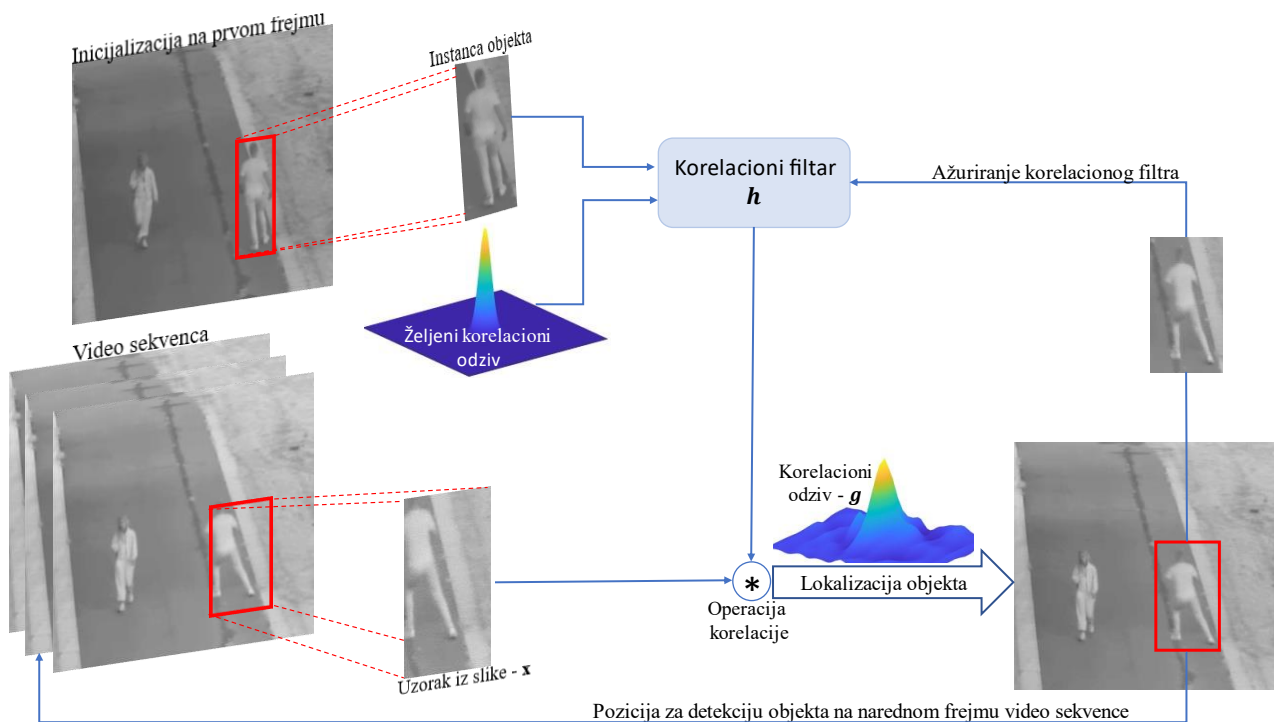
pri čemu je  $\hat{x} = \mathcal{F}(x)$ ,  $\hat{h} = \mathcal{F}(h)$ ,  $*$  označava kompleksnu konjugaciju, a simbol  $\odot$  predstavlja operator množenja (element po element).

Na slici 4.1 je prikazana opšta struktura algoritama za praćenje pokretnih objekata zasnovanih na korelacionim filtrima. Kako se korelacioni filtar dizajnira u odnosu na željeni korelacioni odziv, on se takođe može interpretirati i kao diskriminativni regresorski model.

Dok je kod standardnog pristupa podudaranja šablona filtar dizajniran tako da efikasno odgovara slici iz trening skupa, u praksi se pokazuju veoma loše performanse u scenarijima gde objekat ima razlike u izgledu koje se ne vide u trening skupu, što je veoma česta pojava tokom praćenja objekta. Efikasna lokalizacija, odnosno prepoznavanje objekta koji se prati na novim frejmovima video sekvence zahteva dizajniranje filtra sa većom tolerancijom na izobličenja u odnosu na izgled objekta u osnovnom obučavajućem skupu. Pored toga, filtar u standardnom pristupu podudaranja šablona nije dizajniran da kao rezultat daje oštar pik koji je pogodan za detekciju objekta. Oštar pik u korelacionom odzivu rezultuje lokalizacijom objekta sa većom preciznošću, dok na



testnim slikama na kojima se objekat razlikuje od izgleda objekta na pojedinačnim trening slikama neće proizvesti korelacioni odziv sa jasno izraženim pikom.



Slika 4.1 Struktura algoritama za praćenje objekata u video sekvenci zasnovanih na korelacionim filtrima.

## 4.1 Fundamentalni korelacioni filtri

Najraniji korelacioni filtri nisu bili dizajnirani za primenu u praćenju pokretnih objekata u video sekvencama, ali je princip njihovog dizajna uticao na kasniji dizajn korelacionih filtara koji su prilagođeni i uspešno primenjeni u problemu video praćenja. Tokom osamdesetih i devedesetih godina dvadesetog veka predloženi su brojni pristupi za dizajniranje korelacionih filtara. Filtri su obučavani na primerima ciljnog objekta sa varijacijama u izgledu i sa postavljenim restriktivnim ograničenjima tako da filtri uvek proizvode pikove u korelacionom odzivu iste visine. Neki od značajnijih radova koji su predstavljeni u ovom periodu obuhvataju sintetičke diskriminacione funkcije [136], [137] sintetičke diskriminacione funkcije minimalne varijanse [138], filtre minimalne prosečne korelacione energije [139], filtre optimalnog kompromisa [140], [141], sintetičke diskriminacione funkcije minimalne srednje kvadratne greške [142].

### 4.1.1 Sintetičke diskriminacione funkcije

Polazeći od notacije gde  $x_1, x_2, \dots, x_N$  označavaju  $N$   $d$ -dimenzionalnih vektora, odnosno  $N$  slika sa po  $d$  piksela, gde je broj piksela  $d$  značajno veći od broja slika u trening skupu  $N$ , kod konvencionalnih sintetičkih diskriminacionih funkcija (eng. *Synthetic discriminant functions* - SDF) kompozitna slika, odnosno filtar  $h$  se gradi kao linearna kombinacija slika iz trening skupa:

$$h = \sum_{i=1}^N a_i x_i \quad (4.3)$$

gde se koeficijenti  $a_i$  biraju tako da se zadovolji zahtev:

$$h^T x_j = u_j; \quad j = 1, 2, \dots, N \quad (4.4)$$

pri čemu je  $u_j$  željena vrednost korelacionog odziva, a  $T$  označava operaciju transponovanja. Kada sve trening slike pripadaju istoj klasi,  $u_j$  može biti konstantna vrednost, odnosno za primer dve klase može imati dve različite vrednosti ( $\pm 1$ ). Ako se SDF filtar  $h$  može pronaći, tada se može garantovati da će rezultat korelacije ulazne slike  $x_j$  i dobijenog filtra  $h$  biti željeni izlaz  $u_j$ .

Iz jednačina (4.3) i (4.4) se mogu odrediti koeficijenti  $a_i$ :

$$\sum_{i=1}^N a_i x_i^T x_j = u_j; \quad j = 1, 2, \dots, N \quad (4.5)$$

Jednačina (4.5) se može zapisati i u vektorskom obliku:

$$Ka = u \quad (4.6)$$

gde je  $a = [a_1, a_2, \dots, a_N]^T$ ,  $u = [u_1, u_2, \dots, u_N]^T$ , a elementi  $K_{ij}$  matrice  $K$  predstavljaju elemente proizvoda vektora  $x_i^T x_j$  iz (4.5). Ako je matrica  $K$  invertibilna, tada se rešenje jednačine (4.6) može dobiti kao:

$$a = K^{-1}u \quad (4.7)$$

Može se pokazati da rešenje postoji, odnosno da je matrica  $K$  punog ranga i postoji inverzna matrica  $K^{-1}$  ako i samo ako originalni trening skup od  $N$  slika  $x_1, x_2, \dots, x_N$  formira linearno nezavisni skup vektora [138].

Kompozitna slika, odnosno filtar  $h$ , se može dobiti i u kompaktnijoj formi ako se definiše matrica  $X$  kao matrica dimenzija  $d \times N$ , gde  $i$ -ta kolona ove matrice predstavlja vektor  $x_i$ . Tada se filtar  $h$  može predstaviti relacijom:

$$h = Xa \quad (4.8)$$

pri čemu vektor koeficijenata  $a$  treba da zadovolji relaciju:

$$(X^T X)a = u \quad (4.9)$$

Oдавde se vidi da je matrica  $X^T X$  zapravo identična matrici  $K$  iz (4.6). Stoga, iz jednačina (4.8) i (4.9) se filtar  $h$  može izraziti kao:

$$h = X(X^T X)^{-1}u \quad (4.10)$$

## 4.1.2 Sintetičke diskriminacione funkcije minimalne varijanse

Sintetičke diskriminacione funkcije prikazane u sekciji 4.1.1 su konstruisane tako da se obezbedi da vrednost kros-korelacije ima željenu vrednost konstante  $u_j$  za svaku ulaznu sliku  $x_j$  koja pripada trening skupu. Međutim, kada se na ulaz dovede slika koja ne pripada trening skupu, ili predstavlja zašumljenu sliku iz trening skupa, izlaz neće imati vrednost željene konstante. Za praktičnu implementaciju od interesa jesu samo filtri kod kojih se varijacija izlaza u ovim slučajevima može učiniti dovoljno malom. Stoga, sintetičke diskriminacione funkcije minimalne varijanse imaju za cilj, kao što im i ime kaže, minimizaciju varijanse na izlazu kada se na ulaz sistema dovede zašumljena slika iz trening skupa, kojom se takođe može modelovati i ulazna slika koja ne pripada trening skupu. Filtar sintetičke diskriminacione funkcije minimalne varijanse predložio je Kumar 1986. godine u radu [138].

Zašumljena slika iz trening skupa može se predstaviti kao  $x_j + n$ , gde  $x_j$  predstavlja originalnu nezašumljenu sliku iz trening skupa, a  $n$  vektor aditivnog stacionarnog šuma nulte srednje vrednosti. Uzimajući u obzir jednačinu (4.4) korelacioni odziv je u slučaju zašumljene slike izražen kao:

$$h^T(x_j + n) = u_j + h^T n; \quad j = 1, 2, \dots, N \quad (4.11)$$

gde je član  $h^T n$  u (4.11) rezultat prisustva šuma. Za ovaj član se može pokazati da su srednja vrednost:

$$E\{h^T n\} = 0 \quad (4.12)$$

i varijansa:

$$\text{Var}\{h^T n\} = E\{h^T n n^T h\} = h^T E\{n n^T\} h = h^T C h \quad (4.13)$$

gde  $E\{\cdot\}$  označava matematičko očekivanje, a  $\text{Var}\{\cdot\}$  varijansu. U (4.13)  $C$  označava  $d \times d$  kovarijacionu matricu vektora šuma  $n$ . Uzima se da je pozitivno definitna, čime se garantuje postojanje njene inverzne matrice. Dalje se za izvođenje filtra koji minimizira varijansu u (4.13) i pritom zadovoljava zahtev iz (4.4), uvodi funkcional [138]:

$$\phi = h^T C h - \sum_{i=1}^N 2\lambda_i (h^T x_i - u_i) \quad (4.14)$$

gde koeficijenti  $\lambda_i$  predstavljaju Lagranžove multiplikatore. Izjednačavanjem izvoda ovog funkcionala po  $h$  sa  $\mathbf{0}$  (nula vektor), dobija se da optimalni filter zadovoljava sledeću relaciju:

$$C h_{opt} = \sum_{i=1}^N \lambda_i x_i \quad (4.15)$$

Uzimajući u obzir pretpostavku da je kovarijaciona matrica šuma  $C$  invertibilna, optimalni filter sintetičke diskriminacione funkcije minimalne varijanse ima oblik:

$$h_{opt} = C^{-1} \left[ \sum_{i=1}^N \lambda_i x_i \right] = \sum_{i=1}^N \lambda_i C^{-1} x_i \quad (4.16)$$

dok uslov (4.4) sada poprima oblik:

$$\sum_{i=1}^N \lambda_i x_i^T C^{-1} x_j = u_j; \quad j = 1, 2, \dots, N \quad (4.17)$$

Iz (4.16) se vidi da optimalni filter sada ne predstavlja linearnu kombinaciju slika iz trening skupa. Međutim, ako se slike iz trening skupa transformišu množenjem sa  $C^{-1}$ , tada se optimalni filter može posmatrati kao linearna kombinacija transformisanih slika iz trening skupa, ali se koeficijenti ove linearne kombinacije  $\lambda_i$  izvode koristeći član  $x_i^T C^{-1} x_j$  iz (4.17) [138]. Odatle se, prateći notaciju za vektore  $X$  i  $u$  iz sekcije 4.1.1, dobija optimalni filter u zatvorenoj formi:

$$h_{opt} = C^{-1} X (X^T C^{-1} X)^{-1} u \quad (4.18)$$

Iako je filter sintetičke diskriminacione funkcije minimalne varijanse dizajniran kako bi se prevazišao problem šuma na slikama, ovaj filter i dalje u korelacionom odzivu ne daje oštre pikove. Kako bi se unapredila mogućnost razlikovanja pika od bočnih lobova u korelacionoj ravni, dizajniran je filter minimalne prosečne korelacione energije [139].

### 4.1.3 Filter minimalne prosečne korelacione energije

Iz potrebe da se u korelacionom odzivu izdvoji jasan i oštar pik, odnosno potisnu bočni lobovi, predložen je filter minimalne prosečne korelacione energije (eng. *Minimum Average Correlation Energy* - MACE) [139]. MACE filter ima za cilj da kontroliše celokupnu ravan korelacionog odziva, odnosno minimizira energiju u korelacione ravni koja uključuje i bočne lobove. Prosečna energija korelacionog odziva za sliku iz trening skupa koja ima  $d$  piksela je definisana na sledeći način:

$$E_i = \sum_{l=1}^d |g_i(l)|^2 \quad (4.19)$$

Prelaskom u Furijeov domen i korišćenjem Parsevalove teoreme, (4.19) dobija oblik:

$$E_i = \frac{1}{d} \sum_{l=1}^d |\hat{g}(l)|^2 = \frac{1}{d} \sum_{l=1}^d |\hat{h}(l)|^2 |\hat{x}_i(l)|^2 \quad (4.20)$$

gde  $\hat{h}$  i  $\hat{x}_i$  predstavljaju filter i sliku iz trening skupa u Furijeovom domenu, respektivno. Odavde sledi da se prosečna korelaciona energija za ceo trening skup može dobiti kao:

$$E = \frac{1}{d \cdot N} \sum_{i=1}^N (\hat{h}^* \hat{x}_i) (\hat{x}_i^* \hat{h}) = \hat{h}^* \left[ \frac{1}{d \cdot N} \sum_{i=1}^N \hat{x}_i \hat{x}_i^* \right] \hat{h} = \hat{h}^* D \hat{h} \quad (4.21)$$

gde je  $D = \left[ \frac{1}{d \cdot N} \sum_{i=1}^N \hat{x}_i \hat{x}_i^* \right]$  dijagonalna matrica dimenzija  $d \times d$  čiji elementi na dijagonali predstavljaju prosečnu spektralnu gustinu snage slika iz trening skupa, a simbol \* označava operaciju kompleksne konjugacije [139], [143].

Za minimiziranje prosečne korelacione energije (4.21) uz ograničenja vezana za korelacioni odziv data u frekvencijskom domenu, treba rešiti optimizacioni problem:

$$\begin{aligned} \min_{\hat{h}} \quad & \hat{h}^* D \hat{h} \\ \text{Za} \quad & \hat{X}^* \hat{h} = u \end{aligned} \quad (4.22)$$

pri čemu  $\hat{X}$  u (4.22) predstavlja matricu čije su kolone vektorske reprezentacije slika iz trening skupa u Furijeovom domenu. Prateći istu metodologiju rešavanja kvadratnog optimizacionog problema uz linearna ograničenja kao i u sekciji 4.1.2, primenom Lagranžovih multiplikatora, optimalno rešenje za filter minimalne prosečne korelacione energije izraženo je kao:

$$\hat{h}_{opt} = D^{-1} \hat{X} (\hat{X}^* D^{-1} \hat{X})^{-1} u \quad (4.23)$$

MACE filtri su prvi filtri koji su pokušali da kontrolišu celokupnu ravan korelacionog odziva i postavili osnovu za razvoj korelacionih filtara u frekvencijskom domenu. Rezultat dizajna ovakvog tipa filtara jeste da oni generalno proizvode veoma oštre pikove u korelacionoj ravni. Međutim, dizajn MACE filtara ima i svojih nedostataka. Prvi nedostatak jeste osetljivost ovih filtara na prisustvo šuma u slici. Takođe, MACE filtri često imaju veoma veliku osetljivost na varijacije unutar iste klase. Kako bi se postigao kompromis između pristupa filtara sintetičke diskriminacione funkcije minimalne varijanse koji potiskuju šum, ali ne generišu oštre korelacione pikove i pristupa MACE filtara kod kojih su pikovi u korelacionoj ravni jasno izraženi, ali imaju osetljivost na prisustvo šuma, predložen je filter sintetičke diskriminacione funkcije optimalnog kompromisa [140], [141].

#### 4.1.4 Sintetičke diskriminacione funkcije optimalnog kompromisa

Filter optimalnog kompromisa (eng. *Optimal trade-off filter*) ima za cilj da napravi balans između optimizacije kriterijuma prosečne korelacione energije i kriterijuma minimalne varijanse šuma na izlazu [140], [141]. Stoga, ovaj tip filtara minimizira prosečnu korelacionu energiju za datu varijansu šuma i obrnuto.

Kriterijum prosečne korelacione energije dat je u (4.22) kao:

$$E_1 = \hat{h}^* D \hat{h} \quad (4.24)$$

dok je kriterijum za minimizaciju varijanse šuma na izlazu dat u (4.13), a može se izraziti u Furijeovom domenu kao:

$$E_2 = \hat{h}^* C \hat{h} \quad (4.25)$$

Optimalni kompromis između ova dva kriterijuma se može postići kreiranjem kriterijuma koji predstavlja težinsku sumu kriterijuma  $E_1$  i  $E_2$ , uz zadovoljenje ograničenja iz (4.4):

$$\begin{aligned} \min_{\hat{h}} \quad & \alpha \hat{h}^* D \hat{h} + \beta \hat{h} C \hat{h} \\ & \hat{X}^* \hat{h} = u \end{aligned} \quad (4.26)$$

pri čemu su  $\alpha$  i  $\beta$  nenegativni koeficijenti [143]. Kao i u sekcijama 4.1.2 i 4.1.3, primenom metode Lagranžovih multiplikatora, rešenje za filter optimalnog kompromisa se može izraziti u zatvorenoj formi:

$$\hat{h} = (\alpha D + \beta C)^{-1} \hat{X} (\hat{X}^* (\alpha D + \beta C)^{-1} \hat{X})^{-1} u \quad (4.27)$$

Postavlja se pitanje kako odrediti koeficijente  $\alpha$  i  $\beta$ . Kako nije moguće da se minimizuju oba kriterijuma, poželjno je da se minimizira kriterijum  $E_2$  za svaku moguću vrednost kriterijuma  $E_1$ . Ovo se može postići minimiziranjem funkcije  $E(\eta)$  [140], [143]:

$$E(\eta) = E_2 - \eta E_1 - \hat{h}^* X L \quad (4.28)$$

gde je  $\eta$  Lagranžov multiplikator koji ograničava  $E_1$  na datu fiksnu vrednost, a  $L$  predstavlja vektor od  $N$  Lagranžovih multiplikatora, koji se odnose na linearna ograničenja korelacionih odziva za  $N$  slika iz trening skupa. Parametar  $\eta$  ima vrednosti iz opsega  $(-\infty, 0]$ , gde za najnepovoljniju vrednost  $E_1$  vrednost parametra  $\eta$  iznosi 0, pri čemu se optimizuje samo kriterijum  $E_2$ , dok za najbolju vrednost  $E_1$  parametar  $\eta$  uzima vrednost  $-\infty$ , pri čemu se i optimizuje samo kriterijum  $E_1$  [140]. Ako se parametar  $\eta$  definiše kao:

$$\eta = \frac{\mu - 1}{\mu}, \quad \mu \in [0, 1] \quad (4.29)$$

i  $\Gamma = \mu L$ , tada se kriterijum (4.28) svodi na optimizaciju kriterijuma  $E(\mu)$ :

$$E(\mu) = \mu E_2 + (1 - \mu) E_1 - \hat{h}^* X \Gamma \quad (4.30)$$

Variranjem parametra  $\mu$  u opsegu od 0 do 1, optimizacija kriterijuma  $E(\mu)$  se pomera sa minimiziranja kriterijuma  $E_1$  ka minimizovanju kriterijuma  $E_2$ . U slučaju filtera sintetičkih diskriminacionih funkcija, filter optimalnog kompromisa se može izračunati kao [140]:

$$\begin{aligned} \hat{h} &= B^{-1} \hat{X} (\hat{X}^* B^{-1} \hat{X})^{-1} u \\ B &= \mu D + (1 - \mu) C \end{aligned} \quad (4.31)$$

U mnogim slučajevima, poželjan je izbor vrednosti parametra  $\mu$  blizak 1. Tako će dizajnirani filteri optimalnog kompromisa imati oštrije pikove u korelacionoj ravni, a istovremeno će biti robusni na šum što je slučaj kod SDF filtera minimalne varijanse.

#### 4.1.5 Sintetičke diskriminacione funkcije minimalne srednje kvadratne greške

Filter sintetičke diskriminacione funkcije minimalne srednje kvadratne greške ima za cilj da obezbedi direktnu kontrolu nad oblikom korelacionog odziva. Ovo se postiže minimiziranjem srednje kvadratne greške između rezultujućeg korelacionog odziva i unapred specificiranog željenog korelacionog odziva proizvoljnog oblika [142].

Između željenog korelacionog odziva  $c_i$  i rezultujućeg korelacionog odziva  $g_i$  (za sliku  $x_i$  koja sadrži  $d$  piksela i filtra  $h$ ), srednje kvadratna greška  $E_i$  se može definisati kao:

$$E_i = \sum_{l=1}^d |c_i(l) - g_i(l)|^2 = \frac{1}{d} \sum_{l=1}^d |\hat{c}_i(l) - \hat{x}_i^*(l)\hat{h}(l)|^2, i = 1, 2, \dots, N \quad (4.32)$$

gde  $\hat{c}_i$ ,  $\hat{h}$ , i  $\hat{x}_i$  predstavljaju željeni korelacioni odziv, filter i sliku u Furijeovom domenu, respektivno, a simbol  $*$  se odnosi na operaciju kompleksne konjugacije. Srednje kvadratna greška  $E_i$  predstavlja meru koliko dobro korelacioni odziv  $g_i$  aproksimira željeni korelacioni odziv  $c_i$  za koji se smatra da je kompletno definisan za sve slike u skupu. Može se primetiti da za slučaj  $c_i = 0$ , srednje kvadratna greška  $E_i$  iz (4.32) zapravo predstavlja energiju korelacionog odziva  $g_i$ , kao u (4.19), pa se izvodi zaključak da filtri sintetičke diskriminacione funkcije srednje kvadratne greške generalizuju MACE filtre. Za ceo trening skup, srednje kvadratna greška je definisana kao:

$$E = \frac{1}{N} \sum_{i=1}^N E_i \quad (4.33)$$

Uz ograničenja vezana za korelacioni odziv iz (4.4), a data u frekvencijskom domenu:

$$\hat{X}^* \hat{h} = u \quad (4.34)$$

pri čemu  $\hat{X}$  u (4.34) predstavlja matricu čije su kolone vektorske reprezentacije slika sa  $d$  piksela iz trening skupa u Furijeovom domenu. Definisanjem vektora  $\hat{C}_i$ :

$$\hat{C}_i = [\hat{c}_i(1), \hat{c}_i(2), \dots, \hat{c}_i(d)]^T, \quad i = 1, 2, \dots, N \quad (4.35)$$

i dijagonalne matrice  $\hat{\Delta}_i$  dimenzija  $d \times d$ :

$$\hat{\Delta}_i = \text{diag}([\hat{x}_i(1), \hat{x}_i(2), \dots, \hat{x}_i(d)]), \quad i = 1, 2, \dots, N \quad (4.36)$$

jednačina srednje kvadratne greške iz (4.32) dobija oblik:

$$E_i = [(\hat{C}_i - \hat{\Delta}_i^* \hat{h})^* (\hat{C}_i - \hat{\Delta}_i^* \hat{h})], \quad i = 1, 2, \dots, N \quad (4.37)$$

Ako se prema [142] definišu:

$$\begin{aligned} E_c &= \frac{1}{N} \sum_{i=1}^N (\hat{C}_i^* \hat{C}_i) \\ \hat{p} &= \frac{1}{N} \sum_{i=1}^N (\hat{\Delta}_i \hat{C}_i) \\ \hat{M} &= \frac{1}{N} \sum_{i=1}^N (\hat{\Delta}_i^* \hat{\Delta}_i) \end{aligned} \quad (4.38)$$

jednačina srednje kvadratne greške (4.33) dobija oblik [142]:

$$E = [E_c - \hat{h}^* \hat{p} - \hat{p}^* \hat{h} + \hat{h}^* \hat{M} \hat{h}] \quad (4.39)$$

Primenom metode Lagranžovih multiplikatora, rešenje za filter sintetičke diskriminacione funkcije minimalne srednje kvadratne greške dobija se kao:

$$\hat{h} = \hat{M}^{-1}(\hat{p} + \hat{X}w) \quad (4.40)$$

gde  $w$  predstavlja vektor od  $N$  Lagranžovih multiplikatora [142]. Vektor  $w$  se može dobiti iz (4.40) i jednačine ograničenja (4.34):

$$w = (\hat{X}^* \hat{M}^{-1} \hat{X})^{-1} [u - \hat{X}^* \hat{M}^{-1} \hat{p}] \quad (4.41)$$

Zamenom  $w$  iz (4.41) u (4.40), dobija se konačni oblik filtra sintetičke diskriminacione funkcije minimalne srednje kvadratne greške:

$$\hat{h} = \hat{M}^{-1} \hat{p} + \hat{M}^{-1} \hat{X} (\hat{X}^* \hat{M}^{-1} \hat{X})^{-1} [u - \hat{X}^* \hat{M}^{-1} \hat{p}] \quad (4.42)$$

## 4.2 Savremeni korelacioni filtri za praćenje pokretnih objekata

Korelacioni filtri bazirani na sintetičkim diskriminacionim funkcijama se obučavaju na primerima objekata od interesa za praćenje sa različitim izgledom, ali i sa nametnutim čvrstim ograničenjima tako da ovi filtri uvek rezultuju korelacionim ravnima sa pikovima iste visine. Takođe, nametnuta čvrsta ograničenja uzrokuju i to da ovi filtri imaju problema sa tolerancijama distorzija objekta. U [144] je rešenje potraženo uklanjanjem navedenih strogih ograničenja i postavljen je zahtev da filter ima visoku srednju vrednost korelacionog odziva. Međutim, potpunim uklanjanjem ograničenja kao i izostankom iskorišćenja prethodnih informacija koje se odnose na celokupni korelacioni odziv kod trening slika, javlja se problem osetljivosti na šum. Stoga, ovi pristupi u dizajniranju korelacionih filtara za praćenje pokretnih objekata imaju slaba diskriminativna svojstva. Štaviše, potreba i način obučavanja ovih filtara čini ih nepodesnim za onlajn praćenje objekata. Koristeći adaptivni princip obučavanja, filter minimalne sume kvadrata grešaka [98] postigao je značajna poboljšanja robusnosti i efikasnosti praćenja, a usled jednostavnosti pristupa ima i mogućnost rada u realnom vremenu.

### 4.2.1 Filter minimalne sume kvadrata grešaka

U algoritmu minimalne sume kvadratnih grešaka (eng. *Minimum Output Sum of Squared Error* - MOSSE) korelacioni filter  $h$  se izračunava na osnovu trening skupa slika  $x_i$  i željenih korelacionih odziva  $g_i$  za svaku sliku iz trening skupa. Iako odziv  $g_i$  može imati bilo koji oblik, najčešće se usvaja 2D Gausova funkcija male varijanse i sa ostrim pikom centriranim tačno na poziciji objekta na odgovarajućoj trening slici. Obučavanje filtra se izvršava u Furijeovom domenu, kako bi se iskoristile prednosti osobine konvolucije koja se u Furijeovom domenu preslikava u operaciju množenja.



MOSSE algoritam kao kriterijum za optimizaciju filtra definiše srednje kvadratnu grešku između rezultujućeg i željenog korelacionog odziva [98]. Prateći notaciju iz (4.2), minimizacioni problem u Furijeovom domenu kod MOSSE algoritma ima oblik:

$$\min_{\hat{h}^*} \sum_{i=1}^N |\hat{x}_i \odot \hat{h}^* - g_i| \quad (4.43)$$

gde  $N$  predstavlja broj slika u trening skupu. Rešenje ovog optimizacionog problema predstavlja optimalni filter MOSSE algoritma i može se dobiti u zatvorenoj formi kao [98]:

$$\hat{h}^* = \frac{\sum_{i=1}^N g_i \odot \hat{x}_i^*}{\sum_{i=1}^N \hat{x}_i \odot \hat{x}_i^*} \quad (4.44)$$

Kako zbog izazova koji se javljaju tokom video praćenja, opisanih u sekciji 3.2, dolazi do veoma čestih promena u izgledu objekta koji se prati, MOSSE algoritam koristi mehanizam adaptiranja filtra trenutnoj situaciji na sceni. Filter  $h$  iz (4.44) je izražen kao odnos imenioca  $B_i$  i brojioca  $A_i$ :

$$\hat{h}^* = \frac{A_i}{B_i}, \quad (4.45)$$

$$A_i = \sum_{i=1}^N g_i \odot \hat{x}_i^*, \quad B_i = \sum_{i=1}^N \hat{x}_i \odot \hat{x}_i^*$$

gde se  $A_i$  i  $B_i$  ažuriraju tokom vremena kao:

$$A_i = (1 - \eta)A_{i-1} + \eta g_i \odot \hat{x}_i^* \quad (4.46)$$

$$B_i = (1 - \eta)B_{i-1} + \eta \hat{x}_i \odot \hat{x}_i^*$$

pri čemu parametar  $\eta$  predstavlja stopu učenja. Predloženi metod ažuriranja filtra daje veći značaj novijim frejmovima video sekvence, dok uticaj izgleda objekta na starijim frejmovima treba eksponencijalno da opada tokom vremena. Vrednost parametra  $\eta$  treba da bude postavljena tako da se filter brzo prilagođava promenama objekta na sceni, ali da istovremeno očuva robusnost filtra. Takođe, u jednačinu filtra  $\hat{h}^*$  se uvodi i regularizacioni parametar  $\lambda$  kako bi se izbeglo deljenje nulom:

$$\hat{h}^* = \frac{\sum_{i=1}^N g_i \odot \hat{x}_i^*}{\sum_{i=1}^N \hat{x}_i \odot \hat{x}_i^* + \lambda} = \frac{A_i}{B_i + \lambda} \quad (4.47)$$

Dizajnom filtra kao i proračunima ažuriranja modela, MOSSE algoritam za praćenje pokretnih objekata pruža mogućnost rada u realnom vremenu, sa impresivnim brzinama procesiranja frejmova [98]. Usled ovakvih mogućnosti rada u realnom vremenu MOSSE filter je postao održiva polazna tačka za razvoj korelacionih filtera za praćenje pokretnih objekata koji mogu biti robusniji i postići bolje performanse u praćenju.

## 4.2.2 Kernelizovani korelacioni filtri

Do izraza za korelacioni filter koji se može koristiti za praćenje objekata može se doći i ako se filter posmatra kao diskriminativni klasifikator, čiji je cilj razlikovanje objekta od pozadinske scene. Za obučavanje takvog klasifikatora potrebno je obezbediti odgovarajući skup uzoraka objekta od interesa, kako bi se što bolje izvršila njegova karakterizacija. Međutim, osnovni princip diskriminativnih metoda jeste da se značaj mora dati i relevantnom pozadinskom okruženju. Stoga, pored skupa uzoraka objekta, treba obezbediti i skup takozvanih negativnih uzoraka. Kao negativni uzorci se najčešće koriste uzorci sa različitih lokacija na slici, ali i translirani i skalirani uzorci samog objekta. Na taj način se odražavaju uslovi u kojima će se klasifikator naći tokom praćenja objekta. Takođe, negativni uzorci se mogu i sintetički generisati. Broj sintetički generisanih uzoraka koji se koriste za obučavanje klasifikatora može biti veoma veliki. Inkorporiranje velikog broja uzoraka poboljšava performanse algoritma, ali posledično utiče na brzinu izvršavanja algoritma, pa se mora naći kompromis između performansi i mogućnosti rada u realnom vremenu. Polazeći od osobina cirkularnih matrica, u Furijeovom domenu se može pokazati da filter za praćenje pokretnih objekata posmatran kao klasifikacioni algoritam ima male računске zahteve iako se broj uzoraka povećava. Naime, bez eksplicitne iteracije kroz hiljade uzoraka, različite relativne translacije osnovnog uzorka objekta se mogu uvesti analitički. Takođe, pokazaće se da postoji prirodna veza između linearne regresije i klasičnih korelacionih filtera, odakle se može doći do izraza za korelacioni filter za praćenje pokretnih objekata u video sekvenci [99].

Polazna tačka jeste grebena regresija, gde je cilj da se pronađe linearna diskriminativna funkcija u obliku:

$$f(z) = w^T z \quad (4.48)$$

koja minimizira kriterijumsku funkciju kvadrata grešaka između uzoraka  $x_i$  i njihovih regresionih ciljeva  $y_i$ :

$$\min_w \sum_i (f(x_i) - y_i)^2 + \lambda \|w\|^2 \quad (4.49)$$

gde je  $\lambda$  u (4.49) regularizacioni parametar koji kontroliše preobučavanje. Parcijalnim izvodom kriterijuma (4.49) po  $w$  i izjednačavanjem sa nulom, rešenje minimizatora  $w$  se dobija u zatvorenoj formi kao:

$$w = (X^T X + \lambda I)^{-1} X^T y \quad (4.50)$$

pri čemu  $X$  predstavlja matricu uzoraka  $x_i$  koji su svrstani u redove,  $y$  je vektor regresionih ciljeva  $y_i$ , a  $I$  označava jediničnu matricu. Posmatrano u Furijeovom domenu, gde su veličine uglavnom kompleksne, za slučaj kompleksne matrice  $X$ , jednačina (4.50) ima oblik:

$$w = (X^H X + \lambda I)^{-1} X^H y \quad (4.51)$$

gde  $X^H$  označava hermitski transponovanu matricu  $X$ ,  $X^H = (X^*)^T$ , a  $X^*$  označava kompleksnu konjugaciju matrice  $X$ . Kako je ovaj postupak računski veoma kompleksan, za potrebe praćenja

objekata u realnom vremenu, pristup u opštem slučaju nije upotrebljiv. Međutim, u slučaju kada matrica  $X$  ima osobine cirkularne matrice, računski postupak je značajno jednostavniji.

#### 4.2.2.1 Ciklična pomeranja i cirkularne matrice

Cilj je da se klasifikator obuči baznim uzorkom objekta, ali i negativnim uzorcima koji predstavljaju pozadinske uzorke ili translirani bazni uzorak. Translirani uzorci se mogu dobiti sintetičkim putem, cikličnim pomeranjem baznog uzorka. Ako se, radi jednostavnosti, posmatra jednodimenzionalni vektor  $x$  koji predstavlja bazni uzorak objekta od interesa za praćenje, negativni uzorci se mogu dobiti primenom operatora cikličnog pomeranja. Ovaj operator predstavlja permutacionu matricu:

$$P = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.52)$$

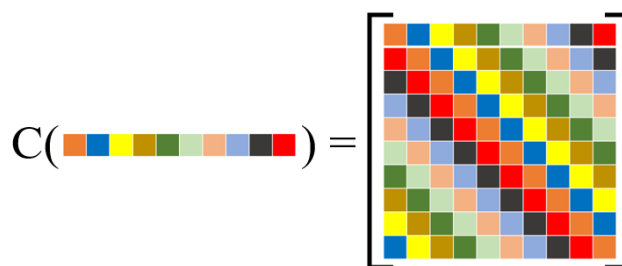
Primena ovog operatora na bazni uzorak  $x$  rezultira pomeranjem elemenata vektora  $x$  za jedno mesto:

$$Px = [x_n \ x_1 \ x_2 \ \dots \ x_{n-1}]^T \quad (4.53)$$

što predstavlja malu translaciju baznog uzorka. Primenom niza pomeranja (4.53), može se dobiti veća translacija baznog uzorka. Niz od  $d$  pomeranja postiže se stepenovanjem matrice  $P$ , pa je translacija elemenata vektora  $x$  za  $d$  mesta data kao  $P^d x$ . Odavde se može zapaziti da se zbog cikličnog svojstva isti bazni vektor  $x$  periodično dobija na svakih  $n$  pomeranja, gde je  $n$  broj elemenata vektora  $x$ . To znači da se potpuni skup pomeraja dobija kao:

$$C(x) = \{P^d x \mid d = 0, 1 \dots n - 1\} \quad (4.54)$$

gde  $C$  označava operator generisanja cirkularne matrice, što je i grafički prikazano na slici 4.2.



**Slika 4.2** Ilustracija rezultata primene operatora cirkularne matrice na bazni jednodimenzionalni vektor  $x$ .

Stoga, može se zaključiti da je za generisanje skupa potrebnog za obučavanje klasifikatora dovoljno poznavanje baznog uzorka  $x$ , a da se ostali translirani uzorci mogu dobiti na potpuno deterministički način na osnovu baznog uzorka. Tako se matrica  $X$  iz (4.50) – (4.51) može predstaviti kao:

$$X = C(x) = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ x_n & x_1 & x_2 & \dots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \dots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \dots & x_1 \end{bmatrix} \quad (4.55)$$

Veoma važno svojstvo cirkularnih matrica jeste osobina da cirkularne matrice nakon primene diskretne Furijeove transformacije postaju dijagonalne [99]. Ova osobina je nezavisna od generišućeg baznog vektora  $x$  i može se izraziti kao:

$$X = F \text{diag}(\hat{x})F^H \quad (4.56)$$

gde  $\hat{x}$  označava diskretnu Furijeovu transformaciju vektora  $x$ . Matrica  $F$  je konstantna matrica i predstavlja takozvanu matricu diskretne Furijeove transformacije. Koristi se za izračunavanje diskretne Furijeove transformacije  $\mathcal{F}$  bilo kog ulaznog vektora:

$$\mathcal{F}(z) = \sqrt{n}Fz \quad (4.57)$$

Osobina cirkularnih matrica da nakon primene diskretne Furijeove transformacije postaju dijagonalne ima velikog praktičnog značaja jer dovodi do značajnog ubrzanja računski zahtevnog proračuna u (4.51).

#### 4.2.2.2 Pojednostavljenje linearne regresije za cirkularne matrice

Nakon generisanja celokupnog trening skupa  $X$  primenom operatora generisanja cirkularne matrice i uzimajući u obzir osobine cirkularnih matrica, proračun izraza (4.51) se može značajno pojednostaviti. Ako se posmatra komponenta  $X^H X$  iz (4.51) i jednačina (4.56), odnosno osobina da cirkularna matrica nakon primene diskretne Furijeove transformacije postaje dijagonalna, dobija se:

$$X^H X = F \text{diag}(\hat{x}^*)F^H F \text{diag}(\hat{x})F^H \quad (4.58)$$

Usled simetričnosti dijagonalnih matrica, primenom hermitskog transponovanja na dijagonalnu matricu  $\text{diag}(\hat{x})$ , kao rezultat ostaje samo konjugovano kompleksni deo  $\hat{x}^*$  [99]. Kako je proizvod  $F^H F$  jednak jediničnoj matrici, izraz (4.58) dobija oblik:

$$X^H X = F \text{diag}(\hat{x}^*)\text{diag}(\hat{x})F^H \quad (4.59)$$

Dalje, proizvod dijagonalnih matrica se svodi na proizvod element po element označen operatorom  $\odot$ , pa se posmatrani izraz svodi na:

$$X^H X = F \text{diag}(\hat{x}^* \odot \hat{x})F^H \quad (4.60)$$

Ako se dobijeni rezultat o dijagonalizaciji cirkularnih matrica iz (4.60) primeni na izraz (4.51), izraz za klasifikator  $w$  u Furijeovom domenu dobija oblik [99]:

$$\hat{w} = \text{diag}\left(\frac{\hat{x}^*}{\hat{x}^* \odot \hat{x} + \lambda}\right)\hat{y} \quad (4.61)$$

pri čemu se i operacija deljenja u (4.61) izvršava element po element. Kako se proizvod dijagonalne matrice i vektora svodi na proizvod element po element elemenata sa dijagonale matrice i elemenata vektora, dalje se izraz (4.61) svodi na:

$$\hat{w} = \frac{\hat{x}^* \odot \hat{y}}{\hat{x}^* \odot \hat{x} + \lambda} \quad (4.62)$$

U prostornom domenu, klasifikator  $w$  se iz (4.62) jednostavno može dobiti inverznom Furijeovom transformacijom.

Za razliku od MOSSE filtra opisanog u sekciji 4.2.1 koji minimizira sumu kvadratnih grešaka između rezultujućih korelacionih odziva slika iz trening skupa  $x_i$  i željenih korelacionih odziva  $g_i$ , izraz (4.62) je definisan za jedan bazni uzorak  $x$ . Ako se za svaki bazni uzorak  $x_i$  definiše cirkularna matrica  $X_i$ , a zatim u izrazu (4.51) matrica  $X$  zameni sa  $X = [X_1 X_2 X_3 \dots]^T$ , tada se može pisati [99]:

$$w = \sum_j \left( \sum_i X_i^H X_i + \lambda I \right)^{-1} X_j^H y \quad (4.63)$$

Dalje važi:

$$w = \left( \sum_i X_i^H X_i + \lambda I \right)^{-1} \left( \sum_i X_i^H \right) y \quad (4.64)$$

odakle se vidi da izraz (4.64) ima isti oblik kao (4.51). Prateći date korake za dijagonalizaciju cirkularnih matrica, može se dobiti izraz za filter u obliku [99]:

$$\hat{w} = \frac{\sum_i \hat{x}_i^* \odot \hat{y}}{\sum_i \hat{x}_i^* \odot \hat{x} + \lambda} \quad (4.65)$$

Izraz (4.65) ima isti oblik kao i izraz za MOSSE filter u (4.47), pri čemu u (4.65)  $y$  predstavlja željeni korelacioni odziv ( $g$  u izrazu (4.47)), a  $w$  korelacioni filter ( $h$  u izrazu (4.47)). Međutim, postoje važne razlike između ova dva filtra [99]. Izraz u (4.65) je dobijen polazeći od opšteg problema grebene regresije i osobina cirkularnih matrica, dok je MOSSE filter izveden od namenski definisane kriterijumske u Furijeovom domenu. Takođe, bitno je uočiti i da je regularizacioni parametar  $\lambda$  uveden u izraz (4.47) samo sa ciljem da se izbegne deljenje nulom.

#### 4.2.2.3 Kernelizacija korelacionih filtara

Kako bi se dobile efikasnije, nelinearne regresione funkcije  $f(x)$  iz (4.48), može se primeniti takozvani kernel trik [145]. Funkcija kernel trika  $k(x, x')$  definisana kao:

$$k(x, x') = \varphi^T(x) \varphi(x') \quad (4.66)$$

na implicitan način mapira uzorke  $x_i$  u višedimenzionalni prostor  $\varphi$  (dualni prostor), bez stvarnog instanciranja vektora u tom prostoru. Sada, u dualnom prostoru, klasifikator  $w$  se može predstaviti kao [99]:

$$w = \sum_i \alpha_i \varphi(x_i) \quad (4.67)$$

gde parametri  $\alpha_i$  predstavljaju parametre za optimizaciju. Ako se definiše kernel matrica  $K$  dimenzija  $N \times N$ , čiji su elementi:

$$K_{i,j} = k(x_i, x_j) \quad (4.68)$$

pri čemu  $N$  predstavlja broj uzoraka  $x_i$  u trening skupu. Sada se rešenje u dualnom prostoru za kernelizovanu verziju grebene regresije:

$$f(z) = w^T z = \sum_i \alpha_i k(z, x_i) \quad (4.69)$$

može dobiti u obliku:

$$\alpha = (K + \lambda I)^{-1} y \quad (4.70)$$

gde je  $\alpha$  vektor čiji su elementi  $\alpha_i$  [99]. Koristeći kernel funkcije za koje se dobija da je matrica  $K$  cirkularna, a zatim uzimajući u obzir osobinu da se u Furijeovom domenu cirkularna matrica može dijagonalizovati, rešenje za vektor  $\alpha$  u Furijeovom domenu dobija se kao:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda} \quad (4.71)$$

U (4.71)  $\hat{k}^{xx}$  predstavlja diskretnu Furijeovu transformaciju prvog reda matrice  $K$ , koja se i dobija cirkularnim pomeranjem  $k^{xx}$ ,  $K = C(k^{xx})$ . Stvarna priroda vektora  $k^{xx}$  jeste kernel autokorelacije (kernel korelacije vektora  $x$  sa samim sobom). Definisano za opšti slučaj dva proizvoljna vektora  $x$  i  $x'$ , kernel korelacija  $k^{xx'}$  jeste vektor čiji su elementi:

$$k_i^{xx'} = k(x', P^{i-1}x) = \varphi^T(x') \varphi(P^{i-1}x) \quad (4.72)$$

gde  $P^{i-1}$  označava permutacionu matricu definisanu u (4.52) [99]. Takođe, operacija deljenja u (4.71) je definisana kao deljenje element po element. Za razliku od (4.70) gde je potrebno računanje kernel matrice dimenzija  $N \times N$ , pri čemu se složenost povećava sa kvadratom broja uzoraka, u (4.71) je potrebno samo računanje autokorelacionog vektora čija se dimenzionalnost povećava linearno sa povećanjem broja uzoraka.

Kako bi za datu cirkularnu matricu  $C(x)$  i odgovarajuća kernel matrica  $K$  bila cirkularna, kernel funkcija mora tretirati sve dimenzije podataka podjednako [99]. Neke od kernel funkcija koje čuvaju cirkularnu strukturu su:

- Kerneli skalarnog proizvoda – linearni, polinomijalni

$$k(x, x') = (x^T x' + a)^b \quad (4.73)$$

$$k^{xx'} = (\mathcal{F}^{-1}(\hat{x}^* \odot \hat{x}') + a)^b \quad (4.74)$$

- Kerneli radijalne osnove – Gausovski kernel

$$k(x, x') = \exp\left(-\frac{1}{\sigma^2} \|x - x'\|^2\right) \quad (4.75)$$

$$k^{xx'} = \exp\left(-\frac{1}{\sigma^2} (\|x\|^2 + \|x'\|^2 - 2\mathcal{F}^{-1}(\hat{x}^* \odot \hat{x}'))\right) \quad (4.76)$$

- Aditivni kerneli – presek,  $\chi^2$
- Eksponencijalno aditivni kerneli

Dualni prostor takođe pruža mogućnost rada sa višestrukim kanalima, odnosno obeležjima poput HOG deskriptora ili obeležjima dobijenih iz modela dubokog učenja. Rad sa višestrukim kanalima se postiže jednostavnim sumiranjem po kanalima u Furijeovom domenu, usled linearnosti diskretne Furijeove transformacije. Ako se vektor  $x$  predstavi kao konkatencija individualnih vektora za  $L$  kanala,  $x = [x_1, x_2, \dots, x_L]$ , tada se po analogiji Gausovski kernel za višestruke kanale može dobiti kao:

$$k^{xx'} = \exp\left(-\frac{1}{\sigma^2} \left(\|x\|^2 + \|x'\|^2 - 2\mathcal{F}^{-1}\left(\sum_l \hat{x}_l^* \odot \hat{x}'_l\right)\right)\right) \quad (4.77)$$

Kako je prikazano na slici 4.1, detekcija objekta na novom frejmu video sekvence se vrši u okolini položaja objekta na prethodnom frejmu. Stoga, regresionu funkciju  $f(z)$  (4.48) je potrebno evaluirati za više testnih uzoraka  $z$  iz okoline položaja objekta na prethodnom frejmu. Ovi uzorci se mogu dobiti cikličnim pomeranjem osnovnog testnog uzorka okoline objekta. Tako se na osnovu kernelizovane korelacije  $k^{xz}$  između osnovnog uzorka (modela objekta)  $x$  i testnog uzorka  $z$  definiše cirkularna kernel matrica  $K^z$ :

$$K^z = \mathcal{C}(k^{xz}) \quad (4.78)$$

Kako bi se korelacioni odziv evaluirao za sve ciklične pomeraje tesnog uzorka, računa se regresiona funkcija (4.69) kao:

$$f(z) = (K^z)^T \alpha \quad (4.79)$$

Primenom metode dijagonalizacije cirkularnih matrica u Furijeovom domenu,  $f(z)$  se efikasnije može izračunati kao:

$$f(z) = \mathcal{F}^{-1}(\hat{k}^{xz} \odot \hat{\alpha}) \quad (4.80)$$

gde  $\mathcal{F}^{-1}$  označava operaciju inverzne Furijeove transformacije. Pozicija koja odgovara maksimalnoj vrednosti funkcije  $f(z)$  iz (4.80) predstavlja poziciju objekta koji se prati na posmatranom frejmu video sekvence.



#### 4.2.2.4 Višestruki kanali i ažuriranje filtra

Za rad sa višestrukim kanalima, odnosno za integraciju višedimenzionalnih obeležja, mogu se primeniti dva pristupa. U prvom pristupu se može primeniti konkatencija obeležja višestrukih kanala, a zatim se vrši obučavanje višekanalnog klasifikatora, odnosno diskriminativnog korelacionog filtra. Ako se uzorak  $x$  sastoji od više kanala obeležja  $x^d$ , pri čemu  $d$  označava indeks kanala,  $d \in \{1, 2, \dots, D\}$ , odnosno može se posmatrati da se na svakoj poziciji  $(m, n)$  uzorka  $x$  nalazi  $D$  - dimenzionalni vektor obeležja. Na isti način se filter  $w$  može proširiti da bude  $D$  - dimenzionalan, označavajući pojedinačne kanale filtra sa  $w^d$ . Tada se korelacija između dve višedimenzionalne funkcije definiše jednostavnim sumiranjem korelacionih odziva za svaki pojedinačni kanal. Filter  $w$  se može dobiti minimiziranjem kriterijumske funkcije:

$$J(w) = \sum_i \alpha_i \left\| \sum_d (x_i^d * w^d - y_i) \right\|^2 + \lambda \sum_d \|w^d\|^2 \quad (4.81)$$

gde se dodatno uvodi težinski koeficijent  $\alpha_i \geq 0$  za svaki uzorak  $x_i$ . Rešenje navedenog kriterijuma se u opštem slučaju ne može dijagonalizovati primenom diskretne Furijeove transformacije. Međutim, za slučaj obučavanja filtra samo jednim uzorkom, dobija se:

$$\hat{w}^d = \frac{\hat{x}^{d*} \odot \hat{y}}{\sum_d \hat{x}^{d*} \odot \hat{x}^d + \lambda} \quad (4.82)$$

Drugi pristup podrazumeva da se za svaki kanal  $d$  uzorka  $x$  izračuna korelacioni filter  $w^d$ , a zatim se agregacijom dobijenih filtera izračunava finalni korelacioni filter:

$$w = \sum_d \beta_d w^d \quad (4.83)$$

gde je  $\beta_d$  težinski koeficijent kojim se reguliše značaj svakog kanala višedimenzionalnih obeležja.

Kako bi se filter bolje prilagodio promenama u izgledu objekta tokom praćenja, ali istovremeno zadržao robusnost u odnosu na iznenadne velike promene, vrši se adaptivno ažuriranje modela izgleda objekta  $\hat{x}_t$  iz (4.62) prema:

$$\hat{x}_t = (1 - \gamma_t) \hat{x}_{t-1} + \gamma_t \hat{x} \quad (4.84)$$

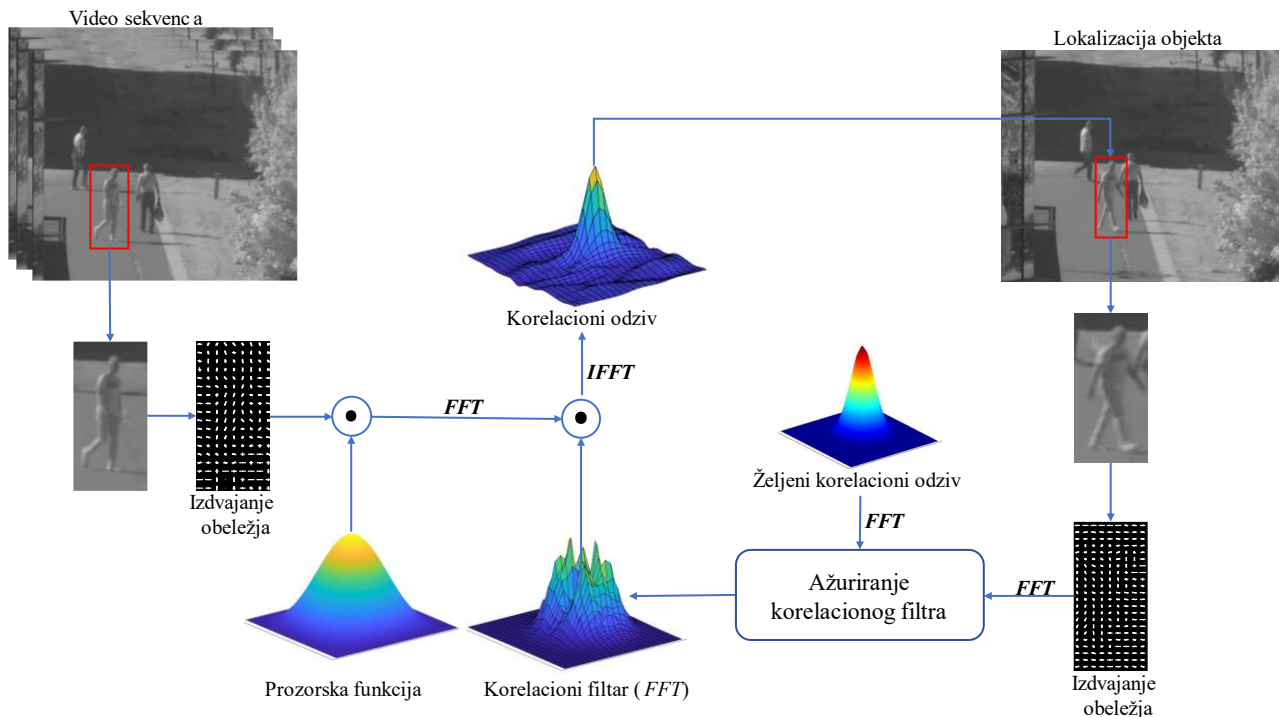
gde je  $t$  indeks trenutnog frejma,  $\hat{x}_{t-1}$  model izgleda objekta iz prethodnog frejma,  $\hat{x}$  uzorak obeležja objekta na trenutnom frejmu, a  $\gamma_t$  predstavlja stopu učenja. Analogno, ažuriranje koeficijenata filtra  $\hat{\alpha}_t$  iz (4.71) je dato kao:

$$\hat{\alpha}_t = (1 - \gamma_t) \hat{\alpha}_{t-1} + \gamma_t \hat{\alpha} \quad (4.85)$$

pri čemu  $\hat{\alpha}_{t-1}$  predstavlja vrednosti koeficijenata iz prethodnog frejma, a  $\hat{\alpha}$  estimaciju koeficijenata (4.71) na trenutnom frejmu.

### 4.2.3 Opšta struktura algoritma za praćenje korelacionim filtrima

Ako se pogleda osnovni princip rada korelacionih filtara i koraci u procesiranju frejmova, osnovna struktura algoritama za praćenje korelacionim filtrima se može predstaviti dijagramom prikazanim na slici 4.3.



**Slika 4.3** Tok procesiranja frejmova i osnovni strukturalni delovi algoritma za praćenje objekta zasnovanog na korelacionim filtrima.

Proces izvršavanja algoritma praćenja započinje inicijalizacijom, pri čemu se definiše okružujući okvir oko objekta od interesa za praćenje. Korelacioni filter se inicijalizuje uzorkom slike sa pozicije definisanog okružujućeg okvira objekta, ali većih dimenzija u odnosu na okružujući okvir objekta kako bi se u model filtra uključile i informacije pozadine i povećale diskriminativne sposobnosti filtra. Dalje se za svaki tekući frejm video sekvence procesiranje vrši uzorkovanjem segmenta slike sa pozicije koja predstavlja estimaciju pozicije praćenog objekta na prethodnom frejmu. Sa datog uzorka slike se vrši ekstrakcija obeležja koja na najbolji mogući način reprezentuju objekat koji se prati. Ekstrahovana obeležja predstavljaju osnovni uzorak za obučavanje korelacionog filtra. Kako se za obučavanje filtra koriste i negativni uzorci, a koji se zarad ubrzanja računskih operacija dobijaju cikličnim pomeranjima osnovnog uzorka, većina originalnog pozadinskog sadržaja je zamenjena sintetičkim cikličnim ponavljanjem originalnog uzorka. S druge strane, ciklična konvolucija uvodi i neželjene granične artefakte. Zbog izobličenja izazvanih cikličnim ponavljanjem osnovnog uzorka, izdvojeni sadržaj i obeležja su pouzdani samo u blizini centra uzorka slike. Zarad potiskivanja diskontinuiteta u graničnim regionima, uzorci se obično množe određenom prozorskom funkcijom, kao što je kosinusna prozorska funkcija. Za tako dobijeni reprezent objekta sa ublaženim graničnim efektima računa se brza Furejeova transformacija (eng. *Fast Furier Transform - FFT*). Između reprezentacije objekta i naučenog korelacionog filtra u Furijevom domenu se računa korelacija, a primenom inverzne brze Furijeove transformacije (eng. *Inverse Fast Furier Transform - IFFT*) dobija se izlazni korelacioni odziv u prostornom domenu. Pozicija pika u dobijenom

korelacionom odzivu predstavlja poziciju objekta na trenutnom frejmu. Sa ove pozicije se ekstrahuju obeležja koja se u skladu sa željenim korelacionim odzivom u Furijeovom domenu koriste za ažuriranje korelacionog filtra. Estimirana pozicija objekta na trenutnom frejmu se koristi i kao pozicija sa koje se u rekurzivnom postupku izvršavanja algoritma vrši ekstrakcija obeležja i detekcija objekta na novom frejmu video sekvence.

#### **4.2.4 Tehnike za savladavanje izazova u praćenju pokretnih objekata**

Korelacioni filtri, a posebno kernelizovani korelacioni filtri pokazuju veoma dobre performanse u praćenju objekata u uslovima bez okluzija, kada je objekat na relativno konstantnoj udaljenosti od kamere, i postižu impresivnu brzinu procesiranja frejmova za rad u realnom vremenu [99], [6]. Međutim, kada se veličina, orijentacija i izgled objekta menjaju, a posebno u uslovima okluzija, performanse praćenja se značajno smanjuju.

Analizom problema u video praćenju izloženih u sekciji 3.1, kao i dostupne literature o poboljšanjima algoritma za praćenje zasnovanim na kernelizovanim korelacionim filtrima, može se uočiti nekoliko glavnih pravaca unapređenja. Pristupi koji koriste karakteristike boja u RGB ili HSV domenu [146] – [149] ili duboka obeležja [150] – [152] (obeležja dobijena iz dubokih neuralnih mreža obučeni na kolor slikama) za modelovanje izgleda objekta nisu od interesa za istraživanje u ovoj disertaciji jer SWIR slika nema informacije o boji, a takođe drugačija je i sama priroda formiranja SWIR slike u odnosu na kolor sliku. Dodatno, usložnjavanje modela i povećavanje broja obeležja utiče na povećanje kompleksnosti algoritma, a posledično na smanjenje brzine procesiranja frejmova. Unapređenja korelacionih filtara za praćenje pokretnih objekata koja se mogu sresti u novijoj literaturi se mogu podeliti u pet kategorija: estimacija veličine objekta, ažuriranje i regularizacija filtra, detekcija okluzija, prevazilaženje okluzija i ponovna identifikacija objekta i primena prediktora. U nastavku ove sekcije će biti detaljnije analizirani pristupi svake od navedenih kategorija, kao i mogućnost primena ovih pristupa za unapređenje korelacionih filtara za praćenje objekata u SWIR video sekvenci.

##### **4.2.4.1 Estimacija veličine objekta**

Udaljavanje ili približavanje objekta kameri rezultira promenom veličine objekta u ravni slike. U svojoj osnovnoj konstrukciji, korelacioni filtri vrše procenu samo promene položaja objekta na slici, ali ne i procenu relativne promene veličine objekta. Izlazna korelaciona ravan ne pruža informacije o promenljivoj veličini objekta koji se prati. Ako model objekta ima fiksnu veličinu, u slučaju povećavanja veličine objekta, izdvojene karakteristike će biti nepotpune, dok se sa smanjenjem veličine objekta u model uvode i pozadinske karakteristike, što značajno može narušiti model izgleda objekta. Precizna procena veličine objekta pruža višestruku korist i direktno utiče na performanse algoritma za praćenje. Procena veličine objekta omogućava algoritmu za praćenje da prilagodi proceduru praćenja kako bi se adaptirao objektu koji se menja, čime se smanjuje mogućnost grešaka u praćenju ili potpunog gubitka objekta. Stoga, značajna pažnja prilikom dizajniranja korelacionih filtara za video praćenje treba da bude posvećena prilagođavanju korelacionog filtra varijacijama veličine objekta.

Jedan od prvih algoritama koji je razmatrao problem estimacije veličine objekta tokom praćenja jeste DSST (eng. *Discriminative Scale Space Tracker*) algoritam [100] metodom iscrpne pretrage. Komponenta za estimaciju veličine objekta se nadovezuje na estimaciju pomeraja objekta

na slici. Definiše se broj skaliranja  $S$ , pa se za svaki parametar skaliranja  $n \in \left\{-\frac{S-1}{2}, \dots, \frac{S-1}{2}\right\}$  sa slike uzorkuju segmenti dimenzija  $a^n w \times a^n h$ , gde je  $a$  faktor skaliranja koji ima vrednost 1.02, a  $w$  i  $h$  predstavljaju dimenzije objekta (širinu i visinu okružujućeg okvira) na prethodnom frejmu video sekvence. Svi uzorci su centrirani na estimiranoj poziciji objekta. Dizajniran je posebni 1D korelacioni filter za estimaciju veličine objekta, pa se kao željeni korelacioni odziv primenjuje 1D Gausova funkcija [100]. Kada se estimira parametar skaliranja objekta, odnosno veličina objekta, filter skaliranja se ažurira. U narednim frejmovima video sekvence, uzorak slike objekta se skalira u skladu sa trenutnim estimiranim parametrom skaliranja kako bi se procenio pomeraj objekta [100]. Takođe, analizirana je združena estimacija pomeraja objekta i veličine objekta obučavanjem 3D korelacionog filtra. Međutim, pokazalo se da je ovakav 3D korelacioni filter dosta sporiji, kao i da performanse ovakvog filtra nisu bolje od sekvencijalne primene estimacije pomeraja i estimacije veličine objekta, odnosno parametra skaliranja. Prateći metodologiju za estimaciju veličine objekta predloženu u DSST algoritmu, mnogo je pristupa u literaturi koji se oslanjaju na sekvencijalnu primenu estimacije pozicije i veličine objekta. U [153], skup skaliranja je definisan kao  $S = \{1, 0.985, 0.99, 0.995, 1.005, 1.010, 1.015\}$ . Na osnovu definisanog skupa skaliranja, segmenti slike se uzorkuju sa dimenzijama u skladu sa parametrima skaliranja definisanim u skupu  $S$ , a zatim se primenom bilinearne interpolacije skaliraju na originalnu veličinu, definisanu prilikom inicijalizacije algoritma, kako bi se mogao primeniti korelacioni filter koji je fiksne dimenzije. Nova veličina objekta se procenjuje na osnovu maksimalnog pika korelacionog odziva svih uzoraka slike. Isti pristup je primenjen u [147], gde je korišćen skup skaliranja  $S = \{0.98, 0.99, 1.0, 1.01, 1.02\}$ , nakon čega je zasebni klasifikator regularizovanih najmanjih kvadrata obučan na skaliranim uzorcima, kako bi se dobio korelacioni filter za procenu veličine objekta. U [154] je primenjen prostor za pretraživanje skalirajućeg parametra, po principu piramidalnog prostora skaliranja, sa brojem skaliranja 31. Optimalni parametar skaliranja u [155] se dobija korišćenjem 1-D korelacionog filtra sa brojem skaliranja  $S=29$ . Slično, a sa ciljem smanjenja broja računskih operacija, skup od tri parametara skaliranja  $S = \{0.9, 1.0, 1.1\}$  je korišćen u [156] za procenu promene veličine objekta u kombinaciji sa poboljšanim DSST algoritmom. Strategija ažuriranja DSST-a je poboljšana kako bi se izbegao problem preobučavanja modela. MKCF (eng. *Multi-kernel Correlation Filter*) algoritam predstavljen u [157] takođe nastoji da proceni parametar skaliranja na efikasniji način vršenjem pretrage u opsegu skaliranja od  $\pm 10\%$  u odnosu na trenutnu veličinu objekta. Dalje, radi ubrzanja procesa estimacije optimalne veličine objekta, MKCF za razliku od prethodnih pristupa vrši skaliranje izdvojenih karakteristika samo jednog uzorka objekta, umesto izdvajanja više uzoraka objekta različitih razmera. U problemu praćenja pešaka u [150], promena veličine objekata je estimirana primenom modela detekcije objekata zasnovanog na dubokom učenju - YOLOv3. U trenutnom frejmu video sekvence, veličina objekta koji se prati dobija se kombinovanjem predikcije okružujućeg okvira objekta iz kernelizovanog korelacionog filtra [99] i okružujućih okvira detektovanih pešaka YOLOv3 modelom koristeći IOU metriku. Za ažuriranje korelacionog filtra se uzima uzorak slike okružujućeg okvira YOLOv3 detekcije za koji je vrednost IOU metrike najveća.

#### 4.2.4.2 Detekcija okluzija

Jedna od često korišćenih metoda u literaturi za detekciju okluzija jeste odnos pika u izlaznom korelacionom odzivu i bočnog loba (eng. *Peak to Sidelobe Ratio* - PSR). Za izračunavanje ovog odnosa se koriste maksimalna vrednost, srednja vrednost i standardna devijacija korelacionog odziva.

U [158], [159] PSR vrednost se poredi sa definisanim pragom, gde PSR vrednost ispod vrednosti praga označava prisustvo okluzije. U [160], [161] PSR se u cilju detekcije okluzija poredi sa dve vrednosti praga. U [160], dva praga su izračunata na osnovu srednje PSR vrednosti,  $\mu_{PSR}$ , tokom prvih deset frejmova praćenja objekta. Za svaki frejm se izračunava razlika između PSR vrednosti u trenutnom frejmu i  $\mu_{PSR}$ , označena kao  $d_{PSR}$ . Zatim,  $d_{PSR} < 0.25\mu_{PSR}$  (donji prag) označava potpunu okluziju,  $d_{PSR} > 0.75\mu_{PSR}$  (gornji prag) označava regularno praćenje (bez prisustva okluzija), dok vrednost  $d_{PSR}$  između dva definisana praga sugerise na prisustvo delimičnih okluzija. Detekcija okluzija u [156] je zasnovana na kombinaciji razlike apsolutnih vrednosti pikova maksimalne i minimalne vrednosti (što daje informaciju o vibracijama u korelacionom odzivu) i kvadratne devijacije korelacionog odziva. Faktor okluzije definisan kao količnik ove dve veličine ima vrednost u opsegu od 10 do 100, pa ovako širok opseg vrednosti omogućava efikasnije određivanje nivoa okluzije. U radu [162] primenjen je čestični filter za procenu stanja okluzivnih uzoraka pozadine umesto procene stanja objekta, koristeći PSR i odnos preklapanja između objekta i uzoraka pozadine iz okruženja objekta. Još jedan čest pristup koji se koristi u literaturi, a koji može odražavati fluktuaciju korelacionog odziva i detektovati okluzije jeste prosečna energija korelacionog pika (*Average Peak Correlation Energy* - APCE). Velika vrednost parametra APCE ukazuje na regularne uslove praćenja, dok niska vrednost APCE ukazuje na prisustvo okluzija ili situaciju gubljenja objekta. U cilju detekcije okluzija u [163], APCE vrednost se razmatra zajedno sa vrednošću pika korelacionog odziva, a takođe se koristi i za ažuriranje modela objekta. Model se ažurira samo kada APCE vrednost premašuje definisani prag  $0.45\mu_{APCE}$ , a kada je vrednost APCE niža od  $0.39\mu_{APCE}$  smatra se da je objekat koji se prati izgubljen, pri čemu  $\mu_{APCE}$  označava srednju APCE vrednost. Prag koji se koristi za procenu prisustva okluzija primenom estimacije APCE vrednosti, kao i za prilagođavanje parametra brzine učenja korelacionog filtra je u [153] postavljen na fiksnu vrednost od 18.7. APCE vrednost normalizovana maksimalnom APCE vrednošću estimiranom tokom praćenja do trenutnog frejma video sekvence se u [148] poredi sa definisanim pragom iz skupa {0.15, 0.45, 0.65} u cilju detekcije okluzija. U [164], detekcija okluzija se posmatra kao klasifikacioni problem i rešava se korišćenjem metode nosećih vektora (eng. *Support Vector Machine* - SVM). Iz korelacionog odziva, segment slike veličine  $11 \times 11$  piksela se izdvaja centriran na poziciji objekta, zatim se vektorizuje, a tako dobijeni 121-dimenzionalni vektor obeležja se klasifikuje binarno sa ciljem određivanja pripadanja klasi okluzija ili ne. SVM je obučen nad skupom podataka VOT2017 koji obezbeđuje podatke o okluzijama. Detekcija okluzija je posmatrana i kao problem testiranja hipoteza [165]. Za trenutni frejm, test hipoteza se primenjuje za vrednost pika korelacionog odziva, koristeći srednju vrednost i standardnu devijaciju korelacionog pika, procenjenih na sekvenci od prethodnih 50 vrednosti pika. U [166], jedan od dva uslova se koristi za odlučivanje o prisustvu okluzija i problema prilikom praćenja. Prvi upoređuje odnos pika korelacionog odziva i prosečne vrednosti pika sa jednim pragom, dok drugi poredi odnos pika i prosečne vrednosti  $5 \times 5$  regiona korelacionog odziva oko pika, ne uključujući vrednost pika, sa drugim pragom. Oba praga su eksperimentalno određena i imaju vrednost 0.6. Da bi se utvrdilo da li je objekat pod okluzijom u [167], računato je Euklidsko rastojanje između detekcije objekta korelacionim filtrom i predviđenih koordinata objekta *Unscented Rauch-Tung-Striebel Smoother* metodom, a zatim upoređeno sa pragom dobijenim kao prosek odnosa pika i bočnih lobova na trenutnom i prethodnom frejmu.

#### 4.2.4.3 Regularizacija i ažuriranje korelacionih filtara

Ciklična konvolucija poboljšava računsku efikasnost, ali s druge strane uvodi i neželjene granične artefakte. Usled periodičnih efekata, većina originalnog pozadinskog sadržaja je zapravo zamenjena cirkularnim pomeranjem osnovnog uzorka slike. Kako se filtar na ovaj način obučava sa manje stvarnih pozadinskih uzoraka, njegova diskriminativna moć može biti ograničena. Zbog izobličenja izazvanih cikličnim ponavljanjem osnovnog uzorka, predviđeni rezultati su pouzdani samo u blizini centra uzorka slike, pri čemu je i veličina oblasti za pretragu ograničena. Kako je prikazano na slici 4.3, pre samog računanja korelacije između uzorka i naučenog filtra vrši se množenje uzorka datom prozorskom funkcijom kako bi se potisnuli diskontinuiteti u pograničnom regionu uzorka. Za dodatno rešavanje problema graničnih artefakata, pristup [101] uvodi prostornu regularizaciju korelacionog filtra. Komponenta prostorne regularizacije je uvedena u kriterijumsku funkciju za filtar višestrukih kanala (4.81) kao:

$$J(w) = \sum_i \alpha_i \left\| \sum_d (x_i^d * w^d - y_i) \right\|^2 + \lambda \sum_d \|f \cdot w^d\|^2 \quad (4.86)$$

Prostorna težinska funkcija  $f, f(k) > 0 \forall k$ , ima zadatak da penalizuje koeficijente filtra  $w^d(k)$  u zavisnosti od njihove prostorne lokacije. Kako bi se fokus filtra  $w$  stavio na region objekta, vrednosti funkcije  $f$  treba da budu veće na pozicijama piksela pozadine, a manje u regionu objekta. Međutim, za razliku od korelacionih filtara koji se obučavaju samo jednim uzorkom iz trenutnog frejma, formulacija sa više slika razbija strukturu cirkularne matrice, što rezultira značajnim smanjenjem računске efikasnosti i mogućnosti rada u realnom vremenu. Dodatno, optimizacija postaje još kompleksnija zbog prostorne regularizacije. Da bi se poboljšala efikasnost formulacije [101], predložen je pristup [107] koji koristi samo jedan uzorak slike za obučavanje filtra, a umesto toga, u kriterijumsku funkciju datu u (4.87) se pored prostorne regularizacije uvodi član  $\mu$  za vremensku regularizaciju kako bi se integrisale informacije iz prethodnih frejmova sekvence. U (4.87)  $w_t$  predstavlja korelacioni filtar u datom frejmu, dok  $w_{t-1}$  označava filtar iz prethodnog frejma.

$$J(w) = \frac{1}{2} \left\| \sum_d (x_t^d * w^d - y) \right\|^2 + \frac{1}{2} \sum_d \|f \cdot w^d\|^2 + \frac{\mu}{2} \|w_t - w_{t-1}\|^2 \quad (4.87)$$

Dok se u [101] i [107] koeficijenti filtra penalizuju izvan regiona objekta, u [168] se uvode stroža ograničenja i koeficijenti filtra izvan regiona objekta se postavljaju na vrednost 0. Formulacija kriterijumske funkcije (4.88) u ovom slučaju se može izraziti uvođenjem binarne maske  $P$ , čime se u potpunosti zanemaruju težine filtra koje odgovaraju pozadinskim karakteristikama.

$$J(w) = \frac{1}{2} \sum_i \|y_i - (P \cdot w) * x_i\|_2^2 + \frac{\lambda}{2} \sum_d \|w\|_2^2 \quad (4.88)$$

Kontekst okoline objekta koji se prati može imati značajan uticaj na performanse praćenja. Posebno u slučaju problema klatera, kontekst okoline je veoma važan za uspešno praćenje. U [169] se u svakom frejmu uzorkuje  $K$  kontekstnih uzoraka  $x_i$  u okolini objekta od interesa  $x_0$ . Odgovarajuće cirkularne matrice ovih uzoraka označene su sa  $X_i$  i  $X_0$ , respektivno. Uzorci okoline se

moгу posmatrati kao negativni uzorci za obučavanje korelacionog filtra. Na osnovu navedenog, cilj je da se obučí filter  $w$  koji rezultuje visokim korelacionim odzivom za uzorak objekta i odziv blizak nuli za uzorke konteksta pozadine. Stoga je u [169] predložena kriterijumska funkcija (4.89) u kojoj se uzorci konteksta okoline dodaju kao regularizator standardnoj formulaciji kriterijumske funkcije korelacionih filtara. Rezultat ove procedure je da uzorak objekta regresira kao u standardnoj formulaciji kontrolisan parametrom  $\lambda_1$ , dok se uzorci konteksta pozadine regresiraju na nule, što je kontrolisano parametrom  $\lambda_2$ .

$$J(w) = \|y - X_0 w\|_2^2 + \lambda_1 \|w\|_2^2 + \lambda_2 \sum_{i=1}^K \|X_i w\|_2^2 \quad (4.89)$$

Pored prostorne i vremenske regularizacije koje se primenjuju kako bi se umanjili granični artefakti ili regularizacije kontekstom zarad boljeg učenja negativnih uzoraka, adaptacija korelacionog filtra u uslovima scene se može postići i primenom različitih heuristika. U [153] se prati APCE vrednost, te u slučaju da ova vrednost bude manja od definisanog praga tokom određenog broja sukcesivnih frejmova, ažuriranje filtra novim uzorcima prestaje, odnosno parametar  $\gamma_t$  u (4.84) i (4.85) se postavlja na vrednost 0. U regularnim uslovima,  $\gamma_t$  ima konstantnu vrednost veću od 0. Parametar ažuriranja korelacionog filtra u [170] se adaptira u zavisnosti od brzine kretanja objekta u ravni slike, pri čemu se sa većom brzinom kretanja objekta i parametar ažuriranja filtra povećava. Princip na kome se temelji ovaj pristup jeste da objekat koji se kreće većom brzinom, brže i menja svoj izgled. U [171] se prati PSR vrednost, ali i razlika u PSR vrednosti između sukcesivnih frejmova. Oba parametra se porede sa adaptivnim pragovima koji se estimiraju tokom vremena na osnovu navedene dve vrednosti u poslednja tri frejma video sekvence. Model izgleda i koeficijenti korelacionog filtra se ažuriraju samo kada su obe vrednosti iznad svojih pragova, pri čemu se koristi konstantna stopa učenja  $\gamma_t$  u (4.84) i (4.85). PSR vrednost za estimaciju adaptivne stope učenja  $\gamma_t$  korišćenjem parametarske težinske funkcije je primenjena u [172]. U [160] se na osnovu razlike PSR vrednosti na trenutnom frejmu i srednje PSR vrednosti vrši odlučivanje o ažuriranju koeficijenata korelacionog filtra. Ukoliko je data razlika veća od definisanog praga, ažuriranje se vrši adaptivnom stopom datom kao parametrizovani količnik PSR vrednosti na trenutnom frejmu i srednje PSR vrednosti.

#### 4.2.4.4 Prevazilaženje okluzija i ponovna identifikacija objekta

U [164], ponovna identifikacija objekta se vrši detekcijom u okolini oblasti poslednjeg pojavljivanja praćenog objekta. Oblast pretrage za ponovnu detekciju se sastoji od 8 susednih preklapajućih prozora iste veličine kao i originalni prozor korelacionog filtra za pretragu. Pod prozorom za pretragu se podrazumeva uzorak slike objekta koji se evaluira korelacionim filtrom. Nova pozicija objekta se određuje iz onog prozora za pretragu čiji pik u korelacionom odzivu ima najvišu vrednost u odnosu na pikove u korelacionim odzivima ostalih prozora za pretragu. Tokom procesa ponovne identifikacije ne vrši se ažuriranje modela objekta. Za ponovnu identifikaciju u [173] je izvršena iscrpna pretraga u kombinaciji sa SVM klasifikatorom. Iscrpna pretraga podrazumeva testiranje svih mogućih prozora za detekciju objekta (uzoraka slike) u definisanoj okolini. Za svaki uzorak slike se ekstrahuje vektor HOG karakteristika, koji se klasifikuje primenom SVM-a. Estimirana lokacija uzorka slike koji je klasifikovan sa najvećom pouzdanošću, a uz uslov

da je vrednost pika korelacionog odziva tog uzorka slike veća od definisanog praga, se koristi kao nova pozicija objekta na datom frejmu. Za ponovnu identifikaciju objekata, SVM se koristi i u [148] i [174]. U [154], dualni onlajn SVM-ovi se obučavaju korišćenjem metoda “pasivno-agresivnog” učenja i onlajn ponovne obuke. Prvi SVM se koristi za grubu selekciju kandidata objekta, dok se drugi primenjuje za poboljšanje tačnosti ažuriranja detektovane pozicije objekta. Za problem ponovne identifikacije u [175] je primenjen *random fern* klasifikator. Od više uzoraka kandidata dobijenih nakon izvršavanja *random fern* klasifikatora, najbolji uzorak se nalazi izračunavanjem korelacije između uzoraka kandidata i skupa uzoraka koristeći klasifikator K-najbližih suseda i Euklidsko rastojanje. Na kraju, konačna odluka o tome da li odabrani uzorci predstavljaju objekat koji treba da se detektuje je zasnovana na kategorijama izdvojenih uzoraka. *Random fern* se takođe koristi za ponovnu identifikaciju objekata u [176]. Ovde se ulazni podaci za *random fern* klasifikator dobijaju uzorkovanjem cele slike pomoću iscrpne pretrage kliznim prozorom, a obuka klasifikatora se odvija istovremeno sa procesom praćenja. U [158], ponovna identifikacija objekta je primenjena nakon detekcije okluzije na osnovu vrednosti PSR-a. Mape korelacionih odziva za ponovnu detekciju i PSR vrednosti su računane za četiri prozora za pretragu oko pozicije objekta na trenutnom frejmu. Nova pozicija objekta se estimira na osnovu maksimalne PSR vrednosti iz ova četiri prozora, bez ažuriranja modela objekta dok je objekat pod okluzijom. Kada se detektuje okluzija u [166], više prozora za pretragu se postavlja oko pozicije gde je objekat izgubljen, bez ažuriranja modela izgleda objekta tokom procesa ponovne identifikacije. Kako bi se računarska kompleksnost smanjila, ovim prozorima za pretragu se daju određeni prioriteti. Usrednjavanjem vektora kretanja objekta, generiše se mapa korelacionog odziva prvog prozora za pretragu koji je pridružen regionu trenutnog vektora kretanja objekta, zanemarujući ostale prozore za pretragu. Višestruki prozori za pretragu se koriste i u [177], gde su prozori raspodeljeni na celom frejmu. Mapa korelacionog odziva se sračunava za svaki prozor, a za objekat se smatra da se nalazi u prozoru sa maksimalnom vrednošću pika korelacionog odziva. U [178], nakon otkrivanja problema u praćenju ili okluzija, objekat kandidat se generiše u oblasti ponovne identifikacije primenom normalizovane kros-korelacije. Zatim se primenjuje model korelacionog filtra fuzionisanih obeležja kako bi se potvrdilo da je to stvarni objekat. Autori u [179] koriste lokalnu i globalnu pretragu za ponovnu identifikaciju objekta. Metod lokalne pretrage treba da detektuje praćeni objekat u okolini poslednje pozicije pojave objekta, gde je klizni prozor dizajniran tako da vrši ciklično pomeranje u lokalnom regionu kako bi se postiglo veće uzorkovanje. Za globalnu pretragu koristi se *EdgeBox* algoritam, sa prihvatanjem 200 predloga sa visokim vrednostima pouzdanosti. Prebacivanje između lokalne i globalne pretrage je zasnovano na prosečnom pomeranju objekta u poslednjih 5 frejmova.

#### 4.2.4.5 Fuzija prediktora i korelacionih filtara

U [180], Kalmanov filter se koristi za predikciju pozicije objekta u trenutnom frejmu dok je objekat pod okluzijom, koristeći naučene informacije u prethodnim frejmovima. Slično, u [181], na osnovu predikcije Kalmanovog filtra, KCF algoritam vrši detekciju objekta i procenu njegove pozicije koja se dalje koristi u koraku estimacije Kalmanovog filtra. Dakle, tokom faze ažuriranja u KCF-u, ispravno stanje se primenjuje u ažuriranju modela izgleda objekta. U elektrooptičkom sistemu za praćenje u [182], u situaciji kada je razlika pozicija objekta između sukcesivnih frejmova veća od definisanog praga udaljenosti, u sledećem frejmu video sekvence Kalmanov filter je zadužen za procenu pozicije praćenog objekta. Usvojen je model objekta konstantne brzine. Kalmanov filter je

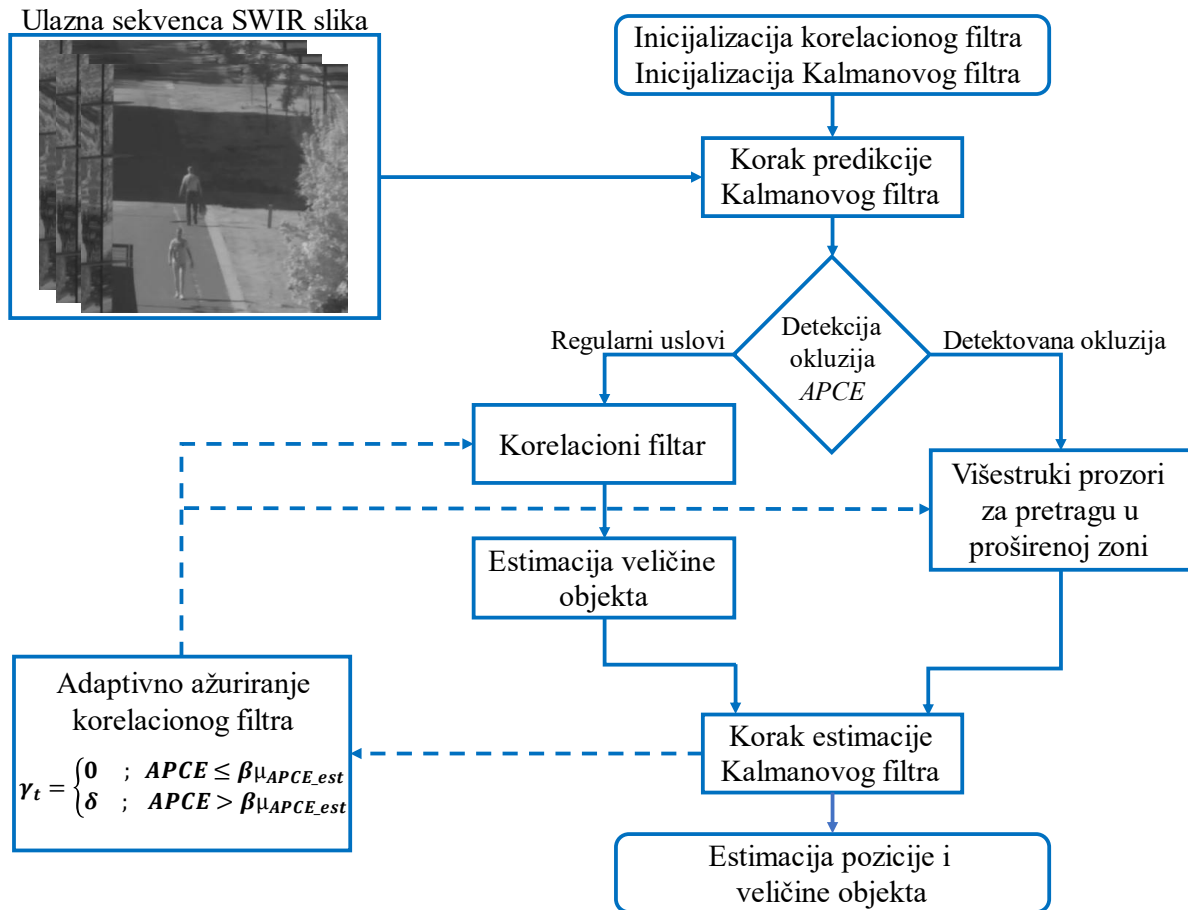


fuzionisan sa KCF algoritmom za transformaciju koordinatnog sistema u [183]. Ovde Kalmanov filter vrši procenu optimalnih 3D koordinata objekta, a zatim se pozicija objekta konvertuje u modifikator pozicije objekta u KCF algoritmu. U [161] je predložen metod predikcije i praćenja objekta za bespilotne letelice (eng. *Unmanned Aerial Vehicle*), gde se Kalmanov filter koristi za predikciju trajektorije objekta i kontrolu kretanja bespilotne letelice kada je objekat pod okluzijom ili izgubljen. Implementiran je standardni Kalmanov filter, model konstantne brzine. Kada se objekat ponovo pojavi, mereno vrednošću pika u korelacionom odzivu, sistem se ponovo oslanja na izlaz korelacionog filtra. Prošireni Kalmanov filter (eng. *Extended Kalman Filter* - EKF) je predložen u [160]. Kada dođe do potpune okluzije, koristeći dostupne informacije o kretanju pre pojave potpune okluzije, predloženi EKF daje predikciju pozicije praćenog objekta. Adaptivne jednačine stanja za procenu ubrzanja i brzine objekta, potrebne u EKF-u, koriste se za predikciju pozicije objekta. U [184], Kalmanov filter se koristi u kombinaciji sa čestičnim filtrom za prevazilaženje okluzija modelovanjem kretanja. U datom pristupu, čestični Kalmanov filter (eng. *Particle Kalman Filter*) koristi prednosti i Kalmanovog filtra i čestičnog filtra – postiže rad u realnom vremenu uz praćenje visoke dinamike objekta. Predikcija čestice stanja objekta se tretira kao položaj praćenog objekta.

### **4.3 Dizajn sistema za praćenje pokretnih objekata zasnovanog na korelacionim filtrima**

Predloženi sistem za praćenje pokretnih objekata zasnovan na korelacionim filtrima ima za cilj da obezbedi dugotrajno praćenje, odnosno kontinuirano praćenje objekta bez dodatnih manuelnih korekcija u izazovnim situacijama tokom praćenja. Još jedan od uslova za mogućnost dugotrajnog praćenja jeste postavljanje kamere na pan-tilt pozicioner koji će pozicionirati sistem tako da objekat konstantno bude centriran u vidnom polju kamere. Stoga je važno da algoritam za praćenje koji generiše kontrolne signale za upravljanje pan-tilt pozicionerom uspešno prati objekat od interesa u video sekvenci i prevazilazi izazove u video praćenju opisane u sekciji 3.2, poput okluzija, promene veličine objekta, te promene u dinamici kretanja. Takođe, pored dobrih performansi praćenja, posebno je važno da algoritam za praćenje nema veliku kompleksnost, odnosno da se može izvršavati u realnom vremenu.

Kako je cilj postići dugotrajno praćenje objekta, pa i u situacijama potpunih okluzija kada na raspolaganju nema nikakvih informacija o objektu, poboljšanja praćenja korelacionim filtrom su prvenstveno fokusirana na model kretanja objekta kako bi se u svakom trenutku mogao generisati upravljački signal pan-tilt pozicioneru. Na osnovu analize u sekciji 4.2.4, uvedeni su i moduli za dodatna poboljšanja u proceni veličine objekta, za detekciju okluzija i neregularnosti u praćenju, za prevazilaženje okluzija i ponovnu identifikaciju objekta nakon perioda pod okluzijom. Predloženi sistem za praćenje prikazan je blok dijagramom na slici 4.4.



**Slika 4.4** Blok dijagram predloženog sistema za praćenje pokretnih objekata u sekvenci SWIR slika.

U predloženom sistemu, korelacioni filtar je fuzionisan sa Kalmanovim filtrom kako bi se poboljšale performanse praćenja objekta. U koraku inicijalizacije, na prvom frejmu video sekvence, Kalmanov filtar se inicijalizuje istim podacima (pozicijom i dimenzijama objekta) kao i model objekta u korelacionom filtru. Na svakom sledećem frejmu, celokupan sistem za praćenje se oslanja na predikciju stanja objekta Kalmanovim filtrom. Centar prozora za pretragu objekta određuje se na osnovu predikcije pozicije objekta Kalmanovim filtrom, pri čemu su dimenzije (širina i visina) prozora za pretragu 3 puta veće od dimenzija samog objekta. Predikcija pozicije objekta Kalmanovim filtrom se procenjuje u bloku za detekciju neregularnosti u praćenju. Glavni zadatak ovog bloka jeste detekcija neregularnosti u praćenju uzrokovanih pojavom drifta, deformacijom objekta, kao i prisustvom okluzija što predstavlja najveći izazov. Takođe, blok za detekciju neregularnosti ima i važnu ulogu u određivanju moda rada sistema. U situacijama kada je praćenje regularno, sistem se oslanja samo na korelacioni filtar koji u tim uslovima pokazuje veoma dobre osobine. Implementiran je korelacioni filtar sa kontekstnom regularizacijom [169] koji minimizira kriterijumsku funkciju (4.89). Rešenje za korelacioni filtar  $w$  koji minimizira kriterijumsku funkciju (4.89) dato je u Furijeovom domenu kao:

$$\hat{w} = \frac{\hat{x}_0^* \odot \hat{y}}{\hat{x}_0^* \odot \hat{x}_0 + \lambda_1 + \lambda_2 \sum_{i=1}^K \hat{x}_i^* \odot \hat{x}_i} \quad (4.90)$$

Rešenje je dobijeno korišćenjem osobine dijagonalizacije cirkularnih matrica  $X_0$  i  $X_i$  iz (4.89) u Furijeovom domenu. U (4.90)  $\hat{x}_0$  predstavlja FFT osnovnog uzorka objekta  $x_0$ ,  $\hat{x}_i$  FFT uzoraka

konteksta  $x_i$  u okolini osnovnog uzorka objekta, a  $\hat{y}$  FFT željenog korelacionog odziva.  $K$  uzoraka okoline objekta,  $(x_1, x_2, \dots, x_K)$ , predstavljaju negativne uzorke za obučavanje korelacionog filtra što značajno utiče na poboljšanje diskriminativnih sposobnosti korelacionog filtra. Operator  $\odot$  u (4.90) označava množenje element po element, pri čemu je i deljenje u (4.90) definisano na isti način. Primećeno je da uvođenje dodatnih regularizacija prikazanih u sekciji 4.2.4.3 poput prostorne i prostorno-vremenske regularizacije ne poboljšava značajno performanse filtra u praćenju objekta u SWIR video sekvenci, a umnogome povećava računsku kompleksnost i vreme procesiranja frejmova. Parametar  $\lambda_1$  kontroliše regularizaciju osnovnog uzorka, dok se uzorci konteksta pozadine regresiraju na nule, što je kontrolisano parametrom  $\lambda_2$ .

U koraku lokalizacije objekta na novom frejmu, ekstrahuju se obeležja - histogram orijentisanih gradijenata (HOG), koja reprezentuju uzorak objekta  $z$  koji treba da se evaluira, odnosno primenjuje se šema praćenja korelacionim filtrima prikazana na slici 4.3. Nakon primene *Hanning* prozorske funkcije na uzorak obeležja,  $z$ , računa se korelacija između naučenog filtra  $\hat{w}$  iz (4.90) i FFT-a obeležja uzorka slike  $\hat{z}$ . Korelacioni odziv se dobija kao:

$$f(z) = \mathcal{F}^{-1}(\hat{w} \odot \hat{z}) \quad (4.91)$$

Pozicija pika u korelacionom odzivu predstavlja procenjenu poziciju objekta na novom frejmu. Dalje se u levoj grani sistema na korelacioni filter nadovezuje blok za estimaciju veličine radi što bolje procene dimenzija objekta.

Sa druge strane, ukoliko se tokom praćenja detektuje poremećaj, a posebno ako je detektovani poremećaj uzrokovan okluzijom, oslanjajući samo na korelacioni filter tehnički je nemoguće nastaviti praćenje ako je objekat zaklonjen potpunom okluzijom ili izađe iz vidnog polja kamere. U mnogim scenarijima u praksi, pokazuje se da objekti tokom određenog perioda imaju konstantnu brzinu kretanja, pa se može pretpostaviti da će se objekat ponovo pojaviti na sceni nakon perioda pod okluzijom. Stoga, Kalmanov filter ima zadatak da obezbedi predikciju oblasti za ponovnu detekciju objekta, prevazilazeći na taj način problem praćenja samo korelacionim filtrom, koji ostaje zaglavljn na poziciji pojave okluzije. Međutim, u nekim slučajevima čak i uz implementaciju prediktivnog algoritma, ponovna detekcija objekta može biti nemoguća, budući da objekat može izvršiti manevar tokom perioda pod okluzijom, pa nakon izlaska iz okluzije može biti izvan predviđene oblasti za pretragu. Stoga, ukoliko se u bloku za detekciju poremećaja detektuje prisustvo okluzije, tok signala prati blok za detekciju objekta u proširenoj zoni primenom višestrukih prozora za pretragu u kombinaciji sa predikcijom Kalmanovog filtra. Pretraga je potrebna samo u određenom broju prozora oko centralnog, na koji se oslanja i bazični korelacioni filter. Kalmanov filter vrši predikciju dinamike kretanja objekta, čime se smanjuje vreme obrade koje bi bilo potrebno za pretragu na celoj slici. Iako primena višestrukih prozora za pretragu uvodi dodatno kašnjenje u sistem, prebacivanje iz režima rada bazičnog korelacionog filtra i režima primene bloka za pretragu u proširenoj zoni, a koji se primenjuje samo kada se detektuje okluzija, celokupan sistem i dalje ima prosečno vreme izvršavanja dovoljno za rad u realnom vremenu.

Iz grane oslonjene na bazični korelacioni filter ili iz grane sa pretragom u proširenoj zoni primenom višestrukih prozora za pretragu, detektovana pozicija i procenjena veličina objekta koriste se kao merenje za korak estimacije Kalmanovog filtra. Izlaz iz bloka estimacije Kalmanovog filtra predstavlja estimaciju pozicije i veličine praćenog objekta na trenutnom frejmu. Na ovaj način se obezbeđuje da sistem neće ostati zaglavljn na poziciji okluzija. Dinamika pokretnog objekta će biti

ispraćena, a kontrolni signali će upravljati pan-tilt pozicionerom tako da objekat ostane u vidnom polju kamere i nakon perioda pod okluzijom.

Na estimiranoj poziciji objekta i sa estimiranim dimenzijama uzorkuje se segment slike za ažuriranje korelacionog filtra, kao i 4 kontekstna uzorka pozadine. Izračunava se korelacioni filter  $\hat{w}$  za dati frejm prema (4.90), a zatim se vrši ažuriranje filtra prema:

$$\hat{w}_t = (1 - \gamma_t)\hat{w}_{t-1} + \gamma_t\hat{w} \quad (4.92)$$

gde  $\gamma_t$  predstavlja interpolacioni faktor ili stopu učenja, a  $\hat{w}_{t-1}$  korelacioni filter u prethodnom frejmu.

U nastavku će biti detaljnije opisani model kretanja objekta, moduli za estimaciju veličine objekta, za detekciju poremećaja i okluzija, kao i modul za ponovnu identifikaciju objekta.

- **Estimacija stanja objekta**

Pored praćenja pozicije objekta, za uspešno praćenje objekta, važna je i procena veličine objekta. Kako bi se aproksimirale promene u poziciji, ali i u veličini objekta između sukcesivnih frejmova, usvojen je linearni model konstantne brzine, pri čemu se stanje objekta modeluje kao:

$$X = \begin{bmatrix} x_c \\ y_c \\ w \\ h \\ \dot{x}_c \\ \dot{y}_c \\ \dot{w} \\ \dot{h} \end{bmatrix} \quad (4.93)$$

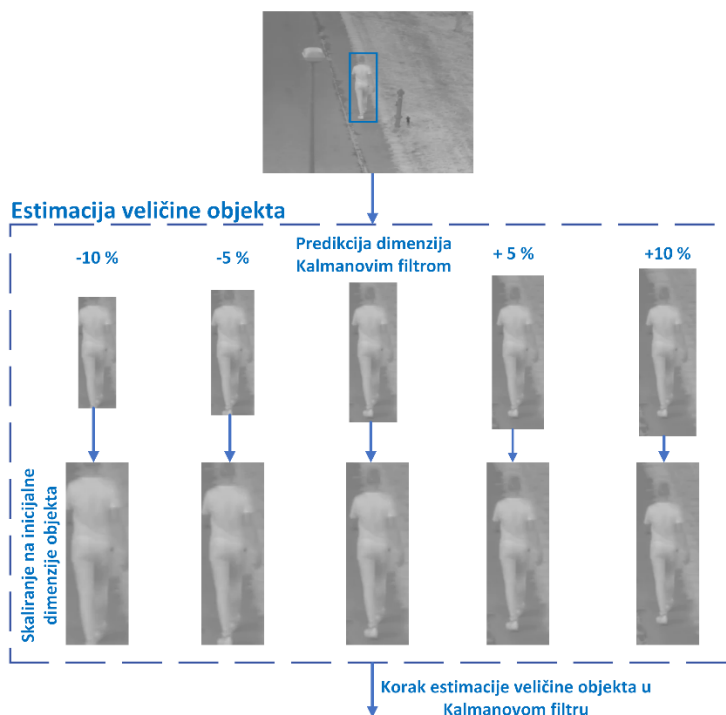
gde  $x_c$  i  $y_c$  predstavljaju poziciju centra objekta u ravni slike u horizontalnom i vertikalnom pravcu, respektivno. Treći element u vektoru stanja objekta,  $w$ , označava širinu okružujućeg okvira objekta, a  $h$  visinu okružujućeg okvira. Poslednje četiri komponente vektora stanja u (4.93) predstavljaju prve izvode, odnosno brzine, prve četiri komponente vektora stanja. Komponenta vektora stanja i njena brzina povezane su modelom konstantne brzine. Za razliku od modela koji se mogu sresti u literaturi o fuziji prediktora i korelacionih filtara, gde se najčešće modeluje samo pozicija objekta, pa se tako u slučaju okluzija može vršiti samo predikcija pozicije objekta, navedeni model pruža mogućnost i estimacije veličine objekta u situacijama kada se o tome ne mogu dobiti nikakve informacije iz slike. Posebno važno je to što model definisan na ovaj način pruža mogućnost procene kako pozicije tako i veličine objekta u regularnim uslovima, ali i tokom kretanja pod okluzijama, pa nakon perioda okluzije sistem ponovnu identifikaciju objekta može započeti sa veličinom objekta koja približnije odgovara njegovoj stvarnoj veličini.

- **Estimacija veličine objekta**

U osnovnoj verziji korelacionih filtara, veličina praćenog objekta je konstantna, odnosno dimenzije okružujućeg okvira objekta su iste kao na prvom frejmu video sekvence. Sa fiksnim dimenzijama okružujućeg okvira, ekstrahovana obeležja su nepotpuna ako se veličina objekta

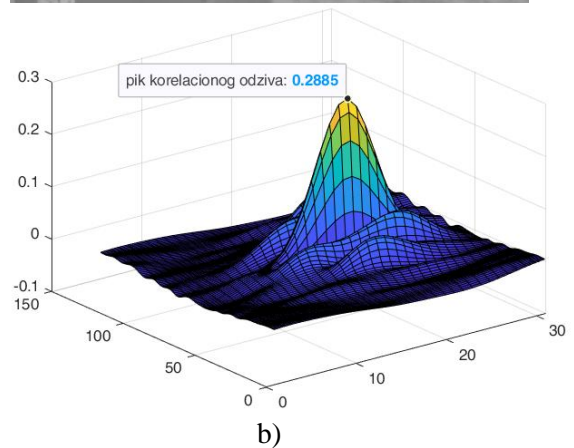
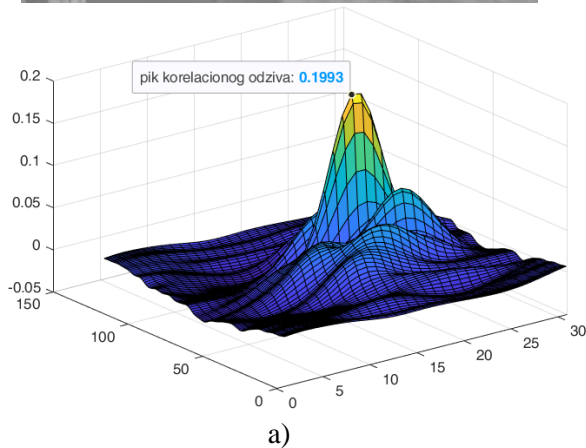
povećava. Sa druge strane, ukoliko se veličina objekta smanjuje, promenljiva obeležja pozadine će biti uvedena u model izgleda objekta i narušiti koeficijente korelacionog filtra.

Kako bi se ispratili varijacije veličine objekta, na trenutnom frejmu video sekvence sistem se prvo oslanja na predikciju dimenzija objekta Kalmanovim filtrom. Vektor stanja definisan u (4.93) pruža mogućnost procene veličine objekta i tokom kretanja praćenog objekta pod okluzijom, što je posebno važno za ponovnu identifikaciju objekta nakon okluzije sa njegovom tačnom veličinom. Pristup da se sistem prvo oslanja na predikciju Kalmanovog filtra je usvojen zbog situacija kada se o objektu ne mogu dobiti nikakve informacije sa slike. Kako je za proračun korelacionog odziva korelacijom uzorka objekta sa trenutnog frejma i filtra potrebno da trenutni uzorak bude dimenzionalno isti kao i filter, definisan je skup skaliranja  $S = \{-10\%, -5\%, +5\%, +10\%\}$ . Pored ekstrahovanog uzorka slike na osnovu predikcije dimenzija objekta Kalmanovim filtrom, ekstrahuju se još četiri uzorka. Ovi uzorci su centrirani na poziciji predviđenoj Kalmanovim filtrom i imaju dimenzije koje su relativne u odnosu na dimenzije predviđene Kalmanovim filtrom prema procentima skaliranja u skupu  $S$ . Ekstrahovani uzorci se skaliraju na inicijalne dimenzije objekta (dimenzije definisane na prvom frejmu), nakon čega se vrši ekstrakcija obeležja. Dimenzije uzorka sa najvećom vrednošću pika korelacionog odziva u korelaciji sa trenutnim modelom objekta, odnosno korelacionim filtrom, se koriste kao trenutna estimacija veličine objekta. Dalje se ove dimenzije primenjuju u koraku estimacije Kalmanovog filtra. Postupak estimacije veličine objekta je prikazan na slici 4.5.



**Slika 4.5** Estimacija veličine objekta.

Na slici 4.6 su prikazani primeri praćenja sa i bez estimacije veličine praćenog objekta. Jasno se može uočiti da je vrednost pika korelacionog odziva kada se koristi estimacija veličine objekta znatno veća u poređenju sa onom dobijenom bez estimacije veličine. Dodatno, može se videti da su bočni lobovi u korelacionom odzivu bez estimacije veličine objekta znatno izraženiji.



**Slika 4.6** a) Praćenje objekta bez estimacije veličine. Vrednost pika korelacionog odziva: 0.1993.  
b) Praćenje objekta uz estimaciju veličine. Vrednost pika korelacionog odziva: 0.2885.

- **Detekcija okluzija i neregularnosti u praćenju**

U svojoj osnovnoj formi, korelacioni filtri nisu u mogućnosti da detektuju neregularnosti u praćenju i situacije gubitka objekta. Kada se objekat izgubi ili bude zaklonjen okluzijom, kao objekat od interesa će se nadalje pratiti pozadina na poziciji gubitka objekta. Čak i kada je pouzdanost izuzetno niska, model izgleda objekta će biti ažuriran obeležjima pozadinskih struktura, što rezultira neuspehom praćenja. Međutim, dobijena mapa korelacionog odziva može se iskoristiti za detekciju neregularnosti u praćenju.

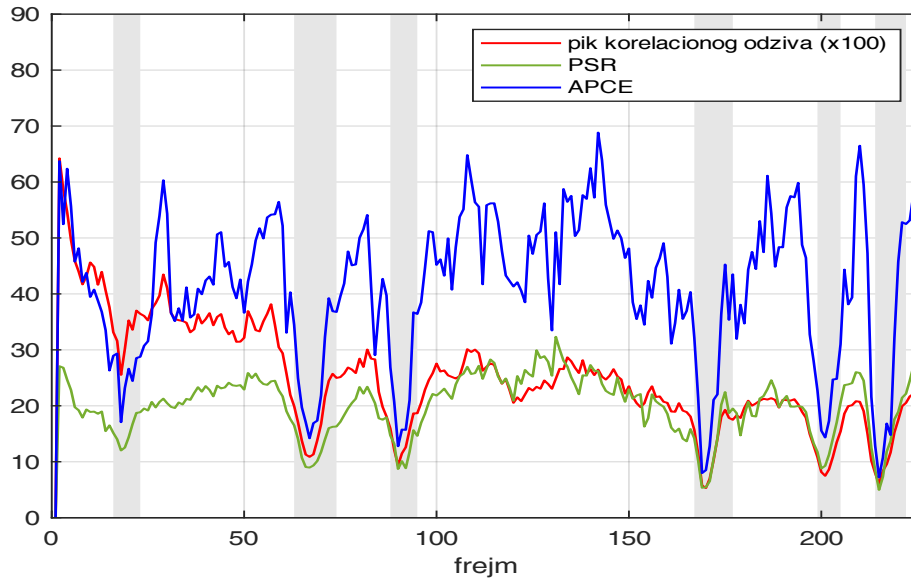
U predloženom sistemu za praćenje kao parametar za detekciju okluzija i neregularnosti u praćenju primenjena je prosečna energija korelacionog odziva (eng. *Average Peak to Correlation Energy* - APCE). Za dati korelacioni odziv  $g$ , APCE se izračunava kao:

$$APCE = \frac{|g_{max} - g_{min}|^2}{\frac{1}{MN} \sum_{m,n} (g(m,n) - g_{min})^2} \quad (4.94)$$

Gde  $g_{max}$  predstavlja vrednosti pika, odnosno maksimalnu vrednost korelacionog odziva,  $g_{min}$  minimalnu vrednost korelacionog odziva, a  $M$  i  $N$  prostorne dimenzije korelacionog odziva. Kretanje APCE vrednosti tokom praćenja objekta u scenariju sa okluzijama prikazano je na slici 4.7. Osenčeni delovi na slici predstavljaju period tokom kog je objekat pod okluzijom. Poređenje je dato sa kretanjem pika korelacionog odziva i PSR vrednosti. PSR vrednost je računata prema:

$$PSR = \frac{g_{max} - \mu}{\sigma} \quad (4.95)$$

gde  $g_{max}$  predstavlja pik korelacionog odziva,  $\mu$  srednju vrednost bočnog loba, a  $\sigma$  označava standardnu devijaciju bočnog loba. Bočni lob je definisan kao deo korelacionog odziva izvan regiona od  $10 \times 10$  piksela oko pozicije pika. Sa slike 4.7 se vidi da APCE ima najveće razlike u vrednosti tokom regularnih uslova praćenja i onih prilikom pojave okluzija, pa se na osnovu APCE vrednosti najbolje može razlučiti kada nastupa okluzija.



**Slika 4.7** Poređenje vrednosti pika korelacionog odziva (skaliran faktorom 100 radi vizualizacije), PSR i APCE vrednosti u scenariju sa okluzijama iz sekvence 5 iz skupa SWIR video sekvenci.

APCE vrednost prvenstveno ima ulogu u promeni moda rada sistema između grane koja se oslanja samo na korelacioni filter za praćenje i grane koja koristi višestruke prozore za pretragu u slučaju pojave okluzija, kao što je prikazano na slici 4.4. Kada je objekat pod okluzijom, vrednost APCE će se brzo smanjiti u poređenju sa vrednostima u regularnim uslovima. Posmatra se APCE vrednost u odnosu na prosečnu APCE vrednost  $\mu_{APCE}$  skaliranu parametrom  $\lambda$ . APCE vrednost manja od  $\lambda\mu_{APCE}$  ukazuje na to da je objekat pod okluzijom i da bi trebalo preći u režim rada gde se primenjuju višestruki prozori za pretragu. Sa APCE vrednošću iznad  $\lambda\mu_{APCE}$ , praćenje se može oslanjati samo na korelacioni filter. Za parametar  $\lambda$  je eksperimentalno usvojena vrednost 0.6.

$$\text{detektovana okluzija} = \begin{cases} 1 & ; APCE \leq \lambda\mu_{APCE} \\ 0 & ; APCE > \lambda\mu_{APCE} \end{cases} \quad (4.96)$$

- **Višestruki prozori za pretragu i ponovnu identifikaciju objekta**

Nakon detekcije okluzija tokom praćenja, aktivira se pretraga objekta u proširenoj oblasti koristeći višestruke prozore za pretragu. Ilustrativni primer je prikazan na slici 4.8. Oslanjanje na predikciju pozicije objekta Kalmanovim filtrom i primenu samo jednog prozora za pretragu može lako dovesti do gubitka objekta pod potpunom okluzijom. Dodatno, ako objekat vrši manevar pod okluzijom, samo jedan prozor za pretragu nije dovoljan i praktično je nemoguća tačna ponovna identifikacija izgubljenog objekta nakon perioda pod okluzijom. Pretragom u proširenoj zoni oko



fiksne pozicije na kojoj je objekat izgubljen, a u slučaju dugotrajne okluzije, ponovna detekcija takođe može biti nepouzdana. Sa predikcijom kretanja objekta Kalmanovim filtrom i primenom višestrukih prozora za pretragu oko predviđene pozicije objekta, ovaj problem se može efikasno rešiti. Kalmanov filter vrši procenu dinamike kretanja objekta, dok višestruki prozori pružaju veći prostor za pretragu kako bi se objekat detektovao i u slučaju manevrisanja tokom okluzije, što rezultira ponovnom detekcijom objekta, ažuriranjem njegovog modela izgleda i nastavkom praćenja nakon okluzije.

Kao što je prikazano na slici 4.8, centralni prozor za pretragu je onaj čiju predikciju pozicije i dimenzija vrši Kalmanov filter. Ostalih osam prozora za pretragu imaju istu veličinu kao centralni prozor i pozicionirani su u okolini centralnog prozora. Kako bi se izbegao problem da se delovi istog objekta nalaze u susednim prozorima u trenutku detekcije, uvedeno je međusobno preklapanje prozora. Preklapanje između susednih prozora iznosi 1/3 po svakoj dimenziji. Svaki od prozora za pretragu se evaluira korelacijom sa trenutnim korelacionim filtrom,  $\hat{w}_t$  iz (4.92), čime se generiše mapa korelacionog odziva za svaki od prozora za pretragu. Detekcije dobijene tokom perioda okluzije uglavnom nisu pouzdane, te ovde nije primenjena estimacija veličine, odnosno sistem se oslanja samo na predikciju dimenzija objekta Kalmanovim filtrom. Detekcija pozicije objekta na trenutnom frejmu zasniva se na vrednosti pikova korelacionih odziva. Na osnovu pozicije pika sa maksimalnom vrednošću vrši se estimacija pozicije objekta. Estimirana pozicija, kao i veličina objekta se dalje koriste u koraku estimacije Kalmanovog filtra za dati frejm.



**Slika 4.8** a) Praćenje objekata u regularnim uslovima primenom jednog prozora za pretragu (— objekat koji se prati, — prozor za pretragu); b) mehanizam višestrukih prozora za pretragu se aktivira nakon detekcije okluzije (— estimacija objekta, — predikcija centralnog prozora za pretragu Kalmanovim filtrom, — okolnih 8 prozora za pretragu sa međusobnim preklapanjem).

Dodatno, kako se model izgleda objekta i koeficijenti korelacionog filtra ne bi narušili tokom perioda pod okluzijom i izbeglo da se obeležja okluzije uvrste u model izgleda objekta, što bi rezultovalo zaustavljanjem praćenja na poziciji okluzije, tokom procesa ponovne identifikacije objekta vrši se adaptivno ažuriranje modela izgleda objekta. APCE vrednost korelacionog odziva finalne procene se koristi za odlučivanje o ažuriranju modela izgleda objekta, odnosno korelacionog filtra. Adaptacija interpolacionog faktora  $\gamma_t$  iz (4.92) se vrši prema:

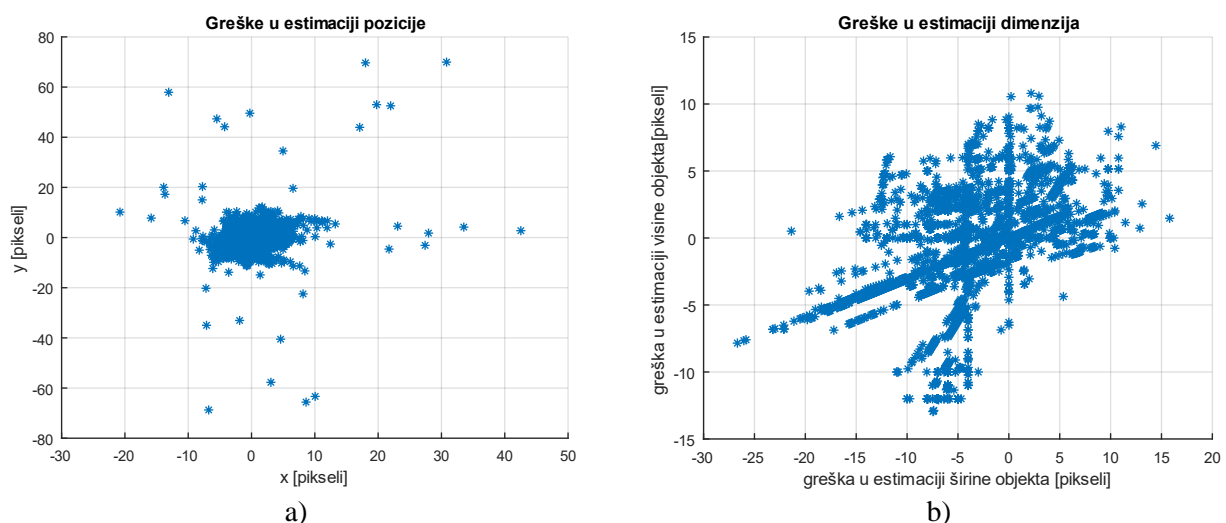
$$\gamma_t = \begin{cases} 0 & ; APCE \leq \beta \mu_{APCE\_est} \\ \delta & ; APCE > \beta \mu_{APCE\_est} \end{cases} \quad (4.97)$$



APCE vrednost korelacionog odziva finalne procene veća od srednje APCE vrednosti finalnih procena  $\mu_{APCE\_est}$  skalirane parametrom  $\beta$  označava da je detekcija objekta na trenutnom frejmu pouzdana, te će se model izgleda objekta ažurirati ispravnim obeležjima. Eksperimentalno je usvojena vrednost parametra  $\beta$  koja iznosi 0.583. Vrednost interpolacionog faktora  $\delta$  iznosi 0.015 [169].

### 4.3.1 Statistička analiza rezultata rada predloženog sistema za praćenje

Za potrebe statističke analize performansi predloženog sistema, iz kreiranog skupa SWIR video sekvenci predstavljenog u sekciji 2.3 korišćeno je 4243 labeliranih frejmova iz 9 nasumično odabranih sekvenci. Analizira se ponašanje sistema u različitim scenarijima praćenja, gde je objekat od interesa za praćenje pešak u pokretu. Referentni podaci su predstavljeni manuelno označenim okružujućim okvirima oko objekta na svakom frejmu, odnosno pozicijom centra pokretnog objekta i njegovim dimenzijama (visinom i širinom). Greška u proceni pozicije objekta i greška u proceni dimenzija objekta analiziraju se u nastavku. Za svaki frejm, greška u poziciji je računata kao razlika između stvarne (*ground-truth*) pozicije i pozicije estimirane predloženim sistemom za praćenje. Greška u dimenzijama je podeljena u dve kategorije: greška procene visine i greška procene širine, takođe merene kao razlika između *ground-truth* dimenzija i onih procenjenih predloženim sistemom za praćenje. Kako bi se analizirao rad sistema na celokupnim izdvojenim sekvencama, kako u regularnim uslovima gde je praćenje uspešno, tako i u svim izazovnim scenarijima, sistem za praćenje je ponovno inicijalizovan svaki put nakon gubitka objekta i prestanka praćenja usled grešaka različitih priroda.

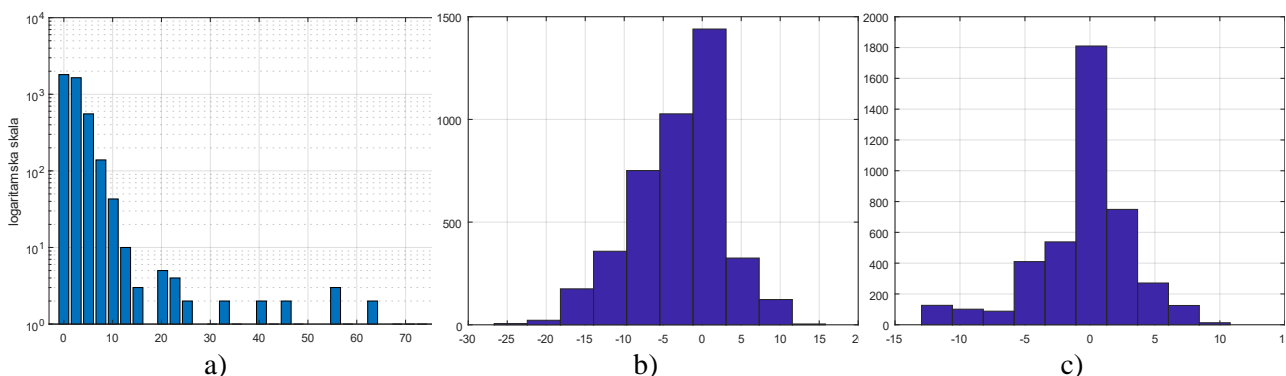


**Slika 4.9** a) Greške u estimaciji pozicije objekta; b) Greške u estimaciji dimenzija objekta.

Slika 4.9 a) daje prikaz greške predloženog sistema u estimaciji pozicije objekta, u vertikalnom i horizontalnom pravcu u ravni slike. Na slici 4.9 b) su prikazane greške u proceni dimenzija objekta. Greške u proceni širine objekta su prikazane po horizontalnoj, a greške u proceni visine po vertikalnoj osi. Može se videti da se većina populacije grešaka procene dimenzija, kako po širini tako i po visini, nalazi unutar jednog klastera. Takođe, većina populacije grešaka u estimaciji pozicije pripada jednom klasteru. Međutim, sa slike 4.9 a) jasno se može uočiti prisustvo grešaka koje značajno odstupaju od većine populacije u klasteru. Ovaj tip grešaka je posebno važno detaljnije analizirati, jer greške velikog intenziteta mogu značajno da poremete sistem i dovedu do gubitka praćenja objekta, ali i gubitka prostora za pretragu i ponovnu detekciju.

Na slici 4.10 a) je prikazan histogram grešaka u estimaciji pozicije. Greške u estimaciji pozicije za prikaz histogramom su računane kao Euklidsko rastojanje između stvarne pozicije objekta i pozicije objekta procenjene predloženim sistemom za praćenje. Da bi se naglasile greške koje dovode do gubitka objekta i prestanka praćenja, histogram je prikazan u logaritamskoj skali. Odavde se odstupanja jasno mogu videti kao repovi u histogramu sa izuzetno visokim vrednostima greške. Broj ovih grešaka nije veliki, ali su njihove vrednosti izuzetno velike u poređenju sa ostatkom populacije, što dovodi do potpunog gubitka objekta i prestanka praćenja. Stoga, ovaj tip grešaka ukazuje na postojanje loših merenja koja se mogu tretirati kao autlajeri [23].

U skladu sa slikom 4.10 a), u prikazu raspodele histogramom, većina populacije ukazuje na postojanje grešaka koje su posledica odstupanja estimirane pozicije objekta od stvarne pozicije, što povećava varijansu praćenja, ali ne dovodi do prekida praćenja. Priroda ovih grešaka proizilazi iz promena u izgledu objekta tokom praćenja, iz promena u dinamici kretanja, kao i zbog promena kako smeru kretanja tako i orijentacije objekta. Takođe, čest je problem nedovoljnog kontrasta između objekta i pozadine na SWIR slici, što dodatno otežava preciznu lokalizaciju objekta na slici. Iz istih razloga dolazi i do pojave grešaka u proceni dimenzija objekta, ali raspodela ovih grešaka nije kontaminirana autlajerima, što se može videti i iz histograma na slikama 4.10 b) i c).



**Slika 4.10** Histogrami grešaka u praćenju predloženog sistema: a) greške u estimaciji pozicije (prema Euklidskom rastojanju), b) greške u estimaciji širine objekta, c) greške u estimaciji visine objekta.

Posmatranjem skupa analiziranih sekvenci i situacija u kojima se javljaju autlajeri u greškama estimacije pozicije, može se zaključiti da su autlajeri rezultat pojave delimičnih ili potpunih, kratkotrajnih i dugotrajnih okluzija. U situacijama kada dođe do okluzija, osnovni korelacioni filter na koji se sistem za praćenje oslanja će ostati zaglavljen na poziciji pojave okluzije i neće moći dalje da prati objekat. Dodavanje prediktora, kao što je standardni Kalmanov filter, omogućava da se praćenje objekta nastavi čak i nakon okluzije, ali samo ako objekat ne vrši manevar ili ne menja dinamiku kretanja tokom perioda pod okluzijom. Primenom višestrukih prozora za pretragu u kombinaciji sa predikcijom standardnog Kalmanovog filtra, moguće je prevazići okluzije i ispratiti potencijalne manevre objekta. Međutim, primena višestrukih prozora za pretragu obezbeđuje šire područje u kom se objekat može detektovati, ali daje prostora i za pojavu velikih grešaka u estimaciji pozicije. U slučaju velike greške u estimaciji pozicije koja će se kroz upravljačke signale propagirati do pan-tilt pozicionera, u realnim aplikacijama sistem koji se nalazi na pan-tilt pozicioneru na taj način može dovesti centar vidnog polja kamere u pogrešnu poziciju i tako trajno izgubiti objekat, kao i prostor u kome bi se objekat potencijalno mogao ponovo detektovati.

Stoga, važno pitanje je kako razlikovati greške u estimaciji pozicije objekta uzrokovane promenama u izgledu objekta i dinamici kretanja objekta usled manevara, od onih koje su nastale

usled pogrešnih merenja odnosno autlajera, te znati kako koje merenje tretirati i postići balans između efikasnosti i robusnosti dizajniranog sistema za praćenje.

Prema razmatranju tehnika fuzije prediktora sa korelacionim filtrima u sekciji 4.2.4.5, standardni Kalmanov filter je najčešće korišćeni prediktor u sistemima za praćenje. Međutim, standardni Kalmanov filter je optimalno rešenje samo pod određenim pretpostavkama. Jedna od ovih pretpostavki je da merni šum ima Gausovu raspodelu. U scenarijima praćenja u realnim aplikacijama, usled prisustva autlajera, pokazuje se da šum merenja predloženog sistema za praćenje može imati raspodelu koja nije Gausova. Dakle, jedna od osnovnih pretpostavki standardnog Kalmanovog filtra nije zadovoljena. Uzimajući u obzir da su stvarna merenja kontaminirana autlajerima i da je standardni Kalmanov filter osetljiv na podatke čija raspodela nije Gausova, u narednom poglavlju su prikazane robusne tehnike estimacije i dizajniran je robusni Kalmanov filter koji se može primeniti kao prediktor u predloženom sistemu za praćenje pokretnih objekata u SWIR video sekvenci.

## 5 Robusna estimacija stanja stohastičkih sistema

Estimacija stanja stohastičkih sistema predstavlja procenu stanja sistema na osnovu dostupnih merenja [185], [186]. Za posmatrane slučajne promenljive  $X$  i  $Y$  koje imaju zajedničku funkciju gustine verovatnoće (FGV)  $f_{X,Y}(x, y)$ , gde  $X$  predstavlja stanje sistema, a promenljiva  $Y$  je merljiva čija je izmerena vrednost u datom trenutku  $y$ , treba estimirati vrednost  $x$  slučajne promenljive  $X$ . Ako se sa  $h(\cdot)$  označi linearna ili nelinearna funkcija preslikavanja, tada se estimacija vrednosti slučajne promenljive  $X$  na osnovu izmerene vrednosti  $y$  slučajne promenljive  $Y$  može izraziti kao:

$$\hat{x} = h(y) \quad (5.1)$$

Stoga, estimator predstavlja funkciju koja merljivi podatak  $y$  preslikava u procenu  $\hat{x}$  nemerljive slučajne veličine  $x$ .

Kako bi se estimator evaluirao, odnosno odredio kvalitet estimacije, definiše se greška estimacije  $e$  kao:

$$e = x - \hat{x} = x - h(y) \quad (5.2)$$

gde  $x$  označava tačnu, a  $\hat{x}$  estimiranu vrednost slučajne promenljive  $X$  za izmerenu vrednost  $y$  slučajne promenljive  $Y$ . Na osnovu definisane greške estimacije, a prema usvojenom kriterijumu kvaliteta,  $J$ , optimalni estimator stanja  $h_{opt}$  se može odrediti iz uslova minimalnog matematičkog očekivanja funkcije gubitka ili rizika  $f(\cdot)$ :

$$\hat{x}_{opt} = h_{opt}(y) = \underset{h}{\operatorname{argmin}} J(e) = \underset{h}{\operatorname{argmin}} E\{f(e)\} \quad (5.3)$$

Posebno, ako je kriterijum  $J$  u (5.3) srednje kvadratni, gde je  $f$  kvadratna funkcija, rešenje optimizacionog problema (5.3) dato je u formi uslovnog matematičkog očekivanja estimirane varijable  $X$  kada je uzeto merenje  $y$  varijable  $Y$ .

$$\hat{x}_{opt} = E\{X|Y = y\} \quad (5.4)$$

Neki od kriterijuma kvaliteta estimatora koji obezbeđuju dobar rad estimatora su nepomerčnost (srednja vrednost greške u (5.2) jednaka je 0,  $E\{e\} = 0$ ), kao i minimalna varijansa greške estimacije.

Jedan od najznačajnijih estimatora stanja stohastičkih sistema, koji takođe predstavlja i jedan od najznačajnijih doprinosa teoriji estimacije, jeste estimator minimalne varijanse greške estimacije pod nazivom Kalmanov filter [46], [47], [185], [186].

## 5.1 Kalmanov filter

Razmatra se linearni dinamički stohastički sistem koji je predstavljen modelom u prostoru stanja:

$$x_{k+1} = F_k x_k + G_k w_k \quad (5.5)$$

$$y_k = H_k x_k + v_k \quad (5.6)$$

U jednačinama (5.5) i (5.6),  $x_k$  predstavlja slučajni  $n$ -dimenzionalni vektor stanja,  $y_k$  je  $m$ -dimenzionalni vektor merenja,  $w_k$  je  $l$ -dimenzionalni vektor šuma stanja, a  $v_k$  je aditivni  $m$ -dimenzionalni vektor šuma merenja. Diskretni trenutak vremena je označen indeksom  $k$ . Štaviše,  $F_k$  predstavlja  $n \times n$  dimenzionalnu matricu prelaza stanja,  $G_k$  je  $n \times l$  dimenzionalna matrica šuma stanja (poremećaja), a  $H_k$  predstavlja  $m \times n$  dimenzionalnu matricu merenja. Inicijalni slučajni vektor stanja,  $x_0$ , je Gausovski sa poznatom srednjom vrednošću  $m_0$  i pozitivnom semidefinitnom kovarijacionom matricom  $P_0$ , definisanim kao:

$$m_0 = E\{x_0\}; P_0 = E\{[x_0 - m_0][x_0 - m_0]^T\} \quad (5.7)$$

gde  $E\{\cdot\}$  predstavlja matematičko očekivanje, a  $T$  označava transponovanje matrice. Takođe, pretpostavlja se da su sekvence  $w_k$  i  $v_k$  beli Gausovi šumovi nulte srednje vrednosti, nekorelisani sa početnim stanjem  $x_0$  za sve indekse diskretnog vremena,  $k$ :

$$E\left\{[x_0 - m_0] \begin{bmatrix} w_k \\ v_k \end{bmatrix}^T\right\} = \mathbf{0}; k = 1, 2, \dots \quad (5.8)$$

gde  $\mathbf{0}$  u (5.8) predstavlja odgovarajuću nula-matricu. Dodatno, uzima se da su sekvence šuma  $w_k$  i  $v_k$  međusobno nekorelisane:

$$E\left\{\begin{bmatrix} w_k \\ v_k \end{bmatrix} \begin{bmatrix} w_j \\ v_j \end{bmatrix}^T\right\} = \text{diag}\{Q_k \delta_{kj}, R_k \delta_{kj}\}; k, j = 1, 2, \dots \quad (5.9)$$

U (5.9)  $\text{diag}\{\cdot\}$  predstavlja blok-dijagonalnu matricu, a  $\delta_{kj}$  označava *Kronecker-ov* delta simbol ( $\delta_{kj} = 0$  ako je  $k \neq j$ , i  $\delta_{kk} = 1$ ). Takođe,  $Q_k$  i  $R_k$  predstavljaju date pozitivne semidefinitne kovarijacione matrice šuma stanja,  $w_k$ , i šuma merenja,  $v_k$ , respektivno. Štaviše, vremenski promenljive matrice  $F_k$ ,  $G_k$  i  $H_k$  su takođe poznate unapred za svaki diskretni vremenski trenutak  $k$ .

Ako je  $\hat{x}_{k|k-1}$  linearna jednokoračna optimalna predikcija trenutnog stanja,  $x_k$ , u minimalnom srednje kvadratnom smislu, a  $P_{k|k-1}$  označava odgovarajuću matricu kovarijanse greške predikcije, tada su jednačine standardnog Kalmanovog filtra izražene kao [47], [187]:

- Korak predikcije

$$\hat{x}_{k|k-1} = E\{x_k | y_1, \dots, y_{k-1}\} = E\{x_k | Y^{k-1}\} = F_{k-1} \hat{x}_{k-1|k-1} \quad (5.10)$$

$$P_{k|k-1} = E\left\{[x_k - \hat{x}_{k|k-1}][x_k - \hat{x}_{k|k-1}]^T\right\} = F_{k-1} P_{k-1|k-1} F_{k-1}^T + G_{k-1} Q_{k-1} G_{k-1}^T \quad (5.11)$$

- Korak estimacije (korekcije)

$$\varepsilon_k = y_k - H_k \hat{x}_{k|k-1} \quad (5.12)$$

$$K_k = P_{k|k-1} H_k^T [H_k P_{k|k-1} H_k^T + R_k]^{-1} \quad (5.13)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \varepsilon_k \quad (5.14)$$

$$P_{k|k} = [I - K_k H_k] P_{k|k-1} \quad (5.15)$$

U jednačinama (5.12) – (5.15),  $\varepsilon_k$  predstavlja rezidual merenja ili inovaciju,  $K_k$  je Kalmanovo pojačanje,  $P_{k|k}$  označava matricu kovarijanse greške estimacije, a  $I$  se odnosi na jediničnu matricu. Početne pretpostavke za inicijalnu procenu stanja  $\hat{x}_{0|0}$  i matricu kovarijanse  $P_{0|0}$  su definisane očekivanjem  $m_0$  i matricom kovarijanse  $P_0$  u (5.7). Rezidual merenja  $\varepsilon_k$  predstavlja nekorelisanu stohastičku sekvencu (beli šum) nulte srednje vrednosti sa kovarijacionom matricom  $S_k$ :

$$E\{\varepsilon_k \varepsilon_j^T\} = S_k \delta_{kj}; S_k = H_k P_{k|k-1} H_k^T + R_k \quad (5.16)$$

Jednačine optimalnog Kalmanovog filtra (5.10) – (5.16) su izvedene primenom Bajesovog pravila pod navedenim pretpostavkama da sve estimacije imaju nezavisne, Gausovski raspodeljene greške [46], [47]. Međutim, jednačine optimalnog Kalmanovog filtra se takođe mogu izvesti bez specifičnih informacija o raspodeli greške, sa izuzetkom pretpostavki srednjih vrednosti i kovarijansi u (5.7) – (5.9), [46], [47]. U tom smislu, Kalmanov filter je optimalni estimator stanja minimalne varijanse greške u familiji linearnih estimatora. Prema tome, može postojati neki nelinearni estimator koji je generalno optimalan, pri čemu ukupna asimptotska greška estimacije dostiže *Cramer-Rao*-vu donju granicu [47].

### 5.1.1 Osobine i ograničenja standardnog Kalmanovog filtra

Karakteristike Kalmanovog filtra se ogledaju u sledećem [185]:

- Kalmanov filter predstavlja linearnu funkciju trenutnog merenja  $y_k$ .
- Predikcija stanja sistema,  $\hat{x}_{k|k-1}$ , zavisi samo od prethodnih merenja  $Y^{k-1} = \{y_0, y_1, \dots, y_{k-1}\}$ , dok estimacija stanja sistema  $\hat{x}_{k|k}$ , eksplicitno zavisi samo od trenutnog merenja  $y_k$ .
- U slučaju vremenski-invarijantnog modela sistema, kovarijacione matrice  $P_{k|k-1}$  i  $P_{k|k}$  se mogu izračunati unapred pre implementacije samog filtra.
- Kalmanov filter generiše rekurzivno uslovno matematičko očekivanje, koje predstavlja optimalnu procenu vektora stanja sistema u smislu minimalno moguće varijanse greške estimacije. Kada su data merenja do tekućeg trenutka  $k$ ,  $Y^k = \{y_0, y_1, \dots, y_k\}$  i ukoliko slučajne veličine  $x_0$ ,  $w_k$  i  $v_k$  imaju Gausovu (normalnu) raspodelu, tada je uslovna funkcija gustine verovatnoće stanja  $x_k$  normalna i ima očekivanu (srednju) vrednost:

$$E\{x_k | Y^k\} = \hat{x}_{k|k} \quad (5.17)$$

i kovarijacionu matricu, koja dostiže *Cramer-Rao*-vu donju granicu:

$$E\{[x_k - E\{x_k|Y^k}][x_k - E\{x_k|Y^k}]|Y^k\} = P_{k|k} \quad (5.18)$$

- Ukoliko veličine  $x_0$ ,  $w_k$  i  $v_k$  nemaju Gausovu raspodelu, tada je Kalmanov filter optimalan samo u klasi linearnih filtera, u smislu minimalne varijanse greške estimacije u okviru date klase.

Optimalnost Kalmanovog filtra sadržana je u pogodnom prediktor-korektor obliku i proračunu matrice pojačanja  $K$  [46], [47]. Da bi se postigle optimalne performanse Kalmanovog filtra, potrebno je obezbediti adekvatan opis dinamike stanja sistema i statistike slučajnih poremećaja. U tom smislu, ako dinamika stanja sistema i povezane opservacije imaju značajne nelinearnosti koje se ne mogu pravilno opisati linearizacijom ili ako osnovne stohastičke sekvence nisu Gausove, performanse Kalmanovog filtra mogu značajno degradirati [30], [31]. Generalno, u slučaju nelinearne dinamike stanja ili kada opservacije nemaju Gausovu raspodelu, dizajn optimalnog estimatora stanja može biti veoma kompleksan [46] – [48], [188]. Posebno, uzevši u obzir da je linearna funkcija merenja, Kalmanov filter je veoma osetljiv na iznenadne pojave veoma velikih realizacija mernog signala. Takve realizacije mernog signala predstavljaju loše podatke ili autlajere. U datom skupu podataka, autlajeri predstavljaju one podatke čije vrednosti značajno odstupaju od vrednosti većine podataka u skupu. Uzročna priroda autlajera je raznolika, što onemogućava predvidivost njihovog pojavljivanja. Mogu se javiti usled grešaka u merenju, usled iznenadnih poremećaja u sistemu, kao i u slučaju nedovoljno preciznog modelovanja sistema. Pojavom autlajera u mernom signalu, procene standardnih statistika posmatranog sistema se mogu značajno promeniti. Stoga, značajna promena standardnih statističkih mera u zavisnosti od situacije ukazuje na pojavu merenja čije vrednosti veoma odstupaju od glavnine merne populacije, odnosno na pojavu autlajera.

Kako FGV mernog šuma često odstupa od Gausove FGV u realnim primenama usled pojave autlajera, standardni Kalmanov filter nije robusno rešenje za praćenje objekata u ovim uslovima [45], [189]. Konkretno, stvarna FGV šuma merenja u mnogim aplikacijama se može predstaviti kao Gausova FGV sa takozvanim otežalim repovima. Naime, stvarna FGV je najpribližnija Gausovoj FGV u centralnom delu, ali sa otežalim repovima u odnosu na Gausovu FGV, koja odgovara, na primer, Laplasovoj FGV [40]. Ovo, posledično, generiše mali procenat autlajera koji kontaminiraju uglavnom Gausovski raspodeljena merenja. Takva FGV se naziva i kontaminiranom Gausovom FGV, pri čemu je kontaminirana FGV simetrična sa nultom srednjom vrednošću i većom varijansom od osnovne Gausove raspodele. Posmatrajući statističku analizu brojnih merenja, stepen kontaminacije je po pravilu od 0.05 do 0.1, što odgovara prisustvu od 5 % do 10 % autlajera u mernom šumu [23]. Čak i samo jedan autlajer u mernim podacima može imati ogroman uticaj na performanse standardnog Kalmanovog filtra. Stoga, postoji dodatni praktični interes za dizajniranje robusne tehnike filtriranja. Koristeći teoriju statističke robusne estimacije, efekat odstupanja (autlajera) u većinski Gausovim opservacijama se može minimizirati. Dakle, robusni Kalmanov filter treba da postigne približno iste rezultate kao i standardni, ako podaci ne sadrže autlajere. S druge strane, u situacijama sa malim ili umerenim procentom autlajera, robusni metod treba da postigne znatno bolje performanse. Prvo svojstvo je poznato kao efikasna robusnost, dok se drugo naziva rezistentna robusnost [40].

Polazeći od praktičnih ograničenja linearnog optimalnog Kalmanovog filtra, rekurzivni robusni estimator se može dalje primeniti kako bi optimalni Kalmanov filter bio robusniji. Dakle,

postoji interes za klasu procedura estimacije koje nisu optimalne, prema određenoj statističkoj meri performanse, ali imaju ograničenu ukupnu grešku estimacije. Familija procedura dinamičke stohastičke aproksimacije nudi razuman izbor, postižući dobre rezultate u mnogim aplikacijama, uključujući estimaciju parametara, optimizaciju, klasifikaciju obeležja i obradu signala [32] – [36]. Kao što je već pomenuto, Kalmanov filter čije se pojačanje razlikuje od optimalnog zbog grešaka u pretpostavljenoj statistici šuma ili zbog neadekvatnog predstavljanja dinamike stanja sistema, može se posmatrati kao algoritam dinamičke stohastičke aproksimacije [46], [32] – [35]. Stoga, ovaj algoritam može predstavljati odgovarajuću zamenu za optimalnu tehniku estimacije kada u praksi nisu ispunjene pretpostavke na kojima se optimalna tehnika estimacije zasniva.

## 5.2 Robusna estimacija stanja dinamičkih sistema

Pojam robusnosti u kontekstu ove disertacije predstavlja neosetljivost estimacione procedure na odstupanja od pretpostavki koje su usvojene prilikom dizajniranja estimatora. Od interesa je prvenstveno neosetljivost, odnosno robusnost na iznenadne velike realizacije mernog signala koje predstavljaju autlajere. Ova robusnost zapravo predstavlja neosetljivost na odstupanja od usvojene raspodele, gde je za standardni Kalmanov filter usvojena Gausova raspodela šuma merenja, a usled prisustva autlajera u mernom signalu u realnim situacijama, stvarna raspodela šuma merenja odstupa od Gausove, u smislu otežinjenih repova [23].

Kako prisustvo autlajera u mernim podacima može da dovede do degradacije performansi standardnog Kalmanovog filtra, prva je pomisao da se dizajnira metod i kriterijum za odbacivanje autlajera iz mernih podataka, odnosno cenzurisanje opservacija. Zatim bi se standardni Kalmanov filter mogao primeniti na skupu cenzurisanih mernih podataka iz kojeg su eliminisani autlajeri. Međutim, često nije lako jasno razdvojiti autlajere od ostatka merne populacije. Štaviše, ako bi se standardni Kalmanov filter primenio na skupu podataka iz kog su eliminisani autlajeri (pretpostavljajući da originalni skup mernih podataka ima normalnu Gausovu raspodelu kontaminiranu autlajerima), takav skup zapravo neće imati normalnu raspodelu. Razlog za to je što će uvek postojati statističkih grešaka usled pogrešnog odbacivanja korektnih podataka i pogrešnog zadržavanja autlajera u mernom skupu [40]. Stoga, estimator poput standardnog Kalmanovog filtra koji podrazumeva normalnu raspodelu mernih podataka nije uvek primenljiv na skup podataka iz kog su nekom heuristikom odbačeni autlajeri. Performanse takve dvokoračne estimacione procedure mogu biti značajno manje od robusne estimacione procedure. Robusni estimator može da napravi finiju diskriminaciju između potpunog prihvatanja i potpunog odbacivanja nekog podatka iz mernog skupa.

Stoga, za usvojeni model procesa za koji se pretpostavlja da predstavlja dobru aproksimaciju realnog sistema, ali svakako ne može biti apsolutno tačan, robusni estimator bi trebalo da postigne dobru efikasnost (blisku optimalnoj). Mala odstupanja od pretpostavki modela treba neznatno da naruše performanse estimatora. Najvažniji zadatak koji robusni estimator treba da ispuni je da spreči da veoma velika odstupanja od usvojenih pretpostavki modela degradiraju sistem u potpunosti.

U statističkoj literaturi postoje najmanje četiri definicije robusnosti: rezistentna, efikasna, mini-maksna i kvalitativna robusnost [23], [37] – [40]. Rezistentna robusnost se odnosi na neosetljivost estimatora na pojedinačne autlajere. Drugim rečima, za estimator se može reći da je



rezistentno robusan ako velike promene u malom procentu populacije mernih podataka rezultuju malim promenama u estimacijama koje se zasnivaju na ovim mernim podacima. Takođe, u praksi se postavlja i zahtev neosetljivosti estimatora na grupisane autlajere. Efikasna robusnost postavlja zahtev da estimator ima veliku efikasnost u odnosu na minimalno moguću grešku estimacije koja odgovara *Cramer-Rao*-voj donjoj granici. Zahtev je postavljen u slučaju da realni šum merenja ima normalnu raspodelu, ali i u slučaju mernog šuma čija raspodela odstupa od normalne, odnosno ima otežale repove. Mini-maksna robusnost pretpostavlja poznavanje familije funkcija raspodela kojoj pripada stvarna raspodela šuma, ali i postojanje najnepovoljnije funkcije raspodele unutar date familije koja o veličini koja se estimira nosi minimalnu količinu Fišerove informacije. U tom slučaju, estimator koji zadovoljava mini-maksnu robusnost ima najmanju moguću varijansu greške za najnepovoljniju raspodelu unutar date familije. Mini-maksni robusni estimator se projektuje tako da ima optimalne performanse za najnepovoljniji mogući slučaj u posmatranoj situaciji. Za svaku drugu raspodelu unutar date familije, greška estimacije ovog tipa robusnog estimatora neće biti veća od granične vrednosti koja je postavljena u odnosu na najnepovoljniju raspodelu. Kvalitativna robusnost postavlja zahtev da male promene u mernim podacima rezultuju malim promenama u estimacijama koje se zasnivaju na ovim mernim podacima. Pod malim promenama se podrazumevaju velike promene u malom delu populacije mernih podataka, ali i male promene u velikom delu populacije mernih podataka. Kako kvalitativna robusnost sama po sebi nije dovoljna, potrebno je da se zahtevi ovog tipa robusnosti prošire zahtevima efikasne robusnosti [185], [23]. Mini-maksna i kvalitativna robusnost su u matematičkom smislu veoma složene, pa se u praksi najčešće projektuje estimator koji je rezistentno i efikasno robusan.

Robusni estimatori stanja se mogu podeliti u tri glavne grupe [23], [37] – [40], [190]:

- M-estimatori: zasnovani na kriterijumu maksimalne verodostojnosti (eng. *maximum likelihood* - ML). Predloženi su od strane Hubera [40] i predstavljaju najčešće korišćene robusne estimatore.
- R-estimatori: izvedeni su na osnovu rang testova [23], [37], [190].
- L-estimatori: zasnovani su na linearnoj kombinaciji redne statistike [23], [37], [190].

Navedene tri glavne grupe estimatora su definisane na različite načine, ali imaju isti cilj - da eliminišu uticaj autlajera i smanje osetljivost estimatora na raspodele sa otežalim repovima. Takođe, pokazano je da ove tri klase estimatora asimptotski postaju ekvivalentne [39], [190].

Rezistentna i efikasna robusnost označavaju praktičnu robusnost. Projektovanje estimatora koji zadovoljava i rezistentnu i efikasnu robusnost može se postići primenom Huberovog M-robusnog pristupa budući da potiče od koncepta optimalne maksimalne verodostojnosti, što ga čini prirodnijim i jednostavnijim za implementaciju [23], [37], [40]. U tom smislu, estimator ne mora biti optimalni ML estimator, već ima zadatak da ga aproksimira na način da se postignu ciljevi praktične robusnosti estimacione procedure.

### 5.3 Metod dinamičke stohastičke aproksimacije

Metode stohastičke aproksimacije prvobitno su bile primenjene za procenu parametara u regresionim modelima [32] – [35], [46]. U tom smislu, pretpostavimo da nepoznati parametar  $\theta_0$  treba estimirati iz opservacija predstavljenih u formi linearne regresije:

$$y_k = h_k \theta_0 + e_k \quad (5.19)$$

gde je  $h_k$  vremenski promenljiva varijabla, pri čemu je  $k$  indeks diskretnog vremena, dok je  $\{e_k\}$  slučajna sekvenca nulte srednje vrednosti sa ograničenom varijansom. Definisanjem nove funkcije, u svakom koraku, takozvane regresione funkcije u obliku:

$$g_k(\theta) = h_k \theta_0 - h_k \theta \quad (5.20)$$

problem određivanja  $\theta_0$  se svodi na pronalaženje rešenja sledeće linearne algebarske jednačine:

$$g_k(\theta) = 0 \quad (5.21)$$

Oduzimanjem člana  $h_k \hat{\theta}_k$  sa obe strane znaka jednakosti u jednačini (5.19), gde  $\hat{\theta}_k$  predstavlja estimaciju  $\theta_0$  u diskretnom trenutku  $k$ , uz korišćenje jednačine (5.20) i definisanjem reziduala merenja:

$$m_k = y_k - h_k \hat{\theta}_k \quad (5.22)$$

dobija se ekvivalentna relacija za jednačinu merenja u (5.19):

$$m_k = g_k(\hat{\theta}_k) + e_k \quad (5.23)$$

Regresiona funkcija u (5.20) je nepoznata, tako da nije moguće izračunati njenu vrednost za dati argument  $\hat{\theta}_k$ . Međutim, dostupna je zašumljena opservacija  $m_k$  iz (5.22). Dakle, procena parametra  $\theta_0$  se svodi na problem aproksimacije rešenja (5.21) iz zašumljenih podataka merenja,  $m_k$ , u (5.22). Primenom iterativnog numeričkog algoritma:

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \gamma_k m_k = \hat{\theta}_k + \gamma_k (y_k - h_k \hat{\theta}_k) \quad (5.24)$$

se rekurzivno može estimirati nepoznati parametar,  $\theta_0$ , gde je  $\{\gamma_k\}$  skalarna, pozitivna i opadajuća sekvenca koja teži nuli kada indeks vremena  $k$  teži beskonačnosti. Ovaj pristup je sličan numeričkom zadatku traženja rešenja za funkciju  $g_k$  (5.21), koji se može rešavati nekim od standardnih numeričkih algoritama [191]. Kako opservacije  $m_k$  sadrže slučajne greške,  $e_k$ , takozvani aditivni beli šum, algoritam (5.24) se naziva stohastičkom aproksimacijom [32] – [35], [46]. Da bi se izvedeni rezultat proširio na problem procene skupa  $n$  konstantnih parametara uređenih u  $n \times 1$  vektor  $\theta_0$ , skalarna linearna jednačina merenja iz (5.19) se mora preformulisati u:

$$y_k = h_k^T \theta_0 + e_k \quad (5.25)$$

gde je  $h_k$  poznati  $n \times 1$  dimenzionalni vektor. Štaviše, vektor merenja  $y_k$  se može procesirati razmatranjem pojedinačnih komponenti, jednu po jednu, u formi (5.25). Po analogiji sa (5.24), algoritam višedimenzionalne stohastičke aproksimacije dobija oblik:

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \Gamma_k (y_k - h_k^T \hat{\theta}_k) \quad (5.26)$$

gde težinski faktor  $\Gamma_k$  predstavlja  $n \times 1$  dimenzionalni vektor [32] – [35], [46]. Izbor vektora pojačanja u (5.26) treba da obezbedi dobre performanse konvergencije i dat je kao:

$$\Gamma_k = \frac{h_k}{\sum_{i=1}^k \|h_i\|^2}; \quad \|h_i\|^2 = h_i^T h_i \quad (5.27)$$

gde  $\|\cdot\|$  označava Euklidsku normu. Forma (5.27) je preuzeta iz problema optimalne procene parametara metodom najmanjih kvadrata [192].

Po analogiji sa slučajem konstantnih parametara u (5.19) i (5.25), metod stohastičke aproksimacije iz (5.24) i (5.26) se može proširiti na procenu vremenski promenljivih varijabli, na osnovu opservacije koje su date u nelinearnoj vektorskoj formi:

$$y_k = \zeta_k(\theta_k) + e_k \quad (5.28)$$

gde nepoznati vektor parametara  $\theta_k$  koji treba estimirati varira od uzorka do uzorka, dok je  $\zeta_k(\cdot)$  data vektorska funkcija vektorskog argumenta, a stohastička sekvenca  $\{e_k\}$  predstavlja beli šum nulte srednje vrednosti [32] – [35], [46]. Štaviše, varijacije u sekvenci  $\{\theta_k\}$  obično su predstavljene stohastičkom nelinearnom vektorskom diferencnom jednačinom:

$$\theta_{k+1} = \varphi(\theta_k) + w_k \quad (5.29)$$

gde je  $\{\varphi_k(\cdot)\}$  poznati skup vektorskih funkcija vektorskog argumenta, dok je stohastička sekvenca  $\{w_k\}$  beli šum nulte srednje vrednosti, koji nije u korelaciji sa šumom merenja  $\{e_k\}$  iz (5.28) [32], [46]. Po analogiji sa (5.26), algoritam dinamičke stohastičke aproksimacije dat je u obliku:

$$\hat{\theta}_{k+1} = \bar{\theta}_{k+1} + \Gamma_k [y_k - \zeta_k(\bar{\theta}_{k+1})]; \quad \bar{\theta}_{k+1} = \varphi(\hat{\theta}_k) \quad (5.30)$$

i primenjen je za procenu  $\theta_k$  u (5.29), koristeći merenja  $y_k$  iz (5.28). Sada je najnovija procena  $\hat{\theta}_k$  iz (5.26), zamenjena jednokoračnom prediktovanom promenom,  $\bar{\theta}_{k+1}$ , u  $\theta_k$  u sadašnjem trenutku  $k$ , koristeći pretpostavljeni model procesa (5.29). Generalno, nije moguće dokazati konvergenciju sekvence  $\{\hat{\theta}_k\}$  iz (5.30), ali se u praksi može očekivati ograničena greška estimacije u statističkom smislu [193]. Rekurzivni algoritam (5.30) izgleda kao estimacija minimalne varijanse, (5.12) – (5.14), u optimalnom Kalmanovom filtru, ali primetna razlika je u tome što matrica pojačanja  $\Gamma_k$  nije optimalna u statističkom smislu, za razliku od Kalmanovog pojačanja u (5.13).

Jednačina (5.30) ima isti oblik kao optimalni Kalmanov filter u (5.10) – (5.14), pošto je trenutna estimacija izračunata kao linearna kombinacija prethodne poznate estimacije (predikcije) i reziduala merenja [46] – [48], [187], [188]. Stoga, bilo koji Kalmanov filter sa pogrešnom sekvencom pojačanja se može posmatrati kao metoda dinamičke stohastičke aproksimacije u (5.30). Ovo ukazuje na mogućnost da se optimalni linearni Kalmanov filter učini robusnijim, korišćenjem algoritma M-robusne dinamičke stohastičke aproksimacije, zasnovanog na dodatnoj proceduri optimizacije za izračunavanje matrice pojačanja, po analogiji sa optimalnim Kalmanovim filtrom (5.13) – (5.15).

Pošto je takav proces optimizacije nelinearan, tehnika statističke linearizacije se može iskoristiti za izvođenje aproksimativne procedure estimacije tipa minimalne varijanse sa poboljšanom brzinom konvergencije [194].

## 5.4 M-robustni Kalmanov filter

Analogno slučaju sa vremenski promenljivim parametrima u (5.28) i (5.29), metod dinamičke stohastičke aproksimacije se može proširiti na izračunavanje robusnih rekurzivnih estimacija stanja dinamičkog stohastičkog sistema predstavljenog sa (5.5), pod pretpostavkom da su merenja iz (5.6) skalarna. Kao što je ranije pomenuto, slučaj višedimenzionalnih merenja u (5.6) se može razmatrati na isti način procesiranjem pojedinačnih komponenti, jednu po jednu. Ovaj pristup pretpostavlja da se komponente vektora merenja u (5.6) mogu procesirati sekvencijalno, kao nekorelisane skalarne opservacije. U tom smislu, potrebno je redefinisati vektor merenja u (5.6) tako da odgovarajuće greške merenja ili komponente vektora šuma, budu međusobno nekorelisane. Ovo, zauzvrat, rezultira dijagonalnom formom kovarijacione matrice šuma merenja  $R_k$  iz (5.9). Pogodan metod stabilne numeričke dekompozicije koji se često koristi u praksi je *Cholesky* faktorizacija ili njena modifikacija, takozvana UD-dekompozicija [47], [188].

Oslanjajući se na koncept M-robustne estimacije, robustifikacija Kalmanovog filtra se može zasnivati na minimizaciji nelinearne M-robustne kriterijumske funkcije (indeks performanse, faktor cene, rizika ili gubitka),  $\rho$ . Da bi se primenio ovaj robusni pristup na rekurzivnu procenu dinamičkog stanja sistema, uvodi se M-robustni indeks performanse u formi vremenski promenljivog funkcionala:

$$J_k(\bar{x}) = E \left\{ \rho \left( \frac{\varepsilon_k(\bar{x})}{s_k} \right) \middle| \bar{x}, Y^k \right\} \quad (5.31)$$

pri čemu je  $E\{\cdot | \bar{x}, Y^k\}$  uslovno matematičko očekivanje kada su dati jednokoračna predikcija stanja,  $\bar{x}$ , kao i raspoloživi merni skup do tekućeg trenutka,  $Y^k = \{y_0, y_1, \dots, y_k\}$ , gde je  $y$  izlaz sistema u (5.6) [34], [45], [194] – [196]. Polazeći od (5.31), može se definisati porodica rekurzivnih estimatora tipa dinamičke stohastičke aproksimacije, minimizirajući M-robustni indeks performanse (5.31) u svakom koraku:

$$\hat{x}_k = \bar{x}_k - \Gamma_k g_k(\bar{x}_k); \quad g_k(\bar{x}_k) = \nabla_{\bar{x}} J_k(\bar{x}_k) \quad (5.32)$$

gde je  $\Gamma_k$  matrično pojačanje,  $\bar{x}_k$  predstavlja jednokoračnu predikciju vektora stanja  $x_k$  koji se estimira, a  $\hat{x}_k$  predstavlja tekuću estimaciju. Član  $\nabla_{\bar{x}} J_k(\bar{x}_k)$  označava gradijentni vektor determinističkog M-robustnog kriterijuma (5.31). Uzimajući u obzir (5.31), dobija se:

$$\nabla_{\bar{x}} J_k(\bar{x}_k) = \frac{\partial J_k(\bar{x}_k)}{\partial \bar{x}} = -\frac{1}{s_k} E \left\{ \psi \left( \frac{\varepsilon_k}{s_k} \right) \middle| \bar{x}_k, Y^k \right\} H_k^T \quad (5.33)$$

pri čemu  $\psi(\cdot)$  predstavlja prvi izvod nelinearne kriterijumske funkcije  $\rho(\cdot)$  u indeksu performanse iz (5.31), koji se u robusnoj statistici naziva funkcija uticaja jer predstavlja meru osetljivosti estimatora na autlajere koji kontaminiraju mernu populaciju [39], [185]. Štaviše,  $\partial(\cdot)/\partial x =$

$\{\partial(\cdot)/\partial x_1 \cdots \partial(\cdot)/\partial x_n\}^T$  označava operator parcijalnog izvoda, gde je  $x$  vektor kolona dimenzija  $n \times 1$ . Analogno sa (5.12), rezidual merenja definisan je kao:

$$\varepsilon_k = \varepsilon(\bar{x}_k) = y_k - H_k \bar{x}_k \quad (5.34)$$

gde je  $y_k$  izlaz sistema, a  $H_k$  matrica merenja iz (5.6). Pored toga, veličina  $s_k$  obezbeđuje estimacije invarijantne na skaliranje i predstavlja faktor normalizacije (skaliranja), kao procenu standardne devijacije greške (reziduala) merenja iz (5.34).

Generalno, uslovno matematičko očekivanje iz (5.33) se ne može odrediti, pa se analogno sa (5.23) može aproksimirati trenutnim uzorkom. Dakle, nepoznato matematičko očekivanje u (5.33) se može estimirati u svakom diskretnom trenutku,  $k$ , trenutnom realizacijom slučajnog argumenta. Ovo rezultira vektorom stohastičkog gradijenta:

$$m_k = -\frac{1}{s_k} \psi\left(\frac{\varepsilon_k}{s_k}\right) H_k^T \quad (5.35)$$

Slično (5.24), zamenom (5.35) u (5.32), porodica M-robusnih rekurzivnih estimatora stanja tipa dinamičke stohastičke aproksimacije poprima oblik:

$$\hat{x}_k = \bar{x}_k + \frac{1}{s_k} \Gamma_k H_k^T \psi\left(\frac{\varepsilon_k}{s_k}\right) \quad (5.36)$$

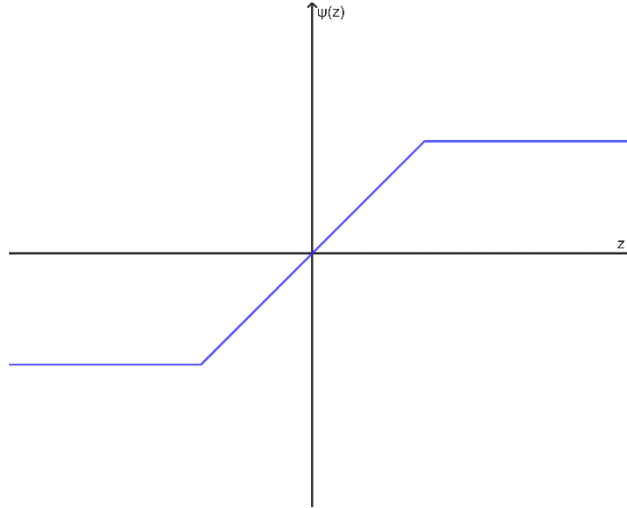
Uloga kriterijumske funkcije  $\rho$  u (5.31) je da obezbedi praktičnu robusnost estimacione procedure (5.36). Da bi se postigle takve performanse, M-robusna funkcija uticaja,  $\psi = \rho'$ , mora biti ograničena i kontinualna funkcija [23], [37] – [40]. Ovo, posledično, dovodi do toga da pojedinačni, ali i grupa autlajera neće imati značajnog uticaja na estimaciju stanja (5.36), zadovoljavajući zahtev rezistentne robusnosti. Pored toga, da bi se zadovoljila efikasna robusnost, estimaciona procedura (5.36) mora da radi dobro kada opservacije imaju Gausovu raspodelu, ali i u slučaju kada je Gausova raspodela kontaminirana autlajerima. Polazeći od Gausove raspodele šuma merenja sa otežalim repovima, Huber je predložio da M-robusna  $\rho$ -funkcija bude kvadratna u sredini i da raste sporije od kvadratne za veće apsolutne vrednosti argumenta i tako ispunjava oba zahteva praktične robusnosti [40].

$$\rho(z) = \begin{cases} \frac{z^2}{2} & ; |z| \leq \Delta \\ \Delta|z| - \frac{\Delta^2}{2} & ; |z| > \Delta \end{cases} \quad (5.37)$$

Odgovarajuća M-robusna funkcija uticaja je monotono neopadajuća nelinearnost tipa zasićenja:

$$\psi(z) = \rho'(z) = \begin{cases} z & ; |z| \leq \Delta \\ \Delta \operatorname{sgn}(z) & ; |z| > \Delta \end{cases} \quad (5.38)$$

gde je  $\Delta$  parametar koji kontroliše efikasnu robusnost. Izbor  $\Delta = 1.5$  često daje prihvatljive rezultate u praksi, a takav postupak je poznat kao Huberov 1.5 M-robusni pristup [40]. Huberova M-robusna funkcija uticaja je prikazana na slici 5.1. Nelinearna obrada podataka primenom saturacione funkcije (5.38) poznata je u statističkoj literaturi kao *winsorization* (eng. *winsorization*) [23], [37] – [40].



**Slika 5.1** Huberova M-robusna funkcija uticaja ( $1 < \Delta < 3$ ).

Kao što je već pomenuto, statistička analiza je pokazala da različiti merni podaci sadrže, po pravilu, od 5 do 10 procenata autlajera [23]. U tom smislu, često se pretpostavlja da su merenja generisana miksom Gausovih raspodela:

$$p_{\delta}(\cdot) = (1 - \delta)N(\cdot | 0, \sigma_n^2) + \delta N(\cdot | 0, \sigma_o^2); \quad 0 \leq \delta \leq 1, \quad \sigma_n^2 = 1, \quad \sigma_o^2 \gg \sigma_n^2 \quad (5.39)$$

gde je  $\delta$  stepen kontaminacije,  $\sigma_n^2$  je varijansa većinskih opservacija u skupu podataka generisanih standardnom Gausovom raspodelom nulte srednje vrednosti i jedinične varijanse,  $N(\cdot | 0, 1)$ , dok je  $\sigma_o^2$  varijansa autlajera generisanih normalnom raspodelom sa nultom srednjom vrednošću,  $N(\cdot | 0, \sigma_o^2)$ . Ovakva raspodela je takođe poznata kao  $\delta$ -kontaminirana normalna raspodela [23], [37] – [40], [195]. Posebno, za klasu raspodela (5.39), sa proizvoljnom simetričnom raspodelom nulte srednje vrednosti, umesto Gausove  $N(\cdot | 0, \sigma_o^2)$ , najnepovoljnija raspodela  $p_0$ , u smislu minimalne Fišerove informacije je Gausova raspodela u sredini i Laplasova, ili dvostruko eksponencijalna, na repovima. Funkcija uticaja,  $\psi = \rho'$ , za datu funkciju verodostojnosti  $\rho(\cdot) = -\ln p_0(\cdot)$ , je nelinearnost tipa zasićenja prikazana u (5.38) [40]. Primeri klasa raspodela koje se obično koriste u inženjerskim problemima, kao i izvođenje najnepovoljnijeg slučaja raspodele u okviru unapred određene klase, mogu se naći u literaturi [195].

Uloga matrice pojačanja  $\Gamma_k$  u (5.36) je da kontroliše brzinu konvergencije. Stoga, analogno optimalnom Kalmanovom filtru, uvodi se dodatni kriterijum aproksimativne minimalne varijanse:

$$J_1(\Gamma_k) = \text{Trace} P_k; \quad P_k = E\{\tilde{x}_k(+) \tilde{x}_k^T(+)\}; \quad \tilde{x}_k(+) = x_k - \hat{x}_k \quad (5.40)$$

gde je  $P_k$  matrica kovarijanse greške estimacije, a  $\text{Trace}$  označava trag matrice. Minimizacija kriterijuma (5.40) po matrici pojačanja  $\Gamma_k$  predstavlja složen nelinearni problem, pa se aproksimativno optimalno rešenje može dobiti primenom tehnike statističke linearizacije [46] – [48]. Polazeći od  $\psi$ -funkcije iz (5.38), čiji je argument uzorak sekvence skaliranih reziduala merenja  $\{\varepsilon_k/s_k\}$  prema (5.36) sa simetričnom FGV koja pripada klasi (5.39), primenom metoda statističke linearizacije dobija se sledeća aproksimacija funkcije uticaja:

$$\psi\left(\frac{\varepsilon_k}{s_k}\right) \approx \alpha \frac{\varepsilon_k}{s_k}; \quad \alpha = \frac{1}{\sigma_o^2} E\{\psi(z)z\} \quad (5.41)$$

pri čemu je  $\alpha$  optimalni koeficijent statističke linearizacije, u srednje kvadratnom smislu, dok je  $\sigma_z^2$  varijansa argumenta  $z$  (više detalja je dato u sekciji 5.4.2) [194].

Posebno, ako je u (5.41) saturaciona funkcija  $\psi(\cdot)$  definisana kao (5.38), tada koeficijent  $\alpha$  zavisi od linearnog segmenta funkcije  $\psi(\cdot)$ , praga saturacije  $\Delta$  i varijanse  $\sigma_z^2$ . Generalno, za male vrednosti  $\sigma_z$  u poređenju sa  $\Delta$ -vrednostima, verovatnoća pojave zasićenja je mala, što dovodi do vrednosti koeficijenta  $\alpha$  bliskim jedinici. Štaviše, za veće vrednosti  $\sigma_z$ , vrednost koeficijenta  $\alpha$  je znatno manja od jedan, zbog veće verovatnoće pojave zasićenja. Dakle, za unapred određenu funkciju  $\psi$  i različite vrednosti  $\sigma_z$  dobija se skup koeficijenata  $\alpha$  iz intervala  $[0, 1]$ . Dodatno, pošto normalizovani rezidual u (5.41) ima jediničnu standardnu devijaciju koja je manja od praga  $\Delta$ , vrednost odgovarajućeg koeficijenta  $\alpha$  je bliska jedinici. Zamenom (5.41) u (5.36) dobija se sledeća relacija za estimaciju stanja statistički linearizovanim M-robusnim Kalmanovim filtrom:

$$\hat{x}_k = \bar{x}_k + K_k \varepsilon_k; K_k = s_k^{-2} \alpha \Gamma_k H_k^T \quad (5.42)$$

Standardna devijacija  $s_k$  reziduala merenja  $\varepsilon_k$  iz (5.34) može biti definisana iz varijanse (5.16), dajući:

$$s_k = S_k^{1/2} = (H_k M_k H_k^T + R_k)^{1/2} \quad (5.43)$$

pri čemu je  $M_k$  matrica kovarijanse greške predikcije definisana u (5.11), odnosno:

$$M_k = E\{\tilde{x}_k(-)\tilde{x}_k^T(-)\}; \tilde{x}_k(-) = x_k - \bar{x}_k \quad (5.44)$$

Konkretno, za model šuma merenja u (5.39), varijansa šuma izražena je kao:

$$R_k = (1 - \delta)\sigma_n^2 + \delta\sigma_o^2 \quad (5.45)$$

Generalno, stepen kontaminacije  $\delta$  u praksi nije tačno poznat i ne može se adekvatno odrediti iz reziduala merenja [23], [37] – [40], [195]. Kao što je ranije pomenuto, razuman izbor u praksi je usvajanje vrednosti stepena kontaminacije  $\delta$  u intervalu od 0.05 do 0.1, što odgovara prisustvu autlajera u mernim podacima u meri od 5 do 10 procenata. Štaviše, standardna devijacija autlajera  $\sigma_o$  takođe nije poznata, ali je značajno veća od jedinične nominalne standardne devijacije  $\sigma_n$ . Uzimajući u obzir (5.5), (5.6), (5.34), (5.42) – (5.44), zajedno sa pogodnim uprošćenjima, dobija se aproksimativno optimalno rešenje, minimiziranjem usvojenog kriterijuma u (5.40) (više detalja je dato u sekciji 5.4.1):

$$\Gamma_k = M_k; P_k = (I - K_k H_k) M_k \quad (5.46)$$

Dodatno, polazeći od (5.36) – (5.46), statistički linearizovan M-robusni rekurzivni estimator stanja tipa dinamičke stohastičke aproksimacije se može definisati na sledeći način:

$$\hat{x}_k = \bar{x}_k + K_k s_k \psi_H \left( \frac{\varepsilon_k}{s_k} \right); K_k = s_k^{-2} \alpha M_k H_k^T \quad (5.47)$$

Rezidual  $\varepsilon_k$ , faktor skaliranja  $s_k$ , Huberova funkcija uticaja  $\psi_H$  i koeficijent  $\alpha$  su definisani u (5.34), (5.43), (5.38) i (5.41), respektivno, dok je kovarijaciona matrica greške estimacije  $P_k$  data u (5.46). Relacije (5.46) i (5.47) su slične rekurzijama ažuriranja merenja (5.12) – (5.15) u koracima estimacije

optimalnog linearnog Kalmanovog filtra. Kako su koraci predikcije i estimacije u optimalnom Kalmanovom filtru nezavisni, korak estimacije se može robustifikovati kombinovanjem rekurzivnog procesa robusne estimacije (5.46) i (5.47), umesto rekurzija ažuriranja merenja (5.12) – (5.15), sa jednokoračnom predikcijom iz (5.10) – (5.11), u cilju izvođenja nove M-robusne verzije optimalnog Kalmanovog filtra. U tom smislu, relacije (5.10) i (5.11) iz koraka predikcije linearnog optimalnog Kalmanovog filtra definišu i korak predikcije u M-robusnom statistički linearizovanom Kalmanovom filtru, dajući jednokoračnu predikciju  $\bar{x}_k$  i odgovarajuću kovarijacionu matricu  $M_k$ :

$$\bar{x}_k = F_{k-1}\hat{x}_{k-1}; M_k = F_{k-1}P_{k-1}F_{k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T \quad (5.48)$$

gde su  $F_{k-1}$  i  $G_{k-1}$  matrica prelaza stanja sistema i matrica šuma stanja iz (5.5), dok je kovarijaciona matrica šuma stanja  $Q_{k-1}$  definisana u (5.9).

Nažalost, u većini slučajeva statistika šuma merenja nije tačno poznata i u takvim okolnostima optimalni koeficijent linearizacije  $\alpha$  iz (5.41) se ne može odrediti. Stoga, koeficijent  $\alpha$  se može aproksimirati fiksnim koeficijentom  $\alpha_f$  koji je definisan relacijom:

$$\alpha_f = E\left\{\frac{\psi(z)}{z}\right\}; \frac{\psi(z)}{z} \approx \psi'(z) \quad (5.49)$$

gde je  $\psi'(\cdot)$  prvi izvod ili nagib funkcije  $\psi(\cdot)$ .

Posebno, za slučaj Huberove  $\psi_H$  funkcije iz (5.38), relacija (5.49) se može iskoristiti za objašnjenje fizičkog značenja fiksnog koeficijenta  $\alpha_f$ , kao i za estimaciju njegove vrednosti. Polazeći od (5.38) i (5.49), dobija se:

$$\alpha_f = E\{\psi'_H(z)\} = \int_{|z| \leq \Delta} p(z) dz = 1 - \delta \quad (5.50)$$

gde je  $p(\cdot)$  nepoznata FGV šuma merenja iz (5.39). Ovde se pretpostavlja da realna FGV skaliranog reziduala  $\varepsilon_k/s_k$  iz (5.47) takođe pripada datoj klasi FGV iz (5.39) (više detalja je dato u sekciji 5.4.2). Integral u (5.50) jednak je verovatnoći da su opservacije generisane nominalnom standardnom Gausovom FGV, koja odgovara linearnom delu  $\psi_H$  funkcije iz (5.38), sa nagibom  $\psi'_H$  jednakim jedinici. U skladu sa (5.39), ova verovatnoća se može proceniti pomoću (5.50), koristeći pretpostavljeni stepen kontaminacije  $\delta$ . Treba napomenuti da se izračunavanje fiksnog koeficijenta  $\alpha_f$  u (5.50) može zasnivati i na najgorem slučaju FGV,  $p_0$ , u okviru date klase (5.39), što rezultira u:

$$\alpha_f = E\{\psi'_H(z)\} = \int_{|z| \leq \Delta} p_0(z) dz = 2(1 - \delta) \operatorname{erf}(\Delta) \quad (5.51)$$

gde je  $\operatorname{erf}$  funkcija greške [40], [195]. Ovo rešenje je asimptotski jednako (5.50), pošto je vrednost  $\operatorname{erf}$  funkcije blizu 0.5 za dovoljno veliki argument.

Dakle, relacije (5.50) i (5.51) definišu fiksne i izvodljive aproksimacije optimalnog koeficijenta statističke linearizacije iz (5.41), predstavljajući aproksimaciju srednjeg nagiba Huberove M-robusne funkcije uticaja  $\psi_H$ , koja je definisana u (5.38).

Dakle, u odsustvu autlajera, što odgovara stepenu kontaminacije  $\delta = 0$ , vrednost koeficijenta  $\alpha$  u (5.50) ili (5.51) jednaka je jedinici, pri čemu se robusno pojačanje  $K_k$  iz (5.47) svodi na optimalno Kalmanovo pojačanje u (5.13). Pošto je robusna funkcija uticaja  $\psi_H$  u (5.38) u svom linearnom



režimu, koji odgovara linearnoj funkciji uticaja optimalnog Kalmanovog filtra, robusni rekurzivni estimator (5.47) se ponaša kao optimalni Kalmanov filter. Međutim, u prisustvu autlajera, koeficijent  $\alpha_f$  u (5.50) ili (5.51) opada sa povećanjem vrednosti stepena kontaminacije, dodatno smanjujući vrednosti matrice robusnog pojačanja  $K_k$ . Pošto funkcija uticaja  $\psi_H$  sada radi u svom režimu saturacije, kombinacija ova dva efekta potiskuje uticaj autlajera na robusne rekurzivne estimacije u (5.47).

S druge strane, varijabilna aproksimacija optimalnog koeficijenta statističke linearizacije iz (5.41) data je kao:

$$\alpha(\varepsilon_k/s_k) = \alpha_k = \begin{cases} \frac{\psi_H(\varepsilon_k/s_k)}{\varepsilon_k/s_k} & ; \varepsilon_k \neq 0 \text{ and } s_k \neq 0 \\ 1 & ; \varepsilon_k = 0 \text{ or } s_k = 0 \end{cases} \quad (5.52)$$

gde je matematičko očekivanje iz (5.50) zamenjeno trenutnim uzorkom. Ova aproksimacija predstavlja trenutni nagib Huberove M-robusne funkcije uticaja iz (5.38) koji je aproksimativno jednak nuli ili jedinici.

Dakle, u odsustvu autlajera promenljivi koeficijent  $\alpha_k$  iz (5.52) ima jediničnu vrednost, tako da se matrica robusnog pojačanja  $K_k$  u (5.47) svodi na optimalno Kalmanovo pojačanje iz (5.13). Pošto Huberova funkcija uticaja iz (5.38) radi u linearnom režimu, robusni rekurzivni estimator u (5.47) se ponaša kao optimalni linearni Kalmanov filter. S druge strane, u prisustvu autlajera, promenljivi koeficijent  $\alpha_k$  ima vrednost blisku nuli, značajno smanjujući robusno pojačanje  $K_k$  u (5.47), dok je funkcija uticaja  $\psi_H$  sada ograničena na svoj režim saturacije, čime se efikasnije potiskuje uticaj autlajera u poređenju sa primenom fiksnog koeficijenta u (5.50) ili (5.51).

### 5.4.1 Statistički linearizovan M-robusni Kalmanov filter

Polazeći od (5.5) – (5.9), za grešku predikcije iz (5.44) dobija se relacija:

$$\tilde{x}_k(-) = F_{k-1}\tilde{x}_{k-1}(+) + G_{k-1}w_{k-1} \quad (5.53)$$

gde je greška estimacije  $\tilde{x}_k(+)$  definisana u (5.40), a varijabla  $w_k$  predstavlja uzorak belog šuma stanja iz (5.5). Uzimajući u obzir (5.5) – (5.6), (5.34) i (5.42), greška estimacije iz (5.40) je data kao:

$$\tilde{x}_k(+) = \tilde{x}_k(-) - K_k H_k \tilde{x}_k(-) - K_k v_k \quad (5.54)$$

gde je varijabla  $v_k$  uzorak belog šuma merenja iz (5.6). Analogno linearnom optimalnom Kalmanovom filtru, početni uslov u (5.9) garantuje da su greške predikcije i estimacije nepomerene što rezultira u:

$$E\{\tilde{x}_k(-)\} = 0; \quad E\{\tilde{x}_k(+)\} = 0; \quad k = 1, 2, \dots \quad (5.55)$$

Dokaz je zasnovan na matematičkoj indukciji [46] – [48], [187], [188]. Pored toga, prema hipotezama (5.7) – (5.8), slučajne greške  $\tilde{x}_k(+)$  i  $\tilde{x}_k(-)$  nisu u korelaciji sa šumom stanja i merenja, što se formalno može zapisati u matematičkoj formi kao:

$$E\{\tilde{x}_k(+)w_k^T\} = 0, \quad E\{\tilde{x}_k(-)v_k^T\} = 0 \quad (5.56)$$

gde 0 u (5.56) označava nula-matricu. Usled (5.54) i (5.56), kovarijansa greške predikcije u (5.40) se može izračunati u svakom koraku rekurzivnom relacijom:

$$P_k = M_k - K_k H_k M_k - M_k H_k^T K_k^T + R_k K_k K_k^T \quad (5.57)$$

gde je  $R_k$  varijansa šuma merenja iz (5.9). Zamenom pojačanja  $K_k$  iz (5.42) u (5.57) i izračunavanjem traga matrice, za aproksimativni kriterijum minimalne varijanse (5.40) dobija se sledeći izraz:

$$J_1(\Gamma_k) = \text{Trace}M_k - 2s_k^{-2}\alpha\text{Trace}\Gamma_k H_k^T H_k M_k + s_k^{-4}\alpha^2 R_k \text{Trace}\Gamma_k H_k^T H_k \Gamma_k \quad (5.58)$$

gde je  $s_k$  faktor skaliranja definisan u (5.43). U izvođenju (5.58) koristi se činjenica da treći matrični član u (5.57) predstavlja transponovani drugi matrični član, odakle potiče i faktor 2 u drugom sabirku u (5.58), pošto matrica i njena transponovana matrica imaju isti trag. Naredni korak u izvođenju se zasniva na primeni pravila za parcijalni izvod traga proizvoda matrica [46] – [48], [187], [188].

$$\frac{\partial}{\partial A} \text{Trace}BAC = B^T C^T; \quad \frac{\partial}{\partial A} \text{Trace}ABA^T = 2AB \quad \text{za } B = B^T \quad (5.59)$$

Poređenjem (5.59) sa (5.58), zaključuje se da je  $B = I$ ,  $A = \Gamma_k$ , i  $C = H_k^T H_k M_k$  za drugi član u (5.58), dok je  $A = \Gamma_k$ , i  $B = H_k^T H_k$  za treći član. Uzimajući u obzir ove ekvivalencije, parcijalnim diferenciranjem (5.58) i izjednačavanjem rezultujuće matrične jednačine sa nula-matricom, dobija se sledeća algebarska jednačina:

$$\frac{\partial J_1(\Gamma_k)}{\partial \Gamma_k} = -2s_k^{-2}\alpha M_k H_k^T H_k + 2s_k^{-4}\alpha^2 R_k \Gamma_k H_k^T H_k = 0 \quad (5.60)$$

Matrična jednačina (5.60) zahteva dalja uprošćavanja kako bi se došlo do izvodljivog suboptimalnog rešenja za  $\Gamma_k$ . Prvo, pošto je vrednost optimalnog koeficijenta statističke linearizacije  $\alpha$  u (5.41) blizu jedinice, u drugom članu (5.60)  $\alpha^2$  se može zameniti sa  $\alpha$ . Štaviše, korišćenjem (5.43) dalje sledi:

$$s_k^{-2}R_k = (R_k^{-1}H_k M_k H_k^T + 1)^{-1} \approx 1 \quad (5.61)$$

Naime, prvi član u zagradi u (5.61) je proporcionalan “nesigurnosti“ predikcije, izraženoj kovarijansom  $M_k$  u (5.44) i (5.48), ali obrnuto proporcionalan srednjoj snazi šuma merenja  $R_k$  u (5.9). Štaviše, predloženi nelinearni filter, zasnovan na tehnici vinzorizacije iz (5.38) i (5.47), ispunjava zahtev efikasne robusnosti. U tom smislu, on je skoro efikasan kao optimalni linearni Kalmanov filter u slučaju čisto Gausovih opservacija, a zadržava dobru efikasnost sa pojavom autlajera unutar Gausovih opservacija. Prema tome, kovarijansa greške estimacije  $P_k$  u (5.48) je mala, tako da je nesigurnost predikcije,  $M_k$ , direktno proporcionalna totalnoj srednjoj snazi šuma stanja  $Q_k$  iz (5.9). Štaviše, srednja snaga šuma merenja  $R_k$ , određena na osnovu (5.45), značajno je veća od srednje snage šuma stanja  $Q_k$ . Kao posledica toga, jednačina (5.61) se svodi na jediničnu vrednost. Uz usvojene aproksimacije, relacija (5.60) poprima oblik:

$$\frac{\partial J_1(\Gamma_k)}{\partial \Gamma_k} \approx -2s_k^{-2}\alpha(M_k - \Gamma_k)H_k^T H_k = 0 \quad (5.62)$$

Imajući u vidu (5.42) i (5.62) sledi:

$$M_k = \Gamma_k; \quad K_k = s_k^{-2}\alpha M_k H_k^T \quad (5.63)$$

Konačno, zamenom (5.63) u (5.57), dobija se kovarijansa greške estimacije izražena kao:

$$P_k = M_k - \alpha s_k^{-2} M_k H_k^T H_k M_k - \alpha s_k^{-2} M_k H_k^T H_k M_k + \alpha^2 s_k^{-4} R_k M_k H_k^T H_k M_k \quad (5.64)$$

Primenom dve pomenute aproksimacije,  $\alpha^2 \approx \alpha$  i (5.61), kao i relacije za pojačanje  $K_k$  iz (5.63), relacija (5.64) se svodi na (5.46).

Konačno, počevši od aproksimiranja  $\alpha$  sa  $\alpha^2$ , relacija (5.47) se dobija zamenom  $\Gamma_k$  iz (5.46) u (5.42) i uključivanjem u tako dobijenu jednačinu  $\psi$ -funkcije iz (5.38) umesto njene statističke linearne aproksimacije iz (5.41). Dakle, ako se koristi varijabla  $\alpha$  iz (5.52), koja ima vrednost blizu nule ili jedan, razumno je primeniti navedenu aproksimaciju  $\alpha^2 \approx \alpha$ . S druge strane, fiksna aproksimacija  $\alpha$  u (5.50) ili (5.51) jednaka je verovatnoći regularnih opservacija, a za mali ili umereni stepen kontaminacije  $\delta$  iz (5.39), verovatnoća je blizu jedinice, što opravdava prethodnu aproksimaciju. Naravno, vrednost  $\alpha$  opada sa povećanjem vrednosti stepena kontaminacije, smanjujući vrednosti faktora pojačanja  $K_k$  u (5.47) čime se potiskuje uticaj autlajera.

## 5.4.2 Optimalni koeficijent statističke linearizacije

Statistička linearizacija je tehnika statističke aproksimacije, čiji je osnovni princip da se data vektorska funkcija  $\psi(z)$  slučajnog vektorskog argumenta  $z$ , aproksimira linearnom matičnom formom:

$$\psi(z) = \alpha z + \beta \quad (5.65)$$

Parametri  $\alpha$  i  $\beta$  su matični koeficijenti koji treba da se odrede. Analogno problemu estimacije, definisanjem funkcije greške:

$$e = \psi(z) - \beta - \alpha z \quad (5.66)$$

ovi koeficijenti se mogu izračunati minimiziranjem kriterijuma srednje kvadratne greške:

$$J(\alpha, \beta) = \text{Trace} E\{e e^T\} = E\{e^T e\} = E\{\|e\|^2\} \quad (5.67)$$

Ovde je *Trace* trag matrice,  $\|\cdot\|$  označava Euklidsku normu, dok  $E\{\cdot\}$  predstavlja matematičko očekivanje, [46] – [48], [187], [188]. Zamenom (5.66) u (5.67) i promenom redosleda linearnih operatora, *Trace* i  $E\{\cdot\}$ , zajedno sa primenom pravila iz (5.59), dobija se:

$$\frac{\partial J(\alpha, \beta)}{\partial \beta} = 2E\{\psi(z) - \beta - \alpha z\} \quad (5.68)$$

Izjednačavanjem (5.68) sa nulom, dalje sledi:

$$\beta = E\{\psi(z)\} - \alpha E\{z\} \quad (5.69)$$

Štaviše, zamenom koeficijenta  $\beta$  iz (5.69) u (5.66) i (5.67), i ponovnim diferenciranjem tako dobijene jednačine po matičnom koeficijentu  $\alpha$ , koristeći pravila (5.59), dobija se:

$$\frac{\partial J(\alpha)}{\partial \alpha} = 2E\{\alpha \tilde{z} \tilde{z}^T + [E\{\psi(z)\} - \psi(z)] \tilde{z}^T\}; \quad \tilde{z} = E\{z\} - z \quad (5.70)$$

Izjednačavanjem (5.70) sa nulom i rešavanjem rezultujuće jednačine, dalje sledi:

$$\alpha = [E\{\psi(z)z^T\} - E\{\psi(z)\}E\{z^T\}]P_z^{-1}; P_z = E\{\tilde{z}\tilde{z}^T\} \quad (5.71)$$

Ako je  $\psi(z)$  vektorska funkcija višedimenzionalnog argumenta  $z$ , onda statistički linearizovana rešenja za (5.69) i (5.71) zahtevaju rešavanje višedimenzionalnih integrala, što sledi iz definicija odgovarajućih višedimenzionalnih matematičkih očekivanja. Ovo pretpostavlja da je zajednička FGV za komponente slučajnog vektora  $z$  data unapred. Štaviše, najčešće usvajana zajednička FGV je multidimenzionalna Gausova FGV. Proračun može biti mnogo pojednostavljen za nelinearnosti sa malim brojem promenljivih argumenata. Konkretno, ako su i funkcija  $\psi(z)$  slučajnog argumenta  $z$  i slučajna promenljiva  $z$  skalarni, sa nultim srednjim vrednostima, onda se  $\beta$  iz (5.69) i  $\alpha$  iz (5.71) svode na skalarnu determinističku veličinu, date kao:

$$\alpha = \frac{E\{\psi(z)z\}}{\sigma_z^2}; \beta = 0 \quad (5.72)$$

pri čemu je  $\sigma_z^2$  varijansa slučajnog argumenta  $z$ . Tačnije, jednačina (5.72) pretpostavlja da je argument  $z$  slučajna promenljiva nulte srednje vrednosti sa simetričnom FGV, dok je  $\psi$  neparna realna funkcija skalarnog argumenta  $z$ .

Konkretno, druga pretpostavka je ispunjena za Huberovu robusnu funkciju uticaja  $\psi_H(\cdot)$  iz (5.38). Štaviše, slučajni argument  $z$  u (5.72) odgovara normalizovanom skaliranom rezidualu merenja  $\tilde{\varepsilon}_k = \varepsilon_k/s_k$  iz (5.41), gde je rezidual  $\varepsilon_k$  dat sa (5.34), pri čemu je njegova standardna devijacija  $s_k$  izračunata u (5.43). Dakle, zamenom trenutnog merenja  $y_k$  iz (5.6) u (5.34), dobija se:

$$\tilde{\varepsilon}_k = \varepsilon_k/s_k = (H_k \tilde{x}_k(-) + v_k)/s_k; s_k = (H_k M_k H_k^T + R_k)^{1/2}; R_k = 1 \quad (5.73)$$

gde je greška predikcije  $\tilde{x}_k(-)$  slučajna promenljiva nulte srednje vrednosti sa kovarijansom  $M_k$  izračunatom u (5.48). Analogno optimalnom Kalmanovom filtru, greška  $\tilde{x}_k(-)$  u (5.73) je Gausovski raspodeljena ako oba bela Gausova šuma  $w_k$  i  $v_k$ , kao i vektor početnog stanja  $x_0$  zadovoljavaju pretpostavke (5.6) – (5.9). Stoga, može se pisati:

$$\tilde{x}_k(-) \sim N(\cdot | 0, M_k) \quad (5.74)$$

Posmatrajući (5.74), slučajna promenljiva  $H_k \tilde{x}_k(-)/s_k$  u (5.73) takođe ima Gausovu FGV definisanu sa:

$$H_k \tilde{x}_k(-)/s_k \sim N(\cdot | 0, H_k M_k H_k^T / s_k^2) \quad (5.75)$$

Štaviše, šum merenja  $v_k$  je ograničen na familiju mešovitih Gausovih FGV iz (5.39), što daje jediničnu nominalnu varijansu  $R_k$  u (5.73). Ovo, posledično, rezultira time da skalirana slučajna promenljiva  $v_k/s_k$  u (5.73) ima sledeću FGV:

$$v_k/s_k \sim (1 - \delta)N(\cdot | 0, 1/s_k^2) + \delta N(\cdot | 0, \sigma_0^2/s_k^2) \quad (5.76)$$

Dakle, normalizovani rezidual  $\tilde{\varepsilon}_k$  u (5.73) je definisan zbirom Gausovih slučajnih promenljivih sa nultom srednjom vrednošću, tako da je njegova uslovna FGV  $p(\tilde{\varepsilon}_k | Y^{k-1})$  definisana konvolucijom između osnovnih Gausovih FGV. Pored toga, pošto konvolucija između Gausovih FGV takođe daje Gausovu FGV, sa odgovarajućom srednjom vrednošću i kovarijansom, iz (5.73), (5.75) i (5.76) se dobija uslovna FGV skaliranog reziduala za data prethodna merenja:

$$p(\tilde{\varepsilon}_k|Y^{k-1}) = N(\cdot|0, H_k M_k H_k^T / s_k^2) \otimes \{(1 - \delta)N(\cdot|0, 1/s_k^2) + \delta N(\cdot|0, \sigma_0^2/s_k^2)\} \quad (5.77)$$

gde  $\otimes$  označava operator konvolucije [187]. Kako je konvolucija linearni operator, relacija (5.77) se može zapisati kao:

$$p(\tilde{\varepsilon}_k|Y^{k-1}) = (1 - \delta)N(\cdot|0, (H_k M_k H_k^T + 1)/s_k^2) + \delta N(\cdot|0, (H_k M_k H_k^T + \sigma_0^2)/s_k^2) \quad (5.78)$$

Uzimajući u obzir (5.43), varijansa druge normalne FGV u (5.78) se može aproksimirati sa:

$$\frac{\sigma_0^2(\sigma_0^{-2}H_k M_k H_k^T + 1)}{R_k(R_k^{-1}H_k M_k H_k^T + 1)} \approx \sigma_0^2; R_k = 1 \quad (5.79)$$

Desna strana aproksimacije u (5.79) sledi iz relacije (5.61), što daje:

$$\sigma_0^{-2}H_k M_k H_k^T + 1 \approx R_k^{-1}H_k M_k H_k^T + 1 \approx 1 \quad (5.80)$$

Zamenom (5.79) u (5.78) dobija se:

$$p(\tilde{\varepsilon}_k|Y^{k-1}) = (1 - \delta)N(\tilde{\varepsilon}_k|0, 1) + \delta N(\tilde{\varepsilon}_k|0, \sigma_0^2) \quad (5.81)$$

Kako je (5.81) simetrična FGV, ispunjena je i prva pretpostavka pod kojom je izveden izraz (5.72), čime se opravdava primena (5.72) u relaciji (5.41).

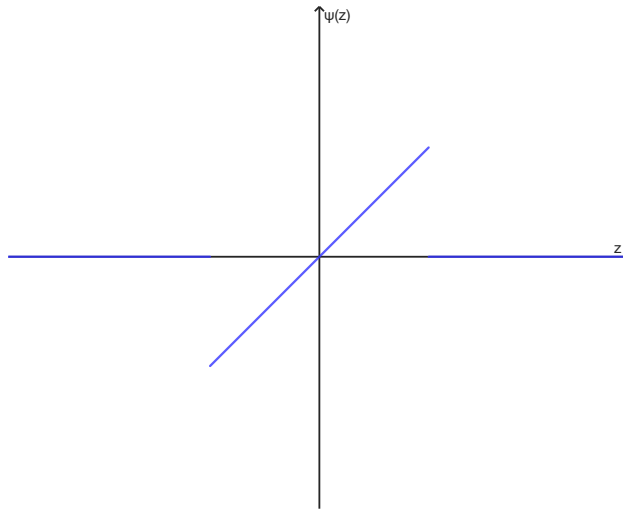
### 5.4.3 Funkcije uticaja

Tokom izvođenja i dizajna robusnog Kalmanovog filtra u prethodnim sekcijama korišćena je Huberova robusna funkcija uticaja  $\psi_H$  definisana u (5.38) i prikazana na slici 5.1. Glavna mana Huberove funkcije uticaja je činjenica da se autlajeri veoma velikog intenziteta nedovoljno potiskuju, odnosno tretiraju se isto kao i autlajeri manjeg intenziteta. Efikasnost estimatora dobijenog primenom ove funkcije uticaja zavisi od odabira praga saturacije Huberove funkcije uticaja  $\Delta$  u (5.38). Međutim, pored Huberove funkcije uticaja, u literaturi se mogu pronaći različite funkcije uticaja koje se mogu iskoristiti za robustifikaciju Kalmanovog filtra. Generalno, funkcija  $\rho$  treba da bude parna konveksna funkcija i da za veće apsolutne vrednosti argumenta ima rast sporiji od rasta kvadratne funkcije. Posebno, ako je  $\rho$  ograničena funkcija dobija se klasa nemonotonih (*redescending*)  $\psi$  funkcija. Funkcije uticaja iz ove klase imaju mogućnost potpunog potiskivanja autlajera veoma velikog intenziteta, ali je njima pridružena funkcija  $\rho$  nekonveksna, što može izazvati probleme u konvergenciji estimacione procedure usled nezadovoljavajućeg kvaliteta procena u početnim koracima.

Neke od funkcija uticaja koje se mogu naći u literaturi su [23], [37] – [40], [190], [195]:

- Nelinearnost odsecanja (slika 5.2)

$$\psi(z) = \begin{cases} z & ; |z| \leq \Delta \\ 0 & ; |z| > \Delta \end{cases} \quad (5.82)$$

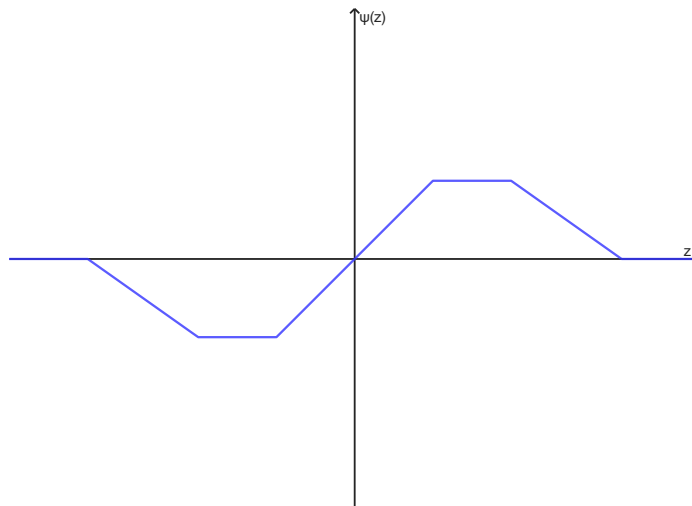


**Slika 5.2** Nelinearnost odsecanja ( $\Delta = 3$ ).

- *Hampel*-ova funkcija uticaja (slika 5.3)

$$\psi(z) = \begin{cases} z & 0 < |z| \leq a \\ a \operatorname{sgn}(z) & a < |z| \leq b \\ a \operatorname{sgn}(z) \frac{c - |z|}{c - b} & b < |z| \leq c \\ 0 & |z| \geq c \end{cases} \quad (5.83)$$

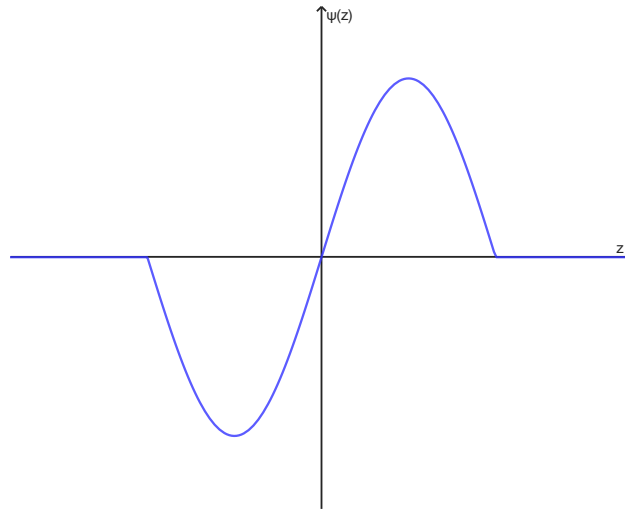
gde su  $a$ ,  $b$ , i  $c$  definisani parametri.



**Slika 5.3** *Hampel*-ova funkcija uticaja ( $a = 1.7$ ,  $b = 3.4$ ,  $c = 8.5$ ).

- *Andrews*-ova funkcija uticaja (slika 5.4)

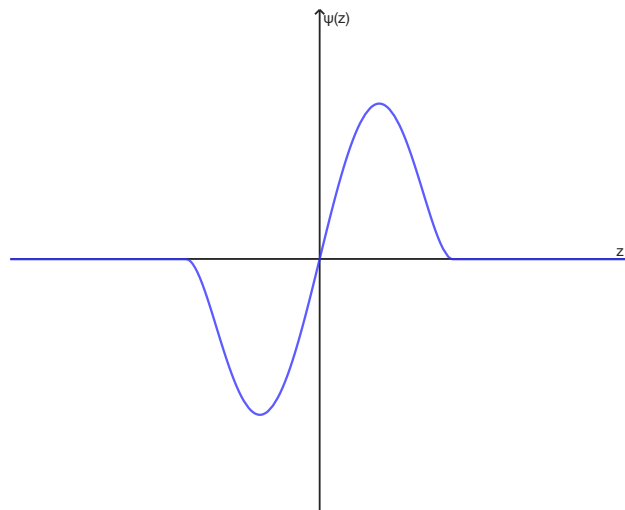
$$\psi(z) = \begin{cases} \sin\left(\frac{z}{\Delta}\right) & ; |z| \leq \Delta\pi \\ 0 & ; |z| > \Delta\pi \end{cases} \quad (5.84)$$



**Slika 5.4** Andrews-ova funkcija uticaja ( $1.5 < \Delta < 2$ ).

- Tukey-jeva funkcija uticaja (slika 5.5)

$$\psi(z) = \begin{cases} z \left(1 - \left(\frac{z}{\Delta}\right)^2\right)^2 & ; |z| \leq \Delta \\ 0 & ; |z| > \Delta \end{cases} \quad (5.85)$$

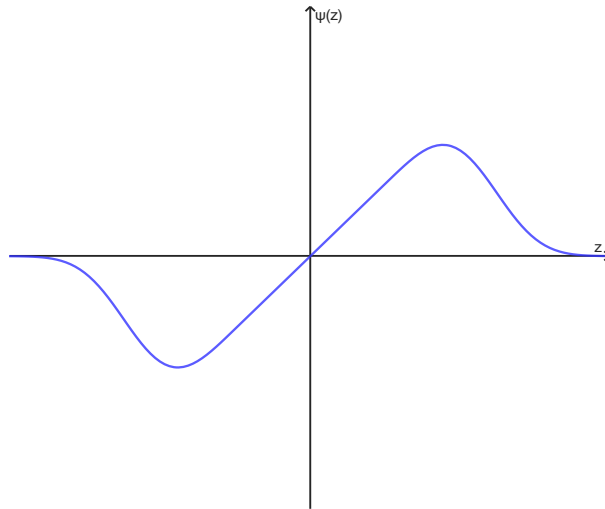


**Slika 5.5** Tukey-jeva funkcija uticaja ( $4.5 < \Delta < 6$ ).

Tukey-jeva robusna funkcija uticaja sa vrednošću parametra  $\Delta = 4.685$  obezbeđuje efikasnost od 95% za normalnu raspodelu.

- Generalizovana Gausova težinska funkcija (slika 5.6)

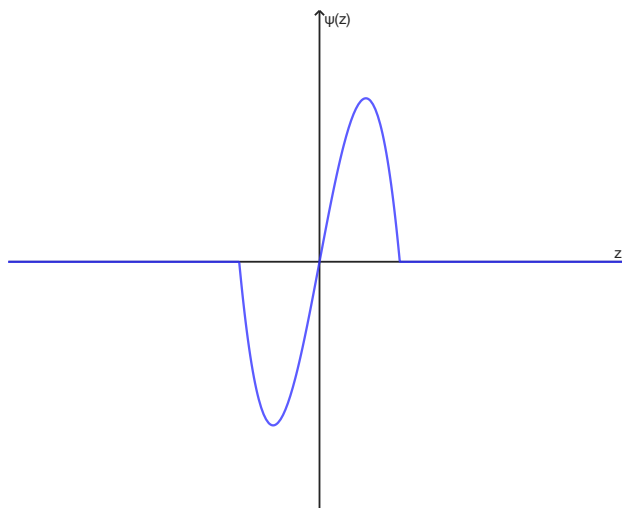
$$\psi_{a,b,c}(z) = \begin{cases} z & |z| \leq \Delta \\ z e^{-\frac{1}{2} \frac{(|z|-\Delta)^b}{a}} & |z| > \Delta \end{cases} \quad (5.86)$$



**Slika 5.6** Generalizovana Gausova težinska funkcija uticaja.

- *Smith*-ova funkcija uticaja (slika 5.7)

$$\psi(z) = \begin{cases} z(\Delta^2 - z^2) & |z| \leq \Delta \\ 0 & |z| > \Delta \end{cases} \quad (5.87)$$



**Slika 5.7** *Smith*-ova funkcija uticaja.

## 5.5 Adaptivni M-robustni Kalmanov filter

Za primenu Kalmanovog filtra pored modela sistema u prostoru stanja, neophodno je poznavati i statistike prvog i drugog reda šuma stanja i šuma merenja. Usvojene pretpostavke za šumove stanja i merenja su da ovi šumovi predstavljaju bele Gausove šumove nulte srednje vrednosti, da su nekorelisani u vremenu, nekorelisani međusobno, kao i da su nekorelisani sa slučajnim početnim stanjem sistema. Kovarijansa šuma merenja  $R_k$  u slučaju video praćenja opisuje preciznost samog algoritma za praćenje objekta u video sekvenci. Kovarijansa šuma stanja  $Q_k$  predstavlja srednju snagu nekontrolabilnih poremećaja u sistemu čije se karakteristike menjaju tokom vremena u zavisnosti od režima rada. Stoga, svrha šuma stanja jeste da opiše nemodelovanu dinamiku sistema,



odnosno da uključi u proces estimacije realne pojave koje se ne mogu potpuno tačno opisati matematičkim modelom.

Prilikom projektovanja Kalmanovog filtra, statistike prvog i drugog reda koje su neophodne za modelovanje šuma je potrebno poznavati unapred. Međutim, ove statistike nije lako precizno odrediti. Potrebno je posedovati dovoljno veliku bazu podataka na kojoj se ove statistike mogu odrediti što često nije slučaj. Takođe, ove statistike se mogu menjati tokom vremena. Sa neadekvatno određenim statistikama šuma, problem optimalne estimacije nije moguće rešiti. Stoga, može biti potrebno dizajnirati adaptivno rešenje, gde se statistike šumova određuju u svakom koraku rada estimatora na osnovu mernih podataka, modela sistema i dobijenih estimacija stanja. Kalmanov filter se na osnovu procena statistika šuma može bolje prilagoditi datoj situaciji u toku praćenja objekta u video sekvenci. Dizajn adaptivnog rešenja koje na osnovu M-robusnih estimacija stanja prilagođava matrice kovarijanse šuma merenja  $R_k$  i šuma stanja  $Q_k$  je prikazan u nastavku ovog poglavlja.

### 5.5.1 Adaptivna M-robusna estimacija statistika šuma merenja

Polazeći od jednačine merenja (5.6), šum merenja  $v_k$  se može estimirati na osnovu relacije:

$$r_k = y_k - H_k \hat{x}_k = H_k(x_k - \hat{x}_k) + v_k \quad (5.88)$$

gde je  $\hat{x}_k$  robusna estimacija (5.47) nepoznatog stanja  $x_k$  [185]. Uzimajući u obzir pretpostavku da je  $v_k$  beli Gausov šum nulte srednje vrednosti i da je  $\hat{x}_k$  nepomerena estimacija  $x_k$ , srednja vrednost greške estimacije je  $E\{\tilde{x}_k\} = E\{x_k - \hat{x}_k\} = 0$ , pa je srednja vrednost slučajnog vektora  $r_k$ :

$$m_r = E\{r_k\} = H_k E\{\tilde{x}_k\} + E\{v_k\} = 0; \quad m_v = E\{v_k\}; \quad m_r = m_v = 0 \quad (5.89)$$

a njegova varijansa se dalje izračunava kao:

$$V_r = E\{r_k r_k^T\} = H_k P_k H_k^T + R_k \quad (5.90)$$

gde je  $P_k$  kovarijaciona matrica robusne estimacije (5.46).

Statistike prvog i drugog reda  $m_r$  i  $V_r$  se mogu i estimirati na bazi realizovanog skupa (prozora) podataka  $r_i, i = 1, 2, \dots, n$  iz (5.88). Uvođenjem faktora skaliranja  $d_r$ , generisanog robusnim medijana estimatorom (*Median Absolute Deviation* - MAD) [23], [37], [185]:

$$d_r = \text{median} \left( \frac{|r_i - \text{median}(r_i)|}{0.6745} \right); \quad i = 1, 2, \dots, n_d \quad (5.91)$$

gde je dužina prozora elemenata  $n_d \in [3, 10]$ , a srednja vrednost  $m_r$  se može estimirati M-robusnim estimatorom tipa otežinjenih najmanjih kvadrata [23], [185], [195]:

$$\hat{m}_r = \frac{\sum_{i=1}^{n_r} \omega_{i0} r_i}{\sum_{i=1}^{n_r} \omega_{i0}} \quad (5.92)$$

pri čemu je robusni težinski koeficijent  $\omega_{i0}$ :

$$\omega_{i0} = \begin{cases} \psi\left(\frac{r_i - \hat{m}_{r0}}{d_r}\right) & ; r_i \neq \hat{m}_{r0} \text{ i } d_r \neq 0 \\ \frac{r_i - \hat{m}_{r0}}{d_r} & \\ 1 & ; r_i = \hat{m}_{r0} \text{ i/ili } d_r = 0 \end{cases} \quad (5.93)$$

Faktor 0.6745 u (5.91) se koristi jer će procenjeni faktor skaliranja, a koji zapravo predstavlja procenu nepoznate standardne devijacije šuma, konvergirati ka stvarnoj standardnoj devijaciji šuma ako je broj uzoraka veliki i ako su uzorci generisani Gausovom raspodelom. U izrazu (5.93),  $\hat{m}_{r0}$  predstavlja početno pogađanje srednje vrednosti  $\hat{m}_r$  iz (5.89) i najčešće se usvaja da je  $\hat{m}_{r0} = 0$ , a  $\psi$  označava robusnu funkciju uticaja. Za procenu  $m_r$  u (5.92) najčešće se koristi skup (prozor) od  $n_r \in [50, 500]$  odbiraka.

Dalje se M-robusna procena varijanse  $V_r$  iz (5.90) izračunava kao:

$$\hat{V}_r = \frac{\frac{1}{n_r} \sum_{i=1}^{n_r} \psi^2\left(\frac{r_i - \hat{m}_r}{d_r}\right)}{\left[\frac{1}{n_r} \sum_{i=1}^{n_r} \psi'\left(\frac{r_i - \hat{m}_r}{d_r}\right)\right]^2} \quad (5.94)$$

za šta se takođe koristi prozor od  $n_r \in [50, 500]$  odbiraka [23], [37], [185], [195].

Konačno, aproksimacijom nepoznate varijanse  $V_r$  iz (5.90) M-robusnom estimacijom  $\hat{V}_r$  iz (5.94), nepoznata varijansa šuma merenja  $v_k$  se može proceniti kao:

$$\hat{R}_k = |\hat{V}_r - H_k P_k H_k^T| \quad (5.95)$$

gde  $|\cdot|$  označava apsolutnu vrednost koja je uvedena kako bi se izbegle potencijalne negativne vrednosti procena u početnim koracima estimacione procedure, kada po pravilu tačnost procena nije zadovoljavajuća.

## 5.5.2 Adaptivna M-robusna estimacija statistika šuma stanja

Analogno postupku za estimaciju statistika šuma merenja u sekciji 5.5.1, mogu se estimirati i statistike prvog i drugog reda šuma stanja  $w_k$  [185].

Polazeći od jednačine stanja (5.5) i jednačine za robusnu estimaciju stanja (5.47), može se proceniti šum stanja  $w_k$  na osnovu definisane veličine  $q_k$ :

$$q_k = [G_k^T G_k]^{-1} G_k^T [\hat{x}_{k+1} - F_k \hat{x}_k] \quad (5.96)$$

Kako se estimacije  $\hat{x}_k$  formiraju na osnovu slučajnog merenja, definisana veličina  $q_k$  se takođe može posmatrati kao slučajna veličina, sa srednjom vrednošću  $m_q$  i varijansom  $V_q$ .

Uz pretpostavku da je  $w_k$  beli Gausov šum nulte srednje vrednosti i da je  $\hat{x}_k$  nepomerena estimacija stanja  $x_k$ , iz (5.96) i (5.5) sledi da je srednja vrednost slučajnog vektora  $q_k$ :

$$m_q = E\{q_k\} = E\{w_k\} = 0; \quad m_w = E\{w_k\}; \quad m_q = m_w = 0 \quad (5.97)$$

pa je varijansa  $V_q$ :

$$V_q = E\{q_k q_k^T\} \quad (5.98)$$

Kao i u sekciji 5.5.1, ove statistike se mogu estimirati na bazi realizovanog skupa (prozora) podataka  $q_i, i = 1, 2, \dots, n$ , iz (5.96). Po analogiji sa sekcijom 5.5.1, uvodi se faktor skaliranja  $d_q$  na bazi MAD estimatora:

$$d_q = \text{median} \left( \frac{|q_i - \text{median}(q_i)|}{0.6745} \right); i = 1, 2, \dots, n_d \quad (5.99)$$

gde  $n_d \in [3, 10]$ , a srednja vrednost  $m_q$  se estimira primenom M-robusnog estimatora:

$$\hat{m}_q = \frac{\sum_{i=1}^{n_q} \omega_{i0} q_i}{\sum_{i=1}^{n_q} \omega_{i0}} \quad (5.100)$$

pri čemu je  $\omega_{i0}$ :

$$\omega_{i0} = \begin{cases} \frac{\psi \left( \frac{q_i - \hat{m}_{q0}}{d_q} \right)}{\frac{q_i - \hat{m}_{q0}}{d_q}} & ; q_i \neq \hat{m}_{q0} \text{ i } d_q \neq 0 \\ 1 & ; q_i = \hat{m}_{q0} \text{ ili } d_q = 0 \end{cases} \quad (5.101)$$

U izrazu (5.101),  $\hat{m}_{q0}$  predstavlja početno pogađanje srednje vrednosti  $\hat{m}_q$  i najčešće se usvaja da je  $\hat{m}_{q0} = 0$ , a  $\psi$  označava robusnu funkciju uticaja. Broj uzoraka  $n_q$  je u opsegu  $[50, 500]$ .

Dalje se M-robusna procena varijanse  $V_q$  iz (5.98) izračunava po analogiji sa (5.94) kao:

$$\hat{V}_q = \frac{\frac{1}{n_q} \sum_{i=1}^{n_q} \psi^2 \left( \frac{q_i - \hat{m}_q}{d_q} \right)}{\left[ \frac{1}{n_q} \sum_{i=1}^{n_q} \psi' \left( \frac{q_i - \hat{m}_q}{d_q} \right) \right]^2} \quad (5.102)$$

za šta se takođe koristi skup od  $n_q \in [50, 500]$  odbiraka.

Na osnovu (5.96), a uzimajući u obzir (5.5) – (5.9) i (5.46) – (5.48), nepoznata varijansa  $V_q$  u (5.98) se može izračunati na osnovu izraza:

$$V_q = [G_k^T G_k]^{-1} G_k^T [F_k P_k F_k^T - P_{k+1}] G_k [G_k^T G_k]^{-1} + Q_k \quad (5.103)$$

Polazeći od (5.103) i aproksimacijom nepoznate varijanse  $V_q$  iz (5.98) M-robusnom estimacijom  $\hat{V}_q$  iz (5.102), nepoznata varijansa šuma stanja  $w_k$  se može aproksimirati kao:

$$\hat{Q}_k = |\hat{V}_q - [G_k^T G_k]^{-1} G_k^T [F_k P_k F_k^T - P_{k+1}] G_k [G_k^T G_k]^{-1}| \quad (5.104)$$

gde je po analogiji sa (5.95) uvedena apsolutna vrednost  $|\cdot|$  kako bi se izbegle potencijalne negativne vrednosti procene varijanse u fazi inicijalizacije.

## 6 Robusni sistem za praćenje pokretnih objekata u sekvenci kratkotalasnih infracrvenih slika

Predloženi sistem za praćenje pokretnih objekata opisan u sekciji 4.3 i prikazan blok dijagramom na slici 4.4 se u svakom koraku procesiranja frejmova video sekvence oslanja na predikciju stanja objekta Kalmanovim filtrom. Lokalizacija objekta korelacionim filtrom na trenutnom frejmu se vrši na poziciji koja se dobija predikcijom Kalmanovog filtra. Dimenzije prozora za pretragu na datoj poziciji se takođe dobijaju iz predikcije Kalmanovog filtra. Sam Kalmanov filter zapravo vrši predikciju dimenzija okružujućeg okvira objekta, ali su prozor za pretragu, odnosno dimenzionalnost korelacionog filtra i dimenzije samog objekta vezani relacijom po kojoj su dimenzije prozora za pretragu veće 3 puta u odnosu na dimenzije okružujućeg okvira objekta. Ova veza korelacionog filtra i predikcije Kalmanovog filtra prvenstveno treba da obezbedi uspešno prevazilaženje problema okluzija u video praćenju. Iz tog razloga, veoma je važna pravilna inicijalizacija samog Kalmanovog filtra.

U koraku inicijalizacije predloženog sistema za praćenje, na prvom frejmu video sekvence, vektor stanja objekta  $X$  iz (4.94) koji se koristi u modelu Kalmanovog filtra se inicijalizuje istim podacima (pozicijom i dimenzijama objekta) kao i model objekta u korelacionom filtru. Prva dva stanja vektora  $X$  označavaju koordinate objekta u ravni slike, dok treće i četvrto stanje predstavljaju širinu i visinu objekta, respektivno, takođe u ravni slike. Ova četiri stanja se inicijalizuju stvarnim (*ground-truth*) podacima kojima se definiše objekat koji je od interesa za praćenje. Ostala četiri stanja, odnosno brzine se inicijalizuju nultim vrednostima, te inicijalni vektor stanja objekta ima oblik:

$$X_0 = \begin{bmatrix} x_{gt} \\ y_{gt} \\ w_{gt} \\ h_{gt} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (6.1)$$

gde  $gt$  u indeksu prva četiri stanja označava *ground-truth* podatke.

Pored inicijalizacije samog vektora stanja objekta, za opisivanje modela Kalmanovog filtra, potrebno je definisati i matrice prelaza stanja  $F$  i šuma stanja  $G$  iz (5.5), kao i matricu merenja  $H$  iz (5.6). U skladu sa definisanim vektorom stanja  $X$ , matrice  $F$  i  $H$  su date kao:

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 & T & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & T & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & T & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & T \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (6.2)$$

gde  $T$  označava periodu odabiranja, za koju se usvaja vrednost 1 obzirom da se praćenje vrši na svakom frejmu video sekvence. Matrica šuma stanja  $G$  je usvojena kao jedinična matrica dimenzija  $8 \times 8$ .

Pored definisanja navedenih matrica, veoma je važna i pravilna inicijalizacija kovarijacionih matrica šuma stanja  $Q$  i šuma merenja  $R$  iz (5.9). U problemu praćenja više objekata po metodologiji praćenja na osnovu detekcija, vektor stanja objekata se u [69], [70], [197] definiše na sličan način kao i vektor stanja (4.94), pri čemu su pored same pozicije objekta u vektor stanja uključene i informacije o dimenzijama objekta. Dok [69] koristi predefinisane vrednosti matrica  $Q$  i  $R$ , u [70] i [197] se inicijalizacija ovih matrica vrši na osnovu inicijalnih vrednosti vektora stanja. Prateći metod inicijalizacije u [197] i primenom predloženih koeficijenta  $\sigma_p = 0.05$  i  $\sigma_v = 0.00625$  iz [70], inicijalizacija matrica  $Q$  i  $R$  na osnovu inicijalnih *ground-truth* vrednosti  $w_{gt}$  i  $h_{gt}$  iz (6.1) je izvršena prema:

$$Q = \text{diag} \left( \begin{bmatrix} (\sigma_p w_{gt})^2 \\ (\sigma_p h_{gt})^2 \\ (\sigma_p w_{gt})^2 \\ (\sigma_p h_{gt})^2 \\ (\sigma_v w_{gt})^2 \\ (\sigma_v h_{gt})^2 \\ (\sigma_v w_{gt})^2 \\ (\sigma_v h_{gt})^2 \end{bmatrix} \right), \quad R = \text{diag} \left( \begin{bmatrix} (\sigma_p w_{gt})^2 \\ (\sigma_p h_{gt})^2 \\ (\sigma_p w_{gt})^2 \\ (\sigma_p h_{gt})^2 \end{bmatrix} \right) \quad (6.3)$$

gde *diag* označava dijagonalnu matricu.

Nakon inicijalizacije Kalmanovog filtra na čiju se predikciju stanja objekta oslanjaju ostali moduli u predloženom sistemu za praćenje, ali i čija estimacija stanja objekta predstavlja izlaz sistema, može se pristupiti analizi rada celokupnog sistema. U nastavku ovog poglavlja data je detaljna analiza performansi praćenja pokretnog objekta u skupu SWIR video sekvenci koji je predstavljen u sekciji 2.3.

## 6.1 Sistem za praćenje korelacionim i standardnim Kalmanovim filtrom

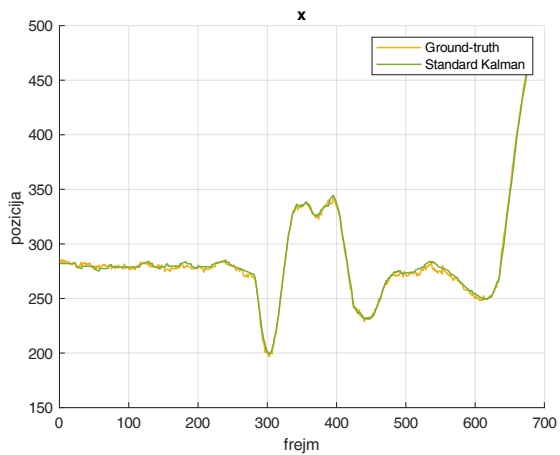
U regularnim uslovima korelacioni filtri pokazuju veoma dobre performanse praćenja objekta. Važna prednost korelacionih filtara jeste brzina rada, posebno zahvaljujući obučavanju na uzorcima koji se dobijaju cikličnim pomeranjima osnovnog uzorka objekta. Osobina dijagonalizacije cikličnih matrica u Furijeovom domenu omogućava da se računaska kompleksnost korelacionih filtara značajno smanji i svede samo na množenje elemenata vektora obeležja u Furijeovom domenu. Dodatno poboljšanje diskriminativnih svojstava korelacionih filtara se postiže uvođenjem regularizacije pozadinskim kontekstom iz okoline objekta, čime se obezbeđuju stvarni negativni uzorci slike za obučavanje filtra. Na taj način se postiže uspešnije praćenje prilikom različitih vrsta deformacija objekta, ali i prilikom pojave klatera. Osnovni razlog fuzije Kalmanovog filtra sa korelacionim filtrom u predloženom sistemu za praćenje jeste mogućnost prevazilaženja okluzija, koje još uvek za same korelacione filtre predstavljaju nepremostivu prepreku za održanje praćenja. Posebno u slučaju potpunih dugotrajnih okluzija, praćenje koje se oslanja samo na korelacioni filter ostaje zaustavljeno na poziciji pojave okluzije. Zbog mogućnosti nastavka praćenja objekta nakon okluzije i u slučaju da objekat vrši manevar, u predloženom sistemu za praćenje je uveden i modul za ponovnu identifikaciju objekta nakon okluzije koji podrazumeva pretragu u široj zoni u odnosu na onu koju obezbeđuje sam korelacioni filter.

Analiza rada predloženog sistema za praćenje primenom standardnog Kalmanovog filtra omogućava uvid u rad sistema u regularnim uslovima, ali i da se prepoznaju situacije koje dovode do velikih grešaka u estimaciji pozicije objekta prikazane u sekciji 4.3.1. Na osnovu ove analize se dalje vrši odlučivanje na koji način se može projektovati sistem sa robusnim estimatorom koji treba da postigne ciljeve praktične robusnosti.

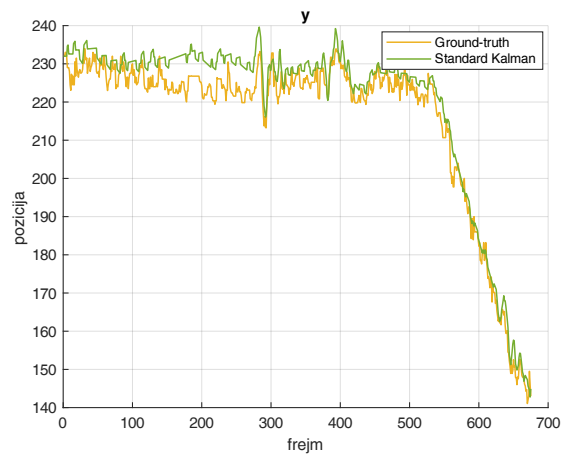
Performanse sistema u regularnim uslovima u sekvenci 14 iz skupa SWIR video sekvenci su prikazane na slici 6.1. Karakteristika analizirane sekvence jeste da u njoj nema prisustva okluzija. Objekat pored promene veličine i orijentacije tokom kretanja vrši i manevar, što se može videti u horizontalnom pravcu kretanja na slici 6.1 a). Sa priloženih grafika se može videti da sistem koji se oslanja na standardni Kalmanov filter u analiziranom scenariju veoma dobro prati pokretni objekat.

Performanse sistema u scenariju sa prisutnim statičnim okluzijama iz sekvence 5 iz skupa SWIR video sekvenci su prikazane na slici 6.2. Osenčeni delovi na slici predstavljaju period tokom kog je objekat pod okluzijom. U ovoj sekvenci objekat ima dominantno kretanje u jednom smeru u horizontalnom pravcu. Sa priloženih grafika se primećuje da sistem uspeva da isprati objekat od početka do kraja sekvence. Međutim, u trenucima pojave okluzija se vidi veliko odstupanje estimirane od stvarne pozicije objekta. U sistemu gde je kamera postavljena na pan-tilt pozicioner, ove greške bi generisale upravljačke signale koji bi sistem pozicionirali tako da se objekat nađe potpuno izvan vidnog polja kamere. Stoga, ovaj tip grešaka u estimaciji pozicije predstavlja greške koje se mogu tretirati kao autlajeri.

Na slici 6.3 su prikazani rezultati analize standardnog Kalmanovog filtra u scenariju iz sekvence 6. Periodi okluzija su označeni osenčenim delovima, a prva okluzija koja se javlja jeste pokretna okluzija.

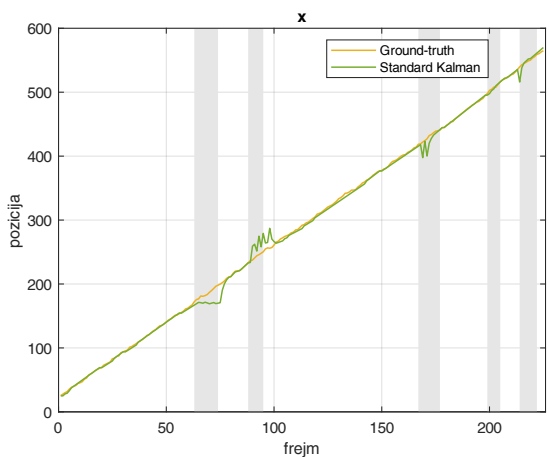


a)

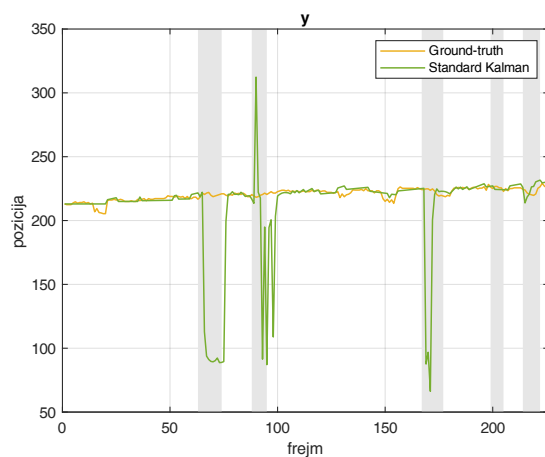


b)

**Slika 6.1** Poređenje stvarne pozicije i pozicije objekta estimirane standardnim Kalmanovim filtrom: a) horizontalni, b) vertikalni pravac.

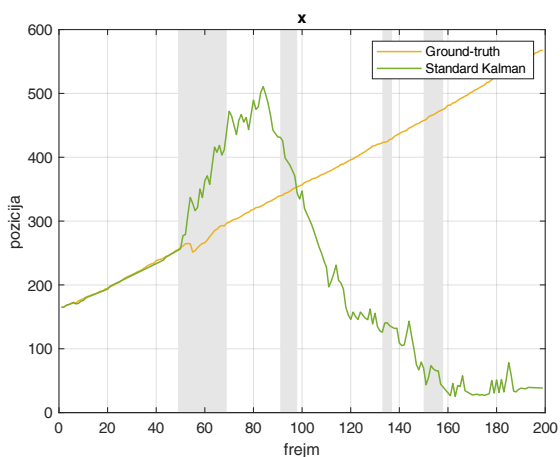


a)

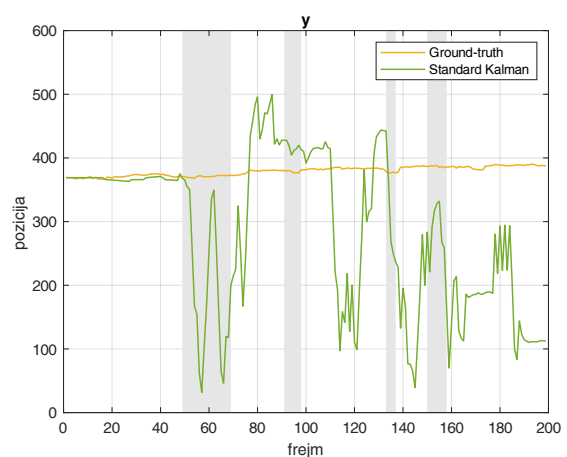


b)

**Slika 6.2** Poređenje stvarne pozicije i pozicije objekta estimirane standardnim Kalmanovim filtrom u scenariju sa statičnim okluzijama: a) horizontalni, b) vertikalni pravac.



a)

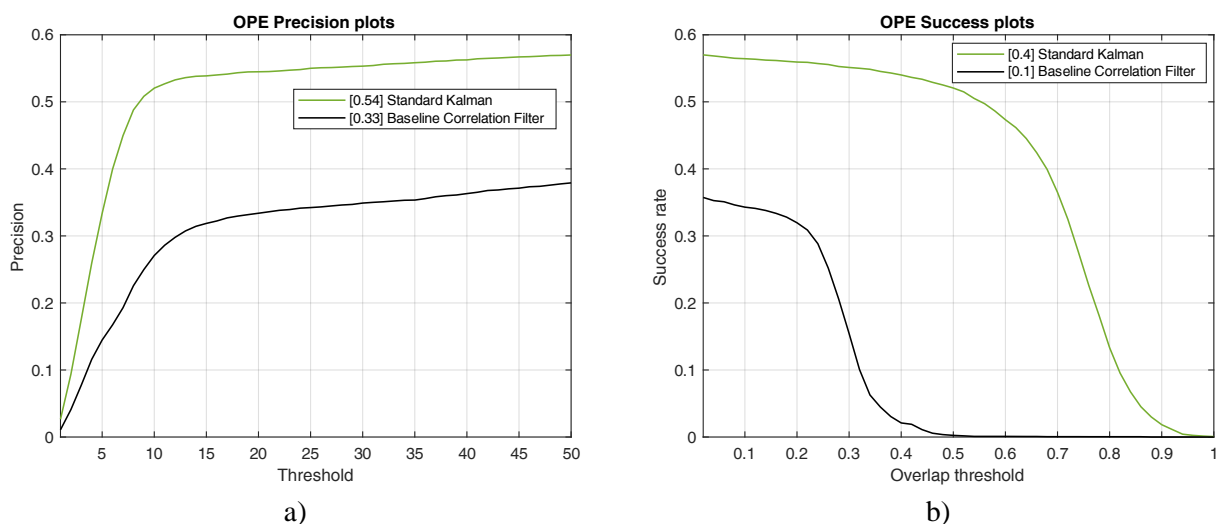


b)

**Slika 6.3** Poređenje stvarne pozicije i pozicije objekta estimirane standardnim Kalmanovim filtrom u scenariju sa dugotrajnom pokretnom okluzijom i više statičnih okluzija: a) horizontalni, b) vertikalni pravac.

Pojavom prve okluzije, greška u estimaciji pozicije objekta se veoma brzo povećava, posebno u vertikalnom pravcu, što se može videti na slici 6.3 b). Za razliku od ishoda u scenariju na slici 6.2, u ovom slučaju nakon pojave okluzije praćenje objekta se više nije nastavilo. Razlog za gubitak praćenja leži u tome što je zona za pretragu pojavom velike greške u estimaciji pozicije značajno izmeštena u ravni slike u odnosu na poziciju objekta, pa je trajno izgubljena mogućnost ponovne detekcije objekta i nastavka praćenja. Kako ova greška u praćenju dostiže veoma veliki intenzitet u kratkom vremenskom intervalu, priroda ove greške se takođe može tretirati kao pojava autlajera.

Iako se primenom predloženog sistema za praćenje sa standardnim Kalmanovim filtrom mogu javiti greške u estimaciji pozicije velikog intenziteta, na sveukupne performanse praćenja se ipak može videti pozitivan uticaj. Poređenjem preciznosti praćenja bazičnog korelacionog filtra sa kontekstnom regularizacijom i sistema sa standardnim Kalmanovim filtrom na slici 6.4 a) i uspešnosti praćenja (eng. *Success plots*) na slici 6.4 b) vidi se da predloženi sistem ima primetno bolje performanse. Evaluacija je izvršena na celokupnom skupu SWIR video sekvenci. Primenjena je takozvana evaluacija jednim prolaskom (eng. *One Pass Evaluation* - OPE) [198], gde se svaka sekvenca evaluira samo jednom inicijalizacijom *ground-truth* okružujućim okvirom na prvom frejmu video sekvence. Grafik preciznosti demonstrira procenat frejmova na kojima je distanca između centra estimiranog i *ground-truth* okružujućeg okvira objekta manja od datog praga. Za rangiranje po ovom kriterijumu je korišćena vrednost preciznosti za prag od 20 piksela. Grafik uspešnosti (*Success plot*) prikazuje procenat frejmova na kojima je preklapanje mereno IOU metrikom između estimiranog i *ground-truth* okružujućeg okvira objekta veće od datog praga. Rangiranje po ovom kriterijumu je izvršeno prema površini ispod dobijenih kriva uspešnosti (eng. *Success rate*). Opisane metrike i kriterijumi će na isti način biti korišćeni za evaluaciju algoritama za praćenje u nastavku ovog poglavlja.



**Slika 6.4** Poređenje: a) preciznosti i b) uspešnosti - na celokupnom skupu SWIR video sekvenci bazičnog korelacionog filtra (sa kontekstnom regularizacijom) i predloženog sistema sa standardnim Kalmanovim filtrom.



## 6.2 Robusni Kalmanov filter

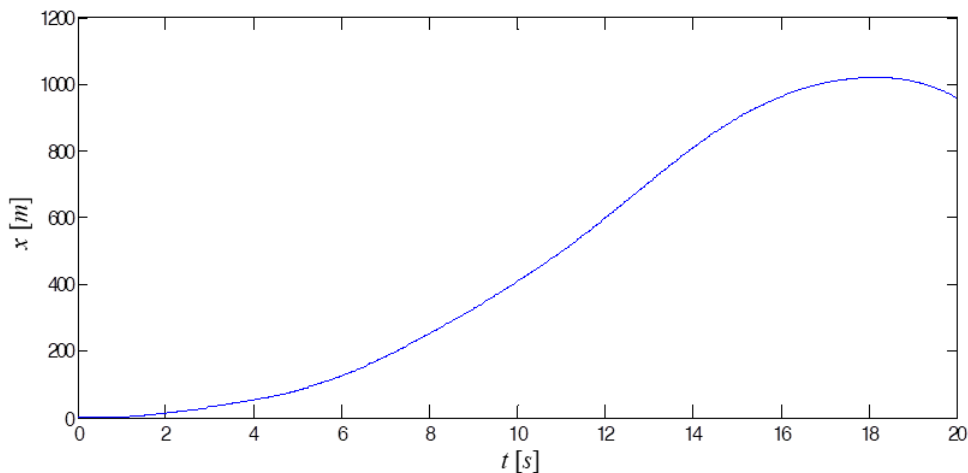
Pre analize performansi celokupnog sistema za praćenje korelacionim i M-robustnim Kalmanovim filtrom, biće izvršena analiza statistički linearizovanog M-robustnog Kalmanovog filtra sa fiksnim i adaptivnim koeficijentom statističke linearizacije.

### 6.2.1 Analiza statistički linearizovanog M-robustnog Kalmanovog filtra

Zarad potvrde osobina M-robustnog Kalmanovog filtra u situacijama kada su u mernim podacima prisutni autlajeri izvršena je analiza rada M-robustnog Kalmanovog filtra na simuliranim podacima. Simulacioni zadatak je modelovanje dinamike stanja sistema, korišćenjem kinematičkih jednačina kretanja [199]. Ako se sa  $x_k$ ,  $v_k$  i  $a_k$  označe pozicija, brzina i ubrzanje objekta, respektivno, u diskretnom vremenskom trenutku  $t_k = kT, k = 0, 1, \dots$ , pri čemu  $T$  označava uniformnu periodu odabiranja, a pod pretpostavkom da je ubrzanje konstantno tokom intervala uzorkovanja  $t_k \leq t < t_{k+1}$ , dobija se sledeći skup jednačina:

$$\begin{aligned}x_t &= x_k + v_k(t - t_k) + \frac{1}{2}a_k(t - t_k)^2 \\v_t &= v_k + a_k(t - t_k) \\a_t &= a_k\end{aligned}\tag{6.4}$$

Model (6.4) sa ubrzanjem koje ima nultu vrednost poznat je kao model konstantne brzine. Stoga, svako kretanje objekta koje se ne može predstaviti modelom konstantne brzine se može smatrati manevrom objekta [199]. Trajektorija kretanja objekta koja je korišćena u simulacijama je prikazana na slici 6.5.



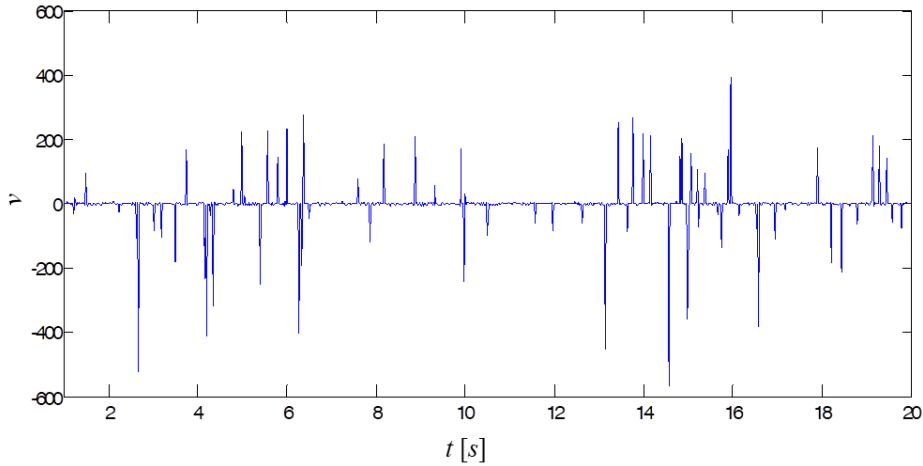
Slika 6.5 Trajektorija objekta korišćenja u simulacijama.

Sekvenca merenja je simulirana korišćenjem linearnog senzora pozicije, predstavljenog kao:

$$y_k = x_k + v_k\tag{6.5}$$

gde sekvenca belog šuma merenja  $\{v_k\}$  u (6.5) ima raspodelu (5.39). Uzorak slučajne promenljive  $v_k$  sa datom raspodelom se može generisati uzimanjem uzorka  $u$  koji pripada (0,1)-uniformnoj raspodeli.

Ako je vrednost ovog uzorka veća od  $\delta$ -vrednosti, uzorak  $v_k$  se generiše iz standardne Gausove raspodele nulte srednje vrednosti i jedinične varijanse. U suprotnom se uzorak  $v_k$  generiše iz kontaminirane Gausove raspodele nulte srednje vrednosti sa pretpostavljenom velikom varijansom  $\sigma_0^2 \gg 1$ . Tipičan šum merenja prikazan je na slici 6.6.



**Slika 6.6** Šum merenja generisan  $\delta$ -kontaminiranom Gausovom raspodelom.

Filtar je predstavljen dvodimenzionalnim, diskretnim, vremenski nepromenljivim modelom u prostoru stanja u obliku (5.5) i (5.6), dat kao:

$$\begin{aligned} x_{k+1} &= Fx_k + Gw_k \\ y_k &= Hx_k + v_k \end{aligned} \quad (6.6)$$

$$F = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad H = [1 \quad 0]$$

Matrica prelaza stanja  $F$  sledi direktno iz (6.4) za  $t = t_{k+1}$  i zanemarivanjem člana ubrzanja, dok se vektor merenja dobija iz (6.5). Sekvenca šuma stanja nulte srednje vrednosti  $w_k$  se uvodi kako bi se kompenzovala nemodelovana dinamika sistema, povezana sa nepoznatim manevrom objekta. Varijansa sekvence šuma stanja  $\{w_k\}$  je  $Q = 0.1$ , a šuma merenja  $\{v_k\}$  je  $R = 1$ . Za periodu odabiranja je usvojeno  $T = 0.02$  s.

Početna estimacija stanja  $\hat{x}_0$  i odgovarajuća kovarijaciona matrica  $P_0$  su izračunati korišćenjem suboptimalne procedure zasnovane na prva dva merenja [199]:

$$\hat{x}_0 = \begin{bmatrix} y_2 \\ \frac{y_2 - y_1}{T} \end{bmatrix}; \quad P_0 = \begin{bmatrix} 1 & 1/T \\ 1/T & Q + 2/T^2 \end{bmatrix} \quad (6.7)$$

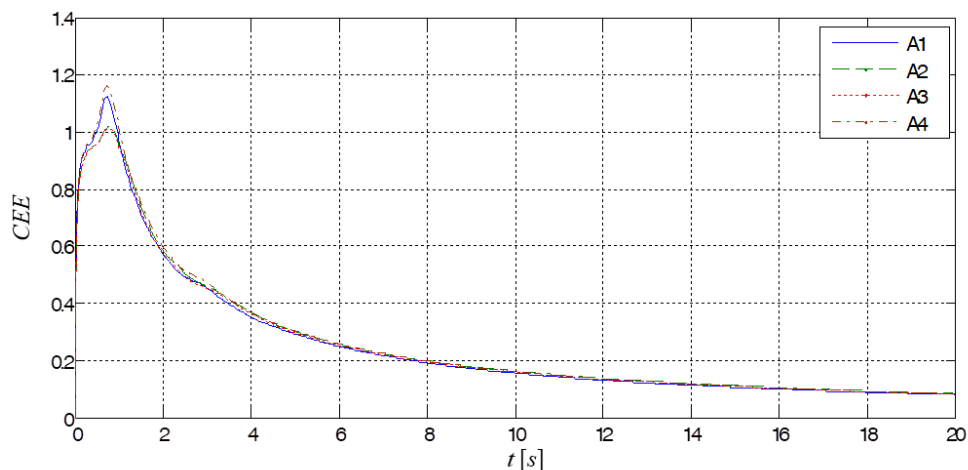
Izvršena je komparativna analiza linearnog optimalnog Kalmanovog filtra (5.10) – (5.16), nadalje označen kao A1; statistički linearizovanog M-robustnog Kalmanovog filtra (5.34), (5.38), (5.41), (5.43), (5.46) – (5.48) sa promenljivim koeficijentom statističke linearizacije iz (5.52), označenim kao A2; statistički linearizovanog M-robustnog Kalmanovog filtra (5.34), (5.38), (5.41), (5.43), (5.46) – (5.48) sa fiksnim koeficijentom statističke linearizacije u (5.50), označen kao A3; i kvazilinearne aproksimacije algoritma A2, zasnovane na linearnoj transformaciji reziduala u (5.47)

umesto nelinearne u (5.38), zajedno sa primenom istog nelinearnog procesiranja reziduala u proračunu adaptivne matrice Kalmanovog pojačanja  $K_k$  kao u A2, označene kao A4.

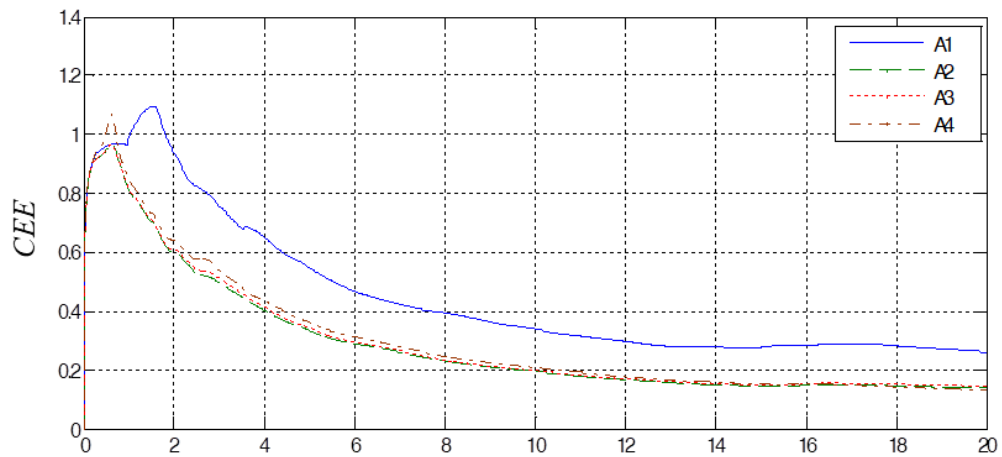
Komparacija performansi analiziranih filtara je izvršena u pogledu estimirane i stvarne pozicije, ali i prema kriterijumu kumulativne greške estimacije (eng. *Cumulative Estimation Error - CEE*):

$$CEE(k) = \frac{1}{k} \sum_{i=1}^k \frac{\|\hat{x}_k - x_k\|^2}{\|x_k\|^2} \quad (6.8)$$

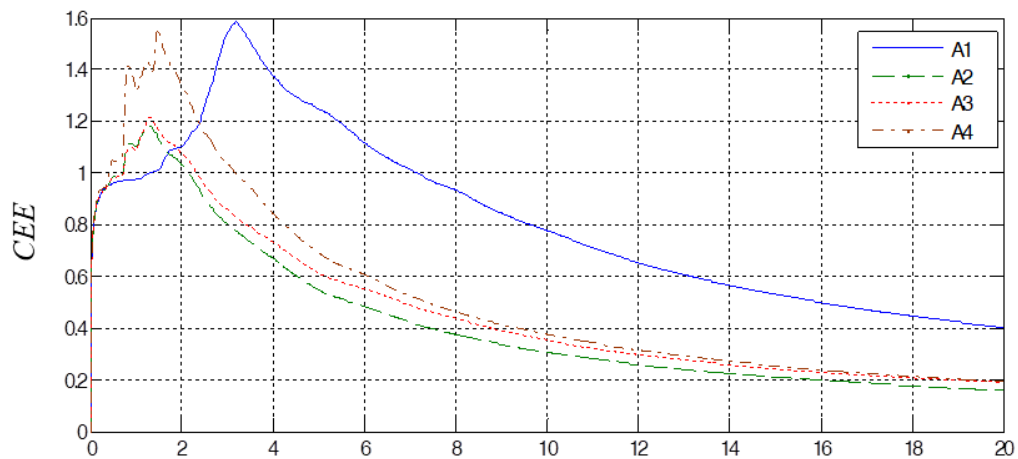
pri čemu  $\|\cdot\|$  označava Euklidsku normu, gde je  $x_k$  stvarna pozicija u trenutku  $k$  sa trajektorije prikazane na slici 6.5. Vrednosti CEE kriterijuma dobijene za posmatrane algoritme i različite realizacije šuma merenja prikazane su na slikama 6.7 – 6.9. Rezultati prikazani na slici 6.7 pokazuju da robustifikovani Kalmanovi filtri A2 – A4 zadovoljavaju zahtev efikasne robusnosti, budući da dobijene vrednosti kriterijuma (6.8) za ove algoritme nisu značajno veće u poređenju sa onim dobijenim za optimalni Kalmanov filter A1 u slučaju merenja koja imaju nekontaminiranu Gausovu raspodelu. Pored toga, estimatori A2 – A4 takođe zadovoljavaju i zahtev rezistentne robusnosti. Ovi estimatori imaju značajno manje vrednosti kriterijuma (6.8) u odnosu na optimalni Kalmanov filter, A1, u prisustvu odstupanja unutar Gausovih opservacija, kao što je prikazano na slikama 6.8 i 6.9. Delovi stvarne i estimirane trajektorije, generisani algoritmima A1 i A2, prikazani su na slici 6.10. Slični rezultati su dobijeni za algoritme A3 i A4. Međutim, analiza performansi estimatora korišćenjem stvarne i estimirane trajektorije nije pogodna, pošto su pozicije objekta na trajektoriji izražene u mnogo većim jedinicama, u poređenju sa vrednostima grešaka u estimaciji. Stoga, rezultati simulacija prema CEE kriterijumu predstavljeni na slikama 6.7 – 6.9 su pokazali da predloženi robusni estimatori A2 – A4 ispunjavaju zahteve praktične robusnosti.



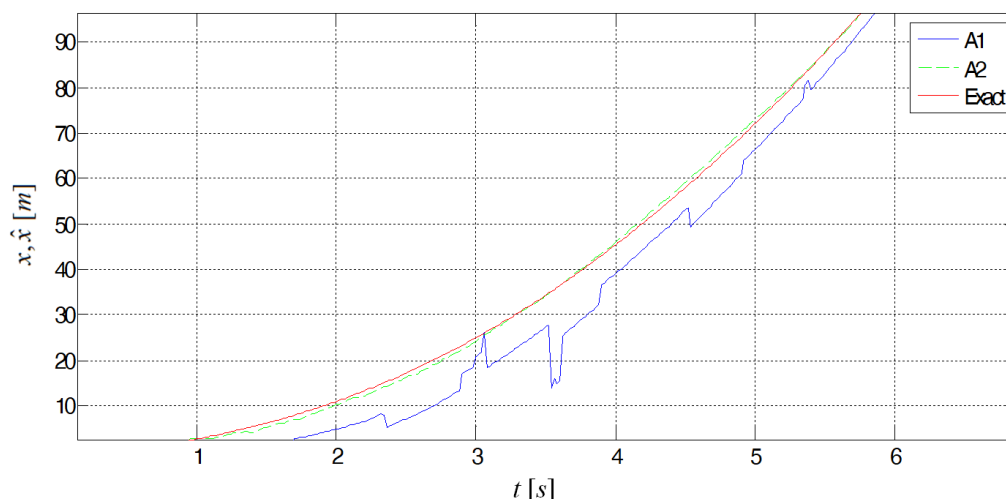
**Slika 6.7** Komparativna analiza estimatora prema CEE kriterijumu u slučaju merenja koja imaju Gausovu raspodelu šuma,  $\delta = 0$  u (5.39).



**Slika 6.8** Komparativna analiza estimatora prema CEE kriterijumu u slučaju merenja koja imaju  $\delta$  - kontaminiranu Gausovu raspodelu šuma (5.39),  $\delta = 0.07$ .



**Slika 6.9** Komparativna analiza estimatora prema CEE kriterijumu u slučaju merenja koja imaju  $\delta$  - kontaminiranu Gausovu raspodelu šuma (5.39),  $\delta = 0.15$ .

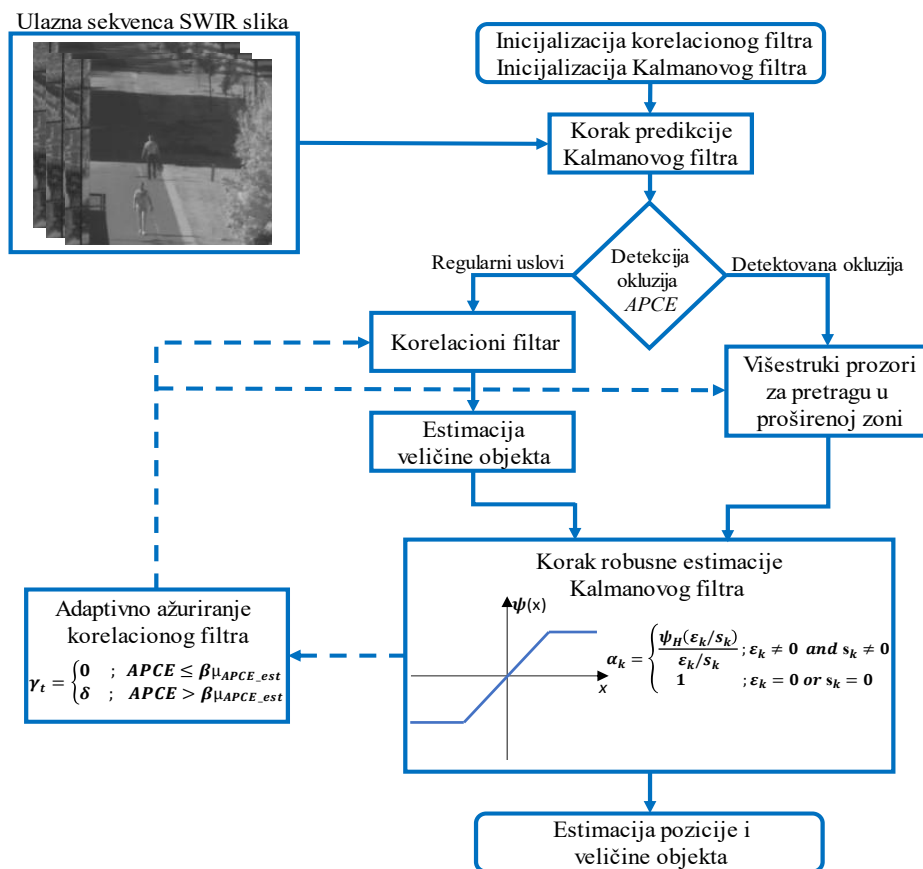


**Slika 6.10** Delovi stvarne trajektorije (*Exact*) i estimiranih trajektorija algoritmima A1 i A2.

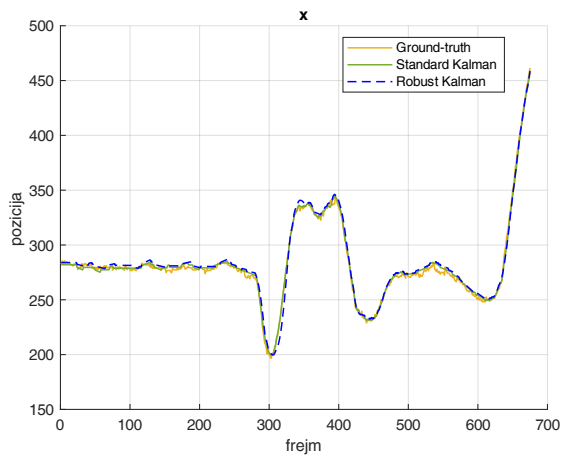
## 6.2.2 Sistem za praćenje korelacionim i M-robustnim Kalmanovim filtrom

U sekciji 6.1 su prikazane karakteristične situacije koje dovode do pojave grešaka velikog intenziteta koje se mogu tretirati kao autlajeri, a čiji se uticaj može umanjiti primenom robusnih tehnika estimacije. Nakon potvrde efikasnosti M-robustnog Kalmanovog filtra u potiskivanju autlajera u sekciji 6.2.1, naredni korak predstavlja dizajn sistema za praćenje pokretnih objekata korelacionim filtrom i M-robustnim Kalmanovim filtrom.

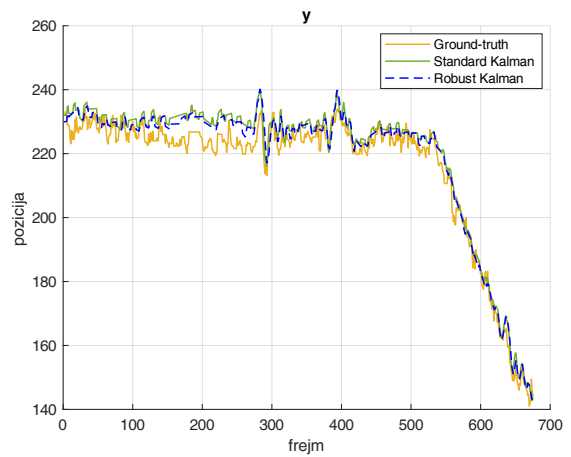
Analiza grešaka u estimaciji pozicije i grešaka u estimaciji veličine praćenog objekta u sekciji 4.3.1 pokazala je da se autlajeri javljaju samo u merenjima pozicije objekta. Kako se prema dizajnu modula za estimaciju veličine objekta u sekciji 4.3 veličina objekta (visina i širina okružujućeg okvira) može promeniti za 5 % ili 10 % između sukcesivnih frejmova, greške veoma velikog intenziteta, odnosno autlajeri se ne mogu javiti u merenjima veličine objekta. Stoga se korak robusne estimacije Kalmanovog filtra primenjuje samo na prve dve komponente vektora stanja objekta iz (4.94) koje predstavljaju poziciju objekta u horizontalnom i vertikalnom pravcu. Kako je u analizi u sekciji 6.2.1 robusni Kalmanov filter sa promenljivim koeficijentom statističke linearizacije pokazao bolje performanse u odnosu na primenu fiksnog koeficijenta, u predloženom sistemu za praćenje pokretnih objekata u SWIR video sekvenci je primenjen robusni Kalmanov filter sa promenljivim koeficijentom statističke linearizacije (5.52). Dodatni razlog za primenu promenljivog koeficijenta statističke linearizacije jeste taj što se procenat kontaminacije podataka autlajerima koji se koristi za određivanje fiksnog koeficijenta statističke linearizacije u praksi ne zna unapred. Na slici 6.11 je prikazan predloženi sistem za praćenje iz sekcije 4.3 sa korakom robusne estimacije stanja objekta.



Slika 6.11 Blok dijagram predloženog sistema za praćenje pokretnih objekata u SWIR video sekvenci sa robusnom estimacijom stanja objekta.

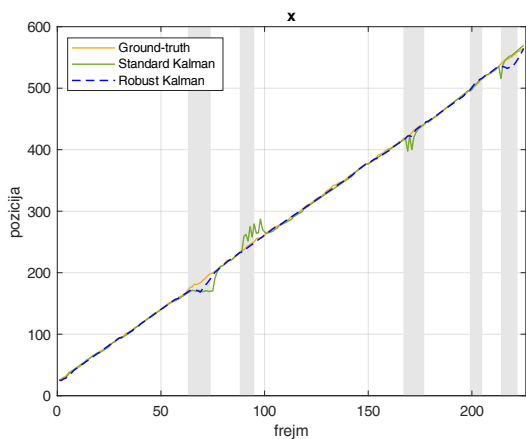


a)

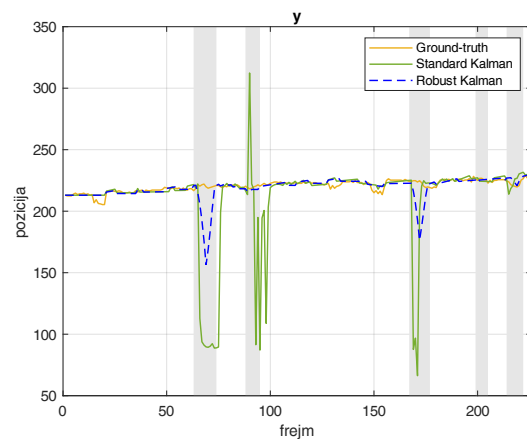


b)

**Slika 6.12** Poređenje stvarne pozicije i pozicije objekta estimirane standardnim i M-robustnim Kalmanovim filtrom: a) horizontalni, b) vertikalni pravac.

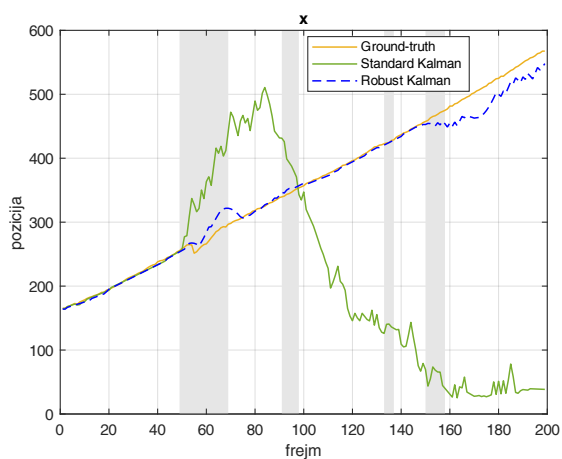


a)

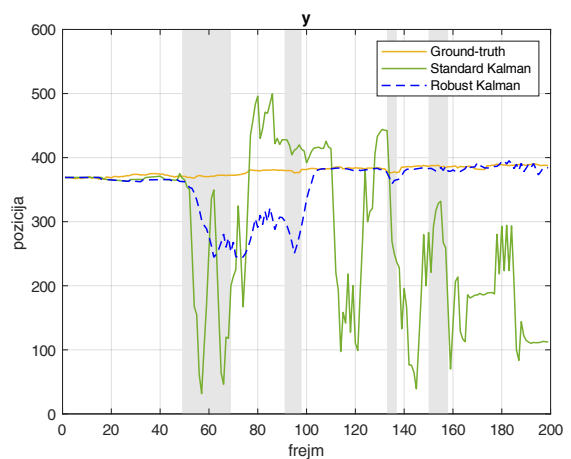


b)

**Slika 6.13** Poređenje stvarne pozicije i pozicije objekta estimirane standardnim i M-robustnim Kalmanovim filtrom u scenariju sa stacionarnim okluzijama: a) horizontalni, b) vertikalni pravac.



a)



b)

**Slika 6.14** Poređenje stvarne pozicije i pozicije objekta estimirane standardnim i M-robustnim Kalmanovim filtrom u scenariju sa pokretnom okluzijom i više stacionarnih okluzija: a) horizontalni, b) vertikalni pravac.

Komparativna analiza performansi M-robusnog Kalmanovog filtra i standardnog Kalmanovog filtra na istim sekvencama na kojima su prikazane performanse standardnog Kalmanovog filtra u sekciji 6.1 prikazane su na slikama 6.12 – 6.14, gde osenčeni delovi označavaju period tokom kog je objekat pod okluzijom. Na slici 6.12 u scenariju bez okluzija, vidi se da je praćenje podjednako uspešno primenom standardnog i robusnog Kalmanovog filtra. Međutim, u scenariju sa okluzijama na slikama 6.13 i 6.14 robusni Kalmanov filter pokazuje značajno bolje performanse u praćenju. Iako u scenariju iz sekvence 5 sistem sa standardnim Kalmanovim filtrom uspeva da isprati objekat do kraja video sekvence, primenom robusnog Kalmanovog filtra se u praćenju ne javljaju velika i nagla odstupanja od stvarne pozicije objekta. U scenariju na slici 6.14, primenom standardnog Kalmanovog filtra praćenje prestaje već nakon pojave prve okluzije, dok robusni metod uz određena odstupanja od stvarne pozicije uspeva da isprati objekat. Rezultati praćenja na frejmovima ovih sekvenci sistemom sa standardnim i robusnim Kalmanovim filtrom prikazani su na slikama 6.15 i 6.16.



**Slika 6.15** Poređenje preciznosti predloženog sistema za praćenje standardnim i M-robusnim Kalmanovim filtrom – sekvenca 5.



**Slika 6.16** Poređenje preciznosti predloženog sistema za praćenje standardnim i M-robusnim Kalmanovim filtrom – sekvenca 6.

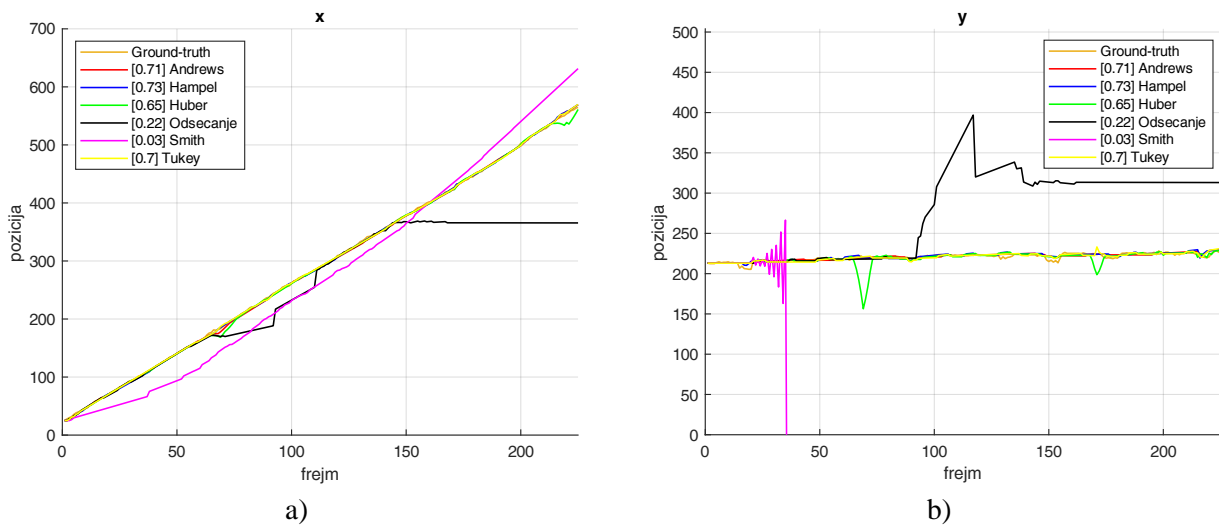
### 6.2.3 Analiza funkcija uticaja

U prethodnoj sekciji su analizirane performanse predloženog sistema za praćenje pokretnih objekata korelacionim filtrom i robusnim Kalmanovim filtrom koji koristi Huberovu funkciju uticaja (5.38). Međutim, primenom Huberove funkcije uticaja, autlajeri različitih intenziteta se tretiraju i potiskuju na isti način. U ovoj sekciji će biti analizirane performanse sistema gde se u koraku robusne estimacije Kalmanovog filtra koriste funkcije uticaja koje imaju mogućnost potpunog potiskivanja autlajera. Robusne funkcije uticaja koje se analiziraju su: nelinearnost odsecanja, *Hampel*-ova, *Andrews*-ova, *Tukey*-jeva i *Smith*-ova funkcija. Definicija, izgled i parametri navedenih funkcija uticaja dati su u sekciji 5.4.3. Komparativna analiza funkcija uticaja izvršena je posmatranjem

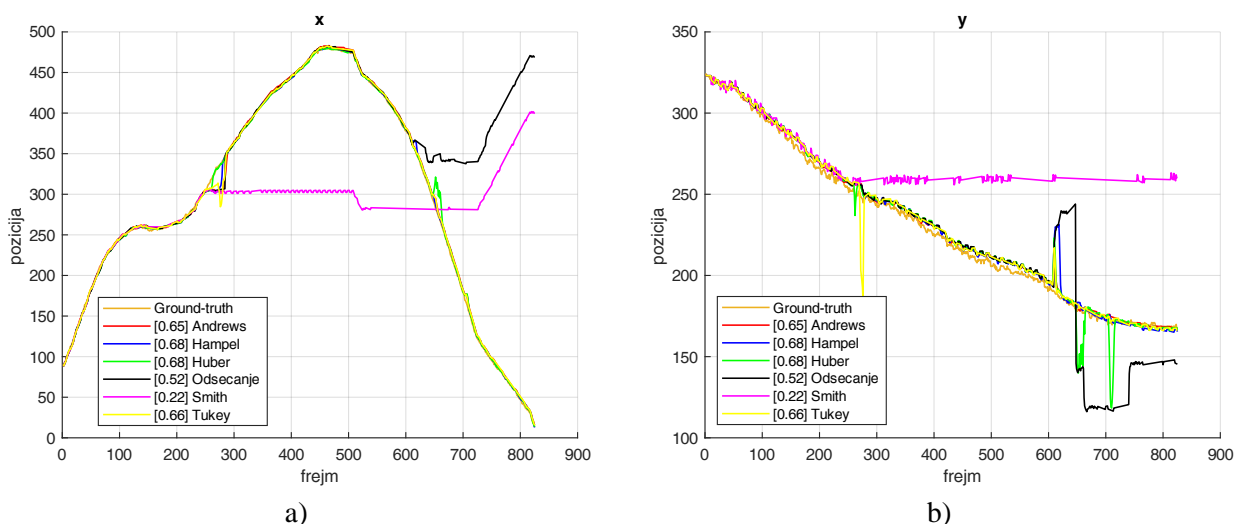


estimirane pozicije objekta u odnosu na stvarnu poziciju u karakterističnim scenarijima praćenja. Rangiranje je vršeno prema *Success rate* metrici.

Na slici 6.17 je prikazan scenario sa statičnim i pokretnim okluzijama, odakle se vidi da sistem primenom *Smith*-ove funkcije uticaja gubi objekat već pri pojavi prve potpune okluzije, dok nešto kasnije se isto dešava i primenom nelinearnosti odsecanja. Primenom ostalih funkcija uticaja praćenje je uspešno do kraja sekvence, pri čemu *Hampel*-ova, *Andrews*-ova i *Tukey*-jeva funkcija pokazuju nešto bolje performanse u odnosu na Huberovu funkciju. Veoma slični rezultati, prikazani na slici 6.18, su dobijeni i za scenario u sekvenci 1, koji pored okluzija uključuje i promenu orijentacije objekta. U ovoj situaciji rezultati dobijeni primenom Huberove funkcije prema *Success rate* metrici su gotovo identični rezultatima dobijenim *Hampel*-ovom, *Andrews*-ovom i *Tukey*-jevom funkcijom.



**Slika 6.17** Komparativna analiza praćenja pokretnog objekta sa različitim funkcijama uticaja u robusnom Kalmanovom filtru u scenariju iz sekvence 5: a) horizontalni, b) vertikalni pravac.

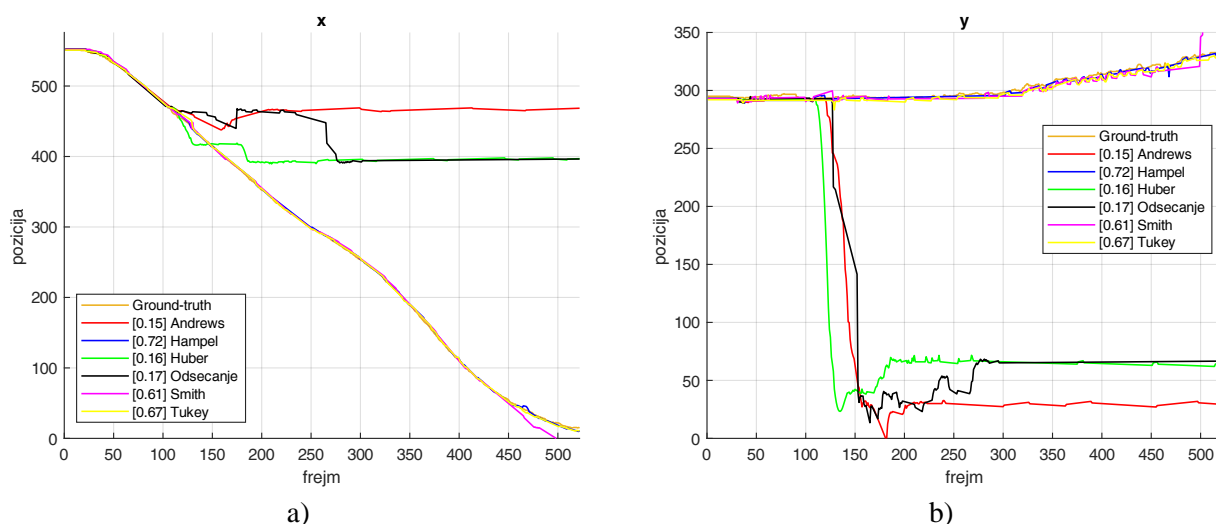


**Slika 6.18** Komparativna analiza praćenja pokretnog objekta sa različitim funkcijama uticaja u robusnom Kalmanovom filtru u scenariju iz sekvence 1: a) horizontalni, b) vertikalni pravac.

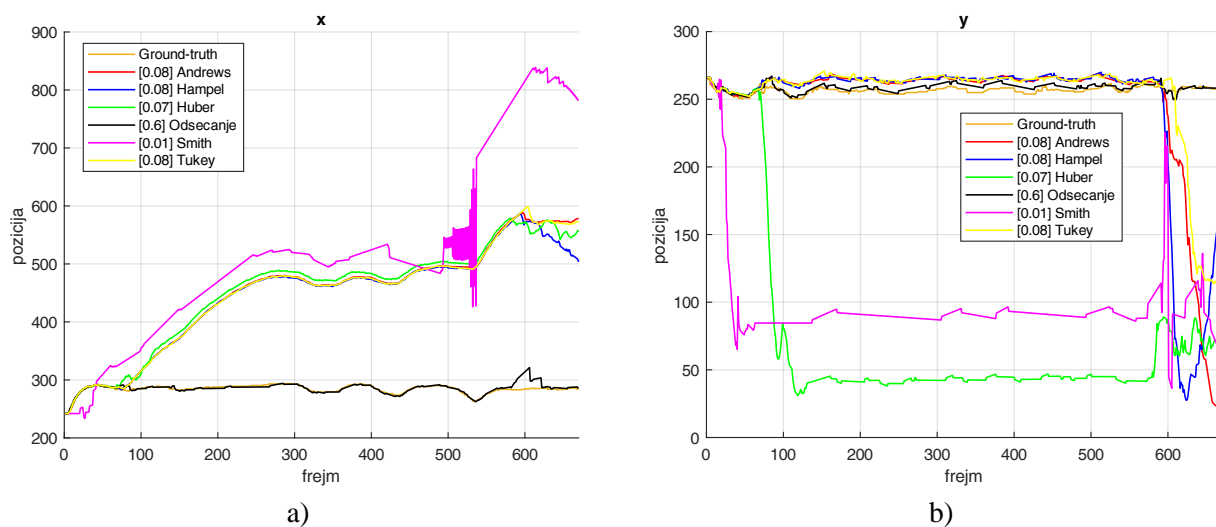
Rezultati analize funkcija uticaja u scenariju iz sekvence 3 koji obuhvata jednu dugotrajnu potpunu okluziju i promenu dinamike kretanja objekta prikazani su na slici 6.19. Primenom nelinearnosti odsecanja, *Andrews*-ove i Huberove funkcije praćenje objekta je prekinuto pojavom



dugotrajne okluzije. Za razliku od prethodna dva slučaja, sada sistem primenom *Smith*-ove funkcije uticaja uspešno prevazilazi okluziju i održava praćenje do kraja video sekvence. Praćenje je uspešno i primenom *Tukey*-jeve i *Hampel*-ove funkcije, koja pokazuje najbolje rezultate praćenja. U sličnom scenariju praćenja iz sekvence 13 koji je snimljen kamerom u pokretu, praćenje je bilo uspešno jedino primenom nelinearnosti odsecanja. Primenom ostalih funkcija uticaja, praćenje je prekinuto pojavom dugotrajne okluzije, što se može videti na slici 6.20.

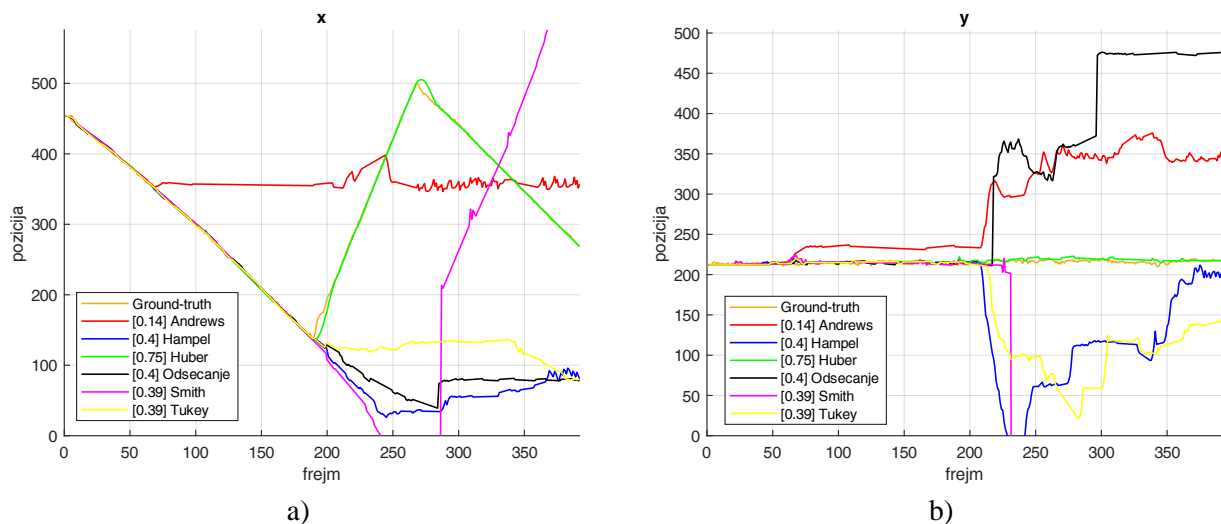


**Slika 6.19** Komparativna analiza praćenja pokretnog objekta sa različitim funkcijama uticaja u robusnom Kalmanovom filtru u scenariju iz sekvence 3: a) horizontalni, b) vertikalni pravac.



**Slika 6.20** Komparativna analiza praćenja pokretnog objekta sa različitim funkcijama uticaja u robusnom Kalmanovom filtru u scenariju iz sekvence 13: a) horizontalni, b) vertikalni pravac.

Na slici 6.21 su prikazani rezultati praćenja u scenariju sa kratkotrajnim delimičnim i potpunim okluzijama bez promene smera kretanja uz naglo pomeranje kamere u jednom trenutku praćenja. Iz prikazanih rezultata se vidi da se jedino Huberova funkcija pokazala uspešnom.



**Slika 6.21** Komparativna analiza praćenja pokretnog objekta sa različitim funkcijama uticaja u robusnom Kalmanovom filtru u scenariju iz sekvence 11: a) horizontalni, b) vertikalni pravac.

Na osnovu analize funkcija uticaja u različitim scenarijima praćenja, može se zaključiti da ni jedna od analiziranih funkcija nije primenljiva u svim scenarijima. Generalno primenom *Smith*-ove funkcije je najčešće dolazilo do divergencije sistema. Primenom ostalih analiziranih funkcija uticaja u određenim scenarijima se potpuno potiskivanje autlajera pokazalo veoma korisno, dok je u drugim to dovodilo do prestanka praćenja. Stoga, kako bi se primenom jedne funkcije uticaja postiglo potpuno ili skoro potpuno potiskivanje autlajera, ali i vršilo ažuriranje Kalmanovog filtra merenjima koja su kvalitetna, u zavisnosti od situacije na sceni, u narednoj sekciji se razmatra sistem za praćenje pokretnih objekata primenom robusnog Kalmanovog filtra sa adaptivnom funkcijom uticaja.

### 6.3 Sistem za praćenje korelacionim filtrom i M-robusnim Kalmanovim filtrom sa adaptivnom funkcijom uticaja

Prilikom projektovanja koraka robusne estimacije stanja pozicije objekta, važno je imati na umu da velika robusnost u odnosu na prisustvo autlajera može smanjiti efikasnost estimatora u regularnim uslovima. Stoga je potrebno postići balans između ova dva zahteva. Greške u estimaciji pozicije uzrokovane manevrom ili bržim kretanjem objekta koji se prati ne predstavljaju autlajere i ne smeju se odbacivati. S druge strane, greške uzrokovane kratkotrajnim ili dugotrajnim okluzijama, odnosno one koje se prvenstveno javljaju kada je aktiviran modul za ponovnu identifikaciju pretragom u proširenoj zoni primenom višestrukih prozora, predstavljaju autlajere čiji uticaj treba umanjiti. Kako bi se omogućio određeni stepen tolerancije prema greškama u poziciji koje su uzrokovane manevrisanjem objekta dok se istovremeno vrši potiskivanje autlajera, robusni Kalmanov filter sa Huberovom funkcijom uticaja je optimizovan primenom adaptivnog praga zasićenja  $\Delta$  u Huberovoj funkciji uticaja. Vrednost parametra  $\Delta$  treba da bude što veća u situacijama bez prisustva autlajera u mernim podacima, odnosno kada je raspodela greške merenja Gausova. Međutim, kada merni podaci sadrže autlajere (raspodela grešaka ima otežale repove), vrednost parametra  $\Delta$  treba da bude što je moguće manja. Stoga, zaključak je da vrednost parametra  $\Delta$  treba da odražava prisustvo autlajera u mernim podacima.

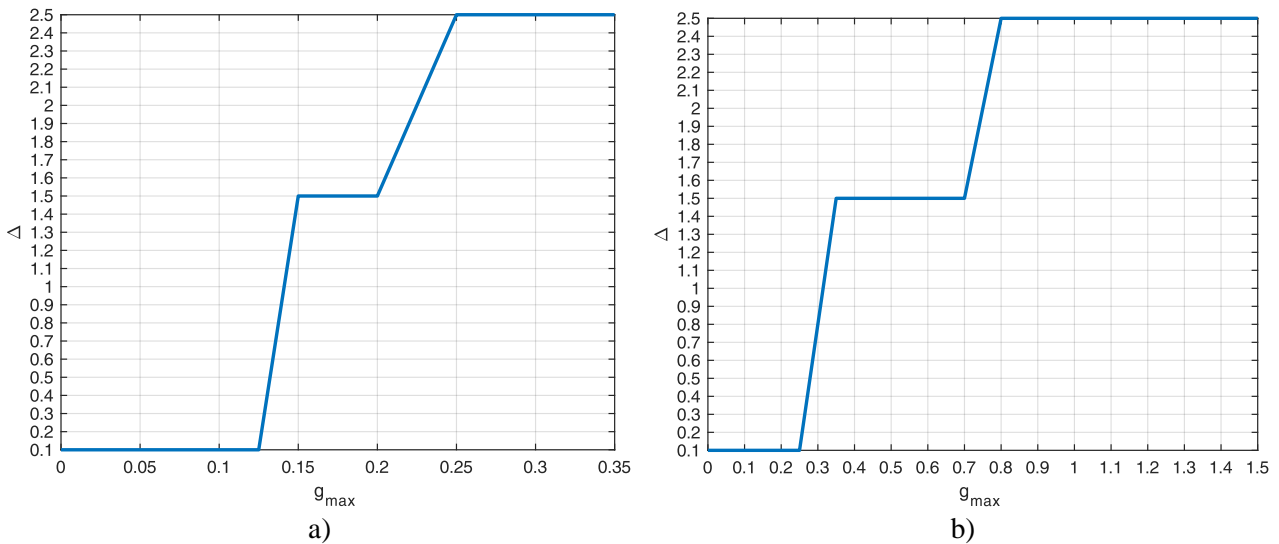
Vrednost parametra  $\Delta$  se određuje u zavisnosti od nivoa kontaminacije podataka, a Huber je predložio vrednosti parametra  $\Delta$  za različite procenite efikasnosti estimatora [40]. Međutim, stepen kontaminacije podataka autlajerima nije unapred poznat jer zavisi od trenutne situacije na sceni. Stepem kontaminacije se može proceniti, a Huber je predstavio korelaciju između stepena kontaminacije i vrednosti parametra  $\Delta$ . U novijoj literaturi [200], težinski faktori izračunati korišćenjem Huberove funkcije uticaja se koriste za procenu nivoa kontaminacije. Parametar  $\Delta$  Huberove funkcije uticaja optimizovan procenjenim nivoom kontaminacije je korišćen u robusnom Kalmanovom filtru za praćenje objekata na termalnoj slici primenom SURF deskriptora. Međutim, kvalitet procene nivoa kontaminacije zavisi od broja uzoraka koji se koriste za procenu [200]. Veći broj uzoraka omogućava precizniju procenu, ali uvodi dodatno kašnjenje.

Kako bi se izbeglo povećanje računске kompleksnosti i uvođenje dodatnog kašnjenja u sistem za praćenje pri optimizaciji parametra  $\Delta$  za ilustraciju situacije na sceni, može se primeniti nov pristup koji koristi informacije iz korelacionog odziva, odnosno vrednost pika izlaznog korelacionog odziva evaluiranog uzorka slike korelacionim filtrom [201]. Veća vrednost pika korelacionog odziva ukazuje na pouzdaniju detekciju objekta. U tom slučaju vrednost parametra  $\Delta$  treba da bude veća, što uvažava veće odstupanje od predikcije pozicije Kalmanovim filtrom i omogućava praćenje manevra i bržeg kretanja objekta. S druge strane, kada je vrednost pika u korelacionom odzivu detektovanog objekta niska, vrednost parametra  $\Delta$  bi trebalo da bude manja. Niža vrednost pika u izlaznom korelacionom odzivu ukazuje na nepouzdanost detekcije, te da dobijena detekcija možda ne predstavlja objekat koji se prati. Dodatno, nepouzdana detekcija koja se nalazi na poziciji koja je znatno udaljena od predikcije pozicije može predstavljati autlajer i posledično dovesti do prekida praćenja. Stoga, na osnovu eksperimentalne analize sprovedene na skupu SWIR video sekvenci, za sistem sa korelacionim filtrom koji je kontekstno regularizovan [169] predlaže se zavisnost parametra  $\Delta$  od vrednosti pika  $g_{max}$  u izlaznom korelacionom odzivu prema relaciji:

$$\Delta = \begin{cases} 0.1 & g_{max} < 0.125 \\ 0.1 + 56 \cdot (g_{max} - 0.125) & 0.125 \leq g_{max} < 0.15 \\ 1.5 & 0.15 \leq g_{max} < 0.2 \\ 1.5 + 20 \cdot (g_{max} - 0.2) & 0.2 \leq g_{max} < 0.25 \\ 2.5 & g_{max} \geq 0.25 \end{cases} \quad (6.9)$$

Grafički prikaz ove zavisnosti dat je na slici 6.22 a). Za primenu sa kernelizovanim korelacionim filtrom [99], ova zavisnost je izražena prema (6.10) i prikazana na slici 6.22 b) [201].

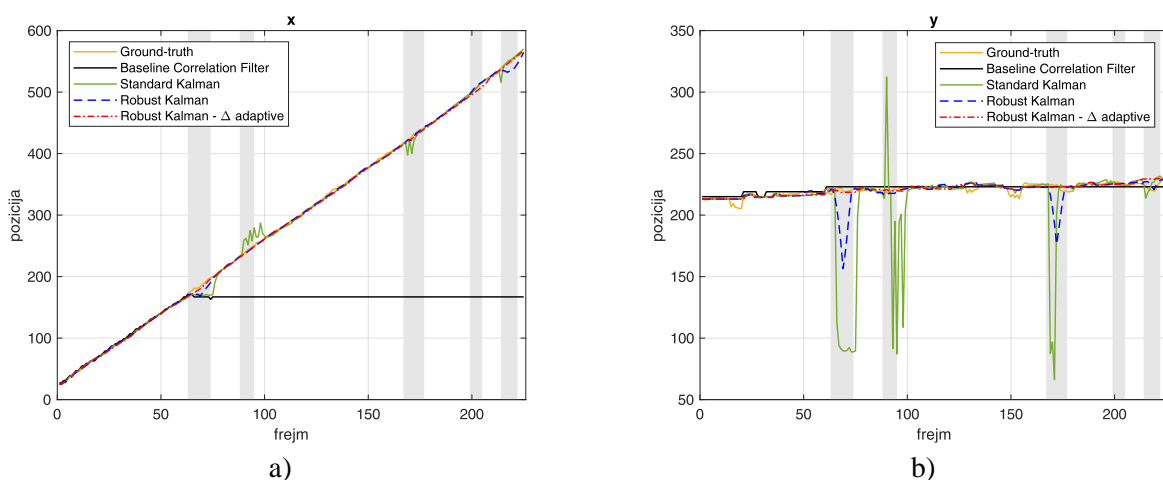
$$\Delta = \begin{cases} 0.1 & g_{max} < 0.25 \\ 0.1 + 14 \cdot (g_{max} - 0.25) & 0.25 \leq g_{max} < 0.35 \\ 1.5 & 0.35 \leq g_{max} < 0.7 \\ 1.5 + 10 \cdot (g_{max} - 0.7) & 0.7 \leq g_{max} < 0.8 \\ 2.5 & g_{max} \geq 0.8 \end{cases} \quad (6.10)$$



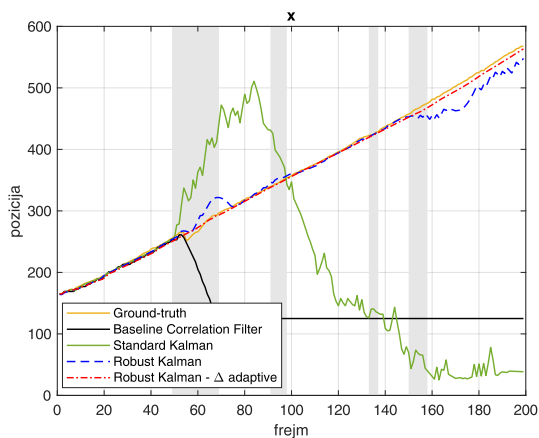
**Slika 6.22** Zavisnost parametra  $\Delta$  u Huberovoj funkciji uticaja od vrednosti pika korelacionog odziva za: a) kontekstno regularizovani korelacioni filtar b) kernelizovani korelacioni filtar.

Uporednom analizom na istim sekvencama sa prisustvom okluzija kao u sekcijama 6.1 i 6.2.2, na slikama 6.23 i 6.24 se mogu videti rezultati praćenja stvarne i estimirane pozicije objekta bazičnim korelacionim filtrom, standardnim Kalmanovim filtrom, M-robustnim Kalmanovim filtrom sa fiksnim parametrom  $\Delta$  u Huberovoj funkciji uticaja i M-robustnim Kalmanovim filtrom sa adaptivnom Huberovom funkcijom uticaja.

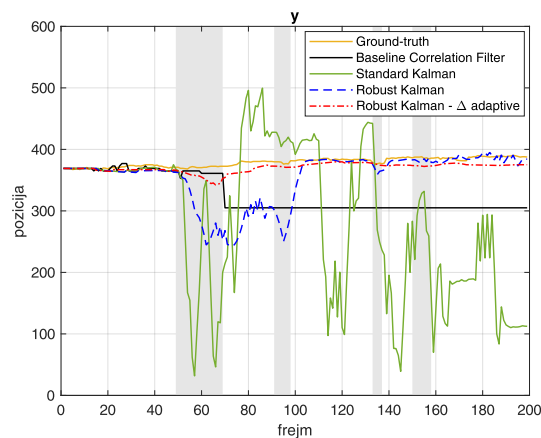
Posmatranjem rada dva robustna Kalmanova filtra, sa fiksnim i adaptivnim parametrom  $\Delta$  u Huberovoj funkciji uticaja, sa slika 6.23 i 6.24 se može videti da oba uspešno prevazilaze sve okluzije i bez gubitka objekta praćenje se održava do kraja sekvence. Međutim, ako se pogleda odstupanje estimirane od stvarne pozicije objekta, vidi se da robustni Kalmanov filtar sa adaptivnom funkcijom uticaja ima manju grešku i lakše savladava okluzije, što je posebno izraženo tokom perioda dugotrajne okluzije. Kod sistema oslonjenog samo na bazični korelacioni filtar, prestanak praćenja nastupa već prilikom pojave prve okluzije, dok primena standardnog Kalmanovog filtra održava praćenje uz velika odstupanja od stvarne pozicije tokom perioda okluzije na slici 6.23, odnosno dolazi do prestanka praćenja prilikom pojave prve okluzije u sekvenci na slici 6.24.



**Slika 6.23** Poređenje stvarne pozicije i pozicije objekta estimirane bazičnim korelacionim filtrom, standardnim, M-robustnim i M-robustnim Kalmanovim filtrom sa adaptivnim parametrom  $\Delta$  u scenariju sa okluzijama: a) horizontalni, b) vertikalni pravac.

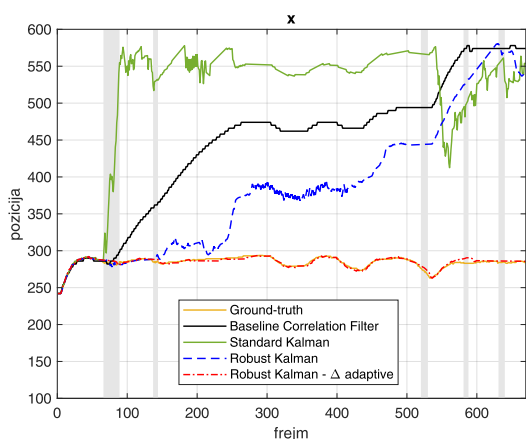


a)

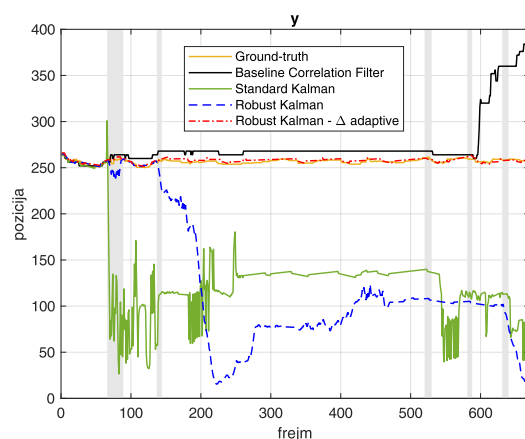


b)

**Slika 6.24** Poređenje stvarne pozicije i pozicije objekta estimirane bazičnim korelacionim filtrom, standardnim, M-robustnim i M-robustnim Kalmanovim filtrom sa adaptivnim parametrom  $\Delta$  u scenariju sa dugotrajnom pokretnom okluzijom i više statičnih okluzija: a) horizontalni, b) vertikalni pravac.



a)

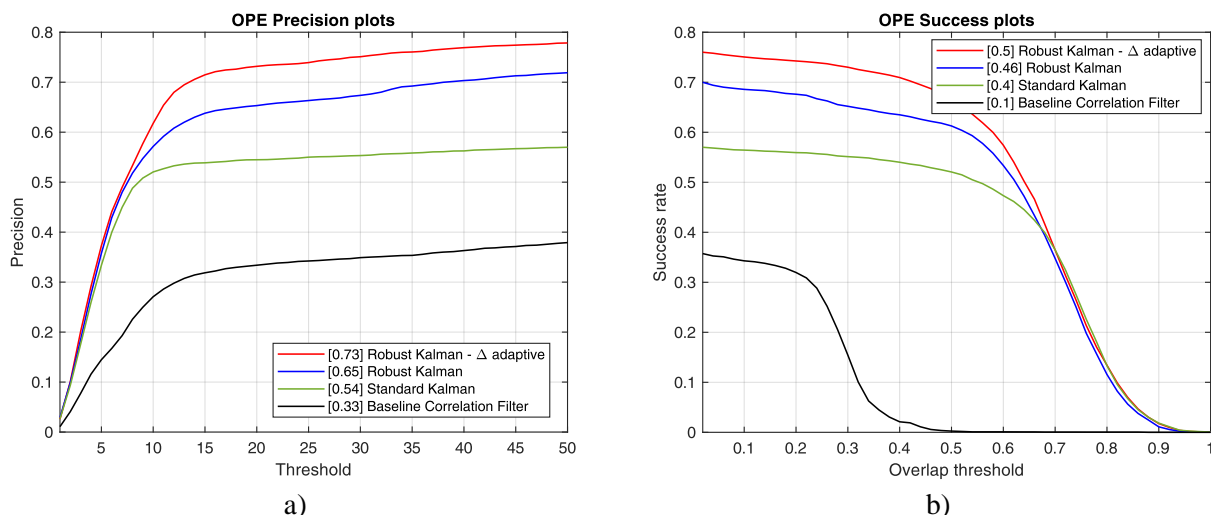


b)

**Slika 6.25** Poređenje stvarne pozicije i pozicije objekta estimirane bazičnim korelacionim filtrom, standardnim, M-robustnim i M-robustnim Kalmanovim filtrom sa adaptivnim parametrom  $\Delta$  u scenariju sa pokretnom okluzijom i više statičnih okluzija u sekvenci 13: a) horizontalni, b) vertikalni pravac.

Analiza rezultata praćenja u sekvenci 13 prikazana na slici 6.25 pokazuje da prilikom pojave druge okluzije u video sekvenci robusni Kalmanov filter sa fiksnom funkcijom uticaja ne uspeva da isprati objekat i dolazi do prestanka praćenja. S druge strane, kada se objekat pojavi nakon okluzije, pouzdanost detekcije se povećava, što u robusnom Kalmanovom filteru sa adaptivnom funkcijom uticaja takođe povećava vrednost parametra  $\Delta$ . Ovo omogućava bržu adaptaciju predloženog robusnog sistema za praćenje i ponovnu identifikaciju praćenog objekta. Rezultat je da estimator brže konvergira i okluzija se efikasnije prevazilazi. U svim analiziranim sekvencama, performanse svih algoritama su skoro identične do pojave prve okluzije u video sekvenci jer je predloženi sistem za praćenje dizajniran da se oslanja na bazični korelacioni filter u regularnim uslovima.

Posmatranjem grafika preciznosti i uspešnosti na slici 6.26 dobijenih za performanse praćenja na celokupnom skupu SWIR video sekvenci, zaključak je da su primenom robusnog Kalmanovog filtra sa adaptivnom funkcijom uticaja značajno poboljšani rezultati praćenja u poređenju sa rezultatima dobijenim za standardni i robusni Kalmanov filter koji koristi Huberovu funkciju uticaja sa fiksnim parametrom  $\Delta = 1.5$ .



**Slika 6.26** Poređenje: a) preciznosti i b) uspešnosti - na celokupnom skupu SWIR video sekvenci bazičnog korelacionog filtra (sa kontekstnom regularizacijom) i predloženog sistema sa standardnim, M-robustnim i M-robustnim Kalmanovim filtrom sa adaptivnim parametrom  $\Delta$  u Huberovoj funkciji uticaja.

Iz perspektive robusnosti, pokazalo se da M-robustni Kalmanov filter sa adaptivnim parametrom  $\Delta$  u Huberovoj funkciji uticaja, potiskuje autlajere u mernim podacima u različitim scenarijima sa tačkom preloma od 25 %. Tačka preloma se definiše kao najveći stepen kontaminacije za koji robusni estimator daje maksimalnu prihvatljivu grešku procene. U predloženom sistemu za praćenje, autlajeri se mogu javiti tokom perioda kada je aktiviran modul za pretragu u proširenoj zoni. U slučaju da period tokom kog je objekat pod okluzijom traje predugo ili se u scenariju praćenja javlja mnogo uzastopnih okluzija, usled čega je pretraga u proširenoj zoni aktivna tokom dužeg perioda, procenat autlajera u mernoj sekvenci se može povećati i dostići tačku preloma. Ovo može dovesti do divergencije Kalmanovog filtra, gubitka objekta i prestanka praćenja.

## 6.4 Sistem za praćenje korelacionim i adaptivnim M-robustnim Kalmanovim filtrom

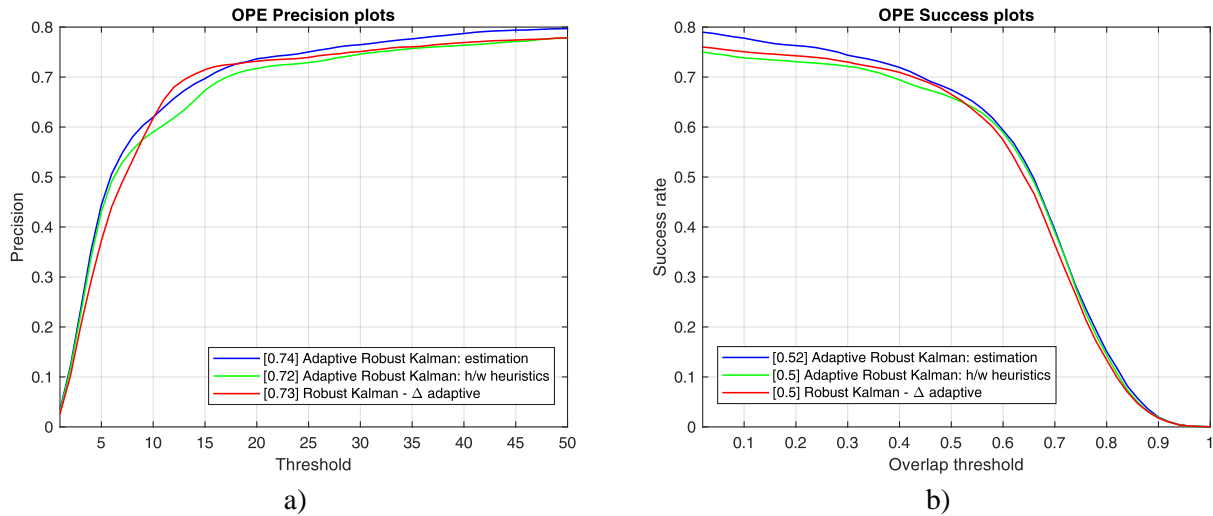
Prilikom projektovanja Kalmanovog filtra potrebno je unapred znati statistike šuma stanja i šuma merenja kako bi se na pravilan način inicijalizovale kovarijacione matrice šuma stanja  $Q$  (5.9) i šuma merenja  $R$  (5.9). Međutim, precizna procena ovih matrica u praksi nije jednostavna. Štaviše, statistike šuma se tokom vremena mogu menjati, pa precizna procena ovih matrica prilikom inicijalizacije ne garantuje da će estimaciona procedura biti optimalna tokom praćenja u celoj video sekvenci. Celokupna analiza rada Kalmanovog filtra u ovom poglavlju izvršena je korišćenjem kovarijacionih matrica šuma stanja i merenja inicijalizovanih prema (6.3). Primenjeni metod inicijalizacije predstavlja dobro rešenje jer ne koristi fiksne vrednosti, već omogućava određeni stepen adaptivnosti korišćenjem vrednosti inicijalnog vektora stanja. Na taj način se sistem prilagođava praćenju konkretnog objekta za koji se vrši inicijalizacija. Međutim, ni inicijalizacija na osnovu ove heuristike ne garantuje procenu kovarijacionih matrica koje su optimalne tokom praćenja objekta u celoj video sekvenci. Stoga će u ovoj sekciji biti analizirano adaptivno rešenje koje omogućava da se kovarijacione matrice šuma određuju u svakom koraku rada estimatora.

Prateći heuristiku inicijalizacije kovarijacionih matrica šuma stanja i merenja iz (6.3), na sličan način se u svakom koraku može vršiti procena ovih matrica [70], [197]. Koristeći estimaciju širine okružujućeg okvira objekta  $w_{t-1}$  iz prethodnog frejma i estimaciju visine iz prethodnog frejma  $h_{t-1}$ , kovarijacione matrice šuma stanja  $Q_t$  i šuma merenja  $R_t$  se za praćenje na trenutnom frejmu video sekvence mogu proceniti prema:

$$Q_t = \text{diag} \begin{pmatrix} (\sigma_p w_{t-1})^2 \\ (\sigma_p h_{t-1})^2 \\ (\sigma_p w_{t-1})^2 \\ (\sigma_p h_{t-1})^2 \\ (\sigma_v w_{t-1})^2 \\ (\sigma_v h_{t-1})^2 \\ (\sigma_v w_{t-1})^2 \\ (\sigma_v h_{t-1})^2 \end{pmatrix}, \quad R_t = \text{diag} \begin{pmatrix} (\sigma_p w_{t-1})^2 \\ (\sigma_p h_{t-1})^2 \\ (\sigma_p w_{t-1})^2 \\ (\sigma_p h_{t-1})^2 \end{pmatrix} \quad (6.11)$$

gde  $\text{diag}$  označava dijagonalnu matricu, a koeficijenti  $\sigma_p$  i  $\sigma_v$  imaju iste vrednosti kao i oni korišćeni prilikom inicijalizacije kovarijacionih matrica šuma u (6.3).

Drugi način adaptacije kovarijacionih matrica šuma stanja i merenja podrazumeva dizajn rešenja koje na osnovu M-robusnih estimacija stanja objekta prilagođava kovarijacionu matricu šuma stanja  $Q_t$  prema (5.104) i šuma merenja  $R_t$  prema (5.95).



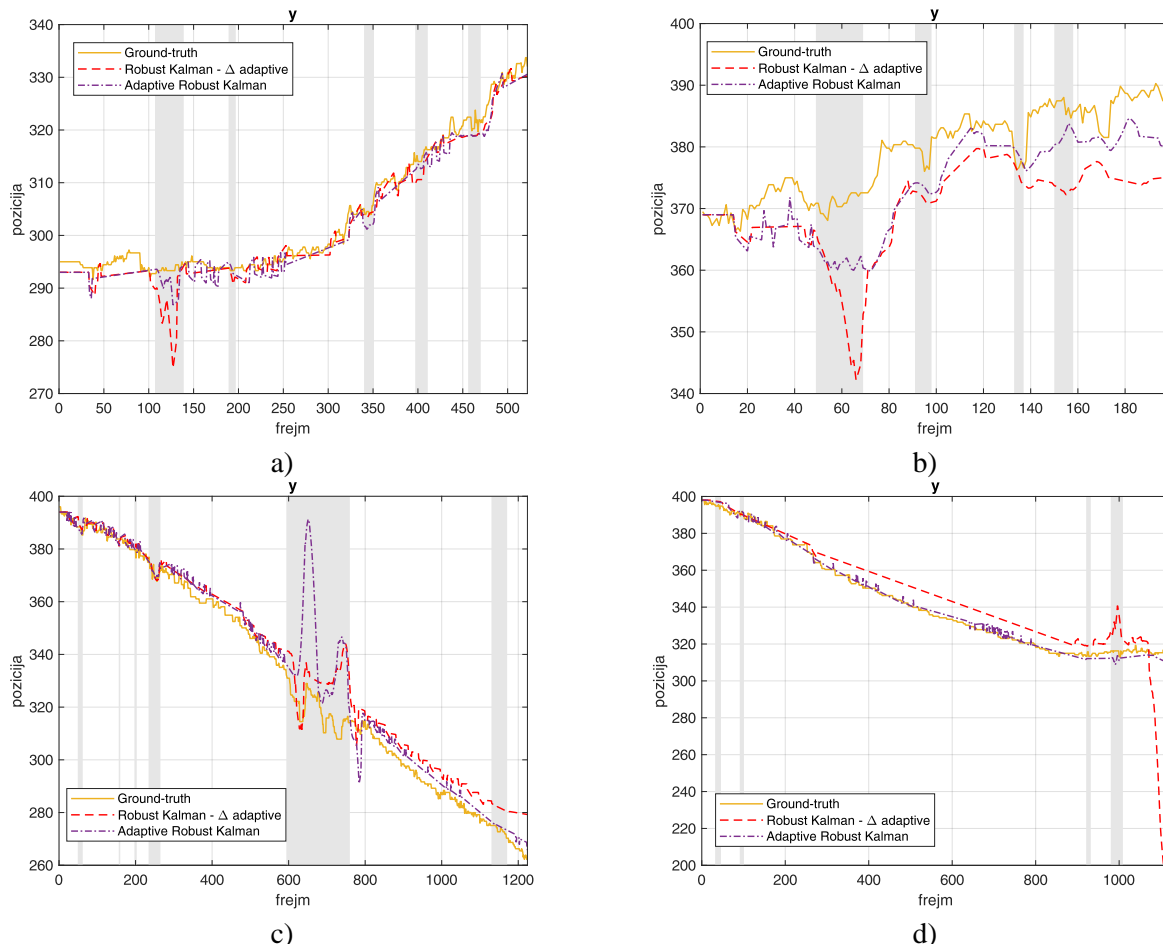
**Slika 6.27** Poređenje: a) preciznosti i b) uspešnosti - na celokupnom skupu SWIR video sekvenci,

M - robusnog Kalmanovog filtra sa adaptivnim parametrom  $\Delta$  i adaptivnog M-robusnog Kalmanovog filtra na osnovu heuristike (6.11) i na osnovu estimacije kovarijacione matrice šuma merenja (5.95).

Eksperimentalna analiza na celokupnom skupu SWIR video sekvenci pokazala je da adaptacija kovarijacione matrice šuma stanja ni prema (6.11) ni prema (5.104) ne doprinosi poboljšanju performansi praćenja sistema. Stoga, za kovarijacionu matricu šuma stanja treba zadržati vrednosti koje su dobijene prilikom inicijalizacije primenom (6.3). S druge strane, pokazuje se da adaptacija kovarijacione matrice šuma merenja doprinosi poboljšanju performansi praćenja. Na slici 6.27 je data komparativna analiza performansi praćenja sistema sa M-robusnim Kalmanovim filtrom sa fiksnim kovarijacionim matricama šuma i adaptivnim parametrom  $\Delta$  u funkciji uticaja, i sistema



sa M-robusnim Kalmanovim filtrom sa adaptivnom kovarijacionom matricom šuma merenja prema (6.11) i (5.95) i adaptivnim parametrom  $\Delta$  u funkciji uticaja. Na osnovu kriterijuma preciznosti (slika 6.27 a)) i uspešnosti (slika 6.27 b)), može se videti da sistem koji koristi adaptaciju kovarijacione matrice šuma merenja na osnovu M-robusnih estimacija stanja objekta prema (5.95) postiže najbolje performanse praćenja. Stoga, za dalje analize biće korišćen adaptivni robusni Kalmanov filter sa adaptivnim parametrom  $\Delta$  u Huberovoj funkciji uticaja koji vrši adaptaciju kovarijacione matrice šuma merenja prema (5.95).



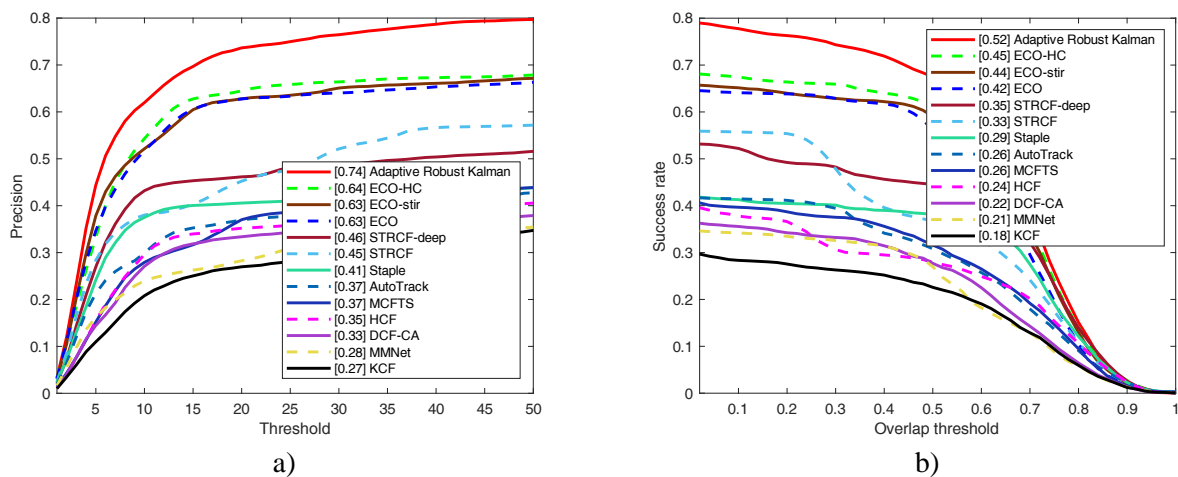
**Slika 6.28** Poređenje stvarne pozicije i pozicije objekta estimirane M-robusnim Kalmanovim filtrom sa adaptivnim parametrom  $\Delta$  i estimirane adaptivnim M-robusnim Kalmanovim filtrom sa adaptivnim parametrom  $\Delta$ : a) sekvenca 3, b) sekvenca 6, c) sekvenca 7, d) sekvenca 9.

Na slici 6.28 su analizirane performanse sistema u scenarijima sa okluzijama u sekvencama 3, 6, 7 i 9 u vertikalnom pravcu. Osenčeni delovi na slikama označavaju periode pod okluzijama. Na slikama 6.28 a) i b) se može videti da je odstupanje estimirane od stvarne pozicije objekta tokom perioda dugotrajne okluzije značajno manje primenom adaptivnog robusnog Kalmanovog filtra. U sekvenci 7 na slici 6.28 c), adaptivni robusni Kalmanov filter iako tokom perioda dugotrajne okluzije ima veću grešku u odnosu na robusni filter sa adaptivnim parametrom  $\Delta$ , on prevazilazi i poslednju okluziju u video sekvenci, što nije slučaj sa robusnim Kalmanovim filtrom sa fiksnom kovarijacionom matricom šuma merenja. Na slici 6.28 d) u sekvenci 9 adaptivni robusni Kalmanov filter preciznije estimira poziciju objekta i uspeva da isprati objekat do kraja video sekvence, dok robusni sa fiksnom kovarijacionom matricom šuma merenja nakon manevra objekta ima konstantnu grešku i gubi objekat nakon poslednje okluzije.



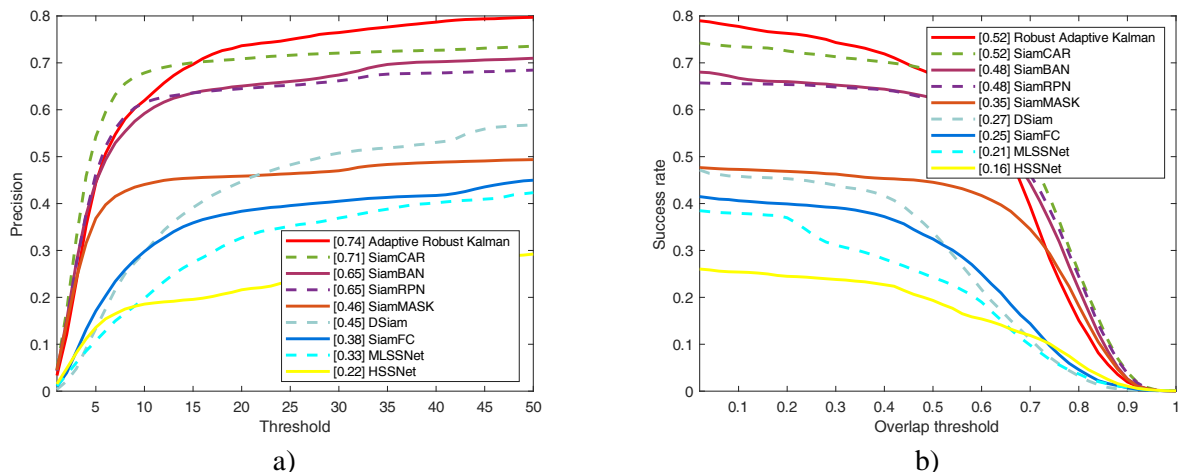
## 6.4.1 Komparativna analiza predloženog sistema i savremenih algoritama za praćenje pokretnih objekata u sekvenci SWIR slika

Kako bi se sagledale performanse predloženog sistema za praćenje objekata u SWIR video sekvenci u odnosu na savremene algoritme za praćenje, analizirane su dve klase algoritama za praćenje pokretnih objekata: diskriminativni korelacioni filtri i duboke Sijamske neuralne mreže. Ove dve klase algoritama su prepoznate kao dominantne paradigme u video praćenju u poslednjoj deceniji [6]. Analizirani su tradicionalni diskriminativni korelacioni filtri: Staple [103], KCF [99], STRCF [108], DCF-CA [169] i AutoTrack [112], kao i diskriminativni korelacioni filtri zasnovani na dubokom učenju obučeni za praćenje objekata u sekvenci kolor slika: HCF [102], ECO [105], ECO-HC [105] i STRCF-deep [108], ali i oni obučeni za praćenje objekata u sekvenci termalnih infracrvenih slika: MCFTS [130], ECO-stir [132] i MMNet [133]. Iz klase dubokih Sijamskih neuralnih mreža analizirani su modeli obučeni za praćenje objekata u sekvenci kolor slika: SiamFC [116], DSiam [118], SiamRPN [119], SiamMASK [121], SiamCAR [123] i SiamBAN [122], kao i oni obučeni za praćenje u termalnom infracrvenom opsegu: HSSNet [96] i MLSSNet [97]. Algoritmi su testirani na svim SWIR video sekvencama iz kreiranog skupa predstavljenog u sekciji 2.3. Rezultati su prikazani na slikama 6.29 i 6.30.



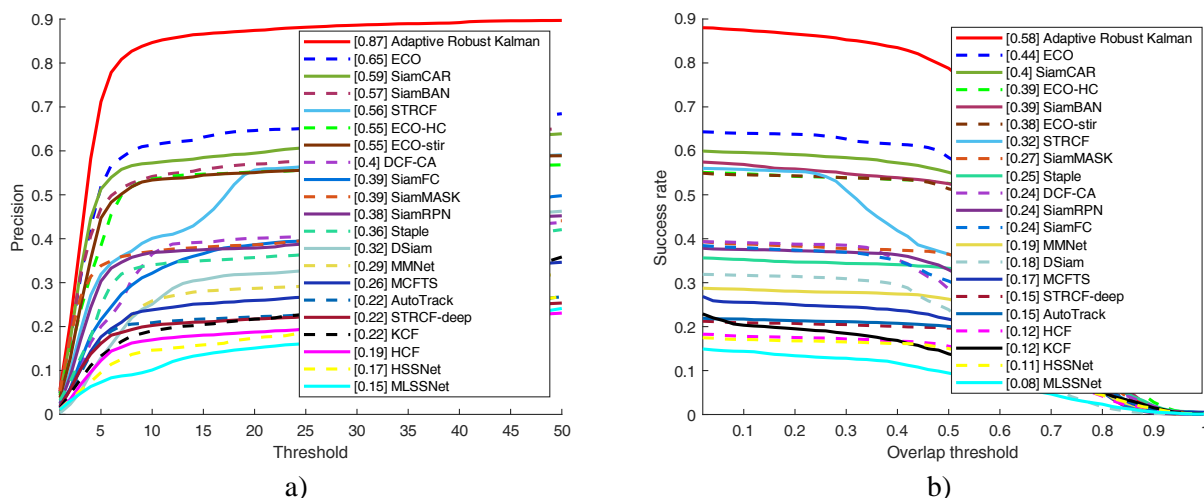
**Slika 6.29** Poređenje: a) preciznosti i b) uspešnosti - na celokupnom skupu SWIR video sekvenci predloženog sistema za praćenje i algoritama baziranih na diskriminativnim korelacionim filtrima.

Sa grafika na slici 6.29 se vidi da predloženi sistem pokazuje značajno bolje performanse u praćenju u odnosu na analizirane algoritme zasnovane na diskriminativnim korelacionim filtrima. U većini slučajeva, a posebno u slučaju dugotrajnih potpunih okluzija, analizirani algoritmi ostaju zaustavljeni na poziciji pojave okluzije. Kada se okluzija javi na početku video sekvence, greška u praćenju se akumulira, što dodatno utiče na ukupne performanse praćenja. Bolje performanse, posmatrano na celokupnom skupu SWIR video sekvenci, predloženi sistem postiže i u odnosu na algoritme zasnovane na Sijamskim neuralnim mrežama, što potvrđuju grafici na slici 6.30. Na graficima preciznosti rangiranje je izvršeno prema preciznosti na pragu od 20 piksela, dok se na graficima uspešnosti rangiranje vrši prema površini ispod kriva uspešnosti.

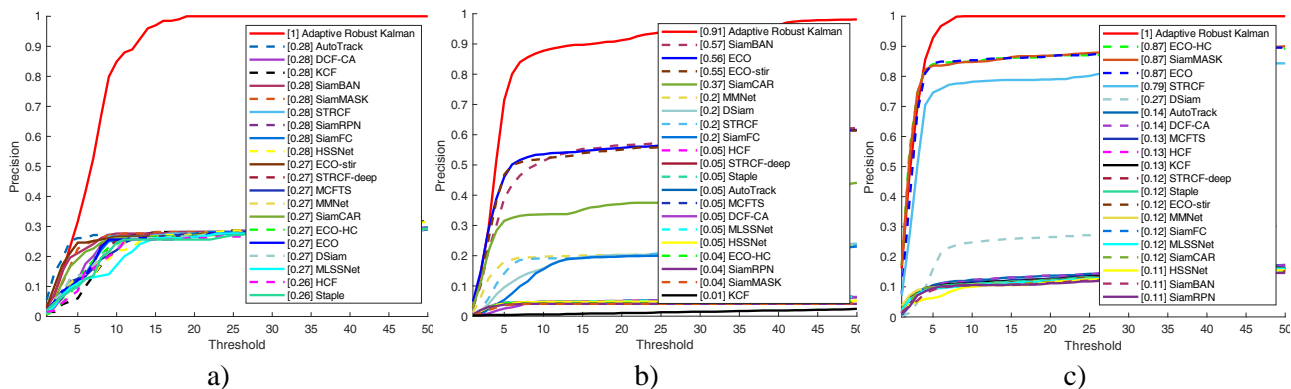


**Slika 6.30** Poređenje: a) preciznosti i b) uspešnosti - na celokupnom skupu SWIR video sekvenci predloženog sistema za praćenje i algoritama baziranih na Sijamskim neuralnim mrežama.

Od posebnog je interesa praćenje objekata malih dimenzija. U klasu sekvenci malih objekata svrstane su sve sekvence na kojima površina objekta ne prelazi 1500 piksela (prema *COCO\_eval* [92] biblioteci za evaluiranje performansi detekcije objekata, u objekte malih dimenzija se svrstavaju objekti do dimenzija  $32 \times 32$  piksela, odnosno do površine od 1024 piksela). Rezultati evaluacije posmatranih algoritama na sekvencama koje su klasifikovane kao sekvence na kojima je objekat koji se prati malih dimenzija su prikazani na slici 6.31. Na graficima preciznosti i uspešnosti na slici 6.31 se vide značajne prednosti predloženog sistema za praćenje u poređenju sa drugim diskriminativnim korelacionim filtrima i Sijamskim neuralnim mrežama. Iz prikazanih rezultata se vidi da analizirani algoritmi koji se prvenstveno oslanjaju na obeležja slike, kada su u pitanju mali objekti u uslovima okluzija, pokazuju slabe diskriminativne sposobnosti za uspešno praćenje. Pored toga, nijedan od algoritama zasnovanih na dubokom učenju nije obučen na SWIR slikama. Oslanjanjem na estimaciju dinamike kretanja objekta adaptivnim robusnim Kalmanovim filtrom sa adaptivnom funkcijom uticaja, zatim primenom metode za detekciju okluzija i pretragom u proširenoj oblasti, moguće je uspešno prevazići okluzije i postići ponovnu identifikaciju objekta, što rezultira uspešnim praćenjem objekta u SWIR video sekvenci.

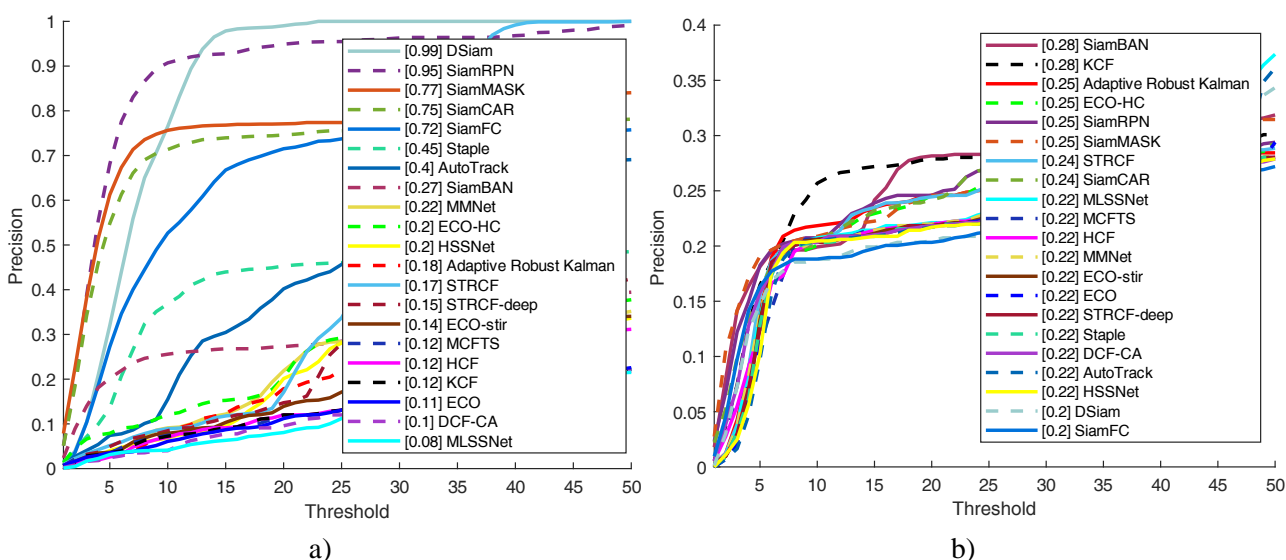


**Slika 6.31** Poređenje performansi predloženog sistema za praćenje i analiziranih algoritama u praćenju malih objekata u SWIR video sekvencama prema a) preciznosti, b) uspešnosti.



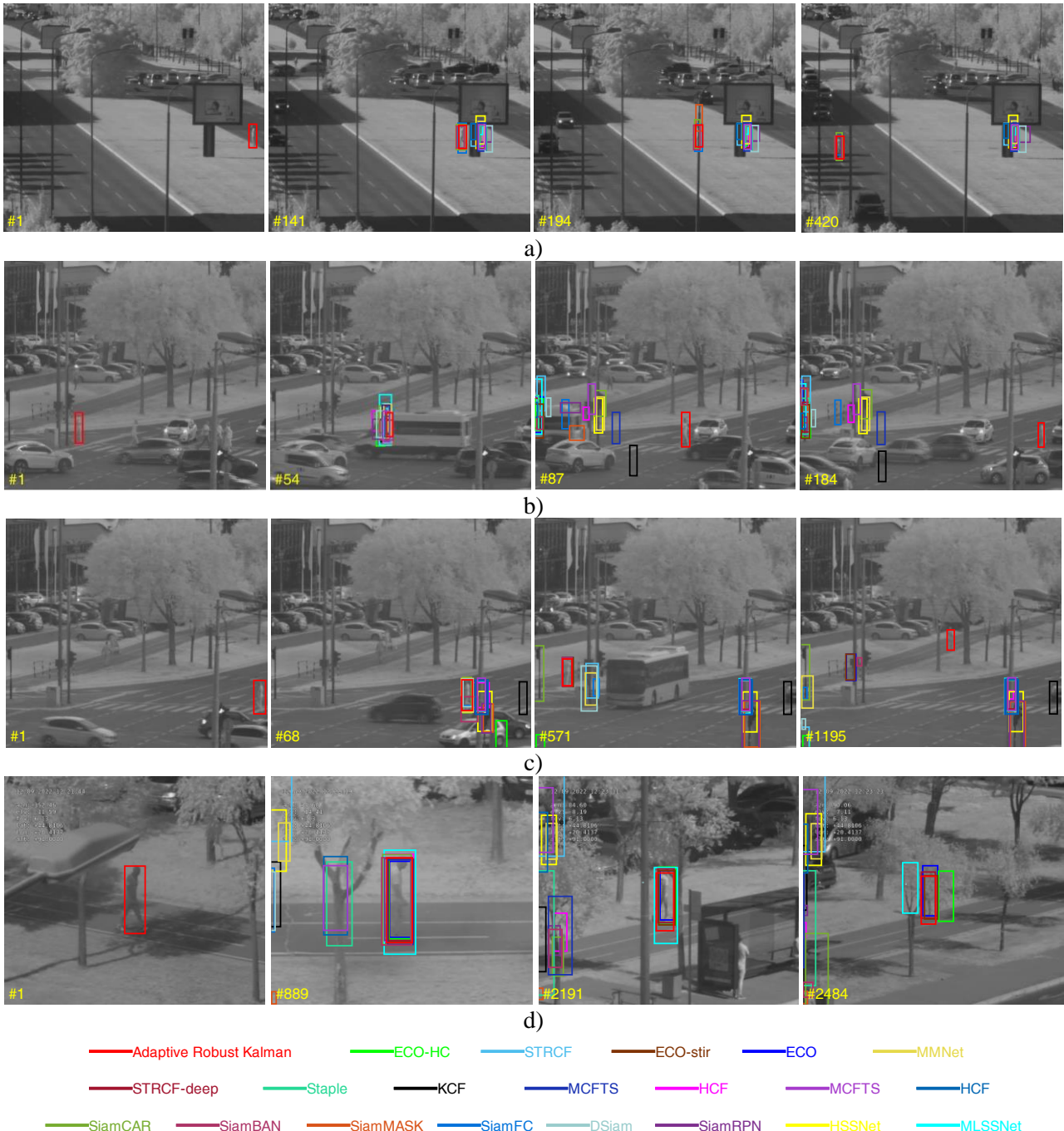
Slika 6.32 Analiza praćenja u izazovnim sekvencama sa prisustvom okluzija:  
a) sekvenca 6, b) sekvenca 7, c) sekvenca 13.

Na slici 6.32 su prikazane performanse praćenja u izazovnim scenarijima praćenja pokretnog objekta malih dimenzija sa promenom dinamike kretanja u prisustvu dugotrajnih statičnih i pokretnih okluzija. Analizirane su sekvence 6, 7 i 13. U analiziranim sekvencama predloženi sistem pokazuje dominantne performanse u poređenju sa ostalim analiziranim algoritmima. S druge strane, na slici 6.33 su prikazane performanse praćenja u sekvencama u kojima predloženi sistem ne uspeva da isprati objekat. Na slici 6.33 a) su prikazani rezultati praćenja u sekvenci 10. U ovom primeru se vidi da su algoritmi zasnovani na Sijamskim neuralnim mrežama pokazali najbolje rezultate. U analiziranoj sekvenci se prati objekat koji se može svrstati u objekte većih dimenzija. Već na početku sekvence, objekat se brzo nađe u senci drugog objekta što značajno utiče na promenu modela izgleda objekta. Korelacioni filtar na koji se predloženi sistem oslanja ne uspeva u takvoj situaciji da detektuje objekat, a kako se opisana situacija događa već na početku video sekvence, Kalmanov filtar na osnovu malog broja merenja ne uspeva dobro da estimira dinamiku kretanja objekta, što rezultira prestankom praćenja. Kako praćenje prestaje već na početku sekvence, greška u praćenju se do kraja video sekvence akumulira, što utiče na loše performanse u analiziranoj sekvenci. Druga sekvenca u kojoj praćenje predloženim sistemom nije uspešno jeste sekvenca 15. U ovoj sekvenci ni jedan od analiziranih algoritama ne uspeva da isprati objekat. Do prestanka praćenja dolazi prilikom pojave dugotrajne okluzije uz dodatno naglo kretanje kamere dok je objekat pod okluzijom. Rezultati praćenja analiziranih algoritama u posmatranoj sekvenci su prikazani na slici 6.33 b).



Slika 6.33 Analiza praćenja u sekvencama sa prisustvom okluzija: a) sekvenca 10, b) sekvenca 15.

Kako bi se jasno demonstrirale performanse predloženog sistema za praćenje u poređenju sa savremenim algoritmima za praćenje, na slici 6.34 su prikazani okružujući okviri generisani analiziranim algoritmima u različitim sekvencama iz skupa SWIR video sekvenci. U svim prikazanim sekvencama može se videti da okružujući okvir estimiran predloženim sistemom zasnovanom na korelacionom filtru i adaptivnom M-robustnom Kalmanovom filtru sa adaptivnom funkcijom uticaja obezbeđuje kontinuirano praćenje objekata u uslovima prisustva različitih tipova okluzija.



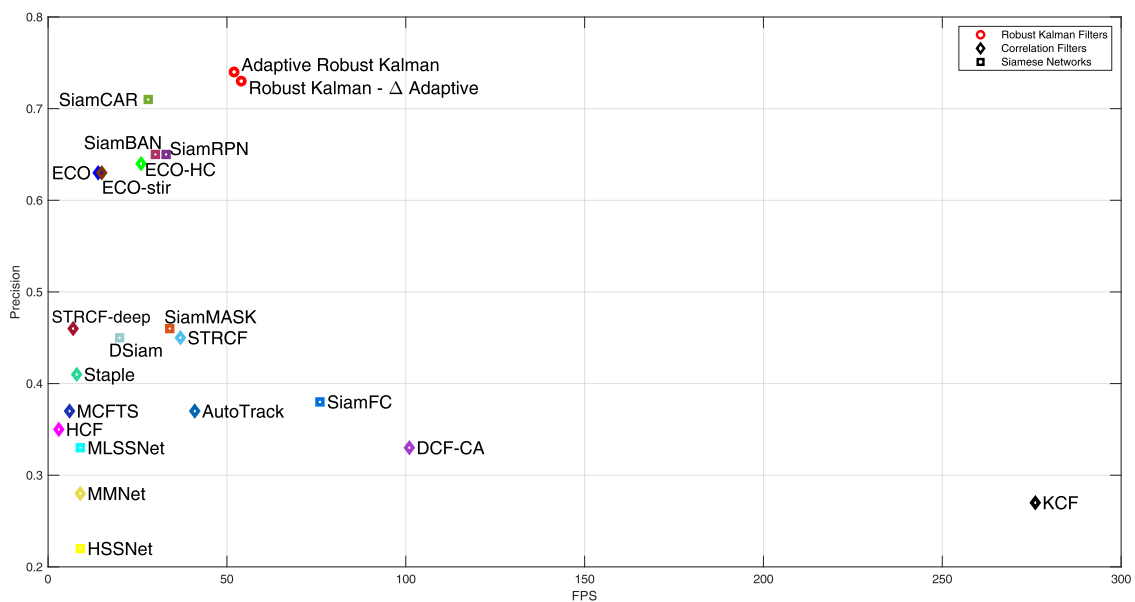
**Slika 6.34** Okružujući okviri estimirani analiziranim algoritmima za praćenje pokretnih objekata: a) sekvenca 3, b) sekvenca 6, c) sekvenca 7, d) sekvenca 16.

Svi analizirani algoritmi za praćenje su evaluirani na PC računaru sa i7 2.6 GHz CPU, 32GB RAM i NVIDIA GeForce RTX 2070 GPU. Poređenje algoritama prema brzini izvršavanja je dato u

tabeli 6.1, kao i na slici 6.35, gde je prikazan odnos brzine izvršavanja i preciznosti. Predloženi sistem za praćenje je implementiran u Matlab softverskom paketu i dostiže brzinu procesiranja frejmova od 54 FPS primenom robusnog Kalmanovog filtra sa adaptivnom funkcijom uticaja, odnosno 52 FPS primenom adaptivnog robusnog Kalmanovog filtra sa adaptivnom funkcijom uticaja. Na osnovu ovih rezultata zaključuje se da sistem ima mogućnost rada u realnom vremenu.

**Tabela 6.1** Poređenje brzina procesiranja frejmova evaluiranih algoritama za praćenje.

Algoritam	Programski jezik	Hardver	Brzina (FPS)
Adative Robust Kalman - $\Delta$ Adaptive	Matlab	CPU	52
Robust Kalman - $\Delta$ Adaptive	Matlab	CPU	54
Staple [103]	Matlab	CPU	8
AutoTrack [112]	Matlab	CPU	41
DCF-CA [169]	Matlab	CPU	101
KCF [99]	Matlab	CPU	276
STRCF [107]	Matlab	GPU	37
HCF [102]	Matlab	CPU	3
ECO [105]	Matlab	GPU	14
ECO-HC [105]	Matlab	GPU	26
STRCF-deep [107]	Matlab	GPU	7
MCFTS [130]	Matlab	GPU	6
ECO-stir [132]	Matlab	GPU	15
MMNet [133]	Matlab	GPU	9
SiamFC [116]	Matlab	GPU	76
DSiam [118]	Matlab	GPU	17
SiamRPN [119]	Python	GPU	29
SiamMASK [121]	Python	GPU	31
SiamCAR [123]	Python	GPU	28
SiamBAN [122]	Python	GPU	28
HSSNet [96]	Matlab	GPU	9
MLSSNet [97]	Matlab	GPU	9



**Slika 6.35** Odnos preciznosti izmerene na celokupnom skupu SWIR video sekvenci i brzine rada predloženog sistema i savremenih algoritama za praćenje.



## 7 Zaključak

Praćenje pokretnih objekata u sekvenci slika predstavlja jednu od najvažnijih komponenti inteligentnih sistema koji kao ulazni senzor koriste kameru. Za sisteme video nadzora posebno je važno da imaju operativne sposobnosti u različitim vremenskim uslovima i uslovima osvetljenja. Zbog toga se, pored kamera koje rade u vidljivom i termalnom infracrvenom spektralnom opsegu, u video sistemima primenjuju i kamere koje rade u kratkotalasnom infracrvenom opsegu, kako bi se obezbedile komplementarne informacije u situacijama i aplikacijama u kojima su termalne kamere i kamere koje rade u vidljivom opsegu neefikasne. Nakon što se obezbedi slika scene sa potrebnim informativnim sadržajem u definisanim uslovima sredine, neophodno je da se dizajnira i sam algoritam za praćenje pokretnih objekata.

Ova disertacija se bavi problemom praćenja jednog pokretnog objekta u sekvenci kratkotalasnih infracrvenih slika. U literaturi je uočen nedostatak pristupa koji se bave ovim problemom, te postoji velika potreba za razvojem i implementacijom rešenja za praćenje jednog pokretnog objekta u sekvenci kratkotalasnih infracrvenih slika. Cilj sistema za praćenje pokretnih objekata jeste postizanje dugotrajnog praćenja u različitim izazovnim situacijama na sceni. U fokusu disertacije su metode za praćenje zasnovane na korelacionim filtrima, prvenstveno zbog svojih dobrih performansi u regularnim uslovima scene, kao i njihove računске efikasnosti zarad mogućnosti rada u realnom vremenu.

Iako korelacioni filtri predstavljaju dobru osnovu za razvoj sistema za praćenje pokretnih objekata u sekvenci SWIR slika, osnovna struktura korelacionih filtara ne omogućava savladavanje izazovnih situacija u video praćenju poput prisustva klatera, kao i promena u izgledu, orijentaciji i dinamici kretanja objekta. Najizazovniji problem jesu okluzije, a posebno potpune dugotrajne okluzije. Struktura korelacionih filtara je takva da pojava okluzija dovodi do zaustavljanja praćenja objekta na poziciji pojave okluzije tako da se model korelacionog filtra dalje prilagođava pozadinskom sadržaju. Stoga je neophodno da se u sistem za praćenje korelacionim filtrom uvede prediktor stanja objekta, kao i adaptivni mehanizam ažuriranja modela korelacionog filtra.

U disertaciji je predložen nov sistem za praćenje pokretnih objekata koji je zasnovan na kombinaciji sledećih podsistema: Kalmanovog filtra kao estimatora stanja objekta; modula za detekciju okluzija i neregularnosti u praćenju objekta; modula za estimaciju veličine objekta i modula za adaptivno ažuriranje kontekstno regularizovanog korelacionog filtra. Predloženi sistem obezbeđuje performanse praćenja koje su značajno bolje od klasičnih sistema za praćenje baziranih samo na korelacionim filtrima. Prevazilaženju potpunih okluzija posebno doprinosi primena Kalmanovog filtra kao prediktora stanja objekta koji se prati, čime je dokazana jedna od polaznih testiranih hipoteza ove disertacije.

Pored osnovnog problema pojave okluzija, objekat tokom perioda pod okluzijom može menjati pravac i smer kretanja, što dodatno otežava ponovnu detekciju objekta nakon okluzije. Stoga, pored primene Kalmanovog filtra za predikciju kretanja objekta, sistem koristi i proširenu zonu pretrage u odnosu na onu definisanu dimenzijama prozora korelacionog filtra, kako bi se poboljšale

moogućnosti ponovne detekcije objekta. Međutim, pretraga u proširenoj zoni, usled nedostupnosti relevantnih informacija o objektu dok je pod okluzijom, može generisati korelacioni odziv čija pozicija pika, koja ukazuje na detektovanu poziciju objekta, značajno odstupa od stvarne pozicije objekta. Analizom rada sistema na kreiranom reprezentativnom skupu SWIR video sekvenci utvrđeno je da je priroda pojave takvih grešaka u detekciji pozicije objekta nepredvidiva (slučajna), pri čemu je intenzitet grešaka veoma veliki, tako da se u statističkom smislu mogu tretirati kao loši podaci ili „autlajeri“.

Kalmanov filter, na čiju se predikciju i estimaciju stanja objekta oslanja projektovani sistem za praćenje, daje optimalne procene stanja linearnog dinamičkog stohastičkog sistema kojim je modelovano kretanje objekta, pri čemu su slučajni šumovi stanja i merenja raspodeljeni po Gausovom zakonu. Međutim, prisustvo autlajera u mernim podacima tokom perioda prevazilaženja okluzija uzrokuje značajna odstupanja od teorijski optimalnih performansi Kalmanovog filtra. Polazeći od ovih praktičnih ograničenja optimalnog Kalmanovog filtra, u disertaciji je predložena nova klasa statistički linearizovanih M-robustifikovanih adaptivnih tehnika Kalmanove filtracije. Predstavljeno je kompletno teorijsko izvođenje estimatora koji ima prediktor-korektor strukturu, pri čemu je korak predikcije projektovan u skladu sa optimalnim Kalmanovim filtrom, dok je rekurzivno ažuriranje merenja (korak korekcije) zasnovano na nelinearnoj robusnoj proceduri dinamičke stohastičke aproksimacije. Ova procedura je generisana minimiziranjem u svakom koraku generalizovanog, vremenski promenljivog Huberovog M-robustnog indeksa performanse. Predloženi proces nelinearne optimizacije primenjuje tehniku statističke linearizacije kako bi se obezbedila suboptimalna robusna verzija matrice Kalmanovog pojačanja, u smislu minimalne totalne varijanse greške estimacije. Pošto određivanje srednje-kvadratnog optimalnog koeficijenta statističke linearizacije pretpostavlja tačno poznavanje statistika šuma merenja, predložene su fiksna i vremenski promenljiva aproksimacija optimalnog koeficijenta.

Teorijska analiza konvergencije predloženih stohastičkih rekurzivnih robusnih algoritama estimacije je zahtevna, kako zbog njihove nelinearne forme, tako i zbog vremenski promenljive i višedimenzionalne dinamike sistema kojim je modelovan objekat koji se prati. Stoga je pre eksperimentalne analize predloženog sistema za video praćenje na realnim podacima, izvršeno testiranje robusnog estimatora na simuliranim podacima praćenja manevrišućeg objekta u jednom pravcu kretanja korišćenjem linearnog senzora pozicije objekta. Rezultati dobijeni eksperimentalnom analizom na realnim sekvencama SWIR slika su potvrdili zaključke koji su proizašli iz kontrolisanog eksperimenta na simulacionim podacima. Pokazano je da fiksni koeficijent statističke linearizacije dovodi do sporijeg smanjenja matrice pojačanja rekurzivnog robusnog estimatora stanja objekta u odnosu na njegovu adaptivnu verziju. Primenom nelinearne transformacije reziduala merenja, koristeći Huberovu M-robustnu funkciju uticaja i robusno izračunavanje matrice pojačanja na osnovu adaptivnog koeficijenta statističke linearizacije, obezbeđuje se dobar kompromis između performansi praćenja i robusnosti estimatora, u smislu efikasnog eliminisanja uticaja autlajera na kvalitet praćenja. Dodatno je razmatrana i primena predložene statistički linearizovane M-robustifikovane adaptivne tehnike filtriranja sa funkcijama uticaja koje nisu monotono neopadajuće te mogu bolje eliminisati uticaj autlajera na kvalitet praćenja objekta. Međutim, robusna kriterijumska funkcija koja odgovara takvoj funkciji uticaja nije konveksna, što može dovesti do problema sa konvergencijom algoritma praćenja tokom inicijalizacije robusnog filtra.

Alternativni način da se postigne efikasnija eliminacija uticaja autlajera na kvalitet praćenja u uslovima Gausovski raspodeljenih merenja koja nisu kontaminirana autlajerima, jeste primena

adaptivne estimacije parametara koji kvantitativno definišu funkciju uticaja. U disertaciji je predložen nov pristup gde se parametar zasićenja Huberove funkcije uticaja prilagođava u zavisnosti od situacije na sceni, koja je izražena korelacionim odzivom. Uspostavljena je originalna relacija između praga saturacije Huberove M-robustne funkcije uticaja i vršne vrednosti korelacionog odziva evaluiranog uzorka slike korelacionim filtrom. Na ovaj način je obezbeđena brža adaptacija robustnog sistema za praćenje u regularnim uslovima, ali i prilikom ponovne detekcije objekta za vreme njegovog manevra tokom perioda okluzija, kada se greške između predikcije pozicije objekta i pozicije detekcije mogu tretirati kao autlajeri, koji se eliminišu robustnom procedurom.

Dodatno unapređenje sistema za praćenje je postignuto adaptacijom statistika šuma merenja u robustnom Kalmanovom filtru. Paralelno sa robustnom estimacijom stanja objekta, na osnovu M-robustnih estimacija stanja objekta izvršeno je prilagođavanje kovarijacione matrice šuma merenja u svakom koraku robustnog Kalmanovog filtra. Važnost adaptacije statistike šuma merenja, uz adaptaciju parametara M-robustne funkcije uticaja, posebno se uočava prilikom praćenja objekata malih dimenzija na slici, kada estimacija dinamike kretanja objekta ima presudnu ulogu u proceni njegove pozicije na sceni.

Sveobuhvatna analiza sistema za praćenje pokretnih objekata u sekvenci kratkotalasnih infracrvenih slika, kombinacijom korelacionog filtra i tehnika robustne adaptivne Kalmanove filtracije, ukazuje da predložena nova metoda video praćenja postiže visoku efikasnost u praćenju objekata u regularnim uslovima, kao i zadovoljavajuću efikasnost u slučaju karakterističnih izazovnih situacija tokom praćenja. Predložena metodologija robustne estimacije uz adaptaciju statistika šuma merenja i adaptaciju parametara robustne funkcije uticaja, primenljiva je i sa drugim tipovima korelacionih filtara, ali i sa metodama praćenja koje mogu pružiti informacije o nivou pouzdanosti lokalizacije objekta.



# Literatura

- [1] A. Cavallaro, E. Maggio: “*Video Tracking: Theory and Practice*” - John Wiley & Sons, 2011.
- [2] M. Kristan, A. Leonardis, J. Matas, et al.: “*The Tenth Visual Object Tracking VOT2022 Challenge Results*” - ECCV Workshops, 2022.
- [3] M. Kristan, J. Matas, A. Leonardis, et al.: “*The Ninth Visual Object Tracking VOT2021 Challenge Results*” - ICCV Workshops, 2021.
- [4] M. Kristan, J. Matas, A. Leonardis, et al.: “*The Eighth Visual Object Tracking VOT2020 Challenge Results*” - ECCV 2020 Workshops, 2020.
- [5] M. Kristan, J. Matas, A. Leonardis, et al.: “*The Visual Object Tracking VOT2015 Challenge Results*” - ICCV Workshops, 2015.
- [6] S. Javed, M. Danelljan, F. S. Khan, M. H. Khan, M. Felsberg, J. Matas: “*Visual Object Tracking with Discriminative Filters and Siamese Networks: A Survey and Outlook*” - IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.
- [7] F. Chen, X. Wang, Y. Zhao, S. Lv, X. Niu: “*Visual Object Tracking: A Survey*” - Computer Vision and Image Understanding, 222, 103508, 2022.
- [8] M. Felsberg, A. Berg, G. Hager, et al.: “*The Thermal Infrared Visual Object Tracking VOT-TIR2015 Challenge Results*” - IEEE International Conference on Computer Vision Workshops, 2015.
- [9] K. Lebeda, S. Hadfield, R. Bowden: “*The Thermal Infrared Visual Object Tracking VOT-TIR2016 Challenge Results*” - ECCV Workshops, 2016.
- [10] Q. Liu, Z. He, X. Li, Y. Zheng: “*PTB-TIR: A Thermal Infrared Pedestrian Tracking Benchmark*” - IEEE Transactions on Multimedia, vol. 22, no. 3, pp. 666-675, 2019.
- [11] Q. Liu, X. Li, Z. He, C. Li, J. Li, Z. Zhou, D. Yuan, J. Li, K. Yang, N. Fan, F. Zheng: “*LSOTB-TIR: A Large-scale High-diversity Thermal Infrared Object Tracking Benchmark*” - Proc. of the 28th ACM International Conference on Multimedia, pp. 3847-3856, 2020.
- [12] M. P. Hansen, D. S. Malchow: “*Overview of SWIR Detectors, Cameras, and Applications*” - Thermosense Xxx, 2008, vol. 6939, p. 69390I, International Society for Optics and Photonics, 2008.
- [13] R. G. Driggers, V. Hodgkin, R. Vollmerhausen: “*What Good is SWIR? Passive Day Comparison of VIS, NIR, and SWIR*” - Proc. of SPIE, Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIV, vol. 8706, pp. 87060L, 2013.
- [14] M. Bertozzi, R. I. Fedriga, A. Miron, J. L. Reverchon: “*Pedestrian Detection in Poor Visibility Conditions: Would SWIR Help?*” - International Conference on Image Analysis and Processing, pp. 229-238, Springer, Berlin, Heidelberg, 2013.
- [15] A. Miron, A. Benschair, R. I. Fedriga, A. Broggi: “*SWIR Images Evaluation for Pedestrian Detection in Clear Visibility Conditions*” - 16th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 354-359, 2013.
- [16] B. E. Lemoff, R. B. Martin, M. Sluch, K. M. Kafka, A. Dolby, R. Ice: “*Automated, Long-Range, Night/Day, Active-SWIR Face Recognition System*” - Infrared Technology and Applications XL, vol. 9070, p. 90703I, International Society for Optics and Photonics, 2014.
- [17] Z. Kandylakis, K. Vasili, K. Karantzalos: “*Fusing Multimodal Video Data for Detecting Moving Objects/Targets in Challenging Indoor and Outdoor Scenes*” - Remote Sensing, vol. 11, no. 4, pp. 446, 2019.

- [18] C. Kwan, B. Chou, A. Echavarren, B. Budavari, J. Li, T. Tran: “*Compressive Vehicle Tracking Using Deep Learning*” - 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pp. 51-56, 2018.
- [19] C. Kwan, B. Chou, J. Yang, T. Tran: “*Compressive Object Tracking and Classification Using Deep Learning for Infrared Videos*” - Pattern Recognition and Tracking XXX, vol. 10995, p. 1099506, International Society for Optics and Photonics, 2019.
- [20] C. Kwan, B. Chou, J. Yang, A. Rangamani, T. Tran, J. Zhang, R. Etienne-Cummings: “*Target Tracking and Classification Using Compressive Sensing Camera for SWIR Videos*” - Signal, Image and Video Processing, vol. 13, no. 8, pp. 1629-1637, 2019.
- [21] C. Kwan: “*Object Tracking and Classification in Videos Using Compressive Measurements*” - Proc. of the 3rd International Conference on Vision, Image and Signal Processing, pp. 1-6, 2019.
- [22] J. Park, J. Hong, W. Shim, D. J. Jung: “*Multi-Object Tracking on SWIR Images for City Surveillance in an Edge-Computing Environment*” - Sensors, vol. 23, no. 14, pp. 6373, 2023.
- [23] V. Barnett, T. Lewis: “*Outliers in Statistical Data*” - John Wiley & Sons, 1994.
- [24] Y. Wu, J. Lim, M. Yang: “*Object Tracking Benchmark*” - IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1834–1848, 2015.
- [25] P. Liang, E. Blasch, H. Ling: “*Encoding Color Information for Visual Tracking: Algorithms and Benchmark*” - IEEE Transactions on Image Processing, vol. 24, no. 12, pp. 5630–5644, 2015.
- [26] A. Li, M. Lin, Y. Wu, M. H. Yang, S. Yan: “*NUS-PRO: A New Visual Tracking Challenge*” - IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 335-349, 2015.
- [27] M. Mueller, N. Smith, B. Ghanem: “*A Benchmark and Simulator for UAV Tracking*” - Proc. of the European Conference on Computer Vision – ECCV, pp. 445 - 461, 2016.
- [28] J. Valmadre, L. Bertinetto, J. F. Henriques, R. Tao, A. Vedaldi, A. W. Smeulders, P. H. Torr, E. Gavves: “*Long-term Tracking in the Wild: A Benchmark*” - Proc. of the European Conference on Computer Vision (ECCV), pp. 670-685, 2018.
- [29] M. Muller, A. Bibi, S. Giancola, S. Alsubaihi, B. Ghanem: “*TrackingNet: A Large-scale Dataset and Benchmark for Object Tracking in the Wild*” - Proc. of the European Conference on Computer Vision (ECCV), pp. 300-317, 2018.
- [30] C. Price: “*An Analysis of the Divergence Problem in the Kalman Filter*” - IEEE Transactions on Automatic Control, vol. 13, no. 6, pp. 699–702, 1968.
- [31] P. Hanlon, P. Maybeck: “*Characterization of Kalman Filter Residuals in the Presence of Mismodeling*” - IEEE Transactions on Aerospace and Electronic Systems, vol. 36, no. 1, pp. 114 –131, 2000.
- [32] A. E. Albert, L. A. Gardner: “*Stochastic Approximation and Nonlinear Regression*” - The MIT Press, Cambridge, MA, 1967.
- [33] G. Saridis, Z. Nikolic, K. Fu: “*Stochastic Approximation Algorithms for System Identification, Estimation, and Decomposition of Mixtures*” - IEEE Transactions on Systems Science and Cybernetics, vol. 5, no. 1, pp. 8–15, 1969.
- [34] S. Stanković, B. Kovačević: “*Analysis of Robust Stochastic Approximation Algorithms for Process Identification*” - Automatica, vol. 22, no. 4, pp. 483-488, 1986.
- [35] I. Tsympkin: “*Adaptation and Learning in Automatic Systems*” - Academic Press, New York, 1971.
- [36] J. M. Mendel: “*Adaptive Learning and Pattern Recognition Systems: Theory and Applications*” - Academic Press, New York, 1970.
- [37] W. N. Venables, B.D. Ripley: “*Modern Applied Statistics with S-PLUS*” - Springer, 2013.
- [38] R. R. Wilcox: “*Introduction to Robust Estimation and Hypothesis Testing*” - Academic Press, 2011.

- [39] F. R. Hampel, E. N. Ronchetti, P. J. Rousseeuw, W. A. Stahel: “*Robust Statistics: The Approach Based on Influence Functions*” - Wiley, 1986.
- [40] P. J. Huber, E. M. Ronchetti: “*Robust Statistics*” - Hoboken, NJ, USA: Wiley, 2009.
- [41] M. A. Gandhi, L. Mili: “*Robust Kalman Filter Based on a Generalized Maximum-Likelihood-Type Estimator*” - IEEE Transactions on Signal Processing, vol. 58, no. 5, pp. 2509–2520, 2010.
- [42] L. Chang, B. Hu, G. Chang, A. Li: “*Robust Derivative-free Kalman Filter Based on Huber's M-estimation Methodology*” - Journal Of Process Control, vol. 23, no. 10, pp. 1555-1561, 2013.
- [43] D. De Menezes, D. Prata, A. Secchi, J. Pinto: “*A Review on Robust M-Estimators for Regression Analysis*” - Computers & Chemical Engineering, vol. 147, 107254, 2021.
- [44] Z. Banjac, Ž. Đurović, B. Kovačević: “*Approximate Kalman Filtering Using Robustified Dynamic Stochastic Approximation Method*” - 26th Telecommunications Forum (TELFOR), pp. 353 –356, 2018.
- [45] Z. Banjac, B. Kovačević: “*Robustified Kalman Filtering Using Both Dynamic Stochastic Approximation and M-robust Performance Index*” - Tech. Gazette, vol. 29, no. 3, pp. 907–914, 2022.
- [46] A. Gelb: “*Applied Optimal Estimation*” - Analytic Sciences Corporation, Cambridge, MA: MIT Press, 2010.
- [47] M. S. Grewal, A. P. Andrews: “*Kalman Filtering: Theory and Practice with Matlab*” - John Wiley & Sons, 2015.
- [48] R. Stengel: “*Stochastic Optimal Control: Theory and Application*” - Wiley, New York, 1986.
- [49] J. Valluru, S. Patwardhan, L. Biegler: “*Development of Robust Extended Kalman Filter and Moving Window Estimator for Simultaneous State and Parameter/Disturbance Estimation*” - Journal Of Process Control, vol. 69, pp. 158-178, 2018.
- [50] M. Murata, H. Nagano, K. Kashino: “*Unscented Statistical Linearization and Robustified Kalman Filter for Nonlinear Systems with Parameter Uncertainties*” - American Control Conference, 2014.
- [51] K. Li, B. Hu, L. Chang, Y. Li: “*Robust Square-root Cubature Kalman Filter Based on Huber's M-estimation Methodology*” - Proc. of the Institution of Mechanical Engineers, Part G, Journal of Aerospace Engineering, vol. 229, no. 7, pp. 1236–1245, 2014.
- [52] B. Stark, M. McGee, Y. Chen: “*Short Wave Infrared (SWIR) Imaging Systems Using Small Unmanned Aerial Systems (sUAS)*” - International Conference on Unmanned Aircraft Systems (ICUAS), pp. 495-501, 2015.
- [53] [http://modtran.spectral.com/modtran\\_home](http://modtran.spectral.com/modtran_home) [Na mreži].
- [54] C. Besikci: “*Extended Short Wavelength Infrared FPA Technology: Status and Trends*” - Quantum Sensing and Nano Electronics and Photonics XV, 10540, pp. 110-125, 2018.
- [55] K. K. Choi, J. Sun, K. Olver: “*Resonator-QWIP FPA Development*” - Proc. of SPIE, vol. 9451, pp. 687 - 698, 2015.
- [56] D. Peric, B. Livada: “*Analysis of SWIR Imagers Application in Electro-optical Systems*”- Proc. of the 4th International Conference on Electrical, Electronics and Computing Engineering, IcETRAN 2017, Kladovo, Serbia, 2017.
- [57] A. Arnulf, J. Bricard, E. Curé, C. Vêret: “*Transmission by Haze and Fog in the Spectral Region 0.35 to 10 Microns*” - Journal of the Optical Society of America, vol. 47, no. 6, pp. 491-498, 1957.
- [58] [https://www.vlatacominstitute.com/\\_files/ugd/510d2b\\_ab410776328144979064c9cfa9bda036.pdf](https://www.vlatacominstitute.com/_files/ugd/510d2b_ab410776328144979064c9cfa9bda036.pdf) [Na mreži].

- [59] D. Perić, B. Livada, M. Perić: “*Analysis and Selection of Components for Active SWIR/NIR Vision Systems*” - Proc. of the 4th International Conference on Electrical, Electronics and Computing Engineering, IcETRAN 2017, Kladovo, Serbia, 2017.
- [60] J. Bentell, P. Nies, J. Cloots, J. Vermeiren, B. Grietens, O. David, A. Shurkun, R. Schneider: “*Flip Chipped InGaAs Photodiode Arrays for Gated Imaging with Eye-safe Lasers*” - TRANSDUCERS, IEEE International Solid-State Sensors, Actuators and Microsystems Conference, pp. 1103-1106, 2007.
- [61] O. Steinvall, F. Berglund, L. Allard, J. Öhgren, H. Larsson, E. Amselem, F. Gustafsson, E. Repasic, P. Lutzmann, B. Göhler, M. Hammer, K. McEwend, K. McEwane: “*Passive and Active EO Sensing of Small Surface Vessels*” - Proc. of SPIE, Infrared Technology and Applications XXXIV, vol. 6940, 2008.
- [62] E. Dvinelis, T. Žukauskas, M. Kaušylas, A. Vizbaras, K. Vizbaras, D. Vizbaras: “*Laser Illumination and EO Systems for Covert Surveillance from NIR to SWIR and Beyond*” - Proc. of SPIE, Electro-Optical and Infrared Systems: Technology and Applications XIII, vol. 9987, pp. 11-18, 2016.
- [63] B. DeCann, A. Ross, J. Dawson: “*Investigating Gait Recognition in the Short-Wave Infrared (SWIR) Spectrum: Dataset and Challenges*” - Biometric and Surveillance Technology for Human and Activity Identification X, vol. 8712, pp. 101-116, International Society for Optics and Photonics, 2013.
- [64] B. E. Lemoff, R. B. Martin, M. Sluch, K. M. Kafka, W. McCormick, R. Ice: “*Long-range Night/Day Human Identification Using Active-SWIR Imaging*” - Infrared Technology and Applications XXXIX, vol. 8704, p. 87042J, International Society for Optics and Photonics, 2013.
- [65] B. Vishnyakov, Y. Blokhinov, I. Sgibnev, V. Sheverdin, A. Sorokin, A. Nikanorov, P. Masalov, K. Kazakhmedov, S. Brianskiy, E. Andrienko, Y. Vizilter: “*Semantic Scene Understanding for the Autonomous Platform*” - The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 43, pp. 637-644, 2020.
- [66] H. Fan, L. Lin, F. Yang, P. Chu, G. Deng, S. Yu, H. Bai, Y. Xu, C. Liao, H. Ling: “*LaSOT: A High-quality Benchmark for Large-scale Single Object Tracking*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5374–5383, 2019.
- [67] L. Huang, X. Zhao, K. Huang: “*GOT-10K: A Large High-diversity Benchmark for Generic Object Tracking in the Wild*” - IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 5, pp. 1562–1577, 2019.
- [68] A. Berg, J. Ahlberg, M. Felsberg: “*A Thermal Object Tracking Benchmark*” - 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2015.
- [69] A. Bewley, Z. Ge, L. Ott, F. Ramos, B. Upcroft: “*Simple Online and Realtime Tracking*” - Proc. of IEEE International Conference on Image Processing, pp. 3464-3468, 2016.
- [70] N. Wojke, A. Bewley, D. Paulu: “*Simple Online and Realtime Tracking with a Deep Association Metric*” - Proc. of the IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017.
- [71] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, X. Wang: “*Bytetrack: Multi-Object Tracking by Associating Every Detection Box*” - Proc. of the European Conference on Computer Vision (ECCV), pp. 1-21, Cham: Springer Nature Switzerland, 2022.
- [72] A. Yilmaz, O. Javed, M. Shah: “*Object Tracking: A Survey*” - ACM Computing Surveys (CSUR), vol. 38, no. 4, 2006.
- [73] M. Pavlović, N. Vlahović, M. Perić, A. Simić, S. Stanković: “*Real-Time Dead Pixels Removal in Thermal Imaging*” - Proc. of the 6th IcETRAN, pp. 345-348, 2019.

- [74] M. Pavlovic, Z. Banjac, B. Kovacevic: “*Digital Video Stabilization Verification Based on Genetic Algorithm Template Matching*” - Advances in Electrical & Computer Engineering, vol. 22, no. 2, 2022.
- [75] S. Puzović, R. Petrović, M. Pavlović, S. Stanković: “*Enhancement Algorithms for Low-light and Low-contrast Images*” - Proc. of the 19th IEEE International Symposium INFOTEH-JAHORINA, pp. 1-6, 2020.
- [76] W. Wang, X. Wei, J. Li, G. Wang: “*Noise Suppression Algorithm of Short-Wave Infrared Star Image for Daytime Star Sensor*” - Infrared Physics & Technology, vol. 85, pp. 382-394, 2017.
- [77] M. N. Chapel, T. Bouwmans: “*Moving Objects Detection with a Moving Camera: A Comprehensive Review*” - Computer Science Review, 38, 100310, 2020.
- [78] B. Garcia-Garcia, T. Bouwmans, A. J. R. Silva: “*Background Subtraction in Real Applications: Challenges, Current Models and Future Directions*” - Computer Science Review, 35, 100204, 2020.
- [79] P. KaewTraKulPong, R. Bowden: “*An Improved Adaptive Background Mixture Model for Real-Time Tracking with Shadow Detection*” - Proc. of the 2nd European Workshop on Advanced Video-Based Surveillance System: Computer Vision and Distributed Processing, pp. 135-144, 2001.
- [80] S. S. Beauchemin, J. L. Barron: “*The Computation of Optical Flow*” - ACM Computing Surveys (CSUR), vol. 27, no. 3, pp. 433-466, 1995.
- [81] D. Sun, X. Yang, M. Y. Liu, J. Kautz: “*Models Matter, So Does Training: An Empirical Study of CNNs for Optical Flow Estimation*” - IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 6, pp. 1408-1423, 2018.
- [82] D. Cizelj, M. Pavlović, T. Unkašević, Z. Banjac: “*Automated Target Tracking Activation Based on Motion Detection*” - Proc. of the 10th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN 2023), Sarajevo, B&H, 2023.
- [83] M. H. Kolekar: “*Intelligent Video Surveillance Systems: An Algorithmic Approach*” - CRC Press/Taylor & Francis Group: Boca Raton, FL, USA, 2019.
- [84] I. Goodfellow, Y. Bengio, A. Courville: “*Deep Learning*” - MIT Press: Cambridge, MA, USA, 2016.
- [85] N. Latinović, I. Popadić, B. Tomić, A. Simić, P. Milanović, S. Nijemčević, M. Perić, M. Veinović: “*Signal Processing Platform for Long-Range Multi-Spectral Electro-Optical Systems*” - Sensors, vol. 22, no. 3, 1294, 2022.
- [86] M. Kim, S. Joung, K. Park, S. Kim, K. Sohn: “*Unpaired Cross-Spectral Pedestrian Detection Via Adversarial Feature Learning*” - Proc. of the 2019 IEEE International Conference on Image Processing (ICIP), pp. 1650–1654, Taipei, Taiwan, 2019.
- [87] M. S. Pavlović, P. D. Milanović, M. S. Stanković, D. B. Perić, I. V. Popadić, M. V. Perić: “*Deep Learning Based SWIR Object Detection in Long-Range Surveillance Systems: An Automated Cross-Spectral Approach*” - Sensors, vol. 22, no. 7, 2562, 2022.
- [88] B. Livada, D. Peric, M. Peric: “*Challenges of Laser Range Finder Integration in Electro-Optical Surveillance System*” - Proc. of the 4th International Conference on Electrical, Electronic, and Computing Engineering (IcETRAN 2017), Kladovo, Serbia, 2017.
- [89] S. Mambo: “*Optimisation and Performance Evaluation in Image Registration Technique*” - Doctoral Dissertation, Université Paris-Est, Créteil, France; Tshwane University of Technology, Pretoria, South Africa, 2018.
- [90] Z. Ge, S. Liu, F. Wang, Z. Li, J. Sun: “*Yolox: Exceeding Yolo Series in 2021*” - arXiv 2021, arXiv:2107.08430, 2021.
- [91] T. Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, C. L. Zitnick, P. Dollár: “*Microsoft COCO: Common Objects in Context*” arXiv 2015, arXiv:1405. [na mreži]: <https://cocodataset.org/>

- [92] <https://github.com/cocodataset/cocoapi/blob/master/PythonAPI/pycocotools/cocoeval.py> [na mreži]
- [93] M. Pavlović, N. Stojiljković, I. Gluvačević, M. Vučetić, M. Perić: “*Real-Time Moving Object of Interest Detection in Multi-Sensor Imaging System*” - Proc. of the 7th International Conference on Electrical, Electronic, and Computing Engineering (IcETRAN 2020), Belgrade, Serbia, 2020.
- [94] S. M. Marvasti-Zadeh, L. Cheng, H. Ghanei-Yakhdan, S. Kasaei: “*Deep Learning for Visual Tracking: A Comprehensive Survey*” - IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 5, pp. 3943-3968, 2021.
- [95] M. Ondrašovič, P. Tarábek: “*Siamese Visual Object Tracking: A Survey*” - IEEE Access, vol. 9, 110149-110172, 2021.
- [96] X. Li, Q. Liu, N. Fan, Z. He, H. Wang: “*Hierarchical Spatial-aware Siamese Network for Thermal Infrared Object Tracking*” - Knowledge-Based Systems, vol. 166, pp. 71-81, 2019.
- [97] Q. Liu, X. Li, Z. He, N. Fan, D. Yuan, H. Wang: “*Learning Deep Multi-level Similarity for Thermal Infrared Object Tracking*” - IEEE Transactions on Multimedia, vol. 23, pp. 2114-2126, 2020.
- [98] D. S. Bolme, J. R. Beveridge, B. A. Draper, Y. M. Lui: “*Visual Object Tracking Using Adaptive Correlation Filters*” - IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2544-2550, 2010.
- [99] J. F. Henriques, R. Caseiro, P. Martins, J. Batista: “*High-speed Tracking with Kernelized Correlation Filters*” - IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 3, pp. 583-596, 2014.
- [100] M. Danelljan, G. Häger, F. Khan, M. Felsberg: “*Accurate Scale Estimation for Robust Visual Tracking*” - British Machine Vision Conference, 2014.
- [101] M. Danelljan, G. Hager, F. Shahbaz Khan, M. Felsberg: “*Learning Spatially Regularized Correlation Filters for Visual Tracking*” - Proc. of the IEEE International Conference on Computer Vision, pp. 4310-4318, 2015.
- [102] C. Ma, J. B. Huang, X. Yang, M. H. Yang: “*Hierarchical Convolutional Features for Visual Tracking*” - Proc. of the IEEE International Conference on Computer Vision, pp. 3074-3082, 2015.
- [103] L. Bertinetto, J. Valmadre, S. Golodetz, O. Miksik, P. H. Torr: “*Staple: Complementary Learners for Real-time Tracking*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1401-1409, 2016.
- [104] M. Danelljan, A. Robinson, F. K. Khan, M. Felsberg: “*Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking*” - Proc. of the European Conference on Computer Vision (ECCV), pp. 472-488, 2016.
- [105] M. Danelljan, G. Bhat, F. Shahbaz Khan, M. Felsberg: “*ECO: Efficient Convolution Operators for Tracking*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6638-6646, 2017.
- [106] A. Lukežić, T. Vojir, L. Cehovin Zajc, J. Matas, M. Kristan: “*Discriminative Correlation Filter with Channel and Spatial Reliability*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6309-6318, 2017.
- [107] F. Li, C. Tian, W. Zuo, L. Zhang, M. H. Yang: “*Learning Spatial-Temporal Regularized Correlation Filters for Visual Tracking*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4904-4913, 2018.
- [108] G. Bhat, J. Johnander, M. Danelljan, F. Shahbaz Khan, M. Felsberg: “*Unveiling the Power of Deep Tracking*” - Proc. of the European Conference on Computer Vision (ECCV), pp. 483-498, 2018.

- [109] G. Bhat, M. Danelljan, L. V. Gool, R. Timofte: “*Learning Discriminative Model Prediction for Tracking*” - Proc. of the IEEE/CVF International Conference on Computer Vision, pp. 6182-6191, 2019.
- [110] M. Danelljan, G. Bhat, F. S. Khan, M. Felsberg: “*Atom: Accurate Tracking by Overlap Maximization*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4660-4669, 2019.
- [111] G. Bhat, M. Danelljan, L. Van Gool, R. Timofte: “*Know Your Surroundings: Exploiting Scene Information for Object Tracking*” - Proc. of the European Conference on Computer Vision (ECCV), pp. 205-221, 2020.
- [112] Y. Li, C. Fu, F. Ding, Z. Huang, G. Lu: “*AutoTrack: Towards High-performance Visual Tracking for UAV with Automatic Spatio-temporal Regularization*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11923-11932, 2020.
- [113] T. Xu, Z. Feng, X. J. Wu, J. Kittler: “*Adaptive Channel Selection for Robust Visual Object Tracking with Discriminative Correlation Filters*” - International Journal of Computer Vision, vol. 129, pp. 1359-1375, 2021.
- [114] D. Yuan, X. Chang, Z. Li, Z. He: “*Learning Adaptive Spatial-temporal Context-aware Correlation Filters for UAV Tracking*” - ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 18, no. 3, pp. 1-18, 2022.
- [115] Y. Liu, H. Yan, W. Zhang, M. Li, L. Liu, “*An Adaptive Spatiotemporal Correlation Filtering Visual Tracking Method*” - Plos One, vol. 18, no. 1, 2023.
- [116] L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, P. H. Torr: “*Fully-convolutional Siamese Networks for Object Tracking*” - Computer Vision – ECCV 2016 Workshops, Lecture Notes in Computer Science, vol. 9914. Springer, 2016.
- [117] D. Held, S. Thrun, S. Savarese: “*Learning to Track at 100 fps with Deep Regression Networks*” - Proc. of the European Conference on Computer Vision (ECCV), pp. 749-765, 2016.
- [118] Q. Guo, W. Feng, C. Zhou, R. Huang, L. Wan, S. Wang: “*Learning Dynamic Siamese Network for Visual Object Tracking*” - Proc. of the IEEE International Conference on Computer Vision, pp. 1763-1771, 2017.
- [119] B. Li, J. Yan, W. Wu, Z. Zhu, X. Hu: “*High Performance Visual Tracking with Siamese Region Proposal Network*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8971-8980, 2018.
- [120] B. Li, W. Wu, Q. Wang, F. Zhang, J. Xing, J. Yan: “*Siamrpn++: Evolution of Siamese Visual Tracking with Very Deep Networks*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4282-4291, 2019.
- [121] Q. Wang, L. Zhang, L. Bertinetto, W. Hu, P. H. Torr: “*Fast Online Object Tracking and Segmentation: A Unifying Approach*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1328-1338, 2019.
- [122] Z. Chen, B. Zhong, G. Li, S. Zhang, R. Ji: “*Siamese Box Adaptive Network for Visual Tracking*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6668-6677, 2020.
- [123] D. Guo, J. Wang, Y. Cui, Z. Wang, S. Chen: “*SiamCAR: Siamese Fully Convolutional Classification and Regression for Visual Tracking*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6269-6277, 2020.
- [124] J. Zheng, C. Ma, H. Peng, X. Yang: “*Learning to Track Objects from Unlabeled Videos*” - Proc. of the IEEE/CVF International Conference on Computer Vision, pp. 13546-13555, 2021.
- [125] Q. Shen, L. Qiao, J. Guo, P. Li, X. Li, B. Li, W. Feng, W. Gan, W. Wu, W. Ouyang: “*Unsupervised Learning of Accurate Siamese Tracking*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8101-8110, 2022.

- [126] X. Chen, B. Yan, J. Zhu, D. Wang, X. Yang, H. Lu: “*Transformer Tracking*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern recognition, pp. 8126-8135, 2021.
- [127] L. Lin, H. Fan, Y. Xu, H. Ling: “*SwinTrack: A Simple and Strong Baseline for Transformer Tracking*” - Advances in Neural Information Processing Systems, vol. 35, pp. 16743–16754, 2021.
- [128] Y. Cui, C. Jiang, L. Wang, G. Wu: “*Mixformer: End-to-end Tracking with Iterative Mixed Attention*” - Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13608–13618, 2022.
- [129] Z. Fu, Z. Fu, Q. Liu, W. Cai, Y. Wang: “*Sparsett: Visual Tracking with Sparse Transformers*” - arXiv preprint arXiv:2205.03776, 2022.
- [130] Q. Liu, X. Lu, Z. He, C. Zhang, W. S. Chen: “*Deep Convolutional Neural Networks for Thermal Infrared Object Tracking*” - Knowledge-Based Systems, vol. 134, pp. 189-198, 2017.
- [131] M. A. Zulkifley, N. Trigoni: “*Multiple-model Fully Convolutional Neural Networks for Single Object Tracking on Thermal Infrared Video*” - IEEE Access, vol. 6, pp. 42790-42799, 2018.
- [132] L. Zhang, A. Gonzalez-Garcia, J. Van De Weijer, M. Danelljan, F. S. Khan: “*Synthetic Data Generation for End-to-end Thermal Infrared Tracking*” - IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 1837-1850, 2018.
- [133] Q. Liu, X. Li, Z. He, N. Fan, D. Yuan, W. Liu, Y. Liang: “*Multi-task Driven Feature Models for Thermal Infrared Tracking*” - Proc. of the AAAI Conference on Artificial Intelligence, vol. 34, no. 7, pp. 11604-11611, 2020.
- [134] H. Zhang, Z. Yin, H. Zhang: “*Thermal Infrared Object Tracking Using Correlation Filters Improved by Level Set*” - Signal, Image and Video Processing, vol. 17, no. 3, pp. 791-797, 2023.
- [135] Q. Liu, D. Yuan, N. Fan, P. Gao, X. Li, Z. He: “*Learning Dual-level Deep Representation for Thermal Infrared Tracking*” - IEEE Transactions on Multimedia, vol. 25, pp. 1269-1281, 2022.
- [136] C. Hester, D. Casasent: “*Multivariant Technique for Multiclass Pattern Recognition*” - Applied Optics, vol. 19, no. 11, pp.1758–1761, 1980.
- [137] D. Casasent: “*Unified Synthetic Discriminant Function Computational Formulation*” - Applied Optics, vol. 23, no. 10, pp. 1620-1627, 1984.
- [138] B. V. K. Vijaya Kumar: “*Minimum-variance Synthetic Discriminant Functions*” - Journal of the Optical Society of America, vol. 3, no. 10, pp. 1579-1584, 1986.
- [139] A. Mahalanobis, B. V. K. Vijaya Kumar, D. Casasent: “*Minimum Average Correlation Energy Filters*” - Applied Optics, vol. 26, no. 17, pp. 3633 3640, 1987.
- [140] P. Refregier: “*Filter Design for Optical Pattern Recognition: Multicriteria Optimization Approach*” - Optics Letters, vol. 15, no. 15, pp. 854-856, 1990.
- [141] P. Refregier: “*Optimal Trade-off Filters for Noise Robustness, Sharpness of the Correlation Peak, and Horner Efficiency*” - Optics Letters, vol. 16, no. 829–832, 1991.
- [142] B. V. K. Vijaya Kumar, A. Mahalanobis, S. Song, S. Sims, J. Epperson: “*Minimum Squared Error Synthetic Discriminant Functions*” - Optical Engineering, vol. 31, no. 5, pp. 915-922, 1992.
- [143] B. V. Kumar, A. Mahalanobis, R. D. Juday: “*Correlation Pattern Recognition*” - Cambridge University Press, 2005.
- [144] A. Mahalanobis, B. V. Kumar, S. Song, S. Sims, J. Epperson: “*Unconstrained Correlation Filters*” - Applied Optics, vol. 33, no. 17, pp. 3751–3759, 1994.
- [145] B. Schölkopf, A. Smola: “*Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*” - MIT Press, 2002.



- [146] F. Chen, W. Xie, T. Xia: “*Target Tracking Algorithm Based on Kernel Correlation Filter with Anti-occlusion Mechanisms*” - Proc. of the 15th IEEE International Conference on Signal Processing (ICSP), vol. 1, pp. 220-225, 2020.
- [147] T. Zhou, M. Zhu, D. Zeng, H. Yang: “*Scale Adaptive Kernelized Correlation Filter Tracker with Feature Fusion*” - Mathematical Problems in Engineering, 2017.
- [148] L. Gan, Y. Ma: “*Long-term Correlation Filter Tracking Algorithm Based on Adaptive Feature Fusion*” - Proc. of Thirteenth International Conference on Graphics and Image Processing (ICGIP 2021), vol. 12083, pp. 253-263, SPIE, 2021.
- [149] L. Liu, T. Feng, Y. Fu: “*Learning Multi-feature Correlation Filter and Saliency Redetection for Long-term Object Tracking*” - Symmetry, vol. 14, no. 5, pp. 911, 2022.
- [150] Y. Zhou, W. Yang, Y. Shen: “*Scale-adaptive KCF Mixed with Deep Feature for Pedestrian Tracking*” - Electronics, vol.10, no. 5, pp. 536, 2021.
- [151] J. Zhang, S. Jiang, Y. Zhang, X. Liu, D. Wang, F. Qiu: “*Long-term Tracking Algorithm Using Deep Features and a Single Shot Multibox Detector*” - Journal of Electronic Imaging, vol. 27, no. 5, pp. 053019, 2018.
- [152] Y. Liu, Y. Liao, C. Lin, Y. Jia, Z. Li, X. Yang: “*Object Tracking in Satellite Videos Based on Correlation Filter with Multi-Feature Fusion and Motion Trajectory Compensation*” - Remote Sensing, vol. 14, no. 3, pp. 777, 2022.
- [153] X. Fang, H. Wang, L. Tao, X. Wang: “*Robust Kernel Correlation Filter Tracker Based on Adaptive Scale and Occlusion Detection*” - Twelfth International Conference on Graphics and Image Processing (ICGIP), vol. 11720, pp. 119-128, SPIE, 2020.
- [154] N. Liang, G. Wu, W. Kang, Z. Wang, D. D. Feng: “*Real-Time Long-Term Tracking with Prediction-Detection-Correction*” - IEEE Transactions on Multimedia, vol. 20, no. 9, pp. 2289-2302, 2018.
- [155] M. Jiang, J. Shen, J. Kong, B. Wang: “*Robust Visual Tracking Based on Kernelized Correlation Filters*” - IEEE International Conference on Information and Automation, pp. 110-115, 2017.
- [156] F. Weiyan, Z. Ting, L. Cong, Z. Xiang: “*A KCF-Based Tracker with Faster Scale Estimation and Effective Occlusion Detection*” - Chinese Control And Decision Conference (CCDC), pp. 762-766, 2019.
- [157] M. Tang, J. Feng: “*Multi-kernel Correlation Filter for Visual Tracking*” - IEEE International Conference on Computer Vision (ICCV), pp. 3038 - 3046, 2015.
- [158] Y. Xu, J. Wang, Y. Li, Z. Miao, Y. Zhang: “*One-Step Backtracking for Occlusion Detection in Real-Time Visual Tracking*” - Electronics Letters, vol. 53, no. 5, pp. 318-320, 2017.
- [159] X. Niu, Y. Qiao: “*Context-Based Occlusion Detection for Robust Visual Tracking*” - IEEE International Conference on Image Processing (ICIP), pp. 3655-3659, IEEE, 2017.
- [160] M. Zolfaghari, H. Ghanei-Yakhdan, M. Yazdi: “*Real-Time Object Tracking Based on an Adaptive Transition Model and Extended Kalman Filter to Handle Full Occlusion*” - The Visual Computer, vol. 36, no. 4, pp. 701-715, 2020.
- [161] H. A. Wei: “*UAV Target Prediction and Tracking Method Based on KCF and Kalman Filter Hybrid Algorithm*” - Proc. of the 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 711-718, IEEE, 2022.
- [162] X. Niu, Z. Cui, S. Geng, J. Yang, Y. Qiao: “*Robust Visual Tracking via Occlusion Detection Based on Depth-Layer Information*” - International Conference on Neural Information Processing, pp. 44-53, Springer, Cham, 2017.
- [163] S. Bai, X. Tang, J. Zhang: “*Research on Object Tracking Algorithm Based on KCF*” - International Conference on Culture-Oriented Science & Technology (ICCST), pp. 255-259, IEEE, 2020.

- [164] I. Karakostas, V. Mygdalis, A. Tefas, I. Pitas: “*Occlusion Detection and Drift-Avoidance Framework for 2D Visual Object Tracking*” - Signal Processing: Image Communication, vol. 90, pp. 116011, 2021.
- [165] Y. Zhang, C. G. Xu, F. Zhang, L. Q. Song: “*Research on Target Loss Early Warning of KCF Algorithm Based on Hypothesis Test*” - Optical Sensing and Imaging Technologies and Applications, vol. 10846, pp. 388-393, SPIE, 2018.
- [166] J. Shin, H. Kim, D. Kim, J. Paik: “*Fast and Robust Object Tracking Using Tracking Failure Detection in Kernelized Correlation Filter*” - Applied Sciences, vol. 10, no. 2, pp. 713, 2020.
- [167] R. Xia, Y. Chen, B. Ren: “*Improved Anti-Occlusion Object Tracking Algorithm Using Unscented Rauch-Tung-Striebel Smoother and Kernel Correlation Filter*” - Journal of King Saud University-Computer and Information Sciences, 2022.
- [168] H. Kiani Galoogahi, T. Sim, S. Lucey: “*Correlation Filters with Limited Boundaries*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4630-4638, 2015.
- [169] M. Mueller, N. Smith, B. Ghanem: “*Context-aware Correlation Filter Tracking*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1396-1404, 2017.
- [170] C. S. Asha, A. V. Narasimhadhan: “*Adaptive Learning Rate for Visual Tracking Using Correlation Filter*” - Procedia Computer Science, vol. 89, pp. 614-622, 2016.
- [171] C. Xiu, Y. Ma: “*Kernel Correlation Filter Tracking Strategy Based on Adaptive Fusion Response Map*” - IET Image Processing, vol. 16, no. 4, pp. 937-947, 2022.
- [172] S. Jeong, J. Paik: “*Partial Block Scheme and Adaptive Update Model for Kernelized Correlation Filters-Based Object Tracking*” - Applied Sciences, vol. 8, no. 8, pp. 1349, 2018.
- [173] J. Wang, H. Yang, N. Xu, C. Wu, Z. Zhao, J. Zhang, D. O. Wu: “*Long-Term Target Tracking Combined With Re-Detection*” - EURASIP Journal on Advances in Signal Processing, 2021(1), pp. 1-16, 2021.
- [174] T. Li, S. Zhao, Q. Meng, Y. Chen, J. Shen: “*A Stable Long-Term Object Tracking Method with Re-Detection Strategy*” - Pattern Recognition Letters, vol. 127, pp. 119-127, 2018.
- [175] J. Qu, C. Tang, Y. Zhang, K. Zhou, A. Razi: “*Long-Time Target Tracking Algorithm Based on Re-Detection Multi-Feature Fusion*” - IET Cyber-Systems and Robotics, vol. 4, no. 1, pp. 38-50, 2022.
- [176] T. V. Kieu, L. P. Nguyen: “*Fast and Robust Long-term Tracking with Effective Searching Model*” - International Journal of Computer and Information Engineering, vol. 15, no. 7, pp. 418-422, 2021.
- [177] Y. F. Liu, Y. He, Q. Tian, J. Yang: “*KCF Tracking Algorithm Based on Outlier Detection*” - Recent Developments in Intelligent Computing, Communication and Devices, pp. 129-135, Springer, Singapore, 2019.
- [178] Q. He, W. Zhang, W. Chen, G. Xie, Y. Yao: “*Target Tracking Algorithm Combined Part-Based and Redetection for UAV*” - EURASIP Journal on Wireless Communications and Networking, 2020(1), pp. 1-17, 2020.
- [179] J. Liu, P. Ye, X. Zhang, G. Xiao: “*Real-Time Long-Term Tracking with Reliability Assessment and Object Recovery*” - IET Image Processing, vol. 15, no. 4, pp. 918-935, 2021.
- [180] L. Yu, C. Fan, Y. Ming: “*A Visual Tracker Based on Improved Kernel Correlation Filter*” - Proc. of the 7th International Conference on Internet Multimedia Computing and Service, pp. 1-4, 2015.
- [181] X. P. Huynh, I. H. Choi, Y. G. Kim: “*Tracking a Human Fast and Reliably Against Occlusion and Human-Crossing*” - Image and Video Technology, pp. 461-472, Springer, Cham, 2015.
- [182] F. Yue, X. Li: “*Improved Kernelized Correlation Filter Algorithm and Application in the Optoelectronic Tracking System*” - International Journal of Advanced Robotic Systems, vol. 15, no. 3, 2018.

- [183] W. Cheng-yun, Z. Zu-xing, S. Chao, G. Yu, Z. Long-yun, J. Yi: “*Kernelized Correlation Filtering Algorithm Fused with Kalman Filter*,” - 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), IEEE, pp. 353-357, 2020.
- [184] M. U. Mustafa, Z. L. Yu: “*A Robust Object Tracking Method for Surveillance Applications to Handle Occlusion*” - International Journal of Computer Theory and Engineering, vol. 13, no. 1, 2021.
- [185] B. Kovačević, Z. Banjac: “*Stohastički Sistemi*” - Akademska Misao, Akademija Inženjerskih Nauka Srbije, 2020.
- [186] B. Kovačević, Z. Banjac, Ž. Đurović: “*Filtracija Stohastičkih Signala - Optimalni, Adaptivni i Robusni Estimatori Parametara i Stanja*” - Akademska Misao, 2017.
- [187] B. Kovačević, Ž. Đurović: “*Fundamentals of Stochastic Signals, Systems and Estimation Theory with Worked Examples*” - Berlin: Springer, 2011.
- [188] B. Lei, G. Xu, M. Feng, Y. Zou, F. Van der Heijden, D. Ridder, D. M. Tax: “*Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*” - John Wiley & Sons, 2017.
- [189] B. Kovačević, Z. Banjac, T. Unkašević: “*Robustified Kalman Filtering Using Both Dynamic Stochastic Approximation and M-Robust Performance Index*” - Kalman Filters - Theory, Applications, and Optimization, eds. A. Khalid, A. I. Sarwat, H. Riggs, IntechOpen, London, UK, 2024.
- [190] J. Jurečková: “*21 M-, L- and R-estimators*” - Handbook of Statistics, vol. 4, pp. 463-485, 1984.
- [191] S. C. Chapra, R. P. Canale: “*Numerical Methods for Engineers*” - McGraw-Hill Education, New York, NY, 2015.
- [192] M. Verhaegen, V. Verdult: “*Filtering and System Identification: A Least Squares Approach*” - Cambridge University Press, Cambridge, 2012.
- [193] T. Young, R. Westerberg: “*Error Bounds for Stochastic Estimation of Signal Parameters*” - IEEE Transactions on Information Theory, vol. 17, no. 5, no.549–557, 1971.
- [194] M. Pavlović, Z. Banjac, B. Kovačević: “*Approximate Kalman Filtering by Both M-robustified Dynamic Stochastic Approximation and Statistical Linearization Methods*” - EURASIP Journal on Advances in Signal Processing, 2023(1), pp. 1-29, 2023.
- [195] B. Kovačević, M. Milosavljević, M. Veinović, M. Marković: “*Robust Digital Processing of Speech Signals*” - Springer, 2017.
- [196] B. Kovačević, Z. Banjac, M. Milosavljević: “*Adaptive Digital Filters*” - Springer Berlin Heidelberg, 2013.
- [197] N. Aharon, R. Orfaig, B. Z. Bobrovsky: “*BoT-SORT: Robust Associations Multi-Pedestrian Tracking*” - arXiv preprint arXiv:2206.14651, 2022.
- [198] Z. Wu, J. Lim, M. H. Yang: “*Online Object Tracking: A Benchmark*” - Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2411-2418, 2013.
- [199] S. S. Blackman, R. Popoli: “*Design and Analysis of Modern Tracking Systems*” - Norwood, MA: Artech House, 1999.
- [200] N. Vlahović, Z. Đurović: “*Robust Tracking of Moving Objects Using Thermal Camera and Speeded Up Robust Features Descriptor*” - International Journal of Adaptive Control and Signal Processing, vol. 35, no. 4, pp. 549-566, 2021.
- [201] M. Pavlović, Z. Banjac, B. Kovačević: “*Object Tracking in SWIR Imaging Based on Both Correlation and Robust Kalman Filters*” - IEEE Access, vol. 11, pp. 63834-63851, 2023.

# Biografija

Miloš Pavlović je rođen 23.06.1995. godine u Zaječaru. Završio je osnovnu školu “Branko Radičević” u Negotinu kao nosilac Vukove diplome. Upisao je gimnaziju u Negotinu koju je završio kao nosilac Vukove diplome i đak generacije. Elektrotehnički fakultet upisao je 2014. godine. Diplomirao je na odseku za Signale i sisteme 2018. godine sa prosečnom ocenom 9,33. Diplomski rad pod nazivom “*Sistem za identifikaciju osoba zasnovan na prepoznavanju govora i lica*” pod mentorstvom prof. dr Željka Đurovića odbranio je u avgustu 2018. godine. Od strane Matematičkog instituta Srpske akademije nauka i umetnosti dobio je nagradu za najbolji studentski rad u oblasti računarstva. Na 5. IcETRAN konferenciji dobio je nagradu za najbolji rad mladog istraživača. Nakon završetka osnovnih studija radio je praksu na Elektrotehničkom Institutu Univerziteta u Drezdenu u Nemačkoj, gde se bavio upravljačkim algoritmima sistema za napajanje u električnim automobilima.

Diplomske akademske – master studije na Elektrotehničkom fakultetu u Beogradu, na modulu za Signale i sisteme upisao je u oktobru 2018. godine, a završio u septembru 2019. godine sa prosečnom ocenom 9,83. Master rad pod nazivom “*Video praćenje automatski detektovanih objekata metodama dubokog učenja*” pod mentorstvom doc. dr Aleksandre Marjanović odbranio je sa ocenom 10.

Doktorske akademske studije upisao je u oktobru 2019. godine na Elektrotehničkom fakultetu u Beogradu na modulu Upravljanje sistemima i obrada signala. Oblast istraživanja Miloša Pavlovića obuhvata upravljanje sistemima, obradu signala, digitalnu obradu slike, kompjutersku viziju i veštačku inteligenciju. Autor je i koautor više od 20 radova objavljenih u domaćim i međunarodnim naučnim časopisima i na međunarodnim konferencijama.

Od decembra 2018. godine zaposlen je u Institutu Vlatacom, na poziciji sistem inženjera. Na ovoj poziciji radio je na razvoju i implementaciji algoritama kompjuterske vizije u okviru elektrooptičkih sistema. Bio je angažovan i na projektu Fonda za nauku Republike Srbije *AI-DECIDE*, gde se bavio distribuiranim multiagentnim sistemima.

## Изјава о ауторству

Име и презиме аутора Милош Павловић

Број индекса 5006/2019

### Изјављујем

да је докторска дисертација под насловом

Праћење покретних објеката у секвенци краткоталасних инфрацрвених слика  
корелационим методама и техникама робусне Калманове филтрације

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

### Потпис аутора

У Београду, 30.03.2024.

Милош Павловић

## Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора Милош Павловић

Број индекса 5006/2019

Студијски програм Електротехника и рачунарство  
(модул Управљање системима и обрада сигнала)

Наслов рада Праћење покретних објеката у секвенци краткоталасних инфрацрвених слика корелационим методама и техникама робусне Калманове филтрације

Ментор др Бранко Ковачевић, професор емеритус

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у **Дигиталном репозиторијуму Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

**Потпис аутора**

У Београду, 30.03.2024.

Милош Павловић

## Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Праћење покретних објеката у секвенци краткоталасних инфрацрвених слика  
корелационим методама и техникама робусне Калманове филтрације

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)
2. Ауторство – некомерцијално (CC BY-NC)
3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)
4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)
5. Ауторство – без прерада (CC BY-ND)
6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.  
Кратак опис лиценци је саставни део ове изјаве).

**Потпис аутора**

У Београду, 30.03.2024.

Милош Павловић

1. **Ауторство.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.