CINTAL - Centro de Investigação Tecnológica do Algarve
Universidade do Algarve

MOVIDE

Modelling Visual Detection

Second annual report
September 2002

Ulrich Bobinger and Hans du Buf

Vision Laboratory
University of Algarve
Campus de Gambelas – FCT
8000-810 Faro, Portugal

Funded by FCT/POCTI/FEDER, project POCTI/36445/PSI/2000

Work requested by	CINTAL Universidade do Algarve, Campus da Penha, 8000 Faro, Portugal tel:+351-289800131, `cintal@ualg.pt`, *www.ualg.pt/cintal*
Laboratory performing the work	VISLAB - Vision Laboratory Universidade do Algarve, FCT, Campus de Gambelas, 8000 Faro, Portugal `ubobing@ualg.pt`, *w3.ualg.pt/~dubuf/vision.html*
Project	MOVIDE Modelling Visual Detection
Authors	Ulrich Bobinger and Hans du Buf
Date	September 2002
Number of Pages	11
Abstract	In this second report we explain the progress in establishing a general model for threshold stimuli. Main focus of this years work was on the quantitative predictions for concentric cosine waves and disk stimuli. All routines for calibration and predictions are now ready in a parallelized version, that runs on a newly assembled cluster with 8 CPUs.
Clearance level	UNCLASSIFIED
Distribution list	VISLAB(1), CINTAL(1)

Contents

1 Resumo

Neste segundo relatório do projecto explicamos o progresso na procura dum modelo genérico de detecção visual de padrões. Baseado no software já descrito no relatório do ano passado, mostramos a flexibilidade do nosso modelo através do cálculo das previsões para polinómios Hermite. Com uma análise detalhada das respostas individuais dos filtros de Gabor, podemos melhorar as previsões para a detecção de discos. A partir destes resultados especificamos os parâmetros dum modelo refinado. A implementação da inibição não local, que efectua a detecção dum estímulo rodeado por flanqueadores, será também discutida. Todos os programas para calibração e previsão estão prontas numa versão paralelizada, que corre num novo cluster com oito CPUs.

2 Abstract

In this second-year report we explain our progress in finding a model of visual pattern detection. Based on the software that was described in the fist-year report we made predictions to other stimuli and show the flexibility of our model by calculating predictions for weighted Hermite polynomes (WHP). With a detailed analysis of the individual responses of Gabor filters we could improve the predictions of threshold data for incremental and decremental disks. From this follows the specifications for some model parameters and conditions for extensions to model refinements. The implementation of nonlocal inhibition, which accounts for the detection of a target surrounded by flankers, will also be discussed. All routines for calibration and predictions are now ready in a parallelized version, that runs on a newly assembled cluster with 8 CPUs.

3 Weighted Hermite Polynomes

Our model should be general in a way that it can account for the detection of various stimulus types. Yang and Reeves (1995) assumed that weighted Hermite polynomes (WHPs) can characterise the receptive fields of certain cortical cells. Hermite polynomes are solutions of equations of the harmonic oscillator, which have been proposed as a mathematical representation of early visual processes (Yang and Reeves, 1991). WHPs form an orthogonal system of functions and the nth order is defined by the following equations:

$$WHP_n(x) = \frac{H_n(x)e^{-\frac{x^2}{2}}}{\sqrt{2^n n!\sqrt{\pi}}}$$

$$H_n(x) = (2x)^n - \frac{n(n-1)}{1!}(2x)^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!}(2x)^{n-4} + ...$$

Under the supervision of Prof. Uwe Mortensen at the University of Münster, detection thresholds for WHPs have been measured in recent years. At ECVP 2002 in Glasgow, Prof. Uwe Mortensen published the results and explained that his own linear model could only be fitted to the data if one assumes a bandpass pre-filter which adjusts to the spectrum of the stimulus pattern.

In the context of the MOVIDE project, we use Gabor filters that model receptive fields of cortical cells, and consider WHPs only as yet a different stimulus set with detection amplitudes that should be predicted by the Gabor model.

Figure 1 illustrates the predictions of our model for the detection thresholds of WHPs for four different stimuli sizes: 0.5, 0.75, 1.0 and 2.0 degree of visual angle.

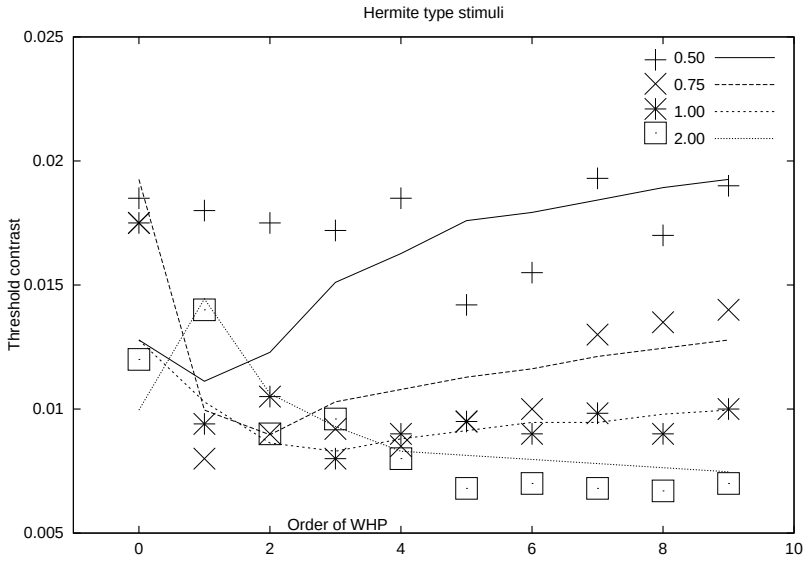

Figure 1: Predictions for Hermite type stimuli for four diffent horizontal widths. Data from Mortensen (2002).

The predictions of our model for WHPs are well correlated with the data for a stimulus size of 1.0 and 2.0 degrees (stars and squares in Fig. 1), but underestimated the threshold values for smaller stimuli and low order of the WHPs . Contrary to the modelling attempts of the group at the University of Münster, we do not have to include any prefilter to predict most of the WHP data correctly.

The introduction of asymmetric filter gain factors, which are needed to account for the differences in detection thresholds, which are currently under investigation, are expected to improve the predictions for smaller stimuli sizes.

4 Analysis of Gabor filter responses to disk threshold detection data

The proposed detection model consists of a bank of Gabor filters which are weighted to account for the CSF for concentric cosine stimuli. This calibration step is discussed in the first annual report of MOVIDE. The calculated predictions for the detection thresholds of circular disks proved only suitable for disks with diameters less than 0.1°.

In order to address this problem we calculated the responses of single Gabor filters to disk stimuli. If these responses are plotted versus the disk diameter (d) times the Gabor frequency (f), all the resulting curves overlap and demonstrate the characteristic features. Figure 2 shows this relationship for the maximum of Gabor filter responses over a wide range of $d \times f$

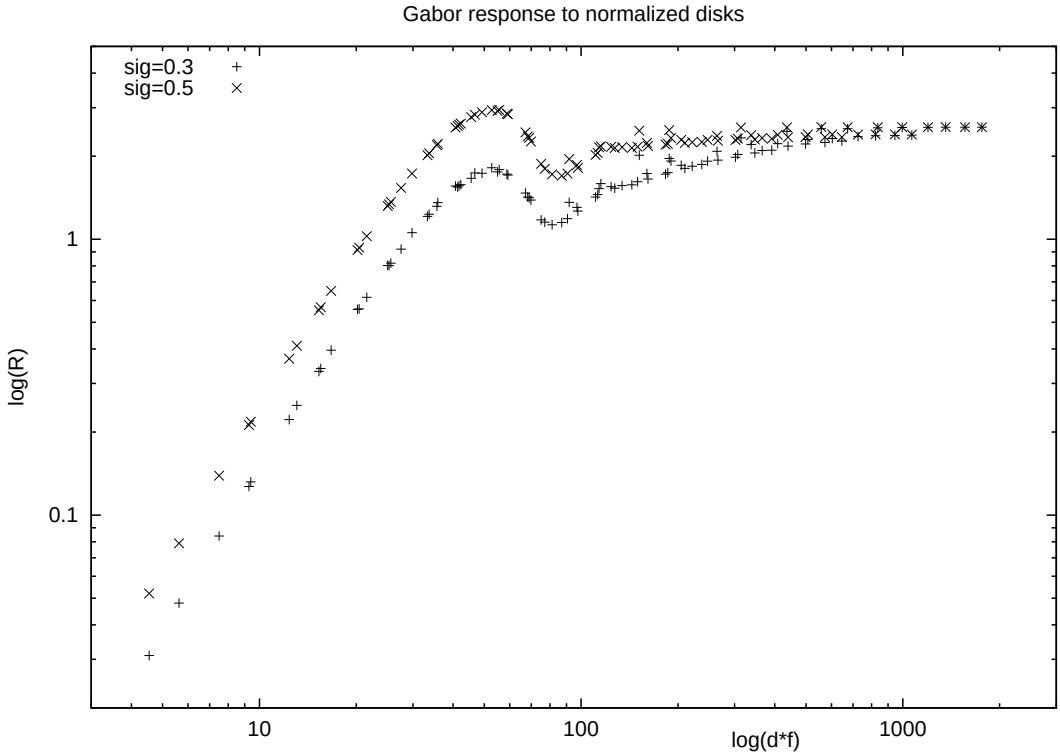

Figure 2: Gabor responses to disk stimuli for two different values of orientational spreads, $\sigma_o = 0.3$ and $\sigma_o = 0.5$

The relationship between the Gabor filter response and the normalised disk size (with respect to the Gabor filter width ($\lambda = 1/f$) depends on σ_o, the Gaussian type spread in the orientation of the Gabor filter. From this we can deduct an upper bound for σ_o. The model with $\sigma_o = 0.5$ exhibits a maximum and a dip, the response for larger disks being constant but less than the first maximum. A similar shape of the amplitude response curve was calculated by du Buf (1992) using a 1-dim model. As this maximum is not reflected by the data, we conclude that the value of σ must be smaller than 0.5 for a correct model.

As shown in Figure 2 the response curve with $\sigma = 0.3$ has also a maximum and a dip, but in this case the maximum is lower than the asymptotic response for large disks. The feature of the dip, which is not found in measured disk data will be addressed in Section 4.2.

In the following sections we will study the consequences of this general feature of Gabor responses to disks for the modelling.

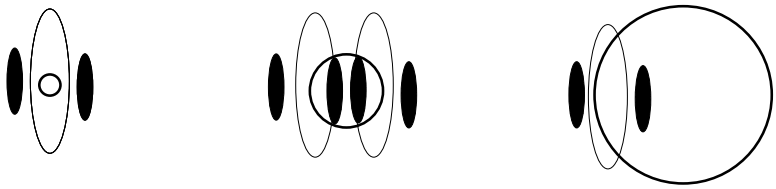

Figure 3: In Fig. 2 three regions can clearly be distinguished: from left to right: small disks, region of interference and large disks

4.1 Small disks

In a region where the disk diameter is smaller than the Gabor wavelength (left in Fig. 3) the response grows linear with the disk area, which is consistent with measured data. However, the absolute values for the contrasts in this region (up to about 30 min^2 for Gabor filter frequencies of 2-4 cpd), determine the exponents of the nonlinear summations after the filter stage in the model. Best approximations could be found $4 \leq \alpha \leq 5$ and $3 \leq \beta \leq 4$ in the case of the simplest model. These restrictions do not hold for modified models, which include nonlocal interactions, inhibition or local summations. In the case of a model with an inhibitory mask (see Fig. 6 in section 5) the best approximations could be found with $\alpha = 2$, $\beta = 3$ and $\alpha = 4$, $\beta = 5$ for decremental and incremental disks, respectively.

4.2 Destructive interference

Figure 2 shows that the response of individual filters becomes smaller when $d \times f$ is between 80 and 110. This dip is not reflected in the measured data. Our multifilter model predictions in this region give nevertheless a good aproximation. The reason is that the first maximum of a filter with a high frequency in the filter bank coincides with the response dip of the filter with the next lower frequency. The following graph (Fig. 4) demonstrates this relationship for calibrated channel responses.

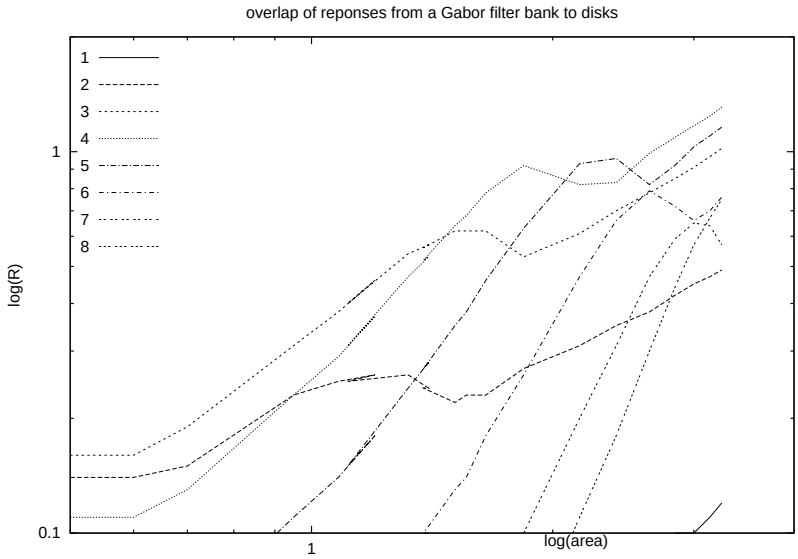

Figure 4: Response curves of individual Gabor filters of a calibrated filter bank with 8 frequencies. The numbers in the legend indicate the frequency bands, 1 being the highest. Crests and troughs of consecutive channels coincide.

4.3 Large disks

For disks with diameters larger than 0.1° the detection amplitude is independent of the disk size. Contrary to this experimental finding, the filter responses of individual Gabor channels always increase with the disk size. This increase has two sources: the curvature of the disk edges becomes smaller and the total length of the edge increases.

The effect of the increasing length is strongest for linear spatial summations ($\alpha = 1$) and vanishes in case of maximum detection, which results in an asymptotic approach to a constant value, as can be seen in Fig. 2.

On the other hand, maximum detection leads to a sharpening of the response of individual channels, which contradicts the smooth shape of the CSF. The best fitted values of the spatial summation index ($\alpha \geq 5$) are fairly close to maximum detection and are not the major contribution to the observed increasing response.

The decreasing curvature of the edge of larger disks causes an intrinsic increase of the response of Gabor filters to disks, as two-dimensional Gabor filters are elongated in the direction orthogonal to the direction of the frequency-modulated part (see Fig. 3 right). Straight-line edges produce a maximum response as the long side of the receptive field is fully excitated, while curved edges only partly cover the ecxitatory center part.

In search of a model that could account for the constant threshold of large disks, we introduced a spatial weighting function of the form

$$sw(r) = 1.0 - \eta * r$$

with r as the radius from the center of the fovea in $min = pixel$. The best estimation of $\eta = 0.022$ leads to the predictions shown in Figure 5

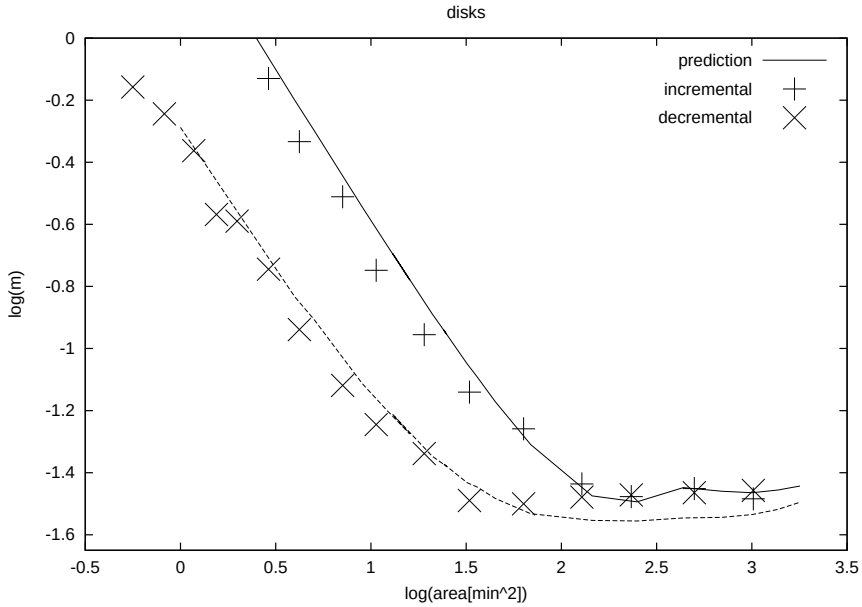

Figure 5: Predictions for incremental and decremental disks using a weighting function $w(r) = 1.0 - 0.022r$. Data from du Buf, 1987.

It can be seen that the predictions of the disk data with an empirical weighting function are satisfactory, but the question remains whether such a function is caused by some underlying process such as nonlocal inhibition. This will be addressed in the next section.

5 Lateral and collinear inhibition

The detection threshold of a pattern depends not only on its contrast and spatial frequency but also on the surrounding. In classic studies Polat and Sagi (1993, 1994) suggested that detection of a target may be influenced by adjacent flanks. The data they presented indicate the existence of facilitation of detection in case of near flankers and inhibitory effects which are maximal at a distance of 2 - 3 λ between target and flankers. Whether the frequency ($1/\lambda$) is the critical parameter to describe facilitaton has been challenged by Levi et al. (2002), who claim that the size of the target rather than the frequency determines the spatial interactions. In a recent study Yu (2002) found that facilitation is frequency tuned and stronger for a cross-orientated configuration.

We extented our model to account for these results by implementing a multiplicative 2-D inhibitory mask ($im_f(x, y)$) for each frequency channel. The response suppression at a position (x, y) is assumed to be proportional to the response at a distance:

$\delta x = 2\lambda\cos(\theta)$ and $\delta y = 2\lambda\sin(\theta)$

The suppression can then be specified by the equation:

$$im_f(x,y) = \quad 1.0 - \frac{w_l}{max(|R(x,y,f)|)} \sum_\Theta |R(x \pm \delta x, y \pm \delta y, f, \Theta)| \qquad \text{collinear}$$

$$- \frac{w_c}{max(|R(x,y,f)|)} \sum_\Theta |R(x \pm \delta x, y \pm \delta y, f, \Theta + \frac{\pi}{2})| \qquad \text{lateral}$$

In our model extension the strength of inhibition has two parameters w_l and w_c which are divided by the maximum response of each channel $|R(x,y,f)|$ in order to balance the effect in case of channels. Maximum collinear inhibition occurs, when the orientation of the target Θ equals the orientation of the flanks θ. Maximum lateral inhibition when they are perpendicular $\Theta = \theta + \frac{\pi}{2}$.

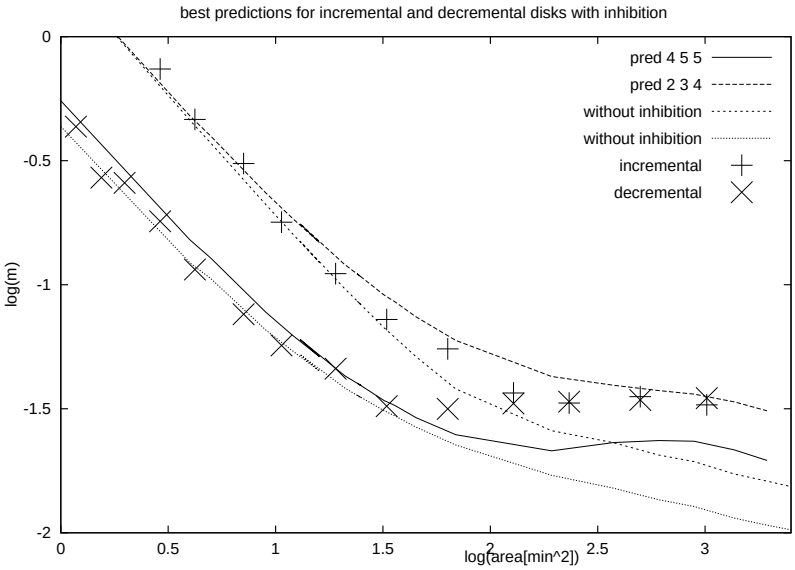

Figure 6: Best prediction for disk threshold data including inhibitation

After the implementation of the inhibitory mask in our algorithm, we used again the threshold data for disks to test the predictions of the modified model. As the disk stimuli are symmetric, they are not decisive with regard to the type of inhibition. In Fig. 6 the best predictions with $w_l = 0.1$ and $w_c = 0.1$ are shown.

In comparison to the predictions of the simple model some improvenemts can be found. The slope of the predictions for large disks becomes less dominant and the good aproximations for small disks could be preserved.

The predictions of the model including a multiplicative inhibitory mask are not as good as the ones of a simple model with linear radial attenuation (Fig. 5). In particular for large disks the predictions still overestimate the sensitivity of the visual system. Model modifications based on different responses for ON and OFF channels, which are now being tested, can hopefully further improve our predictions.

6 Implementation of the model calculations on a new super computer

Due to parallel computing activities in the VISION laboratory we have access to a newly installed cluster with AMD XPs at 1.5 GHz and 512MB DDR in 8 nodes named Calhau. On April 23 the 8 nodes achieved a peak of 5.8 GFLOPS. On May 8 the first production run concerned the calculation of the threshold predictions for 10 disks of different size, which needed 2.5 hours on the SGI Origin (using 1 R10K CPU). The same results were calculated in only 5.6 minutes on the Calhau. After explicitly parallelising the loops in the main program for model calculations, a set of predictions can now be calculated in 2.5 minutes. More details on Calhau can be found at the site: *http://w3.ualg.pt/˜dubuf/calhau.html*.

7 References

du Buf, J. M. H. (1987) Spatial characteristcs of brightness and apparent-contrast perception, PhD thesis, Technical University of Eindhoven, the Netherlands.

du Buf, J. M. H. (1992) Modelling spatial vision at the threshold level. Spatial Vision 6, 25-60.

du Buf, J.M.H. and Bobinger, U. (2002) In search of the Holy Grail: a unified spatial-detection model. Perception Suppl. 31,137 • Project publication •

Mortensen, U. and Meinhardt, G. (2002) Basic coding mechanisms: adaptive or fixed? Perception Suppl. 31,141

Polat, U. (1999). Functional architecture of long-range perceptual interactions. Spatial Vision, 12, 143-162.

Polat, U., and Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. Vision Research, 33, 993-999

Yang, J. and Reeves, A. (1995) Bottom-up visual image processing probed with weighted Hermite polynomials. Neural Networks 8, 669 - 691.

Yang, J. and Reeves, A. (1991) The harmonic oscillator model of early visual image processing. Proc. Visual Communications and Image Science. SPIE, Vol 1606, 520-530.

Yu, C., Klein, S.A., Levi, D. M.(2001). Surround modulation of perceived contrast and the role of brightness induction. Journal of Vision, 1(1), 18-31,

Yu, C. (2002). Contrast detection can be facilitated by cross-oriented surround stimuli. Journal of Vision, 2(3) 243-255

In search of the Holy Grail: a unified spatial-detection model

J M H du Buf, U Bobinger
(Vision Laboratory, University of Algarve, Campus de Gambelas - FCT, P 8000 Faro, Portugal; e-mail: dubuf@ualg.pt; *http://w3.ualg.pt/~dubuf/vision.html)*

Although there exist some models that can predict detection curves for specific patterns like periodic gratings, the development of more advanced models is important because we can try to reason about the processes involved. In an earlier paper (du Buf, 1992 Spatial Vision 6 25 - 60) it was shown that a multichannel model, once calibrated with a sine-wave contrast sensitivity function (CSF), can predict CSFs of other gratings like square-wave, but not detection curves of disks. Recently we developed a modular software package that allows us to easily modify nonlinear models based on complex Gabor filters, first calibrating them with CSFs of 1-D sine-wave gratings, circular (co)sine gratings, or circular Bessel functions, and then predicting detection thresholds of other patterns. Simple models based on a nonlinear summation of Gabor filter responses can predict correct CSFs of many periodic gratings (eg square-wave, square-wave with missing fundamental, trapezoidal), as well as weighted Hermite polynomials. They can also predict many contrast interrelation functions obtained by superimposing specific patterns like lines and edges on a (co)sine background. Correct detection thresholds of disks can only be predicted when including other processes, like collinear and lateral inhibitions. One possible conclusion is that different patterns might be detected at different levels in the visual system.

[Supported by FCT project MOVIDE: POCTI/36445/PSI/2000 and FEDER.]