Constructing a metrics for blur perception with blur discrimination experiments

Chien-Chung Chen*a, Kuei-Po Chen*a, Chia-Huei Tseng*a, Sheng-Tzung Kuo*b, Kuei-Neng Wu*b

* Department of Psychology, National Taiwan University, Taipei, Taiwan
*b Center for Measurement Standards, Industrial Technology Research Institute, Hsin-Chu, Taiwan

ABSTRACT

In this study, we measured blur discrimination threshold at different blur levels. We found that the discrimination threshold first decreased and then increased again as reference edge width blur increased. This dipper shape of the blur discrimination threshold vs. reference width functions (TvW) functions can be explained by a divisive inhibition model. The first stage of the model contains a linear operator whose excitation is the inner product of the image and the sensitivity profile of the operator. The response of the blur discrimination mechanism is the power function of the excitation of the linear operator divided by the sum of the divisive inhibition and an additive factor. Changing mean luminance of the edge has little effect on blur discrimination except at very low luminance. When luminance is low, the blur discrimination was higher at small reference blur than those measured at medium to high luminance. This difference diminished at large reference blur. Such luminance effect can be explained by a change in the additive factor in the model. Reducing contrast of the edge shifted the whole TvW function up vertically. This effect can be explained by the decrease of gain factors in the linear operator. With these results, we constructed a metric for blur perception from the divisive inhibition we proposed and tested in this study.

Keywords: MPRT, motion picture response time, divisive inhibition, luminance contrast, normalization.

1. INTRODUCTION

The ability to present a clear and sharp image is a desirable feature of any imaging device. However, due to all sorts of technical and fiscal constraints, a sharper display can also mean higher production cost. Hence, it has been a challenge to find the balance between the display quality (sharpness of the image) and production cost. We propose to take the performance of the human visual system into account while seeking for such balance because display quality is heavily dependent on human viewers’ perception. To accomplish this goal, it is helpful to have a good quantitative metrics for sharpness or blur (i.e., the opposite of sharpness) to evaluate the quality of a visual display.

A metrics for blur is of particular importance for TFT-LCD displays. In addition to the blur caused by the limitation of pixel size, TFT-LCD displays also has a serious problem of motion blur. That is, when showing motion pictures on a TFT-LCD display, the edges that move from one frame to the other show a blur of haziness worse than those on a still image. Much of the motion blur is due to the slow response time of the TFT-LCD monitors. Suppose that a pixel of a TFT-LCD display is required to change from luminance level x in one frame to level y in the next to signal the moving of a sharp edge. Due to the slow time course of changing luminance, such operation may take more than one frame to finish. As a result, instead of changing from x to y, that pixel can only reach a level somewhere between x and y, and can achieve level y only after several frames. For an observer, it is equivalent that the pixel is signaling a blur edge with luminance changing gradually from frame to frame.

Currently, moving picture response time (MPRT) is a widely accepted measurement of motion blur on a TFT-LCD display. The MPRT procedure has a camera tracking a moving bright bar or edge on a dark LCD screen and taking pictures. The motion blur is then defined as the width of luminance transition from dark (e.g., 10% of the maximum luminance) to bright (e.g., 90% of the maximum luminance). The implicit assumption of MPRT is that an
observer tends to be attracted by a moving edge and thus the fixation of the eyes moves with the moving edge. Hence, it is possible to measure the motion blur on a still captured image.

This procedure, however, may not be sufficient motion blur assessment because of its ignorance of human observers' perception. The physical measurement of blur is not necessary to have a linear relationship with the psychological percept of blur. Hence, we believe that an understanding of the psychological response to blur will contribute greatly to construct a better metrics of display quality. In our study, we will first show that the psychological response to blur has a nonlinear relationship with MPRT. A good metrics of blur should consider this nonlinearity. The purpose of this study is to characterize this nonlinearity and to construct a model for blur perception and in turn establish a metrics for blur.

1.1 Discrimination and response nonlinearity

Pattern discrimination has been a well established method to estimate the nonlinear properties in the human visual system\(^3-7\). In a typical pattern discrimination experiment, the experimenter presents a reference pattern with intensity \(I\) in the visual dimension of interest and a target pattern with intensity \(I+\Delta I\). The task of the observer is to tell the difference between a reference pattern and a target pattern. The discrimination threshold, or the just-noticeable difference, between the target and the reference is defined as the differential intensity \(\Delta I\) that allows the observer to tell the difference between the reference and the target with certain percentage of correctness.

A typical result of discrimination experiments is the "dipper" shaped target threshold vs. reference intensity (TvI) function\(^3-5,7\). That is, as reference intensity increases, the discrimination threshold first decreases (facilitation) and then increases (masking) as shown in figure 1a. The TvI function reflects the response characteristics in the visual system. As shown in Figure 1b, in order for an observer to detect the difference, the differential intensity \(\Delta I\) has to be large enough to exceed the response evoked by the reference pattern by a certain amount, defined as one unit. Suppose that the response function is accelerating near a reference intensity (e.g., \(I_1\) in Figure 1b), it would require less differential intensity (\(\Delta I_1\)) to increase the response by one unit than the absolute threshold (\(I_0\), the threshold measured when the reference has zero intensity). On the other hand, when the response function is decelerating near a reference intensity (e.g., \(I_2\) in Figure 1), it would take greater differential intensity (\(\Delta I_2\)) to evoke the same amount of response. Thus, the discrimination threshold at a certain reference intensity is inversely proportional to the slope of the response function at that reference intensity. Therefore, with a systematic measurement of discrimination threshold at various reference intensities, we can estimate the slope of the response function and in turn the response function itself to a scalable constant.

1.2 luminance effect

In current study, the luminance effect on blur discrimination is also investigated. Current MPRT standard defines the blur as the distance between 10% and 90% luminance change. However, it does not specify the luminance level where this change occurs. Hence, MPRT may be measured as, say, the distance between two pixels with 1 and 9 cd/m\(^2\) respectively on an edge between 0 and 10 cd/m\(^2\) or the distance between 10 to 90 cd/m\(^2\) on an edge between 0 and 100 cd/m\(^2\). As a result, MPRT measurements from different institutions may not based on the same luminance levels. This is potentially problematic for a comparison between different institutions. Since blur is essentially a loss of high spatial frequency information in the visual system\(^8\) and it is known that that luminance can affect the human contrast sensitivity

---

**Figure 1.** The relation between the TvI function and the internal response.
function², it is possible that luminance can also affect blur perception. If so, the MPRT measured at one institute may not be comparable with that measured at a different institute. A similar inconsistence also occurred in the luminance contrast dimension. One institution may measure MPRT with an edge going, say, from 0 to 255 level of an 8-bit graphic DAC, or from the lowest luminance level to the highest a display can produce while the other may measure it with an edge going from, say, 128 to 255 level, or from an intermediate to the maximum luminance. This generates a potential obstacle to utilize MPRT measurement taken from different institute as a general index of display quality.

To rectify this problem, we investigated the luminance effect and the contrast effect on human blur responses. The aim is to understand human blur responses at different luminance and contrast levels, so that this understanding can be used to establish an adequate transformation which equates MPRT measured at one luminance level to that measured at another. We will offer a model that accounts for our experimental data on the luminance effect of blur discrimination as well as the contrast effect observed by other scientists¹⁰.

2. METHOD

2.1 Stimulus

The stimuli were edges within a rectangular bar with 10.7 degree (640 arc minute) width and 0.7 degree (40 arc minute) height. The edge stimuli \( L(x) \) were defined as

\[
L(x) = L_0 + \Delta L \times (\Phi(x, 0, \sigma) - 0.5) \times 2
\]

where \( \Phi(x, 0, \sigma) \) was the cumulative Gaussian function with mean 0 and scale parameter \( \sigma \), \( L_0 \) was the mean luminance, and the \( \Delta L \) was the difference between the maximum and the mean luminance on the stimulus. Figure 2 shows an example of our stimulus. The origin \((x=0)\) point was placed as the center of the display and was marked by a fixation point. The ratio \( \Delta L/L_0 \) determined the luminance contrast of the edges. We used four different mean luminance levels \( L_0 \), 0.26, 2.58, 5.16 and 25.8 cd/m² while keeping the contrast \( \Delta L/L_0 \) at 1.0 throughout the experiment. The scale parameter, or “standard deviation” of the Gaussian function, \( \sigma \) defined the amount of blur in the image. Notice that MPRT is defined as the distance between 10% and 90% luminance change of the edge. Hence the edge width, \( w \), defined in the same way as the MPRT of a captured image, was 2.56\( \sigma \) as derived from the standard Gaussian cumulative distribution function. Nine reference edge widths, ranging from 0 (“sharp edge”) to 0.525 degree visual angle were used in current study.

2.2 Procedure

The images were displayed on a 19-inch EIZO LCD monitor (FlexScan S1910) with 1280(H) \times 1024(V) resolution and 60 Hz refreshing rate. The distance between observers and the monitor was 204 cm. At this distance, the width of a pixel was 0.5’ visual angle. We used the LightMouse photometer¹¹ to measure the full-detailed input-output intensity function
of the monitor. This information allowed us to compute linear lookup table settings to linearize the output within 0.2%. The experiment was conducted in a dark room.

The blur discrimination threshold was measured with a temporal 2AFC paradigm. As shown in Figure 3, in each trial, a fixation point was first appeared in the center of the display for 500 ms. After fixation, the observers were shown two 200-ms long intervals successively separated by a 500-ms break. One interval contained a reference blur stimulus and the other interval contained a test stimuli. The order of the reference and the test was randomized from trial to trial. A noise mask was presented in the break period to avoid apparent motion or other local artifacts that may provide extra cues for the observer.

In each trial, the task of an observer was to press a key to indicate which interval contained a fuzzier edge. We used Quest threshold seeking algorithm to determine the discrimination threshold at 75% percentage correct level. There were 40 trials for each threshold measurement. The threshold \( \Delta w \) was defined as \( w_{\text{test}} - w_{\text{ref}} \) where \( w_{\text{ref}} \) was the width of the reference edge while \( w_{\text{test}} \) was the width of the test edge at which 75% of the time observers succeeded in making the correct judgment. Each datum point reported was an average of 4 to 8 repeated measures. We randomized the sequence in which the reference width was presented in each threshold measurement.

All measurements described above are repeated at four luminance levels: 25.8, 5.16, 2.58, and 0.26 cd/m\(^2\). The luminance of the display was controlled by applying a neutral density (ND) filter between the eye and the display. The visual angle of the ND filters was 10.7°, matches the width of the display. The transmittance rates of the ND filters were 20%, 10%, and 1%. With the no-ND filter condition, we had four mean luminance levels. The mean luminance level without the ND filters was 25.8 cd/m\(^2\). Thus, the three ND filtered conditions had mean luminance 5.16, 2.58 and 0.26 cd/m\(^2\), which corresponded to 20%, 10%, and 1% transmittance rate respectively.

Three female and one male observers participated in this study. Their average age was 24.7 years old. All observers had corrected to normal visual acuity (20/20).

3. RESULT

Figure 4 shows the relationship between the blur discrimination threshold (\( \Delta w \)) and the reference edge width at different luminance levels averaged across 4 observers. The smooth curves are the fits of the model discussed below. All blur discrimination threshold vs. reference width (TvW) functions show a dipper shape. That is, the discrimination threshold first decreased and increased again as the reference width increased. At the mean luminance level 25.8 cd/m\(^2\), the observer could tell the difference between a sharp edge (\( w_{\text{ref}} = 0 \)) and a blur edge of only 0.03° visual angle. As the width of the reference edge increased, the blur discrimination threshold decreased 2 folds before increased again. The best blur discrimination thresholds were in par with the known limited imposed by the distance between two foveal photoreceptors, or 0.016°. Notice that there is no linear relationship between reference blur and blur discrimination threshold.

The mean luminance levels at 2.58 cd/m\(^2\) and above had little effect on blur discrimination. At the same reference width, the difference between discrimination thresholds measured at different luminance levels was never statistically significant at the alpha level of 0.05. The TvW function measured with the lowest luminance level was

![Figure 4](image-url)
luminance, on the other hand, was different from the rest. When the reference width was small, the discrimination threshold increased two-fold from that measured at higher luminance. This difference diminished as the reference width. Such convergence of TvW functions implies that, when luminance did have an effect, its effect was a constant regardless the reference blur. Hence, as blur increased, compared with the effect of blur itself, the contribution from luminance became less important. As a result, the TvW function got closer on a log-log plot as reference blur increases.

Luminance above 2.58 cd/m$^2$ has little effect on blur discrimination. Since most MPRTs are measured above this level, those measurements should be commensurable with each other. On the hand, some did measure MPRT at low luminance level. This may not be the best practice if the compatibility is a concern.

4. DISCUSSION

4.1 Model

We proposed a divisive inhibition model to account for our result. This model incorporated the notion of divisive inhibition or normalization \cite{4,5,13-18} that has successfully explained a wide range of discrimination data in pattern vision. However, since the divisive inhibition model was originally developed to account for contrast detection and discrimination, it needed to be modified to account for blur discrimination.

Figure 5 showed a diagram of our model. This model contains several stages. The first stage is a linear operator with a spatial sensitivity profile $f(x, y)$. The excitation of this linear operator to an image $L(x, y)$ is given as

$$E' = \sum_x \sum_y f(x, y) \cdot L(x, y). \quad (2)'$$

Since we only considered one-dimensional edge in the experiment, Eq. (2)' can be reduced to

$$E' = \sum_x f(x) \cdot L(x). \quad (2)$$

The sensitivity profile of the linear operator is designed to simulate the receptive field of a neuron \cite{19,20} in the primary visual cortex. That is, the sensitivity profile contained an excitatory region sided by an inhibitory region. Georgeson et al. \cite{8} further suggested that an even-symmetric arrangement of such sensitivity profile is able to encode different blur level. We used a difference of two Gaussian functions to construct the even-symmetric linear receptive field. That is, the sensitivity profile

$$f(x) = k \cdot [\exp(-(x-u)^2/s^2) - \exp(-(x+u)^2/s^2)] \quad (3)$$

where $k'$ is a constant controlling the gain of the linear operator; $u$ and $s$ are the location ("mean") and the scale ("standard deviation") parameters respectively. The two Gaussian functions are entirely symmetric about the origin, or the center of the linear operator, except sign. Figure 6 shows examples of such difference-of-Gaussian (DOG) sensitivity profile with different location parameters. The location parameters also define the loci of the maximum and the minimum point of the sensitivity profile. Thus, the distance between the maximum and the minimum points is $2u$. Hence, the peak spatial frequency of this linear operator is $1/4u$. The scale parameters were set to $1.2u$ in our computation. Such scale parameter allows each individual Gaussian function reaching its half height at the origin.
Figure 7 shows the excitations of the DOG linear operators to edges with different width computed by plugging Eq. (1) & (3) in Eq. (2). For the convenience of visualization and comparison, the excitation shown in Fig. 7 has been normalized by the squared root of energy in the linear operators. That is, the constant k was set at $\sum f(x)^{2}^{0.5}$. In each linear operator, the excitation decreases nonlinearly with edge width. The shape of the decrement looks like an inverse of a sigmoid shape. In addition, the edge width shows a greater effect on the excitation of the low high spatial frequency linear operators (i.e., smaller u) than on the excitation of the low spatial frequency ones (larger u). Such behavior is consistent with notion that blur means attenuation of power in the high spatial frequency components of an edge. Hence, blurring a sharp edge should reduce the response of visual mechanisms and such reduction should be greater in high spatial frequency channels than low spatial frequency ones. Such inverted sigmoid shape can be described by the expression

$$E' = k \cdot \tan^{-1}\left(\frac{w}{(h \cdot u)}\right) / \left(\frac{(\sigma / (h \cdot u))}{\sigma / (h \cdot u)}\right)$$

(4)

where w is the width of the edge as defined in Eq. (1), u defines the peak spatial frequency of the linear operator as shown in Eq. (3) and h = $\left(\frac{2}{\pi}\right)^{0.5}$ is constant. Such expression, shown as dotted curves in Fig. 7, fits the linear operator excitations with an overall deviation less than 1%.

The excitation of the linear operator is halfwave-rectified to produce the rectified excitation $E_j$

$$E = \max(E', 0)$$

(5)

where max denotes the operation of choosing the greater of the two numbers.

The response of the blur discrimination mechanism is the rectified linear operator excitation raised to the power p and then divided by a divisive inhibition term I, limited at low levels by an additive constant z. That is,

$$R = E^p / (I + z)$$

(6)

The divisive inhibition input is a nonlinear combination of the rectified excitations of all relevant mechanisms, that is, those have non-zero response to the stimuli, given by

$$I = \Sigma m_n E_n^q$$

(7)

Such nonlinear summation tends to be dominated by the largest term in the summation. This largest number usually comes from the one with largest excitation, E. Hence

$$I = S_i \cdot E^q$$

(7)

where $S_i$ is the sensitivity of the mechanism to the divisive inhibition input.
In our two-interval forced choice experiment, there were two images in each trial: the reference and the test. Let \( R_{\text{ref}} \) and \( R_{\text{test}} \) be the responses to the reference and the test respectively. The difference in response is

\[
D = R_{\text{test}} - R_{\text{ref}}
\]  

(8)

For an observer to perceive the difference between these two stimuli in a trial, the difference of the responses produced by the reference and the test should be greater enough to overcome the noise in the visual system\(^3\). That is, the width difference \( \Delta w \) is at threshold if \( D \) reaches a critical value. In our model fitting, this critical value is designated as 1.

4.2 Luminance effect

The parameters to be estimated in the model includes the gain factor, k and the scale factor, s, of the linear operators; the exponents p and q; the sensitivity to inhibition \( S_i \) and the additive constant \( z \). To reduce redundancy, the gain factor \( k \) was fixed at 100. As discussed above, in our data, the difference between \( T_vW \) function diminished at high reference width. This suggest that the luminance effect was a constant regardless the reference blur. The additive parameter \( z \) in the model fits the bill. Hence, we allowed \( z \) to change with luminance level while other parameters, p, q, \( S_i \) and s, were the same for all luminance level. The best fit of the model are shown as smooth curves in Figure 4. This model fits the data well. The RMSE was 0.061 log unit, smaller than the averaged standard error of 0.128 log unit.

4.3 Contrast effect

Our model can also account for the luminance contrast effect. Since we did not manipulate edge contrast in our experiment, we fit our model to the data collected by Hess et al.\(^{10}\). The symbols in Figure 8 show the data of Hess while the smooth curves are the fits of our model. Their experiment was quite similar to ours. They asked their observers to detect the difference between two edges of different blur levels. And then, they measured the \( T_vW \) functions at different luminance contrast levels. The dipper shaped \( T_vW \) functions were quite similar to ours. The difference is that the contrast effect was to shift the \( T_vW \) functions vertically on a log-log plot and showed no convergence at large edge width.

It is known that the excitation of the linear operator increased with contrast\(^{21}\). Hence, the contrast effect is best modeled as a change in the gain parameter \( k \). We set \( k \) to be 100 for the 100% contrast condition and allowed \( k \) to vary with other contrast. The other parameters, including \( z \), were the same for all contrast conditions. As shown in Figure 8, the model captures all aspects of the contrast effect in the data. The RMSE of the model fit was 0.082 log unit, in par with the fit to the luminance data. Figure 9 shows how the gain parameter changed with luminance (circles). Such effect

![Figure 8](https://example.com/figure8.png)  
Figure 8. Contrast effect on blur discrimination. The data points are taken from Hess et al.\(^{10}\) while the smooth curves are the fits of our model.

![Figure 9](https://example.com/figure9.png)  
Figure 9. The contrast effect on gain factor k. The circles are best fit values of k at different edge contrast c. The dotted line denoted the values of 100 • \( \Phi(c, 0, 6) \) at those edge contrasts.
was well described by a cumulative Gaussian function of contrast $100 \cdot \Phi(c, 0, 6)$ (dotted line).

5. CONCLUSION

In this study, we measured blur discrimination threshold at different reference edge widths. We found that the discrimination threshold first decreased and then increased again as reference blur increased. This dipper shape of the TvW functions can be explained by a divisive inhibition model. The first stage of the model contains a linear operator whose excitation is the inner product of the image and the sensitivity profile of the operator. The response of the blur discrimination mechanism is the power function of the excitation of the linear operator divided by the sum of the divisive inhibition and an additive factor.

Changing mean luminance of the edge has little effect on blur discrimination except at very low luminance. When luminance is low, the blur discrimination threshold was higher at small reference blur than those measured at medium to high luminance levels. This difference diminished at large reference blur. Such luminance effect can be explained by a change in the additive factor in the model. Reducing contrast of the edge shifted the whole TvW function up vertically. This effect can be explained by the decrease of gain factors in the linear operator.

Based on the best fit of the model and eq. (4) and (6), we can conclude that the response of the human visual system to an edge with width $\sigma$ is

$$R_w = \frac{E^{0.84}}{(0.02 \cdot E^{0.74} + 0.9)}$$

where $E = 100 \cdot \tan^{-1}(\sigma/(0.08)) / (\sigma/0.08))$ for edges with medium to high luminance and 100% contrast. For contrast less than 100%, $E$ is about $100 \cdot \Phi(c, 0, 6)$. This model thus associates the physical blur with the perceived blur at different luminance and contrast levels.

ACKNOWLEDGEMENT

Support by grants from the Center for Measurement Standards, the Industrial Technology Research Institute, Taiwan,7407EF1200 to CCC and 7407EF1300 to CHT.

REFERENCES


