Supplementary Formulas for Video Quality Assessment for Computer Graphics Applications

Tunç Ozan Aydın* Martin Čadík* Karol Myszkowski* Hans-Peter Seidel*

MPI Informatik

1 Introduction

In this supplementary material we present fundamental formulas that were left out from the main publication for brevity. Note that all formulas can be found in referenced articles, we merely recollected them for completeness and ease of implementation.

2 JND Space

The JND space nonlinearity accounts for lower sensitivity of the photoreceptors at low luminance, where the luminance L is transformed using a transducer function constructed from the peak detection thresholds [Mantiuk et al. 2005]. One can construct such a transducer function from the following recursive formula:

$$T_{inv}[i] = T_{inv}[i-1] + cvi(T_{inv}[i-1])\; T_{inv}[i-1] \quad \text{for } i = 2..N, \tag{1}$$

where $T_{inv}[1]$ is the minimum luminance we want to consider (10^{-5} cd/m^2 in our case). The actual photoreceptor response R is found by linear interpolation between the pair of i values corresponding to particular luminance L.

The contrast versus intensity function cvi used in the recursive formula above estimates the lowest detection threshold at a particular adaptation level:

$$cvi(L_a) = \left(\max_{\mathbf{x}} \left[CSF^S(\mathbf{x}, L_a)\right]\right)^{-1}, \tag{2}$$

where CSF^S is the contrast sensitivity function and $\mathbf{x}$ are all its parameters except adaptation luminance. If perfect local adaptation is assumed, then $L_a = L$.

3 Static Contrast Sensitivity Function

The static contrast sensitivity function (CSF^S) describes the sensitivity of the visual system as a function of spatial frequency and adaptation luminance. In our implementation we use the CSF proposed by Daly [1993]:

$$CSF^S(\rho, L_a, \theta, i^2, d, c) = P \cdot \min\left[S_1\left(\frac{\rho}{r_a \cdot r_c \cdot r_\theta}\right), S_1(\rho)\right], \tag{3}$$

where

$$\begin{aligned}
r_a &= 0.856 \cdot d^{0.14} \\
r_c &= \tfrac{1}{1+0.24c} \\
r_\theta &= 0.11\cos(4\theta) + 0.11 \\
S_1(\rho) &= \left[\left(3.23(\rho^2 i^2)^{-0.3}\right)^5 + 1\right]^{-\frac{1}{5}} \cdot \\
&\quad A_l \epsilon \rho e^{-(B_l \epsilon \rho)} \sqrt{1 + 0.06 e^{B_l \epsilon \rho}} \\
A_l &= 0.801\left(1 + 0.7\, L_a^{-1}\right)^{-0.2} \\
B_l &= 0.3\left(1 + 100\, L_a^{-1}\right)^{-0.15}.
\end{aligned} \tag{4}$$

*e-mail: {tunc, mcadik, karol, hpseidel}@mpi-inf.mpg.de

The parameters are:

- ρ – spatial frequency in cycles per visual degree,
- L_a – light adaptation level in cd/m^2,
- θ – orientation,
- i^2 – stimulus size in deg^2 ($i^2 = 1$),
- d – distance in meters,
- c – eccentricity ($c = 0$),
- ϵ – constant ($\epsilon = 0.9$), and
- P – absolute peak sensitivity ($P = 250$).

Note that the formulas for A_l and B_l contain the corrections found after the correspondence with the author of the original publication.

Since the filter function depends on the local luminance of adaptation, the same kernel cannot be used for the entire image. To speed up computations, the response map R is filtered six times assuming $L_a = \{\ 0.001,\ 0.01,\ 0.1,\ 1,\ 10,\ 100\ \}$ cd/m^2 and the final value for each pixel is found by the linear interpolation between the two filtered maps closest to the L_a for a given pixel.

4 Spatiotemporal Contrast Sensitivity Function

The spatio–temporal CSF^T on the other hand models the variation of contrast sensitivity at a fixed adaptation luminance (100 cd/m^2 in this case) as a function of spatial and temporal frequencies. In our work we derive spatiotemporal CSF from the following spatiovelocity CSF formula [Daly 1998]:

$$\begin{aligned}
&CSF^T(\rho, v) = \\
&c_0\left(6.1 + 7.3|log\left(\tfrac{c_2 v}{3}\right)|^3\right) c_2 v\left(2\pi c_1 \rho\right)^2 exp\left(-\tfrac{4\pi c_1 \rho(c_2 v + 2)}{45.9}\right),
\end{aligned} \tag{5}$$

where

- ρ is the spatial frequency in cycles per visual degree,
- v is the retinal velocity in degrees per second,
- $c_0 = 1.14$,
- $c_1 = 0.67$, and
- $c_2 = 1.7$.

The last three coefficients ensure that the CSF^T for $v = 0.15$ is close to the CSF^S for $L_a = 100$ cd/m^2. The relation between spatio-temporal frequency ω and retinal velocity is $\omega = v\rho$ assuming the retina is stable.

5 Cortex Transform for Images

The 2D Cortex Transform [Daly 1993] is a collection of the band-pass and orientation selective filters. The band-pass filters are computed as:

$$dom_k = \begin{cases} mesa_{k-1} - mesa_k & \text{for } k = 1..K-2 \\ mesa_{k-1} - base & \text{for } k = K-1 \end{cases} \quad (6)$$

where K is the total number of spatial bands and the low-pass filters $mesa_k$ and $baseband$ have the form:

$$\begin{aligned} mesa_k = & \begin{cases} 1 & \text{for } \rho \le r - \frac{tw}{2} \\ \frac{1}{2}\left(1 + \cos\left(\frac{\pi(\rho - r + \frac{tw}{2})}{tw}\right)\right) & \text{for } r - \frac{tw}{2} < \rho \le r + \frac{tw}{2} \\ 0 & \text{for } \rho > r + \frac{tw}{2} \end{cases} \\ base = & \begin{cases} e^{-\frac{\rho^2}{2\sigma^2}} & \text{for } \rho < r_{K-1} + \frac{tw}{2} \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (7)$$

where

$$r = 2^{-k}, \quad \sigma = \frac{1}{3}\left(r_{K-1} + \frac{tw}{2}\right) \text{ and } \quad tw = \frac{2}{3}r. \quad (8)$$

The orientation selective filters are defined as:

$$fan_l = \begin{cases} \frac{1}{2}\left(1 + cos\left(\frac{\pi|\theta - \theta_c(l)|}{\theta_{tw}}\right)\right) & \text{for } |\theta - \theta_c(l)| \le \theta_{tw} \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

where $\theta_c(l)$ is the orientation of the center, $\theta_c(l) = (l-1)\cdot\theta_{tw} - 90$, and θ_{tw} is the transitional width, $\theta_{tw} = 180/L$. The cortex filter is formed by the product of the dom and fan filters:

$$B^{k,l} = \begin{cases} dom_k \cdot fan_l & \text{for } k = 1..K-1 \text{ and } l = 1..L \\ base & \text{for } k = K. \end{cases} \quad (10)$$

References

DALY, S. 1993. The Visible Differences Predictor: An algorithm for the assessment of image fidelity. In *Digital Images and Human Vision*, MIT Press, A. B. Watson, Ed., 179–206.

DALY, S. J. 1998. Engineering observations from spatiovelocity and spatiotemporal visual models. SPIE, B. E. Rogowitz and T. N. Pappas, Eds., vol. 3299, 180–191.

MANTIUK, R., DALY, S., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2005. Predicting visible differences in high dynamic range images - model and its calibration. In *Human Vision and Electronic Imaging X*, vol. 5666 of *SPIE Proceedings Series*, 204–214.