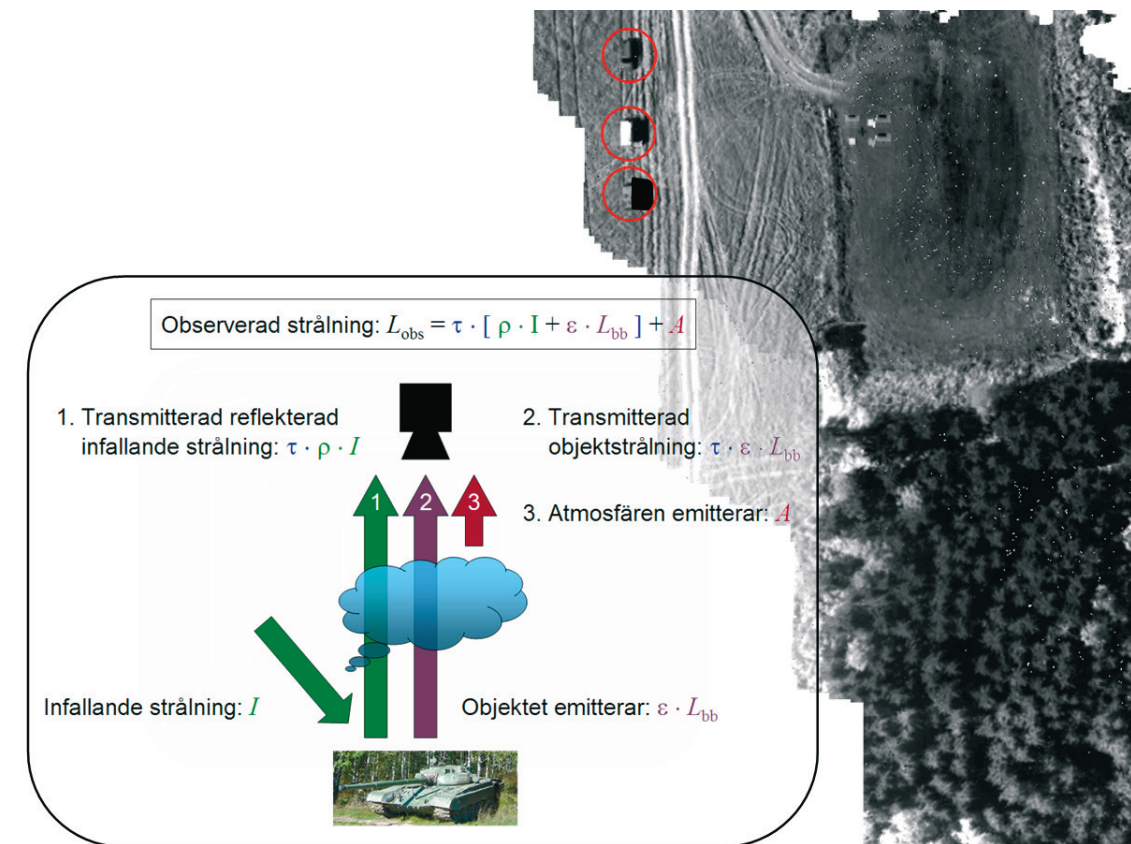




Analys av multi/hyperspektrala bilder –framsteg och resultat

JÖRGEN AHLBERG, MIKAEL LUNDBERG, JOAKIM RYDELL



FOI är en huvudsakligen uppdragsfinansierad myndighet under Förvarsdepartementet. 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
Informationssystem
Box 1165
581 11 Linköping

Tel: 013-37 80 00
Fax: 013-37 81 00

www.foi.se

FOI-R--2895--SE
ISSN 1650-1942

Användarrapport
December 2009

Informationssystem

Jörgen Ahlberg, Mikael Lundberg, Joakim Rydell

Analys av multi/hyperspektrala bilder – framsteg och resultat

FOI-R--2895--SE

Titel	Analys av multi/hyperspektrala bilder – framsteg och resultat
Title	Analysis of multi/hyperspectral images – progress and results
Rapportnr/Report no	FOI-R--2895--SE
Rapporttyp Report Type	Användarrapport User Report
Månad/Month	December
Utgivningsår/Year	2009
Antal sidor/Pages	31 p
ISSN	ISSN 1650-1942
Kund/Customer	Försvarsmakten
Projektnr/Project no	E53064
Godkänd av/Approved by	Ola Kärvell

FOI, Totalförsvarets Forskningsinstitut
Avdelningen för Informationssystem
Box 1165
581 11 Linköping

FOI, Swedish Defence Research Agency
Information Systems
Box 1165
SE-581 11 Linköping

Sammanfattning

Rapporten beskriver tre olika bildbehandlingsmetoder för tre olika ändamål. Alla tre är nödvändiga för att designa sensorsystem för att uppnå den eftertraktade förmågan att kunna spana mot ett område och detektera avvikelser, förändringar samt specifika spektrala signaturer (material).

Den första metoden handlar om *registrering* eller samkalibrering, det vill säga att hitta den rumsliga transformationen mellan två bilder som visar (ungefär) samma område. Korrekt registrering är en förutsättning för förändringsdetektion, mosaikbildning, bildfusion och superupplösning. Den beskrivna metoden tar avstamp i kända metoder och anpassar dem till militära behov.

Den andra metoden handlar om *atmosfärskorrektion* för att kunna studera spektrala signaturer hos de mål man spanar mot. Atmosfärskorrektion är en förutsättning för att kunna detektera specifika signaturer eller klassa pixlar efter material.

Den tredje metoden är en informationsteoretisk metod för *sensorprestandavärdering* och avser att ge ett användbart mått obereonde av detektionsmetod.

Nyckelord

hyperspektral, multispektral, bildbehandling, registrering, mosaik, informationsteori, atmosfärskorrektion

Abstract

This report describes three different image processing methods with three different purposes. All three are necessary for the design of sensor systems to enable the sought-for capability to, during reconnaissance of an area, detect anomalies, changes, and specific spectral signatures (materials).

The first method is for *registration*, *ie* to find the spatial transformation between two images depicting (approximately) the same area. Correct registration is needed for change detection, mosaicking, image fusion, and super resolution. The described method exploits known methods and adapts them to military needs.

The second method is for *atmospheric correction* in order to be able to study the spectral signatures of targets in the reconnaissance area.

The third method is an information theoretic measure for *sensor performance assessment* and is intended to give a useful measure independent of detection methods.

Keywords

hyperspectral, multispectral, image processing, registration, mosaic, information theory, atmospheric correction

Innehåll

1 Inledning	5
1.1 Multi- och hyperspektrala bilder	5
1.2 Översikt	6
2 Registrering / samkalibrering av heterogena sensorer	7
2.1 Bakgrund	7
2.2 Metodik	8
2.2.1 Förbehandling	8
2.2.2 Mjukvaruramverk för registrering	9
2.2.3 Egenskapsbaserad metod	9
2.3 Resultat	11
2.3.1 Olika våglängdsband (sensorfusion)	11
2.3.2 Olika sensorer (sensorfusion)	11
2.3.3 Olika tidpunkter (förändringsanalys)	11
2.3.4 Mosaikbildning (kartering)	15
2.4 Framtid	15
3 Atmosfärskorrektion i hyperspektralt termiskt infrarött	17
3.1 Bakgrund	17
3.2 Framsteg och resultat	18
3.3 Framtid	19
4 Informationsmått för sensorprestandavärdering	21
4.1 Bakgrund	21
4.2 Metod	21
4.3 Resultat	23
4.4 Framtid	24
5 Slutsatser	29
5.1 Registrering	29
5.2 Atmosfärskorrektion i hyperspektralt termiskt infrarött	29
5.3 Informationsmått för sensorprestandavärdering	29
Litteraturförteckning	31

1 Inledning

Den här rapporten beskriver framsteg och resultat inom bildanalys i FoT-projektet ASSM, ”Adaptiva spektrala spanings- och måligenkänningsensorer” som pågår 2009–2011. Projektets mål är:

1. Att föreslå, ta fram, värdera och demonstrera sensorer, sensorkombinationer och sensortekniker som förbättrar förmågan till upptäckt, klassificering och identifiering av markmål;
2. Att ta fram, värdera och demonstrera metoder för automatisering av och/eller operatörsstöd för ovanstående (t.ex automatiska måldetektionsmetoder);
3. Att ta fram förslag till design och effektivt nyttjande av multi-/hyperspektrala sensorer och sensortekniker i specifika applikationer.

Den här rapporten behandlar alltså de framsteg som hittills gjort för att uppnå mål (2).

På lång sikt är projektets syfte att föreslå sensorsystem med förbättrad förmåga till upptäckt, klassificering och identifiering av markmål (vid såväl mark-till-markspaning som flygspaning). Upptäckt av kamouflerade eller delvis skydda mål är projektets fokus, men tekniken kan även tillämpas på, t.ex, förändringsanalys och mindetektion.

1.1 Multi- och hyperspektrala bilder

En *multispektral* sensor är en elektro-optisk sensor (en kamera) som registrerar bilder i flera våglängdsband. En vanlig digitalkamera kan sägas vara multispektral (våglängdsband för rött, grönt resp. blått), men med ett tämligen litet antal (tre) våglängdsband. Trots detta är det uppenbart att man kan se skillnad på olika material i en färgbild, t.ex kan man se skillnaden mellan asfalt och grönt gräs. Genom att välja de våglängder man tittar på med viss omsorg, och dessutom gärna titta i fler än tre olika band, kan man upptäcka skillnader som det mänskliga ögat inte kan uppfatta, t.ex mellan grön kamouflagefärg och grön vegetation. Om man dessutom väljer sina våglängdsband i det termiskt infraröda (IR) området (mellanvågs- eller långvågs-IR) gör man sig oberoende av dagsljus eller aktiv belysning och erhåller alltså mörkerkapacitet.

Om man i stället för *flera* band vill ha *många* band får man byta till andra sensorteknologier, och gå över till *hyperspektrala* sensorer. Medan en multispektral sensor typiskt har 4–10 band har en hyperspektral sensor från några dussin till några hundra band, till priset av betydligt lägre bildtakt eller att sensorn är skannande i stället för stirrande (dvs tittar på en linje i stället för en hel bild i taget).

Både multi- och hyperspektrala sensorer finns på FOI, och beskrivs i, t.ex., [13].

För att använda sig av multi/hyperspektrala bilder krävs som regel olika former av mer eller mindre avancerad bildanalys. Till stor del kan man förstås utnyttja den uppsjö av metoder som finns tillgänglig för ”vanliga” bilder, men vissa problem måste lösas på speciella vis. Den här rapporten beskriver utveckling av sådana metoder som pågår på FOI, och för varje metod förklaras syftet och vad som gör metoden speciell för den här sensortypen.

1.2 Översikt

Rapporten innehåller tre kapitel som var och ett tar upp en bildanalysmetod. Varje kapitel tar upp bakgrund och motivering, följt av en metodbeskrivning, resultat samt framtid. Metodbeskrivningarna är något mer tekniska (men på en populärvetenskaplig nivå) och kan hoppas över av den som mest är intresserad av resultat och framtid.

Kapitel 2 behandlar registrering av bilder, dvs hur man samkalibrerar/linjerar bilder med varandra för att skapa mosaiker/kartor eller för förändringsdetektion. Problemet kan vara lätt eller svårt beroende på omständigheterna – i det enklaste fallet är bilderna från samma sensor och (nästan) samma tidpunkt (jämför s.k. ”stitching” i konsumentkameror), i det svåraste är sensorerna olika (olika band, olika upplösning, ...) och bilderna från olika tidpunkter.

Kapitel 3 handlar om hur man kan estimerar atmosfärs- och målparametrar vid flygspaning med en hyperspektral sensor i långvågs-IR (8–12 μm våglängd). I de våglängderna är atmosfären inte helt transparent, och dessutom en egen källa till strålning. Man vill därför automatiskt kunna kompensera för atmosfärens inverkan och dessutom estimerar parametrar hos det mål man spanar mot för att i förlängningen klassificera materialet (Sten? Gräs? Kamouflagefärg?) och dess temperatur.

Kapitel 4 innehåller ett mer teoretiskt resonemang för att värdera prestanda hos sensorer.

I kapitel 5 sammanfattas resultaten och slutsatserna från kapitel 2–4.

2 Registrering / samkalibrering av heterogena sensorer

2.1 Bakgrund

Vid spaning mot markmål med en flygburen sensor är det ofta viktigt eller till och med nödvändigt att *registrera* bilder till varandra, det vill säga att relatera bildernas koordinater till varandra i varje pixel. Som tidigare nämnts, är korrekt registrering en förutsättning för förändringsdetektion, mosaikbildning, bildfusion och superupplösning. De vanligaste exemplen på registrering är s.k. ”stitching” [8] för att skapa panoramabilder, som finns implementerat i vanliga konsumentkameror, samt sammanslagning av flygbilder till en större mosaik. För att kunna sätta ihop bilderna måste man förstås veta *hur* de ska sättas ihop, det vill säga man måste veta hur den ena bilden ska förskjutas, skevas, förstoras, förminskas, etc. för att den ska passa ihop med den andra. Flera bildexempel ges senare i detta kapitel.

Civil bildbehandlingsforskning driven av konsumentmarknaden har på senare år fört tekniken snabbt framåt (med hjälp av metoder som SIFT [12], SURF [6] och multi-band blending [7, 9]), och problemet kan tyckas vara löst. Det finns dock ett antal aspekter som den civila marknaden *inte* intresserar sig för, och där vi måste bedriva egen forskning och utveckling.

- Exempelvis en sensor som MultimIR [13] tar en bild i ett våglängsband, nästa bild i ett annat, osv.. Sekvensen av bilder ska registreras så att man får en stor mosaikbild med alla våglängdsbanden i varje pixel. Eftersom skillnaden i utseende mellan de olika banden är rätt drastisk, klarar de kommersiellt tillgängliga metoderna inte av det.
- Ett annat exempel är när vi flugit över ett område med två olika sensorer, t.ex en ”vanlig” kamera och en IR-sensor. Att automatiskt registrera dessa bilder till varandra är betydligt svårare än att registrera bilder från samma typ av sensor.
- Ett tredje exempel är att det dessutom kan ha gått lång tid mellan spaningsflygningarna, eller att vi vill registrera mot existerande data, såsom kartunderlag eller satellitbilder.
- Slutligen vill vi också kunna registrera data från helt olika sensortyper till varandra, t.ex en hyperspektral sensor (hög spektral upplösning, låg spatiell upplösning) med en 3D-laser (hög spatiell upplösning, 3D-information).

Förmågan att hantera dessa fall är nödvändig för att designa framtida mångsidiga spaningssystem för kartering, målföljning och förändringsdetektion.

I avsnitt 2.2 beskrivs tekniken som använts och utvecklats inom projektet. Avsnittet är ganska tekniskt och kan hoppas över för den som vill gå direkt till avsnitt 2.3 där ett antal bildexempel illustrerar resultaten i några av de ovan beskrivna exemplen. Slutligen beskrivs, i avsnitt 2.4 den planerade vidareutvecklingen.

2.2 Metodik

Bildregistreringsproblemet kan formuleras så att man givet två bilder (en referensbild och en testbild), vill estimeras den transform som överför den ena bilden (testbilden) till den andra bildens (referensbilden) koordinatsystem. Transformen inkluderar typiskt skalning, rotation och translation, men kan även inkludera skevning, perspektivförändring, kompensation för linsdistorsion etc. Man kan också uttrycka det som att man vill hitta positionen och riktningen (och kanske andra parametrar) för testbildens sensor (i förhållande till referensbildens sensor).

Man använder sig av två typer av metoder, pixelbaserade och egenskapsbaserade. Pixelbaserade metoder använder bildernas pixelvärden direkt för att hitta den transform som maximerar ett likhetsmått, t.ex korrelationen mellan bilderna. Egenskapsbaserade metoder försöker hitta ett antal lokala egenskaper (nyckelpunkter) i bilderna och matcha dem mot varandra.

Fördelar med egenskapsbaserade metoder är att de inte kräver någon initialisering, det vill säga de kan hitta rätt transform även om bilderna är kraftigt transformerade i förhållande till varandra (och man inte känner till hur). De är dessutom beräkningsmässigt snabba.

Fördelen med pixelbaserade metoder är att de kan göras generella så tillvida att man inte behöver ha någon särskild förkunskap om ingående sensordata, utom möjligen hur man gör om sensordata till en 2D-bild. Nackdelen är att man måste ha ett ganska bra initialt estimat av transformen, det vill säga sensorernas position.

2.2.1 Förbehandling

I många fall när det gäller att jämföra bilder som kommer från samma sensor eller sensortyp behövs ingen förbehandling av data för att kunna tolka informationen på samma sätt. I svårare fall där man har att göra med data från olika typer av sensorer krävs troligtvis någon slags förbehandling av data för att kunna underlätta registreringen av bilderna mot varandra. Förbehandlingen är tänkt som ett första steg där man konverterar inbilderna för att underlätta den efterföljande beräkningen att estimeras en transform (förskjutning, skalning, rotation etc.) som gör att bilderna beskrivs i samma koordinatsystem. Om man exempelvis vill registrera en färgbild bestående av tre färgband (röd, grön och blå) mot en gråskalebild med endast ett band behöver man kanske förbehandla färgbilden genom att konvertera även denna till gråskala. Man kan

även tänka sig fallet då man vill registrera en 2D-bild mot 3D-data. För att kunna åstadkomma en sådan registrering krävs att 3D-datat förbehandlas genom en transformation till en 2D-bild, sedd från (ungefär) samma riktning som 2D-sensors bild.

2.2.2 Mjukvaruramverk för registrering

Vi har påbörjat utvecklandet av ett generellt ramverk som kan användas både på homo- och heterogena sensordata. Är sensordata inte bilddata (utan, t.ex, 3D-data från en skannande laser) måste information om hur data görs om till en 2D-bild från en given tittriktning och position medfölja data. Systemet försöker sedan optimera fram den transform som maximerar den ömsesidiga informationen mellan de båda dataseten (bilderna). Detta är alltså en pixelbaserad metod. I de fall det är tillämpligt används i stället en egenskapsbaserad metod.

Optimeringsproblemet för den pixelbaserade metoden har visat sig vara svårare än väntat, medan den implementerade egenskapsbaserade metoden visat sig vara effektiv även på heterogena sensordata. Nedan redogörs därför för den senare.

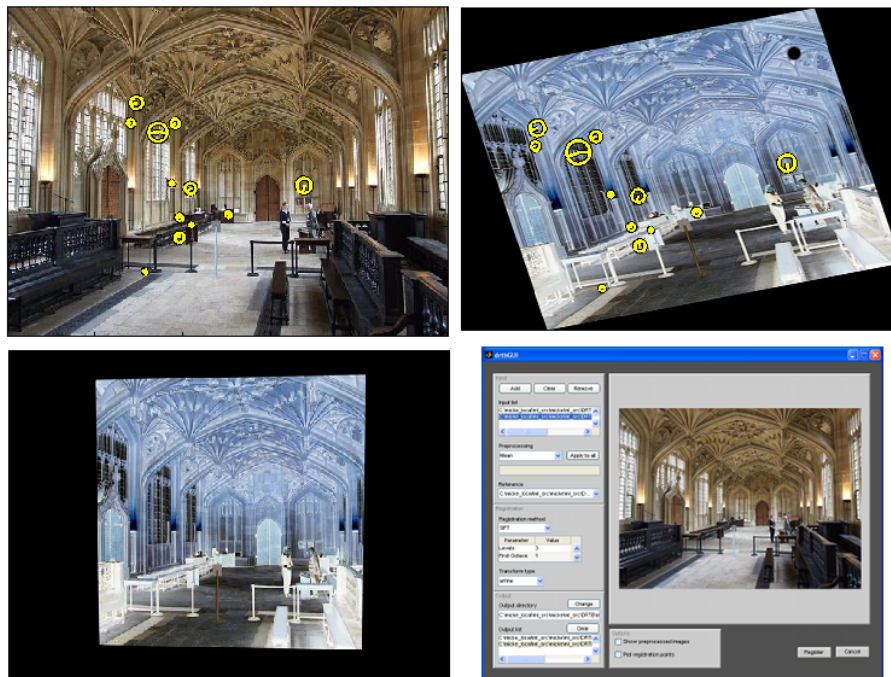
2.2.3 Egenskapsbaserad metod

Den valda metoden bygger på den automatiska ”stitching”-algoritmen beskriven i [9]. Nedan följer en kort beskrivning av det implementerade registreringsverktyget.

2.2.3.1 Nyckelpunktsberäkning

Egenskapsbaserade metoder bygger på att man beräknar stabila nyckelpunkter i data och därefter jämför nyckelpunkterna för testbilden och referensbilden för att hitta *matchande* nyckelpunkter. Därefter esitimeras en transformation som åstadkommer att positionerna för testbildens nyckelpunkter överensstämmer med positionerna för referensbildens nyckelpunkter.

I projektet har vi tagit fram en egenskapsbaserad metod som använder nyckelpunkter beräknade med SIFT (Scale Invariant Feature Transform). SIFT-punkter beräknas med en algoritm som publicerades 1999 av David Lowe [12] och är populär i ”stitching”-mjukvara [8]. Algoritmen hittar extrempunkter i en skal-rymd (en 3D-rymd skapad genom omskalningar av bilden) och för var och en av dessa beräknas en orientering. Genom att i den aktuella skalan och orienteringen beräkna en egenskapsvektor (deskriptor) bestående av de lokala gradienterna kan man generera nyckelpunkter med egenskaper som är invarianta mot skalning, orientering, belysning och translation. För att ytterligare stärka invariansen mot belysning beräknas nyckelpunkter för såväl originalbilderna som bilderna då gråskalevärdena inverteras.



Figur 2.1: Exempel på registrering. Testbilden (övre höger) har skapats genom att beskära, skala, skjuva och invertera färgerna i referensbilden (övre vänster). I dessa bilder visas de matchande nyckelpunkter som algoritmen funnit. Resultatet av registreringen (nedre vänster) visar hur testbilden sedan transformerats för att passa referensbilden. Nere till höger visas det framtagna användargränssnittet.

2.2.3.2 Nyckelpunktsmatchning och transform

Efter att nyckelpunkter har beräknats i den aktuella testbilden och referensbilden, utförs en matchning mellan nyckelpunkterna. Matchningen av nyckelpunkter görs med hjälp av sökmetoden *RANSAC* (RANdom SAMple Consensus, [11]) som hittar de nyckelpunkter som är förenliga med en på förhand vald modell av den transformation som krävs för att få överensstämmelse mellan bilderna. Metoden hittar när den är framgångsrik såväl de matchande nyckelpunkterna som den sökta transformen. Slutligen transformeras testbilden så att den hamnar i samma koordinatsystem som referensbilden. Figur 2.1 visar ett registreringsexempel med matchande nyckelpunkter utmärkta.

2.3 Resultat

Följande stycken visar några exempel på resultat genererade med bildregistreringsverktyget.

2.3.1 Olika våglängdsband (sensorfusion)

Figur 2.2 visar ett exempel från MultimIR som samlar in data i fyra våglängdsband i kortvågs- och mellanvågs-IR. Sensorn arbetar genom att sekventiellt avbilda markområdet: Först tas en bild i det första våglängdsbandet, sedan det andra, det tredje, det fjärde och därefter börjar man på det första våglängdsbandet. Detta fortsätter tills hela det önskade området mätts in. Eftersom data vanligtvis samlas in från en flygande plattform hinner sensorn flytta sig mellan varje insamlad bild. Figuren visar de fyra våglängdsbanden före och efter att banden registrerats med det första våglängdsbandet som referensbild. I exemplet är det tydligt att de olika bilderna skiljer sig åt vad det gäller intensiteten hos olika objekt, och tydliga svart-vita kanter i en bild kan vara vit-svarta eller ha försvunnit i en andra. I detta fall har ingen förbehandling av bilderna gjorts för att underlätta registrering. SIFT-metodens inbyggda robusthet mot intensitetsskillnader p.g.a. att den använder gradienter samt vårt tillägg att även beräkna nyckelpunkter för bilderna då deras gråskala inverteras har gett en god matchning mellan bilderna.

2.3.2 Olika sensorer (sensorfusion)

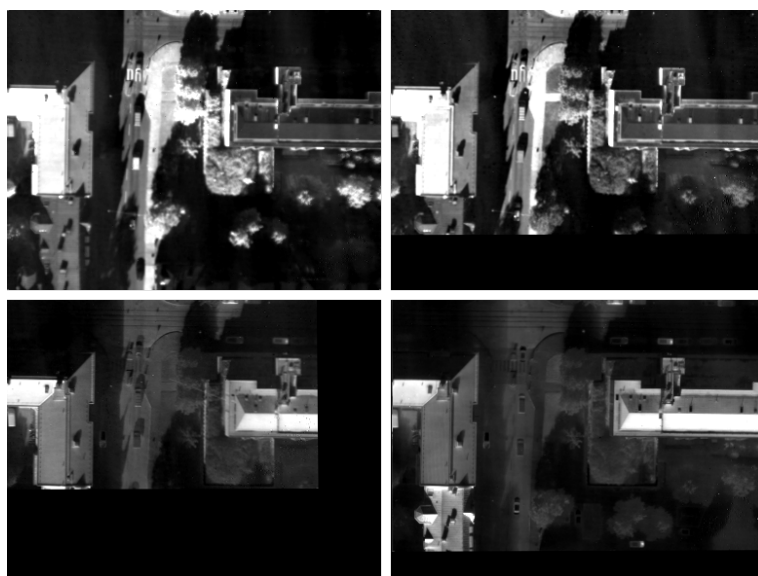
Figur 2.3 visar ett exempel med två flygfoton från samma markområde. Skillnaden mellan bilderna är att de samlats in med två olika sensorer. Den första bilden (Figur 2.3(a)) är en färgbild och den andra är en gråskalebild (Figur 2.3(b)). Färgbilden visar ett mindre område och kräver en transformation där bilden roteras, skalas och translateras för passa ihop med gråskalebilden. För att utföra registreringen förbehandlas färgbilden till gråskala genom att medelvärdesbilda de tre separata färgkanalerna. Figur 2.3(c) visar resultatet efter den automatiska registreringen.

2.3.3 Olika tidpunkter (förändringsanalys)

Registreringsexemplet i Figur 2.4 visar data från sensorn Emerald [13] insamlade vid två tidpunkter. I båda bilderna finns ett par fordon, men dessa har förflyttats mellan tidpunkterna för bilderna. Genom att registrera bilderna mot varandra är det möjligt att producera en förändringsbild (Figur 2.4(c)) som markerar förändringar mellan bilderna. Ljusa punkter visar fordon som finns i Figur 2.4(a) men inte Figur 2.4(b) och vice versa för mörka objekt.



(a) Oregistrerade bilder

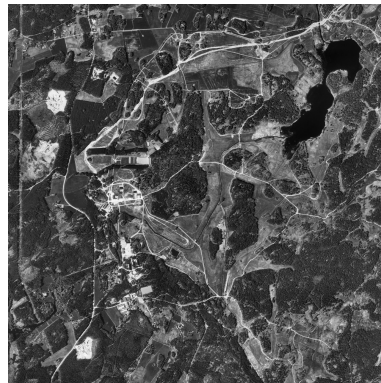


(b) Registrerade bilder

Figur 2.2: Registrering av bilder från olika våglängdsband i kortvågs- och mellanvågs-IR.



(a) Oregistrerad färgbild



(b) Referensbild i svart-vitt



(c) Registrerad bild överlagrad på referensbild

Figur 2.3: Registering av ett flygfoto i färg mot ett flygfoto i gråskala. Flygfoton: Copyright Lantmäterieverket Gävle 2001. Ref. no. L2002/308



(a) Kl. 08:36



(b) Kl. 09:23



(c) Förändringsbild

Figur 2.4: Exempel på hur man kan analysera förändringar mellan bilder insamlade vid olika tidpunkter.

2.3.4 Mosaikbildning (kartering)

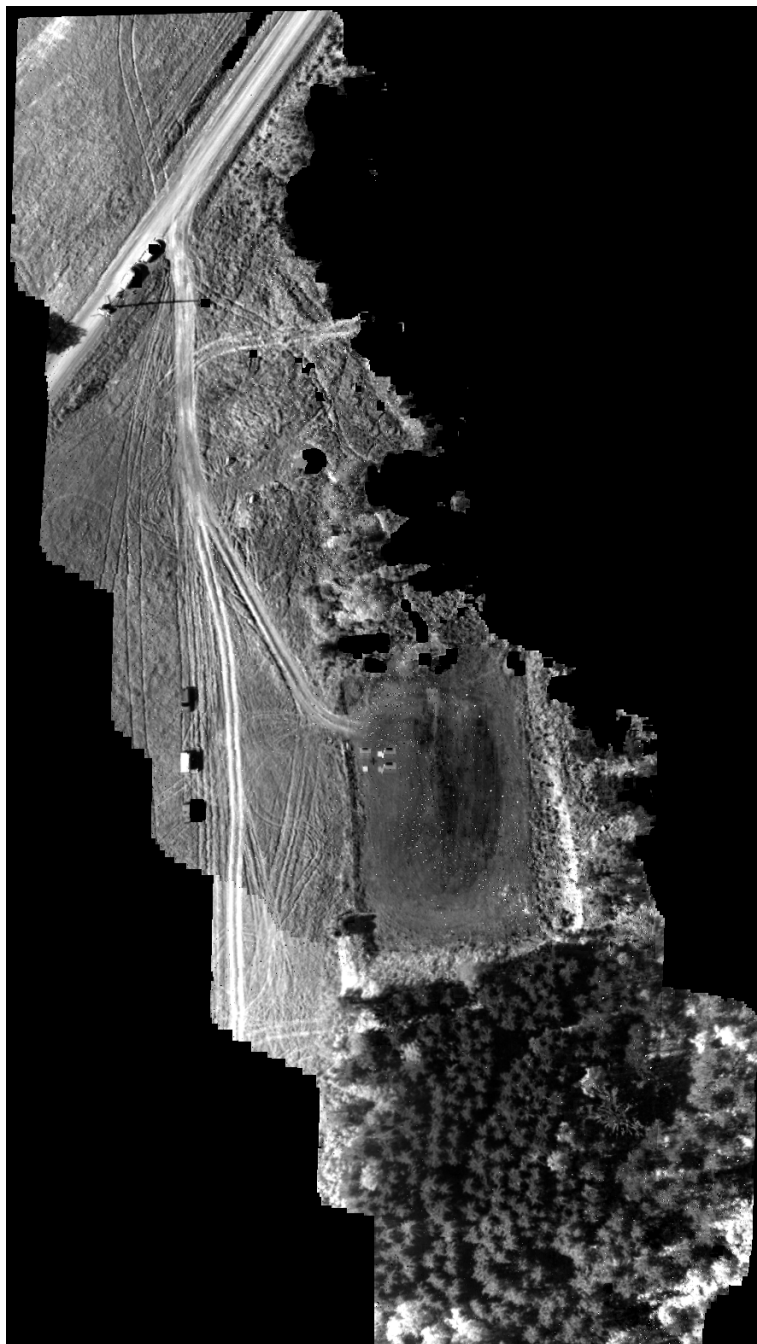
En annan viktig applikation som registreringsverktyget ska utföra är att producera sammanhängande mosaiker från många delbilder, det vill säga ”stitching”. Figur 2.5 visar ett våglängdsband i en mosaik bestående av ca 50 st delbilder från MultimIR.

2.4 Framtid

Föregående bildexempel visar att vårt registreringsverktyg har potential att lösa de problem som den civila marknadens bildregistreringsverktyg inte fokuserat på. Flera av de önskade funktionerna är implementerade och fungerar. Däremot krävs en vidareutveckling av metodiken och ramverket för att få ett mer strömlinjeformat registreringsverktyg.

Några av de uppgifter som behöver vidareutvecklas och implementeras är:

- Automatiskt val av vilka bilder som ska höra ihop i en mosaik. I dagsläget produceras en mosaik genom att utgå från en bild. Därefter registreras nästa bild, som förutsätts ha ett överlapp med den föregående bilden, mot den första och en ny sammanslagen bild skapas, och så vidare.
- För *visualisering* bör sammanslagning av bilder i en mosaik utföras på ett sätt så att skarvarna inte syns, t.ex genom ”multi-band bledning”. Det framtagna verktyget lägger enbart den nya överlappande bilden över den gamla, utan att blanda bilderna på något sätt.
- För *analys* bör sammanslagning av bilder i en mosaik utföras på ett sätt så att superupplösning uppnås genom att utnyttja att bildernas bildpunkter ligger något olika. Ju fler överlappande bilder vi har, desto högre upplösning får vi alltså.
- Mera djupgående studier av vilka förbehandlingar som krävs för att data från väldigt skilda sensortyper ska kunna registreras mot varandra. Var går gränsen för vad som är möjligt med den nuvarande metoden?
- Vidareutveckling av användargränssnittet för att på ett enkelt sätt kunna styra alla önskvärda funktioner.
- Olika metoder för att hantera bilder med flera band. Vilka/vilket band ska användas som referensbild? Hur ska multispektrala bilder förbehandlas på bästa sätt för att få en bra registrering?



Figur 2.5: Mosaik av ca 50 st MultimIR-bilder.

3 Atmosfärskorrektion i hyperspektralt termiskt infrarött

3.1 Bakgrund

Tänk dig en flygfarkost som spanar mot markmål. Flygfarkosten håller sig på några hundra meters till ett par kilometers höjd och är utrustad med en spaningssensor. Sensorn är hyperspektral för att kunna se skillnad på olika material och opererar i långvågsinfrarött (8–12 μm) för att ha dygnet-runtkapacitet.

Alla material karaktäriseras av sitt emissivitetsspektrum ϵ , ett värde mellan 0 och 1 som för varje våglängd berättar hur mycket värmestrålning materialet emitterar vid just den våglängden. Om man kan mäta emissiviteten kan man alltså låta datorn slå upp vilket material man tittar på, och således för varje pixel i bilden få veta om det är, t.ex, en ”pansarpixel”, ”granpixel” eller ”kamouflagenätpixel”.

Tyvärr kan man inte mäta emissiviteten direkt, eftersom:

- Den emitterade strålningen även är temperaturberoende (emissiviteten multipliceras med Plancks svartkroppsfunktion för den givna temperaturen).
- Materialet även reflekterar strålning (det vi vanligen ser utomhus är ju reflekterad solstrålning), beroende på materialets reflektans ρ (som är lika med $1 - \epsilon$).
- Atmosfären inte är transparent, utan har en transmission som varierar från 0 (helt ogenomskinlig) till 1 (helt genomskinlig).
- Atmosfären också strålar (emitterar).

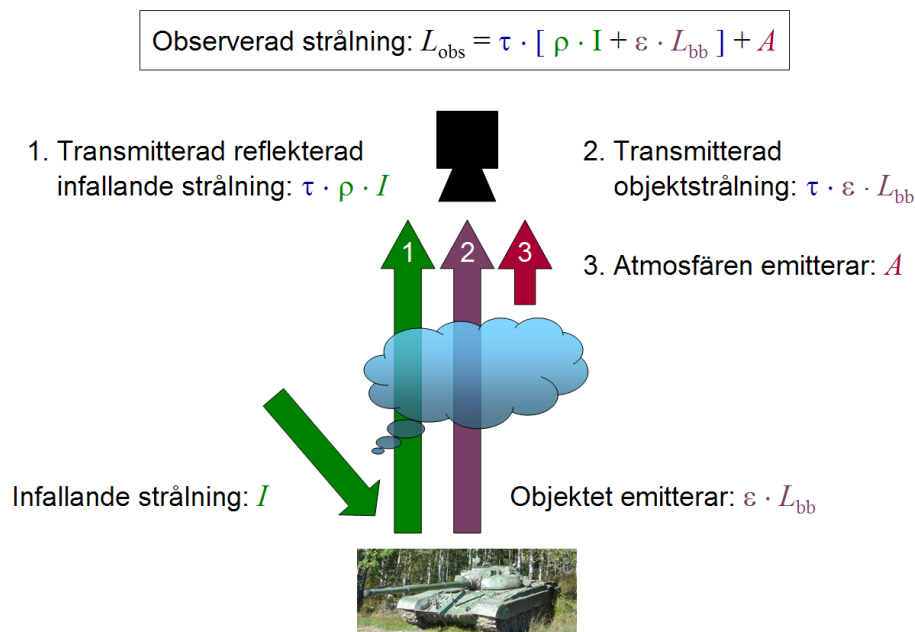
För korta avstånd är atmosfären tämligen transparent, men i vårt scenario, där målet befinner sig någon kilometer bort, gör den stor skillnad. Det vi faktiskt ser med vår sensor är alltså

$$L_{obs} = \tau \cdot (\rho \cdot I + \epsilon \cdot L_{bb}) + A, \quad (3.1)$$

där τ är atmosfärstransmissionen, L_{bb} är Plancks svartkroppsfunktion [5], I är den infallande strålningen (t.ex från himlen), A är strålningen från atmosfären in mot sensorn och ρ är materialets reflektans. Det hela illustreras i Figur 3.1.

Alltså: Om vi visste värdena på A , τ , I samt objektets temperatur skulle vi kunna räkna ut ϵ , och därmed klassificera objektet. Vi skulle alltså kunna materialbestämma varje pixel i bilden, även i mörker.

Lyckligtvis går A , τ och I att simulera med programvaror som MODTRAN [10], men då krävs att man känner parametrar som lufttemperatur och luftfuktighet. Objektets temperatur känner vi inte heller (visste vi det skulle vi



Figur 3.1: 1. Strålning från omvärlden reflekteras i objektet och transmitteras genom atmosfären. 2. Objektet emitterar värmestrålning som transmitteras genom atmosfären. 3. Atmosfären emitterar även den värmestrålning. Sensorn mäter den summerade strålningen L_{obs} .

nog inte behöva spana mot det), men det är här finessen med en hyperspektral sensor kommer in – med vissa begränsningar på hur emissivitetsfunktionen ϵ ser ut kan man estimerar såväl den som objektets temperatur. Och de okända atmosfärsparametrarna!

3.2 Framsteg och resultat

Redan 2006 utvecklade vi metoder för att estimerar atmosfärsparametrarna från bilddata [2]. Genom att anta att det vi tittar på är en *gråkropp* (emissiviteten ϵ har samma värde i alla våglängder) så kan vi samtidigt estimerar atmosfärsparametrarna, objektets emissivitet, samt objektets temperatur. Objekt som inte är gråkroppar kommer att ge stora modellfel, men så länge vi kan hitta ett antal gråkroppspixlar i bilden, t.ex vegetation, kan vi göra ett någorlunda bra estimat av atmosfärsparametrarna. Studien 2006 visade att 10–20 spektralband (med måttligt sensorbrus) behövs för att göra estimatet. Som resultat kan man därför till stor del ta bort effekterna av atmosfären, och analysera objekten man spanar mot. Vi har tidigare visat hur vi kan se skillnad mellan viss kamouflagefärg och såväl naturlig bakgrund som andra typer

av varma objekt [3].

Under 2009 återupptogs arbetet med syftet att införa förhandskunskap i beräkningen. I den tidigare beräkningsmetoden släpptes parametrar såsom atmosfärstemperatur fria. En uppdaterad metod har tagits fram där man kan ange estimat av atmosfärens temperatur och fuktighet samt sensors brusnivå – kunskap som det är högst sannolikt att man har. Estimat av objektgenskaperna bör därför kunna förbättras, men hittills saknas laboriemätningar som kan bekräfta detta.

3.3 Framtid

I den nära framtiden ska den uppdaterade beräkningsmetoden dokumenteras, valideras med mätdata, samt utsättas för internationell granskning och publicering. På längre sikt ska metoden anpassas för stora scener, det vill säga estimat ska göras över hela bilder eller utefter flygstråk, vilket i sin tur möjliggör materialklassificering och måldetektion vid spaningsuppdrag. Metoden ska vara i drift inför fältförsök inom EDA-projektet DUCAS [1], fältförsök som ska genomföras i Belgien under 2011.

4 Informationsmått för sensorprestandavärdering

4.1 Bakgrund

För att avgöra en sensors detektionsförmåga har många olika mått definierats. Ursprungligen användes manuella mått, där en operatör skulle avgöra om han kunde urskilja ett antal svart-vita fält på en avbildad tavla. Med automatiska detektionsalgoritmer blir det i stället mer relevant att avbilda ett antal kända mål och föra statistik på hur stor andel av målen som detekterades vs. antalet falska detektioner. Vanligen plottas de mot varandra i ROC-kurvor, se Figur 4.1.

ROC-kurvor medför två problem. För det första är det inte entydigt vad som är bra. Två kurvor kan t.ex. mycket väl korsa varandra (Figur 4.1(a)). För att lösa det används olika mått som detektionssannolikhet vid en given högsta tolererad falsklarmssannolikhet (Figur 4.1(b)). Ett annat mått är AUC ("Area Under Curve"), vanligen definierad som ytan under ROC-kurvan i ett visst intervall av falsklarmssannolikhet, t.ex. 0–10% (Figur 4.1(c)).

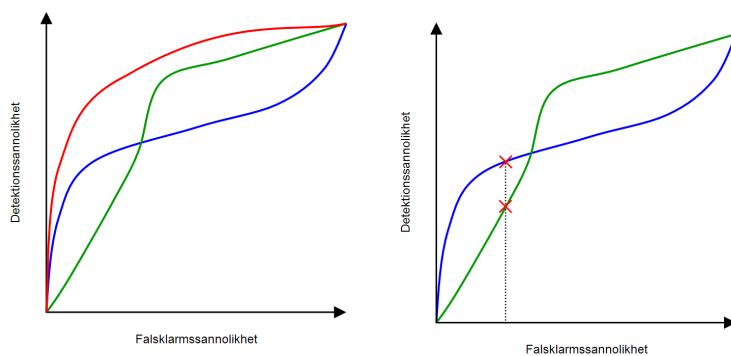
Ett annat problem är att det man mäter med ROC-kurvan är hela *systemets* prestanda, medan man ofta vill veta *sensorn*s prestanda, det vill säga vad som vore möjligt att åstadkomma givet en optimal detektionsalgoritm. Ett sådant mått skulle göra det möjligt att värdera sensorkonfigurationer, samt även att värdera systemets potential – om prestanda för systemet är långt ifrån den optimala prestandan för en given sensor bör man lägga mer energi på att utveckla detektionsalgoritmer.

Vi föreslår därför ett informationsteoretiskt mått som givet en hyperspektral bild från en given sensorkonfiguration och en målmask (ett facit, vilka pixlar/spektrala signaturer är mål respektive bakgrund) ger en kurva som talar om hur mycket information sensorn ger om huruvida en pixel är mål eller bakgrund. Måttet går dessutom att relatera till ROC-kurvan och pekar ut vilken punkt på kurvan som är den mest informationsrika, det vill säga vilken punkt som därmed bör jämföras med konkurrerande sensorkonfigurationer.

4.2 Metod

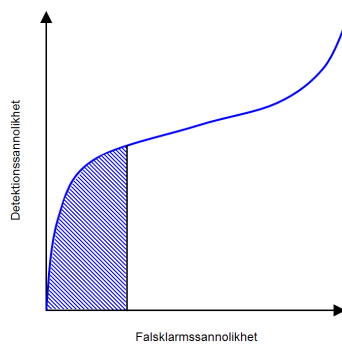
Vi betraktar en pixel från en hyperspektral sensor. Pixeln visar antingen (en liten del av) ett mål eller bakgrunden, vilket vi kan betrakta som en binär variabel (0 eller 1) som beskriver pixelns tillstånd. Beroende på vilket tillstånd den har, har pixeln en spektral signatur hämtad från en målmängd eller en bakgrundsmängd. Är mängderna helt åtskilda kan vi i teorin göra en perfekt mål/bakgrundsdiskriminering. I verkligheten är de oftast överlappande.

Den spektrala signaturen utsätts sedan för ett antal transformer. Den ska



(a) Korsande ROC-kurvor. Den röda kurvan beskriver ett bättre system (högre detektionssannolikhet för varje falsklarmssannolikhet), men är den gröna eller blå kurvan bäst?

(b) Detektionssannolikhet kan jämföras vid en given falsklarmssannolikhet.



(c) "Area Under Curve" för ett givet falsklarmssannolikhetsintervall.

Figur 4.1: ROC-kurvor ("Receiver Operating Characteristics")

exempelvis passera genom luften (se föregående kapitel) och sedan mätas av sensorn (vilket medför effekter som sensorbrus och kvantisering (digitalisering)). Mål- respektive bakgrundmängderna av digitaliserade observerade signaturer är alltså det sensorn ger, och en måldetektionsalgoritm försöker lista ut vilken mängd en varje sådan signatur tillhör. Naturligtvis försöker man även göra en del trick för att förenkla detektionsalgoritmens uppgift, exempelvis genom sådan atmosfärskompensation som beskrev i förra kapitlet, men vi betraktar här hela systemet från mål/bakgrund till digital signatur, men utan detektionsmetod.

Som vi visat tidigare [4] kan vi från testdata estimerar hur mycket *information* en sensor ger om en pixels tillstånd (mål/bakgrund). Resultatet blir ett värde mellan noll och en bit, och därmed går även två olika sensorkonfigurationer att jämföra genom att beräkna informationsmättet för båda.

Vi utökar nu resonemanget genom att sortera alla signaturer i vår testdatamängd efter hur ofta de förekommer i målmängden dividerat med hur ofta de förekommer i bakgrundmängden. Vi beräknar därefter informationsmättet för bara den första signaturen, sedan för den första och den andra, osv., tills man har använt alla tillgängliga signaturer. Man får då typiskt en kurva som i Figur 4.2(a). Kurvan har en tydlig maxpunkt, som liksom det tidigare föreslagna måttet kan användas för att karakterisera sensorprestanda.

Det intressanta är att varje punkt på "informationskurvan" motsvarar en punkt på en ROC-kurva (se Figur 4.2(b)). Det går också att för varje punkt på en ROC-kurva associera ett informationsmått. Vi har alltså ett godhetsmått på varje punkt på ROC-kurvan!

Så hur gör vi? Jo, i vår testdatamängd har vi ett antal spektrala signaturer. Vi börjar med att plocka de signaturer som bara finns i målmängden och klassificera dem som just mål. Genom att klassificera endast dessa signaturer som mål (och övriga som bakgrund) får vi antagligen en låg detektionssannolikhet, men en falsklarmssannolikhet lika med noll. Vi fortsätter med att plocka signaturer i fallande ordning efter kvoten mellan hur ofta signaturen förekommer i målmängden och hur ofta den förekommer i bakgrundsmängden. För varje steg kan vi beräkna vårt informationsmått (en punkt på informationskurvan) samt en punkt på ROC-kurvan. ROC-kurvan blir den optimala, det vill säga den bästa som teoretiskt går att uppnå på den givna testdatamängden med en spektral detektor.

4.3 Resultat

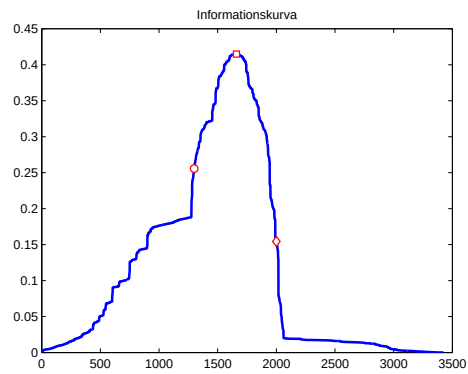
Vi exemplifierar här med testdata från en 32-bands hyperspektral sensor. Bilden (Figur 4.3) visar ett antal minor utlagda i naturlig bakgrund (mestadels gräs). Vi har även tillgång till en målmask som visar vilka pixlar som är målpixlar och vilka som är bakgrundspixlar. Från bilden sorterar vi alla spektrala signaturer (enligt ovan), och börjar med att klassificera den första signaturen som mål och alla andra som bakgrund. Och så vidare. För varje

steg plottar vi informationsmättet, och får informationskurvan i Figur 4.2(a), samt ROC-kurvan (Figur 4.2(b)). Vi ser att när vi nått en detektionssannolikhet på 0.74 (och en falsklarmssannolikhet på 0.046) har vi maximerat informationsmättet *för den aktuella testdatamängden och sensorkonfigurationen*. De motsvarande detektionsbilderna finns i Figur 4.4, och för det mänskliga ögat ser antagligen den första detektionsbilden bäst ur. Detta beror på att vi lägger in spatiella resonemang och automatiskt detekterar ett antal objekt, medan detektionssannolikheten här anges per pixel och inte per objekt.

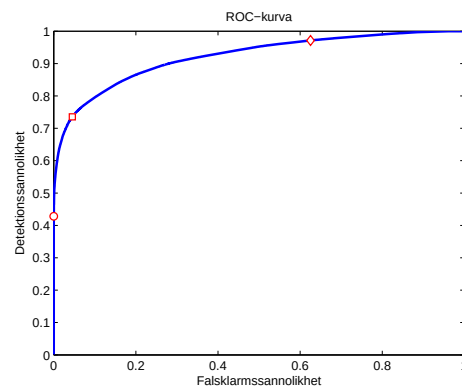
Vi kan nu jämföra med en annan sensorkonfiguration. I Figur 4.5 ser vi detektionsbild och informationskurva för samma indata, men vi har bara använt åtta av de 32 spektralbanden. Resultatet blir ett informationsinnehåll på maximalt 0.30 bit per pixel (i stället för 0.41). Exemplet är övertydligt eftersom det går att se på detektionsbilden att prestanda är sämre, men illustrerar ändå det faktum att vi har *en* siffra som anger sensors prestanda, och som applicerade detektionsmetoder kan benchmarkas mot.

4.4 Framtid

Det föreslagna måttet kommer att användas vid sensorprestandavärderingar i framtida experiment, och skall även utsättas för internationell granskning och publicering. Det föreslagna måttet kommer även att användas för att benchmarka existerande och kommande detektionsmetoder.



(a) Informationskurva. De tre markeringarna visar informationsinnehåll (I) på 0.26, 0.41 respektive 0.15 bit per pixel. x-axeln visar antalet spektrala signaturer som tagits med i beräkningen.



(b) De tre markeringarna visar detektionssannolikheterna (P_D) 0.43, 0.74 respektive 0.97 samt falsklarmssannolikheterna (P_F) 0.00018, 0.046 respektive 0.63.

Figur 4.2: Informationskurva och motsvarande ROC-kurva.



(a) Tre band ur den hyperspektrala bilden visualiserade som en färgbild.



(b) Målmasken ("facit"). Vita pixlar är minor, svarta är bakgrund.

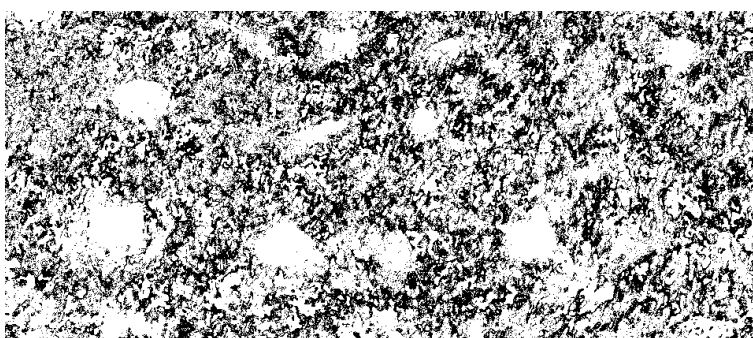
Figur 4.3: Testdatamängden.



(a) $P_D = 0.43$. $P_F = 0.00018$. $I = 0.26$.

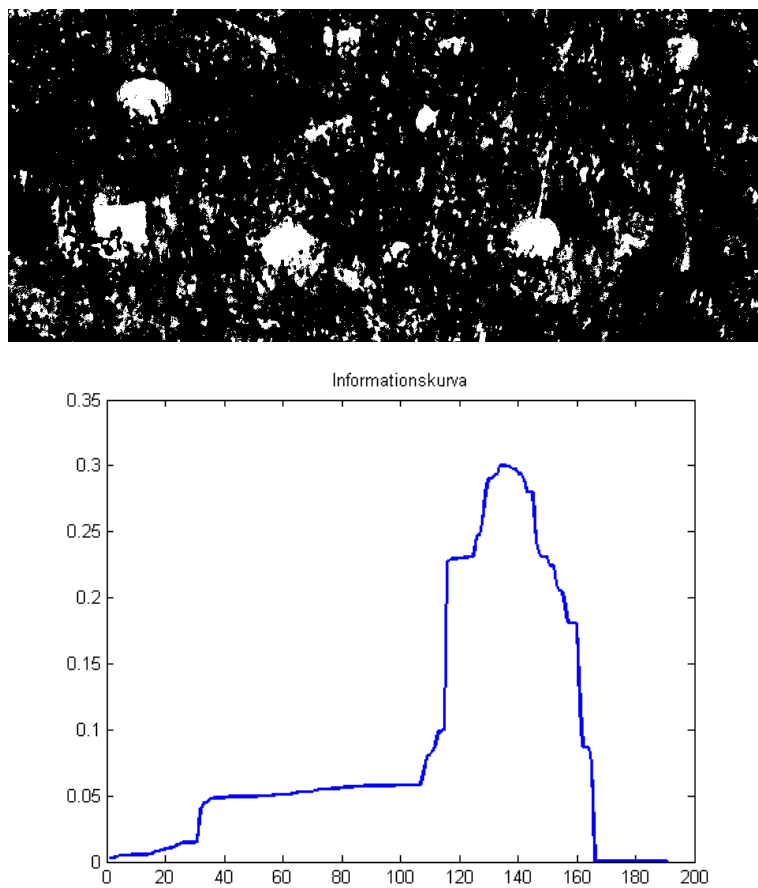


(b) $P_D = 0.74$. $P_F = 0.046$. $I = 0.41$.



(c) $P_D = 0.97$. $P_F = 0.63$. $I = 0.15$.

Figur 4.4: Detektionsbilder motsvarande markeringarna i kurvorna i Figur 4.2. Mittenbilden är alltså den som ger mest information om mål/bakgrundstillståndet.



Figur 4.5: Detektionsbild (vid maximalt informationsinnehåll) och informationskurva för en sensor med åtta band.

5 Slutsatser

Vi har i den här rapporten beskrivit pågående arbete på tre olika bildanalysmetoder för tre olika ändamål. Alla tre är nödvändiga för att designa sensorsystem för att uppnå den eftertraktade förmågan att kunna spana mot ett område och detektera avvikelser, förändringar samt specifika signaturer.

Den generella slutsatsen är att vi har kommit några steg på vägen, men att flera delförmågor återstår att studera och/eller utveckla.

5.1 Registrering

Registrering (samkalibrering) är en förutsättning för att kunna utnyttja data över samma område från flera sensorer. En metod för registrering (samkalibrering) av heterogena sensorer har implementerats och resultat visats. Metoden bygger på teknik utvecklad för konsumentmarknaden, och fungerar väl för visuella och infraröda bilder i olika våglängder. Vidare studier behövs för att möjliggöra sensordatafusion för att åstadkomma superupplösning (spatiellt eller spektralt), för att förenkla mosaikbildning, och för att på bästa sätt hantera andra typer av sensordata (t.ex hyperspektralt, linjeskannad data och 3D-laser).

5.2 Atmosfärskorrektion i hyperspektralt termiskt infrarött

En metod för att estimeras atmosfärens vatteninnehåll och temperatur i en scen har visats, vilket medför att man kan korrigera för atmosfärens inverkan och räkna ut de ”verkliga” signaturerna hos objekt på marken. Det möjliggör i sin tur att man kan detektera specifika signaturer och klassa pixlar efter material. Sådan signaturklassificering är ett kommande arbete i projektet.

5.3 Informationsmått för sensorprestandavärdering

Ett mått för att värdera en sensors prestanda givet en scen med kända mål har beskrivits. Måttet har två fördelar jämfört med den vanligast metoden (ROC-lurvor); dels ger måttet *en* siffra och inte en hel kurva, dels är det oberoende av måldetektionsmetod.

Litteraturförteckning

- [1] J. Ahlberg, *DUCAS statusrapport*, FOI Memo 2983, December 2009.
- [2] J. Ahlberg, *Estimation of atmosphere and object properties in hyperspectral longwave infrared data*, Vetenskaplig rapport FOI-R--2095--SE, Oktober 2006.
- [3] J. Ahlberg and I. Renhorn, *Analysis of Hyperspectral Longwave Infrared Imagery*, Vetenskaplig rapport FOI-RH--0493--SE, December 2005.
- [4] J. Ahlberg and I. Renhorn, "An information-theoretic approach to band selection," *Proceedings of SPIE Vol. 5811, Targets and Backgrounds XI: Characterization and Representation*, pp. 15–23, Orlando, FL, April 2005.
- [5] M. Alonso and E. J. Finn, *Physics*, Addison-Wesley, 1992.
- [6] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding (CVIU)*, Vol. 110, No. 3, pages 346–359, 2008.
- [7] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics*, Vol. 2, No. 4, pages 217–236, 1983.
- [8] M. Brown and D. Lowe, *AutoStitch :: a new dimension in automatic image stitching*, <http://people.cs.ubc.ca/~mbrown/autostitch/autostitch.html>
- [9] M. Brown and D. Lowe, "Automatic Panoramic Image Stitching using Invariant Features," *International Journal of Computer Vision*, Vol. 74, No. 1, pages 59–73, 2007.
- [10] MODTRAN®5, www.modtran.org.
- [11] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography," *Communications of the ACM*, Vol. 24, pages 381–395, 1981.
- [12] D. Lowe, "Object Recognition from Local Scale-Invariant Features," *Proceedings of the International Conference on Computer Vision (ICCV)*, pages 1150–1157, Corfu, Greece, September 1999.
- [13] T. Svensson et al., *Multi- och hyperspektral spaning 2006–2008 — slutrapport*, Användarrapport FOI-R--2642--SE, FOI, December 2008.