

Aktiv inferens för autonomt ihållande övervakning

JOHAN SCHUBERT, FARZAD KAMRANI, TOVE GUSTAVI



Johan Schubert, Farzad Kamrani, Tove Gustavi

Aktiv inferens för autonomt ihållande övervakning

Titel	Aktiv inferens för autonomt ihållande övervakning
Title	Active Inference for Autonomous Persistent Surveillance
Rapportnr/Report no	FOI-R--5965--SE
Månad/Month	Augusti
Utgivningsår/Year	2026
Antal sidor/Pages	45
ISSN	1650-1942
Uppdragsgivare/Client	Försvarets materielverk
Forskningsområde	Ledningsteknologi
FoT-område	Inget FoT-område
Projektnr/Project no	E380392
Godkänd av/Approved by	Emil Hjalmarson
Ansvarig avdelning	Cyberförsvar och ledningsteknik

Bild/Cover: Johan Schubert

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk, vilket bl.a. innebär att citering är tillåten i enlighet med vad som anges i 22 § i nämnd lag. För att använda verket på ett sätt som inte medges direkt av svensk lag krävs särskild överenskommelse.

This work is protected by the Swedish Act on Copyright in Literary and Artistic Works (1960:729). Citation is permitted in accordance with article 22 in said act. Any form of use that goes beyond what is permitted by Swedish copyright law, requires the written permission of FOI.

Sammanfattning

För att undersöka om den relativt nya forskningsinriktningen aktiv inferens lämpar sig för drönarstyrning utvecklar och testar vi en ruttplaneringsmetod baserad på aktiv inferens för autonom styrning av en intelligent agent. Agenten har till uppgift att upprätthålla en dynamisk lägesbild över ett givet område, vilket i det aktuella scenariot innebär att agenten kontinuerligt behöver rekognoscera området för att upptäcka och hålla reda på både fasta och rörliga målobjekt. Uppgiften kräver att agenten balanserar sök- och spårningsbeteenden samtidigt som den uppdaterar den dynamiska lägesbilden, representerad av en evidenskarta, med ny information. Agenten navigerar i en simulering genom att ta upprepade steg mot den punkt på evidenskartan som minimerar den så kallade variationella fria energin (VFE). I beräkningen av variationell fri energi ingår flera komponenter. Den generativa modellen för aktiv inferens använder Dempster-Shafer-teori och en gaussisk sensormodell för att ge agenten indata. Den generativa processen använder en bayesiansk metod för att uppdatera en a posteriori-sannolikhetsfördelning. I det aktuella scenariot beräknas den variationella fria energin för det område som för tillfället befinner sig inom sensors detektionsradie, som avvikelsen mellan en pignistik sannolikhetsfördelning per gridcell, baserad på evidenskartan, och en a posteriori-sannolikhetsfördelning för målobjektet, samt graden av överraskning vid mottagande av nya observationer. Den VFE-baserade metoden hanterar utmaningen att balansera sökning och målföljning och gör det möjligt för agenten att söka igenom stora områden samtidigt som den följer de identifierade målobjekten. Resultaten från 1000 simuleringar visar att den VFE-baserade styralgoritmen överträffar alternativa metoder för långvarig målobservation med i genomsnitt 51–69 % enligt olika mått. Detta sker dock till priset av en sämre förmåga att upptäcka alla målobjekt i området.

Nyckelord: aktiv inferens, frienergi-principen, Dempster-Shafer-teori, autonom agent, ruttplanering, ihållande övervakning.

Summary

We develop an active inference route-planning method for autonomous control of an intelligent agent. The aim is to reconnoiter a geographical area while balancing search and tracking behaviors to maintain a dynamic situational awareness map. To achieve this, we construct an evidence map that reflects our current understanding of the situation, integrating both positive and “negative” sensor observations of possible target objects collected over time, and diffusing the evidence across the map as time progresses. The generative model of active inference uses Dempster-Shafer theory and a Gaussian sensor model to provide input to the agent. The generative process employs a Bayesian approach to update a posterior probability distribution. We calculate the variational free energy across all locations within the sensor radius by assessing the divergence between a pignistic probability distribution of the evidence map and a posterior probability distribution of the target object, including the level of surprisal associated with receiving new observations. Using the variational free energy, we direct the agent’s movement in simulation by taking incremental steps toward a position that minimizes the variational free energy. This approach addresses the challenge of balancing exploration and exploitation, allowing the agent to search large areas while efficiently tracking identified targets. Across 1000 simulation trials, the approach surpasses benchmarks in target observation by an average of 51 %–69 % across various metrics for a 100-meter sensor radius.

Keywords: Active inference, free energy principle, Dempster-Shafer theory, autonomous agent, route-planning, persistent surveillance.

Innehållsförteckning

1	Inledning	7
1.1	Avgränsningar	7
1.2	Läsanvisningar	7
2	Problemformulering och metodansats	9
2.1	Problemformulering	9
2.2	En aktiv inferensmetod för ihållande övervakning	9
2.3	Notationer.....	10
3	Relaterat arbete	11
3.1	Ihållande övervakning och rörelsestyrning	11
3.2	Aktiv inferens för uppskattning och styrning	11
3.3	Aktiv inferens för navigering, utforskning och banditproblem	12
3.4	Aktiv inferens med flera agenter, organisatorisk anpassning och komplexa uppgifter	12
3.5	Vårt bidrag.....	12
4	Aktiv inferens och variationell fri energi	14
4.1	Den generativa modellen	14
4.2	Den generativa processen	14
4.3	Den variationella fria energin.....	15
5	Dempster-Shafer-teorin	16
6	Världsmodell och anpassning av sensordata.....	17
6.1	Världsmodell	17
6.2	Tilldelning av grundläggande tilltro från sensors utdata	18
7	Styrning av en intelligent agent med variationell fri energi	20
7.1	Den generativa modellen	20
7.2	Den generativa processen	23
7.3	Variationell fri energi.....	24
7.4	Val av handling.....	26
8	Implementation	27
8.1	Metod för styrning med VFE.....	27
8.2	Simuleringsmiljö	28
8.3	Grundscenario och parametrar	28
8.4	Beräkningskomplexitet	29
9	Experiment och resultat.....	30
9.1	Kvalitativ analys.....	30
9.2	Kvantitativ analys	31
9.2.1	Grundläggande tilltro i förhållande till verkliga målpositioner	32
9.2.2	Upptäckts- och observationstider.....	35
9.3	Diskussion om resultaten	38
10	Slutsatser	39

11	Framtida arbete	40
12	Referenser	41
	Bilaga A Sensor	44
	A.1 Sensoregenskaper	44
	A.2 Implementering av sensormodellen	44

1 Inledning

Aktiv inferens är en metod för autonomt beslutsfattande. Det kan också ses som ett generiskt ramverk, baserat på principen att ett system naturligt strävar efter att minimera den överraskning som uppstår när systemet tar emot ny information, t.ex. från sensorer. Ett grundläggande antagande i aktiv inferens är att det finns två alternativa handlingsätt för ett system som strävar efter att uppnå samstämmighet mellan verkligheten och den egna lägesbilden: antingen uppdateras systemets interna representation av lägesbilden när ny information tas emot, eller så vidtas åtgärder mot omgivningen för att förändra den, så att den överensstämmer med agentens uppfattning.

Denna rapport undersöker aktiv inferens [1–4] för autonom styrning av en intelligent agent, en obemannad flygfarkost, vars uppdrag är att skapa och bibehålla en korrekt lägesbild över tid. Studien syftar till att avgöra hur aktiv inferens kan användas för att autonomt styra en UAV-svärm och upprätthålla en lägesbild inom ett specifikt område. För att undersöka detta utför vi simuleringar med 1000 scenarier, där vi jämför aktiv inferens med två referensstrategier. Varje scenario inkluderar fyra fasta och tre rörliga målobjekt. En svärm av tio UAV:er söker och följer målen, medan sensorräckvidden varierar systematiskt. Vi genomför en omfattande kvantitativ analys med flera olika mått för att utvärdera de olika styrstrategierna.

Att minimera överraskningar är i sig omöjligt. Istället strävar systemet (i det här fallet agenten) efter att vidta de bästa möjliga åtgärderna genom att välja den åtgärd som minimerar *variationell fri energi* (VFE). VFE är ett informationsteoretiskt koncept som kvantifierar divergensen mellan två sannolikhetsfördelningar samt graden av överraskning när nya observationer anländer. I det aktuella scenariot representerar den första fördelningen agentens uppfattning av verkligheten och baseras på agentens *evidenskarta*, medan den andra utgörs av en a posteriori-sannolikhetsfördelning givet den aktuella observationen. När dessa två fördelningar stämmer överens är VFE låg. Vid aktiv inferens försöker agenten positionera sig så att dess observationer stämmer överens med dess modell av verkligheten.

Rapporten har tagits fram och arbetet har genomförts på uppdrag av FMV.

1.1 Avgränsningar

Notera att vi i den här rapporten modellerar aktiv inferens med hjälp av VFE snarare än *förväntad fri energi* (EFE), med planering över potentiella trajektorier (policyer). Vi utforskar VFE:s förmåga att utföra ihållande övervakning genom att använda girigt handlingsurval för att minimera ett lokalt VFE-fält inom sensors nuvarande räckvidd och därmed sänka beräkningskostnaderna jämfört med EFE.

1.2 Läsanvisningar

Kapitel 2 beskriver problemformuleringen. Kapitel 3 granskar relaterat arbete. Kapitel 4 introducerar grunderna i aktiv inferens och VFE. Kapitel 5 ger en kort översikt över Dempster-Shafer-teorin – ett ramverk för att hantera osäkerhet vid informationsfusion. Kapitel 6 beskriver världsmodellen och den förenklade sensormodellen som har använts i simuleringarna. I kapitel 7 beskrivs matematiskt hur man styr en agent via aktiv inferens genom att minimera VFE. Detta kapitel beskriver även (i) evidenskartan, d.v.s. den matematiska representationen av agentens lägesbild, (ii) modellen för hur informationsdiffusion påverkar informationen i evidenskartan över tid, samt (iii) metoden som använts för att uppdatera evidenskartan baserat på positiva och “negativa” sensorobservationer. Kapitel 8 ger en översikt över implementeringsdetaljerna, valet av vägpunkt (*eng.* waypoint) samt metodens låga beräkningskomplexitet. Kapitel 9 ger både en kvalitativ och en detaljerad kvantitativ analys av vår VFE-baserade styralgoritms egenskaper. Resultaten baseras på 1000 simulerade scenarier, där den VFE-baserade styrningen jämförs dels med

två styralgoritmer som baserats på var sin av de två delkomponenterna i VFE (för att undersöka hur de två komponenterna var för sig bidrar till egenskaperna vid VFE-baserad styrning), dels med två enkla grundstrategier för drönarstyrning som använts som referens. Kapitel 10 sammanfattar de viktigaste resultaten i detta arbete.

2 Problemformulering och metodansats

2.1 Problemformulering

Inom området autonoma system kallas uppgiften att kontinuerligt övervaka ett specifikt område under en obestämd period för ihållande övervakning. Verkliga tillämpningar av ihållande övervakningssystem inkluderar övervakning av områden runt kritisk infrastruktur och militära anläggningar för att upptäcka potentiella inkräktare med ont uppsåt. Det kan också användas för att övervaka populära stränder och förhindra drunkningstillbud samt för att övervaka vilda djur i känsliga naturområden för att stödja bevarandet. UAV:er är särskilt väl lämpade för denna typ av övervakning på grund av deras vidvinkelvyer, hastighet och den relativa avsaknaden av hinder i deras operativa miljöer, vilket underlättar banplaneringen för den autonoma agenten.

Denna artikel handlar om ett scenario där autonoma UAV:er kontinuerligt övervakar ett utvalt område för att upprätthålla en korrekt och uppdaterad dynamisk lägesförståelsekarta. UAV:erna har till uppgift att upptäcka både stationära och rörliga mål och helst spåra deras ungefärliga platser även när de är utanför sensorernas räckvidd. Rörliga mål måste ofta återbesökas för att förhindra att följningen förloras. Dessutom måste varje del av det utsedda området observeras regelbundet, utan att någon region lämnas oövervakad under längre perioder. Även om vi modellerar en svärm av UAV:er med hjälp av en fysisk rörelsemodell fungerar en UAV som ledare (eller agent) och ansvarar för alla beräkningar samt för valet av vägpunkter för att minimera VFE.

I scenarier med ihållande övervakning definieras problemet med banplanering för en autonom agent vanligtvis av en uppsättning specifika mål och begränsningar, var och en med sina egna prioriteringar. Olika metoder kan användas för att implementera den faktiska banplaneringen eller rörelsestyrningen (som diskuteras i kapitel 7). Denna artikel undersöker specifikt möjligheten att använda aktiv inferens med VFE för banplanering. Den primära forskningsfrågan som behandlas i denna artikel kan formuleras enligt följande:

Kan aktiv inferens med VFE tillämpas för att lösa problemet med rörelsestyrning för en svärm av UAV:er, där en ledare med en enda sensor (agenten) utför ihållande övervakning av flera mål över ett specificerat tvådimensionellt område?

I experimentet modellerar vi en svärm av tio UAV:er där endast ledaren är utrustad med en sensor, utför all bearbetning, konstruerar lägesförståelsekartan och väljer en gemensam vägpunkt för hela svärmen.

2.2 En aktiv inferensmetod för ihållande övervakning

För att effektivt implementera aktiv inferens för rörelsestyrning i ett ihållande övervakningssammanhang måste flera nyckelkomponenter specificeras.

En generativ modell representerar alla möjliga tillstånd och deras övergångar för att hantera dessa överväganden. Information om dessa tillstånd härleds från observationer över tid, vilket gör att vi kan uppdatera vår förståelse av det aktuella tillståndet. Vi modellerar osäkerhet med hjälp av Dempster-Shafer-teorin [5, 6]. I denna modell representeras tillståndet stokastiskt; varje möjlig övergång av målets grundläggande tilltro (*eng.* basic belief) från ett tillstånd (t.ex. position) till nästa specificeras av övergångssannolikheter. Tillstånden uppdateras dynamiskt vid varje tidssteg med hjälp av en diffusionsalgoritm. Denna modell indikerar den aktuella grundläggande tilltron till att minst ett målobjekt existerar i varje möjligt tillstånd (se kapitel 7.1).

En generativ process kompletterar den generativa modellen genom att hantera nya observationer från agentens sensor. Vi använder en bayesiansk metod i den här processen. Den

uppdaterar sannolikheterna för alla positioner inom agentens sensorradie. Denna uppdatering beräknar en ny sannolikhet för varje position baserat på den aktuella sensormodellen, samtidigt som den tar hänsyn till de befintliga sannolikheterna (se kapitel 7.2).

Eftersom vi inte kan minimera den överraskning som agenten upplever när den tar emot nya observationer direkt, minimerar vi istället VFE, som fungerar som en övre gräns för den överraskningen. VFE definieras som summan av divergensen mellan två sannolikhetsfördelningar samt graden av överraskning och fungerar som vår målfunktion för att styra agentens rörelser. För varje plats inom sensorradien jämför vi den pignistiska sannolikheten från den generativa modellen med sannolikheten som härleds från observationerna i syfte att minimera deras divergens (se kapitel 7.3).

Vid varje tidssteg beräknar vi VFE för alla platser inom sensorns räckvidd. Agenten tar sedan ett steg med en fast längd i den riktning som minimerar VFE. Denna metod gör det möjligt för agenten att agera autonomt och sträva efter att upprätthålla en korrekt lägesförståelsekarta enligt vår modell.

2.3 Notationer

Tabell 1. Variabler och definitioner.

Variabel	Värde	Definition
G	–	rutnät
c_{xy}	–	cell xy i G
Θ	$\{T, F\}$	utfallsrum
2^Θ	$\{\{T\}, \{F\}, \{T, F\}\}$	potensmängden av Θ
T	–	boolesk sann , indikerar närvarande mål
F	–	boolesk falsk , indikerar att inget mål finns
ϑ	$\{T, F\}$	tillstånd
ϕ	$\{T, F\}$	observation
p	$[0, 1]$	sannolikhet hos en observation
q	$[0, 1]$	sannolikhet för ett tillstånd
t	$\{0, 1, \dots, 500\}$ s	tidssteg
$q_{xy}^t(\vartheta)$	$[0, 1]$	sannolikhet för ϑ i c_{xy} vid t
$p_{xy}^t(\vartheta)$	$[0, 1]$	a priori-sannolikhet för ϑ i c_{xy} vid t
$p_{xy}^t(\phi)$	$[0, 1]$	sannolikhet för att erhålla en observation ϕ i c_{xy} at t
$p_{xy}^t(\phi \vartheta)$	$[0, 1]$	Likelihood, upptäcktssannolikhet för ϕ givet ϑ i c_{xy} vid t
$p_{xy}^t(\vartheta \phi)$	$[0, 1]$	a posteriorsannolikhet för ϑ givet ϕ i c_{xy} vid t
$n_{xy}^t(\vartheta)$	$[0, 1]$	grundläggande tilltro för ϑ i c_{xy} vid t
$m_{xy}^t(\phi)$	$[0, 1]$	grundläggande tilltro för ϕ in c_{xy} at t
F_{xy}	$[0, \infty)$	variationell fri energi i c_{xy}
D_{KL}	$[0, \infty)$	Kullback-Leibler-divergens
s_r	$\{10, 20, \dots, 100\}$ m	sensorradie
σ_T	0.7	max grundläggande tilltro för T vid positiv ϕ från sensorpositionen
σ_F	0.6	max grundläggande tilltro för F vid "negativ" ϕ från sensorpositionen
σ_{pos}	0.7	max grundläggande tilltro för T vid positiv ϕ från målposition
σ_T	$\{7.5, 15, \dots, 75\}$ m	sigma för T vid positiv ϕ från sensorposition
σ_F	$\{5, 10, \dots, 50\}$ m	sigma för F vid "negativ" ϕ från sensorposition
σ_{pos}	10 m	sigma för T vid positiv ϕ från målposition
d	$[0, 100]$ m	avstånd från sensorposition till målposition
N_{UAV}	10	antalet UAV:er
$N_{mål}$	7	antalet mål
N_{celler}	220 400	antalet celler i G
N_{sim}	1000	antalet simuleringar
$a(agent.c_{xy})$	–	aktion som flyttar agenten vid c_{xy}
ϕ	$[0, 1]$	utvärderingspoäng

3 Relaterat arbete

3.1 Ihållande övervakning och rörelsestyrning

Under årens lopp har olika metoder för att hantera styrningsproblemet vid ihållande övervakning föreslagits. Detta kapitel beskriver kortfattat några exempel.

Hari et al. [7] undersöker ett ihållande övervakningsuppdrag för en enda UAV som har till uppgift att upprepade gånger besöka n mål med samma prioritet. Eftersom målen är fasta, reduceras problemet till ett klassiskt handelsresandeproblem. Brown och Anderson [8] betraktar ett mer komplext maritimt övervakningsscenario som inkluderar dynamiska begränsningar för UAV:er och formulerar ett flermåloptimeringsproblem för att maximera informationsvinsten och minimera bränsleförbrukningen. Genomförbara och optimala lösningar erhålls genom att kombinera en bangenereringsmetod med en partikel-svårloptimeringsalgoritm. I likhet med problemet som diskuteras här introducerar Hübel et al. [9] en dynamisk *informationskarta* som är utsatt för informationsförfall och använder den för att härleda en gradientbaserad styrning som driver en grupp agenter att kontinuerligt uppdatera sin lägesförståelse genom att övervaka ett område. Dessutom föreslår författarna en tidsvarierande densitetsfunktion som kan integreras i styrningsalgoritmen för att modellera rörliga intresser. Ramasamy och Ghose [10] föreslår en metod för att öka sannolikheten att autonoma agenter observerar rörliga mål under övervakningsuppdrag. Om man antar att sannolikheten att observera ett mål är ojämnt fördelad över ett övervakat område, tilldelar Ramasamy och Ghose en *viktighetsgrad* till varje rutnätspunkt i en diskretiserad karta där UAV:n i scenariot har upptäckt ett mål. Vikten av en rutnätspunkt beror på antalet detektioner vid den punkten och ökar chansen att UAV:n besöker den platsen igen. I annat arbete som dokumenterats i litteraturen används förstärkningsinlärning för ihållande övervakningsstyrning. Chen et al. [11] använder en förstärkningsinlärningsmetod med flera agenter för att lära sig policyer för varje agent i ett team som har till uppgift att kontinuerligt övervaka en tvådimensionell miljö med stationära hinder. För att uppnå detta modelleras problemet så att ett straff tillämpas vid varje tidssteg om någon punkt i miljön lämnas utan övervakning. Mishra et al. [12] presenterar en annan metod baserad på förstärkningsinlärning för ihållande övervakning.

3.2 Aktiv inferens för uppskattning och styrning

Aktiv inferens kopplar samman perception och handling genom VFE. Teoretiska kopplingar till klassisk uppskattning och styrning visar att minimering av VFE ger mål som kombinerar graden av överraskning med styrningskostnader, vilket leder till ett linjärt-kvadratisk-gaussiskt beteende i linjärt-gaussiska miljöer [13]. I praktiken har aktiv inferens tillämpats för tillståndsuppskattning av en quadcopter med hjälp av dynamisk förväntansmaximering (DEM). DEM är ett perceptionsschema inspirerat av hjärnan, baserat på en datadriven modellinlärningsalgoritm [14]. Aktiv inferens har också använts för adaptiv manipulationsstyrning i industrirobotar, även när miljöns dynamik inte är explicit modellerad [15]. Aktiv inferens har tillämpats för adaptiv styrning av robotarmar med hjälp av multimodala perception-action-baserade och variationella autoencoder-baserade tillståndsrepresentationer. Denna metod kräver ingen dynamisk eller kinematisk modell av roboten [16]. Dessutom har aktiv inferens använts för att utveckla en momentregulator som integrerar obearbetade bilddata (*eng.* raw vision) och proprioception i en strömlinjeformad design för en Franka Emika Panda-robot med 7 frihetsgrader, kapabel att anpassa sig online till förändringar i dynamiken och mänskliga störningar [17]. Den har också använts för feltolerant styrning i närvaro av sensorfel, vilket ger en opartisk tillståndsuppskattning och förenklar handlings-specifikationen [18]. Dessa resultat stöder vår användning av VFE för sluten-loop-styrning med en *mjuk sensormodell* (*eng.* soft-output sensor model).

3.3 Aktiv inferens för navigering, utforskning och banditproblem

För mobila agenter möjliggör aktiv inferens inom hierarkiska generativa modeller mål-orienterad navigering med topologiskt konsistenta kartor och praktiska robotutplaceringar [19]. Modulära aktiva inferenssystem stöder flexibel måldriven navigering, hinder-undvikande och högkonfidensvägval, med stark nollskottsgeneralisering (*eng.* zero-shot generalization) till nya miljöer [20]. Nyfikenhetsdrivet lärande för robotuppgifter, baserat på principen om fri energi, använder bayesianska neurala nätverk för att representera epistemisk osäkerhet och modellera komplexa beteenden [21]. Dessutom har retrospektiv (resterande) överraskning (*eng.* residual surprise) introducerats som ett beräkningsmässigt element i aktiv inferens, vilket fungerar som en nedre gräns för EFE [22]. I beslutsfattande utökar kontextuella flerarmade banditer (CMAB) det klassiska banditproblemet genom att betinga handlingsvalet med observerad kontextuell information. Nyligen utvecklade modeller baserade på aktiv inferens för CMAB [23, 24] använder EFE för att vägleda kontextbaserat åtgärdsval och balansera utforskning och exploatering under osäkerhet – ett beteende som vi strävar efter att reproducera i vår metod genom att minimera ett rumsligt frienergifält över evidenskartan.

3.4 Aktiv inferens med flera agenter, organisatorisk anpassning och komplexa uppgifter

I multiagentmiljöer används aktiv inferens för att designa adaptiva organisationer där roller och strukturer förändras genom att minimera fri energi på teamnivå [25–27]. För komplexa robotuppgifter kan aktiv inferens kombineras med beteendeträd för att möjliggöra kontinuerlig planering och robust utförande med färre noder [28]. Samtidigt utforskas icke-modulära, kognitivt inspirerade arkitekturer för aktiv inferens för att förbättra robustheten mot okända indata [29]. Nyligen genomförda studier utvidgar aktiv inferens till autonom körning genom att införliva handlingsorienterade priors som kopplar samman perception och styrning, vilket leder till ett mer människolikt och kollisionsundvikande beteende [30]. I socialt interaktiva och multiagentscenarier används aktiv inferens för kinestetisk interaktion mellan människa och robot, där metapriors justerar följsamhet och motkrafter under fysisk kontakt [31]. Empatiska och socialt följsamma agenter anpassar sina rörelser efter omgivande individer och kontextuella normer [32], vilket demonstrerar ramverkets flexibilitet för fri energi i koordinations- och interaktionsuppgifter. Aktiv inferens används även för hierarkiska, förkroppsligade perceptions-handlings-loopar, där flera sensomotoriska vägar associerade med olika kroppsdelar (*eng.* body parts) kombineras dynamiskt, vilket gör det möjligt för roboten att omkonfigurera styrningen och endast aktivera de leder som behövs för en given uppgift [33]. Vår metod skiljer sig genom att vi operationaliserar VFE över en rumslig Dempster-Shafer-evidenskarta och en deterministisk mjuk sensormodell och sedan vägleder agenten att röra sig i den riktning som minimerar detta mål per cell vid varje steg.

3.5 Vårt bidrag

Till skillnad från övervakningsstyrning [7–12] introducerar vi en Dempster-Shafer-baserad evidenskarta med diffusion och en deterministisk mjuk sensormodell för att generera cell-för-cell-evidensvärde. Vi styr svärmen av UAV:er genom att minimera ett nätverksbaserat frienergimål som använder Kullback-Leibler-divergensen [34] för att jämföra en långsiktig tilltro, uttryckt som en pignistisk fördelning [35], med en momentan a posteriori-sannolikhet härledd från den senaste observationen – samt graden av överraskning – snarare än att enbart maximera kartans intensitet eller minimera återbesökstiden. Till skillnad från styrningsmetoder baserade på aktiv inferens och EFE-planeringsmetoder [15–33] erbjuder

vår metod en beräkningsmässigt effektiv minimering över ett VFE-fält som (i) separerar observations- och tillståndsevidens samt (ii) möjliggör en effektiv balans mellan sökning och spårning för ihållande övervakning.

Handlingsval vid aktiv inferens betraktas generellt som planering över policyer, såsom vägpunktsbanor, genom att minimera EFE. Detta arbete använder dock en beräknings-effektiv girig styrstrategi. Ett lokalt VFE-fält beräknas över potentiella vägpunkter för att välja nästa destination inom det aktuella sensorområdet. Systemet går sedan framåt med en fast steglängd mot punkten med lägsta VFE och upprepar processen. Denna metod ger en snabb, reaktiv styrenhet som är väl lämpad för kontinuerlig övervakning.

4 Aktiv inferens och variationell fri energi

Den generativa modellen och den generativa processen är stokastiska. Låt utfallsrummet vara mängden av alla möjliga tillstånd, betecknat med $\Theta = \{T, F\}$, där T betecknar närvaron av minst ett målobjekt och F representerar frånvaron av målobjekt. I detta kapitel använder vi de bayesianska sannolikheterna p (för observationen ϕ) och q (för tillståndet ϑ).

4.1 Den generativa modellen

Den generativa modellen räknar upp alla möjliga tillstånd och övergångar [1–3]. I detta sammanhang representeras tillstånd av rutnätsceller, där var och en motsvarar en möjlig position på en karta. Agenten och målobjekten kan växla mellan tillstånd och röra sig från vilken rutnätscell som helst till de närmaste angränsande cellerna.

Sannolikheten i den generativa modellen betecknas med q_{xy}^t , där ϑ är tillståndet och q_{xy}^t är sannolikheten för en rutnätscell c_{xy} vid tidpunkten t . I varje rutnätscell har vi

$$q_{xy}^t(\vartheta = T) + q_{xy}^t(\vartheta = F) = 1. \quad (1)$$

Vi bestämmer sannolikheterna q_{xy}^t för alla rutnätsceller c_{xy} i vår evidenskarta, baserat på alla observationer p_{xy}^t som samlats in under hela scenariot. Dessa observationer utgör vår förståelse av situationen. Kapitel 7 beskriver hur p_{xy}^t och q_{xy}^t beräknas vid spaning med en agent.

4.2 Den generativa processen

I den generativa processen beräknar vi den posteriora sannolikheten för målnärvaro givet observationerna. Vid varje tidssteg har vi en a priori-sannolikhet $p_{xy}^t(\vartheta)$. Denna sannolikhet är lika med den posteriorsannolikhet som beräknades vid föregående tidssteg.

Vi sätter

$$p_{xy}^t(\vartheta) = p_{xy}^{t-1}(\vartheta|\phi). \quad (2)$$

Dessutom initierar vi a priori-sannolikheten för tiden 0 enligt

$$p_{xy}^0(\vartheta = T) = \varepsilon. \quad (3)$$

Med hjälp av Bayes teorem kan vi beräkna den posteriora sannolikheten $p_{xy}^t(\vartheta|\phi)$ i rutnätscellen c_{xy} för det aktuella tidssteget t , baserat på a priori-sannolikheten.

Vi har

$$p_{xy}^t(\vartheta|\phi) = \frac{p_{xy}^t(\phi|\vartheta)}{p_{xy}^t(\phi)} p_{xy}^t(\vartheta), \quad (4)$$

där $p_{xy}^t(\phi|\vartheta)$ representerar sannolikheten, vilket indikerar detektionssannolikheten för en observation $\phi = T$, bestämd av sensormodellen, när vi har ett målobjekt $\vartheta = T$ i rutnätscellen c_{xy} vid tid t . I detta sammanhang representerar $p_{xy}^t(\phi)$ sannolikheten för att erhålla en observation. Sannolikheten kan beräknas enligt

$$p_{xy}^t(\phi) = p_{xy}^t(\phi|\vartheta = T) \cdot p_{xy}^t(\vartheta = T) + p_{xy}^t(\phi|\vartheta = F) \cdot p_{xy}^t(\vartheta = F), \quad (5)$$

där $p_{xy}^t(\vartheta = T)$ betecknar den a priori-sannolikhet som diskuterats tidigare, och

$$p_{xy}^t(\vartheta = F) = 1 - p_{xy}^t(\vartheta = T). \quad (6)$$

Dessutom representerar $p_{xy}^t(\varphi|\vartheta = T)$ detektionssannolikheten för sensormodellen, med

$$p_{xy}^t(\varphi|\vartheta = F) = 1 - p_{xy}^t(\varphi|\vartheta = T). \quad (7)$$

4.3 Den variationella fria energin

Baserat på den generativa modellens $q_{xy}^t(\vartheta)$ och den generativa processens $p_{xy}^t(\vartheta|\varphi)$ kan vi nu beräkna VFE F_{xy}^t för alla rutnätsceller c_{xy} vid tidpunkten t .

Vi har

$$\begin{aligned} F_{xy}^t &= D_{KL}[q_{xy}^t(\vartheta) \parallel p_{xy}^t(\vartheta|\varphi)] - \ln[p_{xy}^t(\varphi)] \\ &= \left[\sum_{\vartheta \in \{T, F\}} q_{xy}^t(\vartheta) \cdot \ln \left(\frac{q_{xy}^t(\vartheta)}{p_{xy}^t(\vartheta|\varphi)} \right) \right] - \ln[p_{xy}^t(\varphi)], \end{aligned} \quad (8)$$

där D_{KL} betecknar Kullback-Leibler-divergensen [34]. Termen $-\ln[p_{xy}^t(\varphi)]$ representerar graden av överraskning från en observation, $q_{xy}^t(\vartheta)$ är sannolikheten enligt den generativa modellen, och $p_{xy}^t(\vartheta|\varphi)$ är den posteriora sannolikheten enligt den generativa processen, beräknad vid varje tidpunkt baserat på Bayes sats. Slutligen är $p_{xy}^t(\varphi)$ sannolikheten för en observation som härleds i ekvation (5).

Vid aktiv inferens väljs den åtgärd som minimerar F_{xy}^t . Denna åtgärd kan innebära att man rör sig till c_{xy} , vilket minimerar F_{xy}^t eller i riktning mot c_{xy} .

5 Dempster-Shafer-teorin

I Dempster-Shafer-teorin [5, 6] tilldelas en proposition en tilltro genom en grundläggande tilltrostilldelning, betecknad med m . Propositionen representeras av en delmängd A av en uttömmande mängd ömsesidigt uteslutande möjligheter, kallad ett utfallsrum (*eng.* frame of discernment), betecknat med Θ .

Den grundläggande tilltrofunktionen m definieras som en funktion från potensmängden av Θ till intervallet $[0, 1]$,

$$m: 2^\Theta \rightarrow [0,1], \quad (9)$$

där $m(\emptyset) = 0$ och

$$\sum_{A \subseteq \Theta} m(A) = 1. \quad (10)$$

Här representerar $m(A)$ den grundläggande tilltron som har tilldelats A .

Om vi får ytterligare information om samma hypotes från en annan källa, kombinerar vi de två grundläggande tilltrofunktionerna för att bilda en mer omfattande förståelse. Detta uppnås genom att beräkna den ortogonala kombinationen med hjälp av Dempsters regel. Låt B vara en delmängd av m_1 och C vara en delmängd av m_2 . Kombinationen av m_1 och m_2 resulterar i en ny grundläggande tilltrofunktion $m_1 \oplus m_2$, där

$$m_1 \oplus m_2(A) = K \cdot \sum_{B \cap C = A} m_1(B) \cdot m_2(C), \quad (11)$$

där K är en normaliseringskonstant.

6 Världsmodell och anpassning av sensordata

I detta kapitel beskrivs hur olika objekt och storheter i UAV-svärmens simuleringsvärld modelleras och representeras. Kapitlet ligger till grund för kapitel 7, där den VFE-baserade styralgoritmen som används av UAV-svärmens ledare härleds.

6.1 Världsmodell

I vårt scenario modelleras världen som en tvådimensionell karta över ett specificerat övervakningsområde. Kartan, det vill säga vår världsmodell (*eng.* world model), är uppdelad i ett rutnät av celler som indexeras med hjälp av ett koordinatsystem. En position i kartan betecknas med ett cellindex i formatet c_{xy} , där x och y är heltal. Positionen för agenten (svärmens ledare) eller för ett målobjekt anges alltid med ett (enskilt) cellindex, eftersom varje cell antas vara tillräckligt stor för att rymma en UAV eller ett målobjekt i sin helhet. Det finns inte heller några hinder i världsmodellen som hindrar flera objekt från att finnas i samma cell. Användningen av en tvådimensionell (istället för en tredimensionell) karta är en medveten förenkling, motiverad av att UAV:erna i experimenten befinner sig på konstant höjd och att alla händelser av intresse i det scenario som undersöks därmed utspelar sig i xy -planet.

Ett målobjekt i scenariot kan antingen representera ett stationärt objekt av intresse eller ett rörligt objekt, såsom en människa eller ett markgående fordon. För enkelhetens skull används i experimenten endast två fördefinierade rörelsemönster för målobjektet: stillastående eller rätlinjig rörelse med konstant hastighet. Hastigheterna som används för de rörliga målobjekten i experimenten är valda för att stämma överens med hur snabbt en människa rör sig i terrängen.

Agentens lägesbild representeras av en dynamisk evidenskarta G (baserad på världsmodellen), strukturerad som ett rutnät av *celler*, där varje cell är associerad med en grundläggande tilltro till målnärvaron i enlighet med Dempster-Shafer-teorin (kapitel 5). Varje cell motsvarar ett specifikt segment av det övervakade geografiska området eftersom ett rutnät läggs över kartan. Således representerar varje cell en distinkt del av övervakningszonen. Mer specifikt är varje cell c_{xy} associerad med två evidensbaserade värden [35]: $n_{xy}^t(\vartheta = T) \in [0,1]$, vilket representerar den grundläggande tilltron vid tidpunkten t att minst ett målobjekt finns i cellen c_{xy} , $n_{xy}^t(\vartheta = F) \in [0,1]$, vilket representerar den grundläggande tilltron att inget målobjekt finns i cellen c_{xy} vid tidpunkten t . Tilltro tilldelas enskilda celler och utgör cellens relativa *viktighetsgrad*, men detta ska inte tolkas som att de tillsammans utgör en täthetsfunktion (med tilltro per ytenhet) som summerar till 1.

Observera att den grundläggande tilltron i varje cell c_{xy} är en egen fördelning med ett eget utfallsrum. Detta designval, att använda ett rutnät med ett utfallsrum per rutnätscell och inte ett större utfallsrum över alla rutnätsceller där tilltron kan tilldelas olika delmängder av utfallsrummet, görs för att undvika den kombinatoriska beräkningsexplosion som uppstår vid hantering av alla snitt mellan delmängder som uppstår när diffusion av olika observationer börjar överlappa. Formlerna för att uppdatera evidenskartan till följd av informationsdegradering respektive nya observationer finns i kapitel 7.

Agentens sensor antas vara en intelligent bildsensor, utformad för markspaning. När agenten är aktiv tar sensorn, med förutbestämda tidsintervall, flygbilder över marken under sig och använder programvara för bildanalys i realtid för att bearbeta dem. I det implementerade scenariot används en förenklad sensormodell, som använder information om agentens och målobjektens aktuella positioner på kartan, samt vissa antaganden om sensoregenskaper, för att generera *sensordata* i form av:

- En vektor med uppskattningar av aktuella positioner (cellindex i kartan) för detekterade målobjekt (att ett objekt detekterats som ett målobjekt innebär att det

av bildbehandlingsprogramvaran först har *identifierats* som ett objekt i bilden och sedan specifikt klassificerats som ett fördefinierat målobjekt).

- En vektor som för varje detekterat målobjekt anger bildanalysprogramvarans skattade konfidens för att objektet – som alltså har klassificerats som ett målobjekt – verkligen är ett målobjekt.

Detaljer om sensormodellen som använts, de förenklingar som införts i modellen samt om själva implementationen som använts i experimenten återfinns i bilaga A. Förenklingar i modelleringen och implementationen som särskilt bör noteras är:

- Alla målobjekt som hamnar inom sensorns räckvidd detekteras med sannolikhet 1.
- Sensorn genererar inga falska målobjektdetektioner.
- Positionsuppskattningarna för målobjekten, genererade av den implementerade sensormodellen, är brusfria.
- Undviker Monte Carlo-sampling genom deterministiska mjuka observationer.

I kapitel 6.2 beskrivs hur utdata från sensorn, i kombination med antaganden om en verklig bildsensorfunktion, används för att tilldela grundläggande tilltrosvärden för *Mål* och *Inget Mål* till enskilda celler i evidenskartan.

6.2 Tilldelning av grundläggande tilltro från sensorns utdata

I det här kapitlet beskrivs hur sensorns observationer – i kombination med rimliga antaganden – används för att tilldela grundläggande tilltrosvärden för *Mål* och *Inget Mål* till enskilda celler i kartan över simuleringsområdet. Värdena används i nästa steg för att uppdatera evidenskartan – alltså agentens lägesbild – med ny information, vilket beskrivs i kapitel 7.

Enligt världsmodellen (kapitel 6.1) representerar evidenskartan (lägesbilden) en över tid ackumulerad grundläggande tilltro till *närvaro*, $n_{xy}^t(\vartheta = T)$, respektive *icke-närvaro*, $n_{xy}^t(\vartheta = F)$, av målobjekt. För att skilja denna ackumulerande grundläggande tilltro från en grundläggande tilltro som baseras endast på nyligen gjorda observationer inför vi beteckningarna $m_{xy}^t(\varphi = T)$ och $m_{xy}^t(\varphi = F)$ för de senare.

För att förklara hur värdena på $m_{xy}^t(\varphi = T)$ och $m_{xy}^t(\varphi = F)$ beräknas, börjar vi med att betrakta specialfallet där exakt ett målobjekt ($k = 1$) har detekterats av sensorn vid tidpunkten t . Utdata från sensorn består i detta fall av de uppskattade positionskoordinaterna för målobjektet ifråga ($x_{k=1}, y_{k=1}$), samt den tillhörande klassificeringskonfidensen $C_{T,k=1}$. Således tilldelar vi i cell c_{x_1,y_1} den grundläggande tilltron $m_{x_1,y_1}^t(\varphi = T) = C_{T,1}$. Eftersom vi vet att positionsuppskattningen är oprecis väljer vi också att tilldela positiva (≥ 0) grundläggande tilltrosvärden för *Mål* i cellerna som omger c_{x_1,y_1} . Vi hänvisar till dessa grundläggande tilltrosvärden, ett per cell, som observationer (för att vara konsekventa med standardnomenklaturen för fri energi [3]). De grundläggande tilltrosvärdena i cellerna runt cell c_{x_1,y_1} är inställda för att matcha en tvådimensionell gaussisk funktion centrerad i c_{x_1,y_1} med standardavvikelse σ_{pos} (given som information associerad med sensorn) och maximalt värde $C_{T,1}$ vid c_{x_1,y_1} . Tilldelningen av grundläggande tilltrosvärden associerade med detekteringen är begränsad till cellerna på ett avstånd $\leq 2\sigma_{pos}$ från c_{x_1,y_1} . Alla celler utanför $2\sigma_{pos}$ -radien, såväl som alla celler utanför sensorns räckvidd, tilldelas ett grundläggande tilltrosvärde om $m_{x_1,y_1}^t(\varphi = T) = 0$. Om det finns flera måldetekteringar samtidigt definieras separata grundläggande tilltrofunktioner $m_{xy,k}^t(\varphi = T)$ för $k = 1, \dots, K$, motsvarande fallet med $k = 1$ som beskrivs ovan. Med $d_{xy,k}$ som betecknar avståndet mellan cell c_{xy} och det detekterade målet k , och $d_{xy,s}$ som betecknar avståndet mellan cell c_{xy} och sensorn, får vi:

$$m_{xy,k}^t(\varphi = T) = \begin{cases} C_{T,k} \cdot e^{-\frac{d_{xy,k}^2}{2\sigma_{pos}^2}}, & (xy)|d_{xy,s} \leq s_r, d_{xy,k} \leq 2\sigma_{pos}^2 \\ 0, & \text{för övrigt} \end{cases} \quad (12)$$

De grundläggande tilltrofunktionerna $m_{xy,k}^t(\varphi = T)$ kombineras med Dempsters regel [5]. Vi har,

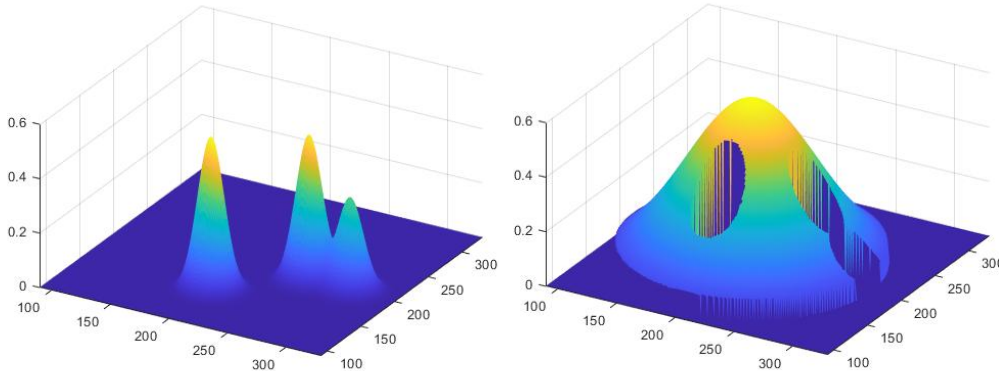
$$m_{xy}^t(\varphi = T) = \bigoplus_{k=1}^K m_{xy,k}^t(\varphi = T), \quad (13)$$

där $\oplus$ är Dempsters regel.

Den nuvarande sensormodellen ger inte explicita bevis för *Inget Mål*, men avsaknaden av målobservationer i en cell inom sensorns räckvidd kan betraktas som indirekta bevis på att cellen saknar mål. Vi modellerar detta genom att tilldela positiva (≥ 0) grundläggande tilltrosvärden för *Inget Mål*, $m_{xy}^t(\varphi = F)$, till alla celler som befinner sig inom sensorns räckvidd *men inte inom $2\sigma_{pos}$ -radien för någon positiv måldetektering*. Dessa grundläggande tilltrosvärden modelleras av en gaussisk funktion, centrerad vid sensorns position, med en lämpligt vald standardavvikelse σ_F och ett maximalt värde C_{Fmax} (se figur 1).

Med $d_{xy,k}$ och $d_{xy,s}$ definierade som tidigare får vi:

$$m_{xy}^t(\varphi = F) = \begin{cases} C_{Fmax} \cdot e^{-\frac{d_{xy,k}^2}{2\sigma_{pos}^2}}, & (xy)|d_{xy,s} \leq s_r, d_{xy,k} \geq 2\sigma_{pos}^2 \forall k \\ 0, & \text{för övrigt} \end{cases} \quad (14)$$



Figur 1. Kombinerade grundläggande tilltrostilldelningar $m_{xy,k}^t(\varphi = T)$ för *Mål* associerade med tre observationer (till vänster), och grundläggande tilltrostilldelningar $m_{xy}^t(\varphi = F)$ för *Inget Mål* (till höger). För positioner inom sensorns radie, $m_{xy}^t(\varphi = F)$, förutom de nära de tre målen.

7 Styrning av en intelligent agent med variationell fri energi

Detta kapitel beskriver vår implementering av VFE, där den generativa modellen representeras som en Dempster-Shafer-evidenskarta med grundläggande tilltro och dess pignistiska transform producerar sannolikhetsfördelningen som används vid VFE-minimering.

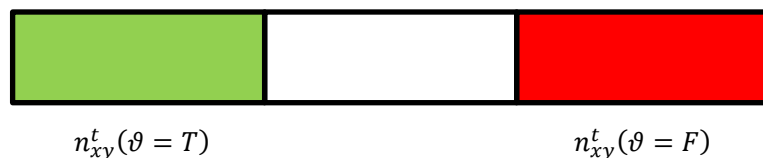
I kapitel 7.1 förklarar vi den generativa modellen, inklusive dess representation och alla formler som behövs för att uppdatera lägesförståelsekartan. Detta innebär att sensor-observationer sammanfogas i kartan och att kartans osäkerhet sprider sig till angränsande celler över tid. I kapitel 7.2 beskriver vi den generativa processen och förklarar hur den posteriora fördelningen beräknas med hjälp av Bayes sats, utifrån a priori-fördelningen samt sensor- och observationsmodellen. Kapitel 7.3 integrerar sedan dessa komponenter och illustrerar hur man kombinerar de pignistiska och posteriora fördelningarna vid beräkningen av VFE. Slutligen beskriver kapitel 7.4 den handlingsregel som agenten använder för att tilldela varje medlem i UAV-svärmen en ny gemensam vägpunkt.

7.1 Den generativa modellen

Detta arbete använder evidenskartan G som sin världsmodell. I den generativa modellen representerar vi osäkerhet med hjälp av en grundläggande tilltrofunktion, betecknad med n (för tillstånd), enligt definitionen i kapitel 6.1.

Följande villkor gäller för de grundläggande tilltrofunktionerna: $n_{xy}^t(\vartheta = T), n_{xy}^t(\vartheta = F) \geq 0$, och $n_{xy}^t(\vartheta = T) + n_{xy}^t(\vartheta = F) \leq 1$ för alla celler c_{xy} vid alla tidpunkter t . Paret $(n_{xy}^t(\vartheta = T), n_{xy}^t(\vartheta = F))$ representerar den grundläggande tilltron för cell c_{xy} vid tidpunkten t , med osäkerhet definierad som $1 - n_{xy}^t(\vartheta = T) - n_{xy}^t(\vartheta = F)$.

I figur 2 representerar det gröna området $n_{xy}^t(\vartheta = T)$; det röda området $n_{xy}^t(\vartheta = F)$; och det vita området osäkerheten.



Figur 2. Illustration av evidens och osäkerhet i cell c_{xy} . Grönt indikerar $n_{xy}^t(\vartheta = T)$, rött indikerar $n_{xy}^t(\vartheta = F)$ och vitt visar kvarvarande osäkerhet $1 - n_{xy}^t(\vartheta = T) - n_{xy}^t(\vartheta = F)$. Detta Dempster-Shafer-par per cell utgör grunden för evidenskartan G .

Osäkerheten $1 - n_{xy}^t(\vartheta = T) - n_{xy}^t(\vartheta = F)$ kan betraktas som en genuin osäkerhet i cell c_{xy} , som agenten bör sträva efter att minimera över alla celler c_{xy} hela tiden för att upprätthålla en aktuell lägesförståelsekarta.

Istället för att tilldela en enda evidens för en positiv detektion till den *mängd* av rutnätsceller som ligger inom den antagna sensordetekteringsradien, tilldelar vi evidens till varje enskild rutnätscell inom den radien. Denna metod erbjuder en mer detaljerad representation, eftersom varje rutnätscell c_{xy} tilldelas ett värde, samtidigt som den avsevärt minskar beräkningskomplexiteten för att uppdatera dessa värden som svar på nya sensormätningar.

Vår representation använder en gaussisk fördelning med en standardavvikelse lika med sensorns radie, vilket säkerställer att mittcellen får det högsta värdet. Denna metod ger en uppsättning distinkta grundläggande tilltrofördelningar – en för varje rutnätscell. Resultatet ger en tillfredsställande upplösning och en relevant spridning i den evidensbaserade lägesförståelsekartan. Diffusionsmodellen som beskrivs i ekvationerna (15) och (16)

representerar osäkerhet i målrörelse och producerar en dynamisk, evidensbaserad lägesförståelsekarta.

Eftersom den observerade miljön är dynamisk, minskar värdet av äldre information med tiden. I evidenskartan visas denna minskning av säkerheten genom spridningen av grundläggande tilltro till angränsande celler. Ekvation (15) beskriver denna utåtriktade spridning av närvaroevidens som en slumpmässig process. Omvänt begränsar ekvation (16) den inåtriktade spridningen av brist på evidens, vilket innebär att ett område anses vara tomt om dess omgivande celler är tomma.

Oavsett om agenten observerade tillståndet för cellen c_{xy} mellan tidpunkterna t och $t + 1$, uppdateras de grundläggande tilltrofördelningarna $n_{xy}^t(\vartheta = T)$ och $n_{xy}^t(\vartheta = F)$ i cellen enligt ekvationerna (15) och (16). Denna process är fristående från sammanslagningen med nya observationer, vilket hanteras i ekvationerna (17) och (18). Uppdateringen för den grundläggande tilltron ges av:

$$n_{xy}^{t+1}(\vartheta = T) = 1 - \left(1 - \frac{n_{xy}^t(\vartheta = T)}{N + 1}\right) \cdot \prod_{(ij) \in \{\text{grannar till } c_{xy}\}} \left(1 - \frac{n_{ij}^t(\vartheta = T)}{N + 1}\right) \quad (15)$$

där N är antalet grannar till cellen c_{xy} . Detta innebär att den grundläggande tilltron $n_{xy}^t(\vartheta = T)$ för cell c_{xy} , delas mellan cellen, dess grannar och sig själv (dvs. mellan nio rutnätsceller i fallet med ett kvadratisk rutnät), och sedan kombineras med hjälp av Dempsters regel. Detta orsakar en diffusion av den grundläggande tilltron $n_{xy}^t(\vartheta = T)$, över ett allt större område över tid (på ett sätt som liknar en explosion).

Antag att $n_{xy}^t(\vartheta = T) > 0$ i cell c_{xy} . Då kommer en del av den grundläggande tilltron $n_{xy}^t(\vartheta = T)$ att spridas till sina N grannar. Omvänt, om $n_{ij}^t(\vartheta = T) > 0$ i intilliggande celler, sprids en del av den grundläggande tilltron till cell c_{xy} .

Vidare har vi

$$n_{xy}^{t+1}(\vartheta = F) = n_{xy}^t\left(\frac{1}{N+1}\right)(\vartheta = F) \cdot \prod_{(ij) \in \{\text{grannar till } c_{xy}\}} n_{ij}^t\left(\frac{1}{N+1}\right)(\vartheta = F) \quad (16)$$

där $n_{xy}^{t+1}(\vartheta = F)$ indikerar att cell c_{xy} är tomt vid tidpunkten $t + 1$. För att cell c_{xy} ska vara tomt vid nästa tidssteg (dvs. $n_{xy}^{t+1}(\vartheta = F)$), så måste alla angränsande celler c_{xy} och c_{xy} självt vara tomma vid tiden t , vilket betyder $n_{xy}^t(\vartheta = F) = 1$ för alla c_{xy} och alla grannar till c_{xy} . Ekvation (16) kontraherar den grundläggande tilltron $n_{xy}^t(\vartheta = F)$ inåt (på ett sätt som liknar en implosion). Våra empiriska resultat tyder på att detta leder till en stabil osäkerhetsdynamik.

Notera att formlerna är utformade för att vara oberoende av rutnätets form. Att byta till en hexagonal topologi från den kvadratiske som används i denna implementation kräver bara att parametern N , som anger antalet närliggande celler i ekvationerna (15) och (16) (se kapitel 7.1), justeras – från 8 till 6.

Sensormätningar uppdaterar kontinuerligt evidenskartan enligt ekvationerna (17) och (18) nedan. Om vi får positiva observationer kombineras den grundläggande tilltron $m_{xy}^{t+1}(\varphi = T)$ som tilldelats av ekvation (13), med det befintliga grundläggande tilltrosparet $(n_{xy}^{t+1}(\vartheta = T), n_{xy}^{t+1}(\vartheta = F))$, härlett av ekvationerna (15) och (16) med hjälp av Dempsters regel [5]. Observera att $m_{xy}^{t+1}(\varphi = T) = 0$ för c_{xy} utanför $2\sigma_{pos}$ av alla målpositioner.

Vi har

$$n_{xy}^{t+1*}(\vartheta = T) = \frac{n_{xy}^{t+1}(\vartheta = T) + m_{xy}^{t+1}(\varphi = T) \cdot (1 - n_{xy}^{t+1}(\vartheta = T) - n_{xy}^{t+1}(\vartheta = F))}{1 - m_{xy}^{t+1}(\varphi = T) \cdot n_{xy}^{t+1}(\vartheta = F)} \quad (17)$$

och

$$n_{xy}^{t+1*}(\vartheta = F) = \frac{(1 - m_{xy}^{t+1}(\varphi = T)) \cdot n_{xy}^{t+1}(\vartheta = F)}{1 - m_{xy}^{t+1}(\varphi = T) \cdot n_{xy}^{t+1}(\vartheta = F)}. \quad (18)$$

Det antas att ekvationerna (17) och (18) alltid tillämpas efter ekvationerna (15) och (16) vid tidpunkten $t + 1$. De grundläggande tilltrofunktionerna $n_{xy}^{t+1}(\vartheta = T)$ och $n_{xy}^{t+1}(\vartheta = F)$ i högerleden av ekvationerna (17) och (18) representerar resultaten från diffusionsuppdateringen vid tidpunkten $t + 1$. Däremot är värdena $n_{xy}^{t+1*}(\vartheta = T)$ och $n_{xy}^{t+1*}(\vartheta = F)$ på vänster sida, markerade med en asterisk (*), de kombinerade resultaten av $n_{xy}^{t+1}(\vartheta = T)$ och $n_{xy}^{t+1}(\vartheta = F)$, fortfarande vid samma tidpunkt $t + 1$.

Vi får "negativa" observationer i rutnätsceller c_{xy} inom sensors radie men utanför $2\sigma_{pos}$ från alla målpositioner, och tilldelar grundläggande tilltro $m_{xy}^{t+1}(\varphi = F)$ som kombineras med $(n_{xy}^{t+1}(\vartheta = T), n_{xy}^{t+1}(\vartheta = F))$ med hjälp av Dempsters regel.

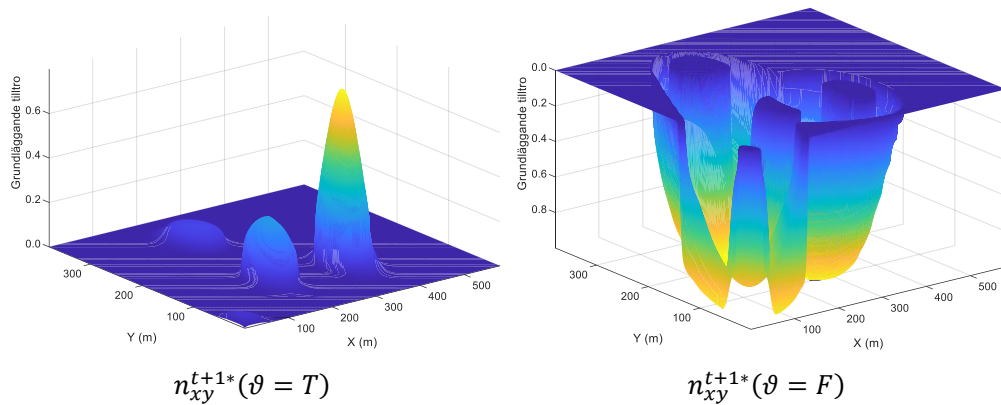
Vi har

$$n_{xy}^{t+1*}(\vartheta = T) = \frac{n_{xy}^{t+1}(\vartheta = T) \cdot (1 - m_{xy}^{t+1}(\varphi = F))}{1 - n_{xy}^{t+1}(\vartheta = T) \cdot m_{xy}^{t+1}(\varphi = F)} \quad (19)$$

och

$$n_{xy}^{t+1*}(\vartheta = F) = \frac{n_{xy}^{t+1}(\vartheta = F) + m_{xy}^{t+1}(\varphi = F) \cdot (1 - n_{xy}^{t+1}(\vartheta = T) - n_{xy}^{t+1}(\vartheta = F))}{1 - n_{xy}^{t+1}(\vartheta = T) \cdot m_{xy}^{t+1}(\varphi = F)}. \quad (20)$$

Se figur 3 för ögonblicksbilder av evidensfördelningarna $n_{xy}^{t+1*}(\vartheta = T)$ och $n_{xy}^{t+1*}(\vartheta = F)$ efter beräkning av ekvationerna (17)–(20).



Figur 3. Figuren illustrerar evidensfördelningen $n_{xy}^{t+1*}(\vartheta = T)$ i varje cell c_{xy} över ytan omedelbart efter att ett fjärde målobjekt detekterats i ett enkelt markscenario (till vänster). Vid tidsteget $t + 1$ visas evidensfördelningen $n_{xy}^{t+1*}(\vartheta = F)$ (till höger). Som illustration visas fördelningen $n_{xy}^{t+1*}(\vartheta = F)$ upp och ner i figuren.

7.2 Den generativa processen

I den generativa processen använder vi en bayesiansk metod med sannolikheter betecknade med p . När vi får en ny observation är vårt mål att beräkna den posteriora fördelningen, $p_{xy}^{t+1}(\vartheta|\varphi)$. Vi måste bestämma denna sannolikhetsfördelning för alla rutnätsceller, c_{xy} , för att uppnå detta. Detta kräver en sensormodell i form av en sannolikhetsfördelning $p_{xy}^{t+1}(\varphi|\vartheta)$ och en observationsmodell $p_{xy}^{t+1}(\varphi)$, samt en a priori-fördelning $p_{xy}^{t+1}(\vartheta)$.

Vi börjar med a priori-fördelningen. Det är viktigt att notera att a priori-fördelningen vid varje tidpunkt är densamma som den posteriora fördelningen vid föregående tidpunkt.

Vi har

$$p_{xy}^{t+1}(\vartheta) = p_{xy}^t(\vartheta|\varphi). \quad (21)$$

Inledningsvis har vi satt

$$p_{xy}^0(\vartheta = T) = \varepsilon = \frac{N_{mål}}{N_{celler}}, \quad (22)$$

där $N_{mål}$ är det antagna antalet mål i det undersökta området och N_{celler} är antalet celler i området.

Med hjälp av en sannolikhetsfördelning och vår sensormodell tilldelar vi sannolikheter till varje rutnätscell c_{xy} inom sensors radie, baserat på två gaussiska fördelningar.

För att beräkna sannolikheten för en observation, betecknad som $p_{xy}^t(\varphi)$, kan vi uttrycka den enligt följande:

$$p_{xy}^{t+1}(\varphi) = p_{xy}^{t+1}(\varphi|\vartheta = T) \cdot p_{xy}^{t+1}(\vartheta = T) + p_{xy}^{t+1}(\varphi|\vartheta = F) \cdot p_{xy}^{t+1}(\vartheta = F), \quad (23)$$

där $p_{xy}^{t+1}(\vartheta = T)$ är den a priori-sannolikhet som härletts i ekvation (21) och

$$p_{xy}^{t+1}(\vartheta = F) = 1 - p_{xy}^{t+1}(\vartheta = T). \quad (24)$$

Dessutom beräknas sannolikhetsfördelningen $p_{xy}^t(\varphi|\vartheta = T)$ med hjälp av en gaussisk fördelning, liknande den metod som beskrivs i bilaga A och ekvation (35).

Eftersom sensorn returnerar ett deterministiskt värde definierar vi sannolikheten direkt utifrån det. Därför är identifieringen i ekvation (25) bara en förkortning snarare än en slutsats om tillståndet,

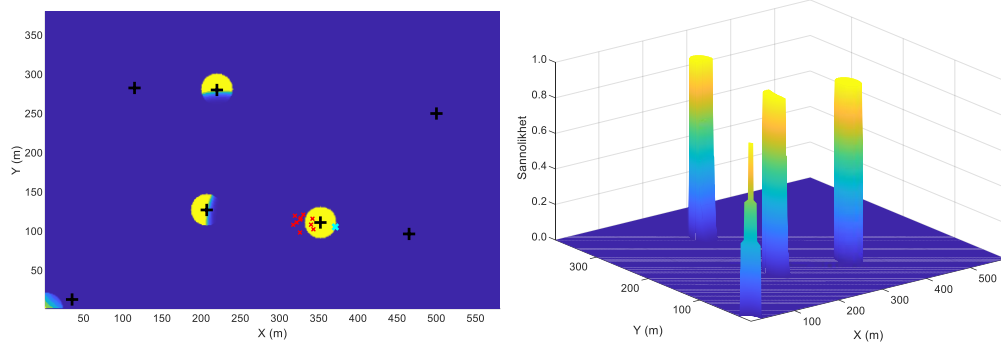
$$\begin{aligned} p_{xy}^{t+1}(\varphi|\vartheta = T) &\triangleq n_{xy}^{t+1}(\vartheta = T), \\ p_{xy}^{t+1}(\varphi|\vartheta = F) &= 1 - p_{xy}^{t+1}(\varphi|\vartheta = T). \end{aligned} \quad (25)$$

Med hjälp av de tre sannolikhetsfördelningarna (a priori $p_{xy}^{t+1}(\vartheta)$, likelihood $p_{xy}^{t+1}(\varphi|\vartheta)$ och observation $p_{xy}^{t+1}(\varphi)$) kan vi beräkna den önskade posteriora fördelningen med hjälp av Bayes teorem.

Vi har

$$p_{xy}^{t+1}(\vartheta|\varphi) = \frac{p_{xy}^{t+1}(\varphi|\vartheta)}{p_{xy}^{t+1}(\varphi)} p_{xy}^{t+1}(\vartheta). \quad (26)$$

Figur 4 visar den posteriora fördelningen i jämförelse med figur 3.



Figur 4. Den posteriora fördelningen $p_{xy}^{t+1}(\vartheta|\varphi)$, beräknad efter den senaste observationen med hjälp av Bayes regel, ekvation (26), a priori, ekvationerna (21)–(24), och gaussisk sensorlikhet, ekvation (25), vid samma tidssteg som i figur 3, för jämförelse med den sammanslagna föreställningen som visas där. Vänster och höger visar samma fördelningar: 2D till vänster och 3D till höger.

7.3 Variationell fri energi

Den grundläggande tilltrofördelningen $n_{xy}^{t+1*}(\vartheta = T)$ representerar fördelningen från den generativa modellen som vi strävar efter att jämföra med den posteriora sannolikhetsfördelning som härleds från den aktuella observationen $p_{xy}^{t+1}(\vartheta|\varphi)$. Det är värt att notera att $n_{xy}^{t+1*}(\vartheta = T)$ omfattar värden över alla delmängder av utfallsrummet Θ , inklusive Θ självt (där $|\Theta| = 3$), medan $p_{xy}^t(\vartheta|\varphi)$ är begränsad till värden associerade med elementen i Θ (där $|\Theta| = 2$). Därför måste vi omvandla den grundläggande tilltrofördelningen till en sannolikhetsfördelning för att möjliggöra en jämförelse mellan de två fördelningarna. Denna omvandling uppnås genom en pignistisk transformation [35], definierad enligt följande:

$$\text{BetP}(\omega) = \sum_{\omega \in X} \frac{m(X)}{|X|} \quad (27)$$

där $\omega \in \Theta$ och $X \subseteq \Theta$.

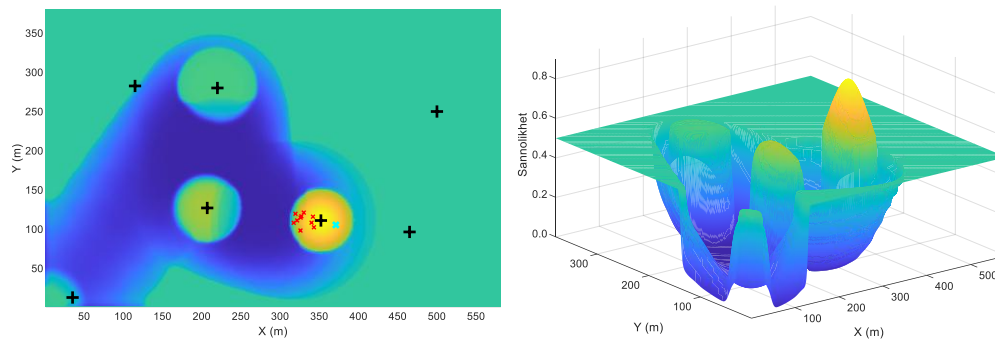
Vi har

$$\begin{aligned} q_{xy}^{t+1*}(\vartheta = T) &= \frac{n_{xy}^{t+1*}(\vartheta = T)}{|\{T\}|} + \frac{n_{xy}^{t+1*}(\vartheta = \Theta)}{|\Theta|} \\ &= n_{xy}^{t+1*}(\vartheta = T) + \frac{1 - n_{xy}^{t+1*}(\vartheta = T) - n_{xy}^{t+1*}(\vartheta = F)}{2} \\ &= \frac{1}{2} [1 + n_{xy}^{t+1*}(\vartheta = T) - n_{xy}^{t+1*}(\vartheta = F)] \end{aligned} \quad (28)$$

och

$$\begin{aligned} q_{xy}^{t+1*}(\vartheta = F) &= \frac{n_{xy}^{t+1*}(\vartheta = F)}{|\{F\}|} + \frac{n_{xy}^{t+1*}(\vartheta = \Theta)}{|\Theta|} \\ &= n_{xy}^{t+1*}(\vartheta = F) + \frac{1 - n_{xy}^{t+1*}(\vartheta = T) - n_{xy}^{t+1*}(\vartheta = F)}{2} \\ &= \frac{1}{2} [1 + n_{xy}^{t+1*}(\vartheta = F) - n_{xy}^{t+1*}(\vartheta = T)], \end{aligned} \quad (29)$$

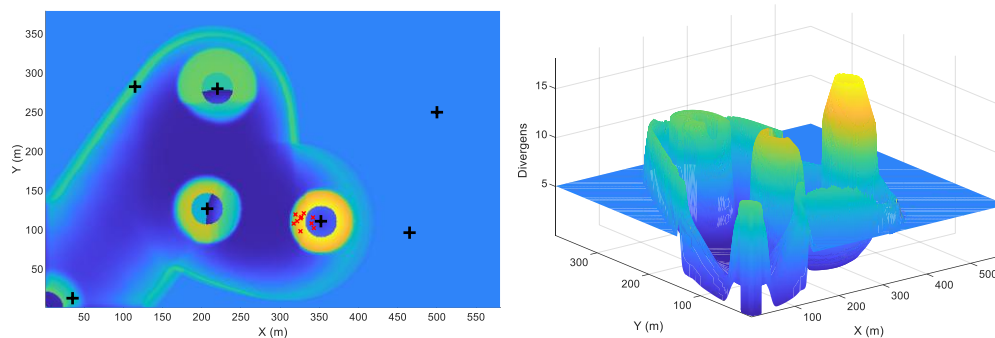
där $q_{xy}^{t+1*}(\vartheta)$ representerar en pignistisk sannolikhet härledd från den pignistiska transformationen av den grundläggande tilltrofördelningen $n_{xy}^{t+1*}(\vartheta)$. I ekvationerna (28) och (29) har vi $|\{T\}| = |\{F\}| = 1$ och $|\Theta| = 2$. Figur 5 illustrerar den transformerade pignistiska sannolikhetsfördelningen av evidensfördelningen som visas i figur 3.



Figur 5. En pignistisk sannolikhetsfördelning $q_{xy}^{t+1*}(\vartheta)$. Vänster och höger visar samma fördelningar: 2D till vänster och 3D till höger.

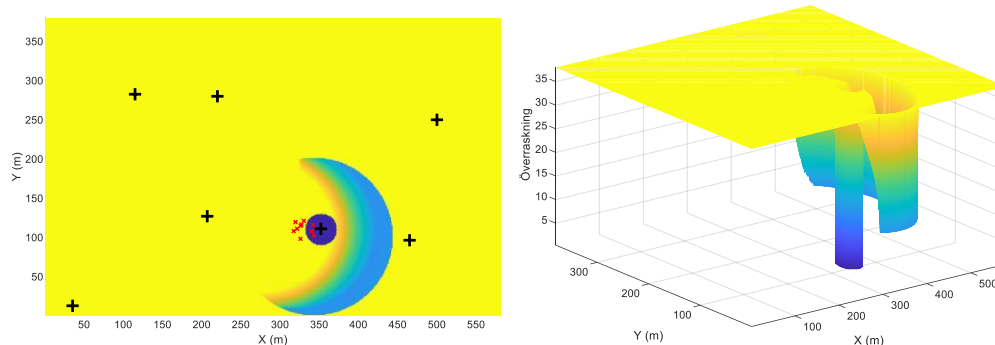
Vi kan nu beräkna VFE för varje rutnätscell c_{xy} inom sensorradien genom att mäta divergensen mellan $q_{xy}^{t+1*}(\vartheta)$ och $p_{xy}^{t+1}(\vartheta|\varphi)$, tillsammans med graden av överraskning som är representerad av $p_{xy}^{t+1}(\varphi)$.

Figur 6 illustrerar divergensen mellan den pignistiska sannolikhetsfördelningen $q_{xy}^{t+1*}(\vartheta)$ från den generativa modellen och den posteriora fördelningen $p_{xy}^{t+1}(\vartheta|\varphi)$ från den generativa processen.



Figur 6. Divergensen mellan modellen och den faktiska observationen. Cellvis $D_{KL}[q_{xy}^{t+1*}(\vartheta) \parallel p_{xy}^{t+1}(\vartheta|\varphi)]$. Detta är den första termen i ekvation (30) inom sensors radie ($r = 100$ meter), vilket illustrerar skillnaden mellan modellens tilltro och a posteriori. Vänster och höger visar samma divergens: 2D till vänster och 3D till höger.

Figur 7 visar graden av överraskning, indikerad som $-\ln[p_{xy}^{t+1}(\varphi)]$.



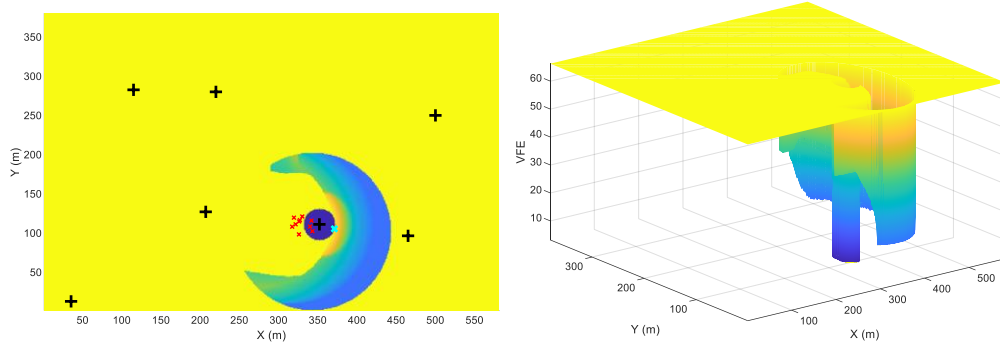
Figur 7. Graden av överraskningen $-\ln[p_{xy}^{t+1}(\varphi)]$, som representerar den andra termen i ekvation (30) visas inuti sensors fotavtryck. Utanför fotavtrycket, där $p(\varphi) = 0$, är graden av överraskning oändlig. I visualiseringen har vi ritat graden av överraskning utanför sensorradien som lika med den högsta graden av överraskning inom sensorradien (eftersom vi inte kan rita oändligt höga värden). Både vänster och höger visar graden av överraskning: 2D till vänster och 3D till höger.

Vi har

$$\begin{aligned}
 F_{xy} &= D_{KL}[q_{xy}^{t+1*}(\vartheta) \parallel p_{xy}^{t+1}(\vartheta|\varphi)] - \ln[p_{xy}^{t+1}(\varphi)] \\
 &= \left[\sum_{\vartheta \in \{T,F\}} q_{xy}^{t+1*}(\vartheta) \cdot \ln\left(\frac{q_{xy}^{t+1*}(\vartheta)}{p_{xy}^{t+1}(\vartheta|\varphi)}\right) \right] - \ln[p_{xy}^{t+1}(\varphi)], \quad (30)
 \end{aligned}$$

där D_{KL} är Kullback-Leibler-divergensen [34].

VFE F_{xy} beräknas för alla rutnätsceller inom sensorradien och minimeras sedan över dessa. Slutligen visas VFE i figur 8.



Figur 8. VFE, betecknad som F_{xy} , för alla rutnätsceller c_{xy} . F_{xy} beräknas med hjälp av ekvation (30). I visualiseringen har vi ritat F_{xy} utanför sensorradien som lika med det högsta värdet inom sensorradien. I visualiseringen har vi ritat VFE utanför sensorradien som lika med det högsta VFE:t inom sensorradien (eftersom vi inte kan rita oändligt höga värden). Vänster och höger visar båda VFE: 2D till vänster och 3D till höger.

7.4 Val av handling

Agenten väljer en handling a när den anländer till sin nuvarande vägpunkt på plats $c_{x_1y_1}$. Handlingen innebär att gå vidare mot en ny vägpunkt vid rutnätscellen $c_{x_2y_2}$, som placeras i riktningen $c_{x_0y_0}$ vilket minimerar F_{xy} vid ett fast avstånd lika med halva sensorradien. Om $c_{x_0y_0} = c_{x_1y_1}$ så är riktningen odefinierad och waypointen förblir vid den aktuella positionen, dvs. $c_{x_2y_2} = c_{x_1y_1}$. Vi har

$$\begin{aligned}
 a \left(agent.c_{x_1y_1} \middle| \begin{aligned} (x_0, y_0) &= argmin_{xy} F_{xy} \\ (x_2, y_2) &= (x_1, y_1) + (x_0 - x_1, y_0 - y_1) \cdot \frac{s_r}{2\sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2}} \end{aligned} \right) \\
 \rightarrow agent.c_{x_2y_2}, \quad (31)
 \end{aligned}$$

där $agent.c_{x_1y_1}$ är agentens aktuella position och $agent.c_{x_2y_2}$ är den uppdaterade positionen. Ledaren tilldelar denna waypoint till svärmen och lågnivåkontrollen flyttar drönarna stegvis mot den. Följaktligen, genom att använda aktiv inferens flyttar vi gradvis agenten mot rutnätscellen $c_{x_2y_2}$, och rör oss därmed i riktning mot den position som minimerar F_{xy} .

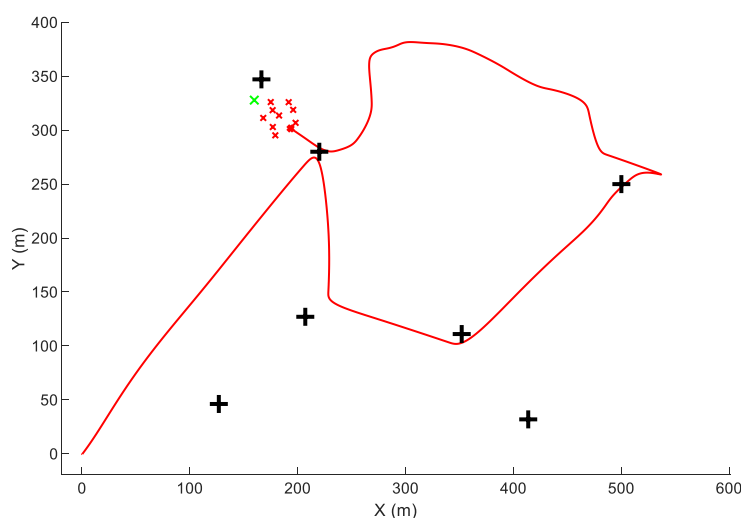
8 Implementation

Att implementera agentstyrning med hjälp av fri energi-principen och aktiv inferens kräver att man utför beräkningar enligt beskrivningen i kapitel 7. Emellertid måste flera viktiga implementationsdetaljer och designval beaktas. I detta kapitel beskrivs först den implementerade metoden för UAV-styrning baserad på VFE (kapitel 8.1). Därefter beskrivs den använda simuleringsmiljön (kapitel 8.2) och grundscenariot som använts vid försök och utvärdering av den VFE-baserade agentstyrningen (kapitel 8.3). Kapitlet avslutas med några ord om beräkningskomplexitet (kapitel 8.4).

8.1 Metod för styrning med VFE

Den generativa modellen och processen utvärderas vid varje steg i simuleringen. Det är viktigt att notera att VFE endast beräknas för positioner inom sensors radie; i denna simulering utvärderas vi sensorradier från 10 till 100 meter. Denna begränsning är inte bara ett designval utan också en grundläggande aspekt av modellen. Enligt ekvation (30) definieras VFE som summan av Kullback-Leibler-divergensen mellan två sannolikhetsfördelningar, $D_{KL}[q_{xy}^{t+1*}(\vartheta) \parallel p_{xy}^{t+1}(\vartheta|\varphi)]$, och överraskningsfaktorn, som beräknas som $-\ln[p_{xy}^{t+1}(\varphi)]$. Här representerar $p_{xy}^{t+1}(\varphi)$ sannolikheten för observationen. Utanför sensors radie är $p_{xy}^{t+1}(\varphi)$ lika med noll, vilket gör att $-\ln[p_{xy}^{t+1}(\varphi)]$ närmar sig oändligheten. Denna observation belyser att graden av överraskning blir oändlig när man konfronteras med det oobserverade. Därför begränsar vi VFE-bedömningen till positioner inom sensors räckvidd. Utanför denna gräns är sannolikheten odefinierad i avsaknad av observationer, så vi beräknar inte VFE i dessa områden.

Ett annat designval gäller valet av nästa vägpunkt. När positionen med det lägsta VFE har identifierats, beräknar vi en vektor från agentens nuvarande position till denna punkt. Den nya vägpunkten placeras på ett avstånd av $0,5 s_r$ längs denna vektor (dvs. en steglängd på 50 m när sensorradien är 100 m) (se figur 9). Värdet $0,5 s_r$ valdes som en heuristisk kompromiss: det håller vägpunkten inom sensorradien och producerar tillräcklig överlappning mellan på varandra följande vägpunkter, samtidigt som det fortfarande tillåter meningsfullt rumsligt framsteg innan nästa omplanering. Agenten förflyttar sig därefter stegvis mot denna plats. Observationerna och lägesförståelsekartan uppdateras vid varje simuleringssteg. En ny vägpunkt utses när den nuvarande har nåtts.



Figur 9. Banan, vägpunkterna och positionen för den minsta VFE:n. Röda kryss ("x") representerar UAV:er. Ledaren (markerad med en röd punkt "•") väljer en ny vägpunkt längs en vektor som pekar mot punkten med den minsta VFE:n (markerad med ett grönt kryss "x"). De svarta kryssen ("x") indikerar de exakta platserna för målen (sett ur ett globalt perspektiv), som agenten inte känner till direkt, utan endast genom sensorobservationer. Det gröna kryss ("x") markerar alltså $\argmin F_{xy}$, som agenten använder för att ställa in nästa vägpunkt.

Agenten samlar in sensorobservationer och uppdaterar den probabilistiska kartan G var 0,2 sekunder. När den VFE-baserade svärmstyrstrategin träder i kraft, bestäms punkten med lägst fri energi. En vägpunkt sätts sedan till 50 % av sensorns radie från agentens position, orienterad mot den lägsta fria energipunkten. Denna vägpunkt, som anges av agenten, delas av alla UAV:er. När vägpunkten nås upprepas processen: en ny minsta fria energipunkt identifieras och cykeln fortsätter tills scenariot avslutas. Samma steglängdsregel tillämpas på styrstrategier baserade på VFE, divergens och överraskning. Strategierna för slumpmässigt och fast spår följer sina respektive fördefinierade rörelseregler.

8.2 Simuleringsmiljö

Vi genomför UAV-styrningsexperiment med aktiv inferens i MATLAB. För de genomförda experimenten användes en modifierad version av en dynamisk multiagentsimulator (MADS), som utvecklats vid FOI. I MADS-simulatorens modelleras de fysikaliska rörelseekvationerna för en grupp UAV:er. Simulatorens hanterar också omvandlingen av högnivå-styrparametrar (t.ex. hastighet och rotation) till hårdvarunära styrkommandon för enskilda UAV:er. I MADS-simulatorens finns även ett grundläggande beteende för svärmstyrning implementerat. Detta beteende baseras på den inom styrteorin klassiska fjäddämpningsprincipen (*eng.* spring-damper) och fungerar så att det dels håller samman en grupp UAV:er, dels förhindrar att de kolliderar med varandra. I experimenten simuleras en grupp UAV:er, där en av UAV:erna (som vi fortsättningsvis refererar till som agenten) fungerar som ledare och ansvarar för all planering, medan de övriga följer ledaren med hjälp av en modifierad version av simulatorens svärmstyrning. Dessa *följare* kan bära på sensorer och bidra till att bygga upp en mer detaljerad och välgrundad lägesbild, men används inte för detta i de rapporterade experimenten. Däremot kan deras existens påverka resultaten i experimenten, eftersom de utvärderingsmått som använts har beräknats för svärmens ledare, som påverkas av gruppens svärmstyrningsbeteende.

8.3 Grundscenario och parametrar

Experimenten som beskrivs i kapitel 9 utgår från ett grundscenario som har implementerats med vissa valda parameterinställningar. Scenariot och viktiga parametervärden beskrivs kort i det här avsnittet.

Experimenten utgår från den tidigare introducerade scenariokartan med $1\text{ m} \times 1\text{ m}$ rutnätsceller över det tvådimensionella spaningsområdet på $580\text{ m} \times 380\text{ m}$. I varje simulering finns det fyra fasta och tre rörliga målobjekt i spaningsområdet. De fasta målobjekten placeras slumpmässigt i spaningsområdet inför varje ny simulering. De tre rörliga målobjekten rör sig med hastigheter på 1, 1,5 respektive 2 m/s och startar vid simuleringens början från slumpmässiga positioner längs kanterna av spaningsområdet, varpå varje målobjekt rör sig i en rak linje mot en slumpmässigt tilldelad position på den motsatta sidan av området. Målen hastigheter är medvetet valda så att vissa sannolikt lämnar simuleringsområdet innan 500-sekundersscenariot är slut. Detta kan ibland leda till att observerade mål förloras, vilket tvingar svärmen att söka nya. I sällsynta fall kan alla tre målen lämna området innan scenariot är färdigt.

UAV-svärmen börjar alltid vid tidpunkten $t = 0$ i det nedre vänstra hörnet av scenariokartan. Svärmen traverserar inledningsvis en bana som passerar de fyra fasta målobjekten enligt den kortaste vägen – detta för att UAV-svärmen ska få möjlighet att samla in initiala sensorobservationer som används för att börja bygga upp den gemensamma operativa lägesbilden, som representeras av den probabilistiska kartan G (i grundscenariot är sensorradien satt till 100 m, men den varierar i vissa experiment, se kapitel 9). När svärmen har nått det fjärde fasta målet växlar ledar-UAV:n till den i förväg valda styrstrategin, som sedan används under resten av scenariot. UAV:s rörelseriktningar och hastigheter bestäms under hela scenariot av en grundläggande vägpunktsbaserad styralgorithm i MADS-simulatorens (kapitel 8.2), som i varje tidssteg i simuleringen tar den nästa gällande vägpunkten som indata och styr agenten i fråga mot den. Som tidigare nämnts kan även det

inbyggda svärmstyrningsbeteendet i MADS-miljön påverka UAV:ernas rörelser. Som en följd av detta implementationsval varierar UAV:ernas hastigheter under simuleringarna. Deras genomsnittliga hastighet över alla scenarier i de genomförda experimenten har beräknats till 4,23 m/s.

I simuleringarna används ett tidssteg på 0,2 s, det vill säga att varje scenario uppdateras var 0,2 s (i simuleringssvärldens tid) med nya sensorobservationer och förflyttningar av både UAV:er och rörliga målobjekt. Simuleringarna i experimenten varar i 500 s (även detta i simuleringssvärldens tid).

8.4 Beräkningskomplexitet

Exklusive visualiseringsrutiner är en konservativ övre gräns för tids- och beräkningskomplexiteten per steg i vår planerare

$$O(N_{UAV}^2 + N_{mål}N_{celler} + N_{celler} + s_r^2), \quad (32)$$

där N_{UAV} är antalet UAV:er i svärmen, $N_{mål}$ är antalet målobjekt, N_{celler} är antalet rutnätsceller och s_r är sensorradien uttryckt i rutnätsceller. I praktiken tar varje simuleringsssteg i ett 580×380 -rutnät med tio UAV:er, sju målobjekt och en sensorradie på 10–100 celler (dvs. 10–100 m), i genomsnitt 0,12 sekunder, vilket motsvarar 300 sekunders beräkningstid för ett 500-sekundersscenario med fem simuleringsssteg per sekund (dvs. snabbare än realtid) på en vanlig bärbar dator (MATLAB; Intel i7; 32 GB RAM).

9 Experiment och resultat

I detta kapitel beskrivs det simuleringsbaserade experimentet som har genomförts för att undersöka egenskaper och prestanda hos den VFE-baserade styrstrategin. Experimenten har utformats för att särskilt undersöka beteenden som är viktiga vid autonom ihållande övervakning och har genomförts med fem olika styrstrategier. Utöver den VFE-baserade styrningen har de två VFE-komponenterna, divergens och graden av överraskning, implementerats var för sig – för att möjliggöra en analys av deras respektive inverkan på den VFE-baserade styrningen. Vidare undersöks två referensstrategier: dels en variant av slumpmässig gång (*eng.* random walk), dels en strategi där agenten kontinuerligt loopar en i förväg designad fast trajektoria för effektiv avsökning av hela det fördefinierade spaningsområdet.

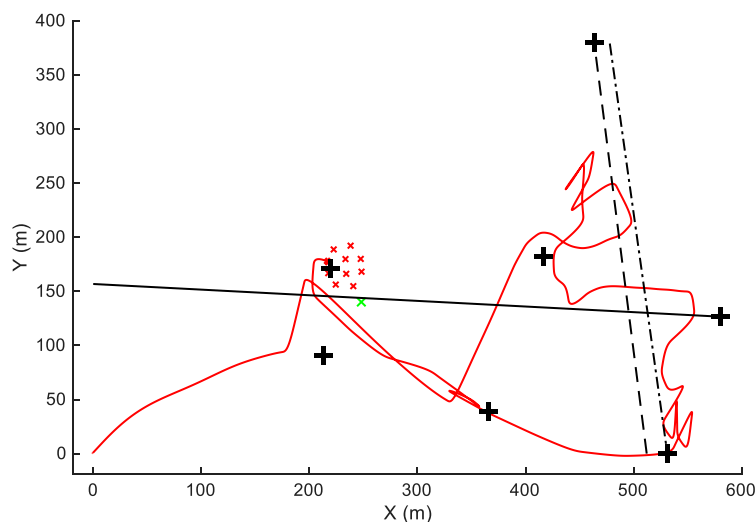
De fem styrstrategierna utvärderas utifrån UAV:ernas förmåga att detektera, följa och korrekt positionsbestämma målobjekten som befinner sig inom det givna spaningsområdet. Resultaten visar att den VFE-baserade styrningen överträffar de andra strategierna när det gäller att följa målobjekt över tid, medan den fasta trajektorien är mest effektiv för att upptäcka alla målobjekt inom spaningsområdet. Av de två VFE-komponenterna har divergens en starkare effekt på prestandan än graden av överraskning.

I avsnitt 9.1 ges en kvalitativ bedömning av den VFE-baserade styrstrategin utifrån en genomförd simulering [36]. I avsnitt 9.2 presenteras en kvantitativ jämförelse mellan den VFE-baserade rörelsestyrningen och de fyra andra styrstrategierna, där resultaten belyser viktiga egenskaper hos den VFE-baserade styrningen. I avsnitt 9.3 diskuteras de erhållna resultaten.

9.1 Kvalitativ analys

Manuell analys av UAV:ernas bana under ett flertal simuleringar med VFE-styrningen bekräftar att den i stora drag fungerar ungefär som förväntat. Efter att UAV-svärmen har passerat de fyra fasta målen uppvisar den ett beteende som växlar mellan sökning och spårning medan den fortsätter att utforska området.

För att visa hur UAV:ernas växlande beteende kan yttra sig utgår vi från en genomförd simulering. I figur 10 visas scenariokartan med agentens (UAV-svärmens ledare) och de rörliga målobjektens trajektorier för det valda 500-sekundersscenariot.



Figur 10. Överblick över agentens (röd linje) och de tre rörliga målobjektens (svarta linjer) trajektorier under ett 500 sekunder långt scenario. Röda kryss ("x") markerar UAV:erna i svärmen. Röd punkt ("•") markerar svärmens ledare. Svarta kryss ("+") visar målobjektens positioner. Gröna kryss ("x") markerar den position som vid aktuell tidpunkt i scenariot har lägst variationell fri energi (VFE). Samtliga positionsangivelser gäller för simuleringstid $t = 500$ (d.v.s. vid slutet av scenariot).

Det som sker i scenariot kan tolkas på följande sätt: När det fjärde fasta målet vid koordinaterna (420, 180) nås, aktiveras VFE-styrningen. UAV-svärmen detekterar ett norrgående målobjekt som rör sig med 1,5 m/s (svart streckad linje) och följer det norrut. På vägen möter svärmen ett långsammare södergående målobjekt som rör sig med 1 m/s (svart streck-punktad linje), och svärmen svänger då söderut för att i stället förfölja det. Strax därefter anpassar svärmen rutten och närmar sig det närliggande fjärde fasta målobjektet för närmare inspektion. Därefter detekterar svärmen ett snabbt östgående målobjekt som rör sig med 2 m/s (svart heldragen linje) och följer det österut genom sökområdet. Svärmen möter sedan det långsamt rörliga målet igen och följer det söderut till punkten (540, 0), där det långsamma målet lämnar sökområdet. Svärmen rör sig då västerut och börjar söka: först återupptäcker den det tredje fasta målet vid koordinaterna (370, 40), sedan det andra vid koordinaterna (220, 170).

Denna sekvens av handlingar illustrerar hur svärmen, med hjälp av en VFE-baserad styralgorithm, kan växla mellan att observera olika målobjekt och smidigt övergå från spårning till sökning.

Slutsatser från en kvalitativ analys av ett begränsat antal scenarier ska tolkas med försiktighet, eftersom det inte är säkert att UAV-beteenden som uppvisas i just dessa scenarier är representativa för beteendemönster som kan framgå av mer omfattande försök med större variation i parametrarna. Kvalitativ analys kan dock ge viktiga insikter, till exempel om hur agenten prioriterar mellan olika beteenden.

9.2 Kvantitativ analys

I detta avsnitt används en simuleringsbaserad metod för att undersöka om aktiv inferens kan hantera utmaningen att styra en UAV-svärm, där en ledare med en enda sensor (agenten) genomför ihållande övervakning av flera mål (se avsnitt 2.1).

Ihållande övervakning kan i många fall betraktas som ett multimåloptimeringsproblem. I scenariot som betraktas här är det övergripande målet för UAV-svärmen att – över tid – upprätthålla så korrekta positionsestimat som möjligt för alla målobjekt inom ett givet spaningsområde. På grund av sensorns begränsade räckvidd måste agenten dock växla mellan att söka efter odetekterade målobjekt och att återbesöka kända målobjekt för att uppdatera sig om deras exakta positioner. För att förbättra situationsförståelsen kan ytterligare åtgärder, såsom att följa efter och observera ett rörligt målobjekts aktivitet under en viss tidsperiod, också vara nödvändiga. Oavsett vilken styrstrategi agenten använder behöver den generera lämpliga beteenden och, när det behövs, skifta mellan dem på ett effektivt sätt.

Ett sätt att kvantitativt utvärdera hur bra en styralgorithm presterar under ihållande övervakning är att genomföra simuleringar och mäta hur mycket målobjektens positioner i svärmens världsmodell avviker från de faktiska målobjektens positioner. Denna utvärderingsmetod gör det möjligt att jämföra olika styralgoritmer, men den ger begränsad insikt i UAV:ernas beteende under simuleringen. Eftersom detta är ett utforskande arbete har vi valt att inte fastställa ett enda övergripande utvärderingsmått, utan i stället ta fram flera utvärderingsmått (se beskrivningar i avsnitt 9.2.1 och 9.2.2) som är utformade för att mäta i vilken utsträckning en styralgorithm uppfyller olika (ibland motstridiga) delmål inom ihållande övervakning. Följande delmål utvärderas kvantitativt i experimenten:

- Upprätthållande av en hög grundläggande tilltro till målnärvaro vid de faktiska målpositionerna (avsnitt 9.2.1).
- Tillförlitlig detektion av målobjekt i simuleringsområdet (avsnitt 9.2.2).
- En jämn fördelning av observationstiden mellan de rörliga målen (avsnitt 9.2.2).
- Spårningskontinuitet under en sammanhängande tidsperiod (avsnitt 9.2.2).

I avsnitt 9.2.1 och 9.2.2 presenteras experimentupplägg, utvärderingsmått samt de erhållna resultaten för de fem styrstrategierna VFE, divergens, graden av överraskning, slumpmässig gång och fast trajektor. Utvärderingsmått är i sig tolkningsbara, men används i detta arbete främst för att jämföra egenskaperna hos de olika styralgoritmerna.

9.2.1 Grundläggande tilltro i förhållande till verkliga målpositioner

Det huvudsakliga målet för agenten som leder UAV-svärmen är att upprätthålla aktuella uppskattningar av målens positioner i varje simulering. Även om detta kan verka vara ett enkelt framgångsmått, visar en närmare granskning att det är mer komplext att bedöma måluppfyllelsen. Att förlita sig på en karta av grundläggande tilltro för att representera världen har sina begränsningar. Till exempel skiljer den inte mellan mål; alla observationer behandlas lika, oavsett vilket mål de gäller. Detta gör det svårt att analysera positionsuppskattningsfel för enskilda mål. Dessutom är det utmanande att exakt bestämma hur många mål som finns inom det undersökta området på kartan med den här metoden. Det är också värt att överväga om mätningen av ett måls positionsuppskattningsfel förblir meningsfull om det överskrider ett visst tröskelvärde, varvid uppskattningen *anses otillförlitlig*. I praktiken kan det vara mer relevant att övervaka i vilken utsträckning en målpositionsuppskattning är *tillräckligt nära* den faktiska målplatsen.

I enlighet med de tidigare punkterna använder vi ett förenklat mått för att utvärdera måluppfyllelsen genom att mäta stödet för faktiska målpositioner över tid. Istället för att beräkna positionsuppskattningar och deras fel summerar detta mått de grundläggande tilltrosuppfattningarna som placerats i kartceller inom en radie på 10 meter från den verkliga positionen. Metoden ger ett mer tillförlitligt mått än att enbart förlita sig på en exakt bedömning av målplatsen. För enkelhetens skull refererar vi till området inom 10 meter från mål j vid tidpunkten k som A_j^k . Därför definieras utvärderingspoängen enligt följande:

$$\phi = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \frac{1}{N_{rörliga}} \sum_{j=1}^{N_{rörliga}} \frac{1}{t_{stop}(i,j) - t_{start}(i)} \sum_{k=t_{start}(i)}^{t_{stop}(i,j)} \frac{1}{|A_j^k \cap G|} \sum_{(x,y) | c_{xy} \in A_j^k \cap G} n_{xy}^t(\vartheta = T), \quad (33)$$

där N_{sim} anger antalet simuleringar, $N_{rörliga}$ är antalet rörliga mål, $t_{start}(i)$ är tiden i scenario nummer i då styrstrategier aktiveras för resten av scenariot (efter att ha nått det fjärde fasta målet), $t_{stop}(i, j)$ är sluttiden i scenario i då mål j lämnar området (eller den är 500 om det förblir inom simuleringens område i slutet av simuleringen), A_j^k är området med en radie på 10 meter runt mål j vid tidpunkten k , och $A_j^k \cap G$ är den del av A_j^k som ligger inom G . Observera att $|A_j^k \cap G| < |A_j^k|$ då målet befinner sig inom 10 meter från simuleringens områdes gräns.

Det föreslagna måttet tar inte hänsyn till falska indikationer på målnärvaro på platser där det inte finns något mål. Detta förenklade mått bygger på antagandet att den mjuka sensormodell som används i simuleringarna inte genererar falska observationer. En mer avancerad version av detta mått skulle kunna utformas genom att justera:

- $\phi_{mål}^* = \phi$,
- $\phi_{\neg mål}^* = \phi$, med den sista summeringen i ekvation (33) utvärderad över $(x, y) | c_{xy} \notin A_j^k \cap G$ istället för $(x, y) | c_{xy} \in A_j^k \cap G$,

och definiera utvärderingspoängen som

$$\phi^* = \frac{\phi_{mål}^*}{\phi_{\neg mål}^*}. \quad (34)$$

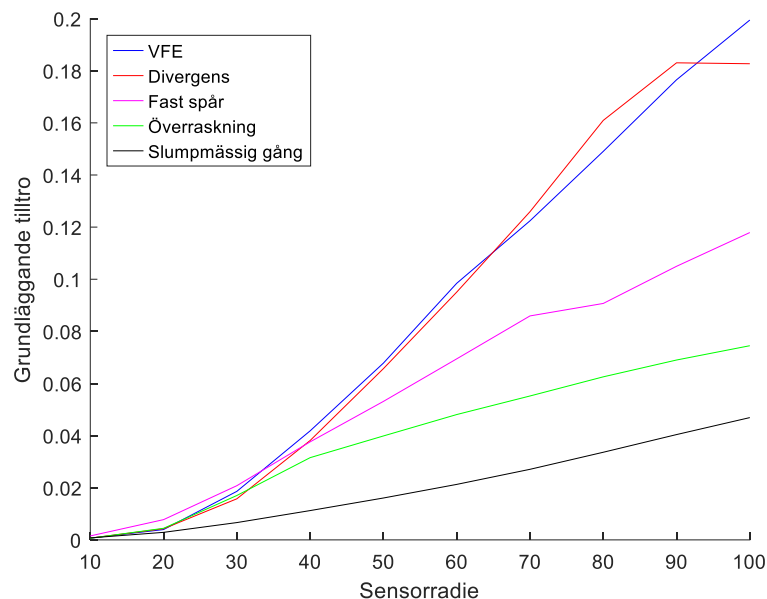
En hög utvärderingspoäng, ϕ^* , antyder att den positiva grundläggande tilltron i världsmodellen fokuserar på målens verkliga positioner snarare än att vara jämnt fördelade över området. Poängen säger dock inte hela historien. Det kan fortfarande finnas falska måldetekteringar som, om de ges tillräckligt med grundläggande tilltro, kan få agenten att vidta suboptimala åtgärder.

I simuleringarna beräknades utvärderingspoängen, ϕ , för fem olika styrstrategier: VFE, divergens, graden av överraskning, slumpmässig gång och fast spår. Resultaten sammanfattas i tabell 2 och figur 11. Det är tydligt att ingen av metoderna fungerar bra med små sensorradier, men fastspårningsmetoden, som är utformad för att systematiskt skanna hela området efter mål, har en liten men tydlig fördel. Med fastspårsmetoden som referens kan vi se att VFE-baserad styrning vid sensorradier på 40 meter eller mer ger högre grundläggande tilltro till nära verkliga målpositioner. För mindre sensorradier visar fastspårsmetoden en viss fördel. När sensorradien ökar förbättras VFE:s prestation i jämförelse med referensen, eftersom VFE i högre grad använder längre sensorradier för att upptäcka nya mål under manövrering, till skillnad från referensen som inte aktivt manövrerar utifrån sensordata. Specifikt, vid en sensorradie på 100 meter, överträffar VFE-referensen med 69,2 %. Tabellen visar att divergenskomponenten främst påverkar VFE:s prestanda, men skillnaden mellan VFE och divergens vid 100-metersradien återspeglar också ett visst bidrag från minimering av graden av överraskning. Styrstrategier som enbart förlitar sig på graden av överraskning eller slumpmässig gång är däremot ineffektiva.

Tabell 2. Utvärderingspoängen, ϕ , är summan av den grundläggande tilltron för alla rutnätsceller inom 10 meter från något av de tre rörliga målen. Den beräknas för de fem styrningsfunktionerna och för olika sensorradier. Vi ser att VFE överstiger riktmärket för fast spår vid sensorradier på 40 meter eller mer.

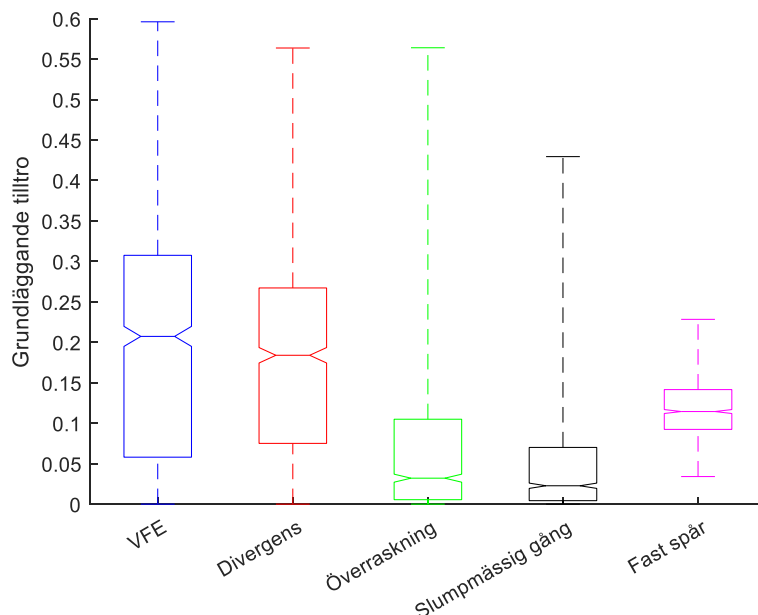
Sensor radie	VFE	Divergens	Graden av överraskning	Slumpmässig gång	Fast spår	VFE / Fast spår
10 m	0,0007	0,0007	0,0006	0,0008	0,0015	0,4667
20 m	0,0040	0,0043	0,0044	0,0029	0,0078	0,5128
30 m	0,0186	0,0158	0,0170	0,0066	0,0208	0,8942
40 m	0,0418	0,0381	0,0315	0,0112	0,0376	1,1117
50 m	0,0678	0,0657	0,0399	0,0161	0,0531	1,2768
60 m	0,0984	0,0951	0,0481	0,0213	0,0695	1,4158
70 m	0,1224	0,1259	0,0552	0,0271	0,0859	1,4249
80 m	0,1491	0,1610	0,0626	0,0336	0,0907	1,6439
90 m	0,1765	0,1830	0,0690	0,0404	0,1050	1,6810
100 m	0,1995	0,1827	0,0745	0,0469	0,1179	1,6921

De olika scenarierna i simuleringarna varierar, vilket innebär att vår utvärdering både jämför dessa styrstrategier och, i viss mån, de möjliga resultat som kan uppnås i sådana scenarier.



Figur 11. Utvärderingspoängen, ϕ , är summan av den grundläggande tilltron för varje sekund för tre banor med rörliga mål. Poängen är medelvärdet av 1000 scenarier och sensorradier från 10 m till 100 m. Figuren jämför VFE-styrningsfunktionen (blå) med divergens (röd), graden av överraskning (grön), slumpmässig gång (svart) och ett fast spår (magenta).

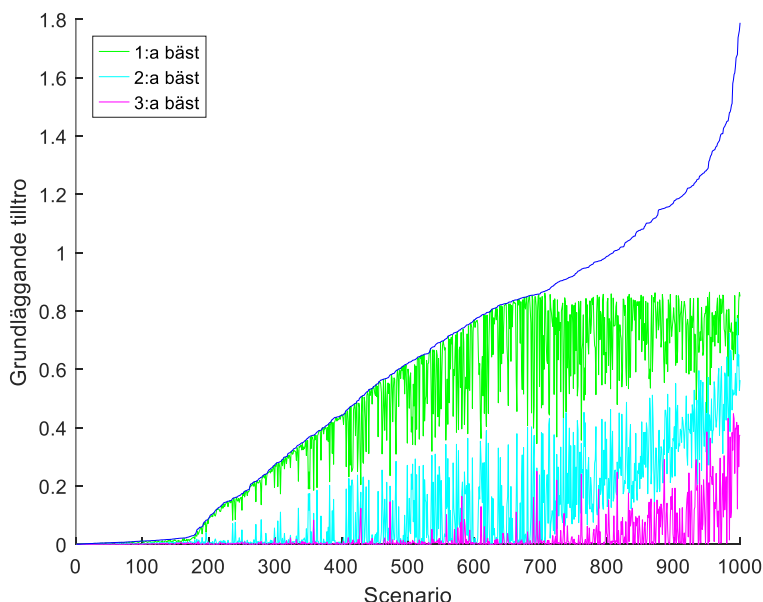
Data i tabell 2 och figur 11 visar medelvärden från 1000 slumpmässigt genererade scenarier. För att visa variationen i utfall mellan dessa scenarier använder vi boxplottar (figur 12) som visar medianen samt den andra och tredje kvartilen. Boxarna visar interkvartilintervallet; whiskers visar de återstående nedre och övre intervallen enligt den använda boxplottkonventionen. Höga medianvärden för VFE och divergens tyder på god prestanda, men deras stora boxar och långa övre whiskers tyder också på hög varians. Däremot visar överraskningskomponenten och slumpmässig gång både låg prestanda och hög varians. Den fasta spårstyrningen har en relativt god median jämfört med andra styrstrategier, men är särskilt märkbar för sin låga varians.



Figur 12. Boxplott som visar utvärderingspoängen, ϕ . Resultaten är baserade på 1000 scenarier med en sensorradie på 100 meter. Boxarna visar den andra och den tredje kvartilen för alla scenarier.

Dessutom utvärderar vi, utifrån den grundläggande uppfattningen om målnärvaro, tre rörliga mål oberoende av varandra. Eftersom dessa mål enbart skiljer sig åt i hastighet och inte är kopplade på något annat sätt, tilldelar vi dem nummer baserade på agentens förmåga att

observera varje mål. I figur 13 visas det mål med den högsta observerade grundläggande tilltron i varje scenario—definierat som det mål med den högsta tilltron i alla rutnätsceller inom 10 m från dess position—markerat i grönt. Det näst mest observerade målet visas i cyanblått och det tredje i magenta. De 1000 scenarierna är ordnade efter den övergripande grundläggande uppfattningen att minst ett mål observeras. Detta visar att drygt hälften av VFE:s prestanda kommer från att detektera det mest synliga rörliga målet, medan mindre än hälften beror på att spåra de näst- och tredje mest observerade målen. Det illustrerar hur VFE balanserar mellan att spåra mål som redan är i sikte och att söka efter nya.



Figur 13. Grundläggande tilltro per sekund för tre banor för rörliga mål i över 1000 scenarier, med en sensorradie på 100 meter och med hjälp av VFE-styrningsfunktionen. Den blå linjen representerar den totala grundläggande tilltron (summan av de tre tilltrofördelningarna) för alla mål, medan de gröna, cyanblå och magenta linjerna visar de första, andra respektive tredje bäst observerade rörliga målen. Dessa fyra kurvor är ordnade från lägst till högst baserat på den totala summan av den grundläggande tilltron till alla mål (blå). X-axelindexet visar samma scenario för alla fyra kurvorna.

9.2.2 Upptäckts- och observationstider

Vi analyserar härnäst upptäcktsfrekvenser och observationstider för olika kontrollstrategier, specifikt för fallet med en sensorradie på 100 m. Vi börjar med att jämföra upptäcktsfrekvenserna mellan de fem strategierna.

Tabell 3 visar att VFE detekterar alla tre mål minst en gång i 46,4 % av scenarierna och även två mål med en sannolikhet på 38,3 %. Ett mål anses vara detekterat om sensorn är inom 100 meter från målets position under minst ett simuleringsteg (0,2 sekunder). Tabellen tyder på att strategier utan aktiv målspårning är mer benägna att upptäcka nya mål, vilket förklarar den höga upptäcktsfrekvensen för *fast spår*. Dessutom uppvisar VFE och divergens ett märkbart högt antal detekteringar av två eller fler mål.

Tabell 3. Andelen scenarier där olika antal rörliga mål detekteras över 1000 scenarier, med en sensorradie på 100 meter för de fem styrfunktionerna.

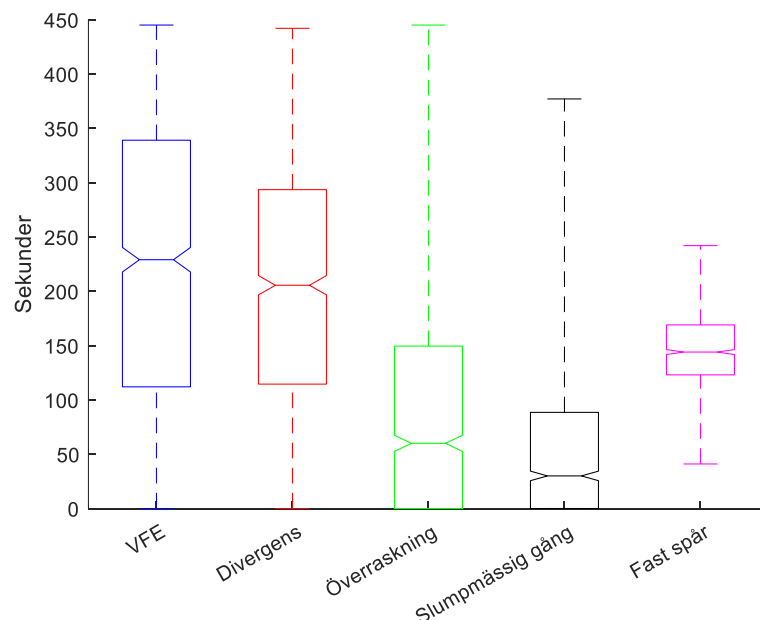
Antal mål	VFE	Divergens	Graden av övertäckning	Slumpmässig gång	Fast spår
0	0,017	0,017	0,036	0,051	0
1	0,136	0,131	0,205	0,205	0
2	0,383	0,374	0,404	0,409	0,011
3	0,464	0,478	0,355	0,335	0,989

En granskning av observationstiderna i tabell 4 visar att VFE konsekvent har de längsta genomsnittliga observationstiderna över 1000 scenarier och överträffar de andra styrstrategierna. VFE observerar i genomsnitt minst ett rörligt mål inom sensors räckvidd i 219,8 sekunder, vilket överträffar alternativen. Ett separat studium av alla scenarier visar att det i genomsnitt finns minst ett rörligt mål kvar i området i 374,0 sekunder efter att styrningen har aktiverats. Därför ägnas 58,8 % ($= 219,8 / 374,0$) av den tillgängliga tiden åt att spåra mål, medan resten används för att utforska nya. Dessutom är VFE särskilt effektivt på att målfölja sitt bäst följda mål, medan divergens presterar bättre när det gäller att följa sitt näst bästa mål.

Tabell 4. Genomsnittlig observationstid i sekunder för minst ett rörligt mål inom sensors räckvidd, för vart och ett av de tre rörliga målen, över 1000 scenarier med en sensorradie på 100 meter, för de fem kontrollfunktionerna. Dessa siffror kan omfatta flera observationssegment och återspeglar inte nödvändigtvis en kontinuerlig observation.

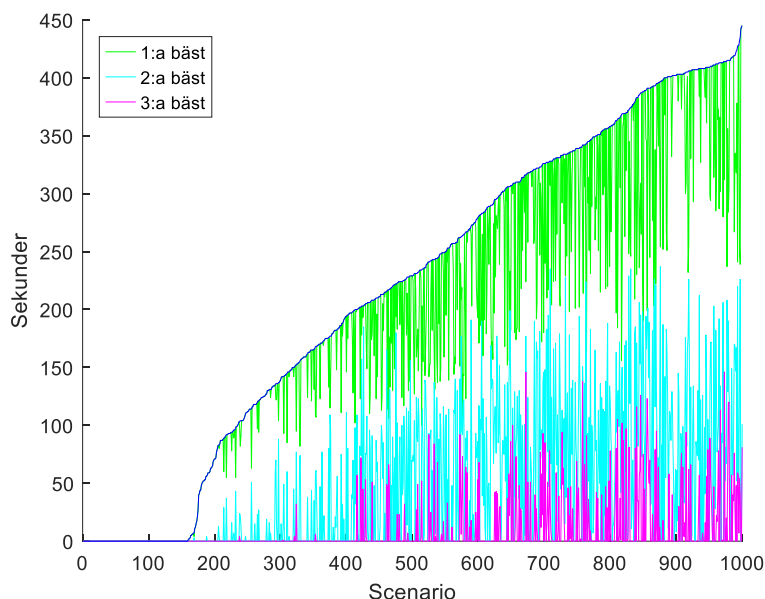
Observation i sekunder	VFE	Divergens	Graden av överraskning	Slumpmässig gång	Fast spår
Alla rörliga mål	219,8	199,3	92,1	54,6	145,5
1:a bäst	195,2	170,2	78,5	46,1	75,8
2:a bäst	48,5	56,1	19,6	13,3	56,0
3:a bäst	9,4	12,6	3,8	2,4	35,7

Figur 14 visar hur observationstiderna är fördelade för fem olika styrstrategier för alla rörliga mål. Den breda variansen speglar skillnader i svårighetsgrad mellan scenarierna. Boxplotten visar att VFE generellt presterar bättre än de andra strategierna, även om den också har den högsta variansen. Boxarna visar medianen samt de övre och undre kvartilerna för varje strategi i alla scenarier. Som tidigare nämnts klarar VFE sig bättre än divergens tack vare sin överraskningskomponent. VFE presterar också bättre än de andra strategierna, främst tack vare divergensen. Den slumpmässiga strategin fungerar dåligt ungefär 75 % av tiden, medan prestandan för riktmärket med fast spår oftast ligger i den första eller andra kvartilen av VFE.



Figur 14. Boxplott som visar observationstid i sekunder för något av de tre rörliga målen över 1000 scenarier, med en sensorradie på 100 meter, för de fem styrfunktionerna. Dessa värden kan omfatta flera observationssegment och innebär inte nödvändigtvis kontinuerlig observation. Boxarna anger den andra och den tredje kvartilen för alla scenarier.

Slutligen analyserar vi observationstiderna för de tre målen med hjälp av VFE i alla scenarier. Figur 15 visar scenarierna sorterade efter den totala observationstiden för de tre målen. Varje scenario visar målet med flest observationer i grönt, med en ojämn linje, det näst mest observerade målet i cyanblått och det tredje mest observerade målet i magenta. Den blå linjen visar den sammanlagda observationstiden för alla tre målen. Resultatet visar att det oftast observerade målet i varje scenario vanligtvis dominerar, men det näst mest observerade är också relevant. Enligt tabell 4 utgör det mest observerade målet 77,1 % av observationerna, medan det näst mest observerade står för 19,2 %. Dessa resultat stämmer överens med dem i figur 13, som analyserade den grundläggande tilltron till observationerna för varje mål. Där hade det mest observerade målet en grundläggande tilltro till målnärvaron på 78,1 %, medan det näst mest observerade hade 18,1 %.



Figur 15. Observationstiden i sekunder för vart och ett av de tre rörliga målen över 1000 scenarier, med en sensorradie på 100 meter. Den jämför VFE-styrfunktionen för alla mål (blå) och de tre främsta rörliga målen, rangordnade efter observationstid: 1:a (grön), 2:a (cyanblå) och 3:e (magenta). Alla fyra är sorterade från lägst till högst, baserat på det totala antalet sekunder som observerats för alla mål (blå). X-axelns index motsvarar samma simulering för alla fyra kurvor. Siffrorna kan representera flera observationssegment och behöver inte nödvändigtvis avspejla en kontinuerlig observation.

Målobserveringstiden visar hur väl olika styrstrategier presterar vid spårning, men kan bestå av flera observationssegment som summeras till den totala observationstiden och kanske inte helt speglar spårningskonsistensen. En analys av den genomsnittliga längden på det längsta kontinuerliga observationssegmentet för vart och ett av de tre målen visar variationer i spårningsprestanda mellan de mål som observeras mest, näst mest och minst. Som visas i tabell 5 håller VFE-styrstrategin en kontinuerlig observation av sitt bäst följda mål i genomsnitt 130,7 sekunder, vilket överträffar de andra fyra strategierna. Divergens når den längsta genomsnittliga varaktigheten för sitt näst bästa mål, medan metoden med fast spår är mest effektiv för att observera sitt minst spårade mål. Jämfört med tabell 5 och tabell 4 är den längsta kontinuerliga observationstiden för det bäst spårade målet, som en procentandel av den totala observationstiden: 67,0 % för VFE, 72,9 % för divergens, 81,9 % för graden av överraskning, 69,0 % för slumpmässig gång och 48,2 % för fast spår över alla scenarier. Detta visar hur ofta varje styrstrategi återbesöker varje mål, med graden av överraskning fokuserad på spårning, fast spår på sökning och VFE som balanserar båda.

Tabell 5. Genomsnittlig maximal kontinuerlig observationstid i sekunder för tre rörliga mål över 1000 scenarier, med en sensorradie på 100 meter för de fem styrfunktionerna. Siffrorna anger den genomsnittliga längden på det längsta kontinuerliga observationssegmentet i alla scenarier.

Observation i sekunder	VFE	Divergens	Graden av övertäckning	Slumpmässig gång	Fast spår
1:a bäst	130,7	124,0	64,3	37,7	36,6
2:a bäst	31,1	43,9	15,2	10,8	26,5
3:a bäst	4,3	9,6	2,6	1,9	17,6

9.3 Diskussion om resultaten

Den kvantitativa analysen visar att VFE generellt presterar bättre än riktmärket för fast spår vid sensorradier på 40 meter eller längre. Mer specifikt, vid en radie på 100 meter, överträffar VFE riktmärket med 69,2 % när det gäller grundläggande tilltro till målnärvaro (se tabell 2) samt med 51,1 % när det gäller observationsvaraktighet (se tabell 4). VFE presterar också bättre när det gäller att observera de två mest följda målen, med kontinuerliga observationstider på 161,8 sekunder (= 130,7 + 31,1) respektive 63,1 sekunder (= 36,6 + 26,5), jämfört med fast spår (se tabell 5). Dock innebär denna höga grad av kontinuerlig spårning en avvägning: VFE kan detektera minst två mål i 84,7 % av fallen, medan fast spår uppnår en detekteringsgrad på 100 %. Resultaten visar att VFE spårar mål under längre perioder än riktmärkena, men identifierar färre mål än det fasta spåret. En betydande del av VFE:s effektivitet är knuten till divergenskomponenten, där även graden av övertäckning spelar en roll, vilket illustreras av boxplott-diagrammen i figurerna 12 och 14.

10 Slutsatser

En simulering av en intelligent agent med VFE som styrstrategi visar dess förmåga att autonomt avgöra när den ska engagera sig i spaning och när den ska spåra specifika mål. Denna beslutsprocess baseras helt på data som samlats in via agentens sensor. Under arbetet navigerar agenten från sin nuvarande plats till den angivna rutnätscellen som minimerar VFE. VFE kvantifierar graden av överensstämmelse mellan agentens uppfattning av verkligheten och agentens aktuella observationer. Minimering av VFE förbättrar agentens förståelse och anpassningsförmåga till sin omgivning. Dessutom tar varje simuleringssteg i genomsnitt cirka 0,12 sekunder i den testade konfigurationen, vilket resulterar i en beräkningstid på 300 sekunder för ett 500-sekundersscenario på standardhårdvara. Detta visar en snabbare beräkning än i realtid i den utvärderade skalan, men skalbarheten för större svärmar, tätare målfält och större kartor har ännu inte fastställts. En kvalitativ bedömning av agentens beteende under hela simuleringen visar lovande resultat. Kvantitativa data som samlats in under simuleringarna indikerar att agenten överträffar riktmärkena för målobservation. VFE överträffar betydligt riktmärket för fast spår och visar en ökning med 69,2 % i den grundläggande tilltron till målnärvaro samt en ökning med 51,1 % i målobservationstid. Agentens förmåga att balansera sökning och målföljning innebär betydande operativ autonomi och effektivitet.

11 Framtida arbete

Framtida arbete kommer initialt att utvidga tillvägagångssättet till att omfatta:

- Aktiv inferens baserad på förväntad fri energi (EFE) för att planera kortsiktig förflyttning av en svärm av UAV:er, baserad på en långsiktig analys av möjliga agentbanor (policyer), samt jämföra VFE- och EFE-prestanda i ihållande övervakning gentemot de två metodernas beräkningskostnader.

Efterföljande utvidgningar omfattar:

- Hierarkisk modellering av tre olika nivåer med olika samverkande objektfunktioner för multikriteriebeslutsfattande:
 - Agerande för ihållande övervakning, hierarkisk nivå 1 (i denna rapport),
 - Agerande för egenskydd, hierarkisk nivå 2 (framtida arbete), och
 - Beslut om insats, hierarkisk nivå 3 (framtida arbete).
- Förfinad modellering av övergångssannolikheter i världsmodellen mellan olika tillstånd hos målobjekt och UAV-svärmen.
- Svärmsplittring för svärmar utan utpekad ledare som agerar individuellt.

12 Referenser

1. Parr, T., Pezzulo, G., Friston, K. J. (2022). *Active inference: The free energy principle in mind, brain, and behavior*. The MIT Press, Cambridge, MA. doi:10.7551/mitpress/12441.001.0001
2. Namjoshi, S. V. (2026). *Fundamentals of Active Inference*. The MIT Press, Cambridge, MA.
3. Buckley, C. L., Kim, C. S., McGregor, S., Seth, A. K. (2017). The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology* **81**:55–79. doi:10.1016/j.jmp.2017.09.004
4. Schubert, J. (2021). Aktiv inferens. FOI-R--5221--SE. [Online] Available: <https://foi.se/rapporter/rapportsammanfattning.html?reportNo=FOI-R--5221--SE>
5. Dempster, A. P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society. Series B* **30**(2):205–247. doi:10.1111/j.2517-6161.1968.tb00722.x
6. Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press, Princeton, NJ. doi:10.1515/9780691214696
7. Hari, S. K. K., Rathinam, S., Darbha, S., Kalyanam, K., Manyam, S. G., Casbeer, D. (2020). Optimal UAV route planning for persistent monitoring missions. *IEEE Transactions on Robotics* **37**(2):550–566. doi:10.1109/TRO.2020.3032171
8. Brown, A., Anderson, D. (2020). Trajectory optimization for high-altitude long-endurance UAV maritime radar surveillance. *IEEE Transactions on Aerospace and Electronic Systems* **56**(3):2406–2421. <https://doi.org/10.1109/TAES.2019.2949384>
9. Hübel, N., Hirche, S., Gusrialdi, A., Hatanaka, T., Fujita, M., Sawodny, O. (2008). Coverage control with information decay in dynamic environments. *IFAC Proceedings* **41**(2):4180–4185. doi:10.3182/20080706-5-KR-1001.00703
10. Ramasamy, M., Ghose, D. (2016). Learning-based preferential surveillance algorithm for persistent surveillance by unmanned aerial vehicles. In *Proceedings of the 2016 International Conference on Unmanned Aircraft Systems*, pp. 1032–1040. IEEE, Piscataway, NJ. doi:10.1109/ICUAS.2016.7502678
11. Chen, J., Baskaran, A., Zhang, Z., Tokekar, P. (2021). Multi-agent reinforcement learning for visibility-based persistent monitoring. In *Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2563–2570. IEEE, Piscataway, NJ. doi:10.1109/IROS51168.2021.9635898
12. Mishra, M., Poddar, P., Agrawal, R., Chen, J., Tokekar, P., Sujit, P. B. (2025). Multi-agent deep reinforcement learning for persistent monitoring with sensing, communication, and localization constraints. *IEEE Transactions on Automation Science and Engineering* **22**:2831–2843. doi:10.1109/TASE.2024.3385412
13. van de Laar, T., Özçelikkale, A., Wymeersch, H. (2021). Application of the free energy principle to estimation and control. *IEEE Transactions on Signal Processing* **69**:4234–4244. doi:10.1109/TSP.2021.3095711
14. Bos, F., Meera, A. A., Benders, D., Wisse, M. (2022). Free energy principle for state and input estimation of a quadcopter flying in wind. In *Proceedings of the 2022 International Conference on Robotics and Automation*, pp. 5389–5395. IEEE, Piscataway, NJ. doi:10.1109/ICRA46639.2022.9812415
15. Pezzato, C., Ferrari, R., Hernández Corbato, C. (2020). A novel adaptive controller for robot manipulators based on active inference. *IEEE Robotics and Automation Letters* **5**(2):2973–2980. doi:10.1109/LRA.2020.2974451
16. Meo, C., Lanillos, P. (2021). Multimodal VAE active inference controller. In *Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2693–2699. IEEE, Piscataway, NJ. doi:10.1109/IROS51168.2021.9636394
17. Meo, C., Franzese, G., Pezzato, C., Spahn, M., Lanillos, P. (2023). Adaptation through prediction: multisensory active inference torque control. *IEEE Transactions on Cognitive and Developmental Systems* **15**(1):32–41. doi:10.1109/TCDS.2022.3156664

18. Baioumy, M., Pezzato, C., Ferrari, R., Corbato, C. H., Hawes, N. (2021). Fault-tolerant control of robot manipulators with sensory faults using unbiased active inference. In *Proceedings of the 2021 European Control Conference*, pp. 1119–1125. IEEE, Piscataway, NJ. doi:10.23919/ECC54610.2021.9654913
19. Çatal, O., Verbelen, T., Van de Maele, T., Dhoedt, B., Safron, A. (2021). Robot navigation as hierarchical active inference. *Neural Networks* **142**:192–204. doi:10.1016/j.neunet.2021.05.010
20. Scholz, F., Gumbsch, C., Otte, S., Butz, M. V. (2022). Inference of affordances and active motor control in simulated agents. *Frontiers in Neurorobotics* **16**:881673. doi:10.3389/fnbot.2022.881673
21. Kawahara, D., Ozeki, S., Mizuuchi, I. (2022). A curiosity algorithm for robots based on the free energy principle. In *Proceedings of the 2022 IEEE/SICE International Symposium on System Integration*, pp. 53–59. IEEE, Piscataway, NJ. doi:10.1109/SII52469.2022.9708819
22. Katahira, K., Kunisato, Y., Okimura, T., Yamashita, Y. (2020). Retrospective surprise: A computational component for active inference. *Journal of Mathematical Psychology* **96**:102347. doi:10.1016/j.jmp.2020.102347
23. Wakayama, S., Ahmed, N. (2023). Active Inference for autonomous decision-making with contextual multi-armed bandits. In *Proceedings of the 2023 IEEE International Conference on Robotics and Automation*, pp. 7916–7922. doi:10.1109/ICRA48891.2023.10160593
24. Wakayama, S., Candela, A., Hayne, P., Ahmed, N. (2025). Active inference for bandit-based autonomous robotic exploration with dynamic preferences. *IEEE Transactions on Robotics* **41**:3841–3851. doi:10.1109/TRO.2025.3577041
25. Levchuk, G., Pattipati, K., Fouse, A., Serfaty, D. (2017). Application of free energy minimization to the design of adaptive multi-agent teams. In *Proceedings SPIE 10206, Disruptive Technologies in Sensors and Sensor Systems*, 102060E. doi:10.1117/12.2263542
26. Levchuk, G., Fouse, A., Pattipati, K., Serfaty, D., McCormack, R. (2018). Active learning and structure adaptation in teams of heterogeneous agents. In *Proceedings SPIE 10653, Next-Generation Analyst VI*, 1065305. doi:10.1117/12.2305875
27. Levchuk, G., Pattipati, K., Serfaty, D., Fouse, A., McCormack, R. (2019). Active inference in multiagent systems: Context-driven collaboration and decentralized purpose-driven team adaptation. In *Artificial Intelligence for the Internet of Everything*, pp. 67–85. Academic Press, Cambridge, MA. doi:10.1016/B978-0-12-817636-8.00004-1
28. Pezzato, C., Corbato, C. H., Bonhof, S., Wisse, M. (2023). Active inference and behavior trees for reactive action planning and execution in robotics. *IEEE Transactions on Robotics* **39**(2):1050–1069. doi:10.1109/TRO.2022.3226144
29. Baltieri, M., Buckley, C. L. (2019). Nonmodular architectures of cognitive systems based on active inference. In *Proceedings of the 2019 International Joint Conference on Neural Networks*, pp. 1–8. doi:10.1109/IJCNN.2019.8852048
30. Nozari, S., Krayani, A., Marin, P., Marcenaro, L., Gomez, D. M., Regazzoni, C. (2024). Exploring action-oriented models via active inference for autonomous vehicles. *EURASIP Journal on Advances in Signal Processing* **2024**:92. doi:10.1186/s13634-024-01173-9
31. Sawada, H., Ohata, W., Tani, J. (2025). Human-robot kinaesthetic interactions based on the free-energy principle. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **55**(1):366–379. doi:10.1109/TSMC.2024.3481251
32. Matsumura, T., Esaki, K., Yang, S., Yoshimura, C., Mizuno, H. (2024). Active inference with empathy mechanism for socially behaved artificial agents in diverse situations. *Artificial Life* **30**(2):277–297. doi:10.1162/artl_a_00416

33. Esaki, K., Matsumura, T., Minusa, S., Shao, Y., Yoshimura, C., Mizuno, H. (2024). Dynamical perception-action loop formation with developmental embodiment for hierarchical active inference. In Buckley, C. L., et al. *Active Inference. IWA I 2023. Communications in Computer and Information Science*, vol 1915, pp. 14–28. Springer, Cham. doi:10.1007/978-3-031-47958-8_2
34. Kullback, S., Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics* 22(1):79–86. doi:10.1214/aoms/1177729694
35. Smets, P., Kennes, R. (1994). The transferable belief model. *Artificial Intelligence* 66(2):191–234. doi:10.1016/0004-3702(94)90026-4
36. Schubert, J., Kamrani, F., Gustavi, T. (2025). Active inference for an intelligent agent in autonomous reconnaissance missions. *Presented at the 6th International Workshop on Active Inference (IWA I 2025)*, Montreal, Canada, 15–17 October 2025, pp. 1–24, arXiv:2510.17450. doi:10.48550/arXiv.2510.17450

Bilaga A Sensor

A.1 Sensoregenskaper

Agentens sensor antas vara en intelligent bildsensor, utformad för markspaning. När agenten är aktiv tar sensorn flygbilder med förutbestämda tidsintervaller och använder programvara för bildanalys i realtid för att bearbeta dem. Bildanalysprocessen antas bestå av tre huvudsteg:

1. Lokalisering av möjliga målobjekt i bilderna.
2. Klassificering av målkandidaterna, där bildbehandlingsalgoritmen returnerar (för varje målkandidat): (i) en klassificering som *Mål* (T) eller *Inget Mål* (F) och (ii) ett värde i intervallet $[0, 1]$ som anger klassificeringskonfidensen.
3. Endast för målkandidater klassificerade som *Mål*: Uppskattning av målkandidaternas positioner. För enkelhetens skull antar vi att bildsensorn har direkt tillgång till agentens GPS-koordinater och dess orientering, vilket gör det möjligt för sensorn att beräkna positionsuppskattningar direkt i det globala koordinatsystemet.

Följaktligen består utdata från den intelligenta bildsensorn av:

1. Uppskattade positioner för alla objekt som klassificeras som *Mål* (om några).
2. För varje objekt som klassificeras som *Mål* (om något) gäller den grundläggande tilltron att objektet i fråga faktiskt motsvarar ett verkligt mål, dvs. klassificeringskonfidensen.

Dessutom antas sensorn ha följande egenskaper:

- Sensorn har ett begränsat synfält inom vilket den tillförlitligt kan detektera och klassificera mål. På marknivå beror storleken på det cirkulära tvådimensionella området som sensorn kan observera på både sensorns egenskaper och agentens flyghöjd. För fasta flyghöjder kan radien för det observerade området antas vara konstant och betecknas med s_r .
- Det finns en icke-försumbar risk att sensorn missar att detektera mål som finns inom det observerade området.
- Målpositionsuppskattningarna, beräknade av den intelligenta bildsensorn, är något oprecisa. De är dock kända för att vara symmetriskt fördelade runt de verkliga målpositionerna med varians, σ_{pos}^2 .

Avsnitt A.2 förklarar hur sensormodellen implementeras i simuleringsmiljön. Avsnitt 6.2 visade hur sensorutdata omvandlas till det indataformat som krävs av algoritmerna som beräknar VFE över det övervakade området.

A.2 Implementering av sensormodellen

En förenklad modell av den intelligenta bildsensorn som beskrivs i avsnitt A.1 implementeras i simuleringsmiljön och används i experimenten. Modellsensorn genererar utdata direkt utifrån sin kunskap om de faktiska målpositionerna i simuleringsmiljön. Dess beteende baseras på förutsättningen att den verkliga bildsensorn är tränad för objektklassificering i flygbilder och att en verklig sensors förmåga att lokalisera och klassificera potentiella målobjekt i en bild främst beror på avståndet mellan sensorn och objektet. De mest anmärkningsvärda förenklingarna i sensorimplementeringen är:

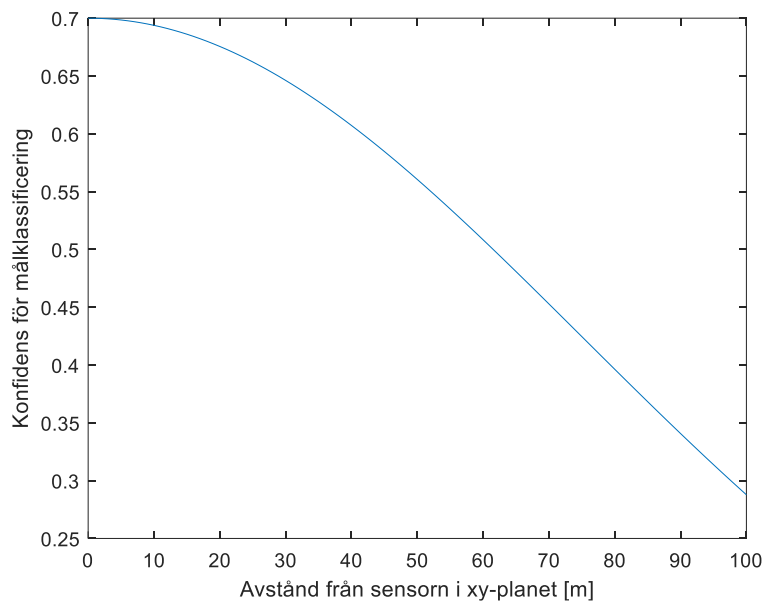
- Alla mål som hamnar inom sensorns räckvidd detekteras med sannolikhet 1.
- Sensorn genererar inga falska måldetektioner.
- Målpositionsuppskattningarna, genererade av den implementerade sensormodellen, är brusfria.

Utdata från modellsensorn genereras enligt följande:

- Uppskattade positioner för de detekterade målen ersätts med de faktiska målpositionerna.
- Istället för klassificeringskonfidenser som genereras i målklassificeringsprocessen returnerar mockup-sensorn *simulerade klassificeringskonfidenser* för varje detekterat mål, $k = 1, \dots, K$. En gaussisk funktion används för att modellera minskningen av klassificeringskonfidensen $C_{T,k}$ som en funktion av avståndet d_k mellan sensorn och ett detekterat mål, vilket visas i ekvation (35). Vi har,

$$C_{T,k}(d_k) = C_{Tmax} \cdot e^{-\frac{d_k^2}{2\sigma_T^2}} \quad (35)$$

där σ_T bestämmer minskningshastigheten för klassificeringskonfidensen (se figur 16), medan $C_{Tmax} \leq 1$ är en skalfaktor som bestämmer det högsta uppnåeliga värdet för $C_{T,k}$.



Figur 16. En gaussisk funktion modellerar minskningen av klassificeringskonfidensen i takt med att avståndet från sensorn ökar. Kurvan har $a_T = 0,7$, $\sigma_T = \frac{3}{4}s_r$ (med $s_r = 100$ m) och motsvarar positiv målklassificering.

I verkligheten påverkar förhållanden och terräng sannolikheten för objekt-detektering, klassificeringsprecision och klassificeringskonfidens. Beroendet av $C_{T,k}$ på sådana externa faktorer modelleras dock inte i den simulerade sensorn.

