

A VISIBILITY MODEL FOR QUALITY ASSESSMENT OF DIMMED IMAGES

Tai-Hsiang Huang¹, Chen-Tai Kao², Yi-Chia Chen³, Su-Ling Yeh³, and Homer. H. Chen¹

¹Graduate Institute of Communication Engineering
²Department of Electrical Engineering, ³Department of Psychology
National Taiwan University, Taipei, Taiwan

ABSTRACT

Dimming the backlight of a liquid crystal display (LCD) reduces the dynamic range and intensity of a displayed image and hence has a profound impact on the perceived image quality. To reflect the actual degradation of image quality perceived by human eye, we take visibility in addition to contrast into account and develop a visibility model for image quality assessment. The proposed visibility model also considers factors related to the display and the ambient light to accommodate the viewing condition. Experimental results are shown to demonstrate the performance of the proposed model.

Index Terms—Image quality assessment, dim backlight

1. INTRODUCTION

Dimming the backlight of a liquid crystal display (LCD) saves the battery power of a portable multimedia device but considerably degrades the image quality largely due to the reduction of image intensity and dynamic range. An example of the image quality degradation is shown in Fig. 1, where the dark image regions become too dark to be visible to human eye when the image is illuminated with dimmed backlight.

The decision of the backlight intensity can be considered a tradeoff between power saving and quality degradation. The amount of power saving can be computed by directly measuring the power usage of the display device; however, there is no obvious way for a computer to measure the difference in visual quality between images illuminated with different backlight intensities. Although a human observer can easily choose which image looks better, running an extensive subjective test to quantify the amount of visual quality difference is labor intensive and impractical. Therefore, there is a need for a computational scheme to measure such quality difference.

Many existing image quality assessment metrics, such as the high dynamic range visual difference predictor (HDR-VDP) [1], its advanced version HDR-VDP-2 [2], the structure similarity index (SSIM) [3], the visual information fidelity metric (VIF) [4], and the visual signal to noise ratio

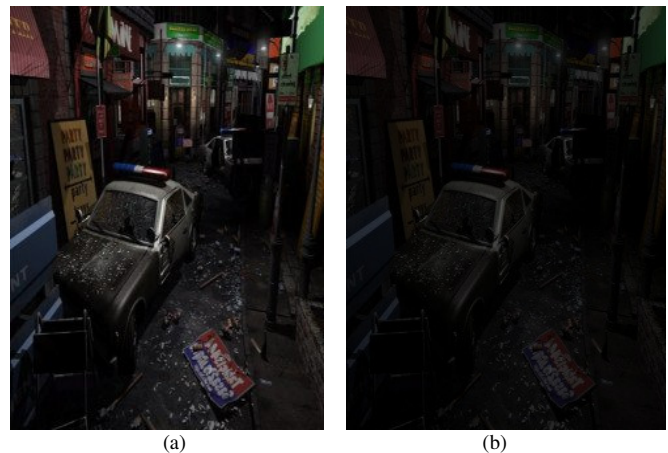


Fig. 1. Illustration of the effect of LCD backlight on perceptual image quality: (a) 100% backlight and (b) 10% backlight.

(VSNR) [5], are designed to quantify distortions introduced in the process of image compression, image processing, image capturing, or noisy communication. These metrics typically work under the presumption that the test image and the reference image have the same dynamic range. Neglecting this presumption and directly applying such metrics to images illuminated with different backlight intensities would lead to incorrect quality assessment.

A number of quality assessment metrics are designed for images with different dynamic ranges. For instance, the dynamic range independent metric (DRIM) [6] evaluates the quality of an image based on its contrast loss due to the reduction of dynamic range and image intensity. But, being a perceptual attribute of an image, the contrast is a meaningful measurement only when the pixels under evaluation are visible. For invisible pixels such as those in Fig. 1 under dim backlight, the DRIM type of values become meaningless because the contrast of such pixels is not detectable by human eye.

We believe visibility is an important attribute of image quality, especially for the dim backlight scenario described above. Therefore, in addition to contrast, we include pixel visibility as part of the factors of image quality assessment. Our goal is to measure the visibility difference of an image under two lighting conditions: full backlight and dimmed backlight.

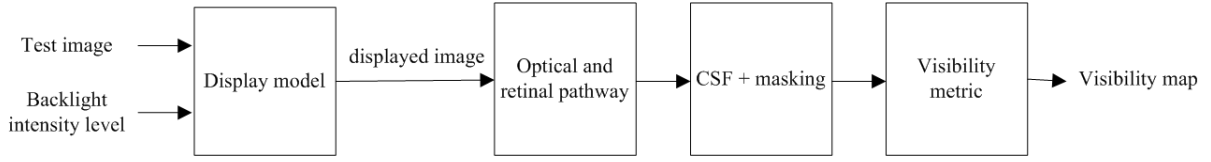


Fig. 2. The architecture of the proposed model.

Given an image and a backlight intensity level, our proposed visibility model estimates the visibility of each image pixel and expresses it as a probability. The output of the model is a visibility map containing the probability of visibility of each pixel. The overall architecture of the model mimics the anatomy of the human visual system (HVS), including optical pathway, retinal pathway, contrast sensitivity, and contrast masking of the HVS. The proposed model also considers the effects of the display model and the ambient light, which are important but often ignored in previous image quality assessment metrics.

The remainder of the paper is organized as follows. We describe the psychological properties of HVS in Section 2 and the proposed model in Section 3. The experimental setup is described in Section 4, and the experimental results are presented in Section 5. Finally, the conclusion is drawn in Section 6.

2. CONTRAST SENSITIVITY AND CONTRAST MASKING

Most of the research in vision science indicates that human visual perception varies according to the viewing condition. Specifically, image details are not detectable by HVS when their contrast is low. The contrast threshold of visual stimuli is found to depend on spatial frequency of the stimuli, which is known as contrast sensitivity function (CSF) [7], which was found to vary with the luminance levels [8]. Mantiuk *et al.* [2] systematically assessed the contrast sensitivity of different spatial frequencies in various luminance levels and constructed a model that covered more conditions than previous studies, and, is thus more suitable for the prediction of visibility. We used the CSF model they proposed to predict the visibility of each pixel of an image.

Besides the research on sensitivity to visual stimuli with specific spatial frequencies, vision scientists also discovered a phenomenon named contrast masking [9], in which the visibility of a sine-wave grating signal is shown to be attenuated by a simultaneously presented grating mask with different spatial frequencies. When the frequencies of the target and the mask are similar [9], and the contrast of the mask is high [10], the maximum masking effect is reached. However, when the contrast of the mask is low, there is a facilitation effect on target detection known as the pedestal effect [11]. Foley [12] modified the model proposed by Legge *et al.* [10], capturing both the masking and the facilitation properties of visual system in detecting signals of

different spatial frequencies, and at the same time considered the different facilitation effects in various mask sizes and selective facilitation effects in the same grating orientation. The model is supported by behavioral data and is consistent with physiology findings [13].

In this work, in addition to the CSF model [2], we use Foley's contrast masking model [12] to identify pixels that are perceived differently from their surrounding pixels.

3. VISIBILITY MODEL

The architecture of the proposed visibility model is shown in Fig. 2. We first model the optical and retinal pathway of HVS, and then the contrast sensitivity and contrast masking properties of HVS. Finally, the probabilities of visibility of image pixels are computed through a psychometric function. Details of each component in the architecture are described in this section.

3.1. Display model

Although it is the displayed image that should be evaluated, in practice, however, the image available for quality assessment is the source image. Therefore, we estimate the luminance of the displayed image from the following display model [14]:

$$L_d = L_b + p^\gamma (L_m - L_b) + L_a, \quad (1)$$

where L_d is the luminance value of the displayed image, L_b is the luminance of the black level of the display, L_m is the maximum luminance of the display, p is the normalized pixel value of the source image, γ is the gamma parameter, and L_a is the ambient light reflected from the display panel. L_a is related to the ambient illuminance E_a (in lux) by

$$L_a = \frac{k}{\pi} E_a, \quad (2)$$

where k is the reflectivity of the display panel. Note that the values of all parameters are measured with respect to the actual display used.

3.2. Optical and retinal pathway

The model proposed by Mantiuk *et al.* [2] is adopted to model the characteristic of the optical and retinal pathway of

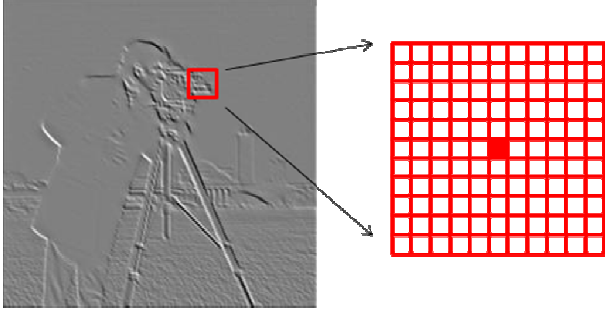


Fig. 3. Illustration of the contrast determination: The contrast of the central pixel in the cropped block shown on the right is the difference between the pixel value of the central pixel and the mean of the remaining pixels in the cropped block.

HVS. The architecture of this model mimics the anatomy of these pathways. Step by step, it models the disability glare effect [15], the expected fraction of light sensed by each type of photoreceptors (L-, M-cones and rods), and the luminance masking of these photoreceptors. The output of this model is a joint cone and rod achromatic response A , which is computed by:

$$A = R_L + R_M + R_R, \quad (3)$$

where R_L , R_M , and R_R , respectively, denote the responses of L-cones, M-cones, and rods. The equal weighting of R_L and R_M is motivated by the fact that L- and M- cones contribute approximately equally to the perception of luminance. Although a more accurate model should also consider inhibitive interactions between rods and cones, this model is a sufficient approximation for our purpose.

3.3. Contrast sensitivity function and contrast masking

We use CSF [2] along with the contrast masking model [12] to predict the visibility of each pixel in an image.

Since the contrast sensitivity and the contrast masking mechanisms of HVS are selective to narrow ranges of spatial frequencies and orientations [16], we adopt the steerable pyramid [17] for image decomposition. Specifically, the input image is decomposed into four orientation bands and the maximum possible number of spatial frequency bands at the given image resolution. Let $B_{f,o}$ denote the f -th spatial frequency band and o -th orientation of the steerable pyramid, the contrast $C_{f,o}(x, y)$ of each pixel $B_{f,o}(x, y)$ is obtained by

$$C_{f,o}(x, y) = B_{f,o}(x, y) - A(B_{f,o}(x, y)), \quad (4)$$

where A is an operator that computes the mean of the neighboring pixels of $B_{f,o}(x, y)$ in the 11×11 window centered at $B_{f,o}(x, y)$ (See Fig. 3).

The contrast sensitivity $T_{f,o}$ for each spatial frequency and orientation band is computed as in [2]:

$$T_{f,o} = \sqrt{C^{2p}(f, o, L_A) + M^2(f, o, L_A)}, \quad (5)$$

where C is the contrast sensitivity, M accounts for the masking effect, p is a tuning factor empirically determined to be 3.5 [2], and L_A is the adapting luminance. Different from the original method [2] that uses the photopic luminance ($R_L + R_M$) of each pixel to approximate L_A , we use the photopic luminance of the actually measured ambient illuminance level E_a to approximate L_A .

Considering the contrast sensitivity $T_{f,o}$, we compute the normalized contrast $N_{f,o}(x, y)$ for each pixel as:

$$N_{f,o}(x, y) = \frac{C_{f,o}(x, y)}{T_{f,o}}. \quad (6)$$

3.4. Visibility metric

We compute the probability of visibility for every pixel in the input image by combining the information from each frequency-orientation band. This process is the same as that described in HDR-VDP-2 [2]. Specifically, the normalized contrast values are first transformed to the corresponding probability values through a psychometric function:

$$P_{f,o}(x, y) = 1 - \exp(\log(0.5)N_{f,o}(x, y)), \quad (7)$$

which is derived on the basis of the report from Daly [18]. After that, the probability values are summed across all bands and orientations to obtain the overall probability for all orientation and frequency-selective mechanisms of HVS. The summation is computed as:

$$P_{map} = 1 - \prod_{(f,o)} (1 - P_{f,o}). \quad (8)$$

4. EXPERIMENTAL SETUP

Considering the display model calibration (see Section 3.1), we have to determine the values of γ , L_m , and L_b , before applying (1) to compute the displayed pixel luminance. A ViewSonic VX912 display with backlight intensity ranging from 45.9 to 161.9 cd/m^2 is used. We measure the luminance for each normalized pixel value by a luminance meter (LAIKO DT-101), and determine γ , L_m , and L_b by linear regression under the uniform backlight assumption that every pixel in the display are the same. As shown in Fig. 4, the resulting γ , L_m , and L_b are 2.25, 191.33 and 0.25, respectively.

In our experiments, we use the proposed visibility model to estimate the visibility loss of image details due to the dim backlight of the LCD. The reference and target images are the same test image but illuminated with full backlight level and 10% backlight level respectively. The difference between the estimated probabilities of visibility

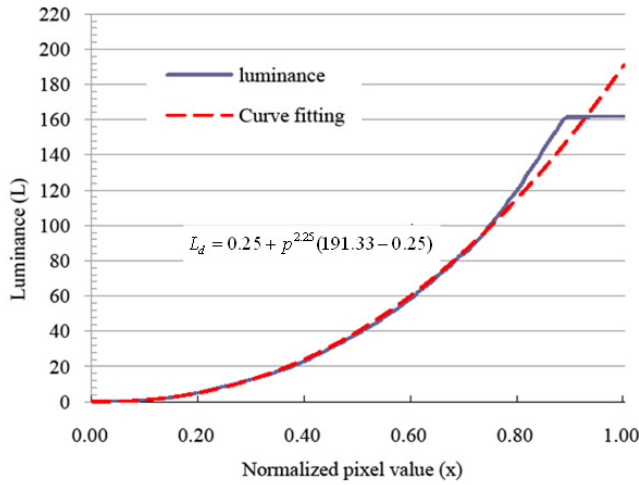


Fig. 4. Determining the maximum backlight L_m , luminance of the black level of the display L_b , and gamma correction coefficient γ by linear regression. In this example, $L_m=191.33$, $L_b=0.25$ and $\gamma=2.25$.

of the two images is used to model the visibility loss of image details.

The viewing condition is as follows: 1) the distance between the viewer and the LCD display is 0.5 m; 2) the display diagonal size is 19 inch; 3) the resolution of the display is 1280 x 1024 (in pixel); 4) all test images have the same size, 320 x 240 (in pixel).

5. RESULTS AND DISCUSSION

The results of three test images are shown in Figs. 5–7. Test images with full backlight are shown in Figs. 5(a)–7(a), while those with 10% backlight in Figs. 5(b)–7(b). Figs. 5(c)–7(c) show the visibility difference map, with colors blue, cyan, green, yellow, and red corresponding to visibility difference values 0, 0.1, 0.2, 0.3, 0.4 or above, separately. In Figs. 5(d)–7(d), we highlight pixels with high visibility difference, which in turn suggests high probability of quality loss, as red dots.

It can be seen that, illuminated with dim backlight, Fig. 5(b) suffered great details loss, especially in dark areas. At the top-left corner and the right side of Fig. 5(a), for instance, the texture visible in the background becomes invisible in Fig. 5(b). This fact is accurately predicted by the visibility difference map (see Fig. 5(c)), in which the corresponding dark areas exhibit high probability of visibility loss. Red dots in Fig. 5(d) denote visible pixels that have high probability of becoming invisible when illuminated with dim backlight.

Note that while the tilted rod on the bottom-right corner of Fig. 5(a) is almost invisible in Fig. 5(b), it should not be considered as having great visibility loss. The reason is that even in full backlight, the rod’s texture is too dark to be clearly discerned, so the rod’s being invisible while being

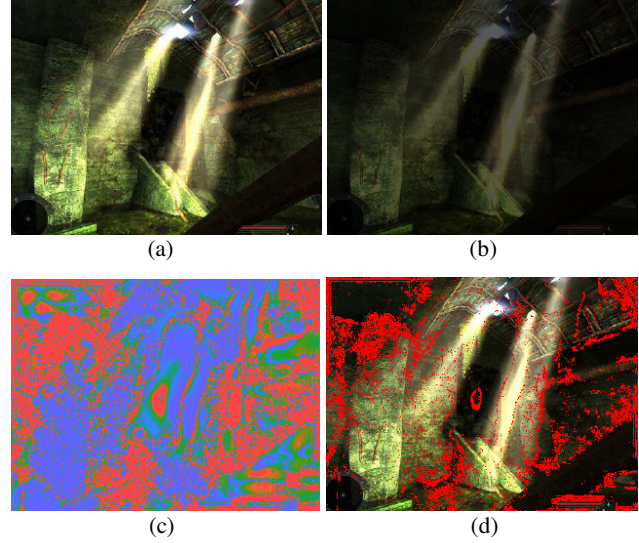


Fig. 5. Prediction of visibility loss for a screenshot from the game “Far Cry”. (a) Original image illuminated with full backlight. (b) Original image illuminated with dim backlight. (c) Visibility difference map. Colors blue, cyan, green, yellow, and red correspond to visibility difference values 0, 0.1, 0.2, 0.3, 0.4 or above, separately. (d) Prediction of visibility loss due to dim backlight. All visible pixels that become invisible are marked in red.

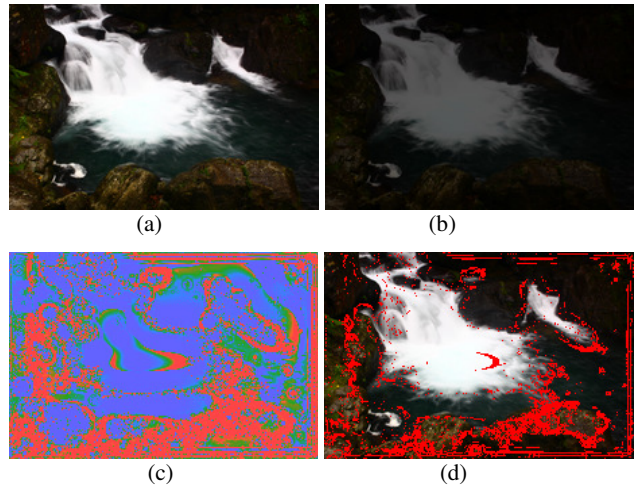


Fig. 6. Prediction of visibility loss for “Pond”. (a) Original image illuminated with full backlight. (b) Original image illuminated with dim backlight. (c) Visibility difference map. Colors blue, cyan, green, yellow, and red correspond to visibility difference values 0, 0.1, 0.2, 0.3, 0.4 or above, separately. (d) Prediction of visibility loss due to dim backlight. All visible pixels that become invisible are marked in red.

illuminated with dim backlight does not ruin the image quality. Since we aim to target the quality loss caused by the visibility loss instead of the visibility itself, invisible pixels that are invisible in the first place should not be considered. It can be seen in Fig. 5(d) that our method faithfully reflects this fact.

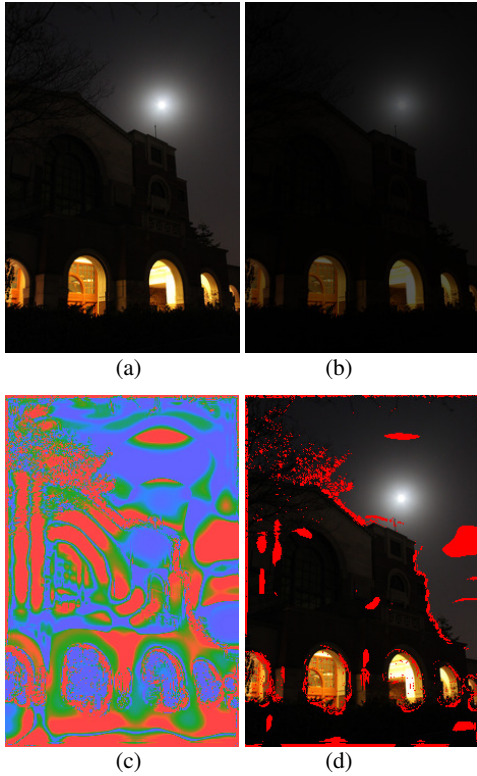


Fig. 7. Prediction of visibility loss for “Moonlight”. (a) Original image illuminated with full backlight. (b) Original image illuminated with dim backlight. (c) Visibility difference map. Colors blue, cyan, green, yellow, and red correspond to visibility difference values 0, 0.1, 0.2, 0.3, 0.4 or above, separately. (d) Prediction of visibility loss due to dim backlight. All visible pixels that become invisible are marked in red.

For other test images, the method also works properly. In Fig. 6(a), the rock visible at the bottom side becomes hardly visible in Fig. 6(b). This fact is in accordance with the predicted result in Fig. 6(d). Also, at the top-left corner of Fig. 7(a) and Fig. 7(b), where the tree branches turn from visible to invisible, the corresponding visibility loss is detected, shown in Fig. 7(d).

Although most of the prediction for visibility loss is accurate, there are times when the prediction fails. For example, see the false prediction right above the moon in Fig. 7(d). These erroneous predictions often occur near the transition of bright and dark area, and can be explained by ringing artifacts. Since our method primarily deals with the image’s pyramid sub-bands, ringing artifacts tend to appear while performing image reconstruction. As a result, these artifacts appear near sharp transitions of the pixel brightness or the boundary of the image (see the right side of Fig. 6(d)).

We also made a performance comparison between the HDR-VDP-2 [2], DRIM [6], and the proposed method. For the HDR-VDP-2 and DRIM metrics, the image illuminated with full backlight is used as the reference image, and the

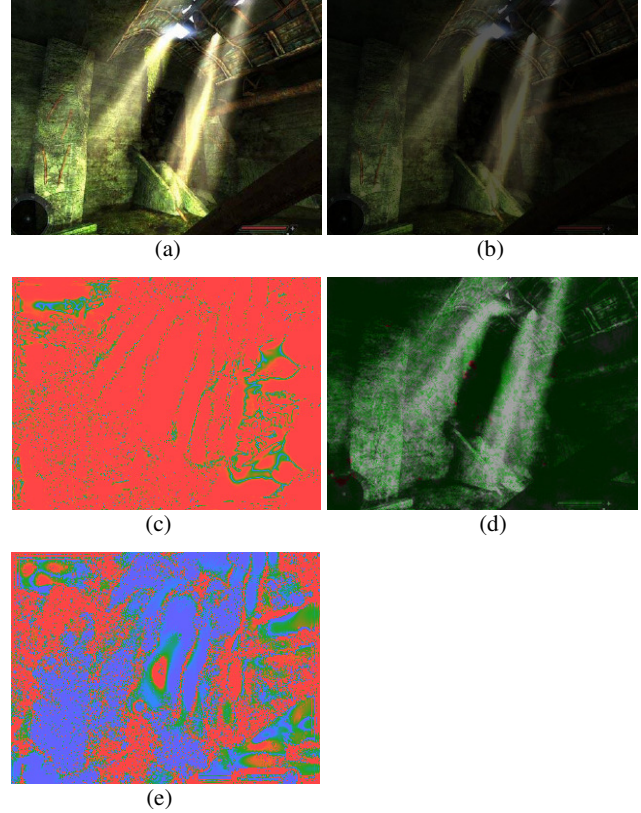


Fig. 8. Performance comparison. The visibility difference between image (a) and image (b) is predicted by metrics (c) HDR-VDP-2 [2], (d) DRIM [6], and (e) the proposed method.

image illuminated with dim backlight is used as the test image. The results are shown in Fig. 8. The luminance value of Fig. 8(b), compared with Fig. 8(a), clearly decreases. This luminance reduction causes details loss in the top-left and the right side of the image, but the details in other parts of the image are still visible. For this pair of images, the HDR-VDP-2 metric detects pixels with luminance reduction (the red pixels in Fig. 8(c)). However, from Fig. 8(c), it is hard to tell if an image region becomes invisible or not. The DRIM metric detects pixels with contrast loss (the green pixels in Fig. 8(d)). It can be seen that most of the detected pixels have low intensity values. This is because the dim pixels lose more contrast than the bright pixels as the backlight dims. From Fig. 8(d), we can see that though the green pixels cover most image regions with details loss, the result is not accurate enough because the contrast loss does not necessary imply the visibility loss. The proposed method, on the other hand, detects the pixels with visibility difference (the red pixels in Fig. 8(e)). The main difference between our result and that of DRIM metric is that the pixels that are invisible even when illuminated with full backlight are excluded in our result. For example, some pixels in the top-left corner of the image are invisible both in Fig. 8(a) and Fig. 8 (b). These pixels are not detected by our method

because the visibility of them does not change due to the dimmed backlight.

6. CONCLUSION AND FUTURE WORK

We have proposed a visibility model to estimate the visibility of images illuminate with dim backlight. The proposed model incorporates two HVS models, the CSF model and the contrast masking model. The former enables our model to estimate the visibility of each pixel in a displayed image, while the latter allows our model to detect invisible pixels caused by high-contrast textures surrounding such pixels. Experimental results show that while the existing methods, HDR-VDP-2 and DRIM, detect the luminance reduction and contrast loss due to the dim backlight, they cannot detect the visibility loss. Our model, in contrast, well predicts the visibility loss of image pixels and thus helps the image quality assessment for dimmed images.

The color vision of HVS is omitted in the proposed model, and thus we plan to extend our model by modeling the human perceptions for different color channels of the image data. We also plan to extend our model to video sequences to make it more general and useful.

7. ACKNOWLEDGEMENT

This project is supported in part by Himax Technologies, Inc., in part by a grant from National Science Council of Taiwan under contract NSC 100-2221- E-002-197-MY3, and in part by a grant from the National Taiwan University under Contracts 10R80919-5 and 10R70609.

8. REFERENCES

- [1] R. Mantiuk, K. Myszkowski, H. -P. Seidel, "Visible difference predictor for high dynamic range images," in *Proc. of IEEE International Conference on Systems*, pp. 2763-2769, 2004.
- [2] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. on Graphics*, vol. 21, no. 3, 2011.
- [3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Processing*, vol. 13, no. 4, pp. 600- 612, 2004.
- [4] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. on Image Processing*, vol. 15, no. 2, pp. 430- 444, 2006.
- [5] D. M. Chandler, S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. on Image Processing*, vol. 16, no. 9, pp. 2284-2298, 2007.
- [6] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H. P. Seidel, "Dynamic range independent image quality assessment," *ACM Trans. on Graphics*, vol. 27, no. 3, pp. 1-10, 2008.
- [7] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiol.*, vol. 197, pp. 551-566, 1968.
- [8] M. A. Losada, R. Navarro, and J. Santamaria, "Relative contributions of optical and neural limitations to human contrast sensitivity at different luminance levels," *Vision Research*, vol. 33, no. 16, pp. 2321-2336, 1993.
- [9] C. F. Stromeyer III and B. Julesz, "Spatial frequency masking in vision: critical bands and spread of masking," *J. Opt. Soc. Am.*, vol. 62, pp. 1221-1232, 1972.
- [10] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Am.*, vol. 70, pp. 1458-1471, 1980.
- [11] J. Nachmias and R. V. Sansbury, "Grating contrast discrimination may be better than detection," *Vision Res.*, vol. 14, pp. 1039-1042, 1974.
- [12] J. M. Foley, "Human luminance pattern-vision mechanisms: masking experiments require a new model," *J. Opt. Soc. Am. A*, vol. 11, pp. 1710-1719, 1994.
- [13] K. K. De Valois and R. H. B. Tootell, "Spatial-frequency specific inhibition in cat striate cortex cells," *J. Physiol.*, vol. 336, pp. 359-376, 1983.
- [14] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 1-15, 2008.
- [15] J. Vos and T. van den Berg, "Report on disability glare," *CIE Research Note*, vol. 135, pp. 1, 1999.
- [16] R. De Valois, D. Albrecht, and L. Thorell, "Spatial frequency selectivity of cells in macaque visual cortex," *Vision research*, vol. 22, no 5, pp. 545-559, 1982.
- [17] E. Simoncelli and W. Freeman, "The steerable pyramid: a flexible architecture for multi-scale derivative computation," in *Proceedings., International Conference on Image Processing*, IEEE Comput. Soc. Press, vol. 3, pp. 444-447, 2002.
- [18] S. Daly, "The Visible Difference Predictor: An Algorithm for the Assessment of Image Fidelity," *Digital Images and Human Vision*, MIT Press. ch., pp. 179-206, 1993.