Automatic characterization of electro-optical sensors with image-processing, using the Triangle Orientation Discrimination (TOD) method

Dirk-Jan J. de Lange*[a], J. Mathieu Valeton[b], Piet Bijl[b]

[a]TNO Physics and Electronics Laboratory, The Hague, The Netherlands

[b]TNO Human Factors Research Institute, Soesterberg, The Netherlands

ABSTRACT

The objective characterization of electro-optical sensors and that of image enhancement techniques has always been a difficult task. Up to now the sensor is characterized using the minimum resolvable temperature difference (MRTD) or the minimum resolvable contrast (MRC). The performance of image enhancement techniques is done by visual inspection by a human observer. Since in more and more cameras some kind of image processing is applied, a more elaborate test is needed that can measure the performance of the combination of the sensor and the image processing. A good candidate is the TOD (Triangle Orientation Discrimination) method, which is developed as an alternative for MRTD and the MRC methods. We are investigating how the standard TOD-method can be extended to cameras with image processing and whether these measurements can be automated.

An algorithm is under development, which is based on the TOD-method and predicts the characterization by human observers of camera-system performances. The algorithm combines the TOD-method, an early-vision model, and an orientation discriminator. The algorithm uses the same images as used in human-observer experiments. After correction for the physical properties of the display and the human eye, the algorithm tries to find the orientation of the stimulus. The algorithm can also predict the performance of only image processing using a simple scene-generator in stead of a camera setup.

Keywords: automatic characterization, sensor performance, image enhancement performance, early-vision, ideal observer, TOD, MRTD.

1. INTRODUCTION

Objective quantification of the effects of image enhancement techniques is still very difficult. Since there is not a single parameter curve, like an MRTD or MRC curve, to describe the performance of a sensor system with some kind of image processing, it is hard to compare cameras and to compare image enhancement techniques. In a collaboration between the TNO Human Factors Research Institute and the TNO Physics and Electronics Laboratory a research program is started to develop a test to characterize the performance of a camera with some kind of image processing. The performance of a camera depends on several factors. First of all the characteristics of the camera itself are important, but the performance also depends on the observation task and the scenario. We limited ourselves to the detection task and a scenario with static objects in a high clutter background. The background will contain both dim and bright areas. It is important to test all the situations in which image processing might introduce artifacts. In the research program we are investigating how the standard TOD-method using human observers can be extended to cameras with image processing and whether these measurements can be automated. This paper describes the algorithm that will be used to predict the characterization by human observers of camera-system performances. The algorithm uses the same images as used in human-observer experiments. After correction for the physical properties of the display and the human eye, the algorithm tries to find the orientation of the stimulus. It is also possible to use this algorithm for the performance prediction of image processing itself by using a simple scene-generator in stead of a camera setup.

* Correspondence: Email: deLange@fel.tno.nl;

In *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XI*, Gerald C. Holst, Editor, Proceedings of SPIE Vol. 4030 (2000) • 0277-786X/00/$15.00

2. TOD-METHOD

The Triangle Orientation Discrimination method (TOD) for the characterization of cameras has been developed by TNO Human Factors Research Institute[1,2] as an alternative for MTRD and MRC methods. In the method the triangular shaped stimuli are recorded by a camera and displayed on a monitor. The task of the observer consists of indicating the orientation of the presented triangle. There are only 4 choices: i.e. up, down, left or right. If the observer is not able to determine the orientation he has to make a guess. The forced choice used in this method makes it relative easy to implement the TOD test in a computer algorithm.

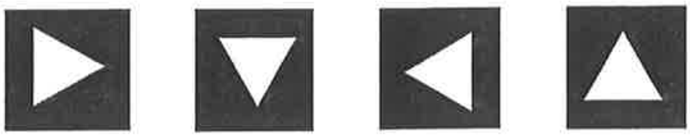

Figure 1: The 4 different stimuli used in the TOD method, i.e. right, down, left and up orientation.

3. GENERAL LAYOUT

The algorithm presented in this paper, uses real images as input. In every step of the algorithm the image is kept in the spatial domain and is not transformed to the spatial-frequency domain, since most image enhancement techniques work in the spatial domain and therefore are difficult to describe in the spatial-frequency domain. Most of the existing models, which predict the performance of a sensor system, use only spatial-frequencies and are thus not suitable for the characterization of cameras with image enhancement techniques.

The purpose of the algorithm is to characterize automatically a camera system with any form of image processing. The camera is placed in a test setup and the stimuli are presented to the camera. The data from the camera is digitized and used as input for the algorithm.
In the algorithm the whole process of an image being displayed on a monitor and observed by the human eye is modeled. The left panel in figure 2 shows the flow diagram of the algorithm. In the Display Model the effects of a TV-monitor are simulated. The model includes the non-linear response and the limited dynamical range of the display. The Early Vision Model calculates how the image from the display is projected on the retina of the human eye. In the next step the projected image on the retina is presented to an orientation discriminator, which tries to determine the orientation of the triangle using correlation techniques. The result of the orientation discriminator is compared to the real orientation of the presented stimulus and in the statistics block the scores of the orientation discriminator are stored as a function of the contrast, size and orientation of the stimulus. If all the different triangles are processed, the algorithm determines the 75%-correct contrast level as a function of triangle size.

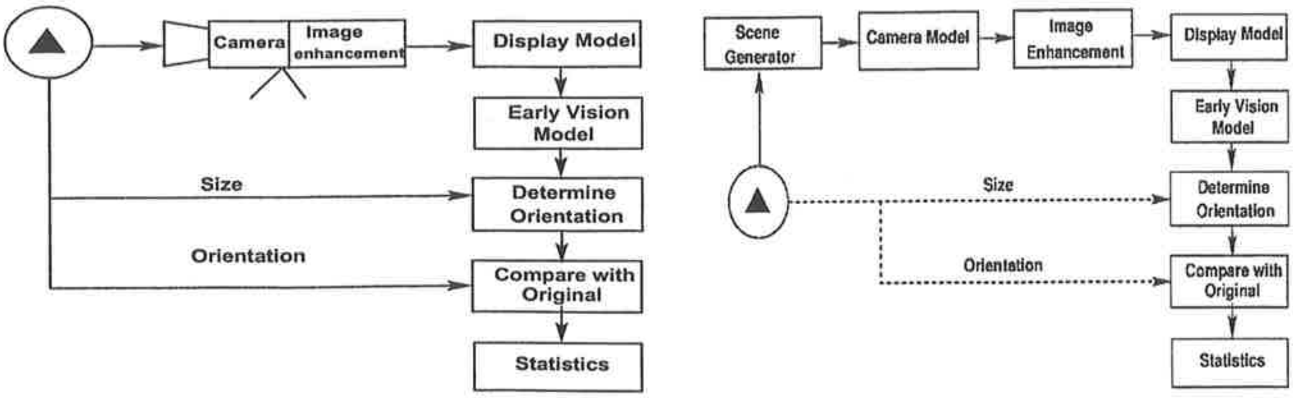

Figure 2: The left panel shows the flow diagram for the evaluation of real camera images, the right panel shows the flow-diagram for the evaluation of only the image enhancement techniques using generated images.

It is also possible to apply the algorithm to generated scenes in stead of real camera recordings, using the flow diagram shown in right panel of figure 2. In this setup it is possible to test the effect of only the image enhancement, which is very useful in the development and tuning of image enhancement techniques. The generated scenes will be digital replicas of the test-cards used in the human observer experiments. Until these test-cards are finalized some simple scenes are used for the tests of the algorithm. The effects of a camera are simulated in the camera model, which can be very simple if the aim of the simulation is to investigate only the image enhancement technique. Therefore the camera model incorporates only the instantaneous-field-of-view (IFOV) of the camera, gain, offset and Gaussian noise.

4. EARLY VISION MODEL

The Early Vision Model calculates how an image is projected on the retina of a human eye. In the model the physical properties of the eye are included, like the dispersion of the light, absorption of the light in the eye and the size of the cones on the retina. The model calculates the mean number of photons absorbed in each cone. The presented model is based on the Early Vision model presented by Geisler and Davila [3]. In this model the following assumptions are made:

- The observer knows where to expect the stimulus. Fovea observation.
- The eye does not move during each observation
- The is no defocusing of the eye during an observation
- Contributions of the cones dominate. Photopic view.
- Gray-value images only. No wavelength dependency.
- Only Poisson noise.

Figure 3 shows the flow diagram of the Early Vision model. The model has the display image as input.

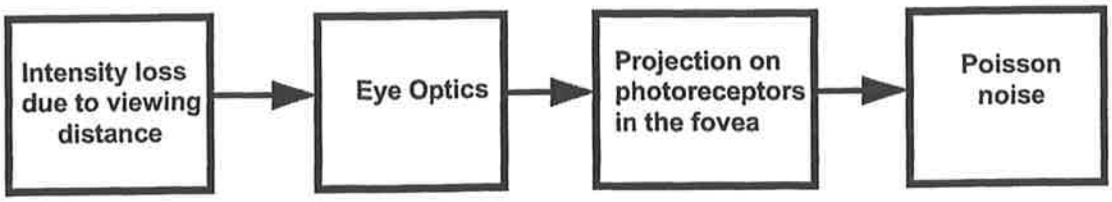

Figure 3: Flow diagram of the Early Vision model.

In the first step the intensity of the light from the monitor is corrected for the distance between the observer and the monitor. In the second step the effect of the optics of the eye are calculated. In the last steps the blurred image is projected on the fovea of the retina and the response of each cone is calculated.

In the eye optics block two effects are simulated. First of all, only a part of the light that falls on the pupil is transmitted. Figure 4a shows the transmittance of the lens as a function of the wavelength[4]. Since there is no wavelength dependency is the model, an average transmittance of 42% is used. After correction for the transmittance the image is blurred with the point spread function (PSF) of the eye optics. The PSF used in this model is derived from the line-spread function presented by Campbel and Gubisch [5] for white light and a pupil diameter of 3 mm. This PSF does not include the effect of diffusely scattered light. The PSF, which is shown in figure 4b, is described by the sum of 2 Gaussian distributions:

$$h(x,y) = ah_1(x,y) + (1-a)h_2(x,y), \tag{1}$$

where $a=0.583$ is a mixing parameter, and $\sigma_1 = 0.443$ arc min and $\sigma_2 = 2.04$ arc min the standard deviations of the Gaussian distributions $h_i(x,y)$.

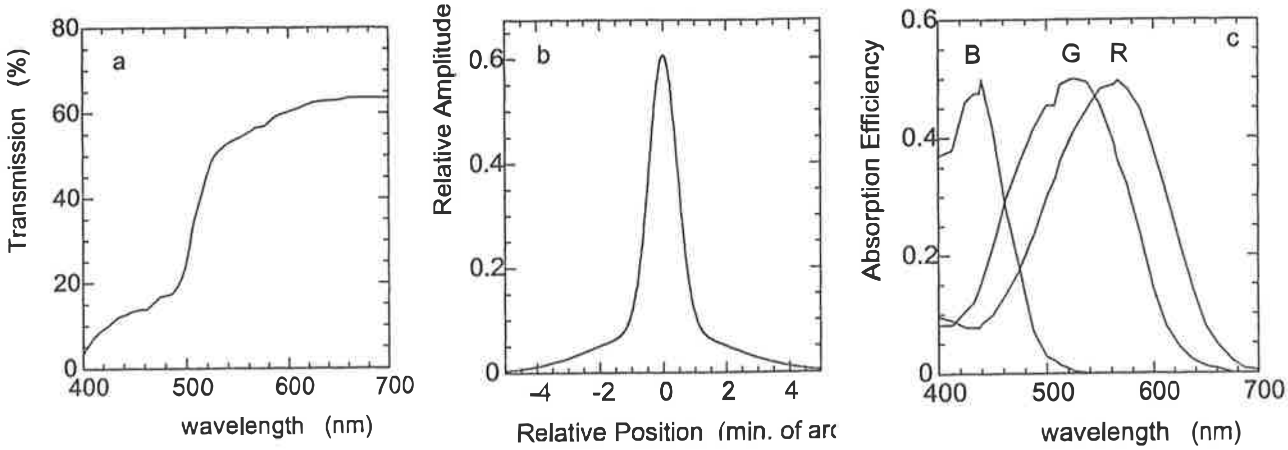

Figure 4: a) Transmission of the eye [4],
b) PSF for white light and a pupil diameter of 3 mm [5],
c) Absorption efficiency of the cones [6].

After correction for the PSF of the lens the image is projected on the retina. Since it is assumed that the observer knows roughly where to expect the next stimulus we only consider here projection on the photoreceptors in the fovea. Geisler and Davila assumed that the photoreceptors where tightly packed on a triangular grid and that the cross-section of the photoreceptors has a circular shape. Figure 5a shows schematically the triangular grid for the photoreceptors used by Geisler and Davila. The lines indicate the $1\sigma_l$, $2\sigma_l$ and $3\sigma_l$ areas of the PSF. A triangular grid with circular cells is very difficult to implement in a computer algorithm. Therefore a square grid with square cells is used in the algorithm, as is shown in figure 5b. The precise shape of the cells is not very important since the standard deviation σ_l of the PSF is 1.5 times larger than the radius of the a cone, which implies that a point light-source is always projected on several cones. The grid size of the square grid is chosen such that the cone density is the same as for the triangular grid.
There are three kinds of cones, which are sensitive to respectively red, blue and green light. In the model, it is assumed that there are no blue-sensitive cones in the fovea and twice as many red-sensitive cones as green-sensitive cones. In figure 4c the spectral sensitivity of each type of cone is shown [6]. Since there is no wavelength dependency is this model we assume an average absorption efficiency of 2 red-sensitive cones and 1 green-sensitive cone, which amounts to 23 %. After correction for the absorption efficiency of each cone Poisson noise is added.

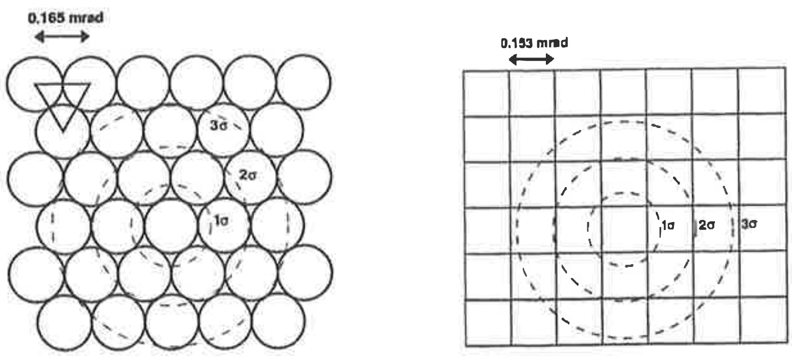

Figure 5: Triangular cone grid used by Geisler and Davila and square cone grid used in the algorithm described in this paper. The lines indicate the $1\sigma_l$, $2\sigma_l$ *and* $3\sigma_l$ *areas of the PSF. Both grids have the same number of cones per area.*

5. ORIENTATION DISCRMNATOR

The orientation discriminator tries to determine the orientation of the triangle from the image on the fovea. The orientation is determined by correlating the image which 4 ideal triangles, as shown in figure 1. The sizes of the ideal triangles are adjusted to match the size of the triangle in the fovea image. The orientation that yields the highest correlation is chosen to be the correct one.

6. STATISTICS

After the determination of the orientations of all the different stimuli, the probability correct is calculated as a function of triangle contrast and size. The binomial probability function gives the probability for ***x*** successes at ***n*** independent determinations with a known chance ***p***. In this experiment we want to determine ***p*** from the measured ***x*** and ***n***. For this we have to use a kind of inverse binomial calculation. The exact formula for the expectation value of ***p*** is given by:

$$\langle p \rangle = (n+1)\int_{p=0}^{1} p\,P_{(x,n)}(p)dp = (n+1)\binom{n}{x}\frac{(x+1)!(n-x)!}{(n+2)!} = \frac{x+1}{n+2}, \tag{2}$$

where $P_{(x,n)}(p)$ is the binomial probability distribution. The variance in ***p*** is given by:

$$\sigma_p^2 = (n+1)\binom{n}{x}\frac{(x+2)!(n-x)!}{(n+3)!} - \langle p \rangle^2 = \frac{(x+1)(n-x+1)}{(n+2)^2(n+3)}. \tag{3}$$

In order to optimize the algorithm the number of measurements ***n*** is kept as low as possible. Therefore, it is essential that the exact formulas, given above, are used and not their well-known approximations for large values of ***n***:

$$\langle p \rangle = \frac{x}{n}, \qquad n >> 1 \tag{4}$$

$$\sigma_p^2 = \frac{\langle p \rangle (1 - \langle p \rangle + 1/n)}{n}. \qquad n >> 1 \tag{5}$$

After the calculation of ***p*** the contrast is determined at which 75% of the orientations were determined correctly. Figure 6 shows an example of the probability correct determined by the algorithm as a function of the stimulus contrast.

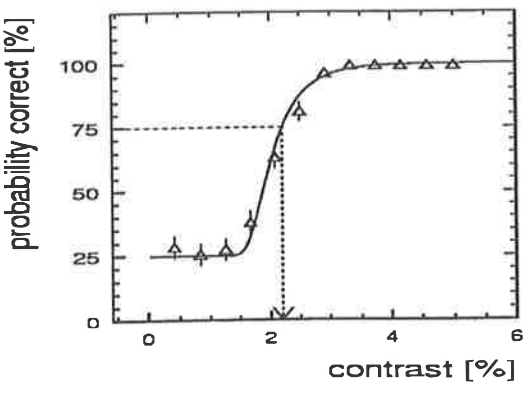

Figure 6: Fraction correct orientation discrimination as a function of the stimulus contrast, as determined by the algorithm. The solid lines represent the fit of a Weibull function to the data. The threshold contrast is defines as the 75% correct point of the Weibull function (about 2.2% in this example).

The line indicates a fit to the data using a Weibull function:

$$p = 0.25 + 0.75 / 1.5^{(C_{75}/C)^{\beta}},\tag{6}$$

where C is the contrast, C_{75} the 75% contrast threshold, and β a parameter that determines the steepness of the curve. The 75 % correct contrast threshold will be determined for different stimulus sizes.

7. RESULTS

As a first test the results of the algorithm were compared to a simple human observer experiment. The experiment, which was performed at the TNO Human Factors Research Institute, involved 1 single observer looking at generated triangles on a computer display. The luminance background level was 22 cd/m^2 and the distance between the observer and the display was 8 m. The duration time of the triangular stimuli was 0.53 seconds. The aim of the experiment was to test the vision model used in the algorithm, and not to test a camera or any image processing. Figure 7 shows the 75 % correct contrast thresholds, denoted by the triangles, versus the stimulus size in arc min. The size of the triangle is defined as the square root of the triangle area. The results of the simulation results are depicted by the crosses after they were multiplied by a factor 9 to match the experimental results. The algorithm is thus able to predict the correct shape of the curve but overestimated the performance of this specific human observer, indicating that the noise in the eye or in the brain was underestimated. The only noise present in the model in Poisson noise in the photon absorption process in the photoreceptors. It is very likely that there are several sources for noise in the eye and the human brain, but since there are no quantitative numbers for these noises they were not explicitly implemented in the model. In conclusion, the relative dependence of the contrast threshold on the stimulus size is in this experiment predicted correctly by the simulation algorithm. A single scaling factor was needed to reproduce the results of the human observer experiment.

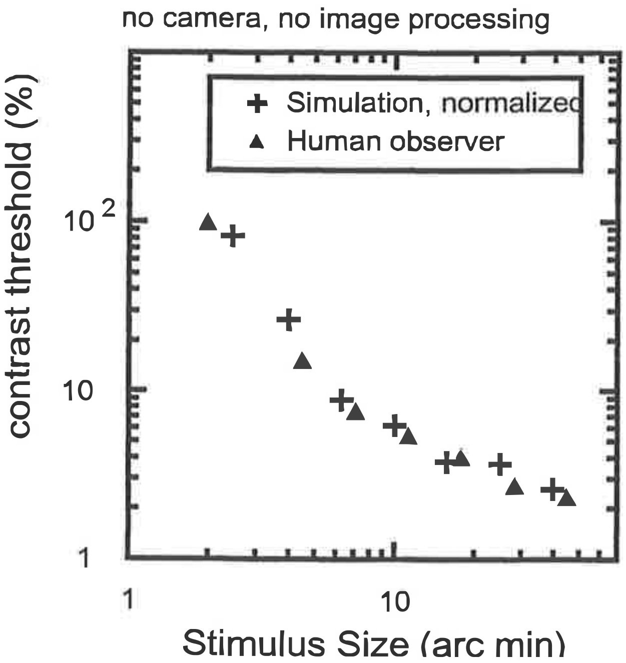

Figure 7: The calculated TOD 75% correct contrast-thresholds as a function of the stimulus size for an observer looking at stimuli at a display 8 meters away at a luminance background level of 22 cd/m^2. The stimulus size is defined as the square root of the stimulus area. The triangles denote the measured threshold results for a human observer, and the crosses denoted the simulation predictions. The simulation results were multiplied by a factor 9.

As a second test of the algorithm the TOD contrast thresholds were calculated for a different implementations of histogram equalization. Figure 8 shows an example of the simple generated scene used in this test and the result after global histogram equalization.

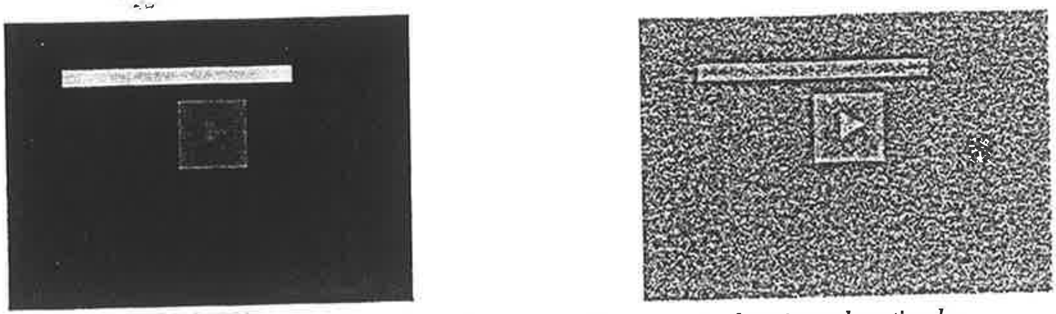

Figure 8: The left panel shows an example of a generated scene with the triangular stimulus. The right panel shows the result after global histogram equalization.

A histogram equalization algorithm calculates the statistics of the image intensities and redistributes the intensities in the images so that the full intensity range is used. The effect of histogram equalization depends on the stimulus size with respect to the image size. If the stimulus is large enough the stimulus becomes better visible, but if the stimulus becomes to small the intensities of the stimulus may become statistically less significant than the background noise. The histogram equalization starts to amplify the noise, which makes it difficult to detect the stimulus. Figure 9 shows the TOD contrast thresholds calculated by the algorithm as a function of the stimulus area. The diamonds show the result if no image enhancement is applied, and the triangles show the result for global histogram equalization using a small background image. For stimulus sizes down to 2.5 pixels the thresholds for histogram equalization are smaller than without image enhancement, indicating that the histogram equalization improved the detection of the stimulus. At smaller sizes the histogram equalization results in higher thresholds, indicating that the detection of the stimulus was made more difficult.

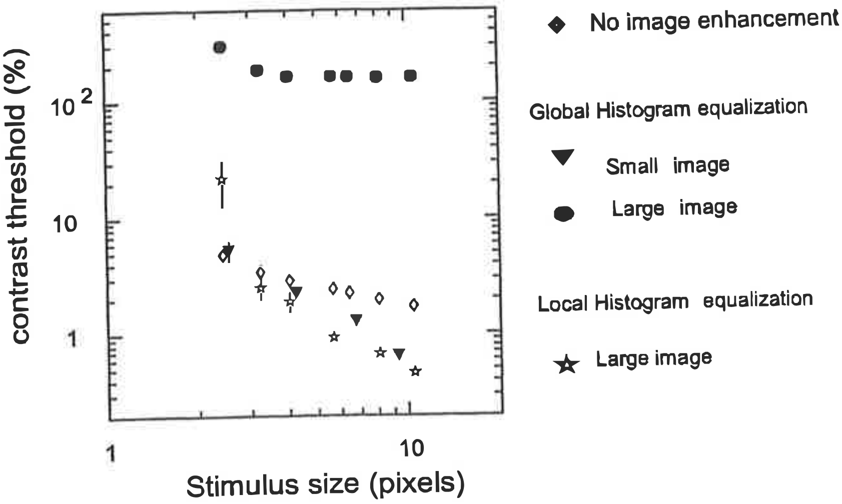

Figure 9: The calculated TOD contrast thresholds as a function of the stimulus size for different implementations of histogram equalizations. The algorithm used a scene generator and a simple camera model.

This becomes even more apparent if a much larger background image is used, denoted by the circles. In this case the global histogram equalization made the detection of the stimulus almost impossible. For large background images the use of local histogram equalization is therefore preferred, which performs its calculation on segments of the image instead of the whole image. The results for local histogram equalization, denoted by the stars, are comparable to the results for global histogram equalization on small images. The simulated results for the different histogram equalization algorithms behave as could be expected based on the nature of the algorithms. This suggests that the total-simulation algorithm predicts the relative difference in behavior of these simple image enhancements correctly. More tests and comparisons with human observer experiments are needed before any statement can be made about the absolute values of the predictions.

CONCLUSIONS

An algorithm is presented, which is able to predict the characterization by human observers of the performances of cameras, which have a kind of image processing incorporated. The algorithm combines the TOD-method, an early-vision model, and an orientation discriminator. The algorithm uses the same images as used in human-observer experiments. The algorithm can also be used to predict the performance of only image enhancement techniques.

The first results indicate that the algorithm is able to reproduce the results of human observer experiments. Furthermore, the algorithm is able to correctly predict the relative difference in behavior of simple image enhancement techniques.

At this point the algorithm can be used to indicate whether, for detection tasks, one image enhancement technique is better than another. The next step in the development of the algorithm is to try to quantify the difference in performance and to design more representative backgrounds. More comparisons with human observer experiments are foreseen in the near future.

ACKNOWLEDGEMENTS

This research was performed under contract with the Royal Netherlands Army.

REFERENCES

[1] P. Bijl en J.M. Valeton, 'Triangle orientation discrimination: the alternative to minimum resolvable temperature difference and minimum resolvable contrast', *Opt. Eng.* **37**, 1976 (1998).

[2] P. Bijl en J.M. Valeton, 'Validation of the new triangle orientation discrimination method and ACQUIRE model predictions using observer performance data for ship targets', *Opt. Eng.* **37,** 1984 (1998).

[3] W.S. Geisler en K.D. Davila, 'Ideal discriminators in spatial vision: two-point stimuli', *Opt. Soc. Am.* A **2**, 1483 (1985).

[4] G. Wyszecki en W.S. Stiles, 'Color Science', New York: John Wiley & Sons (1982).

[5] F.W. Campbell en R.W. Gubisch, "Optical quality of the human eye", *J. Physiology* (London) **186**, 558 (1966).

[6] O. Estévez, "On the fundamental data-base of normal and dichromatic color vision", Ph.D. thesis, University of Amsterdam, the Netherlands 1979.