Calibration of the Visual Difference Predictor for Estimating Visibility of JPEG2000 Compression Artifacts in CT images

Kil Joong Kim*[a], Rafal Mantiuk[b, c], Kyoung Ho Lee[a], Wolfgang Heidrich[b]
[a]Dept. of Radiation Applied Life Science, Seoul National Univ., 28 Yongon-dong, Chongno-gu, Seoul, 110-744, Korea; [b]Dept of Computer Science, Univ. of British Columbia, 201-2366 Main Mall, Vancouver, BC, Canada; [c]School of Computer Science, Bangor Univ., Dean Street, LL57 1UT, United Kingdom

ABSTRACT

Many visual difference predictors (VDPs) have used basic psychophysical data (such as ModelFest) to calibrate the algorithm parameters and to validate their performances. However, the basic psychophysical data often do not contain sufficient number of stimuli and its variations to test more complex components of a VDP. In this paper we calibrate the Visual Difference Predictor for High Dynamic Range images (HDR-VDP) using radiologists' experimental data for JPEG2000 compressed CT images which contain complex structures. Then we validate the HDR-VDP in predicting the presence of perceptible compression artifacts. 240 CT-scan images were encoded and decoded using JPEG2000 compression at four compression ratios (CRs). Five radiologists participated to independently determine if each image pair (original and compressed images) was indistinguishable or distinguishable. A threshold CR for each image, at which 50% of radiologists would detect compression artifacts, was estimated by fitting a psychometric function. The CT images compressed at the threshold CRs were used to calibrate the HDR-VDP parameters and to validate its prediction accuracy. Our results showed that the HDR-VDP calibrated for the CT image data gave much better predictions than the HDR-VDP calibrated to the basic psychophysical data (ModelFest + contrast masking data for sine gratings).

Keywords: visual difference predictor, calibration, CT image, human observer, image compression, JPEG2000

1. INTRODUCTION

Visual difference predictors (VDPs) are computational metrics for estimating visibility of differences between a pair of images.[1, 2] Such predictors are often based on the models of the human visual system (HVS). Since how well the predictors can predict human responses largely depends on the values of the algorithm parameters of the predictors, it is important to calibrate those parameters and to validate the predictors, using an actual human experimental data. For those purposes, many of the VDPs, such as S. Daly's VDP[3] and Watson's model,[4] have used a basic psychophysical data.

However, the basic psychophysical data and typical reference datasets, such as ModelFest,[5] often do not contain sufficient number of stimuli and its variations to test more complex components of a VDP, such as visual masking or spatial integration. Another shortcoming of these datasets is that they tend to focus on a single factor, such as spatial frequency or contrast in isolation, and thus cannot validate if the VDPs produces correct predictions for the combination of several factors. Therefore, complex image data found in practical applications are necessary to calibrate the algorithm parameters as well as to validate their performances.

Meanwhile, lossy image compression has been considered as an effective method to reduce a large amount of medical image data generated by modern computed tomography (CT) scanners. Since medical images have very strict quality requirements, the lossy encoded images must not contain any visible compression artifacts. Several researchers[6-9] showed the feasibility of the use of VDPs in predicting the presence of the artifacts in compressed CT images. Because CT images contain complex structures and image compression is one of the major practical applications, radiologists' experimental data for the compressed CT images would be a suitable test dataset for the VDPs.

In this paper we calibrate the Visual Difference Predictor for High Dynamic Range images (HDR-VDP)[10] using radiologists' experimental data for Joint Photographic Experts Group (JPEG) 2000 compressed CT images. Then we

validate the HDR-VDP in predicting the presence of perceptible compression artifacts. In addition, we investigate the importance of individual components of the HDR-VDP, such as the optical transfer function or visual masking. Fonts

2. MATERIALS AND METHODS

2.1 Data preparation

This study and the acquired data have been previously used to compare HDR-VDP predictions with those produced by PSNR.[6] The data, however, was not used to fine-tune HDR-VDP parameters.

This study included 240 abdomen CT images comprising 120 thin-sections (0.67-mm-thick) and 120 thick-sections (5-mm-thick). Using the two different subsets was to introduce heterogeneity in the test dataset in terms of image noise. CT images with a thinner section have more noise caused by fewer photons contributing to the reconstruction of a voxel.[11] The heterogeneity was considered important in measuring the robustness of the VDP in predicting the perceptible compression artifacts, as it is well known that the compression tolerance is significantly affected by image noise level.[9, 12, 13]

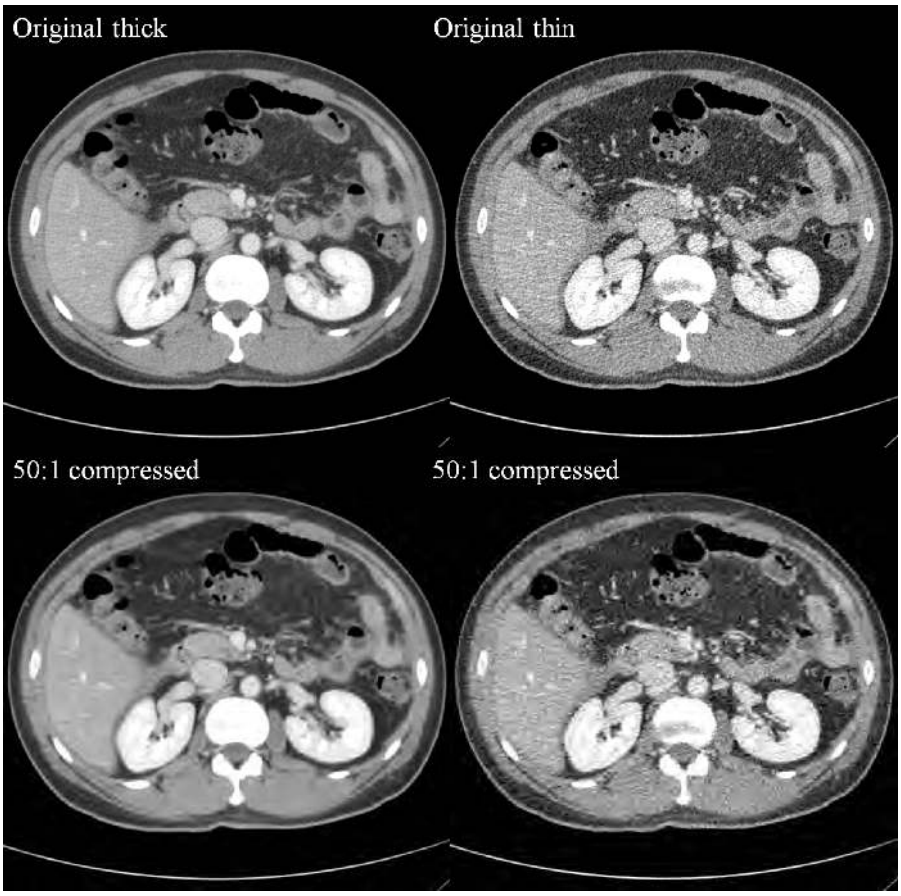

Figure 1. Example of the CT images used in calibrating the HDR-VDP. Top left image is a thick-section image and top right image is the thin-section image at the same z-position. Thin-section image is much noisy. To better present compression artifacts, we used compression ratio of 50:1 for these images, which was not tested in our experiment.

Using a JPEG2000 algorithm (Pegasus Imaging Co., Tampa, FL), each image of thin-section images was compressed to five different compression ratios (CRs): reversible, 4:1, 6:1, 8:1, or 10:1. Similarly, each image of thick-section images was compressed to reversible, 6:1, 8:1, 10:1, or 15:1. These CRs were determined to be below, around, or above the previously reported acceptable compression thresholds.[13] Reversibly compressed images were used as negative controls for the subsequent human visual analysis.

2.2 Human visual analysis

Each of the 1200 compressed images (120 × 5 ratios × 2 subsets) was paired with its original image for the visual comparison. Five radiologists with more than 5 years of clinical experience participated. All radiologists had normal or corrected-to-normal visual acuity. The 1200 pairs were randomly assigned to 20 sessions. The order of the reading sessions was changed for each reader. Sessions were separated by a minimum of 3 hours to minimize reader's fatigue.

Each image pair was alternately displayed on a single monitor in an alternating fashion, which the reader selectively toggled between the two images and could return to the first image as desired. The order of the original and compressed images was randomized. Each reader independently determined if the two images were identical (indistinguishable) or if any detectable difference was present between them (distinguishable). During the visual analysis, the readers were asked to examine an entire image to find any image difference, paying attention to structural details, particularly small vessels and organ edges, and to the texture of solid organs and soft tissues.

Images were displayed using viewing software in a monochrome monitor calibrated according to the Digital Imaging and Communications in Medicine part 14 (Grayscale Standard Display Function).[14] Since the reading distance would affect the readers' sensitivity to the compression artifacts, the reading distance was limited to a range of the readers' habitual clinical reading distances (34-78 cm). Details of the display system and viewing conditions are summarized in Table I. A pixel value of CT image is represented as Hounsfield Units (HU). The number of HU for a material is defined as the relative attenuation coefficient to that of water. Because the number of HU in CT image is represented with 12 bit-depth beyond 8 bits to which general display device is confined, a mapping function (i.e., window setting) is necessary to display these images. For the human visual analysis, the window center and width were set as 20 and 400 HU which were the default values in our clinical practice. The same window setting values were applied to convert the images to 8-bit images in calculating the HDR-VDP to reproduce the display in the human visual analysis.

Table 1. Display system and viewing conditions in human visual analysis.

Display system	
Display resolution	1536 × 2048 pixels
Display size	31.8 × 42.3 cm
Image resolution	1483 × 1483 pixels (stretched using bilinear interpolation from the original image resolution, 512 × 512)
Viewing software	PiViewSTAR (version 5.0.7, SmartPACS, Phillipsburg, NJ)
Luminance	0.42- 398.3 cd/m^2
Viewing conditions	
Ambient room light	30 lux
Reading distance	34-78 cm
Window setting	width, 400 HU; level, 20 HU
Magnification	not allowed
Reading time	not constrained

2.3 HDR-VDP

The HDR-VDP,[10] which is publicly available, is an extension of S. Daly's VDP[3] As modern medical display systems are significantly brighter than general purpose displays, the HDR-VDP that covers a wide range of luminance is more suitable for medical applications than the classical VDP. The HDR-VDP has been reported to accurately reproduce visibility of compression artifacts in CT images[6-9] as compared with the radiologists' data.

The framework of the HDR-VDP is similar to those of other VDPs.[1, 2] First, a pair of images (original and compressed images) is transformed into the luminance format according to the display function, in order to reproduce the real scene shown to observers. Second, while the viewing conditions (Table I) are taken into consideration, the images pass the filters modeling individual components of the HVS, including an optical transfer function (OTF), amplitude nonlinearity, contrast sensitivity function (CSF), multi-band decomposition (i.e., cortex transform), and visual masking (Fig. 2). Third,

a probability-of-detection map is calculated by computing the value of a psychometric function. Each value in that map indicates the probability that a human observer viewing the two images would detect a difference at that pixel location.

We made two modifications of the HDR-VDP for this study. First, the intrer-channel masking effect was taken into consideration based on Foly's work[15] whereas the original HDR-VDP accounts only intra-channel masking. After incorporating the new masking model, the visual difference within a single band is computed as:

$$\Delta C_{k,l}[i,j] = \frac{\left[C1_{k,l}[i,j] - C2_{k,l}[i,j]\right]^p}{\alpha \cdot M_{k,l}[i,j]^q + \beta \cdot X_{k,l}[i,j]^q} \tag{1}$$

where k and l are the band and orientation indices; i and j are pixel indices; $C1$ and $C2$ is the normalized contrast in test and reference image bands (refer to ref [3] for details on notation); and M and X are self and cross-channel masking components. M is computed as: $M_{k,l}[i,j] = min\left(C1_{k,l}[i,j] - C2_{k,l}[i,j]\right)$ (so called mutual masking in the original VDP), and X is the sum of M in all bands except the band k, l; and α, β, p and q are masking parameters, which we want to calibrate.

The second modification is the spatial summation, which is applied before a probability-of-detection map is calculated by summing the probability values across all spatial and orientation selective bands and local neighborhood. A Gaussian filter was applied with a standard deviation (σ). Spatial summation or pooling is a common component of many visual models[4] but was missing in the HDR-VDP.

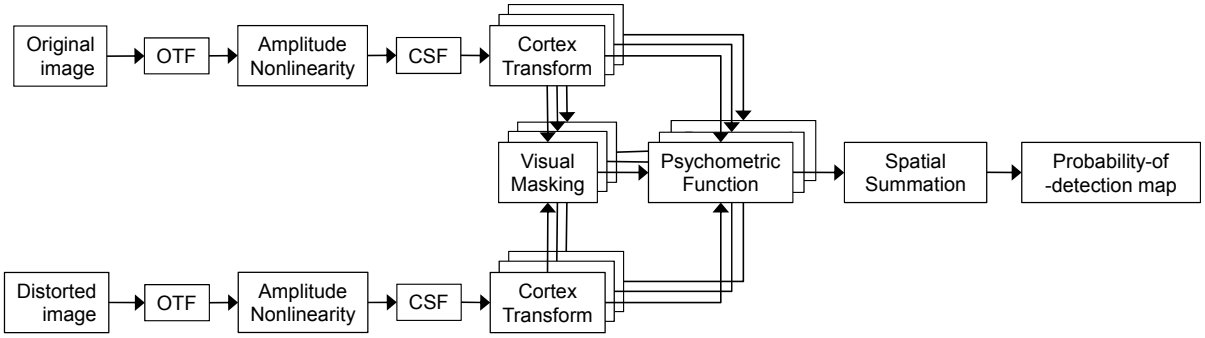

Figure 2. The block diagram of the HDR-VDP.

Since we need a single-valued probability of detection for the entire scene, we compute it as the maximum of the probability map (*Pmax*). However, the drawback of the final probability measure is that due to the asymptotic behavior of the psychometric function (Fig. 3) it will quickly reach its maximum probability value equal one for large super-threshold difference or its minimum value equal zero for contrast below the threshold. To use such a measure for calibration of the HDR-VDP (described in the subsequent section), we need to ensure that it is monotonic with respect to the difference between original and compressed images. This can be achieved by adding a second term that contains the mean square-root error (*MSE*) of the visual channel differences:

$$P_{corr} = P_{\max} + \left(1 - exp\left(-\frac{\left(P_{\max} - 0.5\right)^2}{0.32}\right)\right) \times \sum_{k,l} D[k,l] \tag{2}$$

where $\begin{array}{ll} D[k,l] = MSE[k,l] & if\ MAX(P) > 0.5 \\ D[k,l] = log_{10}(MSE[k,l]) \times 0.01 & if\ MAX(P) < 0.5 \end{array}$,

and where P is a probability map and k and l stand for different band and orientation channels.

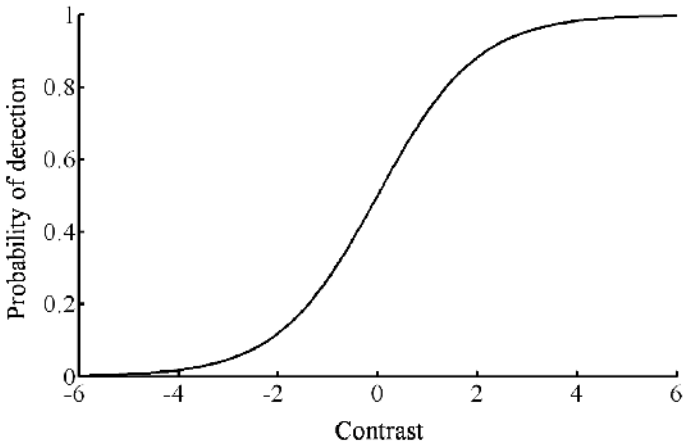

Figure 3. An example of psychometric function.

2.4 Threshold dataset

Since the goal of the HDR-VDP, like other VDPs, is to estimate the visual threshold at which a distortion is just barely visible, a threshold dataset is necessary to calibrate and to validate the HDR-VDP. The calibration would be performed by adjusting the algorithm parameters until the best prediction is found for the threshold dataset as described below.

To derive the threshold dataset, we compressed each of the 240 original CT images to its corresponding threshold CR at which 50% of radiologists would detect compression artifacts. The threshold CR of each image was approximately estimated from the radiologists' responses for its four compressed images (with four CRs) by fitting a psychometric function (Weibull) using the psignifit toolbox (ver. 2.5.6) for Matlab (http://bootstrap-software.org/psignifit) (Fig. 4).

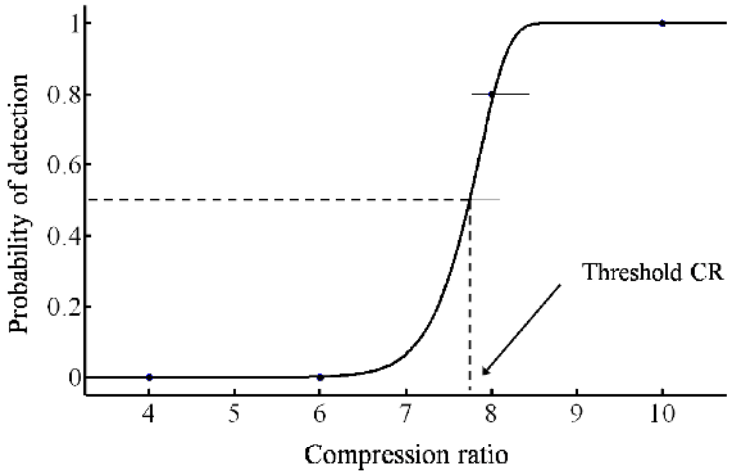

Figure. 4. An example of the procedure to derive the threshold CR of a thin-section CT image. The image was compressed to four different CRs (4:1, 6:1, 8:1, and 10:1) and five radiologists independently determined if the pairs (original and compressed images) were distinguishable or not. A psychometric function is fitted for the four data points. The probability of detection (Y-axis) was defined as the number of radiologists responded as distinguishable divided by five (total number of radiologists). From the fitted psychometric function, the threshold CR was determined as the CR corresponding to 50% probability of detection.

2.5 Calibration of the HDR-VDP

The default algorithm parameters of the HDR-VDP are fitted to Modelfest dataset or derived from the original VDP paper. One of the questions we want to answer was whether these parameters result in good predictions the compressed

CT dataset. To calibrate the HDR-VDP to the CT dataset, we selected five important parameters: the peak sensitivity (denoted as *peak_s*) of the CSF; p, q, α and β parameters in the visual masking; and the standard deviation, σ, of the Gaussian kernel in the spatial pooling.

Of 240 threshold images, 120 images were randomly selected for calibration and the remaining 120 images for validation. Assuming that P_{corr} would be 0.5 for a threshold image if the calibration was perfect, the *prediction_error* for a threshold image was defined as $\left|P_{corr} - 0.5\right|_1$. The five parameters were adjusted until the sum of the prediction_error for the 120 threshold images was minimized. Since the model of the HDR-VDP is not invertible, an iterative optimization method (the simplex search method) was used to find the minimum of the *prediction_error*.

2.6 Validation of the calibration results

To validate the calibration results, we used different procedure, in which contrast factor between original and threshold mages is reduced or amplified until the HDR-VDP result (i.e., P_{corr}) reaches 0.5.

The contrast factor, c, is a multiplier of a difference between the reference image, R, and the test image, T: $T_c = R + C \cdot (T - R)$, where T_c is a test image modulated by the contrast factor. For example, if the calibration is perfect, the contrast equals one, otherwise the contrast factor increases for over-prediction or decreases for under-prediction. The overall metric error for 120 validation dataset was reported as the mean square root contrast difference (denoted as $Cont_{RMS}$) in dB, similarly as in ref [4]. The $Cont_{RMS}$ with (configuration A in Table II) and without (configuration B) calibration was compared.

In addition, we performed receiver operating characteristic (ROC) analysis to test the performance of the HDR-VDP. In this analysis, P_{max} of the HDR-VDP was used. The responses of the five readers were pooled for each of 480 (120 original images of validation dataset × 4 irreversible CRs) compressed images: if three or more readers responded that a pair of images was distinguishable, the pooled response was considered as distinguishable; otherwise, the response was considered as indistinguishable. These pooled readers' responses were used as the reference standard in ROC analysis.

2.7 Investigate the importance of the individual components

To investigate the importance of the individual components, we tested six different configurations of the HDR-VDP by using the same calibration and validation procedures described above: (A) all components of the HDR-VDP and calibrated to the CT data; (B) all components of the HDR-VDP but calibrated to the ModelFest dataset (4) and visual masking data from (15); (C) disabling OTF; (D) single channel model (disabling cortex transform); (E) disabling cortex transform and disabling visual masking; (F) disabling spatial pooling. We compared the results between configurations: (A) vs. (C) to investigate the importance of OTF; (A) vs. (D) to investigate the importance of cortex transform; (A) vs. (E) to investigate the importance of cortex transform and visual masking; (D) vs. (E) to investigate the importance of visual masking; and (A) vs. (F) to investigate the importance of spatial pooling.

3. RESULTS

3.1 Calibration results

The calibration to the CT data increased the performance of the HDR-VDP in terms of $Cont_{RMS}$ (2.51 dB vs. 1.23 dB) (Configuration A vs. B in Table III). The areas under the ROC curve (AUC) of the HDR-VDP calibrated using basic psychophysical data (ModelFest and Folley's masking data) and using CT data were 0.964 and 0.953, respectively, showing significant differences ($p < 0.007$).

3.2 Investigate the importance of the individual components

$Cont_{RMS}$ of the HDR-VDP with all components was not much different from that with no OTF (1.23 dB vs. 1.29 dB). The AUCs with and without OTF were not significantly different (0.964 vs. 0.966, $p = 0.067$) (configuration A vs. C in Table III).

$Cont_{RMS}$ with all components did not tend to be much different from that with no cortex transform (1.23 dB vs. 1.53 dB). The AUCs with and without cortex transform were not significantly different (0.964 vs. 0.960, $p = 0.109$) (configuration A vs. D in Table III). $Cont_{RMS}$ with all components was higher that with no cortex transform as well as without visual masking (1.23 dB vs. 2.46 dB). The AUCs with and without the two components were significantly different (0.964 vs. 0.943, $p < 0.001$) (configuration A vs. E in Table III). In addition, $Cont_{RMS}$ with turning off cortex transform tended to be

higher that without the cortex transform as well as without visual masking (1.53 dB vs. 2.46 dB). The AUCs between those configurations were significantly different (0.960 vs. 0.943, $p < 0.001$) (configuration D vs. E in Table III).

$Cont_{RMS}$ of the HDR-VDP with spatial pooling was not much different from that without spatial pooling (1.23 dB vs. 1.24 dB). However, the AUCs with and without spatial pooling were significantly different (0.964 vs. 0.954, $p = 0.015$) (configuration A vs. F in Table III).

Table 2. Calibrated parameter values of the HDR-VDP.

Configuration	Calibrated parameters of the HDR-VDP						$Cont_{RMS}$ (dB)	AUC
	peak_s	p	q	α	β	σ		
(A) Use all components	4.03	1.82	0.92	280.47	0.00	8.02	1.23	0.964 (0.950, 0.975)
(B) Use default parameters	3.42	2.34	1.62	0.50	147.91	0.20	2.51	0.953 (0.938, 0.966)
(C) Disabling OTF	2.77	1.92	0.98	13.04	8.06	8.25	1.29	0.966 (0.952, 0.976)
(D) Disabling cortex transform	2.03	2.00	27.84	4.94	N/A	6.95	1.53	0.960 (0.945, 0.971)
(E) Disabling cortex transform and visual masking	1.18	N/A	N/A	N/A	N/A	0.97	2.46	0.943 (0.926, 0.957)
(F) Disabling spatial pooling	2.32	3.24	0.73	2859.24	1.79	N/A	1.24	0.954 (0.938, 0.966)

Note—Data in parentheses are the 95% CIs. AUC: area under ROC curve; OTF: optical transfer function.

Table 3. Comparison of configuration results.

Configurations	To investigate	$Cont_{RMS}$ (dB)	AUC
(A) vs. (B)	calibration to CT dataset	1.23 vs. 2.51	0.964 vs. 0.953 ($p < 0.007$)
(A) vs. (C)	OTF	1.23 vs. 1.29	0.964 vs. 0.966 ($p = 0.067$)
(A) vs. (D)	cortex transform	1.23 vs. 1.53	0.964 vs. 0.960 ($p = 0.109$)
(A) vs. (E)	cortex transform and visual masking	1.23 vs. 2.46	0.964 vs. 0.943 ($p < 0.001$)
(D) vs. (E)	visual masking	1.53 vs. 2.46	0.960 vs. 0.943 ($p < 0.001$)
(A) vs. (F)	spatial pooling	1.23 vs. 1.24	0.964 vs. 0.954 ($p = 0.015$)

4. CONCLUSION

In this study, we calibrated the HDR-VDP for the compressed CT images. Both contrast analysis (in terms of $Cont_{RMS}$) and ROC analysis showed that the HDR-VDP calibrated for the CT image set gave much better predictions than the HDR-VDP calibrated to the basic psychophysical data (ModelFest + contrast masking data for sine gratings). This might be due to the fact that the basic psychophysical data may not capture all variations of the stimuli that can be found in complex images. Interestingly, the masking parameters p and q for the HDR-VDP calibrated for the CT dataset are significantly different (p=1.82, q=0.92) from those for the sine-wave masking data (p=2.34, q=1.62). The difference of these parameters determines the slope of the contrast masking function, which is higher (0.9) for the CT dataset than for the sine-gratings (0.72). This is consistent with the learning effect, which causes making to be higher for unfamiliar patterns, such as noise or complex structures in CT images, than for the patterns that can be easily learned (sine gratings) (see p. 195 in ref. 3). Note that the absolute values of these parameters cannot be compared to those reported in the literature (usually p=2.4 and q=2.0), as HDR-VDP uses different contrast units. The prediction errors in terms of ContRMS (1.23 dB) and AUC (0.96) from ROC analysis were very low even though a different set of images was used for validation. This demonstrates that the HDR-VDP is a promising metric for predicting near-threshold visibility in JPEG2000 compressed CT images.

In addition, we investigated the importance of each major component of the HDR-VDP by disabling it and running the calibration procedure. Both contrast analysis and ROC analysis showed consistent improvement in prediction accuracy

after enabling each component, except for the spatial pooling. However, calibration executed on these reduced models often resulted in implausible parameter values, such as q=27.84 or α=280.47, indicating a possible over-training or redundancy of the model. Moreover, although the prediction improvement after adding OTF, spatial pooling or multi-channel decomposition (cortex transform) is minimal for the CT dataset, we found it to be substantial for other datasets. Based on these observations, we cannot claim that any of the HDR-VDP components is redundant for predicting a visibility of a broad range of possible image distortions.

The study presented in this paper demonstrates the importance of using a good variety of stimuli samples when calibrating visual models. The models calibrated using only sine-wave patterns are unlikely to perform well when predicting visibility of blur or noise in images. Creating a visual model that performs consistently well across a broad range of input stimuli may be a daunting task, requiring much more than a few hundred samples for calibration. But this work also shows that it is possible to calibrate a visual model, such as HDR-VDP, to a specific type of distortions, such as JPEG2000 artifacts in CT images, resulting in very good predictions for this class of input images. Therefore, while creating a generic visual model capable of simulating the human visual performance for all kinds of stimuli is probably an unreachable goal at the moment, the specialized models calibrated for a particular task can be reliable tools for a range if imaging applications.

ACKNOWLEDGEMENTS

This work was supported by Engineering Foundation (KOSEF) grant funded by Ministry of Education Science and Technology (MEST), Republic of Korea (No. 2009-0057685).

REFERENCES

[1] M. P. Eckert and A. P. Bradley, "Perceptual quality metrics applied to still image compression," Signal Processing 70, 177-200 (1998).

[2] Z. Wang and A. C. Bovik, Modern image quality assessment, 1st ed. (Morgan & Claypool, San Rafael, 2006).

[3] S. Daly, "The visible differences predictor: an algorithm for the assessment of image fidelity," in Digital images and human vision, edited by A. B. Watson (MIT Press, Cambridge, MA, 1993), pp. 179-206.

[4] A. B. Watson and A. J. Ahumada, Jr., "A standard model for foveal detection of spatial contrast," J Vis 5, 717-740 (2005).

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

[6] B. Kim, K. H. Lee, K. J. Kim, R. Mantiuk, V. Bajpai, T. J. Kim, Y. H. Kim, C. J. Yoon and S. Hahn, "Prediction of perceptible artifacts in JPEG2000 compressed abdomen CT images using a perceptual image quality metric," Acad Radiol 15, 314-325 (2008).

[7] B. Kim, K. H. Lee, K. J. Kim, R. Mantiuk, S. Hahn, T. J. Kim and Y. H. Kim, "Prediction of perceptible artifacts in JPEG2000 compressed chest CT images using mathematical and perceptual quality metrics," Am J Roentgenol 190, 328-334 (2008).

[8] B. Kim, K. H. Lee, K. J. Kim, R. Mantiuk, H. R. Kim and Y. H. Kim, "Artifacts in slab average-intensity-projection images reformatted from JPEG 2000 compressed thin-section abdominal CT data sets," Am J Roentgenol 190, 342-350 (2008).

[9] K. J. Kim, B. Kim, K. H. Lee, R. Mantiuk, H. S. Kang, J. Seo, S. Y. Kim and Y. H. Kim, "Objective index of image fidelity for JPEG2000 compressed body CT images," Med Phys 36, 3218-3226 (2009).

[10] R. Mantiuk, S. Daly, K. Myszkowski and H.-P. Seidel, "Predicting visible differences in high dynamic range images-model and its calibration," Proc Human Vision and Electronic Imaging X, IS&T/SPIE's 17th Annual Symposium on Electronic Imaging, 204-214 (2005).

[11] M. F. McNitt-Gray, "AAPM/RSNA Physics tutorial for residents: topics in CT. radiation dose in CT," Radiographics 22, 1541-1553 (2002).

[12] V. Bajpai, K. H. Lee, B. Kim, K. J. Kim, T. J. Kim, Y. H. Kim and H. S. Kang, "The difference of compression artifacts between thin- and thick-section lung CT lmages," Am J Roentgenol 191, 38-43 (2008).

[13] H. S. Woo, K. J. Kim, T. J. Kim, S. Hahn, B. H. Kim, Y. H. Kim, C. J. Yoon and K. H. Lee, "JPEG 2000 compression of abdominal CT: difference in compression tolerance between thin- and thick-section images," Am J Roentgenol 189, 535-541 (2007).

[14] Digital Imaging and Communications in Medicine (DICOM). Part 14: Gray scale standard display function. medical.nema.org/dicom/2004/04_14pu.pdf. Accessed June 1, 2006

[15] J. M. Foley, "Human luminance pattern-vision mechanisms: masking experiments require a new model," J Opt Soc Am A Opt Image Sci Vis 11, 1710-1719 (1994).