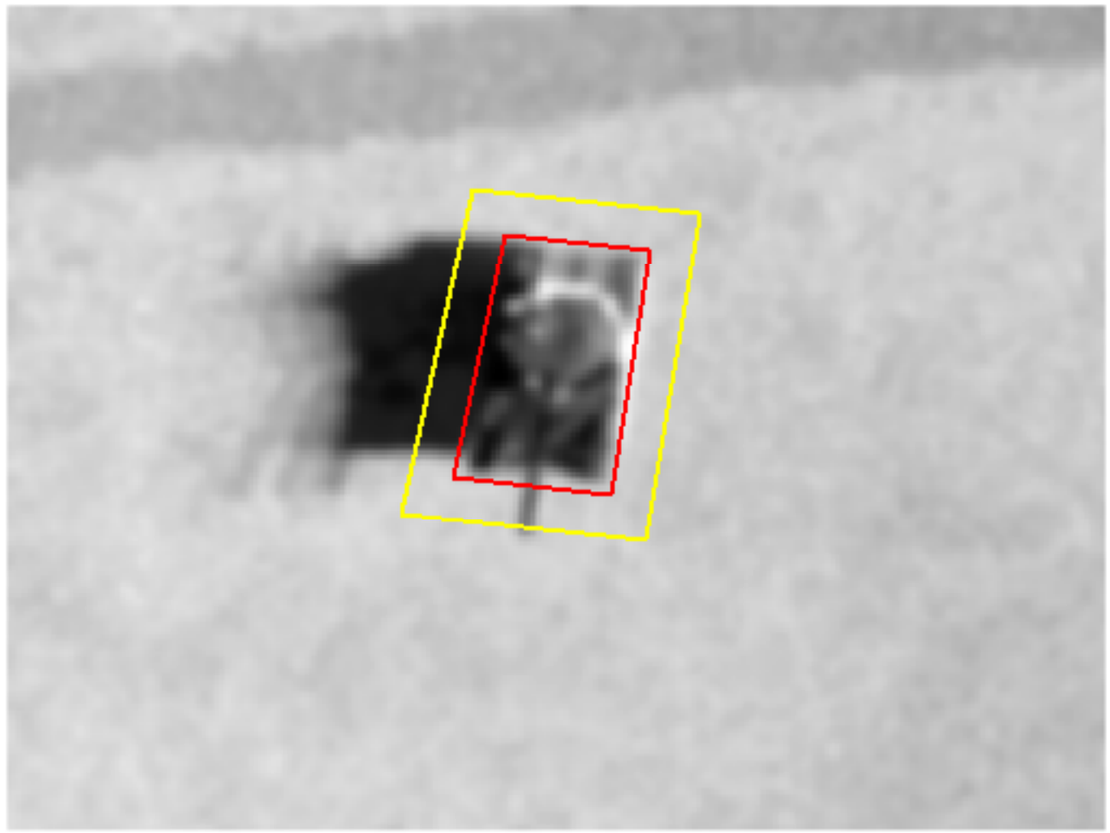
Models used in assessment of optical sensors

STEN NYBERG



FOI, Swedish Defence Research Agency, is a mainly assignment-funded agency under the Ministry of Defence. The core activities are research, method and technology development, as well as studies conducted in the interests of Swedish defence and the safety and security of society. The organisation employs approximately 1250 personnel of whom about 900 are scientists. This makes FOI Sweden's largest research institute. FOI gives its customers access to leading-edge expertise in a large number of fields such as security policy studies, defence and security related analyses, the assessment of various types of threat, systems for control and management of crises, protection against and management of hazardous substances, IT security and the potential offered by new sensors.

FOI
Defence Research Agency
Sensor Systems
P.O. Box 1165
SE-581 11 Linköping

Phone: +46 13 37 80 00
Fax: +46 13 37 81 00

www.foi.se

FOI-R--2290--SE Technical report
ISSN 1650-1942 June 2007

Sensor Systems

Sten Nyberg

Models used in assessment of optical sensors

FOI-R--2290--SE Technical report Sensor Systems
ISSN 1650-1942 June 2007

Issuing organization	Report number, ISRN	Report type
FOI – Swedish Defence Research Agency Sensor Systems P.O. Box 1165 SE-581 11 Linköping	FOI-R--2290--SE	Technical report
	Research area code 4. Sensors and low observables	
	Month year June 2007	**Project no.** E 3071
	Sub area code 45 Low observables	
	Sub area code 2	

Author/s (editor/s)	Project manager
Sten Nyberg	Patrik Hermansson
	Approved by Jörgen Ahlberg
	Sponsoring agency Armed forces
	Scientifically and technically responsible Patrik Hermansson

Report title

Models used in assessment of optical sensors

Abstract

This report contains a description of three models that can be used in the assessment of an optical sensor. The focus is on methods assessing sensor performance, not the technical details of the sensor. The complete chain from target to observer is taken into account. From a lot of parameters an output in terms of recognition probability is obtained. Here recognition is used as a common notion for detection, classification and identification. The first model is a simple model called the Bailey model which assumes an observer is looking at the target. He might be using a binocular or an electro-optical sensor to aid him. The second model is one of latest versions of the well known Night Vision Lab model named NVTherm. It has a very detailed description of the sensor. The third model differs from the preceding ones since it starts with an image of the scene taken with a relevant choice of target, background, atmosphere and sensor.

A small experiment is made where synthetic images generated with the Cameosim computer code are compared to real images captured with an IR sensor. The result may give some guidance concerning important aspects in generating synthetic imagery.

Keywords

Optical sensor, synthetic imagery, performance model, assessment

Further bibliographic information	Language English
ISSN 1650-1942	**Pages** 27 p.
	Price acc. to pricelist

Utgivare FOI - Totalförsvarets forskningsinstitut Sensorsystem Box 1165 581 11 Linköping	**Rapportnummer, ISRN** FOI-R--2290--SE	**Klassificering** Teknisk rapport
	Forskningsområde 4. Sensorer och signaturanpassning	
	Månad, år Juni 2007	**Projektnummer** E 3071
	Delområde 45. Signaturanpassning	
	Delområde 2	
Författare/redaktör Sten Nyberg	**Projektledare** Patrik Hermansson	
	Godkänd av Jörgen Ahlberg	
	Uppdragsgivare/kundbeteckning Försvarsmakten	
	Tekniskt och/eller vetenskapligt ansvarig Patrik Hermansson	
Rapportens titel Modeller för värdering av optiska sensorer		

Sammanfattning

Denna rapport beskriver tre modeller som kan användas i värderingen av en optisk sensor. Fokus ligger på prestanda och inte tekniska detaljer i sensorn. Hela kedjan från mål till observatör beaktas. Från en mängd parametrar fås sannolikheten för detektion och ibland också för klassificering och identifiering. Den första modellen, Baileys modell, utgår från att en observator tittar efter ett mål, eventuellt med kikare eller en elektrooptisk sensor. Den andra modellen är NVTherm från Night Vision Laboratory (NVESD). Den har en mycket detaljerad beskrivning av sensorn. Den tredje modellen skiljer sig från övriga genom att starta med en bild tagen med relevant val av mål. bakgrund, atmosfär och sensor.

Ett litet experiment har gjorts där syntetiska bilder från ett Cameosim program har jämförts med verkliga registreringar från en ir-sensor. Resultatet kan vara vägledande för vilka egenskaper som är viktiga vid skapandet av syntetiska bilder.

Nyckelord

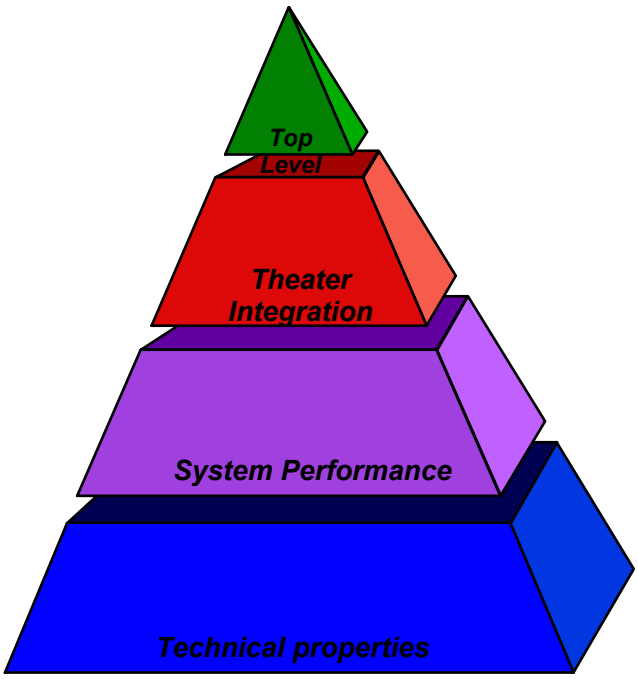
Optisk sensor, syntetiska bilder, prestandamodell, värdering

Övriga bibliografiska uppgifter	**Språk** Engelska
ISSN 1650-1942	**Antal sidor:** 27 s.
Distribution enligt missiv	**Pris: Enligt prislista**

CONTENT

1 INTRODUCTION

When assessing a system to be used for some purpose, it is important to define and understand in which context the system is used. Figure 1 illustrates what might be called the assessment pyramid, describing the hierarchical different levels in the assessment pyramid. A commander might want to know if he will win the next battle. In this case we are on the top level of the pyramid. A common way to find answers in this level is to do realistic field trials. The next level includes factors such as the availability of vehicles and soldiers. Also soft factors are present, like the willingness for soldiers to actually fight. They certainly have an opinion of the strength of the enemy. System performance is the next level in the pyramid. In this report we consider situations where the commander needs some decision support. We need to know if we are visible to the enemy in the current situation, which to a high degree involves the weather for instance. The models in this report try to find a measure of performance, which in some sense should summarize all the parameters in the lowest level. The focus will be on the system performance level.



Factors

Mission accomplishment

Blue/red casualties
Causality rate
Blue/red survivors
Rate of advance
(MOE: measure of effectiveness)

Mean detection range
Time to detection
Probability of detection
Location error to range ratio
(MOP: measure of performance)

Camouflage technical characteristics
Size, shape ground pattern, color, texture, gloss, temperature, spectral reflectance, radar cross section, emissivity
(Data)

Figure 1. The assessment pyramid with its four levels. To the right some examples are given for each level.

Below a few models are described and they are also available as computer programs.
Bailey's model is the oldest but still very relevant to the situation of visual observations with or without electro-optical sensors. The next model, NVTherm, is an US model with a detailed description of the optical sensor which has als been extensively validated. However background and environment are often highly variable, which make it difficult to obtain high precision. The last program, Terrtex, is not a recognition model, but is used to find relevant target and background properties from actual sensor images. Using some assumptions it is possible to obtain an approximate estimation of detection probability.

2 BAILEY'S MODEL

2.1 GENERAL REMARKS

This model [1, 2] describes analytically the capabilities and limitations of a human observer in the task of looking for and finding known or expected fixed objects. The description enables the user to estimate recognition probabilities as a function of the many parameters required to describe a specific situation. The model is tailored to the case of an airborne observer looking at terrain with or without optical aids or electro-optical sensors, but with prior knowledge of the approximate appearance of an object.

The model is structured according to three distinguishable psychophysical processes:

- deliberate search over a fairly well-defined area
- detection of contrasts (a subconscious retino-neural process)
- recognition of shapes outlined by the contrast contours (a conscious decision based on comparison with memory)

In addition, when the observer is viewing a displayed image of a scene, noise is sometimes present which degrades his performance of these three steps. The probability that the three steps a completely successfully, multiplied by a noise degradation factor, gives the probability of target recognition

2.2 MODEL DESCRIPTION

Imaging sensor performance may be estimated and/or evaluated by application of a target detection/recognition model such as that suggested by the Rand Corporation [1,2]; namely;

$$P_r = P_1 * P_2 * P_3 * \eta \qquad (1)$$

where P_r is the probability that a target will be recognized on the display, P_1 is the probability that the observer, searching an area that is known to contain a target, looks with his foveal vision for a specified glimpse time (1/3 s) in the direction of the target, P_2 is the probability that if the displayed target image is viewed foveally for one glimpse period it will, in the absence of noise, have sufficient contrast and size to be detected, P_3 is the probability that if a target is detected, there will be enough detail shown for it to be recognized (again during a single glimpse and in the absence of noise) and η is an overall factor arising from noise.

The probability P_1 is difficult to estimate because it is affected by the solid angle presented to the eye of the search field, by the time available to search it, by the number of confusing elements within the scene, and by the availability of any a priori information as to where to look on the display. The model employs the relation

$$P_1 = 1 - \exp\left[-\left(\frac{700}{G}\right)\cdot\left(\frac{a_t}{A_s}\right)\cdot t\right] \qquad (2)$$

where

a_t = area of target,
A_s = area to be searched,
t = glimpse time (0.3 s) and
G = congestion factor, usually between 1 and 10, for most real imagery of interest.

The probability of detection P_2 at the threshold contrast C_t is by definition 50%. A useful approximation for P_2 at other contrasts C available at the eye is given by

$$P_2 = \frac{1}{2} \pm \frac{1}{2} \cdot \sqrt{1 - \exp\left[-4.2\left(C/C_t - 1\right)^2\right]} \qquad (3)$$

where the minus sign is used when $C<C_t$.

Defining the number of resolution cells as

$$N_r = \frac{L_{\min}}{\alpha R} \qquad (4)$$

where

L_{min} = minimum projected target dimension,
α = angular resolution of the sensor and
R is the target range.

The factor P_3 is given by

$$P_3 = \begin{cases} 1 - \exp\left[-\left(N_r/2 - 1\right)^2\right] & N_2 \geq 2 \\ 0 & N_2 < 2 \end{cases} \qquad (5)$$

The noise factor is

$$\eta = \begin{cases} 1 - \exp\left[-\left(SNR - 1\right)\right] & SNR \geq 1 \\ 0 & SNR < 1 \end{cases} \qquad (6)$$

where SNR is the displayed signal-to-noise ratio.

The accuracy obtained with this model will be no better than 20 to 30 percent.

3 NVTHERM

The model described here has been developed using ideas from John Johnson at former Night Vision Laboratory, now Army Night Vision and Electronic Sensors Directorate (NVESD). The first version of the model was written by Ratches et al [3]. The model has later been continuously developed and is a standard model for the US defense. The history of the model is described in [4, 5]. A newer version, named NVThermIP, has been developed but is not available without permission.

NVTherm is a model estimating the detection probability when using an electro-optical sensor. It was mainly developed for use in the infrared area but is also used at other wavelengths.

3.1 MINIMUM RESOLVABLE TEMPERATURE DIFFERENCE

The common way to characterize an optical sensor is the minimum resolvable temperature difference as a function of spatial frequency, the MRT-function. The main part of it is the properties of the optics, electronics and detector. For the system performance the target, background, atmosphere and the operator have a large influence. It is common to use a two-dimensional MRT.

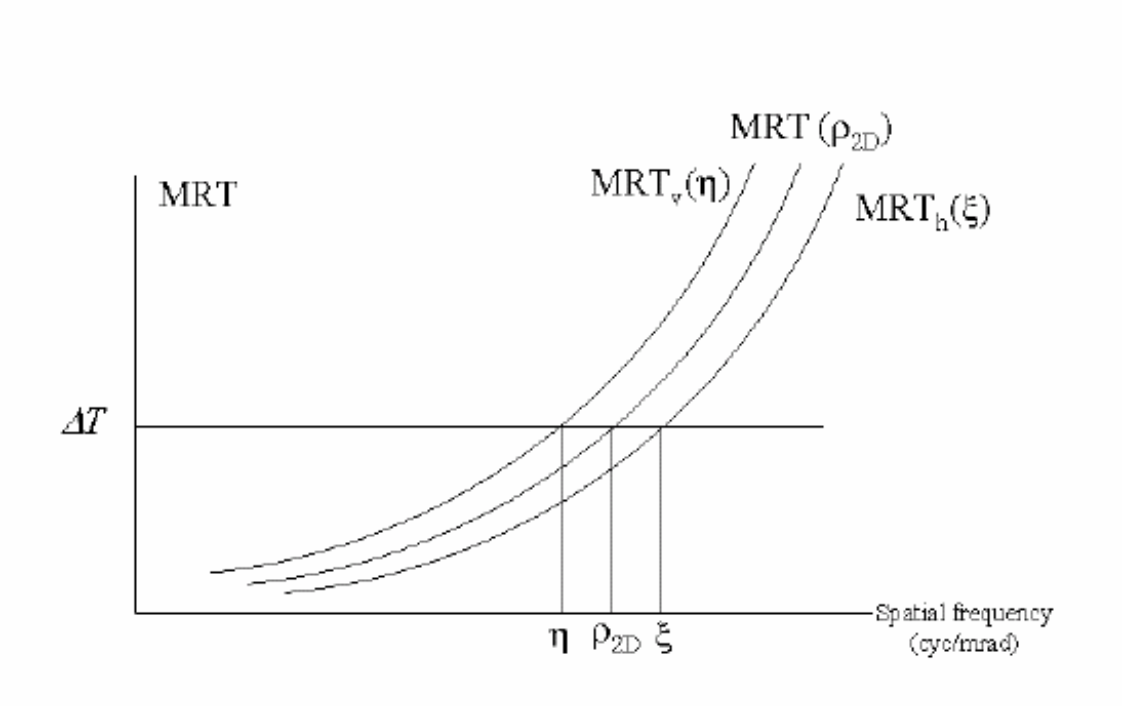


Figure 2. The two-dimensional MRT-function, MRT(ρ_{2D}).

A two-dimensional MRT is determined with the vertical and horizontal MRT-functions as shown in Figure 2.

The spatial frequencies of the horizontal and vertical spatial frequencies give the two-dimensional MRT spatial frequency through the geometrical mean

$$\rho_{2D} = \sqrt{(\xi \cdot \eta)}.$$

The matching MRT is then plotted as a function of the two-dimensional spatial frequency. This new function is the two-dimensional MRT. Note that the conversion is a spatial frequency conversion and no manipulation is performed on the two-differential temperatures.

3.2 PROBABILITY AS A FUNCTION OF RANGE

The procedure for producing a probability of detection, recognition or identification curve is quite simple. Consider the procedure flow as given in Figure 3. Four parameters are needed to generate a static probability of discrimination curve as a function of range: the target contrast, the characteristic dimension, an atmospheric transmission estimate within the band of interest for a number of ranges around the ranges of interest, and the sensor two-dimensional MRT.

The atmospheric transmission is determined (by using an appropriate model) and an equivalent blackbody apparent temperature is calculated based on the atmospheric signal reduction. Once an apparent differential temperature is obtained, the highest corresponding spatial frequency that can be resolved by the sensor is determined. This is accomplished by finding the spatial frequency (on the MRT curve) that matches the target apparent differential temperature. The target load line is the target contrast modified by the atmospheric transmission. The number of cycles across the critical target dimension that can actually be resolved by the sensor at a particular range then determines the probability of discriminating (detecting. recognizing or identifying) the target at that range. The number of cycles are given by

$$N = \rho \cdot \frac{d_c}{R} \qquad (7)$$

where ρ is the maximum resolvable spatial frequency in cycles per milliradian, d_c is the characteristic target dimension in meters, and R is the range from the sensor to the target in kilometers.

The probability of discrimination is determined by using the Target Transfer Probability Function (TTPF) given by

$$P_{rob} = \frac{\left(N/N50\right)^E}{1+\left(N/N50\right)^E} \qquad (8)$$

where $N50$ is the value for 50% probability and E = 3.76. This function is a fit to experimental data and is shown in the lower right graph in Figure 3. It is discussed in [5].

The computation needed to compute probability given the MRT-function is done by the routine Acquire, which is part of NVTherm.

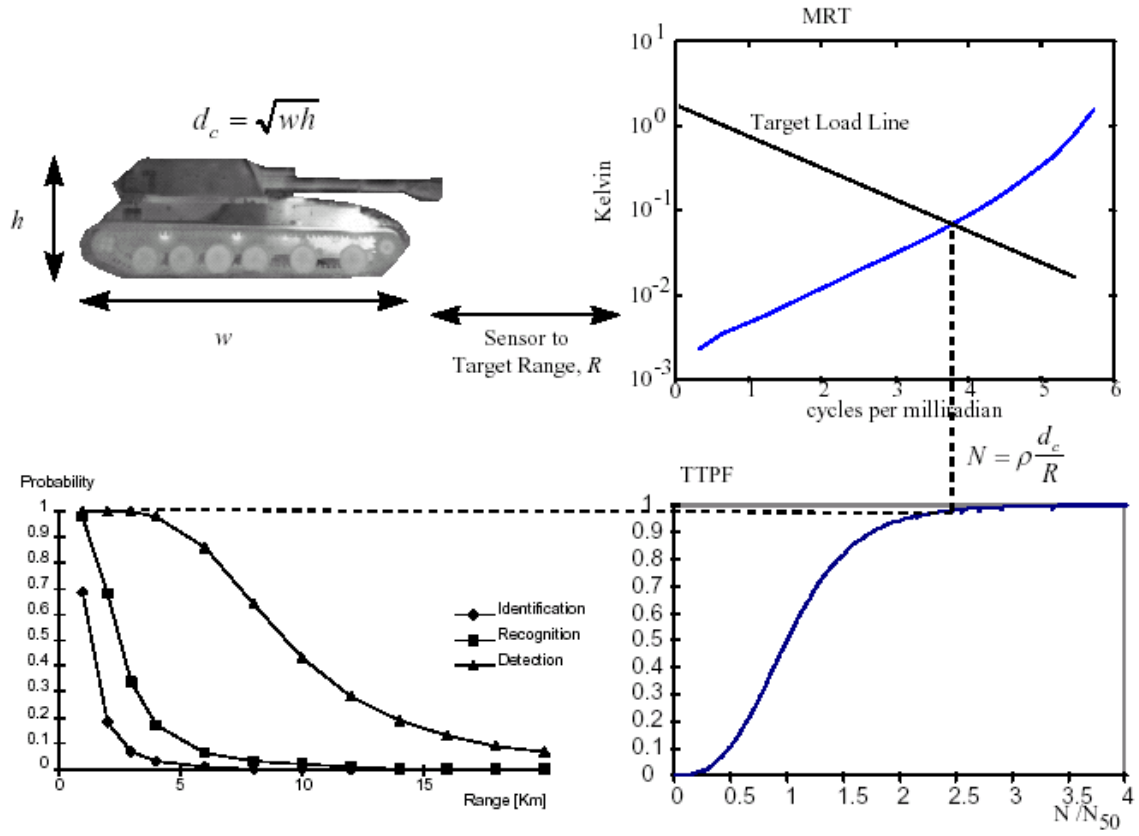


Figure 3. Illustration of the range calculation.

3.3 RANGE PREDICTIONS

The probability of discrimination is determined using the Target Transfer Probability Function (TTPF) given in Section 3.2. The level of discrimination (detection, recognition or identification) is selected and the corresponding fifty percent cycle criterion, N50, is taken. The probability of detection, recognition or identification is then determined with the TTPF for the number of cycles given by Equation 7. The probability of discrimination task is then assigned to the particular range. A typical probability of discrimination curve will have the probability plotted as a function of range. Therefore, the above procedure would be repeated for a number of different ranges.

While the following may be obvious it may nevertheless be good to mention some characteristics that improve probability of detection, recognition, and identification in infrared systems. Improvements are seen with larger targets, larger target-to-background contrast, larger target emissivities, larger atmospheric transmission, smaller MRT values (as a function of spatial frequency), and usually smaller field-of-views (if the target does not have an extremely small differential temperature).

4 TERRTEX

Terrtex is a program that was developed by FOI many years ago with the purpose to obtain a general tool for camouflage assessment and modified several times [6-11]. Terrtex was developed to assess the effectiveness of camouflage on vehicles in mainly forest backgrounds. Recent additions to Terrtex include the calculation of detection probabilities and models for visual search.

The approach proposed in Terrtex is to apply texture descriptors to quantify the similarity between different parts of an image. In addition, other descriptors are used to distinguish man-made object characteristics.

4.1 INTRODUCTION

Developments in the area of signature suppression make it progressively more difficult to recognize targets. In order to obtain a sufficiently low degree of false alarms it is necessary to take into account all available information, such as spatial, spectral, polarimetric and temporal properties. First thing to be further explored is the spatial content since the information is already in the sensor data. Accordingly, there is a genuine need to use spatial properties when analyzing the difference between a target area and a background area. This is more relevant today since modern signature suppression techniques have focused on the reduction of distinct features, such as hot spots in the infrared band. The approach is to apply texture descriptors to characterize the background and more or less camouflaged targets. In addition, other descriptors are used to characterize man made objects. Appropriately selected features should make it easier to locate areas containing vehicles or other man made objects. If time is not critical, an approach using geometrical models is preferable. Given limited time and resolution, one has to rely on measuring selected features. The assumption is that an area with observable targets has different statistical properties than other areas. Statistical properties together with detected specific target features such as straight lines, edges, corners or perhaps reflections from a window have to be combined with methods used in data fusion.

Using texture information together with other kinds of information such as spectral and temporal features makes it possible to address tasks like the assessment of efficiency of signature reduction methods and obscuring countermeasures, as well as supporting the analysis of optical sensors performance used in missions as reconnaissance systems, weapon sights and target seekers.

In the literature several attempts are described to assess the performance of signature suppression techniques [12-16]. However, there is still a need to improve the methods. Often assumptions are made which are difficult to verify. A complication to be addressed is that low contrast objects are often not fully distinguished from the natural background.

Several ways to analyze images make it possible to assess different methods for signature reduction. One method is to visualize the properties of an image region. This can be done in several ways. Examples:

- Display the Wiener spectrum (another name for a two-dimensional power spectrum) for a region of interest. Specific features may show up in such an image.
- Display some relevant image transformations, such as edges or line images.
- Displaying the Wiener spectrum for a small region around every pixel in the image. In this case it is easier to examine local events in the image.
- Compute parameters that describe different features of the Wiener spectrum, like shape and distribution as examples of descriptors.
- Use one or several feature measures to define some kind of similarity measure or the opposite distance measures.
- Compute some measures that combine (non-camouflaged or camouflaged) target and background information.

Visualization of feature images is important because it is sometimes impossible to condense all the information into to a single number. Just as in image quality, color or texture analysis, several dimensions are needed to characterize a situation accurately. However, to validate these measures, there is a big demand for simple figures like detection time or signal-to-noise ratio.

An often-used method for visualizing the similarity of a given set of features is by trying to isolate target from the background. In this case, the image is segmented into target areas and background areas. This is a task, as mentioned earlier, that is not always possible to accomplish.

The ultimate validation is of course to test a method in a real target detection experiment. Using images of the scenes, the process can be simulated with a computer. Having a large enough set of images it is possible to assess probability for detection and false alarm rates etc.

A major problem when assessing camouflage is the lack of a good theory for handling target detection in a cluttered environment. Theoretical work is often limited to the use of normal distributions for the background description. In a low observable situation, this is completely unsatisfactory.

4.2 FEATURES

In order to characterize targets in a background we need to find the differences that are important when it comes to detection, classification and identification. Here we will mostly deal with detection of vehicles. In this case detection means to find an object, some military vehicle, in the terrain.

4.2.1 FEATURE DESCRIPTIONS

A great number of texture measures are described in the literature. It is important to find features that are useful when trying to quantify the difference between targets and background. A good set of features could be designed by means of Gabor functions [17]. These are limited in space and frequency domains. However, computing many Gabor filter responses is computationally quite expensive.

Tamura [18] has studied the relationship between textural features and visual perception. The six features he used were coarseness, contrast, directionality, line likeness, regularity and roughness. He found good correspondence in a ranking test with an implementation of 16 typical digitally computed texture measures. Woodruff [19] has estimated that three features should be enough to characterize normal textures. Texture measures based on the Fourier transform are presented in [20].

In order to characterize the vehicle, several features are computed from the image. A set of Gabor coefficients are used to sample the fourier transform of the input image and they are computed at each pixel. From the local responses invariant features are calculated. The features available are:

- Mean value
- Local deviation
- Edge concentration
- Blob strength
- Low frequency energy
- Medium frequency energy
- High frequency energy

- Fractal dimension
- The fractal error
- Orientation similarity
- Mean curvature.

The features are computed directly in the image domain, many come from Gabor filter banks and the remaining from combinations of first and second order derivatives. The Gabor filters all originate from a basic filter, illustrated in Figure 4.

To find the importance of each feature, linear discrimination analysis is used [21]. We know the position of the object (target and background) features in the feature space. Then it is easy to see how each feature axis is oriented with respect to the discriminant line (linear case). The angle between the axes and this line is an indication of the importance of a feature.

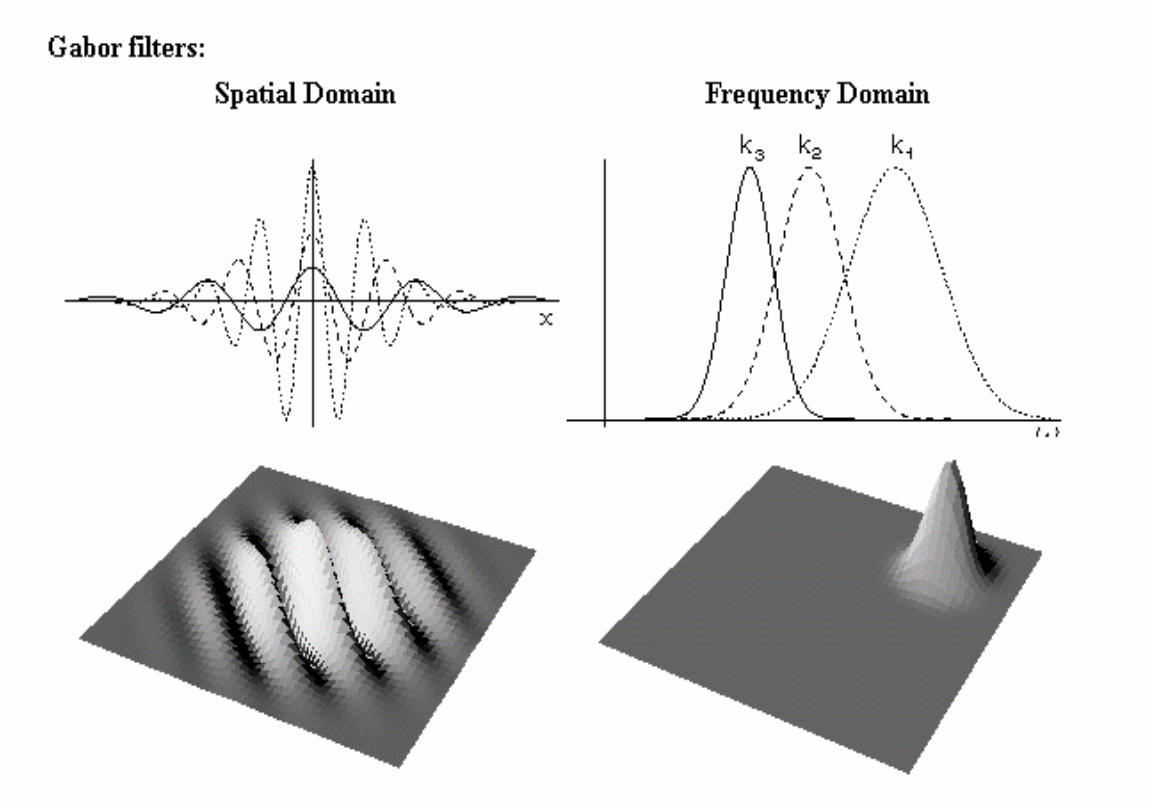


Figure 4. Gabor filters in the spatial domain and in the frequency domain.

4.2.2 FEATURE EXAMPLES

The different background features relate to properties of the Gabor filter. A few examples of feature images are given in Figure 5.

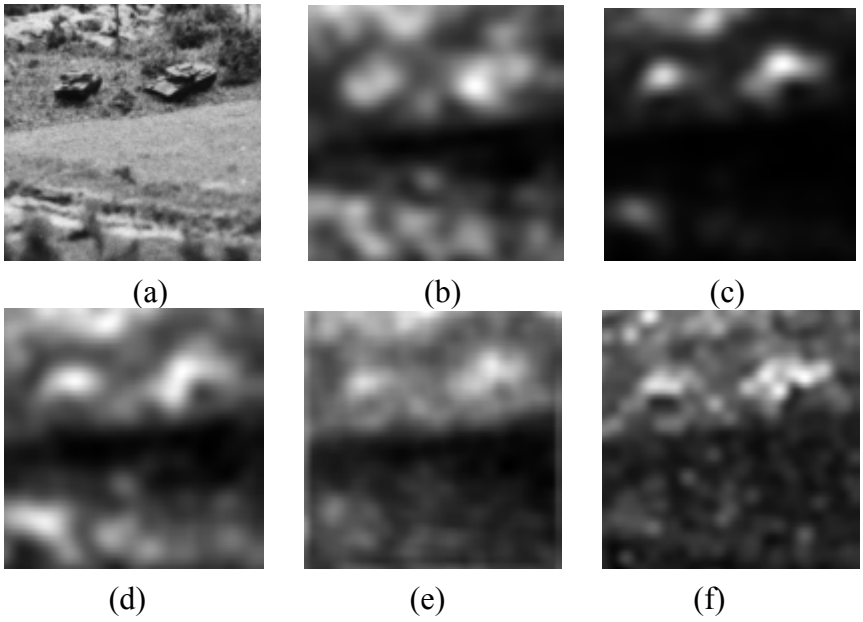


Figure 5. An image (a), local deviation (b), edge concentration (c), laplace blob (d), gabor energy, (e), mean curvature (f).

4.3 MEASURE OF PERFORMANCE

The measure of performance most easily obtained is target detection probability as function of range. Time to detection demands some modeling of the search process and will be dealt with later.

4.3.1 DETECTION PROBABILITY

The distance measure applied to the two object distributions is the Bhattacharrya measure defined as [16, 21]

$$D = -\ln\left[\int \left[p_T(x)p_B(x)\right]^{\frac{1}{2}} dx\right] \qquad (9)$$

where T and B represent the target and the background respectively.

The connection of this measure with detection theory in this area of application is given in [16]. A more general treatment is given in [21]. Figure 6 shows the distributions for target and background.

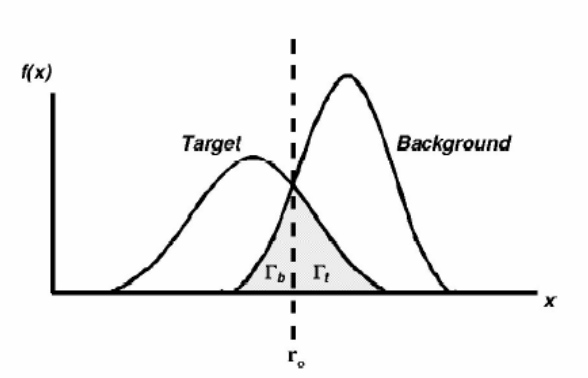


Figure 6. The target and background distributions.

The probability of detection is by definition given by

$$P_{Det} = 1 - P_B \Gamma_T \qquad (10)$$

where P_B is the a priori probability of a random sample being background, Γ_T represents the error of missing a target with r0 as the detection threshold.

Assuming k independent samples over the target we obtain

$$P_{Det} = 1 - P_B \Gamma_T^k \qquad (11)$$

Using $P_B \approx 1$ and $\Gamma_T \approx \Gamma$ gives

$$P_{Det} = 1 - \exp\left(k \ln\{\Gamma\}\right) \qquad (12)$$

which can be reduced to

$$P_{Det} = 1 - \exp(-kD) \qquad (13)$$

Using

$$k(R) = \frac{A}{R^2 \theta^2} \qquad (14)$$

where A is the target area, R = is the range and θ = is the resolution.
We obtain the probability of detection as a function of range

$$P_{Det}(R) = 1 - \exp\left(-\frac{DA}{R^2 \theta^2}\right) \qquad (15)$$

If the distributions are identical, we obtain the integral over a probability distribution function, which by definition is 1. After applying the logarithm, the distance will be 0. If the distributions are totally disjoint, we obtain the logarithm of 0 which symbolically can be defined as minus infinity. Even if the distribution is parameter free, it is difficult to estimate a continuous distribution from a sampled image.

For simplicity and usefulness let us assume that we have a Gaussian multivariate distribution. In this case the Bhattacharrya distance measure is given by

$$D = \frac{1}{4}(\mu_T - \mu_B)^T [\Sigma_T + \Sigma_B]^{-1}(\mu_T - \mu_B) + \frac{1}{2}\ln\left[\frac{\left|\frac{1}{2}(\Sigma_T + \Sigma_B)\right|}{\sqrt{|\Sigma_T||\Sigma_B|}}\right] \qquad (16)$$

where μ_i and Σ_i are mean values and covariance matrices for the distributions P_T and P_B.

Since the covariance matrix is used, many features can be used. A nice feature is the inherent normalization which, for example, allows measurements in different wavelength bands to be combined in a simple way. If the features are independent the distance can be further simplified. We obtain

$$D = \frac{1}{4}\frac{(\mu_T - \mu_B)^2}{\left(\sigma_T^{\,2} + \sigma_B^{\,2}\right)} + \frac{1}{2}\ln\left[\frac{\frac{1}{2}\left(\sigma_T^{\,2} + \sigma_B^{\,2}\right)}{\sqrt{\sigma_T^{\,2}\sigma_B^{\,2}}}\right] \qquad (17)$$

where only mean values and standard deviations are used. For each feature we get a distance. The total distance is simply the sum of the individual Bhattacharrya distances for each feature. This makes it much easier to estimate the distributions.

Since D is a generalized signal-to-noise ratio (SNR), the name GSNR is proposed which simply is defined as *GSNR = 4*D*. The constant 4 makes GSNR more or less equal to the common SNR definition.

4.3.2 DETECTION TIME

Estimating detection time is not straight forward. There are two cases to consider and they are considered in Bailey's model when trying to model the detection probability. As given above (eq. 1 and 2) the probability of detection is given by

$$P_r = P_1 * P_2 * P_3 * \eta$$

where P_1 is a factor related to visual search and is given by

$$P_1 = 1 - \exp\left[-\left(\frac{700}{G}\right)\cdot\left(\frac{a_t}{A_s}\right)\cdot t\right]$$

where,

a_t = area of target,
A_s = area to be searched,
t = glimpse time (0.3 s) and
G = congestion factor, usually between 1 and 10 for most real imagery of interest.

If the target is in the foveal area of the eye, then the time to detect a target is given by the fixation time. This is about 0.3 seconds in general. The time to detection in normal cases is given as the time before we have the target in the foveal field of view. This depends on the size of the image and also on the amount of false alarms present. In Bailey's model the false alarms are contained in the congestion factor.

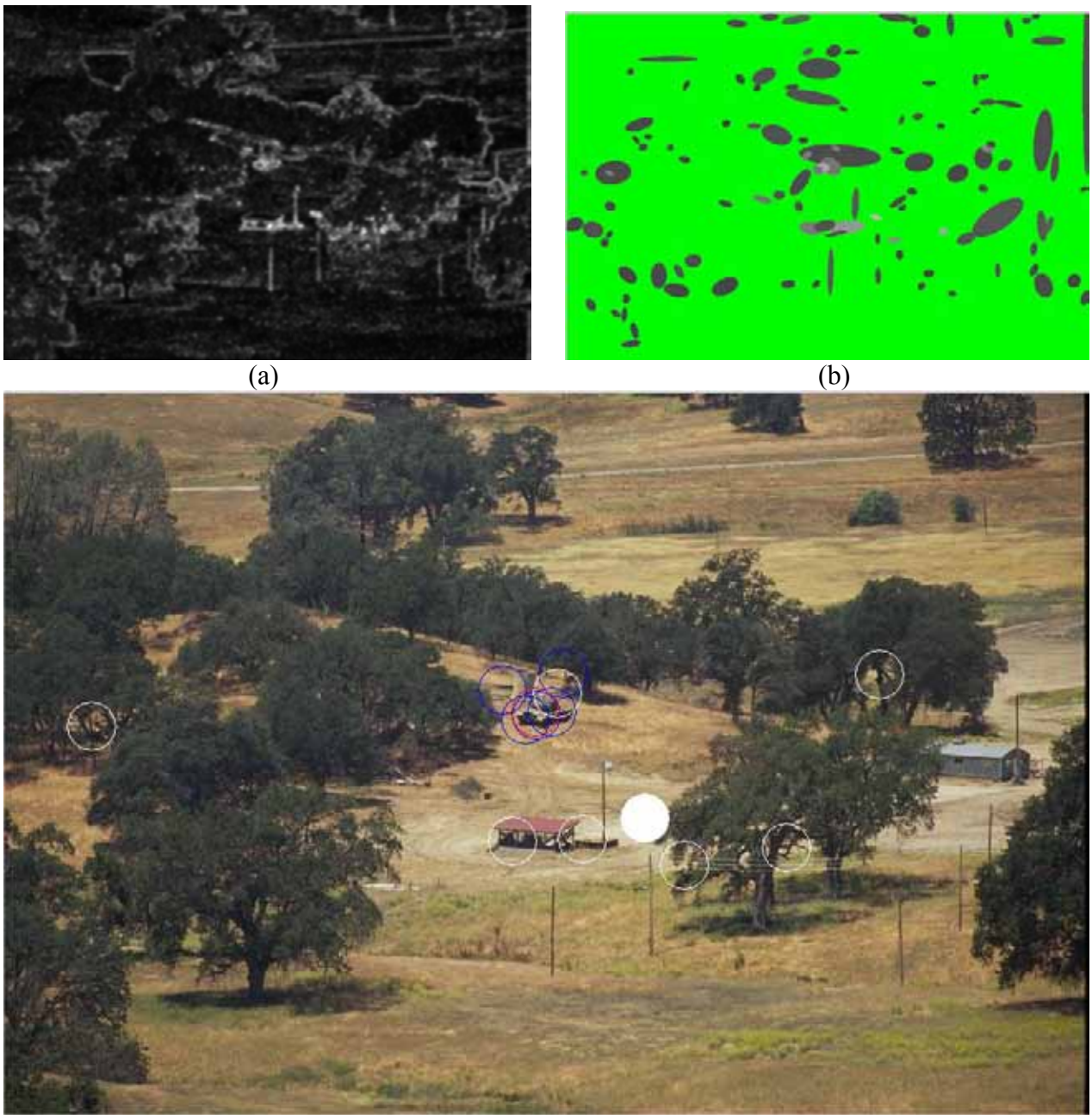
The Night Vision Lab model NVTherm do not consider visual search and only deals with the small time in foveal vision.

In Terrtex some preliminary steps have been taken to estimate the time needed for visual search. Ideas from Laurent Itti [22] have been implemented to a small degree. The first thing is to look for possible false alarms in an image. As in most assessment situations we know where the target is situated. The strength of each potential target is estimated using the features present in the program. Like Itti we combine the features used to a saliency map, which is an image depicting the probability that a target is present at that point. The saliency map is built from the feature images through a fusion process. All the possible alarms are sorted in strength. The somewhat oversimplified model used in the current version of Terrtex assumes that the strongest alarm draws our attention first. We fixate for a small time, about 0.3 s, before the attention is moved to next alarm in the list. This eye movement takes some time. The process stops when we hit the real target. Some steps are shown in Figure 7. This very simple model gives the time to detection as

$$T_{Det} = N_{alarm} * t + Dist * v \qquad (18)$$

where

N_{alarm} = the number of alarms visited before the real target is hit,
t = fixation time for each alarm (about 0.3 s),
$Dist$ = sum of all eye movement distances before the real target is hit,
v = eye movement speed.



(a) (b)

(c)

Figure 7. Illustration the process of estimating detection time: (a) saliency map made from the feature images, (b) blob detection applied to the saliency image, (c) Output image where the positions with a high saliency are marked by circles.

4.4 CALIBRATION

The current version of Terrtex is more or less selfcalibrating, although there are some approximations and assumptions involved. Calibration can be done with a data set received from Toet at TNO [23]. The data set includes 44 images from different forest scenes and complete data from a perception experiment where the probability of detection and the time to detection were measured. A few examples from the data set are shown in Figure 8. Calibration has been applied to an earlier version of Terrtex [9] and the correlation reached between model and experiment was 88 percent. An open question is if the used background is useful in other situations. It might be better to use synthetic images.

Figure 8. Two examples from the image database [23].

5 EXPERIMENT WITH A TARGET SEQUENCE

The purpose of this experiment was to use a model and compare its output when applied to synthetic images with that obtained from registered images. Since we are using images the third model, Terrtex, is used. In order to obtain a reasonable range of detection probabilities it was necessary to reduce the resolution of the images. Another limitation is the difficulty to mark the target borderline if the target area is only a few pixels. Because all the measured QWIP images had target areas of several hundred pixels the images were reduced in size. Then to get a suitable resolution all images were reduced in resolution by a low pass filter of Gaussian shape with a sigma equal to 2. The synthetically images were treated in a similar way. This also had the desired effect that both series of images had the same sensor resolution. Due to difficulties to find a good sequence of QWIP images only a few sizes are available.

5.1 SYNTHETIC IMAGE GENERATION

As field trials and field campaigns can be both expensive and time consuming, there has been an increasing interest in generating synthetic optical imagery using computers [24] in defence applications. The UK developed computer code Camouflage Electro-Optic Simulation (Cameosim) [25] has been used and explored at FOI since early 2003 [26-29]. By feeding Cameosim with information on 3D-terrain and 3D-objects together with material classification and material parameters such as thermal and spectral reflectance properties as well as information about weather conditions, the electro-optic radiation transport equations (RTEs) that affect the scene are solved from first principle physics [24-26]. A sensor is then defined using pixel resolution, field-of-view (fov), spectral response etc, and after (Monte-Carlo based) rendering the scene, 32-bit synthetic output is delivered in form of hyperspectral imagery with wavelengths ranging from 0.3 to 14 μm.

A 2 x 2 km^2 part of the military combat school at Kvarn, North West of Linköping has been modelled starting with a 25 cm resolution laser scan of the area. A material classified polygonised 3D-model has then been imported into Cameosim, where each material has been associated with a number of material properties such as spectral reflectance, heat transfer parameters etc. From the laser scans, extracted information on exact tree positions, tree height, tree width, and tree species (i.e. separating leaf from needle trees) have been used together with 3D-models to model trees in the scene. Spectral reflectance for the scene constituents (such as leafs, grass, gravel, tree bark etc) was measured during a large field campaign in the summer of 2003. The corresponding thermal properties were estimated from available tables and references. By using the link to the internationally recognised atmospheric code MODTRAN [30] for atmospheric modelling (e.g. spectral transmission, path radiance, air temperature, aerosol profiles etc) a complete environment for optical signature simulation in a typical Swedish rural area exists. A military vehicle (T72) was

modelled using the commercial computer code RadThermIR [31]. For more information regarding terrain, vehicle and weather modelling refer to previous work [26-29].

In this work, an in-flight towards a stationary military vehicle (T72) from a range of 1500 m has been simulated in Cameosim. A simple model of the QWIP [32] sensor with a fov of 20°×15°, a Gaussian spectral response with a full width at half maximum (FWHM) of 1.0 µm around 8.5 µm, and 320 × 240 pixels was used. The rendering in Cameosim was carried out using a 15 s in-flight sequence at 10 Hz delivering 150 images. As the sensor model in Cameosim is limited, rendered imagery was further processed to simulate signal degradation due to optics and detector. This was done by estimating modular transfer functions (MTFs) corresponding to diffraction, mean optical aberration and pixel size and pixel pitch [33,34]. As the QWIP sensor normally is run in a mode delivering apparent temperatures, the calculated radiance values in each pixel was converted to apparent temperatures [34]. An Audio/Video Interleaved (AVI) file was then constructed from the image sequence.

5.2 CONTRAST MEASUREMENTS

The feature selected for this experiment was the edge concentration because it is in several cases the most significant feature. The measure was the generalized signal-to-noise ratio. Figure 9 shows how Terrtex is used to calculate the contrast. The part inside the red polygon is compared with the part that is between the two polygons.

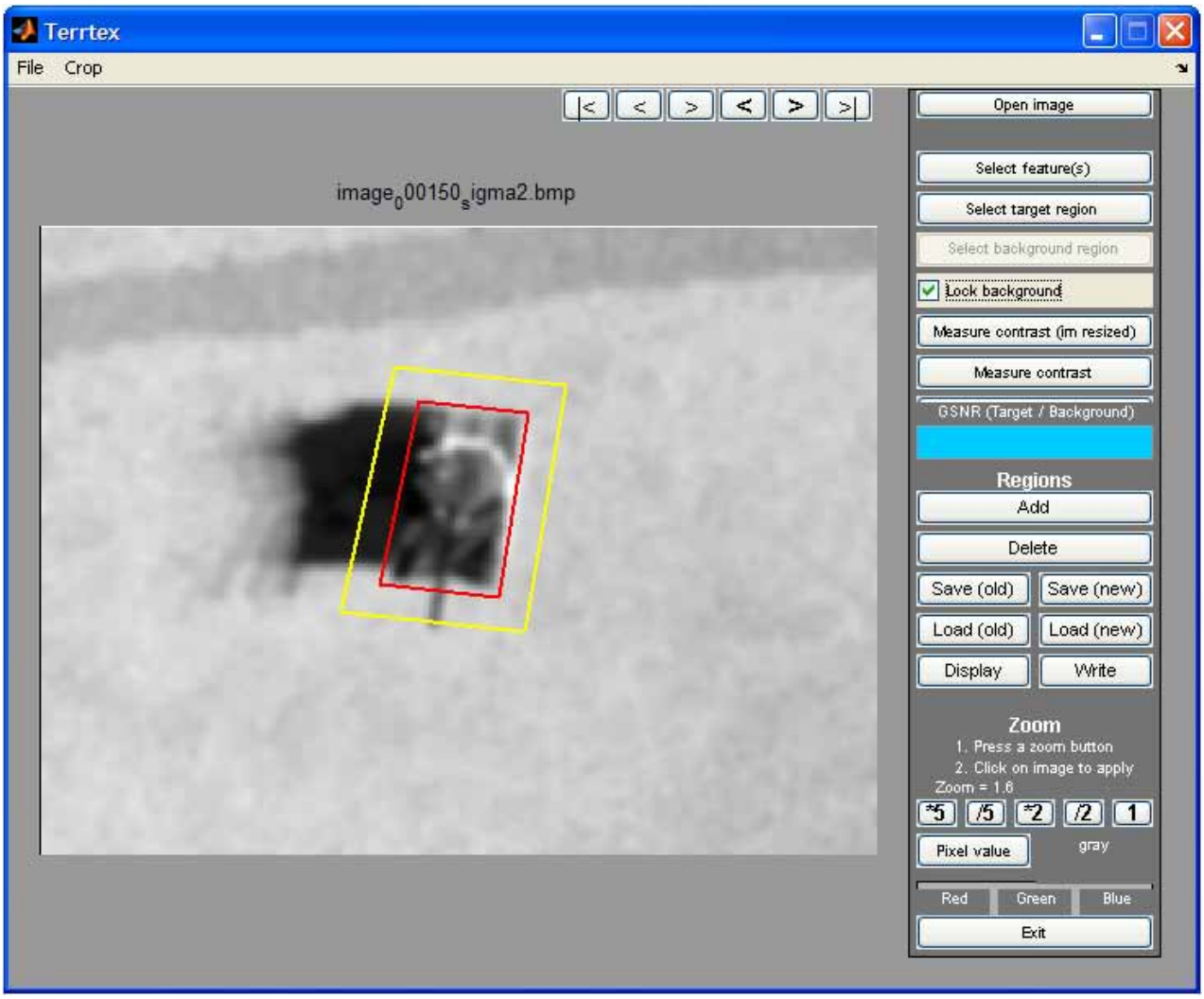


Figure 9. Terrtex used for contrast measurements. The image shows a synthetic image with boxes outlining the target and the background.

Six synthetically generated images and the same number of QWIP images were used. However some of the QWIP images are of approximately the same size. Figure 10 shows the images used.

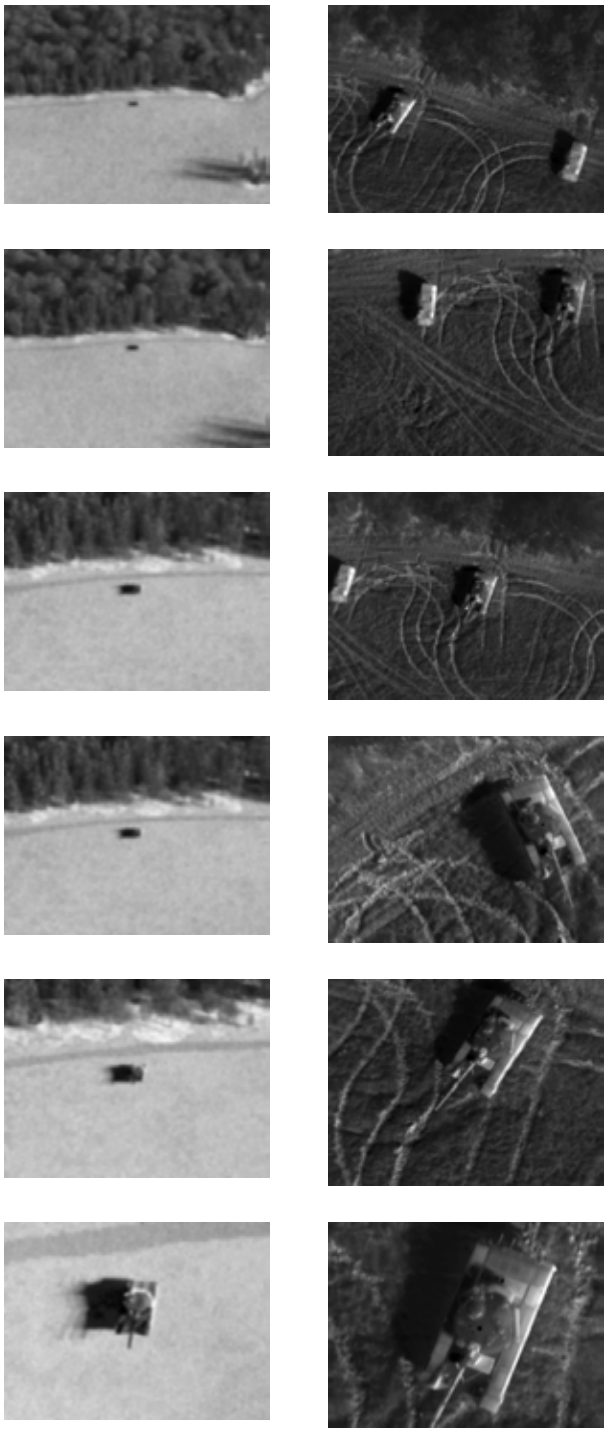


Figure 10. Images used in the experiment. The left column shows the synthetic images while the QWIP images are shown in the right column.

Table 1 and 2 show the results for the synthetic images and QWIP images respectively.

Table 1. Measured generalized signal-to-noise ratios for the synthetic imagery.

Image name	Footprints per target	Edge GSNR	kD (equation 13)	Detection probability
100	2.1	0.811	-0.034	0.35
110	3.6	0.413	-0.319	0.31
120	6.4	0.316	-0.504	0.40
130	8.1	0.270	-0.553	0.42
140	29.8	0.407	-3.03	0.95
150	179.2	2.103	-94.3	1.00

Table 2. Measured generalized signal-to-noise ratios for the QWIP imagery.

Image name	Footprints per target	Edge GSNR	kD (equation 13)	Detection probability
082045	12.0	0.483	-1.45	0.77
082205	12.6	0.537	-1.69	0.82
082326	9.1	1.060	-2.40	0.91
082608	148	0.572	-21.2	1.00
082821	48.4	0.230	-2.81	0.94
082939	59.2	2.159	-26.8	1.00

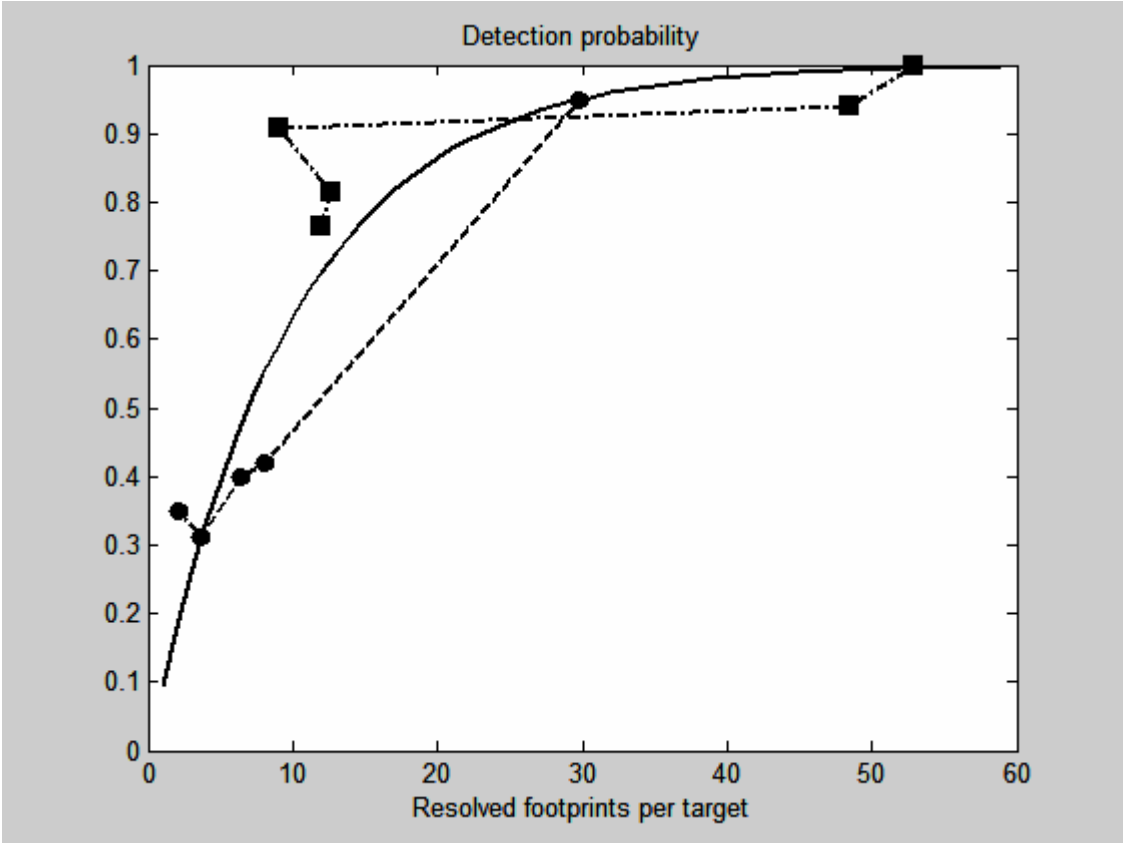


Figure 11. Plot showing the detection probability for synthetic (circle) and QWIP (square) images. The solid lines represents an empirical fitting of the equation 15 in Secion 4.3.1.

As can be seen in Figure 11, the variation in aspect and elevation when registering the QWIP images is not ideal. The goes back to the fact that the QWIP sensor hang in a wire under a helicopter. If a probability function is fitted to data in Figure 11, for both the synthetic imagery

and the QWIP imagery, using equation 15 in Secion 4.3.1, we can see that most observations are quite close to this line. This is interpreted that the two cases, synthetic and registered imagery, both fit the same curve. This indicates, given the uncertainties, that there are no significant difference when comparing the two image sets. An earlier made comparison between synthetic and QWIP images is given in [35].

It is clear that when it comes to signature comparisons great care has to be taken when acquiring test images. Ideally the exact imaging geometry, time of day and weather should be monitored and applied when generating the synthetic images. Similar consideration applies to the registration of the QWIP images.

6 DISCUSSION

By use of image analysis techniques, it is possible to obtain a measure of the similarity between camouflaged targets and the surrounding areas. It is also possible to compare targets having different degrees of camouflage with background areas. The difficult task is the selection of a suitable set of features.

Future work might also include integration of spatial properties with spectral and temporal features. This is necessary if assessment of a given signature suppression measure is to be done. Furthermore, the distance measures have to be ”calibrated”, for example related to recognition distances. Fuzzy logic [36, 37] methods, perhaps implemented with a RBF (Radial basis network) network [38], may be utilized to introduce more knowledge from visual perception than is used in the current investigation. Steps in this direction are the experiments with the Search_2 image data set. The price to be paid for the use of many features is a heavy computation load, a disadvantage that will be less relevant in the future.

Implementing a simple automatic target recognizer (ATR), and performing detection experiments, is one way to obtain a somewhat more objective measure. However, calibration has to be done also in this case. Experiments with the Search_2 image data set indicate that using one single feature will give reasonable results. In these experiments, the features edge concentration and shape seem to give some useful results. The tests indicate that the best result will be obtained using mean and variance based distances.

From the experiment in Section 5, we can see that synthetic imagery may be used instead of registered images. Success demands careful modeling not only of the target and the background but also of the sensor and the atmospheric. This may make computer codes like Cameosim valuable in assessing different parts in an optical system.

7 REFERENCES

[1] Bailey, H. H., “Target detection through visual recognition; A quantitative model,” The Rand Corporation, Memo RM-6158-PR, (AD-721 446), February 1970.
[2] RCA Corporation, Electro Optics Handbook, Technical Series EOH-11, 1974.
[3] J. A. Ratches,”Static performance model for thermal imaging systems,” Opt. Eng., vol. 15, no. 6, Dec. 1976.
[4] J. A. Ratches, R. H. Vollmerhausen, R. G. Driggers,“Target acquisition performance modeling of infrared imaging systems: past, present, and future,” IEEE Sensors journal, vol. 1,no. 1, June 2001.
[5] L. M. Biberman,”Electro-optical imaging: system performance and modeling,”,SPIE Press, Bellingham 2000.
[6] S. Nyberg, M. Uppsäll and L. Bohman, ”An approach to assessment of camouflage methodology,” *Proc. SPIE* 1967,300-307 (1993).

[7] S. Nyberg, M. Uppsäll, "*TERRTEX A Software Package for Texture Analysis",* FOA-R-97-00574-615-SE, National Defence Research Establishment, Division of Sensor Technology, Linköping, Sweden, 1997.

[8] S. Nyberg and K. Schutte, "Assessing Camouflage Methods Using Textural Features," *RTO Meeting Proceedings 45, Search and Target Acquisition*, RTO-MP-45 AC/323(SCI) TP19, NATO Research and Technology Organization, Utrecht, The Netherlands, 10.1-10.10 (1999).

[9] S. Nyberg, and L. Bohman, ” Assessing Camouflage Methods Using Textural Features,” *Opt. Eng.* **40**(9), 1869-1876 (2001).

[10] S. Nyberg, and L. Bohman, “Characterizing low-signature targets in background using spatial and spectral features” in Remote sensing sounders and imagers: instruments, subsystems, and algorithms, Proc. SPIE 5152, 139-149 (2003).

[11] F.M. Gretzmacher, G.S. Ruppert and S. Nyberg, "Camouflage assessment considering human perception data," in Targets and Backgrounds: Characterization and Representation IV, Proc. SPIE 3375, 58-67 (1998).

[12] R. Beichel, G.S. Ruppert and F.M. Gretzmacher, "Fuzzy Logic Approach for the Quantitative Assessment of Camouflage effectiveness in the Thermal Infrared Domain," in *Targets and Backgrounds IV: Characterization, Visualization, and the Detection Process, Proc. SPIE* **4029**, 378-385 (2000).

[13] G.S. Ruppert, R. Beichel and F.M. Gretzmacher, "Robust Measure for Camouflage Effectiveness in the Visual Domain," in *Targets and Backgrounds IV: Characterization, Visualization, and the Detection Process, Proc. SPIE* **4029**, 386-393 (2000).

[14] G. R. Gerhardt and T. J. Meitzler, ”Performance assessment methodology for ground vehicle infrared and visual signature countermeasures (CMs),” *Proc. SPIE* 1687, 334-341 (1992).

[15] C. M. Birkemark, "A methodology for computerized evaluation of camouflage effectiveness and estimation of target detectability," in *Targets and Backgrounds: Characterization and Representation V, Proc. SPIE* **3699**, 229-238 (1999).

[16] C. M. Birkemark, “*CAMEVA, a methodology for estimation of target detectability*,” Opt. Eng. 40(9), 1835-1843 (2001).

[17] T. P. Weldon, W. E. Higgins and D. F.Dunn, "Gabor filter design for multiple texture segmentation," *Opt. Eng*. 35(10), 1- 17 (1996).

[18] H. Tamura, S. Mori and T. Yamawaki, ”Textural features corresponding to visual perception,” IEEE Transaction on systems, man, and cybernetics, VOL. SMC-8, No. 6, 460-473 (1978).

[19] C. Woodruff, ”A proposed methodology for the spatial characterization of foliage backgrounds,” Department of Defense, Technical note MRL-TN-510, AD-A178 747, 1987.

[20] W. D. Stromberg and T. G. Farr, ”A Fourier-based textural feature extraction procedure,” IEEE Transactions on geoscience and remote sensing, VOL. GE-24, NO. 5, 722-731 (1986).

[21] K. Fukunaga, “Introduction to statistical pattern recognition (2nd edition),” Academic Press, New York, 1990.

[22] L. Itti, C. Gold, C. Koch,”Visual attention and target detection in cluttered natural scenes,” Opt. Eng., 40(9), 1784-1793 (2001).

[23] A. Toet, P. Bijl, F. L. Kooi and J. M. Valeton, ”A high-resolution image data set for testing search and detection models,” TNO-report TM-98-A020, April 1998.

[24] D. Filbee et al., “*Modeling of High Fidelity Synthetic Imagery for Defence Applications*,” Proceedings of SPIE Vol. 4718, 2002, s. 12.

[25] Insys Ltd, *User Manuals, CAMEO-SIM v. 5.5.1*, 2006.

[26] T. Winzell, “*Elektro-Optisk Modellering med CAMEO-SIM*,” FOI-RH--68--SE, 2003.

[27] T. Winzell, “*Vehicle modelling within the JP 8.10 THALES program for WP3*,” FOI Memo H186, 2004.

[28] C. Nelsson et al., "*Optical signature modelling – Final report,"* FOI-R--1812--SE, 2005.
[29] T. Winzell, *Modellering av JAS39 Gripens IR-signatur med CAMEO-SIM*, FOI-RH--0297--SE, 2004.
[30] F.X. Kneizy et al., *"The MODTRAN 2/3 Report and LOWTRAN 7 Model*," Phillips Laboratory, Geophysics Directorate, PL/GPOS Hanscom AFB MA 01731-3010, USA, 1996.
[31] See e.g., www.thermoanalytics.com
[32] See e.g., www.acreo.se/templates/Page____996.aspx
[33] T. Winzell, "*Sensor Modelling on Synthetic Input",* FOI Memo 1508, 2005.
[34] T. Winzell, "*Basic Optical Sensor Model,"* FOI-R--2135--SE, 2006.
[35] P. Hermansson, S. Nyberg, E. Andersson, D. Börjesson,"Metods for validating optical signature simulations – progress report," FOI-R—1421—SE, Swedish Defence Research Agency, Division of Sensor System, Linköping, Sweden, 2004.
[36] T. J. Meitzler, E. Sohn, H. Singh and A. Elgarhi, "Predicting search time in visual scenes using the fuzzy logic approach," in *Targets and Backgrounds: Characterization and Representation V, Proc. SPIE* **3699**, 229-238 (1999).
[37] L. Gabrijel, A. Dobnikar and N. Steele, "RBF Neural Networks and Fuzzy Logic Based Control - a Case Study," Proceedings of the 2nd IMACS International Multiconference CESA'98, Hammamet, Tunisia, vol. 2, 284-289 (1998).
[38] S. V. Chakravarthy and J. Ghosh, "Scale-Based Clustering Using the Radial Basis Function Network," IEEE transactions on neural networks, Vol. 7, No 5, September 1996.