

Papers that [reference](#) this paper.

To be published in B. E. Rogowitz and T. N. Pappas, eds., *Human Vision and Electronic Imaging IV*, SPIE Proceedings Vol. 3644, paper 8, 1999.

Nonlinear features in vernier acuity

Erhardt Barth *, Bettina L. Beard, and Albert J. Ahumada, Jr.

NASA Ames Research Center, Moffett Field, CA 94035-1000

*At present: Institute for Signal Processing, Medical University of Lübeck, Germany

Abstract

Nonlinear contributions to pattern classification by humans are analyzed by using previously obtained data on discrimination between aligned lines and offset lines. We show that the optimal linear model (which had been identified by correlating the noise added to the presented patterns with the observer's response) can be rejected even when the parameters of the model are estimated individually for each observer. We use a new measure of agreement to reject the linear model and to test simple nonlinear operators. The first nonlinearity is position uncertainty. The linear kernels are shrunk to different extents and convolved with the input images. A Gaussian window weights the results of the convolutions and the maximum in that window is selected as the internal variable. The size of the window is chosen such as to maintain a constant total amount of spatial filtering, i.e. the smaller kernels have a larger position uncertainty. The results of two observers indicate that the best agreement is obtained at a moderate degree of position uncertainty, about plus or minus one arc min. Finally, we analyze the effect of orientation uncertainty and show that agreement can be further improved in some cases.

Keywords: spatial vision, nonlinear operators, pattern recognition, vernier acuity, visual noise, position uncertainty, nonlinear system identification.

1. Introduction

Human sensory processing is often studied using psychophysical experiments measuring the discriminability between two stimuli. The observer essentially acts like a function that assigns a categorization probability to each external stimulus that might be presented by the experimenter. At last year's meeting Beard and Ahumada described a technique generating an image that represents the observer's stimulus classification rule when that rule is linear in the stimulus image pixels. The technique involves adding white noise to the stimulus image pixels and computing the correlation of the noise amplitude in each stimulus pixel with the observer responses. The sum of the two correlation images for each of the two stimuli gives an estimate of the classification image. If the observer's rule is linear, it is equivalent to cross correlating the classification image with the given stimulus image and assigning a monotonic response probability to each value. The probabilities represent the cumulative distribution of the observer's internal noise. Beard and Ahumada estimated the observer classification image and showed that it was not compatible with some simple models of vernier acuity. Here we present a method for evaluating the adequacy of any computational model of observer performance that is image based. The method is then used to evaluate some linear and nonlinear models for observer performance in the vernier task. Our method rejects a best-fitting linear model for two of three observers, but fails to reject simple nonlinear models.

2. EXPERIMENTAL DATA

The data we will consider were previously obtained and evaluated by two of us. In those experiments, two different patterns were presented: two horizontal lines, both 16 pixels by 3 pixels, that were aligned in one case and offset by one pixel in the other case. Observers, whose acuity was verified to be 20/20 or better, had to decide whether the lines were offset or not. A pixel subtended 0.31 arc min at the observer's eye. To these line patterns, different white-noise images were added at each trial. The noise was of low contrast such as to minimize its effect on the decision rule but strong enough to influence the decision. A total of 11400 trials were presented to three observers. Observer PW ran all trials with a peak noise contrast of 0.25. To increase the number of error trials, observer DF's noise level was changed to 0.35 and observer BLB's level was changed to 0.4 part way through the experiment. The different noise images were sorted by stimulus and observer response into the categories S0R0, S0R1, S1R0, and S1R1, depending on the stimulus (S0 meaning no offset, S1 meaning offset) and the response (R0 meaning that the observer responded "no offset", and R1 meaning "offset" response). The average noise image was then computed with the following signs: $-S0R0 + S0R1 - S1R0 + S1R1$ to obtain a classification image. The values at different pixels in the classification images indicate how well the image-intensity values at a given pixel (and different trials) correlate with the observer's responses (at those trials). The classification images are shown in Figure 1 (top) for three observers. It is difficult to see the form of the images in the noise.

Figure 2

shows a classification image where the noise is reduced three ways: the data for the three observers is combined, the image has been spatially smoothed, and it has been quantized so that pixels not at the central gray level represent correlations that are significantly different from zero.

The main result was that the influence of the noise is indeed significant such that certain models of vernier acuity can be excluded. If the observer were ideal, except for optical blur for example, the resulting classification image would have had only the side of the stimulus that changed when the line was offset.

The simplest assumption would be that observers weighted the intensity values at each pixel with a kernel, which is similar to the classification images, i.e. they based their decision on linear features. A statistical analysis, however, revealed a lack of fit between the optimal linear model and the observer performance, therefore suggesting that nonlinear features are important. The importance of nonlinearities for vision models had been emphasized repeatedly .

3. The Linear Model

The classification images as such fully specify the linear part of the model - see above and . In this section we describe the classification images by a different mathematical function and show how the parameters of that function vary for different observers. The chosen function is

(1)

$$L(x,y) = \frac{e^{xy}}{s_x s_y} e^{-\frac{x^2}{2s_x^2}} e^{-\frac{y^2}{2s_y^2}}$$

i.e., a Gaussian differentiated by x and y and normalized to unit maximum. The reasons for our choice are simplicity, good fit, and some considerations by others based on neurophysiological results and geometrical intuition . The parameters of interest are s_x and s_y . The values, which optimally fit the classification images, are given in Table 1 in pixels and for different observers. The goodness of fit also varies with observers and is characterized by the ratio of the peak kernel value (signal) to the root mean squared error (noise) given in decibels in the last row of Table 1. Observer BLB has the largest kernel size parameters and DF has the best fit. The kernels are shown as gray-level images in Figure 1, in an arrangement that corresponds to the one in Table 1.

Table 1: Spatial parameters of the optimal kernels.

observer	BLB	DF	PW
s_x , arc min	3.1	3.0	2.7
s_y , arc min	1.5	1.4	1.1
peak kernel to rms error ratio, dB	16.3	16.4	15.8

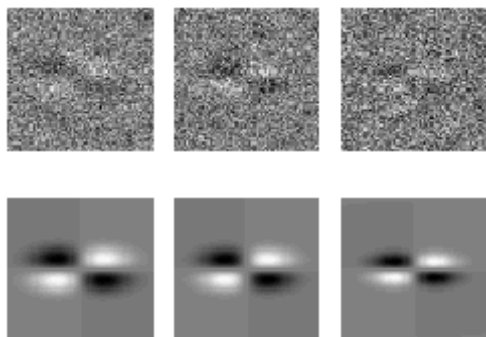


Figure 1: Classification images (top row) and estimated kernels (bottom row) for observers BLB, DF, and PW (from left to right). The images are 64 by 64 pixel large centers of the input images.

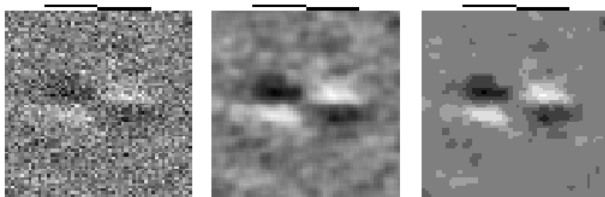


Figure 2: Classification image for all observers (left), same image after spatial smoothing (middle), and significant correlations (right). See text, and for details.

4. Observer Performance

The raw statistical data are obtained by first counting the number of cases in which no signal was presented (aligned pattern S0) and the observer responded "no signal" (R0). This number is denoted by N_{S0R0} . By analogy we obtain the numbers N_{S0R1} , N_{S1R0} , and N_{S1R1} by denoting the case of offset pattern with S1 (vernier offset), and the corresponding response with R1. From these numbers the observer d' 's are obtained by using the inverse of the standard normal cumulative distribution function F^{-1} .

(2)

$$d' = F^{-1}\left[\frac{N_{S1R1}}{N_{S1R0} + N_{S1R1}}\right] - F^{-1}\left[\frac{N_{S0R1}}{N_{S0R0} + N_{S0R1}}\right]$$

$F^{-1}[q]$ was implemented in *Mathematica* as `Quantile[NormalDistribution[0,1], q]`. In Figure 3 the d' 's for the three different observers are plotted as a function of (vertical) optimal-kernel size illustrating that the observer(s) with a higher-resolution kernel had better performance. The images presented to observers DF and BLB had larger noise levels, and these observers performed worse. More interestingly, kernel size seems to adapt to the larger noise levels.

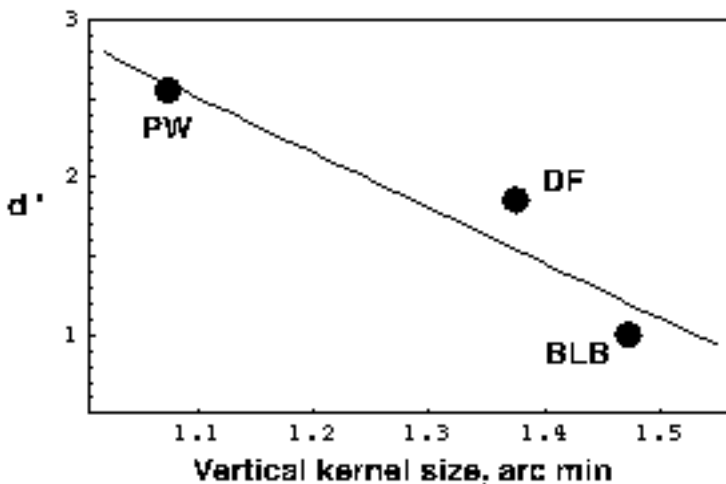


Figure 3: Observer d' plotted versus optimal-kernel size (s_y in arc min). The three data points and a linear fit are shown.

5. Measure of Model-Observer Agreement

Based on a measure of agreement described in it had been concluded that the linear model could not account for all of the present data. In order to identify nonlinear effects, a more sensitive measure of model-observer agreement is desirable. The measure described below was designed such as to make maximal use of the available data. It is based on the four distributions S0R0, S0R1, S1R0, and S1R1 shown in Figure 4. The first step is to compute the features for every noise image as described in Section 6. The (hypothetical) distributions shown on the left are then obtained by sorting the features according to the observer's response. If the input (vernier pattern plus noise) has no offset (S0) and the observer responded "no offset" (R0) for that particular noise image, the number (feature) is attributed to the class S0R0. The three other classes are obtained by analogy.

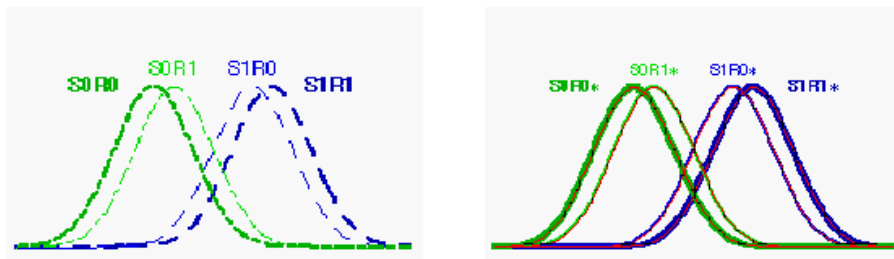


Figure 4: Schematic distributions of features sorted by observer’s response (left) and distributions of features assuming equivalent internal noise (right).

The distributions schematically shown on the right of Figure 4 (S0R0*...) are obtained by sorting the same numbers (features) in a different way. The procedure is illustrated in Figure 5.

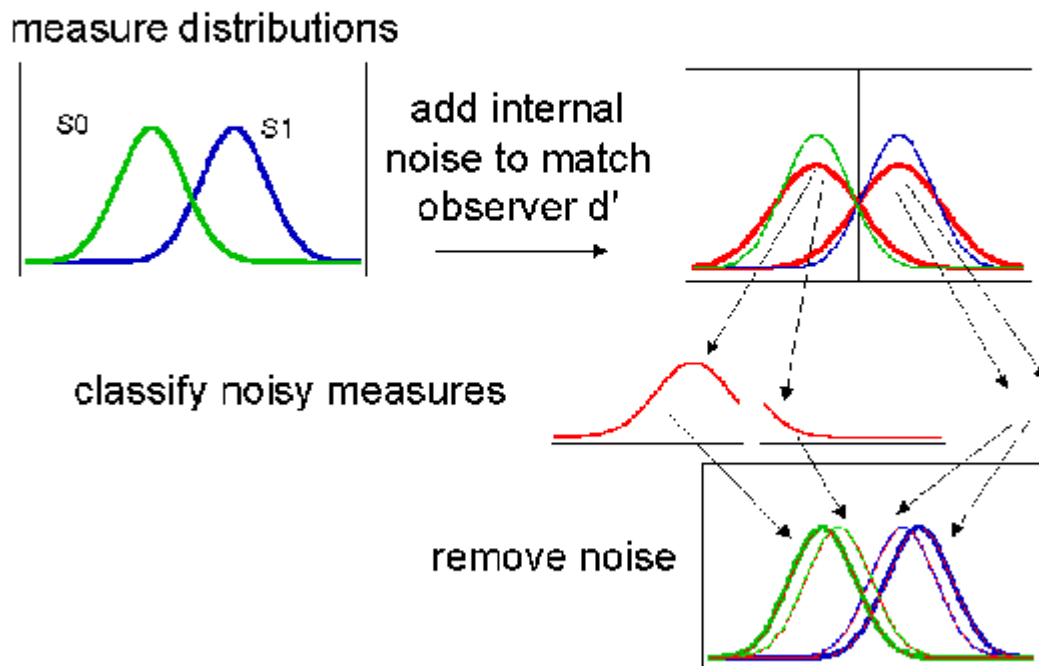


Figure 5: Sorting of features by model response.

The features are first sorted to S0 and S1. It is then assumed that the observer decisions are based on the features plus some internal noise, i.e., the distributions S0 and S1 plus the internal noise. The variance of the internal noise is estimated as

(3)

$$\sigma_N^2 = \frac{\sigma_{S0}^2 + \sigma_{S1}^2}{2} \left(\left(\frac{d'_M}{d'_O} \right)^2 - 1 \right)$$

The observer d' (d'_O) is defined by Equation (2). The model d' (d'_M) is computed directly from the distributions S0 and S1 as the

$$d'_M = \frac{E[S1] - E[S0]}{\sqrt{(\sigma_{S0}^2 + \sigma_{S1}^2)/2}}$$

normalized difference between the means, i.e., $d'_M = \frac{E[S1] - E[S0]}{\sqrt{(\sigma_{S0}^2 + \sigma_{S1}^2)/2}}$. The internal noise, with Gaussian distribution and variance defined by Equation (3), is then added to the model outputs. In the next step, the distributions S0 and S1 are split into S0R0*, S0R1*, S1R0*, and S1R1* by using the features plus noise. The important point here is that the original features are sorted by the magnitudes of the features plus noise. Thus the distributions S0R0, S0R1, S1R0, S1R1 are based on the features sorted by

observer's response. The distributions $S0R0^*$, $S0R1^*$, $S1R0^*$, and $S1R1^*$ are based on the same features but sorted by the "model's response", i.e., by the features plus noise. Finally, the measures for model-observer agreement for the two cases (no offset $S0$ and vernier offset $S1$) are obtained as

$$A_0 = \frac{E[S0R0] - E[S0R1]}{E[S0R0^*] - E[S0R1^*]} ; \quad A_1 = \frac{E[S1R0] - E[S1R1]}{E[S1R0^*] - E[S1R1^*]} . \quad (4)$$

In case of perfect agreement, these measures would be equal to one. Given a set of data, different model elements will influence the measure. The main parameter is the separation (among $S0R0$ and $S0R1$ for the case $S0$, and $S1R0$ and $S1R1$ for case $S1$) that can be achieved with given features, i.e., the nominators in Equation (4). This separation, which is specific to the features, is then related to observer performance by the amount of internal noise. With increasing observer performance (higher d'_O that leads to less internal noise) the agreement decreases because, based on Equation (3), less noise has to be added to the features and the separation of the $*$ distributions increases. In other words, for a given separation between, say, $S0R0$ and $S0R1$ the value of A_0 raises with higher internal noise, i.e., with lower d'_O . Therefore, if the measures defined by Equation (4) are larger than one, the possibility remains that the observers performed worse than they could have for some other reason. If the measure is significantly less than one, the hypothesis that decisions have been based on the features plus some internal noise can be rejected. In this case we would have to increase the amount of internal (model) noise for better agreement, but then, the model d' (now including the internal noise) would be lower than the observer d' . In other words, it is easier to see how performance could be decreased by something not related to the computations involved (features plus internal noise) than it is to see performance increased by such factors, i.e., there is an asymmetry between agreement values below and above one.

Nevertheless, we prefer to observe how the measures of agreement vary with different parameters and different features instead of relying only on the absolute values. This will make certain models more likely than others. To obtain the results shown in the next section, the measures A_0 and A_1 have been computed 200 times with different internal noises and the means are plotted with error bars (plus and minus 1.96 times the square root of the variance).

6. Feature Computation and Position Uncertainty

The input images to the model consist of aligned and offset vernier patterns plus noise, i.e., precisely the images presented to the observers. The results on position uncertainty have been obtained by first filtering the images with one of the kernels shown in Figure 1 depending on the observer to be modeled. The filters were implemented by using the Fast Fourier Transform and the Fourier Transform of the kernel in Equation (1). Then, the results of the filtering in the centers of the images (16x16 pixel) have been multiplied with a Gaussian window. The width of the Gaussian had four different values of 0, 1, 2, and 3 pixels, zero width meaning that only the center pixel was different from zero. After the Gaussian weighting the maximum in the 16x16 window was taken as the feature. This algorithm was chosen as a straightforward model of position uncertainty. For zero width, the procedure is equivalent to taking the scalar product of the input image and the kernel image (the linear model). Depending on the width of the Gaussian window, the filter parameters were estimated such that the result of successive filtering with the kernel (differentiated Gaussian) and the window (Gaussian) would remain constant. As a consequence, for an increasing width of the window, the images were filtered with kernels of decreasing spatial extent. We distinguish two cases. If the window is isotropic (isotropic uncertainty) with increasing diameter of the Gaussian, then the filter kernels will shrink and change their shape (the ratio of height to width). If the kernels are required to keep their shape (and only shrink) then the window must be anisotropic (anisotropic uncertainty). These considerations are simply a consequence of a well-known theorem due to Pythagoras.

7. Orientation Uncertainty

The next nonlinearity we investigated was orientation uncertainty (in addition to position uncertainty). The linear kernels shown in Figure 1 have zero output to a straight, horizontal line, but respond if the straight line is tilted (or curved, or discontinuous). Suppose that the subjects base their decisions on mechanisms that detect the discontinuity introduced by the vernier offset regardless of a small amount of tilt. To make a discontinuity detector that is uncertain about orientation, some type of orientation-sensitive nonlinearity is needed. A possible implementation of a discontinuity detector is to select the minimum response from kernels (1) differing in orientation but preliminary results show that this method does not increase agreement. The results shown in the next section have been obtained by selecting the maximum response as described below. We still use the term "orientation uncertainty" although it refers to the kernel domain.

We have computed the features for seven orientations as shown in Figure 6 (top). As with position uncertainty the results of the filtering were multiplied by a Gaussian window which is now localized in space and orientation as shown in Figure 6 (bottom). The maximum value is then taken as the feature. The kernels had spatial parameters, which are optimal with respect to position uncertainty - see Figure 8.

The results are plotted for constant position uncertainty and as a function of orientation uncertainty (sigma of the Gaussian in deg). The different amounts of orientation uncertainty have been implemented by changing the orientation increment and keeping the seven weights constant (0.32, 0.60, 0.88, 1.0, 0.88, 0.60, 0.32). Since the range of orientations is small, the shape of the kernels has not been changed.

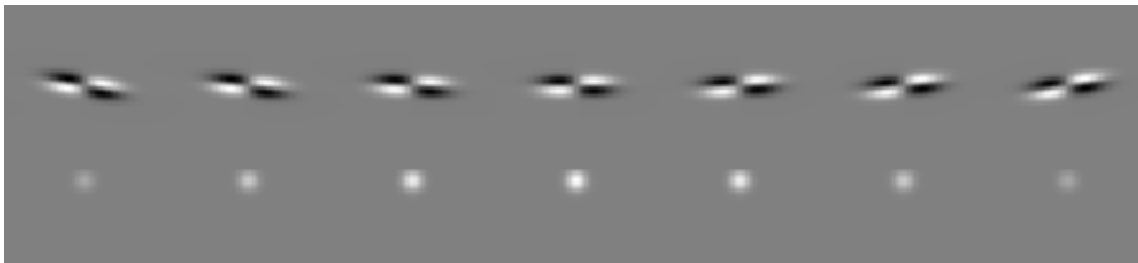


Figure 6: Optimal kernels with seven different orientations (top) and Gaussian windows used to model position and orientation uncertainty. For better illustration an orientation range of plus-minus 6 degrees is shown, which is 3 times the range found to be optimal for observer PW in S0 - see Figure 10.

8. Results

The linear model

Using the kernels shown in Figure 1 (which are optimized for each observer) the first question we address is, in which cases we can reject the linear model (defined by these kernels) based on the measures in Equation (4). As mentioned before, the measures are computed 200 times. For the observer DF the measure A_1 was less than one for all 200 estimations. For observer PW, A_0 was less than one in all cases and A_1 was larger than one in one case (0.5 %). Therefore it is likely that these two observers use some other measure. This is in accordance with previous results.

Isotropic uncertainty

Position uncertainty was modeled as described in Section 6 and the results are shown in Figure 7 for the three observers BLB, DF, and PW (top down) and the two conditions S0 and S1 (left and right). The agreement values are plotted with error bars and a quadratic fit (the outlying value obtained for observer PW at three pixels is not considered because of the poor sampling at that high resolution).

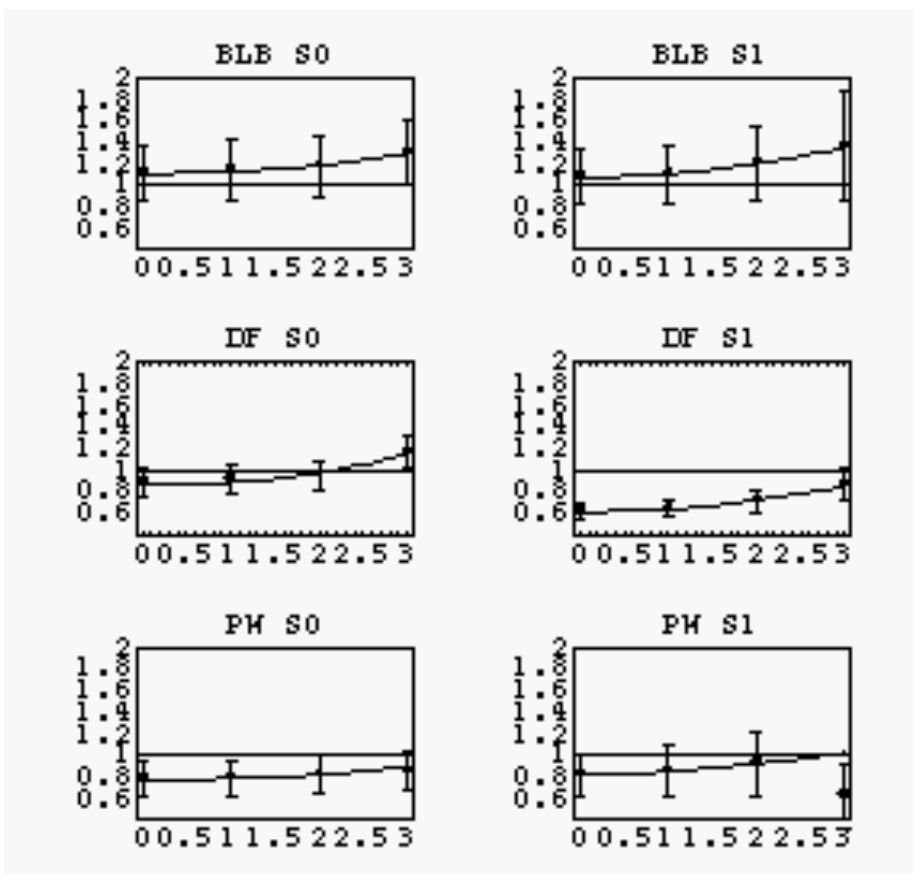


Figure 7: Model-observer agreement for different observers and stimulus conditions, and as a function of isotropic, position uncertainty (measured in pixels).

The consistent trend is that the measures of agreement increase with position uncertainty. In some cases (observers DF and PW) this leads to a better agreement (closer to one). For observer BLB the agreement departs from the optimal value of one, but the

results do not reject position uncertainty.

To better illustrate the meaning of these results, the linear kernel, the Gaussian window of position uncertainty, and the optimal kernel at that uncertainty are shown in Figure 8 for observers DF and PW. The parameter values of the optimal kernels are derived from the quadratic fit, i.e., those values of position uncertainty have been selected, for which the fit of the agreement values equals one. Note that the optimal kernels are similar ($s_y=0.75$ arc min for DF and $s_y=0.67$ for PW) for the two observers. This suggests that a higher image-noise level, as for observer DF compared to PW, has a larger influence on the linear kernel than it has on the optimal kernel, i.e., it increases position uncertainty.

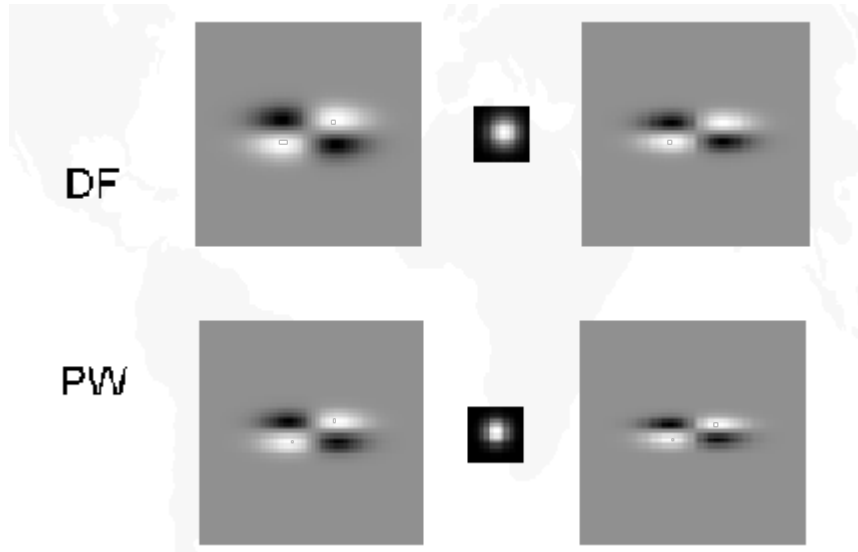


Figure 8: Linear kernel (left), Gaussian for position uncertainty (middle), and the optimal kernel (right) that would yield the linear kernel if convolved with the Gaussian in the middle. All images are shown for the cases DF, S1 (top) and PW, S1 (bottom). The parameters of the Gaussians in the two windows are: $s_x=s_y= 3.7$ pixels (1.14 arc min) for DF, and $s_x=s_y= 2.7$ pixels (0.84 arc min) for PW.

Anisotropic uncertainty

The results shown in Figure 9 have been obtained in analogy to those in Figure 7 but with anisotropic uncertainty. We note that the improvement described for isotropic uncertainty cannot be obtained in this case and we conclude that isotropic uncertainty seems more likely. We should note, however, that the important feature here is not the shape of the Gaussian window used before maximum selection, but the fact that the filter kernel changes its shape in case of isotropic uncertainty - see Discussion below.

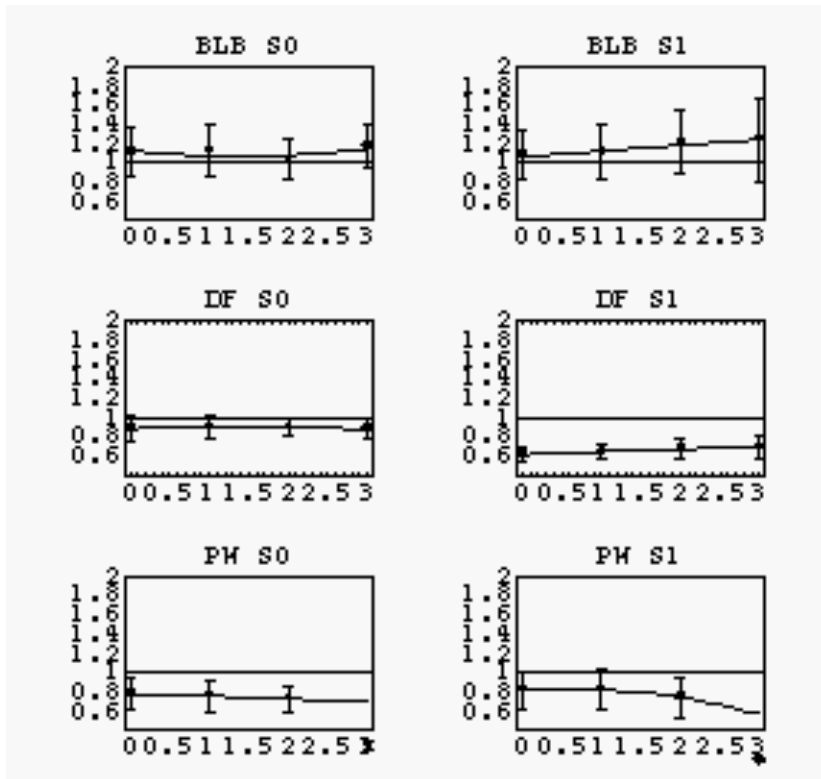


Figure 9: Model-observer agreement for different observers and stimulus conditions, and as a function of anisotropic position uncertainty (measured in pixel).

Orientation uncertainty

The results shown in Figure 7 cannot reject a simple nonlinear model (position uncertainty) but agreement values are still low for subject PW in S0. Results obtained for PW with position and orientation uncertainty are shown in Figure 10. Note that the optimal agreement value of one is reached at an uncertainty of plus-minus two degrees and that the variance is low (0.0002). The optimal kernel (Figure 8 right) was used as described in Section 7. For observer DF, agreement A1 is 0.98 for the optimal kernel and remains constant with orientation uncertainty.

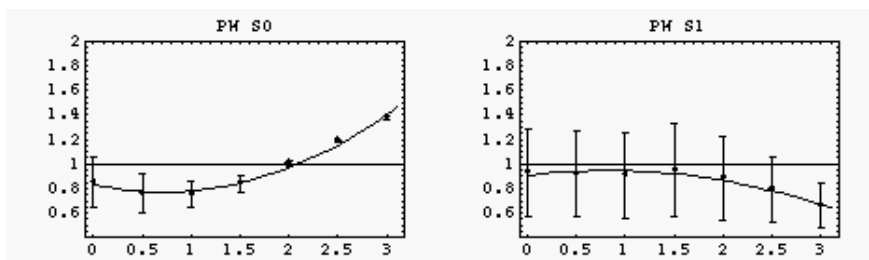


Figure 10: Model-observer agreement for observer PW, and the two stimulus conditions, as a function of orientation uncertainty (sigma of Gaussian in deg).

9. Discussion

We have rejected the linear model suggested by the classification images by using a new measure of agreement and linear kernels with parameters that have been estimated individually for each observer. These parameters show that the observers responding to images with higher noise levels have larger kernels and a lower performance. We have then shown that part of the data under consideration are consistent with the view that the observers base their decision on a measure derived from linear kernels, here estimated as derivatives of Gaussians, and an additional nonlinearity. A possible nonlinearity is position uncertainty. Position uncertainty as implemented here, however, is not the key feature since similar results can be obtained by just using the smaller kernels and picking the center value (instead of the maximum in a region). Since a better agreement is obtained in case of isotropic uncertainty it seems more likely that the shape of the kernel which is actually involved is not only smaller but also thinner (more elongated horizontally than the kernels found as classification images). A further nonlinear effect, which remains beyond position uncertainty, is that the agreements for the stimulus conditions S0 (no offset) and S1 (offset) differ, especially for observer DF.

Improvements obtained with position and orientation uncertainty suggest that observers might not rely on the precise positions of the lines. Therefore, mechanisms that detect the vernier-offset discontinuity become more likely. Further experiments, in which position and orientation of the vernier stimulus are varied, might help to clarify this issue.

In conclusion, two observers seem to make use of kernels that are smaller than the ones derived from the classification images. If these smaller kernels were used as linear kernels by the observers, we should not find the larger classification images. Therefore, a nonlinearity must be involved which "spreads" the size of the kernels. Position and orientation uncertainty are likely candidates with the virtue of simplicity but other nonlinearities are possible and it remains to be seen if the present or future data can further differentiate among them.

10. Acknowledgements

This work has been supported by the Deutsche Forschungsgemeinschaft grant Ba 1176/4-1 to E.B. and NASA RTOP 548-50-12.

11. REFERENCES

1. [Beard, B. L. and A. J. Ahumada](#). A technique to extract the relevant features for visual tasks. In B. E. Rogowitz and T. N. Pappas, eds., *Human Vision and Electronic Imaging III*, SPIE Proceedings Vol. 3299, pp. 79-85, 1998
2. [Ahumada, A.J. and B.L. Beard](#), Response classification images in vernier acuity. *Investigative Ophthalmology and Visual Science*, **39**(4), p. S1109, 1998.
3. Zetsche, C. and E. Barth, Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, **30**, pp. 1111-1117, 1990.
4. Barth, E., C. Zetsche, and I. Rentschler, Intrinsic two-dimensional features as textons. *J. Opt. Soc. Am. A*, **15**(7): p. 1723-32, 1998.
5. Zetsche, C., E. Barth, and B. Wegmann, The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. B. Watson, ed., *Digital images and human vision*. MIT Press: Cambridge, MA, pp. 109-138, 1993.
6. Young, R. A., *The Gaussian derivative model for machine and biological image processing*, Computer Science Dept., General Motors Research Laboratories: Warren, MI 48090, 1985.
7. Koenderink, J.J. and A.J.v. Doorn, Representation of local geometry in the visual system. *Biol. Cybern.*, **55**, pp. 367-375, 1987.
8. Wolfram, S., *Mathematica: A system for doing mathematics by computer. Second Edition*, New York: Addison-Wesley, 1991.