

# Sensor performance as a function of sampling ( $d$ ) and optical blur ( $F\lambda$ )

Piet Bijl & Maarten A. Hogervorst

TNO Defense, Security & Safety: Human Factors, P.O. Box 23, Soesterberg, The Netherlands

Phone: +31 346 356 368, fax: +31 346 353 977

E-mail: piet.bijl@tno.nl, maarten.hogervorst@tno.nl

## ABSTRACT

Detector sampling and optical blur are two major factors affecting Target Acquisition (TA) performance with modern EO and IR systems. In order to quantify their relative significance, we simulated five realistic LWIR and MWIR sensors from very under-sampled (detector pitch  $d \gg$  diffraction blur  $F\lambda$ ) to well-sampled ( $F\lambda \gg d$ ). Next, we measured their TOD (Triangle Orientation Discrimination) sensor performance curve. The results show a region that is clearly detector-limited, a region that is clearly diffraction-limited, and a transition area. For a high contrast target, threshold size  $T_{FPA}$  on the sensor focal plane can mathematically be described with a simple linear expression:  $T_{FPA} = 1.5 \cdot d \cdot w(d/F\lambda) + 0.95 \cdot F\lambda \cdot w(F\lambda/d)$ ,  $w$  being a steep weighting function between 0 and 1. Next, tacticle vehicle identification range predictions with the TOD TA model and TTP (Targeting Task Performance) model were compared to measured ranges with human observers. The TOD excellently predicts performance for both well-sampled and under-sampled sensors. While earlier TTP versions (2001, 2005) showed a pronounced difference in the relative weight of sampling and blur to range, the predictions with the newest (2008) TTP version that considers in-band aliasing are remarkably close to the TOD. In conclusion, the TOD methodology now provides a solid laboratory sensor performance test, a Monte Carlo simulation model to assess performance from sensor physics, a Target Acquisition range prediction model and a simple analytical expression to quickly predict sensor performance as a function of sampling and blur. TTP approaches TOD with respect to field performance prediction.

Keywords: TOD, TTP, Target Acquisition, range prediction, diffraction blur, detector size

## 1. INTRODUCTION

When we perceive scenes and objects with natural vision, our Visual Acuity (VA, i.e. the amount of detail that we can resolve with foveal vision) is limited by the optics blur in our eye. Vision through optics such as binoculars does not essentially differ from that with the unaided eye, and to a certain extent this also holds for conventional scanning Electro-Optical (EO) systems.

The current EO systems of choice, however, have a pixel array sensor or Focal Plane Array (FPA), and these systems are usually more or less under-sampled. This means that the scene information is spatially sampled with a spacing that can be relatively coarse compared to the optics blur of the EO system. The results are i) a different type of degradation, ii) an image that is phase dependent (i.e. the shape of an object image depends on the relative position of the object with respect to the sensor pixels) and iii) the occurrence of aliasing or spurious response. These characteristics make Target Acquisition (TA) or object recognition with an under-sampled imager essentially different from that with a well-sampled imager.

Recent studies<sup>1,2</sup> show that  $F\lambda/d$  is an important parameter with respect to the amount of under-sampling in pixel-based systems.  $F$  is the F-number of the optics,  $\lambda$  is the average spectral wavelength of the sensor system and  $d$  is the detector pitch of the pixels on the Focal Plane Array (FPA). This makes sense because the variable  $F\lambda$  is associated with diffraction blur (e.g., the Airy disk diameter of the optics is  $0.41 \cdot F\lambda$ , and the optical cut-off  $f_0 = (F\lambda)^{-1}$ ) while  $d$  is associated with FPA resolution (Nyquist frequency  $f_N = 1/(2d)$ ). So, for a well-sampled system we expect the smallest resolvable details with the system to be proportional to  $F\lambda$  and for a very under-sampled system to be proportional to detector pitch. A transition from one to the other region is exhibited when  $0.41 < F\lambda/d < 1.0$ <sup>1,2</sup>. In this region, both sampling and blur contribute to the visual performance with the sensor system.

Several promising models exist to predict TA performance with EO sensors: the TOD method and model<sup>3,4,5,6</sup>, TRM3<sup>7,8</sup> and the TTP metric<sup>9,10,11</sup>. They all claim to cope with well-sampled and under-sampled imagers, but they disagree in the relative importance of detector spacing and diffraction blur on the 'effective resolution' or VA with the sensor. In addition, they behave differently in the transition region<sup>12,13</sup>. For example, a quantitative comparison between TOD and TTP (the model underlying the NVTherm thermal imager range prediction model) in 2002 showed huge differences<sup>12</sup>. Predictions with an updated version of the TTP metric (the basis of the NVThermIP model) are much closer to the

TOD<sup>13</sup>: when the performance with blur-limited sensors is set equal, NVThermIP is about 25% more optimistic in range performance prediction for under-sampled imagers than the TOD is. While the TOD essentially remained unchanged over the years, the TTP metric has undergone major changes. Until now, the TTP metric was based on the EB (Equivalent Blur) assumption and neglected part of the aliasing called “in-band aliasing”. The newest version of the TTP metric<sup>11</sup>, indicated as the AAN (Aliasing As Noise) model, does take in-band aliasing into account.

In this study, we try to disentangle the contribution of blur and sampling to the TOD of a sensor system. This will result in a very simple approximation of the sensor’s Visual Acuity. We will perform some example range calculations and compare the results to those with different versions of the TTP metric.

This paper is organized as follows. The experimental method is described in Chapter 2. Results are given in Chapter 3, and modeled using a simple weighted linear sum model based on blur and detector size in Chapter 4. Example range predictions with the model are given in Chapter 5, and the results are discussed in Chapter 6.

## 2. METHODS

### 2.1 TOD SENSOR PERFORMANCE CHARACTERIZATION

The Triangle Orientation Discrimination<sup>3</sup> (TOD) method is a way to quantify sensor performance with a human-in-the-loop. The method provides a sensor performance measure (the TOD curve) and a TA (Target Acquisition) model to calculate range performance in the field (the TOD TA model). A recent overview of the TOD method is described by Bijl & de Vries<sup>9</sup>.

Basically, the TOD curve describes the ability of a human observer using the sensor under test to discriminate between equilateral triangles of four possible orientations (apex Up, Down, Right or Left, see Figure 1). Independent variables are triangle angular size and (thermal or visual) contrast.

Visual Acuity (VA) is one of the special points on the TOD curves and represents the ability to resolve spatial details at the high contrast limit.

A TOD curve can be measured with real or a simulated sensor, and with a human observer or a vision model<sup>5,14</sup>.

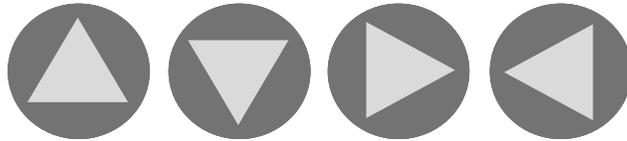


Figure 1 The test pattern or stimulus in the TOD method is an equilateral triangle with one of four possible orientations: apex Up, Down, Left or Right. The observer has to indicate its orientation. Task difficulty depends on test pattern size and contrast. From Bijl & Valetton<sup>9</sup>.

### 2.2 TEST PATTERN GENERATION AND SENSOR SIMULATION

A simulation platform EO-VISTA<sup>15</sup> that includes a triangle test pattern generator, the pcSitoS sensor simulation model<sup>16</sup> and a Human Visual System (HVS) model<sup>5</sup> was used to generate the test patterns and simulate the image degradation through a number of sensors (see 2.4). Choices with the simulation platform are given in section 2.3-2.5.

### 2.3 TEST TARGETS

Test patterns of 7 different sizes around the threshold were generated for each sensor, and each size was generated 30 times with different orientation and a randomized position with respect to the FPA. In order to avoid effects of noise on the results, thermal contrast was high: 20K.

### 2.4 CHOICE OF SENSORS

We selected a number of typical MWIR and LWIR sensors. See Table 2. The FLIR SC2000 uncooled microbolometer is extensively investigated and data with the real and simulated sensor are reported in a series of papers<sup>15,17,18</sup>. The other sensors were selected from Krapels et al. (2007)<sup>2</sup>. Using these sensors is convenient because imagery and model calculations are available.

The result is a set of LWR and MWIR sensors with an  $F\lambda/d$  varying from 0.21 (very under-sampled) to 2.21 (well-sampled), varying in spectral wavelength, aperture, focal length, detector pitch  $d$ , fill factor, and FOV.

## 2.5 TEST AND OBSERVERS

### 2.1.1. Human observers and observer model

Four human observers participated in the experiment: NL, JS, AW and TB.

In addition, the test was run with the HVS model<sup>5</sup> instead of a human observer.

### 2.1.2. Observer experiment

The total number of images was 7 (test pattern sizes) \* 30 (repetitions per size) \* 5 (sensors) = 1050. Each image was presented once to each observer and to the HVS model. Observers had to judge the orientation of the triangle, even if they were not sure. In order to minimize learning effects, the order of the presentations was first randomized, then divided into four blocks and the blocks were presented in different order to each human observer according to a 4 by 4 Latin Square design. Total measurement time per observer is approximately 75 minutes.

The result of the test is a 75% correct threshold triangle size for each sensor and each observer. Each threshold is based upon 210 observations.

## 3. RESULTS

Figure 2 shows the probability vs. triangle size relationship for the five sensors as measured with the four human observers and the HVS model. A Weibull curve is fitted through the data (filled lines: human observers; dashed lines: HVS model). This fit was accepted in all 25 cases. For all observers, threshold sizes at the 75% correct level  $S$  (in mrad) are tabulated in Table 3. VA is defined as  $VA = S^{-1}$ . The accuracy for each threshold estimate is high: standard errors are in the order of 3-4%.

### 3.1 HUMAN OBSERVER THRESHOLDS

One of the four observers (AW) scores significantly lower than the other three: for the five sensors threshold size is 2, 2.5, 3, 6, and 9 times the standard error above the results for the other three observers. This observer is removed from further analysis. In Table 3, we report the weighted average  $S_{human\ observers}$  and standard error in the mean (maximum of internal and external error) for the resulting three observers. Again, the error is small: 3-6%.

### 3.2 HVS THRESHOLDS

Threshold sizes  $S_{HVS}$  for the HVS model are systematically higher than  $S_{human\ observers}$ . The ratio slightly depends on the amount of under-sampling: Below  $F\lambda/d = 1$  (detector limited), the ratio is independent and around 0.77, above  $F\lambda/d = 1$  the ratio increases above 0.88 and seems to approach 1.0, i.e. in the case of a diffraction-limited system the HVS model perfectly predicts human Visual Acuity. The same set of test patterns were presented to observers and HVS model, so the difference in performance cannot be ascribed to coincidental differences in test sets. We conclude that the HVS model comes close to the human acuity for well-sampled systems but needs improvement for under-sampled systems (see Discussion).

In addition, with the HVS model the slopes of the probability versus size curves for low  $F\lambda/d$ -values seem to be shallower than with the human observers (see Figure 2).

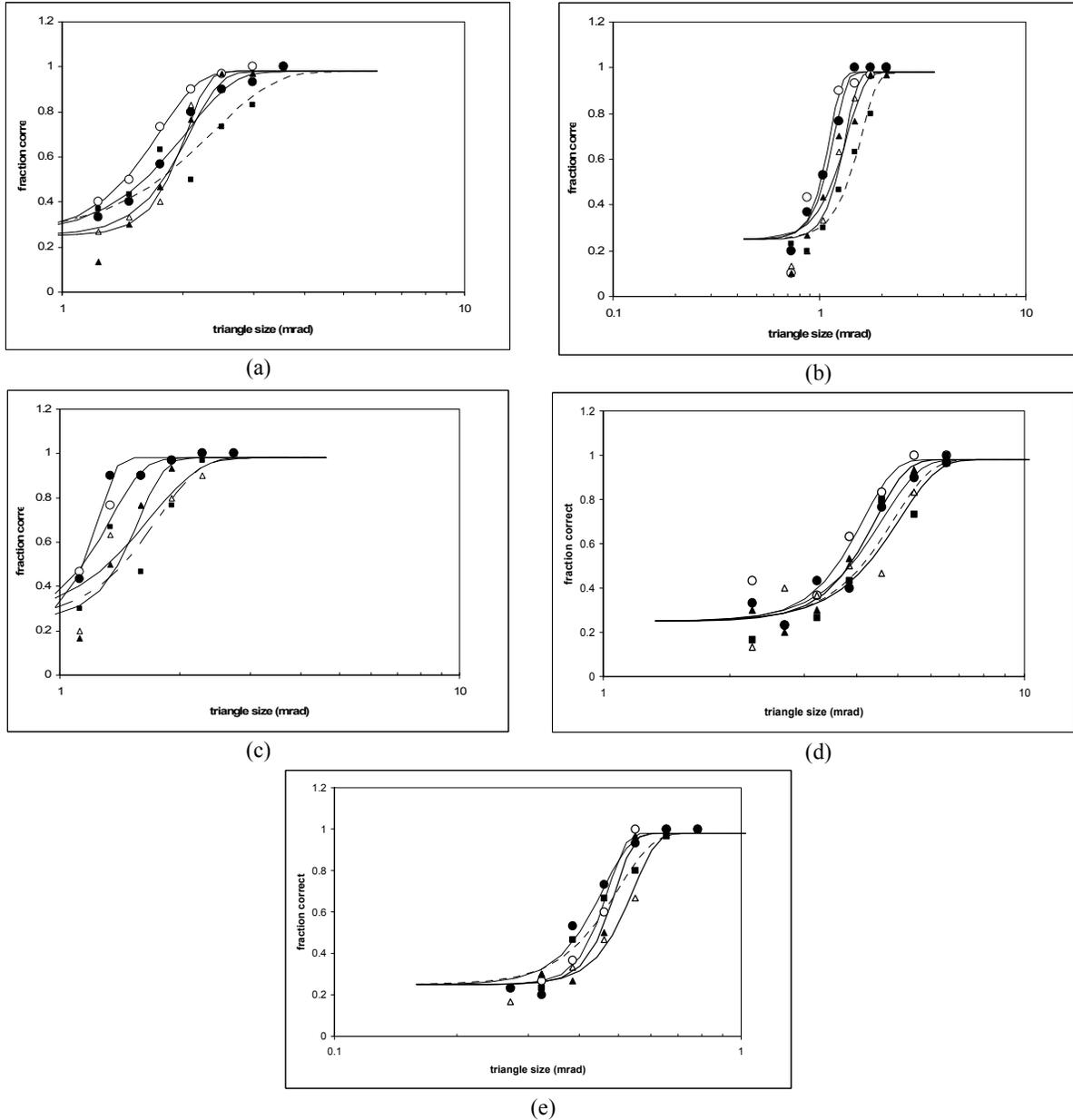


Figure 2 Fraction correct versus triangle size (in mrad) and best fits for the five simulated sensors with the four observers and the HVS model. Target thermal contrast: 20K. (a) FLIR SC2000; (b) ‘Typical LWIR’; (c) TVS 8502; (d) Indigo Phoenix; (e) DRS heavy TWS. Open circles: NL; filled circles: JS; open triangles: AW; filled triangles: TB; filled squares: HVS model. Solid lines: best fits to the human observer data; dashed lines: best fits to the HVS model data.

### 3.3 TRIANGLE TRESHOLD SIZE AS A FUNCTION OF $F\lambda/D$

In Figure 3, triangle threshold size, now expressed in  $\mu\text{m}$  at the Focal Plane, is plotted as a function of  $F\lambda/d$ . Filled triangles show the average results for human observers, while open triangles show the data for the HVS model. Also included in the plot are detector pitch  $d$  (open squares) and  $F\lambda$  (open circles), all in  $\mu\text{m}$ . Error bars indicate the standard error in the mean.

In Figure 3 we see that:

- For small  $F\lambda/d$ , triangle size at the focal plane seems to be proportional to detector pitch  $d$ .

- For large  $F\lambda/d$ , triangle size at the focal plane is close to  $F\lambda$ .

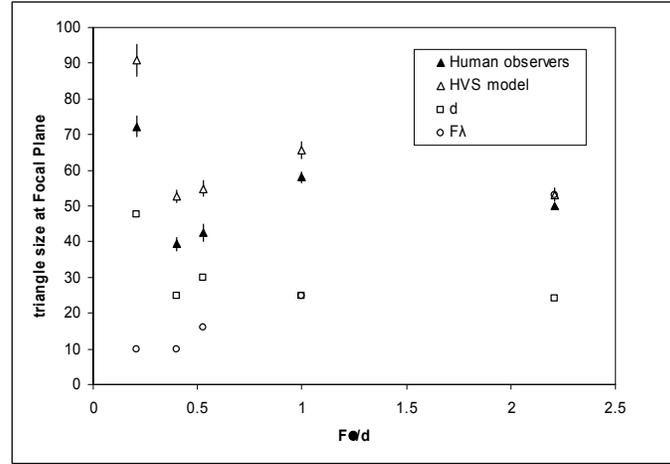


Figure 3 Triangle threshold size at the Focal Plane (in  $\mu\text{m}$ ) as a function of  $F\lambda/d$ . Filled triangles: human observers; open triangles: HVS model. Also plotted are detector pitch  $d$  (open squares) and  $F\lambda$  (open circles).

## 4. MODELING THE RESULTS

### 4.1 WEIGHTED LINEAR SUM MODEL

For small  $F\lambda/d$  we expect detector pitch  $d$  to be the major factor determining TOD acuity. This is in agreement with the results in Figure 3: below  $F\lambda/d=1$  we see that the triangle threshold size nicely follows the detector pitch.

For large  $F\lambda/d$  we expect  $F\lambda$  to be the major factor determining TOD acuity. Above  $F\lambda/d = 2$ , the system is well-sampled and the effect of detector pitch should be very small.

In the intermediate range we expect both  $d$  and  $F\lambda$  to contribute, and the total degradation should be more than the separate contributions from sampling and blur.

We assume the threshold size to be a weighted linear sum of a detector-limited and blur-limited component:

$$(1) T_{FPA} = c_1 \cdot d \cdot w\left(\frac{d}{F\lambda}\right) + c_2 \cdot F\lambda \cdot w\left(\frac{F\lambda}{d}\right)$$

where  $T_{FPA}$  is the TOD triangle threshold size at the Focal Plane.  $T_{FPA}$ ,  $d$  and  $\lambda$  are all in  $\mu\text{m}$ .  $w(x)$  is a weighting function that is close to 1 for  $x \leq 1$ , and rapidly drops to zero for  $x > 2$ . We assume that the weighting function is described by a Weibull function with slope  $\theta$ :

$$(2) w(x) = 1 - 10^{-x^\theta}$$

Further note that:

$$(3) \frac{F\lambda}{d} = \frac{f_{0,detector}}{f_{0,optics}}$$

where  $f_{0,detector}$  and  $f_{0,optics}$  are the MTF cut-off frequencies of the detector and optics, respectively. From (1) it also follows that:

$$(4) VA = \frac{f}{c_1 \cdot d \cdot w\left(\frac{F\lambda}{d}\right) + c_2 \cdot F\lambda \cdot w\left(\frac{d}{F\lambda}\right)}$$

where VA is Visual Acuity through the sensor in  $\text{mrad}^{-1}$ , and  $f$  = focal length in mm

#### 4.2 RESULTS FOR THE TOD

See Figure 4. The weighted linear sum model nicely fits the data, even though a large number of sensor parameters were neglected. Parameter values  $c_1$  and  $c_2$  were optimized to match the data at the lower and the higher limits of  $F\lambda/d$ , while parameter  $\theta$  was optimized for a good match in the middle region. The values are given in Table 1, row 2 (human observer) and row 3 (HVS model). The weighting functions are shown in Figure 5.

The data in Table 1 indicate that:

- For a detector-limited sensor system, TOD triangle threshold size at high contrast is approximately 1.5 times detector pitch  $d$ . This result has been reported in earlier studies
- For a diffraction limited system, TOD triangle threshold size at high contrast is approximately 0.95 times diffraction blur  $F\lambda$
- The HVS model matches the human observer data for diffraction-limited sensors but underestimates performance for under-sampled imagers by a factor of  $1.5/1.9 = 0.79$ . See also the Discussion.

Table 1 Optimum parameter values for the weighted linear sum model

	$c_1$	$c_2$	$\theta$
TOD with human observer	1.5	0.95	5
TOD with HVS model	1.9	0.95	5

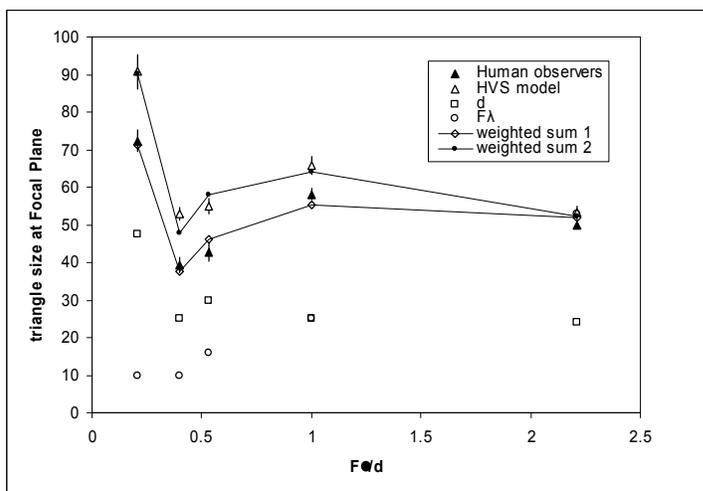


Figure 4 Data from Figure 3, including the weighted sum model predictions (see equation 1) fitting the human observer (weighted sum 1) and the HVS model (weighted sum 2) data. Fit parameters are given in Table 1.

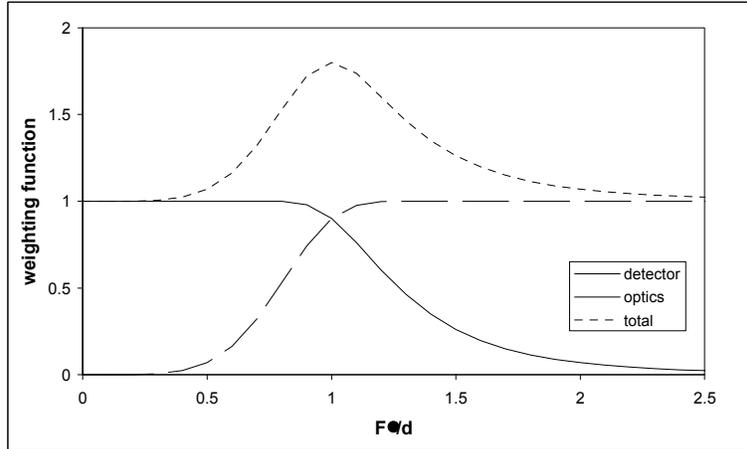


Figure 5 Weighting functions  $w(d/F\lambda)$ ,  $w(F\lambda/d)$  with slope  $\theta = 5$  (equation 2) and their sum as a function of  $F\lambda/d$ .

### 4.3 NVTHERMIP CALCULATIONS

With the sensors listed in Table 2, NVThermIP range calculations were performed. For each sensor, optimum viewing distance was chosen. Using the procedure described elsewhere<sup>13</sup>, the range calculations are calculated back to target size on the Focal Plane Array.  $V_{50}$  was chosen in such a way that for the sensor with the highest  $F\lambda/d$  value (the most well-sampled sensor), predicted triangle size matched that of the TOD measurements. Thus, the ratio between predicted and measured triangle size (and predicted range) is 1.0 for sensor number 5. Predicted range versus  $F\lambda/d$  showed a different behaviour than with the TOD. For example, the ratio is around 1.27 for a typically under-sampled imager (sensor 1 and 2). This is in agreement with the earlier finding<sup>13</sup> that NVThermIP is about 25% more optimistic for under-sampled imagers than TOD compared to well-sampled imagers.

We were able to fit the relationship between predicted range and  $F\lambda/d$  but with different constants  $c_1$  and  $c_2$ , a shallower weighting function ( $\theta = 2$ ) and a transition from detector-limited to blur-limited is around  $F\lambda/d = 0.7$  instead of 1.

## 5. TA RANGE PREDICTIONS WITH THE WEIGHTED LINEAR SUM MODEL

In this section we will show some example calculations with the weighted linear sum model defined in section 4.1. We will use the parameters deduced for the TOD measurements with the human observer (see Table 1), unless stated otherwise. Since the model is deduced for acuity-limited performance, we assume high contrast and no atmospheric losses.

### 5.1 TYPICAL UNDER-SAMPLED SENSOR

VA from equation 4 and  $w(x)$  from equation 2 are the starting point for our calculations, with  $c_1 = 1.5$ ,  $c_2 = 0.95$ , and  $\theta = 5$ . Probability versus range relationship  $P(r)$  is calculated with the following two equations<sup>6,13</sup>:

$$(6) r_{75} = \frac{A}{S_{C_0} \cdot M_{75}} = \frac{A \cdot VA}{M_{75}}$$

$$(5) P(r) = 1 - 4^{-(r_{75}/r)^\beta}$$

where  $r_{75}$  is the target acquisition range (in km) at probability  $P = 0.75$ ,  $\beta$  determines the slope of the relationship and has value  $\beta = 1.25$  in our model<sup>6,13</sup>,  $A$  is the target set characteristic size (in m, see section A.2),  $C_0$  is the target set characteristic inherent RSS contrast ( $\Delta T_0$  for thermal, see section A.2), and  $S_{C_0}$  ( $= VA^{-1}$  in the high contrast limit) is the corresponding triangle threshold size (triangle square-root area in mrad) from the TOD curve. Finally,  $M_{75}$  is a magnification factor between required target and triangle size that quantifies the task difficulty (see section A.2).

For the FLIR SC2000, we have (Table 2):  $d = 47.5 \mu\text{m}$ ;  $F\lambda = 10 \mu\text{m}$  (F/1 optics), and  $f = 36.3 \text{ mm}$ . Typical values for tracked armored vehicle identification with LWIR sensors are<sup>6,19</sup>:  $M_{75} = 14.0$ ,  $A = 3.0 \text{ m}$  and  $\Delta T_0 = 4.7 \text{ K}$ .

For this condition we find that  $VA = 0.51 \text{ mrad}^{-1}$ , in agreement with reported measurements<sup>15</sup>, and  $r_{75} = 110 \text{ m}$ . Probability versus range is shown in Figure 6 (solid line).

## 5.2 RANGE IMPROVEMENT WITH FOCAL LENGTH

Suppose we equip the FLIR SC2000 (section 5.2) with an objective with tenfold focal length ( $f = 363 \text{ mm}$ ). With F/1 optics (lens diameter equals focal length), range increases by the same factor. But if the lens diameter is  $36.3 \text{ mm}$  we have an F/10 optics and blur starts to play an important role.

Again using equations (2) and (4), we end up with  $VA = 3.7 \text{ mrad}^{-1}$  and  $r_{75} = 790 \text{ m}$ , i.e. an effective ID range increase of a factor of 7.2 instead of 10.

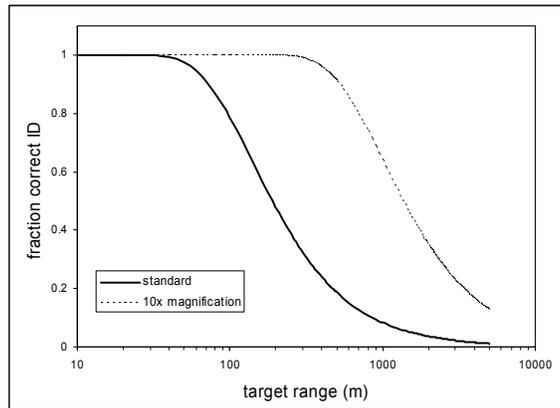


Figure 6 Probability versus range relationship example for target ID with the under-sampled FLIR SC2000 sensor. Range is plotted on a log scale. Solid line: standard optics (section 5.1). Dashed line: optics with 10x higher focal length and F-number (section 5.2). In the second example, diffraction blur plays a significant role and the effective ID range increase is 7.2 times.

## 5.3 TACTICAL VEHICLE IDENTIFICATION DATA

### 5.3.1 Experimental data

Tactical vehicle ID data for well-sampled and under-sampled imagery are used from a series of TTP validation experiments published by Vollmerhausen et al. (2004, 2008)<sup>10,11</sup>. Details on the experiments and sensor specifications are reported in the referred papers. Experiment #25 is performed using a sensor with realistic in-band aliasing, and was set up to determine the effect of display pixel visibility (E-zoom and pixel interpolation) on target ID. Experiment #36 was performed to vary the amount of aliasing by changing the fill factor. Both experiments contained six conditions (indicated by lines 1-6) including normal and extreme parameter values. Figure 7 shows part of the data (open and filled diamonds) obtained with these experiments. Left graph: experiment 25. Right graph: experiment 36. Ranges appear to be rather insensitive to the independent variables in the experiments: if we consider all conditions in experiment 25,  $r_{50} = 1.8\text{-}2.1 \text{ km}$ , i.e. the maximum range variation is 15%. In experiment 36, all ranges are close to  $r_{50} = 0.9 \text{ km}$ .

### 5.3.2 TOD range predictions

On the basis of a number of experiments with well-sampled imagery, the authors of the TTP validation study<sup>10</sup> derived the target set specific parameters required for model predictions:  $V_{50} = 20.8$ , target characteristic size =  $3.1 \text{ m}$ , and target characteristic contrast is high (sufficiently high to be resolution limited). With these parameters and the  $V_{50} \leftrightarrow M_{75}$  conversion derived for well-sampled imagers<sup>13</sup> ( $M_{75} = 0.60 \cdot V_{50} = 12.5$ ) we can make tactical vehicle ID range predictions using the TOD weighted linear sum model derived in Chapter 4. For the sensor in experiment 25, we have  $d = 30 \mu\text{m}$ ,  $F\lambda = 8 \mu\text{m}$  and  $f = 220 \text{ mm}$ <sup>11</sup>, so  $VA = 4.88 \text{ mrad}^{-1}$  and  $r_{75} = 1.21 \text{ km}$ . For the sensor in experiment 36, we have  $d = 25 \mu\text{m}$ ,  $F\lambda = 8 \mu\text{m}$  and  $f = 73.3 \text{ mm}$ <sup>11</sup>, so  $VA = 1.95 \text{ mrad}^{-1}$  and  $r_{75} = 0.48 \text{ km}$ .

The probability vs. range relationships predicted with the TOD TA model are shown in Figure 7 together with the data. Predictions are excellent: for experiment 25, the predictions perfectly match for ID levels above 0.40 and slightly deviate below this level. For experiment 36, ranges at higher probability levels are slightly underestimated (approximately 20% at the 75% correct ID level) but in general the curve matches the data well. The average ratio between measured and predicted range at the 50% correct ID level over the four conditions shown is 1.03 with a standard deviation of 0.12.

Note that there are *no* free parameters: target set parameters were derived for well-sampled imagery in the original study<sup>10</sup> and  $V_{50}$  was converted to  $M_{75}$  for well-sampled imagery<sup>13</sup>. So TOD predictions match tactical vehicle ID data for both well-sampled and under-sampled imagery.

For each experiment, there is only one predicted curve because the weighted linear sum model is based on only three sensor variables and does not differentiate between the experimental conditions within the experiment. As an end-to-end sensor performance measure, the real TOD will make a difference between the experimental conditions.

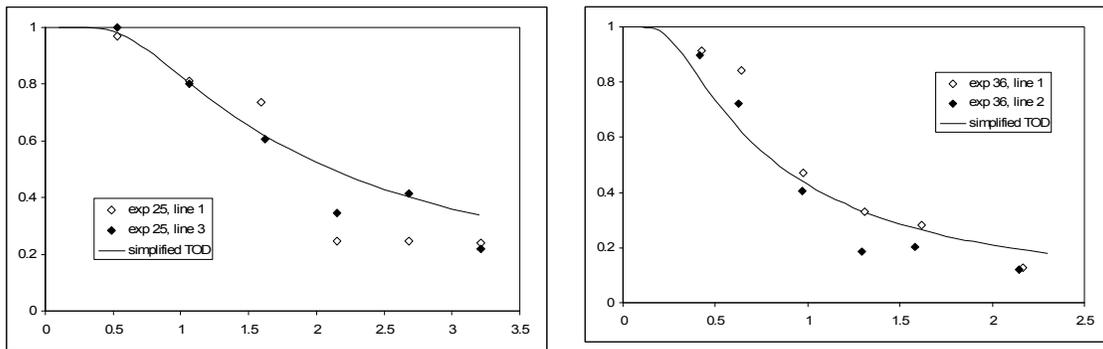


Figure 7 Experimental tactical vehicle ID data (open and filled diamonds) from two TA model validation experiments (# 25 and # 36, see section 5.3.1) with under-sampled imagers published by Vollmerhausen et al. (2004, 2008)<sup>10,11</sup> and TOD predictions (filled lines, see section 5.3.2). The predictions are excellent. Note that there are *no* free parameters: target set specific parameters were derived for well-sampled imagery in the original study<sup>10</sup> and  $V_{50}$  was converted to  $M_{75}$  for well-sampled imagery<sup>13</sup>. So TOD predictions match tactical vehicle ID data for both well-sampled and under-sampled imagery.

### 5.3.3 TTP range predictions

TTP range predictions for experiments 25 and 36 are taken from Vollmerhausen et al. (2008)<sup>11</sup>. The results are shown in Figure 8, together with the observer data (symbols as in Figure 7). Left graph: exp 25; Right graph: exp 36.

Filled and dash-dot lines represent the predictions with the TTP EB (Equivalent Blur) model. Filled lines predict the open diamond data, dash-dot lines correspond to the filled diamond data. In all conditions, the EB model considerably overestimates measured TA range: the average ratio between measured and predicted range at the 50% correct ID level over the four conditions shown is 1.49 with a standard deviation of 0.30.

Dotted and dashed lines represent the predictions with a new TTP metric: the AAN (Aliasing As Noise) model. This model takes into account the effects of in-band aliasing. Dotted lines correspond to the open diamond data, dashed lines to the filled diamond data. With this model, the average ratio between measured and predicted range at the 50% correct ID level over the four conditions shown is 0.94 with a standard deviation of 0.12.

Obviously, the AAN model provides a much better fit to these experimental data than EB does, and TOD and AAN prediction accuracy to these data are comparable.

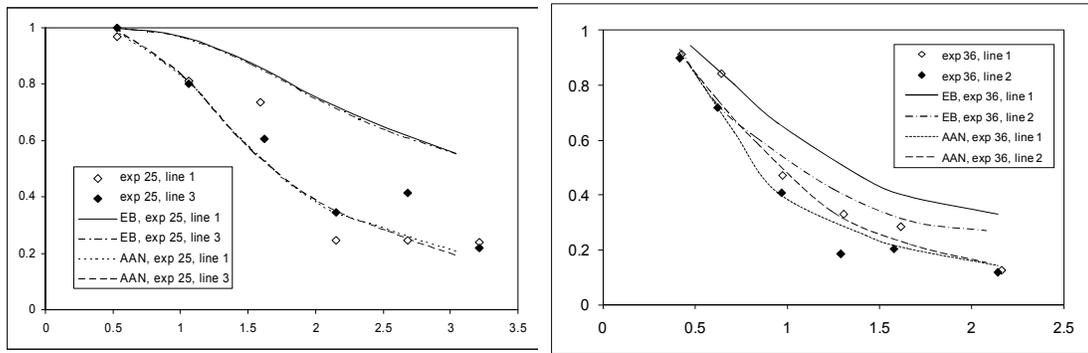


Figure 8 Same experimental data as in Figure 7. In addition, range predictions with the TTP Equivalent Blur (EB) model<sup>10</sup> (filled and dash-dot lines) and the TTP Aliasing As Noise (AAN) model<sup>11</sup> (dotted and dashed lines). See section 5.3.3. The AAN model takes into account in-band aliasing and better predicts the experimental data. Data and model predictions are taken from Vollmerhausen et al. (2008)<sup>11</sup>.

## 6. DISCUSSION AND CONCLUSIONS

### 6.1.1 Analytical TOD model

We derived a very simple analytical expression (the ‘weighted linear sum model’) to estimate the VA (Visual Acuity, the high contrast cut-off of the TOD curve) of a pixel-based Electro-Optical sensor system, based on only three sensor variables: detector pitch  $d$ , optic blur  $F\lambda$ , and focal length  $f$ .

This simple expression calculates the contribution of blur and sampling on sensor performance, and enables a first order prediction of the TA range potential with a given sensor system. The next steps are to extend the equation for an entire TOD curve and to incorporate more sensor parameters. Obvious parameters are spatial and temporal noise, spectral wavelength, fill factor, display characteristics and viewing distance.

Development of such a model makes sense because the most popular range prediction models disagree on the relative importance of sampling and blur and this issue needs more study. In addition, current predictions of a TOD curve are performed with an image-based simulation of the TOD test (see section 2.2), either with a human observer or the HVS model. Though this simulation potentially yields the most accurate results (see also 6.1.2), calculations are time-consuming and not always practical.

### 6.1.2 HVS model accuracy and improvement

The results show that the Human Visual System (HVS) model used in the TOD simulation model accurately predicts human performance with well-sampled imagers but systematically underestimates TOD performance for detector-limited sensors. The difference is approximately 20%, and explains deviations shown earlier<sup>5,13</sup>.

The prediction error can be explained as follows. In order to determine its most probable orientation, the HVS model correlates the neural activity of a degraded triangle with the activities of four reference triangles. In a realistic model, the reference triangles should be degraded as well, but in the current model we use the original test patterns. With an under-sampled imager, the degraded test patterns can have many different shapes very unlike the original triangle because of the phase-dependency. Therefore, taking the highest correlation of the sensor image with reference test patterns degraded in different phases may lead to a result closer to human performance.

### 6.1.3 TOD and TTP range predictions

TA range performance was calculated for tactile vehicle ID with two under-sampled imagers, using target set specific parameters that were derived with well-sampled imagery<sup>10, 11</sup>. The results of the predictions, shown in Figure 7, are excellent, illustrating that the model (and the TOD) correctly weigh blur and sampling for these sensors. In this example, the weighted linear sum approximation of the TOD was used: final range predictions using the simulation model (2.2) need to be performed. This result supports a range of earlier TOD validation studies.

The same ID data set was used by other authors to validate different versions of the TTP metric<sup>11</sup>. The results were that the 2005 TTP metric (based on the Equivalent Blur approximation) overestimates under-sampled range performance. Predictions with the 2008 version however, are very close to the TOD calculations and to the data from the two experiments. The AAN model is the first TTP version that takes into account in-band aliasing, while the earlier versions only calculated out-of-band aliasing. TOD inherently takes into account both in-and out-of-band aliasing.

## 7. CONCLUSIONS

- A very simple analytical expression was derived to quickly calculate sensor performance with a pixel-based EO sensor system
- Tactical vehicle ID studies show that the TOD accurately predicts range performance for well-sampled and under-sampled imagers.
- The results suggest that the TOD triangle test pattern discrimination task is a good representative for a real vehicle ID task: in all validation studies performed so far, the effect of sensor differences on performance for these two tasks has been equivalent.
- The new TTP metric (the AAN model) predictions are close to those with the TOD and to the vehicle ID data. More extensive comparison studies with complex imaging systems are desirable.

## REFERENCES

1. G.C. Holst (2007). Imaging system performance based upon  $F\lambda/d$ . Opt. Eng., Vol. 46.
2. K. Krapels, R. G. Driggers, E. Jacobs, S. Burks, and S. Young (2007). Characteristics of infrared imaging systems that benefit from superresolution reconstruction. Appl. Opt. **46**, 4594-4603
3. Bijl, P. & Valetton, J.M. (1998). TOD, the alternative to MRTD and MRC. Optical Engineering **37**, 7, 1976 - 1983.
4. Bijl, P. & Valetton, J.M. (1999). Guidelines for accurate TOD measurement. SPIE Proceedings, Vol. 3701 14 - 25.
5. Hogervorst, M.A., Bijl, P. & Valetton, J.M. (2001). Capturing the sampling effects: a TOD sensor performance model. SPIE Proceedings Vol. 4372, 62-73.
6. Bijl, P. & Vries, S.C. de, (2009). Visual Acuity and Contrast Sensitivity with compressed motion video. In: Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XX, 7300 (in press).
7. Wittenstein, W. (1999). Minimum temperature difference perceived – a new approach to assess undersampled thermal imagers. Optical Engineering **38**, 5, 773 – 781.
8. Wittenstein, W., Fick, W. & Raidt, U. (1996). Range Performance of Two Staring Imagers - Presentation of the Field Trial and Data Analysis. Proc. SPIE Conf. on Infrared Imaging Systems, 2743, 132, (1996).
9. Vollmerhausen, R., & Driggers, R.G (1999). NVTherm: next generation night vision model. Proc. IRIS Passive Sensors, 1, 121-134.
10. Vollmerhausen, R., & Jacobs, E. (2004). The Targeting Task Performance (TTP) Metric A New Model for Predicting Target Acquisition Performance. Report No. A394224, CNVESD, Ft. Belvoir (VA).
11. Vollmerhausen RH, Driggers RG, and Wilson DL (2008). Predicting range performance of sampled imagers by treating aliased signal as target-dependent noise. Journal of the Optical Society of America. A, Optics, image science, and vision **25**(8):2055-65
12. Bijl, P., Hogervorst, M.A. & Valetton, J.M. (2002). TOD, NVTherm and TRM3 model calculations: a comparison. SPIE Proceedings Vol. 4719, 51-62.
13. Bijl, P. & Hogervorst, M.A.(2007). NVThermIP vs TOD: matching the Target Acquisition range criteria. SPIE Proceedings 6543, pp. 65430C.
14. De Lange, D.J., Valetton, J.M. & Bijl, P. (2000). Automatic characterization of electro-optical sensors with image-processing, using the Triangle Orientation Discrimination (TOD) method. SPIE Proceedings, Vol. 3701, 104-111.
15. Bijl, P., Hogervorst, M.A. & Vos, W. (2008). Modular Target Acquisition model & visualization tool..In: Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XIX, 6941, 69410E.
16. Wittenstein, W. (2006). pcSITOS User's Guide and model description. Report FGAM-FOM, Ettlingen, Germany.
17. Bijl, P., Schutte, K. & Hogervorst, M.A. (2006). Applicability of TOD, MRT, DMRT and MTDP for dynamic image enhancement techniques. SPIE Proceedings 6207, 154-165.
18. Beintema, J.A., Bijl, P., Hogervorst, M.A. & Dijk, J. (2008). Target Acquisition performance: effects of target aspect angle, dynamic imaging and signal processing. In: Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XIX, 6941 1 69410C.

19. Driggers et al. (2006). Current infrared target acquisition approach for military sensor design and wargaming Infrared imaging systems. *Proc. SPIE*, Vol. 6207, pp. 620709.1-620709.17

### APPENDIX 1: EXPERIMENTAL CONDITIONS AND RESULTS

Table 2: Five typical MWIR and LWIR sensors with increasing  $F\lambda/d$

Sensor	FLIR SC2000	'Typical' LWIR	TVS 8502	Indigo Phoenix	DRS heavy TWS
type	LWIR, uncooled microbolometer	LWIR	MWIR, InSb, cooled	LWIR, uncooled	LWIR, uncooled microbolometer
Primary sensor parameters					
Spectral range ( $\mu\text{m}$ )	7.5-13	8-12	3-5	8-12	8-12
Av. optical transmission	0.8	1	1	1	1
Optics blur	calculated	0	calculated	calculated	calculated
Aperture diameter (mm)	36.3	32	7.5	5.2	19.5
Focal length (mm)	36.3	32	30	13	103.3
NETD (K)	0.07	0.05	0.025	0.07	0.092
H detectors	320	1280	256	640	640
V detectors	240	720	256	512	480
H detector pitch ( $\mu\text{m}$ )	47.5	25	30	25	24
V detector pitch ( $\mu\text{m}$ )	47.5	25	30	25	24
H detector size ( $\mu\text{m}$ )	42.5	25	27	23	22
V detector size ( $\mu\text{m}$ )	42.5	25	27	23	22
Frame rate (Hz)	50	60	120	60	30
Deduced parameters					
Average $\lambda$ ( $\mu\text{m}$ )	10	10	4	10	10
f-number	1	1	4	2.5	5.3
Fill factor	0.8	1	0.81	0.85	0.84
$F\lambda/d$	0.21	0.40	0.53	1.0	2.21
HFOV ( $^\circ$ )	24	53.1	14.6	63.2	9
VFOV ( $^\circ$ )	18	31.4	14.6	52.4	6.75
IFOV (mrad)	1.31	0.78	1.0	1.92	0.23

Table 3: Results of the measurements.  $S$  is threshold triangle size @ 20K (mrad).  $VA = S^l$ .

Sensor	FLIR SC2000	Typical LWIR	TVS 8502	Indigo Phoenix	DRS heavy TWS
$S$ (mrad) with NL	$1.81 \pm 0.07$	$1.15 \pm 0.04$	$1.38 \pm 0.05$	$4.24 \pm 0.14$	$0.479 \pm 0.012$
$S$ (mrad) with JS	$2.12 \pm 0.09$	$1.21 \pm 0.04$	$1.27 \pm 0.05$	$4.74 \pm 0.18$	$0.469 \pm 0.016$
$S$ (mrad) with AW	$2.14 \pm 0.09$	$1.38 \pm 0.04$	$1.76 \pm 0.08$	$5.21 \pm 0.19$	$0.555 \pm 0.016$
$S$ (mrad) with TB	$2.08 \pm 0.06$	$1.41 \pm 0.05$	$1.62 \pm 0.05$	$4.55 \pm 0.016$	$0.500 \pm 0.013$
Weighted average $S$ (mrad) with observers NL, JS, TB	$1.99 \pm 0.08$	$1.23 \pm 0.06$	$1.42 \pm 0.08$	$4.47 \pm 0.12$	$0.483 \pm 0.008$
$S$ (mrad) with HVS model	$2.50 \pm 0.124$	$1.65 \pm 0.055$	$1.83 \pm 0.07$	$5.05 \pm 0.18$	$0.514 \pm 0.018$
$S_{observers}/\text{IFOV}$	$1.52 \pm 0.06$	$1.58 \pm 0.07$	$1.42 \pm 0.08$	$2.33 \pm 0.06$	$2.10 \pm 0.03$
$S_{HVS}/\text{IFOV}$	$1.91 \pm 0.09$	$2.12 \pm 0.07$	$1.83 \pm 0.07$	$2.63 \pm 0.09$	$2.23 \pm 0.08$
$S_{observers}/S_{HVS}$	$0.80 \pm 0.05$	$0.75 \pm 0.04$	$0.78 \pm 0.05$	$0.89 \pm 0.04$	$0.94 \pm 0.04$