

FREDRIK NÄSSTRÖM, DAVID BERGSTRÖM,  
FREDRIK BISSMARCK, PER GRAHN,  
DAVID GUSTAFSSON, JÖRGEN KARLHOLM





Fredrik Näsström, David Bergström,  
Fredrik Bissmarck, Per Grahn,  
David Gustafsson, Jörgen Karlholm

# Prestandamått för sensorsystem

Titel	Prestandamått för sensorsystem
Title	Performance measures for sensor systems
Rapportnr/Report no	FOI-R--4139--SE
Månad/Month	November
Utgivningsår/Year	2015
Antal sidor/Pages	44 p
ISSN	1650-1942
Kund/Customer	FM
Forskningsområde	7. Sensorer och signaturanpassning
FoT-område	Sensorer och signaturanpassning
Projektnr/Project no	E54571
Godkänd av/Approved by	Martin Rantzer
Ansvarig avdelning	Sensor och TK-system

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk. All form av kopiering, översättning eller bearbetning utan medgivande är förbjuden.

This work is protected under the Act on Copyright in Literary and Artistic Works (SFS 1960:729). Any form of reproduction, translation or modification without permission is prohibited.

*Omslagets bildcollage: Översta raden från vänster till höger, den vänstra bilden föreställer en Stridsvagn 122 och helikopter 9 (Foto: Försvarmakten) och den högra bilden föreställer JAS 39 Gripen (Foto: Johan Lundahl/Försvarmakten). Nedersta raden från vänster till höger, den vänstra bilden föreställer TGB-16 (Foto: Daniel Rydberg/Försvarmakten) och den högra bilden föreställer Korvett Visby (Foto: Försvarmakten)*

## Sammanfattning

För att kunna bedöma det sammantagna värdet av ett sensorsystem måste sensorsystemets prestanda värderas i taktiskt relevanta situationer. Det är då viktigt att kunna mäta prestanda för både enskilda sensorer och för hela det värderade systemet i olika miljöer, väderförhållanden och tidpunkter.

Denna rapport från projektet ”Värdering av sensorsystem” beskriver metoder och verktyg som kan användas för att mäta prestanda för optroniksystem och radarsystem. Ett urval av dessa metoder och verktyg har testats för att analysera deras användbarhet.

De beskrivna metoderna och verktygen kan användas av Försvarmakten vid värderingsarbeten och studier. De kan även användas för att ge underlag vid utveckling av eller inköp alternativt uppgradering av nya sensor- och systemförmågor.

För att utveckla förmågan att värdera sensorsystem vidareutvecklar projektet också värderingsmetoden COAT/TVS.

Nyckelord: värdering, sensorsystem, sensorprestanda

## Summary

In order to assess the overall value of a sensor system the performance must be measured in tactically relevant situations. It is important to be able to measure performance of both individual sensors and for the entire sensor system in various environments, weather conditions and time periods.

This report from the project "Assessment of sensor systems" describes methods and tools that can be used to measure performance of optronic systems and radar systems. A selection of these methods and tools have been tested to analyze their usefulness.

The described methods and tools can be used by the Swedish Armed Forces in assessments and studies. They can also be used to provide a basis for development, purchase or upgrade of new sensor and system capabilities.

To develop the ability to assess sensor systems the project further develops an assessment method COAT / TVS.

Keywords: assessment, sensor system, sensor performance

## Innehåll

<b>1</b>	<b>Inledning</b>	<b>7</b>
1.1	COAT/TVS-metoden .....	8
<b>2</b>	<b>Estimering av prestanda för optroniksystem</b>	<b>9</b>
2.1	NIIRS.....	9
2.1.1	General Image Quality Equation (GIQE) .....	11
2.2	Kvalitetsmått baserade på statistik för naturliga scener .....	14
2.3	Prestandamått för manuell måldiskriminering.....	20
2.3.1	Johnson-kriteriet .....	20
2.3.2	Targeting Task Performance (TTP).....	21
2.3.3	Thermal Range Model (TRM).....	26
<b>3</b>	<b>Estimering av prestanda för radarsystem</b>	<b>28</b>
3.1	NIIRS för SAR-radarbilder .....	28
3.2	Kvalitetsmått för SAR-bilder.....	30
<b>4</b>	<b>Estimering av prestanda för signalbehandling</b>	<b>32</b>
4.1	Värdering av målföljningsprestanda hos sensorsystem .....	32
4.2	Bayesiansk skattning för estimering av måldiskrimineringsförmåga	36
<b>5</b>	<b>Estimering av prestanda för sensorsystem</b>	<b>38</b>
5.1	Mätningar .....	38
5.2	Simuleringsbaserade metoder .....	38
<b>6</b>	<b>Slutsatser och fortsatt arbete</b>	<b>40</b>
<b>7</b>	<b>Referenser</b>	<b>41</b>

## Förkortningar

CLEAR	CLassification of Events, Activities and Relationships
COAT	Communication AssessmentT
CONOPS	Concept Of Operations
GMTI	Ground Moving Target Indication
GSD	Geometrisk markupplösningen
FOV	Field of view
LWIR	Long-wavelength infrared
MSSLab	MultiSensorSimuleringsLab
MTF	Modulation Transfer Function
MWIR	Mid-wavelength infrared
NIIRS	National Imagery Interpretability Rating Scale
NV-IPM	Night Vision Integrated Performance Model
OSPA	Optimal Subpattern Assignment
ROE	Rules Of Engagement
RER	Relative Edge Response
RPAS	Remotely Piloted Aircraft System
SAR	Synthetic Aperture Radar
SWIR	Short-wavelength infrared
TRM	Thermal Range Model
TTP	Targeting Task Performance
TVS	Taktisk Värdering av Sensorsystem
UAV	Unmanned Aerial Vehicles



# 1 Inledning

En svårighet vid värdering av komplexa militära system är att information om sensor-systemen ofta är otillräcklig. Vid en värdering av befintliga system finns ofta tekniska specifikationer eller spelkort, men för internationella och framtida sensorsystem är sådan information ofta bristfällig. Insamling av information kan göras på flera olika sätt såsom mätningar, simuleringar, analytiska metoder, intervjuer av erfarna personer etc. Dessa metoder att samla in information är alla olika resurskrävande och ger information med olika tillförlitlighet.

Syftet med denna rapport är att beskriva olika metoder och verktyg som kan användas vid värdering av optroniksystem och radarsystem. De beskrivna metoderna och verktygen kan användas vid värderingsarbeten och studier av olika scenarier, typsituationer och målbilder som är intressanta för Försvarmakten. De kan även användas för att ge underlag vid utveckling av eller inköp alternativt uppgradering av sensorsystem, liksom vid behovsinventering av nya sensor- och systemförmågor.

Vid värdering av ett sensorsystem finns det ofta yttre faktorer som påverkar möjligheten till inhämtning av information, såsom krav på tillförlitlighet, kostnad och tidsaspekter, dvs. hur snabbt värderingen måste komma fram. Detta påverkar vilken eller vilka metoder som är lämpliga att använda eftersom olika metoder är olika resurskrävande och resulterar i information med olika kvalitet.

Beroende på vad som skall värderas hos sensorsystemet kommer olika information om dess sensorprestanda att behövas. Sådan information kan till exempel vara sannolikhet för *måldetektion*, *måligenkänning*, *målidentifiering* och *positionering* av mål. För dessa begrepp har vi använt samma definition som används i STANAG No. 3769.

- Detektion (eng. Detection): I bildtolkning ska man kunna upptäcka förekomsten av ett objekt, men utan att kunna känna igen objektet.
- Igenkänning (eng. Recognition): På denna nivå ska målen delas in i mer detaljerade klasser, såsom stridsvagn, lastbil, personbil, spaningsflygplan, etc. För ett fordon ska man kunna bestämma formen och förhållandet mellan längd och bredd.
- Identifiering (eng. Identification): På denna nivå ska målen delas in i modelltyp såsom T-72, M1-A1, MiG-21 J, etc. På artillerisystem ska längden på eldrören, mynningsbroms, rekylbroms, etc. kunna bestämmas för att systemet ska identifieras till en specifik klass.
- Teknisk Analys (eng. Technical Analysis): Teknisk analys är definierad som förmågan att exakt beskriva en funktion, objekt eller komponent.

För att ta ett enkelt exempel så räknas proceduren att dela in människor i klasserna man, kvinna och barn som igenkänning medan identifiering handlar om att se skillnad på Kalle, Stina och Martin.

Kravet på upplösningen hos sensordata skiljer sig åt mycket beroende på vad man studerar, t.ex. igenkänning av en bro kräver en upplösning på ca 5 meter medan igenkänning av artilleri krävs en upplösning på 0.5 meter. Väder och tidpunkt på dygnet har stor påverkan på prestanda för flera sensorer.

Sensorprestanda beror av många variabler som:

- Målobjekt: konfiguration, signatur, kamouflage, skyl, mål på marsch eller stillastående etc.
- Väder: temperatur, molnbas, olika typer av nederbörd etc.

- Ljushöjdhållanden: tid på året och dygnet, aktuellt väder och plats på jorden.
- Vegetation: fördelningar av träd och buskar och dess årstidsförändringar.
- Terräng: bebyggelse, öppen, småbruten, betäckt. Terrängen kan dessutom vara platt, kuperad eller bergig, samt ha inslag av sjöar och vattendrag.
- Datainsamling: observationstid, avstånd, höjd, aspektvinkelförändring på målet under observationen, hastighet på sensorplattformen etc.
- Hot och störning mot sensorn och sensorbäraren.
- Handlingsregler för uppdraget (ROE - Rules Of Engagement)

## 1.1 COAT/TVS-metoden

I detta projekt har en värderingsmetod som kallas COAT/TVS<sup>1</sup> använts. COAT skapades för att möta Försvarens och FMV:s behov av värdering av tekniska kommunikationsfunktioner, tjänster och grad av sårbarhet i olika kommunikationsnätverk. Metoden skapar en spårbar och kvalitetsstämplad värdering av kommunikationssystemet utifrån kundens uttalade behov av tjänster. Metoden utgår ifrån kundens eller uppdragsgivarens frågeställning (behov) och bryter sedan ner densamma till indirekta behov på underliggande teknisk (system)nivå. Detta görs för att finna de relevanta parametrarna, exempelvis tekniska, som krävs för att kunna hantera problemställningen. Grundmetoden vidareutvecklades senare så att den skulle kunna tillämpas på sensorsystem och det är denna utvecklade variant som vi vidareutvecklar i detta projekt.

COAT/TVS består av ett ramverk av processer, verktyg och metoder lämpliga för en behovsorienterad strukturerad värdering av komplexa telekommunikationssystem i ett taktiskt och operativt sammanhang. Metoden värderar helheten – inte bara tekniken – och skapar en struktur och en spårbarhet i arbetet. COAT skall ses som ett ramverk som är enkelt och flexibelt och därför kan metoden användas både för stora och små värderingar och för olika miljöer. Metoden, som redovisas i sin helhet i [1], är modulärt uppbyggd och består av fyra processer:

- Kunddialog
- Planering och modellering
- Bearbetning
- Sammanställning

Därutöver finns en övergripande process för att säkerställa att arbetet dokumenteras på ett tillfredsställande sätt.

I processen Bearbetning bryts kundbehovet ner till underliggande behov. Nedbrytningen representeras naturligt i en trädstruktur (behovsträd) som går från de övergripande kundbehoven, ner till mätbara faktorer. Dessa faktorer kan ibland bestämmas med hög precision, men ibland blir man tvungen att göra grova uppskattningar eller rena bedömningar. När värden erhållits förs dessa sedan uppåt i behovsträdet igen och värderas på varje plan utifrån de antaganden, förutsättningar och vikter de getts.

Tidigare värderingsarbete [2] har visat att det är mycket viktigt att kunna estimeras olika prestandamått för sensorsystem. Av den anledningen har vi i denna rapport fokuserat på olika metoder och verktyg för just detta.

---

<sup>1</sup> COAT/TVS = COmmunication AssessmenT / Taktisk Värdering av Sensorsystem

## 2 Estimering av prestanda för optroniksystem

Det finns något olika traditioner på olika nivåer av informationsinhämtning vad gäller prestandamått för optroniksystem. För taktiska sensorer, vilket inkluderar markbundna och lågflygande EO- och IR-sensorer, används traditionellt sannolikheter för detektion, igenkänning och identifiering som prestandamått. För strategiska sensorer, vilket inkluderar medel- och högtflygande plattformar och satelliter, är det däremot brukligt att använda NIIRS-skalan (National Imagery Interpretability Rating Scale) [3].

### 2.1 NIIRS

För flyg- och satellitburna tillämpningar användes länge *skala* och *upplösning* som mått på prestanda för ett sensorsystem från en viss flyghöjd. Skala uttrycktes ofta i formen av exempelvis 1:1000, som angav att 1 cm i bilden motsvarade 10 m i verkligheten. Upplösning angavs ofta som antalet upplösta parallella linjer som kunde urskiljas på en markmåltavla. Dessa enkla och trubbiga mått hade dock inte förmågan att kunna beskriva tolkningsbarheten hos de bilder som sensorsystemen producerade, pga. att måtten inte fångade in alla aspekter som påverkar. I början av 1970-talet påbörjade därför en amerikansk myndighetskommitté (IRARS) utvecklingen av NIIRS, som var en subjektiv skala som bättre skulle kunna kommunicera den mängd relevant information som en bildtolk kunde och inte kunde extrahera ur en specifik bild. Denna skala har idag blivit en internationell standard för hur både kvalitet i flygbilder anges samt för hur prestanda hos flygande sensorsystem måttsätts och kravställs.

NIIRS är en *subjektiv* uppgiftsbaserad skala för bildkvalitet och består av 10 steg eller nivåer från 0 (sämst) till 9 (bäst) med flera olika tolkningsuppgifter eller kriterier för varje steg. För varje skalsteg ökar svårighetsgraden i de tolkningsuppgifter som kan utföras. En bildanalytiker anger en NIIRS-nivå för en bild och indikerar på detta sätt att de uppgifter som tillhör denna nivå samt de underliggande nivåerna kan utföras i den aktuella bilden. Skalan finns sedan 20 år tillbaka framtagen för de visuella, termiskt infraröda och radaravbildande våglängdsområdena samt i en version för ett multispektralt avbildande system (som har fyra specifika band: RGB + NIR). Noteras bör att visuella NIIRS är framtagen för monokromatiskt avbildande sensorsystem i det visuella området (s.k. pankromatiska sensorsystem) och alltså inte omfattar färgbilder. Skalan formulerades ursprungligen med fokus på militära tolkningsuppgifter men har under åren även kompletterats med kriteriebeskrivningar bättre anpassade mot civila tillämpningar i takt med att fjärranalys fått allt större spridning inom exempelvis kartering, jordbruksövervakning, miljöövervakning, katastrofhantering, osv. [4, 5].

I varje steg på skalan är varje kriterium eller tolkningsuppgift formulerad med två huvudelement: en *tolkningsnivå* (ex. detektion, igenkänning eller identifiering) samt ett *objekt* (ex. människa, bil eller flygplan). Till dessa två kommer ofta (men inte alltid) även en *bakgrundskontext* (ex. stadsbebyggelse, flygfält, osv.) samt ett *kvalificerande objektexempel* (ex. AK47, T-72 eller F-15). Kriterierna bör vara formulerade efter tidsaktuella tolkningsuppgifter och objekt och ett kontinuerligt underhåll av skalorna krävs därför för att hålla dem uppdaterade och användningsbara. De senaste versionerna är från 1994 och 1996 för visuella respektive infraröda EO-sensorer. Tabell 1 visar ett utdrag ur dessa versioner för att exemplifiera NIIRS-skalans uppbyggnad [4].

Ursprungligen fanns även parallellt med nivåerna i NIIRS-skalan en approximativt logaritmisk upplösningsskala, där varje NIIRS-nivå också hade ett mått (eller snarare ett måttintervall) för upplöst avstånd på marken. Detta avståndet, betecknat GRD (Ground Resolved Distance), utgjorde storleken på det minsta mönstret eller minsta avståndet

mellan två objekt som kunde lösas upp och halverades ungefärligen för varje nytt steg uppåt i NIIRS-skalan (nivå 6 krävde ungefärligt dubbel så hög markupplösning som nivå 5, osv.). Begreppet GRD användes flitigt i den äldre analoga våtfilmseran men numera rekommenderas NIIRS-skalan att användas utan denna upplösningsskala, eftersom måttet inte är anpassat till dagens moderna digitalt samplande EO-system.

Tabell 1. Utdrag ur NIIRS-tabeller för visuella och termiskt infraröda EO-sensorer [4].

Level	Visible NIIRS	Infrared NIIRS
0	Interpretability of the imagery is precluded by obscuration, degradation, or very poor resolution	Interpretability of the imagery is precluded by obscuration, degradation, or very poor resolution
1	Detect a medium-sized port facility and/or distinguish between taxi-ways and runways at a large airfield.	Distinguish between runways and taxiways on the basis of size, configuration or pattern at a large airfield. Detect a large (e.g., greater than 1 square kilometer) cleared area in dense forest.
2	Detect large hangars at airfields. Identify an SA-5 site based on road pattern and overall site configuration. Detect large buildings (e.g., hospitals, factories).	Detect large aircraft (e.g., C-141, 707, BEAR, CANDID, CLASSIC). Detect individual large buildings (e.g., hospitals, factories) in an urban area.
3	Identify the wing configuration (e.g., straight, swept, delta) of all large aircraft (e.g., 707, CONCORD, BEAR, BLACKJACK). Identify a large surface ship in port by type (e.g., cruiser, auxiliary ship, non-combatant/merchant).	Distinguish between large (e.g., C-141, 707, BEAR, A300 AIRBUS) and small aircraft (e.g., A-4, FISHBED, L-39). Distinguish between large (e.g., greater than 200 meter) freighters and tankers.
4	Identify all large fighters by type (e.g., FENCER, FOXBAT, F-15, F-14). Detect an open missile silo door.	Identify the wing configuration of small fighter aircraft (e.g., FROGFOOT, F-16, FISHBED). Identify individual closed cargo hold hatches on large merchant ships.
5	Identify radar as vehicle-mounted or trailer-mounted. Identify, by type, deployed tactical SSM systems (e.g., FROG, SS-21, SCUD).	Distinguish between single-tail (e.g., FLOGGER, F-16, TORNADO) and twin-tailed (e.g., F-15, FLANKER, FOXBAT) fighters.
6	Identify the spare tire on a medium-sized truck. Distinguish between SA-6, SA-11, and SA-17 missile airframes.	Distinguish between thermally active tanks and APCs. Identify missile tube hatches on submarines. Identify outdoor tennis courts.
7	Detect the mount for antitank guided missiles (e.g., SAGGER on BMP-1). Identify individual rail ties.	Identify automobiles as sedans or station wagons. Identify the missile transfer crane on a SA-6 transloader.
8	Identify a hand-held SAM (e.g., SA-7/14, REDEYE, STINGER). Identify windshield wipers on a vehicle.	Identify limbs (e.g., arms, legs) on an individual. Detect closed hatches on a tank turret.
9	Identify screws and bolts on missile components. Detect individual spikes in railroad ties.	Identify access panels on fighter aircraft. Identify turret hatch hinges on armored vehicles.

Förutom att NIIRS används av bildanalytiker för att måttsätta kvalitet i bilder används den också av tillverkare av EO-system vid kravställning inför framtagning av ny eller modifierad hårdvara samt av systeminköpare inför upphandling och/eller efter leverans av nya EO-system som ska uppfylla en viss spanings- eller övervakningsförmåga. NIIRS är också ofta en viktig utgångspunkt vid planering av ett spaningsuppdrag, för

säkerställning att uppdragsuppgiften kan lösas med god sannolikhet och med god säkerhet för spaningsplattform och pilot.

### 2.1.1 General Image Quality Equation (GIQE)

I takt med att NIIRS blev ett mer och mer standardiserat mått för bildkvalitet och spaningsförmåga hos ett avbildande system uppstod också ett behov av att kunna *förutsäga NIIRS-nivån* för ett EO-system, redan innan det byggdes och testades eller innan det skickades ut på ett spaningsuppdrag. Bildkvalitetsekvationen GIQE (General Image Quality Equation) utvecklades i slutet av 1980-talet men blev inte officiellt förrän 1994, då den släpptes till allmänheten i syfte att tillhandahålla ett verktyg med vars hjälp prestanda kunde predikteras. Som exempel har EO-sensorer hos både Predator och Global Hawk krävts och tagits fram mha. NIIRS och GIQE. Ekvationen togs fram empiriskt från tester med ett stort antal bilder och bildanalytiker och har validerats och förfinats över åren genom ytterligare perceptionstester. Den senaste versionen av GIQE för visuella NIIRS (version 4.0) är från 1997 och beskrivs enligt [6].

$$NIIRS = 10.251 + a \log_{10} GSD + b \log_{10} RER + 0.656H - 0.344 \frac{G}{SNR}, \quad (1)$$

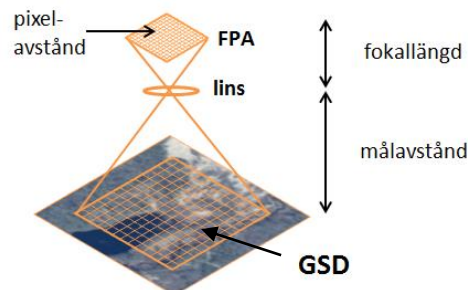
medan den infraröda versionen (som är applicerbar både i MWIR och LWIR) utkom tre år senare (2000): [7]

$$NIIRS = 10.751 + a \log_{10} GSD + b \log_{10} RER + 0.656H - 0.344 \frac{G}{SNR}. \quad (2)$$

I ekv. (1) och (2) är den första termen en konstant och den enda termen som skiljer de två ekvationerna åt. De övriga termerna från vänster till höger kommer sedan i fallande storleksordning från större och mer signifikanta till mindre. Termerna representerar i tur och ordning *geometrisk spatial upplösning* (den andra termen), *optisk spatial upplösning* (den tredje och fjärde termen) samt *signal/brus-förhållandet* (den femte och avslutande termen).  $a$  och  $b$  i ekvationerna är två konstanter som beror på värdet hos  $RER$  (Relative Edge Response), där  $a = 3.32, b = 1.559$  om  $RER \geq 0.9$  och där  $a = 3.16, b = 2.817$  om  $RER < 0.9$ .

I den andra termen ingår  $GSD$  (Ground Sampling Distance), vilken är pixelupplösningen på mark och beräknas geometriskt (se illustration i figur 1) ur målavståndet  $R$ , betraktelsevinkeln  $\alpha$  (vinkeln mellan markplan och siktlinjen), optikens fokallängd  $f$  och pixelavståndet  $p$  enligt

$$GSD = \frac{pR}{f \cos \alpha}. \quad (3)$$



Figur 1. Den geometriska markupplösningen ( $GSD$ ) är pixelavståndet i det avbildande systemet projicerat genom linsen på marken och beror av vald fokallängd, målavstånd och betraktelsevinkeln (här noll grader, dvs. tittandes rakt nedåt).

I den tredje termen används måttet RER, vilken är centrumlutningen hos systemets relativa kantrespons (se figur 2) och kan härledas ur det avbildande systemets MTF (modulationsöverföringsfunktion, se vidare diskussion under avsnitt 2.3) enligt

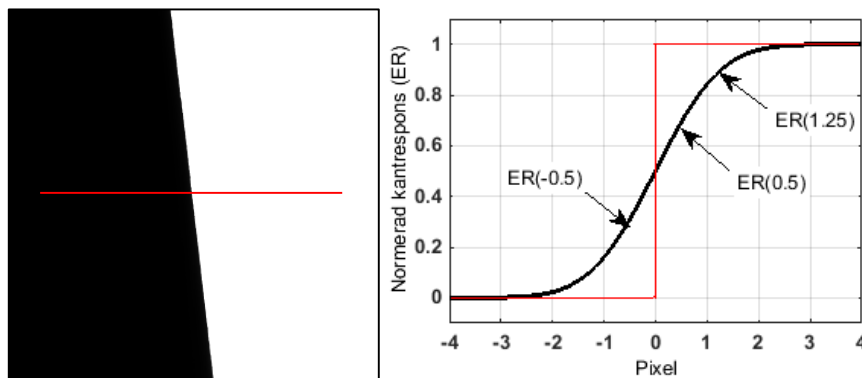
$$RER = ER(0.5) - ER(-0.5). \quad (4)$$

I ekv. (4) är  $ER$  kantresponsen som kan beräknas ur systemets MTF enligt

$$ER(d) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\xi_{cutoff}} \frac{MTF(\xi)}{\xi} \sin(2\pi\xi d) d\xi, \quad (5)$$

där  $d$  är respospositionen från kantens centrum,  $\xi$  är spatiala frekvensen (linje- eller punkttätheten) och  $\xi_{cutoff}$  är systemets cutoff-frekvens (dvs. den högsta möjliga frekvensen som systemet kan överföra).

$H$  i den fjärde termen i ekv. (1) och (2) är ett mått på kantresponsens överskjutning, pga. eventuell pålagd *kantskärpning* i signal- och bildbehandling (s.k. MTF-kompensering eller *MTFC*). Det är definierat som maxvärdet av kantresponsen i intervallet 1 till 3 pixlar från kantcentrum, förutom om kantresponsen är monotoniskt ökande varvid den ges av kantresponsens värde 1,25 pixlar från kanten (se figur 2).



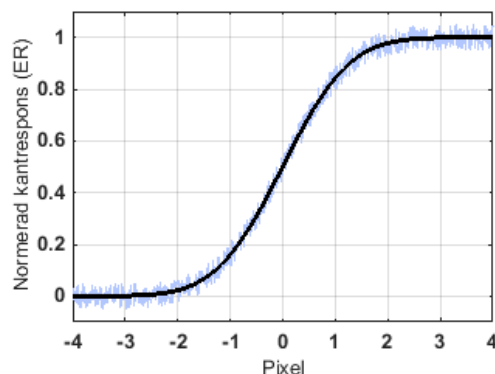
Figur 2. Till vänster visas kantresponsen för ett avbildande system som fås genom att analysera hur kanten mellan två olika konstanta grånivåer avbildas i systemet. Till höger visas att en i verkligheten perfekt skarp kant (röd linje) i ett avbildande system blir utsmetad till en oskarp kant (svart linje) pga. olika kontrastförsämrande effekter.

Om  $GSD$ ,  $RER$  och  $H$  skiljer sig åt i x- och y-led, beräknas dessa som geometriska medelvärden innan de sätts in i (1) och (2).

I den sista termen i ekv. (1) och (2) är  $SNR$  signal-brusförhållandet, dvs. kvoten mellan signalkontrasten mellan mål och bakgrund och signalbruset (se figur 3).  $G$  är en tillkommande brusförstärkningsfaktor, som uppstår pga. eventuell pålagd kantskärpning (s.k. *MTFC*). Om ingen *MTFC* används är den identiskt lika med ett. Den beräknas enligt

$$G = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (MTFC_{ij})^2} \quad (6)$$

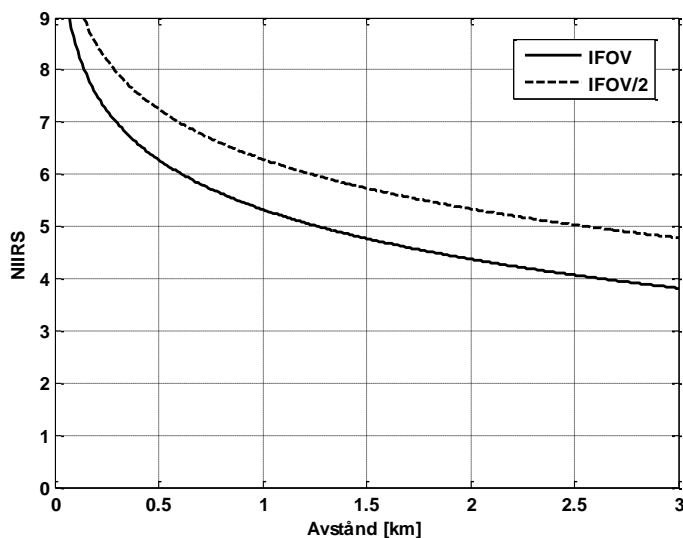
där  $MTFC_{ij}$  är kärnan (matrisen) för kantskärpningen. Den sista termen i GIQE kan vanligtvis försummas om  $SNR > 10$  (och ingen kantskärpning är applicerad), men får en större icke-försumbar negativ effekt då den börjar närma sig detektionsgränsen ( $SNR = 1$ ) och blir så småningom delvis eller helt dominerande om den kraftigt understiger denna gräns.



Figur 3. Signal/brus-förhållandet erhålls genom att ta kvoten mellan medelkantresponsens (svart linje) högsta och lägsta värde (här 0 och 1 pga. normering) och medelvärdet av standardavvikelseerna för variationerna (blå linje) kring dessa värden.

Ekv. (1) och (2) har validerats mot försök med bildtolkare och korrelerar relativt väl med observerade NIIRS. Förklaringsgraden  $R^2$  (ett mått på korrelation mellan 0 och 1, där 1 är perfekt överensstämmelse) är 0.934 för den visuella versionen, med ett standardfel på 0.307. Motsvarande  $R^2$  för den infraröda versionen är 0.8, med ett standardfel på 0.38. För ytterligare detaljer kring GIQE hänvisas till de angivna referenserna (alternativt redogörelsen i [5]).

Figur 4 visar exempel på predikterad NIIRS för två avbildande EO-system i IR-området. I detta fall antogs en VGA-upplösande sensor (640×480 pixlar) med två olika synfält,  $24^\circ \times 18^\circ$  resp.  $12^\circ \times 9^\circ$ , med antagande om en gaussisk punktspridningsfunktion med FWHM (Full Width Half Maximum) lika med 1 pixel. Dessa antaganden innebar RER=0.76, H=1.0 samt SNR=100 (dvs. goda signal/brus-förhållanden) i GIQE-ekvationen. Ingen MTFC antogs applicerad, dvs.  $G=1$ .



Figur 4. Exempel på två IR-avbildande system med predikterad NIIRS-prestanda enligt GIQE som funktion av avstånd (räckvidd). De två kurvorna visar exempel på den NIIRS-prestandavinst (ca 1 NIIRS-steg) som kan erhållas om synfältet halveras (markupplösningen dubblas) och övriga parametrar i GIQE antas opåverkade. Obs! Det minskade synfältet medför dock en fjärdedels yttäckning, vilket förstås måste beaktas i en helhetsvärdering av sensorsystemet.

En nackdel med GIQE i dess nuvarande utformning är att den inte tar hänsyn till eventuell datakomprimering eller det presentationsmedium som används. Komprimering kan ibland krävas för effektiv dataöverföring med en bandbegränsad kommunikationslänk, vilket kan introducera kvalitetsförsämringar i överförd bilddata. Typ av skärm eller display som en bildtolk använder och de ljusförhållanden som råder i rummet där analysen sker kan också påverka slutresultatet av en tolkningsuppgift. En tredje nackdel är att den endast kan hantera stillbilder. På senare tid har dock en kompletterande NIIRS-skala tagits fram för rörliga bildsekvenser, kallad Video-NIIRS eller V-NIIRS [8]. En motsvarande bildkvalitetsekvation är under utveckling och kallas MIQE (Motion Imagery Quality Equation) [9]. Ekvationen är dock inte lika väl validerad som GIQE och har därför inte nått samma status och acceptans ännu.

## 2.2 Kvalitetsmått baserade på statistik för naturliga scener

Att automatiskt kunna bedöma kvaliteten på sensordata och olika typer av degenereringar hos bilder är viktigt för t.ex. automatiska system. Det är idag även möjligt att bedöma vilken typ av degenereringen som en bild har degenererats av. I följande avsnitt ges en översikt om bildbaserade metoder för automatisk bildkvalitetsbedömning (Image Quality Assessment, IQA). Teorin och metoderna som presenteras i avsnittet bygger på bildkvalitetsbedömning hos visuella bilder [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21]. Här presenteras de grundläggande principerna och komponenterna för automatisk bildkvalitetsbedömning. Tillämpningarna som har drivit utvecklingen av teori och metod är civila. Den vanligaste tillämpningen av metoden är bedömning av olika komprimeringsalgoritmers inverkan på bildkvaliteten. Vilken komprimeringsmetod ger bäst resultat? Hur mycket kan bilden komprimeras utan att kvalitén blir alltför lidande? Det finns ingen direkt koppling mellan dessa kvalitetsmått och frågeställningar kring måldetektionsförmågan i en given bild. Istället erbjuder metoden en generell kvalitetsbedömning av bilden och möjligheter till att upptäcka generella kvalitetsämmande särdrag hos bilden så som komprimeringsartefakter, oskärpa eller rörelseoskärpa.

Automatisk bildkvalitetsbedömning delas vanligen upp i tre olika fall baserat på kännedom om referensbilden:

1. Känd referensbild (Full-reference (FR)). Här är referensbilden känd vid bedömningstillfället och avvikelser från referensbilden kan användas vid bedömningen. Ett typiskt fall är kvalitetsbedömning vid bildkomprimering då både den okomprimerade (referensbilden) och den komprimerade bilden som skall bedömas är tillgängliga. En frekvent förekommande metod vid FR är Structural Similarity Index (SSIM) [22].
2. Reducerad kännedom om referensbilden (Reduced-Reference (RR)). Här finns viss information om referensbilden så som närvaro av vissa särdrag.
3. Okänd referensbild (Non-Reference (NR)) kallas också blind kvalitetsbedömning. Här finns ingen information om referensbilden, vilket kräver andra antaganden och ansatser för att kunna bedöma bildkvalitén.

Intressantast vid bedömning av sensordata är blind kvalitetsbedömning då vanligen information om referensbilder saknas. Metoder för bildkvalitetsbedömning med känd referensbild (FR) och reducerad kännedom om referensbilden (RR) kommer inte behandlas vidare i denna rapport.

För att kunna bedöma bilder utan någon kännedom om referensbilden krävs antagande om hur bilden har genererats. För uppsättningar av naturliga bilder finns det statistiska regulariteter som alltid uppfylls. Regulariteterna gäller för en uppsättning av bilder av



tillräckligt stor variation och inte direkt för enskilda bilder. Enskilda bilder av tillräckligt hög upplösning (i pixlar) uppfyller vanligen approximativt även dessa regulariteter, dock varierar parameterestimeringarna mycket. Parameterestimeringarna beror mycket på det konkreta bildinnehållet vid estimering på enskilda bilder. Estimeringar av regularitetsparametrar för enskilda bilder utgör särdrag vid automatisk blind kvalitetsbedömning.

Mönsterigenkänningsmetoder används sedan för att automatiskt rangordna bilderna utifrån kvalitén. Vid rangordningsproblem skall en rankningsfunktion estimeras som bestämmer den relativa ordningen mellan par av bilder. Rankningsfunktionen kan sedan appliceras på en större uppsättning bilder. Träningsdata vid ett rangordningsproblem är par av bilder  $(I_1, I_2)$  med känd inbördes kvalitetsrangordning:  $I_1$  är bättre än  $I_2$  (dvs.  $I_1 >_{Rang} I_2$ ),  $I_2$  är bättre än  $I_1$  (i.e.  $I_2 >_{Rang} I_1$ ) eller  $I_2$  lika bra som  $I_1$  (i.e.  $I_2 =_{Rang} I_1$ ).

### Statistik för naturliga scener (NSS)

Det är känt att det finns statistiska egenskaper som uppfylls av uppsättningar av naturliga bilder. Estimering av parametrar för de statistiska egenskaperna blir snarlika för alla uppsättningar av naturliga bilder. Med naturliga bilder menas i detta sammanhang alla bilder som vanliga människor ser med sina ögon i sin vardag. Det innefattar såväl naturmiljöer som stadsmiljöer. Även bilder tagna inomhus innefattas av begreppet naturliga bilder (även om de flesta databaser med naturliga bilder bara innehåller utomhusbilder). En exakt definition av naturliga bilder finns inte. Naturliga bilder kan kontrasteras mot t.ex. röntgenbilder eller IR-bilder som inte räknas som naturliga bilder.

Dessa statistiska egenskaper gäller för en uppsättning av naturliga bilder och inte för enskilda bilder. En uppsättning av naturliga bilder har en fördelning med avseende på bl.a. scenval och avstånd till objekten i scenen. En uppsättning av naturliga bilder måste innehålla en tillräckligt stor variation med avseende på scenval och avstånd till objekten i scenen för att de statistiska egenskaperna skall uppfyllas. Någon definition om hur uppsättningen av naturliga bilder måste variera saknas. För uppsättningar innehållande allt för enahanda scener på snarlika avstånd till objekten i scenen gäller inte de statistiska egenskaperna. Uppsättningen av bilder måste således vara representativ. Estimering av de statistiska egenskaperna på enskilda bilder med varierade bildinnehåll är således inte oproblemiskt.

### Skalinvarians

Naturliga bilder är *skalinvarianta*. Det finns inte någon naturlig skala för bilden, eftersom alla objekt i princip kan finnas på alla avstånd från sensorn. Om man till exempel beaktar en naturscen innehållande träd så är det troligt att det finns träd på olika avstånd till sensorn. De varierande avstånden genererar likartade objekt, träd, på olika skalor i bilden. Trädens höjd och utbredning i pixlar beror på avståndet till sensorn. En uppsättning naturliga bilder kommer att innehålla objekt på alla skalor; träd på alla avstånd från sensor.

Skalinvariansen implicerar några centrala regulariteter hos naturliga bilder.

En av regulariteterna är att linjär lokal statistik kommer att vara identisk efter subsampling. Om man applicerar ett godtyckligt linjärt filter på en uppsättning av naturliga bilder och därefter beräknar ett normaliserat histogram över filterresponserna. Bilderna lågpasfilteras sedan med t.ex. ett gaussiskt filter och bilderna subsamplas sedan med t.ex. en faktor 2. Samma linjära filter appliceras sedan återigen på uppsättningen av subsamplade naturliga bilder och beräknar det normaliserade histogrammet över filterresponserna. Jämförelse mellan de normaliserade histogrammen från originaluppsättningen av

bilderna och uppsättningen av subsamplade bilder kommer att vara väldigt lika. Lokal linjär statistik är således oberoende av bildens skala [23].

Skalinvarians hos naturliga bilder kan också uttryckas i frekvensrummet. Effektspektrum (eng. Power spektrum) beräknat över en uppsättning av naturliga bilder följer en så kallad "power law" i spatiala frekvenser. Då gäller

$$S(\omega) = \frac{A}{\omega^{2-\eta}}$$

där  $S$  är effektspektrum (power spektrum) estimerat på en uppsättning bilder,  $\omega$  spatiala frekvenser och  $A$  respektive  $\eta$  estimerade konstanter. Estimeringar av  $\eta$  beror på uppsättningen av bilder som används vid estimeringen. På uppsättningar av naturliga bilder innehållande bilder av tillräckligt stor variation estimeras  $\eta$  till  $\approx 0.2$ .

Även effektspektrum för enskilda bilder följer vanligen en "power law" i spatiala frekvenser. Dock beror estimeringen av  $\eta$  på det konkreta bildinnehållet i de enskilda bilderna, och estimeringar av  $\eta$  varierar mycket. Estimeringar av  $\eta$  på bilder innehållande större intensitetshomogena regioner, så som en blå himmel eller vatten, ger små värden på  $\eta$  ( $\eta \approx 0.05$ ). Estimering av  $\eta$  på bilder innehållande småskaliga detaljer (texture), så som buskar och gräs på korta avstånd, ger större värden på  $\eta$  ( $\eta \approx 0.6$ ).

Skalinvarians kan också uttryckas i den spatiala domänen med autokorrelationsfunktionen. Autokorrelationsfunktionen definieras som

$$C(\tau) = E(I(x_0)I(x_0 + \tau))$$

där  $E$  är väntevärdet och  $I$  är bilden. Autokorrelationen uttrycker hur bilden är korrelerad med sig själv som en funktion av avståndet  $\tau$ . Eftersom autokorrelationsfunktionen är fourier transformen av effektspektrumet så implicerar skalinvarians att autokorrelationsfunktionen har följande form

$$C(\tau) = C_1 + \frac{C_2}{|\tau|^\eta}$$

där  $C_1$  och  $C_2$  är estimerade konstanter och  $\eta$  har samma värde som vid estimering av "power law" för effektspektrum.

### Kantpixelsfördelning

En frekvent förekommande modell för hur bilden genereras är att en scen/världen består av diskreta objekt, såsom hus, träd, mark, bilar osv. Dessa diskreta objekt är relativt homogena, de består av samma material under likartade ljusförhållanden. De diskreta objekten utgör homogena regioner med relativt små kontrastvariationer - objekten har likartad färg (intensitet). Till skillnad från de homogena regionerna med små kontrastvariationer, är det vanligen större kontrast vid objektets kant. Den homogena regionen kontrasteras mot en bakgrund med potentiellt helt andra materialegenskaper eller andra ljusförhållanden. Världen består av dels av intensitetshomogena objekt med relativt liten kontrast variation, dels kanter mellan objekten och en bakgrund bestående av andra objekt. Vid kanten av objekten är kontrasten relativt stor – ett diskret hopp i intensitet.

Detta kan formaliseras med hjälp av (spatiala) partiella derivator. De partiella derivatorna är relativt små inom de homogena regionerna då kontrastvariationen är liten. Å andra sidan är de partiella derivatorna stora vid objektets kanter där intensiteten gör ett diskret hopp. Inom de homogena regionerna är de partiella derivatorna små eller nästan 0, emedan de vid objektskanterna är godtyckligt stora. Fördelningen av partiella derivator kan modelleras med en Laplace-fördelning (täthetsfunktion)

$$p(x) = \frac{1}{z} e^{-\frac{|x|^\alpha}{s}}$$

där  $z$  är en normaliseringskonstant. Ibland kallas  $\alpha=1$  för Laplace-fördelning och generaliserad Laplace-fördelning för godtyckligt  $\alpha$ . Liknande fördelningar som ofta används är Weibull och generaliserad Gaussisk-fördelning. Laplace-fördelningen karakteriseras primärt (i förhållande till en normalfördelning) av två egenskaper:

- Större del av sannolikhetsmassan är centrerad runt  $x \approx 0$ .  $x \approx 0$  innebär små absolutvärden på de partiella derivatorna, vilket korresponderar med intensitetshomogena regioner med små eller inga intensitetsvariationer. Således utgörs större delen av bild av en intensitetshomogena regioner. Det är primärt parametern  $\alpha$  som påverkar hur centrerade täthetsfunktionen är kring  $x \approx 0$ .
- Fördelningen har långa svansar. Jämfört med en normalfördelning har en Laplace-fördelning en större sannolikhetsmassa långt ut i "svansarna". Stora  $x$  innebär stora absolut värden på de partiella derivatorna, vilket korresponderar med objektets kanter som har stor intensitetskontrast. Det är primärt parametern  $s$  som påverkar hur stor del av sannolikhetsmassan som utgörs av kantpixlar.

Estimering av parametern  $\alpha$  på en större uppsättning bilder ger vanligen värden  $\alpha \approx 0.7$ .

Fördelningen av de partiella derivatorna på enskilda bilder kan också modelleras med en Laplace-fördelning. För enskilda bilder varierar dock estimering av  $\alpha$  med bildinnehållet. Stora värden på  $\alpha$  för enskilda bilder indikerar att bilden innehåller större intensitetshomogena regioner så som t.ex. himmel, hav eller öken. Små värden på  $\alpha$  för enskilda bilder indikerar att bilden innehåller regioner med finskaliga detaljer (texturer) så som t.ex. gräs och buskar på kortare avstånd [23].

Även fördelningen av de partiella derivatorna för mindre bildpatcher kan modelleras med en Laplace-fördelning. Hypotestestning indikerar att de partiella derivatorna kommer från en Laplace-fördelning för nästan alla bildpatchar [24]. Återigen varierar värdena på  $\alpha$  med bildinnehållet.

## Wavelet-transformationen

Wavelet-transformen är en flerskalig metod som används frekvent inom bildanalys [25, 26]. Wavelet-transformen bygger upp en flerskalig representation av bilden i form av en bildpyramid. Bildpyramiden byggs upp genom att applicera kombinationer av lågpasfilter (L) och högpasfilter (H), och därefter subsampla de filtrerade bilderna. Bilden filtreras först horisontellt, i x-led, och subsamplas i x-led, därefter filtreras bilden vertikalt, i y-led, och subsamplas i y-led. Följande kombinationer används i wavelet-transformen:

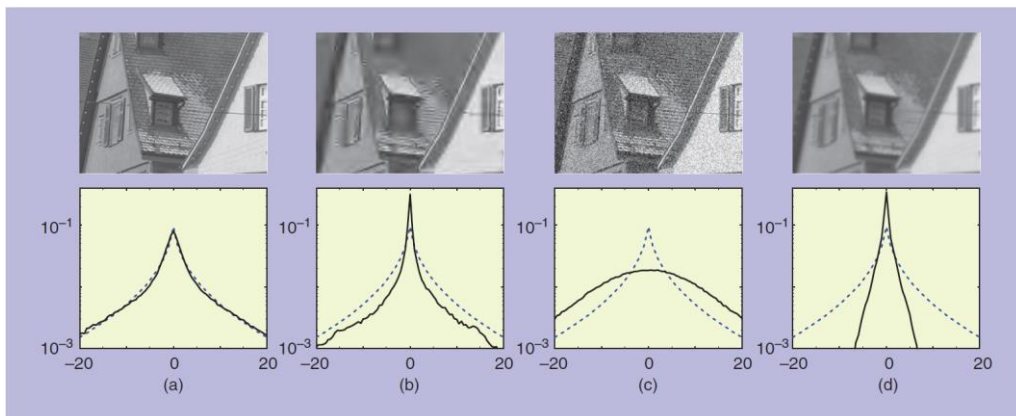
- LL – lågpasfiltrering (L) och subsampling av bilden först horisontellt, och därefter vertikalt.
- HL – först horisontell högpasfiltrering (H) och horisontell subsampling av bilden, följt av vertikal lågpasfiltrering (L) och vertikal subsampling av bilden.
- LH – först horisontell lågpasfiltrering (L) och horisontell subsampling av bilden, följt av vertikal högpasfiltrering (H) och vertikal subsampling av bilden.
- HH – högpasfiltrering (H) och subsampling av bilden först horisontellt, därefter vertikalt.

Wavelet-bilderna – LL, LH, HL och HH – subsamplas vid konstruktionen med en faktor 2, vilket ger fyra bilder med halverad upplösning (dvs. de fyra subsamplade bilderna

har tillsammans samma upplösning som originalbilden). De filtrerade bilderna brukar benämnas subband då de innehåller information från olika band. LL är en nersamplad version av originalbilden, HH innehåller högfrekventa småskaliga detaljer så som brus och texturer, emedan LH och HL tenderar att innehålla horisontell respektive vertikal kantinformation vanligen kallade geometrisk bildinformation. Intensiteterna i de konstruerade bilderna benämns wavelet-koefficienter. Wavelet-transformation tillämpas rekursivt på den subsamplade LL bilden för att en bygga bildpyramid med avtagande spatialupplösning (i.e. ökande skala). Varje rekursiv tillämpning av wavelet- transformen på LL i den föregående transformationen kallas en ny (högre) skala.

Fördelning av wavelet-koefficienterna i subbanden i wavelet-transformationen kan modelleras med en Laplace-fördelning. De olika subbanden i transformen modelleras med en Laplace-fördelning per subband och skala. För naturliga bilder följer koefficienterna i subbanden HL och LH samma fördelning och parameterestimering görs därför vanligen tillsammans. Koefficienterna i LL, som är en nersamplad version av originalbilden, kommer vanligen inte från en Laplace-fördelning och modelleras vanligen inte. För varje skala i wavelet-transformationen estimeras således två parametrar  $\alpha$ : en för wavelet-koefficienterna i subbanden LH/HL och en för wavelet-koefficienterna i subbandet HH [27, 28].

För en uppsättning bilder kan wavelet-koefficienterna modelleras med en Laplace-fördelning. Även för enskilda bilder kan fördelningen av wavelet-koefficienterna modelleras med en Laplace-fördelning, dock bara på skalor som innehåller tillräckligt många koefficienter (dvs. bildupplösningen är tillräckligt stor).



Figur 5. Empirisk distribution av wavelet-koefficienter för (a) original bild (oförändrad), (b) komprimerad med JPEG2000, (c) med adderat gaussiskt brus och (d) oskärpa genom faltning med en gaussisk kärna. Empiriska fördelning av wavelet-koefficienter (solid) respektive en anpassad (ML) Laplace-fördelning (streckade) till data. Wavelet-koefficienterna för original bilden (a) kan modelleras bra med en Laplace-fördelning, emedan anpassning av wavelet-koefficienter till en Laplace-fördelning för de degenererade bilderna är sämre. (Från [12]).

## Rangordningsmetoder

Mönsterigenkänningsmetoder används för att automatiskt kvalitetsrangordna bilderna. Bilderna representeras som särdragsvektor, där särdragen utgörs av parameterestimeringar från statistiska modeller för naturliga bilder (NSS). Det antas att den statistiska avvikelser hos dessa parametrar för de degenererade bilderna avviker så pass signifikant att det kan användas för att automatiskt bedöma bildkvalitén. Olika typer av degenereringar hos bilderna ger upphov till olika typer av statistiska avvikelser för parametrarna. Utan att specificera hur statistiken avviker för olika typer av degenerering antar man att parametrarna innehåller tillräckligt med information för att kunna rangordna bilder.

För att kunna använda träningsbaserade mönsteringekänningsmetoder krävs en relativt omfattande mängd träningssexempel. Träningsmängden utgörs vid rangordning av en mängd ordnade par av bilder  $(I_1, I_2)$  representerade som särdragsvektorer. Rangordningen mellan de ordnade paren är känd:  $I_1$  är bättre än  $I_2$  (i.e.  $I_1 >_{Rang} I_2$ ),  $I_2$  är bättre än  $I_1$  (i.e.  $I_2 >_{Rang} I_1$ ) eller  $I_2$  lika bra som  $I_1$  (i.e.  $I_2 =_{Rang} I_1$ ).

Träningsmängden består av bildpar där en slutanvändare, en expert, har kvalitetsbedömt bildparen och avgjort vilken bild som var av högst visuell kvalitet (eller eventuellt lika hög kvalitet). Träningsmängden utgörs således av par av bilder som manuellt parvis har rangordnats. Varje kombination av bildpar visas för en användare som avgör vilken bild som användaren bedömer ha högst visuell kvalitet. För att få en fullständig rangordning av t.ex. 100 bilder krävs  $100 \cdot 99 / 2 = 4950$  manuella bedömningar. Således får man ett relativt stort antal träningssexempel med ett relativt litet antal bilder, dock krävs vanligen även en stor mängd bilder. Ett attraktivt, men något förrädiskt, alternativ är att betygsätta bilderna med avseende på bildkvaliteten, t.ex. varje bild får ett betyg från 1 till 5. Betygen kan sedan implicit användas för att rangordna bilderna. Således behövs endast 100 bilder betygsättas för att generera en träningsmängd innehållande 4950 bildpar. Det är dock problematiskt att konsistent betygsätta större mängder bilder då betygsättningen tenderar att påverkas av nyligen observerade bilder, t.ex. att betygen divergerar över tiden. Vidare är det svårt att jämföra betygsättningar gjorda av olika bedömare. För att få en konsistent träningsmängd av rangordnade bilder krävs således en parvis bedömning av bilderna. Vidare är det önskvärt att bildparen bedöms av fler än en person, detta för att minska subjektiviteten i kvalitetsbedömningarna. För att minska antalet bildpar som måste rangordnas kan ett begränsat antal par väljas ut för bedömning. De mest informativa paren för bedömning är par av bilder som är kvalitativt likartade. Kvalitetsbedömning mellan väldigt bra och väldigt dåliga bilder ger begränsat med information.

Vid inläring av en rangordningsfunktion används de rangordnade paren. Algoritmen förses med par av särdragsvektorer extraherade från bilderna samt vilken av bilderna som har högst rang. Det finns ett stort antal metoder för rangordningsinläring. Flertalet av de inlärningsbaserade rangordningsmetoderna är utvecklade för websökningar och Information Retrieval (IR) och dessa tillämpningar som är forskningsledande inom rangordning av data. Exempel på frekvent använda inlärningsbaserade rangordningsmetoder: RankSVM [29, 30], RankBoost [31, 32] and GBrank [33, 34].

Vid applicering av den inlärd rangordningsfunktionen finns två tillvägagångssätt:

1. *Komparativ rangordningsfunktion.* Vid applicering förses algoritmen med två bilder för en direkt komparativ kvalitetsbedömning. Algoritmen returnerar information om vilken av de två bilderna som är av högst kvalitet. Vid applicering har rangordningsfunktionen  $r$  form  $r: X \times X \rightarrow C$  där  $C$  utgörs av omdömena: bild1 är av högre kvalitet, bild2 är av högre kvalitet samt bild1 och bild2 är av lika högkvalitet. Rangordningsfunktion är således direkt komparativ vid applicering.
2. *Graderande rangordningsfunktion.* Vid applicering förses algoritmen med en bild för kvalitetsbedömning. Algoritmen returnerar ett betyg för bilden i en inlärd ordinal skala, dvs. en skala som kan användas för rangordning av bilder, men som inte ger någon meningsfull kvantitativ information om kvalitets skillnader mellan olika värden. Vid applicering har rangordningsfunktionen  $r$  formen  $r: X \rightarrow \mathbb{R}$ . Vid rangordning av ett bildpar appliceras således rangordningsfunktionen för gradering, en gång per bild, därefter jämförs graderingarna för att avgöra vilken bild som är av högst kvalitet.

Kvalitetsbedömning med en komparativa rangordningsfunktion kan göras genom att rangordna den aktuella bilden mot en uppsättning av bilder av varierande kvalitet. Den

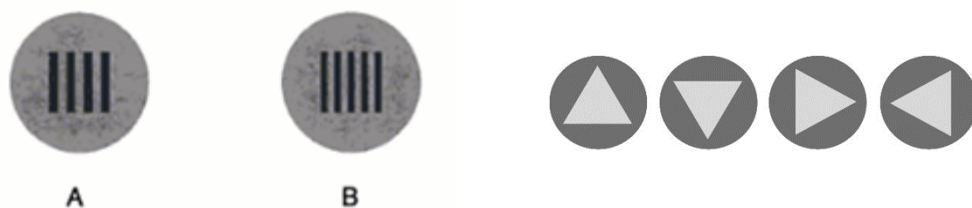
relativa frekvensen av bilder som den aktuella bilden rangordnas högre än kan ses som ett mått på dess kvalitet. Den aktuella bilden ”tävlar” mot ett antal bilder av varierande kvalitet och antalet vinster är ett mått på hur bra bilden är.

Kvalitetsbedömning med en graderad rangordningsfunktion kan göras genom att först gradera en uppsättning bilder av varierande kvalitet. Därefter graderas den aktuella bilden och graderingen jämförs sedan med graderingarna för uppsättningen bilder. Den relativa frekvensen av bilder som den aktuella bilden har graderats högre än kan ses som ett mått på bildens kvalitet.

## 2.3 Prestandamått för manuell måldiskriminering

För taktiska sensorer, där målutpekning ofta står i fokus, har bildkvalitetsmått som exempelvis NIIRS något mindre relevans och mått som sannolikhet för detektion, igenkänning och identifiering används oftare pga. deras mer direkta koppling till målbeaktning och dess effektivitet. Ett antal olika metoder på detta tema har utvecklats genom åren, varav de mest frekvent använda är Johnson-kriteriet (avsnitt 2.3.1), Targeting Task Performance (avsnitt 2.3.2) och Thermal Range Model (2.3.3).

Typiska systemprestandamått som används för EO- och IR-sensorer är MTF (Modulation Transfer Function) för sensorsystemets spatiala upplösningssmått samt NEI (Noise Equivalent Irradiance) eller NETD (Noise Equivalent Temperature Difference) för dess känslighet. Då dessa parametrar beror av varandra och det ofta finns en kompromiss dem emellan har också systemprestandaparametrar tagits fram där båda ingår. MRC (Minimum Resolvable Contrast) och MRTD (Minimum Resolvable Temperature Difference) beskriver den minsta kontrasten eller temperaturskillnaden hos ett visst periodiskt linjemönster (se figur 6) som kan upplösas/särskiljas av en observatör. Denna subjektiva prestanda blir därmed beroende av spatial frekvens hos linjemönstret [35]. Eftersom linjemönster är känsliga för hur mönster och samplande sensorsystem är upplinjerade i förhållande till varandra (dvs. fasskillnaden dem emellan), kan s.k. vinkningsdistorsion uppstå (eng. aliasing) där högfrekvent innehåll beblandas med lågfrekvent. Av denna anledning finns även andra typer av måltavlor framtagna, exempelvis liksidiga trianglar i olika orienteringar (se figur 6), storlekar och kontrast som t.ex. i TOD-metoden (Triangle Orientation Discrimination) [36].



Figur 6. Till vänster visas två olika periodiska linjemönster i en s.k. 4-bar target som används vid bestämning av MRTD/MRC. Till höger visas trianglar med fyra olika orienteringar, som används i TOD-metoden.

### 2.3.1 Johnson-kriteriet

Ett av de första försöken att relatera laborativa sensorprestandamått till prestanda för användare eller observatörer ute i fält gjordes av John Johnson vid U.S. Army Night Vision Laboratory på 1950-talet. Johnson genomförde en serie experiment med testpersoner och bildförstärkare för att undersöka hur bra de lyckades särskilja linjer med varierande linjeavstånd i ett periodiskt mönster över olika avstånd, parallellt med deras förmåga att upptäcka, känna igen och identifiera olika typer av militära målobjekt över samma avstånd. Linjerna på testtavlor hade samma kontrast som målobjekten och

experimenten gav upphov till en tabell som visade hur många linjer som måste kunna upplösas över varje typ av objekt för olika typer av diskrimineringsuppgifter, se exempel i tabell 2. Antalet linjer räknades över den karaktäristiska dimensionen av objektet, som i 1D-fallet utgjordes av den kortaste dimensionen för objektet och i 2D-fallet av kvadratroten ur objektets målarea. En cykel består av en linje och ett mellanrum och  $N_{50}$ -värdena anger antalet cykler som behövs för att med 50 % sannolikhet kunna lösa den specifika uppgiften. Detektion i tabellen anger förmågan att kunna upptäcka att ett militärobject finns i scenen och igenkänning representerar förmågan att kunna klassificera objektet (ex. stridsvagn kontra lastbil). Identifiering innebär att mer noggrant kunna specificera typen av objekt (ex. stridsvagn av typen M1A1 eller T62, osv.).

Tabell 2. Diskrimineringskriterier enligt Johnson [5].

Diskrimineringsuppgift	Antalet 1D-cykler ( $N_{50}$ )	Antalet 2D-cykler ( $N_{50}$ )
Detektion/upptäckt	1.0	0.75
Igenkänning	4.0	3.0
Identifiering	6.4	6.0

För att beräkna sannolikheten för en viss diskrimineringsuppgift för en specifik sensor vid ett specifikt avstånd  $R$  beräknas antalet upplösta/särskiljbara cykler över målobjektet enligt [37]

$$N = \rho_0 \frac{d_c}{R} \quad (7)$$

där  $\rho_0$  är den begränsande frekvensen (eng. limiting frequency) för sensorn och  $d_c = \sqrt{w \cdot h}$  är den karaktäristiska längden/dimensionen för objektet med bredden  $w$  och höjden  $h$ . Den begränsande frekvensen för sensorn mäts vanligen upp i en laborativ mätning av MRTD/MRC men kan också räknas ut teoretiskt.

Sannolikheten ges sedan av den empiriska formeln [37]

$$P(N) = \frac{(N/N_{50})^{2.7+0.7(N/N_{50})}}{1 + (N/N_{50})^{2.7+0.7(N/N_{50})}} \quad (8)$$

där diskrimineringsnivån (upptäckt, igenkänning och identifiering) och dess  $N_{50}$ -kriterium ges ur tabell 2. Även om Johnson i sina experiment utgick från bildförstärkare förfinades metoden över åren till att också omfatta IR-sensorer. Metoden, känd som Johnson-kriteriet, blev snabbt väletablerad och används trots sina brister fortfarande idag som en enkel och approximativ metod.

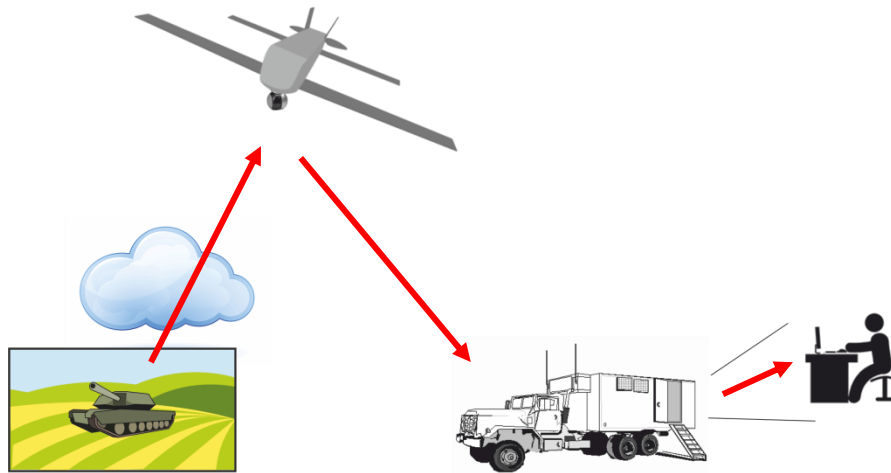
### 2.3.2 Targeting Task Performance (TTP)

Johnson-kriteriet bygger på ett antagande om att bildkvaliteten och prestandan hos det elektrooptiska systemet helt och hållet bestäms av dess egenskaper vid en enskild spatial frekvens, dvs. den högsta upplösta (begränsande) frekvensen. Detta innebär (felaktigt) att en generell kontrastförbättring erhålls för alla spatiala frekvenser, både lägre och högre, då systemets kontrastgivande prestanda förbättras vid just denna frekvens. För samplande (pixelupbyggda) sensorer är dock denna frekvens snarare en indikation på

samlingsfrekvens (pixelavstånd) än generell bildkvalitet. Det är dessutom så att eftersom metriken endast grundar sig på responsen vid en enskild frekvens, så kan den inte prediktera effekterna av exempelvis pålagd kantskärpning, vars syfte är att omforma hela frekvensspektrat i en bild. En annan påtalad brist hos Johnson är att metriken inte kan prediktera inverkan av brus på ett korrekt sätt, vilket yttrar sig i att om systemet är brusbegränsat så antyder metriken (återigen felaktigt) att bättre bildkvalitet och bättre räckviddsprestanda erhålls om endast upplösningen (den begränsande frekvensen) förbättras, dvs. om antalet pixlar ökas [38].

NVESD (US Army Night Vision and Electronic Sensor Directorate) i USA tog 2003 fram en ny värderingsmetod kallad Targeting Task Performance (TTP), som ersätter till Johnson med bättre förutsättningar för att kunna värdera moderna samplande EO- och IR-sensorer. Metoden är validerad genom flertalet perceptionsförsök och finns idag integrerad i exempelvis mjukvaran NV-IPM (Night Vision Integrated Performance Model) från SENSIAC/NVESD. TTP är i likhet med de flesta liknande modeller för observatörsprestanda (GIQE och TRM) en statisk prestandamodell som gäller under förutsättning att observatören har så mycket tid som behövs för att lösa sin uppgift [39].

I likhet med Johnson finns det i TTP ett underliggande antagande att räckviddsprestanda är proportionell mot bildkvalitet. De två metriken predikter dock bildkvalitet på två olika sätt. Medan det i Johnson-kriteriet antas det att den elektrooptiska sensorns prestanda bestäms av dess egenskaper vid den högsta genomsläppta frekvensen tas det i TTP hänsyn till det avbildande systemets kontrastgivandeprestanda vid *samtliga* spatiala frekvenser, något som bäst beskrivs av systemets modulationsöverföringsfunktion (MTF). Systemets MTF bestäms av alla komponenter i avbildningskedjan från scen till sensor till observatör, se Figur 1, där den ursprungliga informationen i form av modulation (vilket är ett annat ord för kontrast) försämras via olika oskärpe- och brusbidrag som introduceras längs vägen.



Figur 7. En avbildningskedja är en serie av fysikaliska länkar från scen till sensor till observatör. Länkarna i kedjan introducerar oskärpa och brus till den observerade bildinformationen.

För ett linjärt och spatialt skiftinvariant avbildningssystem kan den sammanlagda (superponerade) försämringen av kontrast i bilden matematiskt beskrivas av avbildningskedjans s.k. system-MTF, här betecknad  $MTF_{sys}(\xi)$ :

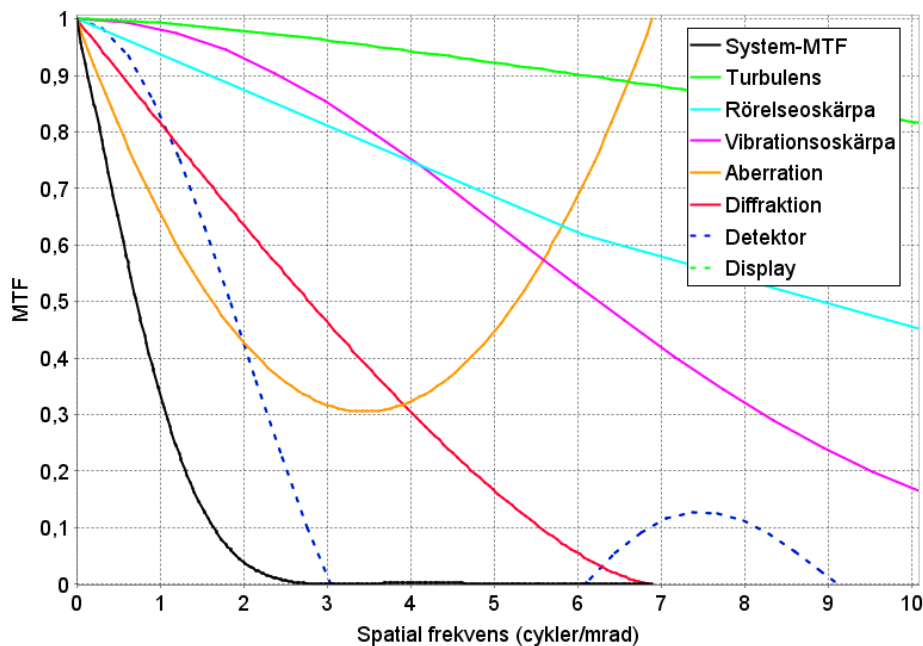
$$MTF_{sys}(\xi) = MTF_{atmosfär}(\xi)MTF_{optik}(\xi)MTF_{detektor}(\xi)MTF_{display}(\xi) \cdots \quad (9)$$

som i frekvensrummet blir en produkt av de enskilda degraderande MTF-bidragen (i den vanliga spatiala domänen beskrivs den av en faltning).  $\xi$  betecknar spatial frekvens,



vilket är ett mått på linje- eller mönstertäthet i en bild där låga frekvenser representerar stora, grovkorniga variationer och höga frekvenser är fina detaljer.

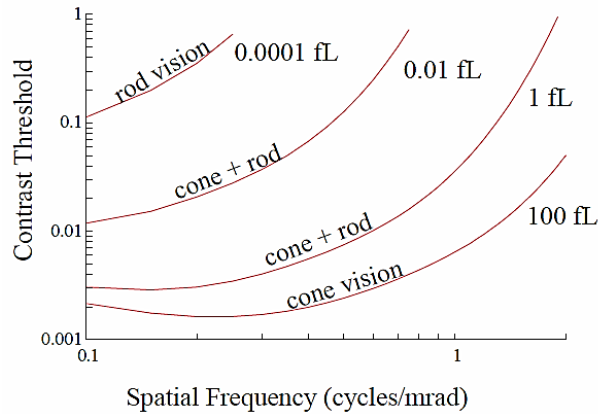
Komponenter i kedjan som inverkar kan exempelvis vara *atmosfär* (pga. turbulens och aerosolspridning), *optik* (pga. diffraktion och aberrationsfel), *detektor* (pga. pixlars storlek och utformning), *display* (pga. pixelstorlek, uppdateringsfrekvens, ljusreflexer, mm.) och *observatörens seende* (pga. diffraktion och synskärpa). Ytterligare försämringar kan även uppstå pga. *rörelse- och vibrationsoskärpa* hos sensorplattformen (och/eller observatören relativt displayen) eller genom nödvändig *komprimering* av data i länken ned till en markstation. Olika komponenter dominerar i olika applikationer och matematiska modeller finns framtagna och integrerade för de flesta av dessa i mjukvaran NV-IPM (Night Vision Integrated Performance Model).



Figur 8. MTF för en typisk EO-sensor bestäms av flertalet olika degraderande bidrag i avbildningskedjan, som exempelvis atmosfär (turbulens), rörelse- och vibrationsoskärpa hos sensorplattform, optik (diffraktion och aberration), detektor och display. Det totala systemets MTF (svart kurva), kan beräknas genom produkten av de ingående komponentbidragen (färgade kurvor).

I TTP ingår också en modell för den mänskliga observatörens förmåga att diskriminera kontraster med olika spatial frekvens. Modellen (Bartens ögonmodell) är empiriskt uppbyggd från olika perceptionsförsök med sinusmönster med olika frekvenser och kontraster. Yttre faktorer såsom brus (från ex. sensor, display och öga) samt luminansnivå (ex. från en display) inverkar också, vilket beror på att det mänskliga ögats stavar och tappor har olika ljus- och färgkänslighet, olika densitet över näthinnan och olika kontrastdiskriminerande egenskaper. Förmåga kan variera mellan individer, förändras med ålder och beror på fysisk och psykisk status hos den mänskliga observatören (trötthet, sjukdomar, osv.), men vanligtvis används en modell utgående från medelresultaten för en stor ensemble testindivider. I TTP används den s.k. *kontrasttröskelfunktionen*  $CTF(\xi)$ , som beskriver den kontrast som en mänsklig operatör kräver för att precis kunna diskriminera skillnaden mellan ett ljust och ett mörkt fält och är en serie av kurvor som funktion av spatial frekvens för olika luminansnivåer. Figur 9 visar exempel, där det kan noteras att ögat tappor sin förmåga att urskilja små kontraster vid samtliga frekvenser då ljusnivån minskar (kurvorna flyttas uppåt då luminansen minskar). Notera också att vid goda ljusförhållanden (då tapparna dominerar seendet) har ögat en bättre

kontrastdiskriminerande förmåga vid medelhöga frekvenser än vid låga frekvenser (CTF-kurvan uppvisar ett lokalt minimum). För att ta hänsyn till att kontrasttröskeln beror av brus multipliceras funktionen med en brusfaktor (som är större än ett) som beräknas som kvadratroten ur summan av respektive bruskillors varianser.

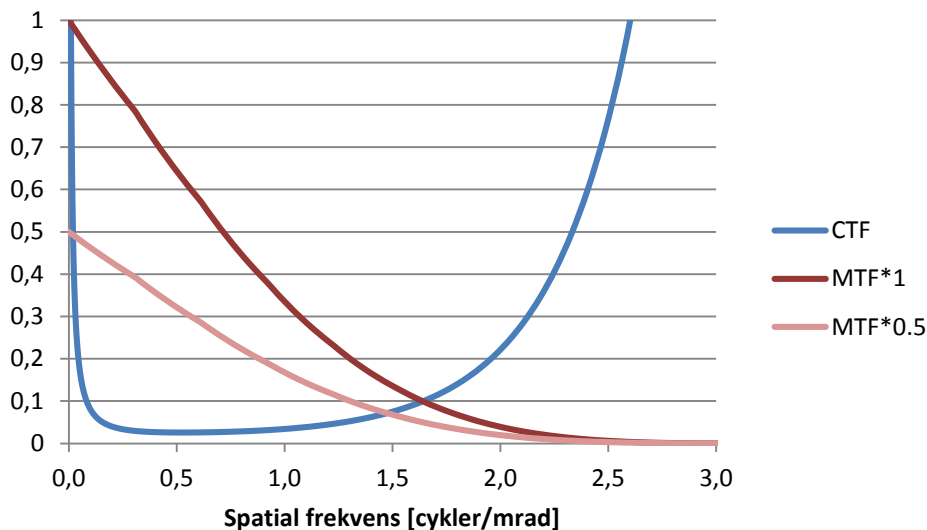


Figur 9. Ögats kontrasttröskelfunktion vid olika ljusnivåer (fL står för foot-Lambert och är ett mått på luminans, där 1 fL motsvarar ungefär  $3.43 \text{ cd/m}^2$ ). Frekvensen där kontrasttröskeln korsar värdet 1 utgör den högsta möjliga spatiala frekvensen som ögat kan uppfatta vid den givna ljusnivån [38].

Bildkvalitetsmättet TTP beräknas sedan enligt integralen

$$TTP = \int_{\xi_{low}}^{\xi_{cut}} \left[ \frac{C_{TGT} MTF(\xi)}{CTF(\xi) \sqrt{1 + Noise(\xi)}} \right]^{1/2} d\xi \quad (10)$$

där  $C_{TGT}$  är den skenbara mål/bakgrunds-kontrasten vid sensorn (som beror av avståndet till målet pga. atmosfärens dämpande inverkan),  $Noise(\xi)$  är det totala systembruset och där integrationen sker mellan de spatiala frekvenser där hela integranden är större än ett. Summeringen sker upp till den högsta möjliga spatiala frekvensen som observatören kan uppfatta, se illustration i Figur 10. Genom ekv. (10) erhålls ett mått på systemets (över)förmåga att avbilda scenen med en kontrast som överstiger den mänskliga perceptionens förmåga att uppfatta den vid olika frekvenser och bildkvalitetsmättet TTP kan alltså ses som ett mått på den överskotts-kontrast som finns i bilden. Om upplösningsskarakteristika skulle skilja sig åt i horisontell och vertikal led (vilket den exempelvis gör för exempelvis skannande sensorer), beräknas TTP som ett geometriskt medelvärde av de två riktningssberoende TTP-värdena.



Figur 10. Exempel på kontrasttröskelfunktion (CTF) och skenbar kontrast (MTF multiplicerad med målkontrast), för två olika kontrastvärden 1 och 0.5. Bildkvalitetsmättet TTP blir proportionell mot arean mellan blå och röd kurva enligt ekv. (10). Detta kan jämföras med Johnson-kriteriet, där kvalitetsmättet utgörs av (den högra) skärningspunkten mellan de två kurvorna.

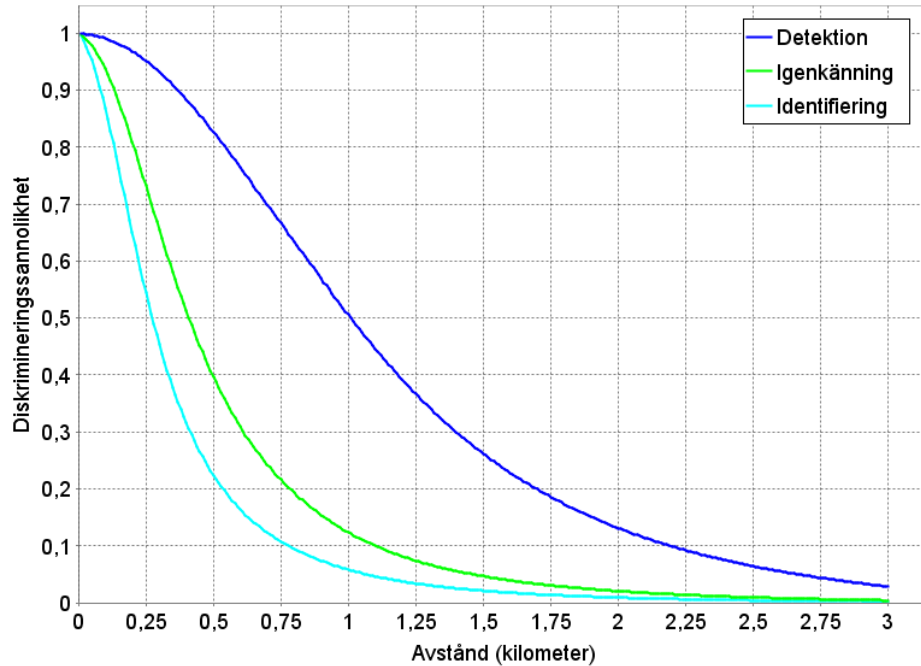
Antalet upplösta/särskiljbara cykler över målobjektet (i TTP betecknad  $V$ ) beräknas därefter, helt ekvivalent med Johnson enligt ekv. (7), enligt

$$V = TTP \frac{d_c}{R} \quad (11)$$

Sannolikheten att kunna utföra diskrimineringsuppgiften som funktion av avstånd beräknas därefter av den empiriska funktionen

$$P = \frac{(V/V_{50})^{A+B(V/V_{50})}}{1 + (V/V_{50})^{A+B(V/V_{50})}} \quad (12)$$

där  $V_{50}$  är antalet cykler på målet som krävs för den aktuella uppgiften och som vanligtvis bestäms/kalibreras ur perceptionsförsök i fält mot det specifika målet och uppgiften. A och B är två empiriskt framtagna konstanter, där de rekommenderade värdena i nuvarande version av TTP är  $A=1.5$  och  $B=0$ . Figur 10 visar exempel på hur typiska räckviddskurvor kan se ut, där sannolikheten P är plottad mot olika målavstånd. Framräknade sannolikheter bör dock inte tas som någon absolut sanning, då det kan finnas stora individuella variationer hur olika observatörer presterar och lämpar sig oftast bättre för relativa jämförelser mellan olika sensorer, mål, bakgrunder, väderförhållanden, osv.



Figur 11. Exempel på räckviddskurvor beräknade enligt TTP-metriken, för diskrimineringsuppgifter i form av detektion ( $V_{50}=2.0$ ), igenkänning ( $V_{50}=7.5$ ) och identifiering ( $V_{50}=13.0$ ).

### 2.3.3 Thermal Range Model (TRM)

Thermal Range Model (TRM) är ett europeiskt initiativ och ett alternativ till TTP och dess föregångare (NVTherm, FLIR92, Johnson, osv.). Modellen är framtagen av Fraunhofer (tidigare FGAN) på 1990-talet som svar på den då rådande bristen av prestandamodeller anpassade för undersamlade termiska avbildningssystem, för vilka pixelstorleken är stor i förhållande till den optiska oskärpan. Matematiskt definieras ofta undersamlade system som avbildningssystem där sammansatt MTF (system-MTF) innan sampling är större än noll vid Nyquist-frekvensen (halva samplingsfrekvensen). Den amerikanska FLIR92-modellen (en av TTP:s föregångare) saknade möjligheten att hantera dessa typer av system där fasskillnaden mellan pixel och infallande ljus spelar en stor roll för huruvida en eller flera linjer löses upp korrekt eller inte. Räckviddsförsök visade klart att fenomenet som s.k. aliasing, där höga spatiala frekvenser i scenen viks ned och beblandas med låga frekvenser vid själva samplingsförfarandet, hade stor effekt för observatörsprestandan. [40, 41] I sammanhanget bör det dock nämnas att TTP (som introducerades i NVTherm) numera har en metod att korrigera eller åtminstone kompensera för detta och TRM och TTP ger ofta (men inte alltid) jämförbara prestandaresultat. Den senaste versionen (2010) av TRM benämns TRM4, som förutom termiska system också innefattar metodik att hantera andra typer av elektrooptiska sensorer [42].

I TRM definieras två nya begrepp som är centrala för metoden. Istället för MTF används AMOP (Average Modulation at Optimum Phase) som har snarlik definition men som beskriver kontrastöverföringsegenskaperna hos systemet då det är optimerat för effekterna av sampling och fasskillnad. Istället för det tidigare använda MRTD definieras begreppet MTDP (Minimum Temperature Difference Perceived) [40] som

$$MTDP(r_p) = \frac{\frac{\pi}{2} SNR_{th} \sqrt{\Psi_z(r_p)}}{AMOP(r_p)} \quad (12)$$

där  $r_p$  är spatial frekvens,  $\Psi_z(r_p)$  är det totala systembruset och  $SNR_{th}$  är tröskeln för signal/brus. För sensorer i den reflektiva domänen används MSDP (Minimum Signal Difference Perceived) med ekvivalent definition, som ersätter det tidigare begreppet MRC.

Med ovanstående modifikationer kan man med TRM räkna ut räckvidd (och diskriminerings sannolikheter) på liknande sätt som i Johnson och TTP, genom att från ekv. (12) och målets storlek beräkna antalet upplösta linjer över målet och jämföra detta med det antal linjer som uppgiften kräver. I likhet med TTP är TRM en statisk prestandamodell där observatören har den tid som behövs för att lösa en uppgift.

### 3 Estimering av prestanda för radarsystem

Det finns många olika typer av radarsystem, men för placering ombord på en RPAS (Remotely Piloted Aircraft System) för spaning mot markmål handlar det om:

- Se och undvika andra flygande farkoster.
- Utföra spaning mot objekt på marken som står still under spaningstiden.
- Utföra spaning mot objekt på marken som befinner sig i rörelse under spaningstiden.

Det första fallet tas inte upp här då det i första hand berör plattformens navigation, utan vi kommer att koncentrera oss på spaningsradar mot mål på marken.

Vid spaning mot stillastående mål använder man sig av en SAR (Synthetic Aperture Radar) där man utnyttjar plattformens egenrörelse för att få bättre upplösning. Man skapar en stor virtuell antenn genom att utnyttja signaler som registreras under en bestämd flygsträcka. Om man antar att den betraktade scenen är stillastående kan man räkna fram bilder med hög upplösning. Man vinner i upplösning, men förlorar i tid. Rörliga mål kommer att smetas ut och få en felaktig position beroende av förflyttningshastighet och riktning, eller kan till och med försvinna helt. Denna svaghet kan i en del fall minskas genom en mer avancerad och tidsödande bearbetning av data. Det behövs extra beräkningssteg som kan detektera samt uppskatta storlek och riktning hos detekterade rörliga mål.

För att spana mot rörliga markmål behövs en GMTI-radar (Ground Moving Target Indication) som är optimerad för rörliga mål, men som tyvärr har andra problem istället. Det finns två huvudprinciper.

- En konventionell avsökande dopplerradar som kan se rörelse i radiell riktning relativt plattformen. Detta ger relativt grov upplösning, men mättiden är mycket kort. Normalt brukar man också filtrera bort de detektioner som anses komma från marken. Vilken hastighetskomponent som då undertrycks beror på plattformens egenhastighet och radarns momentana lobriktning. Den här tekniken fungerar bäst för relativt snabbt rörliga objekt som t.ex. fordon av olika slag, men fungerar sämre för långsamt rörliga objekt då en lång mättid krävs för att säkert detektera dessa.
- DPCA (Displaced Phase Center Antenna) som enkelt kan uttryckas som att man har ett SAR-system med flera antenn-mottagar-kanaler som är placerade utmed plattformens flygriktning. Man kan då räkna fram målens hastighetsvektorer. Vilka hastighetsintervall som kan detekteras avgörs av våglängd, antal och utsträckning av antennkanaler samt av flyghastighet och integrationstid. Denna teknik kräver noggrann kalibrering och mycket beräkningskraft uppe i plattformen för att fungera bra. Ännu så länge finns det få system på marknaden varför erfarenheterna är få.

Vi kommer inte att närmare beskriva metoder för att uppskatta prestanda för detektion av rörliga mål i denna rapport.

#### 3.1 NIIRS för SAR-radarbilder

NIIRS-kriterierna finns också definierade för SAR-bilder enligt tabellen nedan. Precis som i det optiska fallet inkluderas operatörerna i tolkningskedjan. Jämför med det som gäller för optiska sensorer i avsnitt 2.1.

NIIRS-nivå	GRD (m)	Exempel på spaningsförmåga
0	-	Interpretability of the imagery is precluded by obscuration, degradation, or very poor resolution
1	> 9.0	Detect the presence of aircraft dispersal parking areas. Detect lines of transportation (either road or rail), but do not distinguish between
2	4.5 - 9.0	Detect the presence of large (e.g., BLACKJACK, CAMBER, COCK, 707, 747) bombers or transports. Detect a military installation by building pattern and site configuration. Detect large non-combatant ships (e.g., freighters or tankers) at a known port facility.
3	2.5 - 4.5	Detect medium-sized aircraft (e.g., FENCER, FLANKER, CURL, COKE, F-15). Determine the location of the superstructure (e.g., fore, amidships, aft) on a medium-sized freighter.
4	1.2 - 2.5	Distinguish between large rotary-wing and medium fixed-wing aircraft (e.g., HALO helicopter versus CRUSTY transport). Detect individual vehicles in a row at a known motor pool.
5	0.75 - 1.2	Count all medium helicopters (e.g., HIND, HIP, HAZE, HOUND, PUMA, WASP). Distinguish between river crossing equipment and medium/heavy armored vehicles by size and shape (e.g., MTU-20 vs. T-62 MBT).
6	0.40 - 0.75	Distinguish between variable and fixed-wing fighter aircraft (e.g., FENCER vs. FLANKER). Distinguish between small support vehicles (e.g., UAZ-69, UAZ-469) and tanks (e.g., T-72, T-80).
7	0.20 - 0.40	Distinguish, by size and configuration, between a turreted, tracked APC and a medium tank (e.g., BMP-1/2 vs. T-64). Detect a missile on the launcher in an SA-2 launch revetment.
8	0.10 - 0.20	Distinguish the fuselage difference between a HIND and a HIP helicopter. Identify the SA-6 transloader when other SA-6 equipment is present.
9	< 0.1	Detect major modifications to large aircraft (e.g., fairings, pods, winglets). Identify, based on presence or absence of turret, size of aun tube, and chassis configuration, wheeled or tracked APCs by type (e.g., BTR-80, BMP- 1/2, MT-LB, MI 13).

### 3.2 Kvalitetsmått för SAR-bilder

En SAR-bilds kvalitet beror på många olika faktorer förutom själva sensorn som består av sändare och mottagare med tillhörande antensystem. Detta medför att det är svårare att definiera en SAR-version av GIQE (den optiska versionen beskrivs i avsnitt 2.1.1) med motsvarande egenskaper. Bildkvaliteten påverkas av:

- Sensorn och dess egenskaper.
- De eftersökta målen och dess egenskaper.
- Den bakgrundsmiljö som målen befinner sig i.
- De algoritmer som används för att räkna fram bilder samt detektera och klassificera detekterade objekt.

Det finns de som försökt med varierande framgång att ta fram en SAR-version av GIQE, men man får tillsvidare anse att det finns en hel del ytterligare att göra på detta område. Nolan et. al. i [43] tycks ha kommit längst med detta. De föreslår följande:

$$GIQE_{SAR} = a + b \left[ -\log\left(\frac{c}{2BW\cos\Phi}\right) - \log\left(\frac{\lambda}{4\sin\left(\frac{\Delta\theta}{2}\right)}\right) + \log\left(\log([1 + SNR] * 2^{-I(\text{non-coherent})})\right) + \log(I(\text{polarizations})) \right]$$

där  $BW$  är den nominella bandbredden,  $\Phi$  är belysningsvinkel under horisonten som också brukar benämnas depressionsvinkel,  $\Delta\theta$  är den syntetiska aperturvinkeln och  $I$  är en indikatorfunktion som indikerar förekomst av mätparametern inom parentes. Det som inte är enkelt att räkna ut är SNR eller egentligen signalnivå som kommer från målobjekt i förhållande till signalnivå som kommer från bakgrundsobjekt och andra störningar.

En SNR-ekvation för SAR erhålls om man integrerar den så kallade radarekvationen över den koherenta integrationstiden. Om man antar att  $G_t A_r \sigma$  är konstant under integrationstiden och att flygfarten  $v$  är konstant samt att integrationsintervallet är symmetriskt placerat kring tvärsriktningen relativt flygbanan fås SNR-ekvationen enligt

$$\left(\frac{S}{N}\right)_{\max} = \frac{P_{av}}{v} \frac{G_t A_r \sigma}{(4\pi)^2 r^3 k_B T_s L_s} \frac{\Delta\theta + \sin\Delta\theta}{2} \quad (2)$$

där

$P_{av}$  = utsänd medeleffekt på sändarens utgång (W)

$G_t$  = antennförstärkning under sändning (-)

$A_r$  = effektiv antennarea under mottagning (m<sup>2</sup>)

$\sigma$  = målets radartvärsnitt (m<sup>2</sup>)

$T_i$  = koherent integrationstid (s)

$k_B$  = Boltzmann-konstanten = 1.38 10<sup>-23</sup> J/K

$T_s$  = effektiv brustemperatur på mottagarens ingång (K)



$L$  = förluster pga kablage, fasbrus, signalbehandling etc (-)

$r$  = kortaste avståndet (tvärsavståndet) mellan antenn och mål utefter flygbanan

$\Delta\theta$  = aperturvinkel

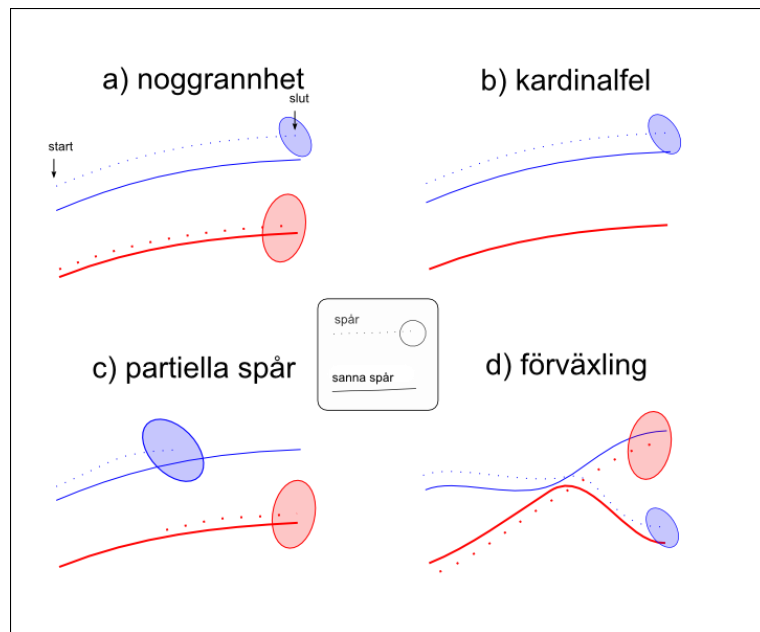
Ekvationen 2 beskriver relationen mellan signalen från ett mål och radarmottagarens egenbrus, vilket inte riktigt är det som söks utan man måste dessutom relatera signalen till klottret (icke önskad signal) för att få en uppfattning om kontrasten mellan mål och bakgrund. Detta kan ske genom att i ovanstående byta ut målets målarea mot något som kallas specifik målarea. Detta är ett mått som försöker sammanfatta medelvärdet av returerna som funktion av marktyp och infalls- och utfallsvinkel, se [44] men observera att denna modell är utvecklad för flygburen radar och den är inte verifierad för SAR-radar.

## 4 Estimering av prestanda för signalbehandling

### 4.1 Värdering av målföljningsprestanda hos sensorsystem

Målföljning (eng. tracking) är en viktig komponent i många militära tillämpningar, såsom spaning, övervakning, bekämpning, etc. I det enkla fallet har man bara ett mål att följa. Då är det trivialt att utvärdera målföljarens prestanda – har man spår på målet eller inte, och med vilken precision. Det blir mer komplicerat när man har multipla spår att följa, då ett antal olika fel kan uppkomma (se nedan).

Ett bra värderingsmått är värdefullt, då man ofta vill kunna göra jämförelser mellan olika algoritmer, eller optimera parameterinställningar. Sådana utvärderingar har visat sig behövas vid flera tillfällen, både för teoretiska och praktiska tillämpningar. Syftet med detta arbete är att hitta ett utvärderingsmått, färdigt eller egenutvecklat, som passar våra behov. Det ska vara så enkelt som möjligt och återspegla verkligheten (den upplevda kvaliteten på målföljningen) som kvalitetsmått.



Figur 12. Exempel på olika typer av spårningsfel. Spår är ritade med prickade linjer, hel-dragna är sanna spår (ground truth). Alla objekt/spår rör sig från vänster till höger. Ellips-erna motsvarar osäkerheten i spårens lägesuppskattning. Se texten för mera förklaring.

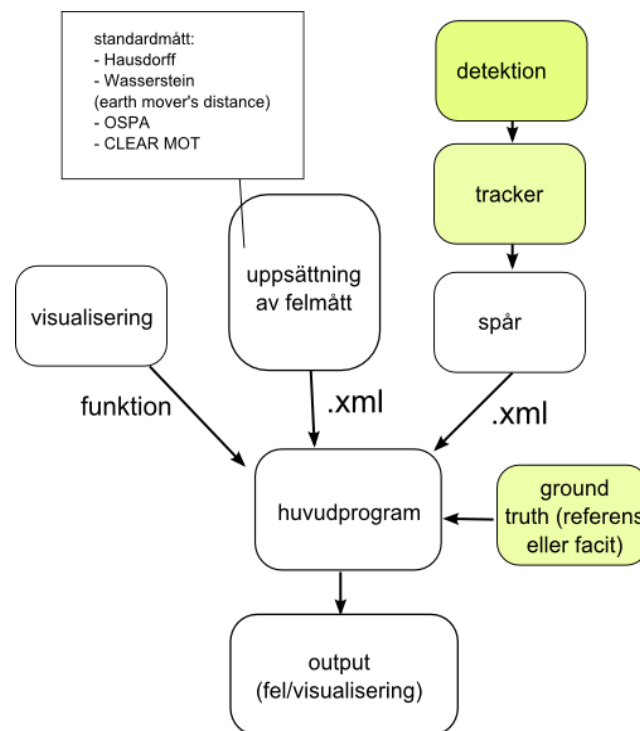
Det finns fortfarande ingen standardiserad metodik för detta [45]. Anledningen till det är svårigheten att jämföra de olika typer av fel som uppkommer vid tracking av multipla mål:

- noggrannhet: avvikelsen av spåret från det sanna målet.
- kardinalfel: när antalet spår är fler än existerande mål (falska positiva) eller färre (sanna negativa).
- partiella spår: spår som bara täcker delar av det sanna målets spår.
- förväxling: när ett spår hoppar mellan flera sanna mål.

- e) associationsfel: om de data som anses höra till det aktuella spåret väljs på ett felaktigt sätt uppstår effekter som t.ex. att man tappat spår eller att de estimerade parametrarna blir instabila eller felaktiga.

Att vikta dessa olika typer av fel är att jämföra äpplen och päron; det är en subjektiv fråga som avgörs av tillämpning och kontext. Forskning och utveckling här handlar därför om att ta fram en bra metodik för värderingsmått, där vikten av olika fel handhas konsekvent, och där måttet är definierat för alla möjliga fall.

Tidiga värderingsmått, t.ex. *Hausdorff distance* och *Wasserstein (earth mover's) distance*, är primitiva och ofullständiga: de innefattar inte alla typer av fel och är odefinierade för en del tillstånd (t.ex. om det inte förekommer sanna mål, men spår har genererats). Detta har motiverat utveckling av bättre mått. Fastän det inte finns någon standardmetodik, så har en del framsteg gjorts inom forskningsområdet nyligen. Vi har funnit två mått av speciellt intresse: *Optimal Subpattern Assignment (OSPA)*-mättet [45] och *CLEAR (Classification of Events, Activities and Relationships) MOT*-mättet (Multiple Object Tracking [46]). OSPA har fördelen att det är ett mått som integrerar samtliga feltyper i en skalär. Det är dock generellt definierat och har 3-5 fria parametrar (dessa kan dock elimineras med rimliga antaganden). Det innefattar också en optimering som har hög komplexitet (dvs. det blir en långsam beräkning om det är många mål). CLEAR MOT är enklare, med bara en fri parameter, men är egentligen två olika mått: ett för *precision* (CLEAR MOTP) och ett för *fel* (CLEAR MOTA).



Figur 13. Komponentbeskrivning av utvärderingsverktyg för multipelspår. De gröna boxarna – detektion, tracker, och ground truth är det som användaren behöver för att tillämpa verktyget. Med hjälp av konfiguration kan användaren också ange vilka mått och vilken visualisering som ska användas.

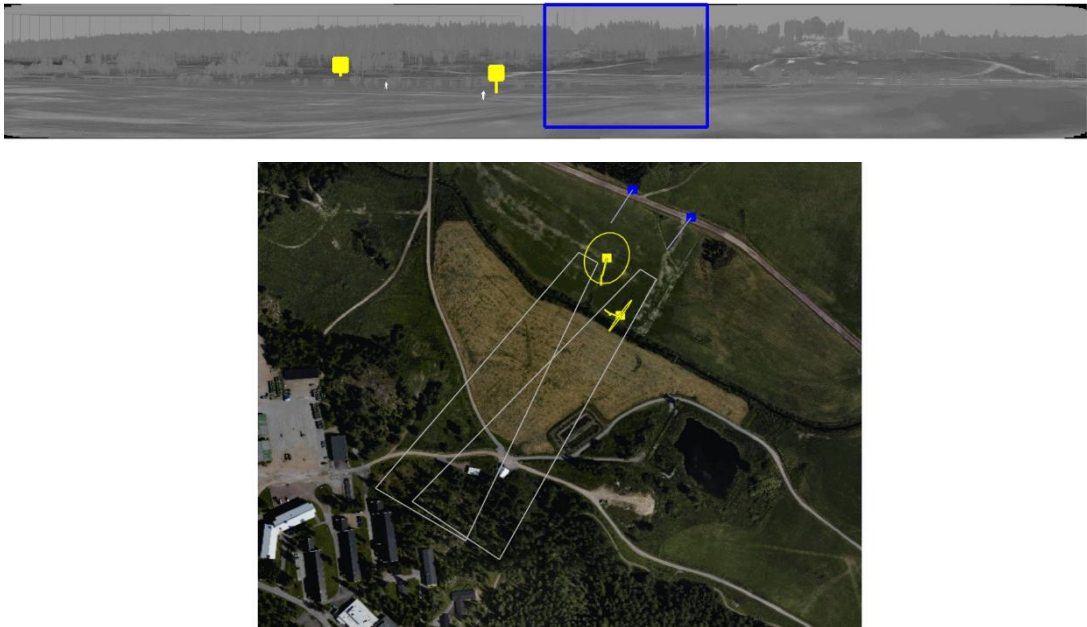
Inom projektet Värdering av sensorsystem har vi tagit fram ett konfigurationsstyrt mjukvaruverktyg (se figur 13) för multipelspårning. Det kan användas för t.ex.

- Jämförelse av prestanda hos olika målföljningsalgoritmer
- Parameteroptimering för målföljningsalgoritmer

- Mätning och spåranalys i en viss tillämpning
- Jämförelse av beteenden hos olika mått
- Visualisering av spårning

En förutsättning är att man har någon typ av referens eller facit för sina spår (eng. *ground truth*), d v s trajektorier från de sanna objekten. Man behöver också spår eller alternativt detektioner. Man kan sedan konfigurera vilka felmått man vill använda sig av, vilken visualisering man vill ha, och eventuellt vilken tracker man vill generera spår med. Mjukvaran är skriven generiskt för att kunna handha vilka typer av spår som helst (t.ex. 2D eller 3D rumskoordinater) och är tänkt att kunna tillämpas för vilken sensorutvärdering som helst där multipelspårning förekommer. En ambition är också att utveckla ett gränssnitt mot simuleringsverktyget MSSLab, som presenteras i avsnitt 5.2, för att enkelt kunna utvärdera spårning i det verktyget. Simuleringar i t.ex. MSSLab är ett värdefullt verktyg för spårningsanalys eftersom man har *ground truth* gratis, till skillnad från riktiga mätningar där det kan vara svårt och dyrt att ta fram en noggrann *ground truth*.

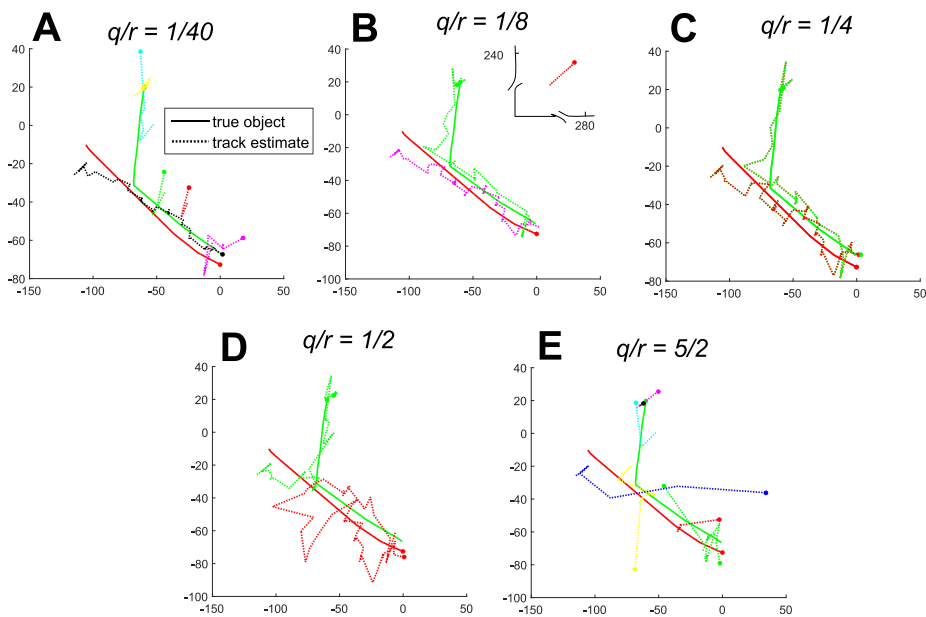
En del utvecklingsarbete kring användargränssnitt, samt testning av verktyget kvarstår.



Figur 14. Ett tillämpat exempel. En IR-kamera panorerar över ett fält med en aktiv persondetektor. Ovan: mosaik av sensors indata. Blå rektangel visar den aktuella sensorvyn. Gula flaggor indikerar aktuella spår. Observera de två personerna i bilden, i närheten av spåren. Nedan: kartvy av panoreringen. Blå rektanglar är sensorplattformens läge, med aktuell sensorriktning indikerad. Sensorvyn är utritad, grå fyrkant. Aktuella spår i gult. (Endast den vänstra sensorn är relevant i den här studien.)

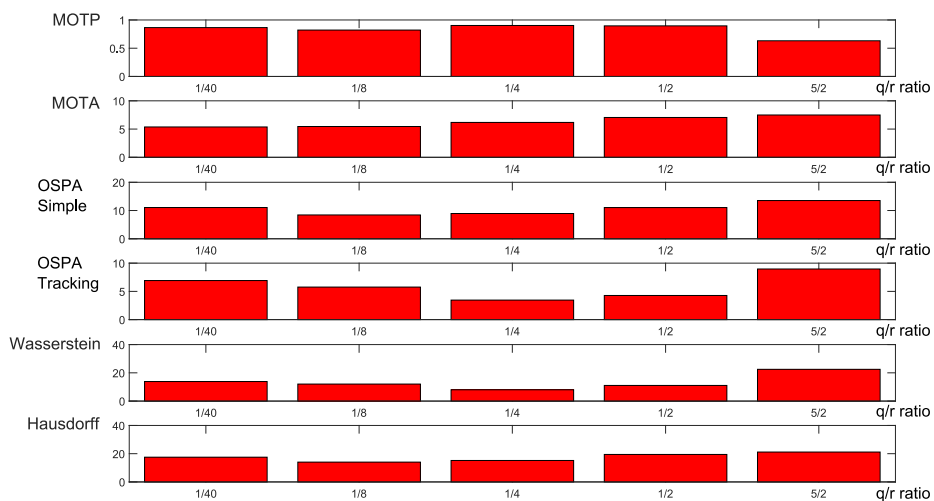
### Ett tillämpningsexempel

Här illustrerar vi ett exempel på hur värderingsmått kan tillämpas. I det valda scenariot panorerar en IR-kamera ut över ett fält, med en aktiv persondetektor (figur 14). Vi vill generera spår utifrån detektionerna, och vi vill veta vilket parametervärde vi ska sätta på kvoten ( $q/r$ ) mellan rörelse- ( $q$ ) och observations- ( $r$ )-modellerna, för att vår tracker ska ge så rimliga tolkningar som möjligt. ( $r^2$  och  $q^2$  är variansen i både  $x$  (longitud) och  $y$  (latitud) för respektive modell. Kvoten  $q/r$  avgör hur tungt förutsägelsen av objektets rörelse ska väga gentemot nya observationer.)



Figur 15. Parametersättningens effekt på målföljningen. 5 olika parametersättningar A-E, med ökande av kvoten  $q/r$  (mer vikt för rörelsemodellen), ger olika spårtolkningar. De sanna objektens rörelse är visade med heldragna linjer (grönt och rött), med punkter i slutet av de sanna spårerna. Spår (tracks) visas med streckade linjer med unik färg för varje spår. I fall C så svarar spårerna mot de sanna objekten - två personer som möts. I D sker en förväxling av spårerna där personerna möts.

Målföljningsalgoritmen är ett Extended Kalman Filter (EKF) med en kamera-observationsmodell (perspektivkamera) som kördes offline på sensordata med 5 olika värden på kvoten  $q/r$ . Vi kan se mycket olika resultat beroende på parametersättning. Vid för låga och höga kvotall klarar inte målföljaren av att associera detektionerna, och därför genereras många spår från samma objekt (figur 15A/E). Ett optimum kan vi se i figur 15C, där spårerna stämmer överens med de sanna objekten. Något lägre tal (figur 15B) ger nästan rätt resultat, medan med något högre (figur 15D) så förväxlas spårerna.



Figur 16. Kumulativt fel (normaliserat med antal tidsteg) för fem olika  $q/r$ -kvoter (svarar mot Figur 15), för sex olika prestandamått. (Samtliga mått har utvärderats med en cutoff distance på 20 m).

Finns det ett kvantitativt mått som speglar de kvalitativa resultat vi kan utvärdera i Figur 15? Vi mätte upp det kumulativa felet i scenariot för 6 olika felmått: CLEAR MOTP & MOTA [46], Optimal Subpattern Assignment (OSPA) ”Simple” (per tidssteg [45], OSPA ”Tracking” [47], Wasserstein- [45] och Hausdorff-avstånd [45]. I figur 16 ser vi att OSPA Tracking och Wasserstein-avstånd stämmer mest väl överens med vår kvalitativa bedömning. Där får vi lägst fel för de ”bästa” spåren ( $q/r = 1/4$ , Figur 15C) och högst för de med ”sämsta” ( $q/r = 1/40$  respektive  $5/2$ , Figur 15A/E).

Exemplet här är bara en illustration – i en skarp tillämpning studerar man många fler parameteruppsättningar, och utsätter trackern för mycket mer sensordata. Vi får dock redan här en fingervisning om att metoden är tillämpbar. Mera detaljerade studier av felmått och hur de ska tillämpas och jämföras pågår.

## 4.2 Bayesianisk skattning för estimering av måldiskrimineringsförmåga

Ett relativt nytt sätt att estimeras sensorprestanda är med hjälp av Bayesianisk skattning. Dessa metoder bygger på statistisk teori och de matematiska grunderna har utvecklats i USA av Ulf Grenander och hans medarbetare med stöd av det amerikanska försvaret. En lärobok i ämnet [48] sammanfattar mycket av arbetet. Andra representativa publikationer är [49] som analyserar orienteringsskattning i termiska IR-bilder, samt [50] som analyserar med vilken säkerhet man vid olika brusnivåer kan diskriminera mellan olika målhypoteser i visuella bilder. En intressant tillämpning redovisas i [51, 52], nämligen en analys av med vilken noggrannhet rymdteleskopet Hubbles orientering kan bestämmas från ett eller flera markbaserade teleskop.

Metodiken är rent teoretisk och bygger på en noggrann matematisk modellering av sensor, scen och målsignatur. Den stora fördelen med metodiken är att resultaten som erhålls är algoritmoberoende i den meningen att den anger prestandagränser som inte kan överträffas av någon algoritm (eller människa). Det betyder att resultatet ger ett direkt mått på informationsinnehållet i sensordata, snarare än prestanda för en enskild algoritm applicerad på sensordata. Behovet av algoritmoberoende mått på sensorprestanda har sin grund i att det i allmänhet inte är möjligt att matematiskt bestämma den optimala algoritmen för att klassificera eller skatta målparametrar ur en viss typ av sensordata.

Bayesianska metoder förutsätter att samtliga företeelser som påtagligt påverkar utfallet av en mätning kan modelleras med statistiska fördelningsfunktioner, samt att dessa kan analyseras analytiskt eller numeriskt. Eftersom detta är ett mycket omfattande arbete bedöms det vara svårt att hantera generellt. Ett viktigt exempel på begränsningar hos metodiken är att bakgrundens (naturens) variationsrikedom inte kan modelleras matematiskt, vilket medför att detektionsprestanda inte kan adresseras utom för speciella väldefinierade miljöer.

Metodiken kan delas upp i fem olika steg:

1. 3D CAD-modeller av målen konstrueras. Modellerna är geometriskt detaljerade för att kunna fånga upp olika sensorfenomen (sensorberoende) och har relevanta materialbeskrivningar. Modellerna är grunden för olika fysikaliska beräkningar, t.ex. termodynamiska beräkningar i IR-fallet.
2. Probabilistisk signaturmodell. I signaturen ingår att matematiskt/statistiskt beskriva hur målets egenskaper (så som de registreras av en sensor) varierar med väder, tid på dygnet, målets tillstånd (t ex motor avstängd, skyld), etc.

3. Probabilistisk sensormodell. I sensormodellen (eller kanske hellre modell för avbildningsprocessen) ingår sensors position relativt målet, distorsion av ideala data beroende på atmosfär, optisk överföringsfunktion, A/D-omvandling i detektorn (mottagaren), sensorbrus i mottagaren.
4. Bayesiansk formulering av estimeringsproblemet. Klassificerings-, och parameterskattningsproblemen formuleras här på ett sätt som en optimal algoritm skulle lösa dem, dvs. så att sannolikheten för felklassificering respektive osäkerheten i det skattade parametervärdet minimeras.
5. Beräkningar, analytiska eller numeriska, av tillhörande, ofta ytterst komplicerade matematiska uttryck. I många realistiska fall leder den Bayesianska formuleringen till så stort beräkningsbehov (t.ex. sökningar i stora databaser) att algoritmen inte kan tillämpas i en skarp situation. Däremot kan metoden användas för att bestämma prestandagränser.

De osäkerhetsmått som metodiken genererar kan användas för att fastställa krav på sensoregenskaper för att en viss prestandanivå ska uppnås i en given tillämpning. Exempel på krav på sensoregenskaper är upplösning och brusnivå. Med osäkerhetsmått kan man sedan exempelvis värdera om en viss sensor eller sensorkombination är lämplig för en målsökare som har vissa krav på precision. Om sensors egenskaper istället är givna, kan analysen användas för att optimera datainsamlingen, exempelvis med avseende på positionering av sensorn. Metodiken kan också användas för att bedöma värdet av att fusionera data från flera sensorer, eller vid jämförelse av informationsinnehållet i data från olika våglängdsband eller från olika sensortyper.

## 5 Estimering av prestanda för sensorsystem

Sensorprestanda beror på scenariovariabler som målobjekt, miljö, tidpunkt, datum, etc. Det betyder att det inte går att bestämma en allmängiltig prestanda för en sensor utan man måste ange under vilket scenarier som prestanda gäller. Man kan anta en viss fördelning av scenarier och få fram ett medelvärde över förekommande scenarier. För framtagning av sensorprestanda för vissa speciella situationer kan både mätningar och simuleringar användas.

### 5.1 Mätningar

Mätningar är en mycket vanlig metod för att för att få fram underlag för estimering av sensorprestanda för befintliga sensorsystem. Med mätningar kan effekter av olika scenariovariabler såsom målobjekt, ljusförhållanden, väder, vegetation etc. relativt enkelt studeras genom att mäta vid olika tidpunkter och vid olika platser. Effekterna av vissa sensorvariabler som t.ex. vissa vädertyper är dock svåra att studera eftersom vädret inte går att styra. Det kan även vara mycket svårt eller kostsamt att få tillgång till speciella målobjekt. Mätningar är även mycket värdefulla för att verifiera simulerade och analytiska resultat. De är också viktiga som underlag för att bestämma materialparametrar som ingår i många högkvalitativa sensorsimuleringsprogram.

De största nackdelarna med mätningar är att det är mycket kostsamt och tar tid att ta fram ett sensorsystem och genomföra mätningar. Detta gäller speciellt om mätningar ska göras från en flygande plattform. Vanligtvis görs därför endast mätningar för några fall för att verifiera prestanda eller simuleringsresultat. Att undersöka prestanda för framtida sensorsystem genom att vidareutveckla ett befintligt sensorsystem brukar även det vara mycket kostsamt.

### 5.2 Simuleringsbaserade metoder

Ett vanligt sätt att estimerar prestanda hos ett system är att använda simuleringsbaserade metoder. Detta förutsätter god kännedom om det aktuella systemet samt validerade modeller för den omgivning som systemet skall operera i. Fördelen med simuleringsbaserade metoder är att när väl systemet är modellerat och omgivningsmodellerna finns, så kan simuleringar göras för en mängd olika fall till en rimlig kostnad och på relativt kort tid.

Ett vanligt sätt att estimerar sensorprestanda är med hjälp av Monte Carlo-metoder. Detta är en klass av metoder som har gemensamt att ett stort antal upprepade försök utförs med slumpmässiga variationer av parametrar för att skatta prestanda hos ett system. Dessa metoder kan användas för att lösa både deterministiska och stokastiska problem där analytiska metoder saknas eller är alltför komplicerade.

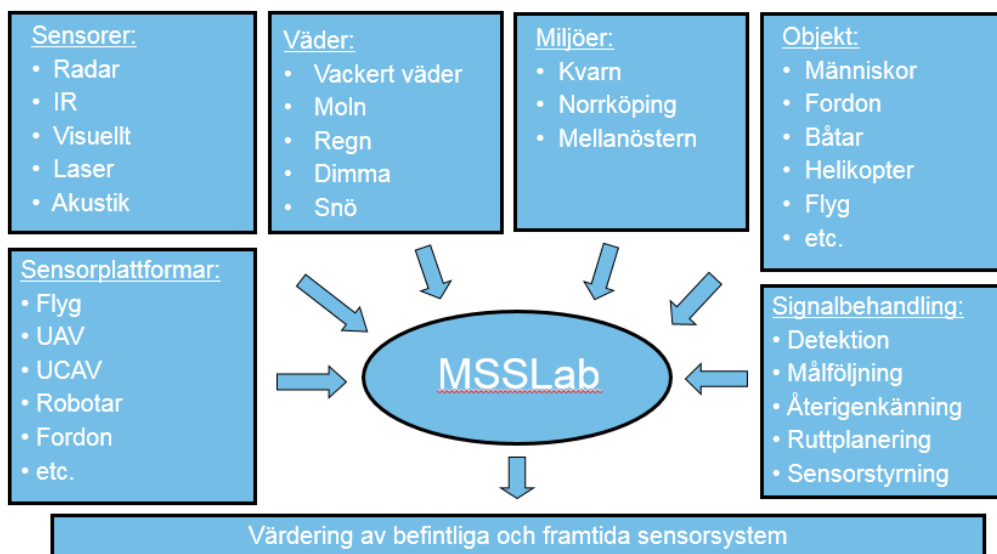
Monte Carlo-metoder inleds med att täthetsfunktioner bestäms. Dessa täthetsfunktioner ska beskriva tillstånden som är möjliga för systemet. En slumpmässig sampling från täthetsfunktionen görs sedan och deterministiska simuleringar görs med dessa indata. Indata kan till exempel vara väder, flygbanor, målets placering och orientering som inte kan antas exakt kända i en verklig situation.

Utfallen från alla individuella beräkningar kan sedan behandlas statistiskt för att skatta systemets prestanda. Sensorprestanda som kan skattas med denna metod är t.ex. sannolikheten att ett mål detekteras av en viss detektionsalgoritm i en viss given situation.



Metoderna ger även möjlighet att värdera alternativa ingrepp i ett system genom att producera jämförbara resultat från experiment med olika parametervärden.

Ett mycket användbart simuleringsverktyg för estimering av sensorprestanda är MSSLab (MultiSensorSimuleringsLab) [53]. Detta är en simuleringsmiljö där avancerade sensorsystem kan simuleras i olika miljöer och vid olika väderförhållanden och tidpunkter. Simuleringsverktyget möjliggör simulering av dynamiska scenarier med rörliga plattformar, sensorbärare, målobjekt etc. Högkvalitativa sensorsimuleringar kan göras med IR, visuellt, laser och radar och dessa sensorsystem kan simuleras både enskilt, som samverkande sensorsystem och som multisensorsystem. Avancerade algoritmer för objekt-detektion, följning, återigenkänning, sensorstyrning, ruttplanering är även integrerade i simuleringsverktyget.



Figur 17. Estimering av sensorprestanda med MSSLab.

## 6 Slutsatser och fortsatt arbete

I denna rapport har vi gjort en sammanställning av olika metoder och verktyg för estimering av prestandamått för både optronik- och radarsystem. För optroniksystem finns det många verktyg för estimering av sensorprestanda, men för radarsystem mot markmål finns det dock inte lika många färdiga verktyg.

Metoder som TTP och TRM har visat sig användbara vid värdering av avbildande elektrooptiska system ur ett användarperspektiv, framför allt för relativa jämförelser. Generellt kan överensstämmelsen även vara relativt god i absoluta termer mellan predikterade och verkliga räckvidder, men detta kräver mycket god kännedom kring aktuella sensorsystem, scenarion (mål/bakgrund, väder, ljusförhållanden, osv.) och diskrimineringsuppgiften. Det finns dock begränsningar med metoderna som exempelvis att de är statiska och inte tar hänsyn till scendynamik (t.ex. målrörelser) eller att bakgrundsklotter inte kan hanteras tillfredsställande. Målrepresentationen är också relativt grov, där detaljer och texturer som kan ha betydelse för diskriminering inte fångas upp på ett realistiskt sätt. I viss mån kan detta hanteras med de inbyggda kalibreringsvärdena i modellerna ( $N_{50}$  och  $V_{50}$ ). För att erhålla bättre överensstämmelse med verkligt uppmätta eller upplevda räckvidder krävs därför ytterligare validering av metoderna, med perceptionsförsök mot verkliga mål i realistiska miljöer, i syfte att kalibrera in dessa värden bättre.

GIQE är en bra modell för prediktion av kvalitetsmått som bildtolkare av flygbilder använder (NIIRS), men har kanske inte samma användbarhet på mark- och sjöarenorna där olika typer av sannolikhetsmått är mer naturliga. Kopplingar mellan GIQE och TTP finns, men bör vidare studeras. Detta gäller även metoder baserade på statistik i bilder.

Ett verktyg för värdering av målföljningsprestanda har tagits fram för att möjliggöra utvärdering av målföljningsalgoritmer för multipla spår och för att kunna göra jämförelser mellan olika algoritmer, eller för att optimera parameterinställningar. Det framtagna verktyget tar hänsyn till både måldetektioner och målspar, och metoden har visat sig vara tillämpbar och ger kvalitetsmått som återspeglar verkligheten. En del utvecklingsarbete kring gränssnittet för verktyget kvarstår.

I det fortsatta arbetet planerar vi att jämföra några olika värderingsmetoder mot COAT/TVS-metoden som har använts hitintills i projektet. Under 2013 och 2014 studerades ett tillämpningsexempel som handlade om värdering av spaningsförmågan för tre relativt olika UAV-system: UAV Örn, MQ-9 Reaper och MQ-8B Fire Scout. I ett fortsatt arbete vill vi jämföra den tidigare värderingen med andra värderingsmetoder såsom: House of Quality (HOQ), Analytisk Hierarki-process (AHP) och Morfologisk analys.

## 7 Referenser

- [1] B. Asp, C. Carling, A. Hunstad, B. Johansson, P. Johansson, J. Nilsson och Å. Waern, "COAT - Användarguide", FOI-R--2409--SE.
- [2] Näsström F., Allvar J., Bergström D., Grahn P., Forsgren R., Hedström J., Persson A., Ulvklo M., "Slutrapport: Simuleringsbaserade metoder för sensorvärdering – Värdering av spaningsförmågan för obemannade system", FOI-R--3993--SE, december 2014.
- [3] Driggers R. G., Kelley M., "National imagery interpretation rating system and the probabilities of detection, recognition and identification", *Optical Engineering*, Vol.36, Nummer 7, 1997.
- [4] Irvine, J. M.; "National Imagery Interpretability Rating Scales (NIIRS): Overview and methodology", *SPIE Vol. 3128*, 1997.
- [5] Leachtenauer, J. C.; Driggers, R. G.; "Surveillance and reconnaissance imaging systems: Modeling and performance predictions", Artech House Publishing, 2001.
- [6] Leachtenauer J. C., Malila W., Irvine J., Colburn L., Salvaggio N. "General image-quality equation: GIQE", *Applied Optics*, Vol. 36, Nummer 32, 1997.
- [7] Leachtenauer J. C., Malila W., Irvine J., Colburn L., Salvaggio N. "General image-quality equation for infrared imagery", *Applied Optics*, Vol. 39, Nummer 26, 2000.
- [8] Motion Imagery Standard Board, "Video-National Imagery Interpretability Rating Scale", Standard MISB ST 0901.2, NGA Case 09-137, 27 Februari 2014.
- [9] Young, D. L., Bakir, T., "Draft Motion Imagery Quality Equation (MIQE)", Civil Commercial Imagery Evaluation Workshop, Fairfax, Virginia, April 2009.
- [10] Mittal A., Moorthy A., och Bovik A., "No-reference image quality assessment in the spatial domain", *IEEE Trans Image Process*, Vol. 21, sidor 4695 - 4708, 2012.
- [11] Mittal A., Moorthy A., och Bovik A., "Making image quality assessment robust", *Proc Int Conf Signals, Systems and Computers*, sidor 1718 - 1722, 2012.
- [12] Mittal A., Muralidhar G., Ghosh J., och Bovik A., "Blind image quality assessment without human training using latent quality factors", *IEEE Signal Process Lett*, Vol. 19, sidor 75 - 78, 2012.
- [13] Mittal A., Soundararajan R., och Bovik A., "Making a completely blind image quality analyzer", *IEEE Signal Process Lett*, Vol. 22, sidor 209 - 212, 2013.
- [14] Moorthy A. och Bovik A., "A two-step framework for constructing blind image quality indices", *IEEE Signal Process Lett*, Vol. 17, sidor 513 - 516, 2010.
- [15] Moorthy A. och Bovik A., "Blind image quality assessment: from natural scene statistics to perceptual quality", *IEEE Trans Image Process*, Vol. 20, sidor 3350 - 3364, 2011.
- [16] Moorthy A. K. och Bovik A. C., "Visual Importance Pooling for Image Quality Assessment", *Selected Topics in Signal Processing*, *IEEE Journal of*, Vol. 3, sidor 193-201, 2009.
- [17] Saad M., Bovik A., och Charrier C., "Blind image quality assessment: a natural scene statistics approach in the DCT domain", *IEEE Trans Image Process*, Vol. 21, sidor 3339 - 3352, 2012.
- [18] Sheikh H., Bovik A., och Cormack L., "No-reference quality assessment using natural scene statistics: JPEG2000", *IEEE Trans Image Process*, Vol. 14, sidor 1918 - 1927, 2005.
- [19] Wang Z., Bovik A., Sheikh H., och Simoncelli E., "Image quality assessment: from error visibility to structural similarity", *IEEE Trans Image Process*, Vol. 13, sidor 600 - 612, 2004.
- [20] Wang Z., Sheikh H., och Bovik A., "No-reference perceptual quality assessment of JPEG compressed images", *Proc IEEE Int Conf Image Process*, sidor 477 - 480, 2002.
- [21] Zhou W. och Bovik A. C., "Reduced- and No-Reference Image Quality Assessment", *Signal Processing Magazine, IEEE*, Vol. 28, sidor 29-40, 2011.

- 
- [22] Dosselmann R. och Yang X., "A comprehensive assessment of the structural similarity index", *Signal, Image and Video Processing*, Vol. 5, sidor 81-91, 2011.
- [23] Gustavsson D., "On Texture and Geometry in Image Analysis," PhD-thesis, Department of Computer Science (DIKU), University of Copenhagen, 2009.
- [24] Yanulevskaya V. och Geusebroek J.-M., "Significance of the Weibull Distribution and its Sub-models in Natural Image Statistics", 2009.
- [25] Mallat S., "A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way", Academic Press, 2008.
- [26] Mallat S. G., "A theory for multiresolution signal decomposition: the wavelet representation", *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, Vol. 11, sidor 674-693, 1989.
- [27] Jingsang H. och Mumford D., "Statistics of natural images and models", i *Computer Vision and Pattern Recognition*, 1999. IEEE Computer Society Conference on., Vol. 1, pp. 1-547, 1999.
- [28] Mumford D. och Desolneux A., "Pattern Theory: The stochastic analysis of real-world signals", A K Peters, Ltd, 2010.
- [29] Joachims T., "Optimizing search engines using clickthrough data", presenterad vid Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, Edmonton, Alberta, Canada, 2002.
- [30] Joachims T., "Training linear SVMs in linear time", presenterad vid Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, Philadelphia, PA, USA, 2006.
- [31] Freund Y., Iyer R., Schapire R. E., och Singer Y., "An efficient boosting algorithm for combining preferences", *J. Mach. Learn. Res.*, Vol. 4, sidor 933-969, 2003.
- [32] Schapire R. E. och Freund Y., "Boosting: Foundations and Algorithms", The MIT Press, 2012.
- [33] Zheng Z., "A regression framework for learning to rank in web information retrieval," State University of New York at Buffalo, 2008.
- [34] Zheng Z., Chen K., Sun G., och Zha H., "A regression framework for learning ranking functions using relative relevance judgments", presenterad vid Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, Amsterdam, The Netherlands, 2007.
- [35] Holst G., "Electro-optical imaging system performance", 5th Ed., JCD Publishing, 2003.
- [36] Bijl P., Valetton J.M.; "TOD, a New Method to Characterize Electro-Optical System Performance", *Proc. SPIE*, Vol. 3377, s. 182-193, 1998.
- [37] Driggers R., Friedman, M. H., Nichols, J. M.; "Introduction to Infrared and Electro-Optical Systems", 2nd Ed., Artech House, 2012.
- [38] Vollmerhausen R. H., Jacobs E., "The Targeting Task Performance (TTP) Metric – A new model for predicting target acquisition performance", Technical report AMSEL-NV-TR-230, NVESD, 2004.
- [39] Vollmerhausen R. H., Jacobs E., Driggers R. G., "New metric for predicting target acquisition performance", *Proc. SPIE*, Vol. 5076, s. 28-40, 2004.
- [40] Wittenstein W., "Thermal range model TRM3", *Proc. SPIE*, Vol. 3436, s. 413-424, 1998.
- [41] Wittenstein W., "Minimum temperature difference perceived—a new approach to assess undersampled thermal imagers", *Optical Engineering*, Vol. 38, Nummer 5, s. 773-781, 1999.
- [42] Gal, R., Wittenstein, W., Weiss, R., Schuberth, W.; "TRM4.v1 Parameter Description and Model Documentation", Fraunhofer IOSB Technical Report, Augusti 2010.
- [43] Nolan A. R., Goley, G. S. och Bakich M., "Performance Estimation of SAR imagery using NIIRS techniques", *Proc. of SPIE* Vol. 8394, 2012.
- [44] Lindquist T., "Markekomodell för radartillämpningar", LM Ericsson, rapport MI/Xe 3043, 1974.

- 
- [45] Schuhmacher D., Vo B.-T., och Vo B.-N., A consistent metric for performance evaluation of multi-object filters. *Signal Processing, IEEE Transactions on*, 56(8), 3447–3457, 2008.
- [46] Bernardin K., och Stiefelhagen R., Evaluating multiple object tracking performance: the CLEAR MOT metrics. *EURASIP Journal on Image and Video Processing*, 2008.
- [47] B. Ristic, B.-N. Vo, D. Clark, and B.-T. Vo, “A metric for performance evaluation of multi-target tracking algorithms,” *Signal Processing, IEEE Transactions on*, vol. 59, no. 7, pp. 3452–3457, 2011.
- [48] Grenander U., Miller M., “*Pattern Theory: From Representation to Inference*”, Oxford University Press. 2007.
- [49] Cooper M. L., och Miller M. I., “Information Measures for Object Recognition Accommodating Signature Variability”, *IEEE Transactions on Information Theory*, 46(5), 1896-1907, 2000.
- [50] Grenander U., Srivastava A. och Miller M. I., “Asymptotic Performance Analysis of Bayesian Target Recognition”, *IEEE Transactions on Information Theory*, 46(4), 1658-1665, 2000.
- [51] Gerwe D. R., Idell P. S., “Cramér-Rao analysis of orientation estimation: viewing geometry influences on the information conveyed by target features”, *Journal of the Optical Society of America A*, 20(5), 797-816, 2003.
- [52] Gerwe D. R., Idell P. S., “Cramér-Rao analysis of orientation estimation: influence of target model uncertainties”, *Journal of the Optical Society of America A*, 20(5), 817-826, 2003.
- [53] Näsström F., Allvar J., Berggren J., Chevalier T., Forsgren R., Habberstad H., Hedström J., Hemström f., Hendeby G., Karlsson M., Persson A., ”Simuleringsbaserade metoder för sensorvärdering – Årsrapport 2011”, FOI-R-3357--SE, december 2011.

FOI är en huvudsakligen uppdragsfinansierad myndighet under Försvarsdepartementet. Kärnverksamheten är forskning, metod- och teknikutveckling till nytta för försvar och säkerhet. Organisationen har cirka 1000 anställda varav ungefär 800 är forskare. Detta gör organisationen till Sveriges största forskningsinstitut. FOI ger kunderna tillgång till ledande expertis inom ett stort antal tillämpningsområden såsom säkerhetspolitiska studier och analyser inom försvar och säkerhet, bedömning av olika typer av hot, system för ledning och hantering av kriser, skydd mot och hantering av farliga ämnen, IT-säkerhet och nya sensorers möjligheter.



FOI  
Totalförsvarets forskningsinstitut  
164 90 Stockholm

Tel: 08-55 50 30 00  
Fax: 08-55 50 31 00

[www.foi.se](http://www.foi.se)