Poster Presentation/ OSA Fall Vision Meeting 2006 at University of Rochester

Modeling ModelFest Data and Luminance Dependent CSFs Based on Implicit Masking

Jian Yang

Applied Vision Research and Consulting, Penfield, NY, USA

www.visionrc.com, www.webibooth.com

ABSTRACT

Usually, models are developed to describe visual performance at a specified background luminance level. As visual performance varies with not just stimulus contrast, but also absolute luminance, it has advantage to include the luminance as a variable in the model. Here I present a luminance-dependent visual image-processing based on the concept of implicit masking (Yang et al, 1995; Makous, 1997), implemented here with a front-end low-pass filter, a retinal local compressive nonlinearity described by a modified Naka-Rushton equation, a cortical representation in the Fourier domain, and a frequency dependent compressive nonlinearity. The model is used to fit CSFs over 7 mean luminance levels (Van Nes and Bouman, 1967), and to fit the Modelfest data (Carney et al, 1999).

Using Minkowski summation over every frequency components at a decision stage, the model results have a RMS error of 0.10 log unit with the CSF data and 0.12 log unit with the Modelfest data. For the Modelfest data, the error comes largely from stimuli #35 (a noise pattern) and #43 (a natural scene), where the model estimates are much lower than the experimental data. After adding (A) a spatial aperture (Watson and Ahumada, 2005) at the front-end and (B) linear frequency summation windows prior to the Minkowski summation, the RMS error for fitting the Modelfest data reduced to 0.06. By just adding either (A) or (B) alone, the resulting RMS error was about 0.08. Nevertheless, adding these components did not improve the fit to the CSF data.

Purpose:

Develop a general-purpose vision detection model, to be used in real applications with diverse viewing conditions.

As contrast thresholds of spatial and temporal patterns can vary with background luminance levels, we intend to include background luminance as a variable in the model to simulate the effect of luminance on pattern detection.

Approach:

(1). Implement the concept of implicit masking (Yang & Makous, 1994; Yang, Qi, and Makous, 1995; Makous, 1997) into a visual image processing model.

(2). Adjust the model components to deliver a good fit to existing data in the literature.

Frequency Masking:

A concept of frequency masking--- assuming that the human visual system can perform image analysis in a frequency domain somehow, a sine-wave is represented by a dot in this domain. Due to spatial inhomogeneities in the visual system, this dot will spread to its surrounds, producing frequency components that were not in the original sine-wave. Figure 1 demonstrates the effect of frequency masking in the frequency domain as an analogy to the perception in the space domain. The center dot represents a frequency masker.

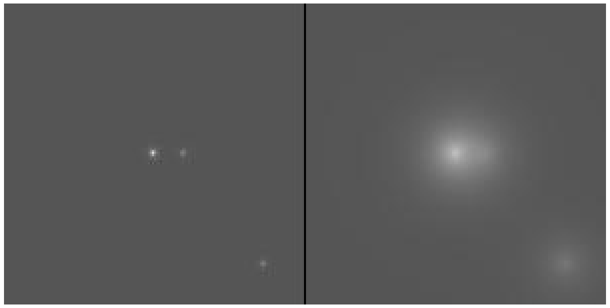

Figure 1. An illustration of frequency components in the spatial frequency domain. The center dot represents a frequency masker. When there is no frequency spread, the visibility for each of the two target dots on its right side is very high (left panel). After adding some spread, the closer target is barely discriminable, while the far dot is still easy to tell (right panel).

Implicit Masking:

It is well-known that the visibility of a spatial target is largely determined by the luminance contrast of the target, but not the absolute luminance of the target. Along this line, traditional vision models assume that the early stage of visual processing is to extract luminance contrast of a visual stimulus as the signal to later visual processing. The potential limitation with such models is that they would not show the effect of absolute luminance on visual detection, if there are any.

Actually, experimentally obtained contrast thresholds do vary with absolute luminance of the stimuli. Among many published reports, Van Nes and Bouman (1967) studied the change of the contrast threshold of a sinusoidal grating with its background luminance and spatial frequency. The results showed that the contrast thresholds are barely affected by the absolute luminance only with coarse patterns at high luminance levels .

What visual functions can be used to simulate the luminance effects? Yang and Makous (1994) proposed that the mean luminance level of a stimulus can be treated in the frequency domain along with all other frequency components of the stimulus, as a component at 0 cpd. As illustrated in figure 1, such a component can act as a frequency masker to other frequency components, especially to those low frequency components which are close to the 0 cpd component in the frequency domain. As the component at 0 cpd is involved in the very nature of any visual stimulus and is not deliberately intended as a stimulus, its desensitizing effects is called *implicit masking* (Makous, 1997).

A Model of Visual Image Processing:

Vision models can have different forms based on the purpose when the models were developed. For example, most of the existing spatial vision models take visual stimuli properties such as spatial frequency, contrast, location of some well-defined spatial patterns as the model inputs. These models have served their purposes very well as to probe the underlying visual mechanisms in processing spatial patterns. In addition, there are also image-based vision models (e.g., Watson & Ahumada, 2005), where the inputs to the models are images or pixels based distributions. Such models can be used in real industry applications to handle with arbitrary target shape.

Our model framework is based on the ideas of implicit masking, modified compressive nonlinear process, and other well-known properties of the visual system that have been used in many models. The model block diagram is shown in Fig. 2.

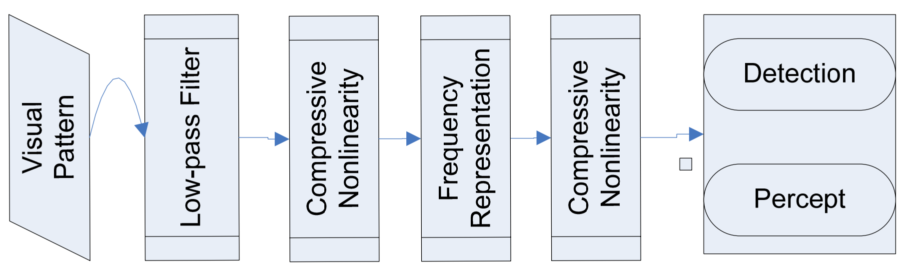

Figure 2. A diagram of visual image processing, which includes a front-end low-pass filter, a retinal nonlinearity, a cortical frequency representation and a frequency-dependent nonlinear process, and finally a decision stage.

Model Computation:

Step 1. Low-pass Filtering: $LPF(f) = Exp(-\alpha f)$

Step 2. Spatial Filtering at Fovea: $I_c(x,y) = \bar{I} + [I(x,y) - \bar{I}]\, Exp(-\frac{x^2+y^2}{2r_f^2})$

Step 3. Retinal Compressive Nonlinearity:

Adaptation pools: $W_g(x,y) = \frac{1}{2\pi r_g^2} Exp(-\frac{x^2+y^2}{2r_g^2})$

Modified Naka-Rushton equation, similar to a divisive normalization process:

$$I_R = \frac{w_0(1+I_0^n)\, I_c^n}{I_g^n + (I_0\, w_0)^n}$$

Step 4. Cortical Compressive Nonlinearity:

Masking pool (or channel): $W_m(f_x, f_y) = Exp[-(f_x^2 + f_y^2)^{0.5}/\sigma]$

The same form of compressive nonlinearity in the spatial frequency domain:

$$T_c = sign(T)\, \frac{w_0(1+T_0^v)|T|^v}{T_m^v + (T_0\, w_0)^v}$$

Step 5. Linear Summation within frequency windows: $\sum T_c$

Step 6. Signal Strength in Detection Stage:

Minkowski summation: $R = \{\Delta f_x\, \Delta f_y\, \Sigma[\, (T_{c\_t} - T_{c\_r})^{\beta}\,]\,\}^{1/\beta}$

Fits to Van Nes & Bouman Data:

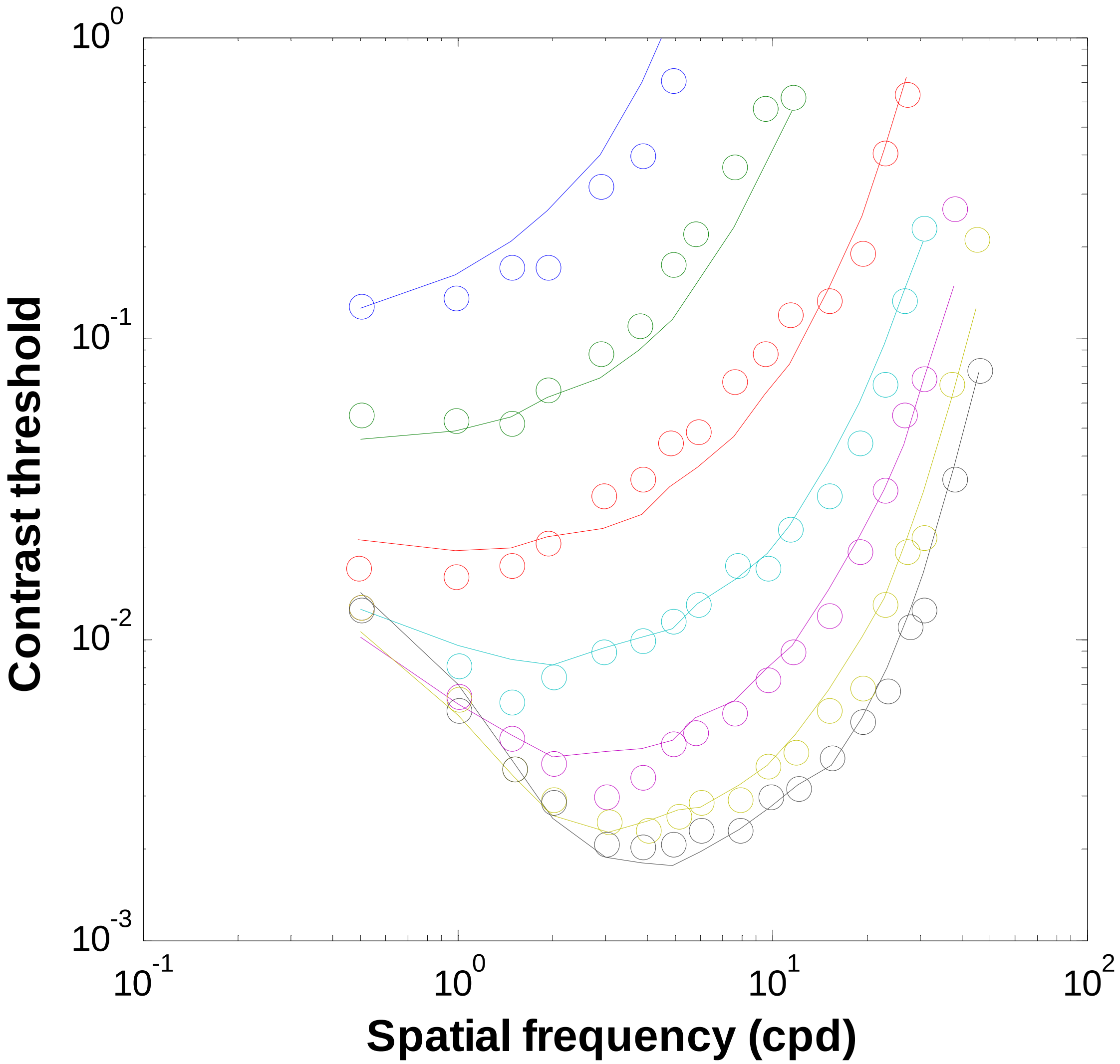

Figure 3. Contrast threshold versus spatial frequency, with mean retinal illuminance ranging from 0.0009 (top) to 900 (bottom) trolands in log steps. The data points are from Van Nes and Bouman and the smooth curves are the fits with current model with a rms error of 0.1 log unit.

Modelfest Stimuli:

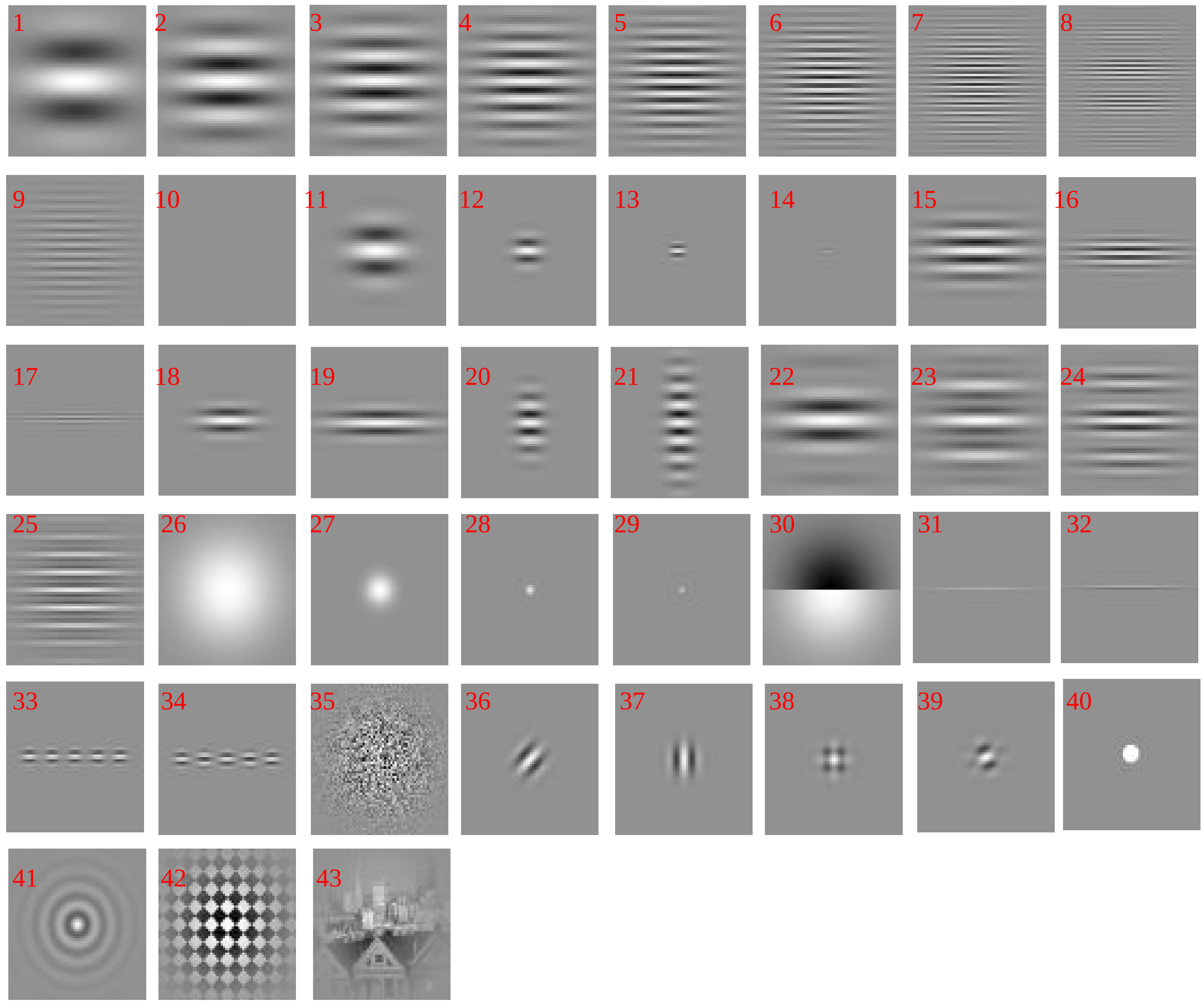

Fig. 4

Fits to Modelfest Data (without steps 2 and 5):

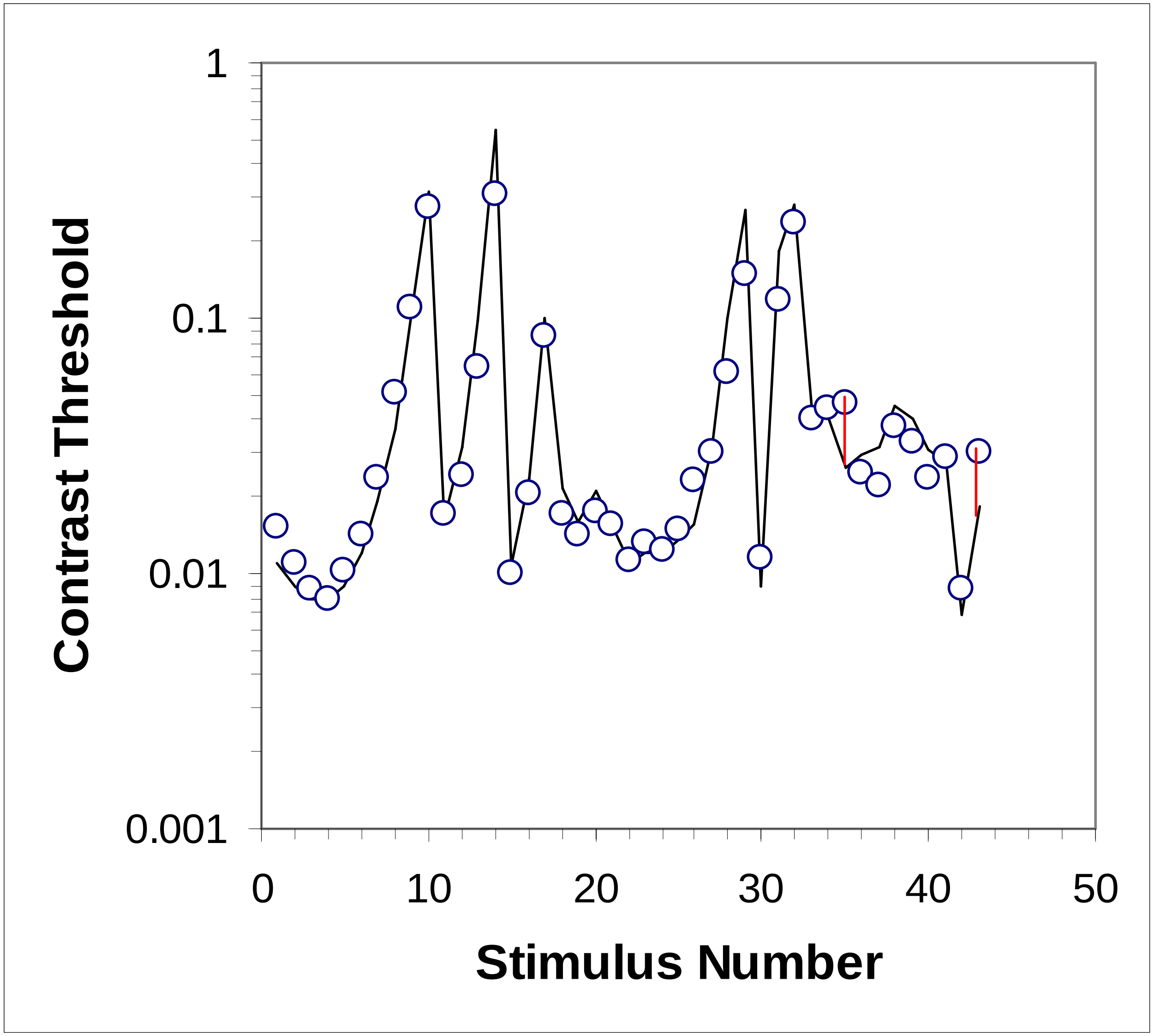

Figure 5. Contrast thresholds of 43 Modelfest stimuli (circles) and the model fits (lines) with a rms error of 0.12 log unit.

Fits to Modelfest Data (after adding steps 2 and 5):

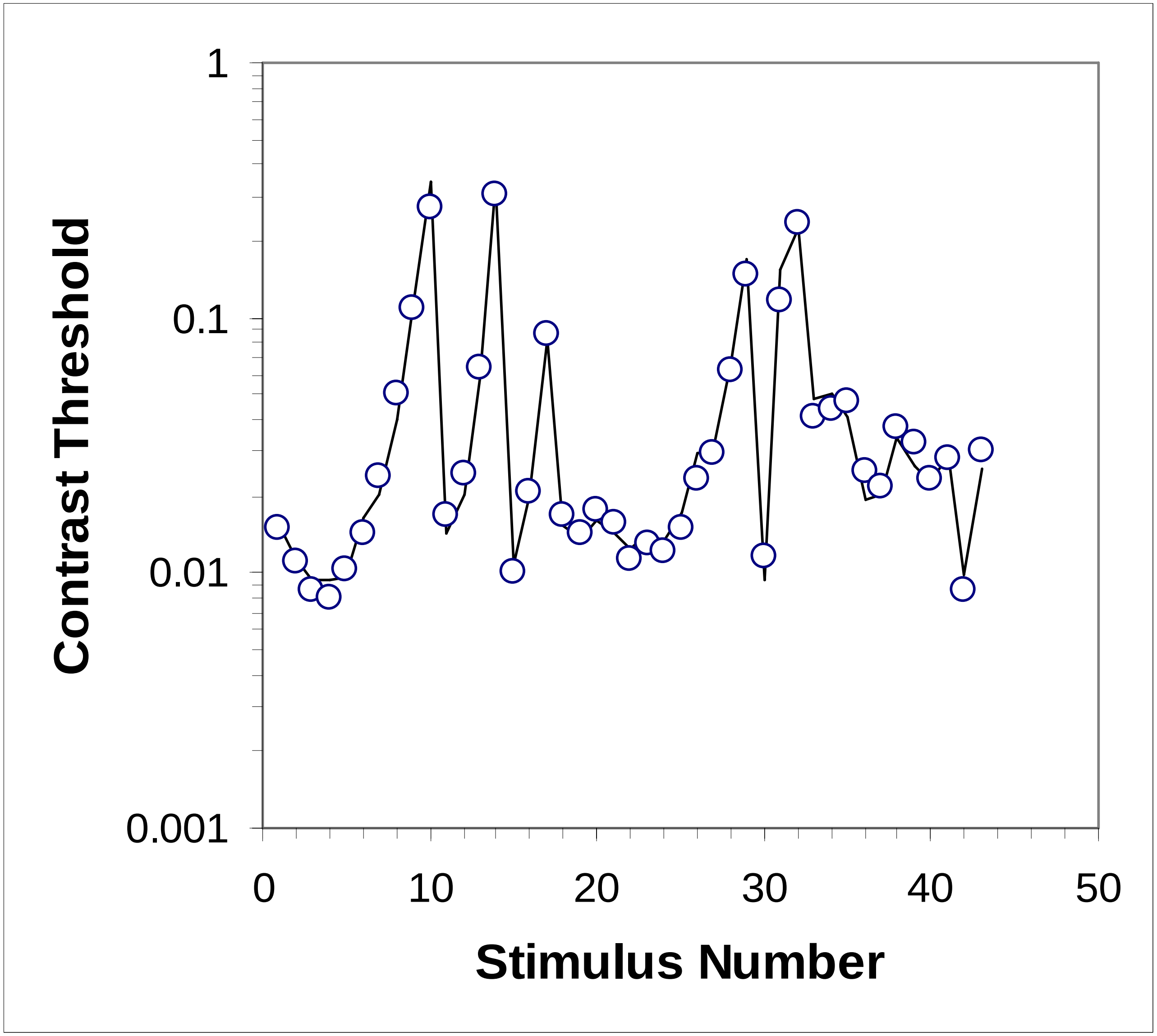

Figure 6. Contrast thresholds of 43 Modelfest stimuli (circles) and the model fits (lines) with a rms error of 0.06 log unit.

Summary:

The term T_m in an equation of Step 4 includes the energy spread of the 0 cpd component. This component is processed in the same way as other frequency maskers, if there are any. Thus, the concept of *implicit masking* is naturally implemented in the image processing framework.

This model can describe the contrast thresholds obtained in two prominent and very different studies, namely the luminance dependent CSFs and the Modelfest data with the same model parameter values, except parameters α which determines the low-pass filter, and R_0 which determines the threshold criterion.

It is important to include (A) a spatial aperture at the front-end and (B) linear frequency summation windows prior to the Minkowski summation. By just adding either (A) or (B) alone, the resulting rms error for fitting the Modelfest data was about 0.08 as compared to 0.06 after including both and 0.12 of none.

A large portion of the materials will be published in EURASIP Journal on Applied Signal Processing, entitled "Simulating visual pattern detection and brightness perception based on implicit masking."

REFERENCES

Carney, T., Klein, S. A., Tyler, C. W., Silverstein, A. D., Beutter, B., Levi, D., Watson, A. B., Reeves, A. J., Norcia, A. M., Chen, C. C., Makous, W., & Eckstein, M. P. (1999). The development of an image/threshold database for designing and testing human vision models. Proc. SPIE, 3644, 542-551.

Makous, W. L. (1997). Fourier models and the loci of adaptation. J. Opt. Soc. Am. A, 14, 2323-2345

Van Nes, F. L. and Bouman, M.A. (1967). Spatial modulation transfer in the human eye. J. Opt. Soc. A, 57, 401-406.

Watson, A. B. and Ahumada, A. J. (2005). A standard model for foveal detection of spatial contrast. J. of Vision, 5, 717-740.

Yang, J. & Makous, W. (1994). Spatiotemporal separability in contrast sensitivity. Vision Research, 34, 2569-2576.

Yang, J., Qi, X. & Makous, W. (1995). Zero frequency masking and a model of contrast sensitivity. Vis. Res., 35, 1965-1978.

Yang, J. (in press). Simulating visual pattern detection and brightness perception based on implicit masking. EURASIP Journal on Applied Signal Processing