AN ADAPTABLE HUMAN VISION MODEL FOR SUBJECTIVE VIDEO QUALITY RATING PREDICTION AMONG CIF, SD, HD AND E-CINEMA

Abstract

A new highly adaptable model for predicting human vision response is presented for enabling an improved method of predicting subjective video quality. The ability to adapt enables comparison of video with dissimilar image sizes, viewing environments, frame rates, video quality classes, etc. (for example HD vs. SD vs. CIF). Model test results are compared with human response. Responses are from stimuli covering both JND (just noticeable differences) and supra-threshold (extending to near the opposite extreme). Supra-threshold responses compared include adaptation behavior exemplified by nonlinear response responsible for significant sensitivity changes, masking and visual illusions. Given the prediction of visible impairments, simulation of the contextual adaptation that occurs during the training portion of ITU-R BT.500 (subjective assessment methodology) is used for predicting DMOS.

1. 0 Introduction

Previously developed models for predicting subjective video quality have one or both of the following: a) objective impairment measurements, each weighted by estimates of subjective annoyance, and summed to produce a single metric per frame and/or sequence [1] [2] and/or b) an attempt to at least partially mimic human vision system response [3] [4] [5] [6] [7] [8] [9]. However, vast stimulus-response data from vision science literature has remained largely unexploited. Models that have been developed within the vision science community typically only address isolated human vision system behaviors, or are parameterized as to not be readily adaptable to the application of video quality as exemplified by [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22].

As a result, current standards and practices for subjective video quality measurement ignore substantial changes in human vision sensitivities. For example, ITU-T J.144 models such as VQM [1] do not take into account differences in viewing conditions, frame rates, resolutions or nominal quality context. And since J.144 was developed for standard definition (SD), using it for predicting subjective video quality ratings of other formats such as CIF or HD is problematic [23]. There have been reports that use of J.144 methods can be problematic even with some SD video if content and/or impairments are quite different from that used for development and verification of the algorithms.
In addition to human vision perceptual response issues, display device differences and the nominal quality of each format also play important roles [23].

Even for a given format, display and viewing conditions, application dependent nominal quality range (for example, set by bit rate or program content nominal complexity) can vary significantly. This has made necessary a) normalizing subjects to the range of quality of video in ITU-R BT.500 [24] and b) the bit rate and frame rate matrix of categories of VQEG impairment ranges and types respectively [25]. As such, the context of nominal quality range affects the sensitivity and quality rating scales such as MOS and DMOS.

Using the simulation approach shown in the system diagram of Figure 1, video quality rating prediction requires a human vision model that can adapt as the human vision system does, to different displays, viewing conditions, video context and formats, along with a method for adapting quality scales as humans do in subjective quality rating training and conditioning.

Such a human vision model has been developed and patented [26] [27] [28] [29]. An overview of its specification, components, calibration and validation are covered in the remaining sections of this paper.

Tektronix

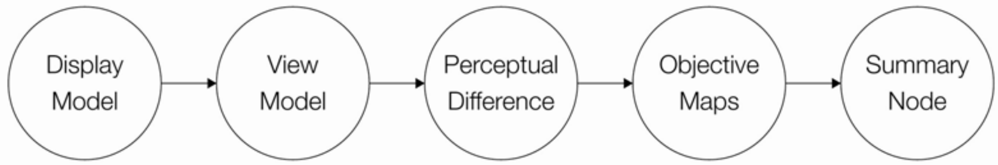

Fig. 1. Simulation system diagram for subjective video quality rating prediction.

2.0 Defining Human Vision Model Response Specifications

The specifications of the model may be formulated from the data available from thousands of published results of experiments in vision science. Stimulus-response pairs become the numeric framework for describing the input-output behavior desired from the human vision model. A good introductory overview of experiments and associated stimulus-response sensitivity analysis is given in [30].

For brevity, the scope of the remaining discussion will be limited to achromatic light (luminance) only. Analysis for chromatic light follows a similar path.

2.1 Adaptation

Consider that responses may be separated into four classes: transparent, linear, fixed (stationary) non-linear and adaptive. Each has counterparts in existing algorithms for predicting subjective video quality ratings.

For example, the point-to-point (pixel) difference terms used in PSNR measurement corresponds to the transparent response class, with each pixel of a reference subtracted from the test, nominally with no prior processing, such as filtering, etc. Hypothetically, for video stimuli with content that corresponds closely with this first class, the point difference used in PSNR would be sufficient for determining relative response. However, in practice, by ignoring the behavior of the other classes PSNR is not generally an accurate predictor of video quality. For example, PSNR interpreted as human vision response to given stimuli is as if human vision responds instantaneously (counter to, for example, [30] [12] [31] [32] [33]) and with perfect acuity (vs. [11] [13] [14] [17] [18] [30] [31] [34] [35] [36] [37] [38] [39] [40] [41]).

Linear response corresponds to cases where fixed (stationary) linear filters applied to the stimulus can mimic, to the degree required, the human vision response. Hypothetically, for video stimuli with content that corresponds closely with the linear class, fixed spatiotemporal filters would be sufficient in determining the sensitivity of perceptual differences. However, linear response does not include such things as the human vision system's change in ability to discern detail in light or dark patches as a function of surround or ambient light level, nor Weber's law. This is as if human vision response had no significant signs of adaptation to mean luminance (vs. [13] [18] [30] [34] [35] [37] [38] [40]), masking (vs. [11] [13] [14] [17] [18] [30] [31] [34] [35] [36] [37] [38] [39] [40] [41]), or sensitivity to image similarities between reference and test (vs. i.e. orientation sensitivity [10] [34] [30]).

Fixed (or stationary) nonlinear response corresponds to cases where two superimposed images (say, for example, two sources of light added), for each point (or pixel) in the image, the response is not equal to the sum of the responses of each image taken alone.

From the measurement standpoint, stimulus-response pairs that are well mimicked using fixed-nonlinear processes such as the commonly used Sobel filter [1, Fig B.10] [2] fall into this class. When combined with linear filtering, this class includes most of the more advanced measurement algorithms of commercially available video quality rating prediction methods. Yet even when combined with filters to account for spatiotemporal response, it does not account for phenomena like flicker vs. brightness, where light of a given intensity appears brighter if it is turned on and off rapidly [42], nor the effect of perceived brightness vs. luminance adaptation [43], nor other temporal aspects of visual illusions such as a phantom third pulse seen when two pulses are used as a stimulus [44].

However, for the general case, human vision response to video is adaptive (dynamically non-linear). Response sensitivities can change by more than an extra order of magnitude beyond the spatiotemporal dynamic range. An example is spatiotemporal contrast sensitivity vs. luminance as illustrated in response for 0.28 nits vs. 91 nits in [45] (Also, see Figures 6, 7 and 8).

These human vision system stimulus-response adaptations correspond to physiological adaptations of various anatomical structures comprising the human vision system. Adaptations include changes in pupil size [46], photoreceptor response [47], and neuron response [47, p22] [70] [48].

The change in sensitivity with average luminance, such as light and dark adaptation, involves the non-linearity that is consistent with many visual perception phenomena. Phenomena such as enhanced brightness with flicker [42], changes in dynamic responses to step increases [43, Fig 10], after-images [49], visual illusions [44] [50] [51] [52] and extreme sensitivity (i.e. photosensitive epilepsy) [54] are consistent with the types of non-linearity that account for most of the already mentioned adaptation [27, equation 3].

Current standards measuring of video quality (predicting subjective quality) generally do not account for adaptation. As a result, proponents of video quality measurement methods for standardization generally submit a different model (or at least different calibration of the same model) for each combination of resolution and frame rate, to cover SD [1], HD [55] or CIF [56]. Also, comparisons among these standards have been considered to be impractical. Even with special versions of these models, without accounting for adaptations, it is still problematic to get good results over ranges such as the combination of low light levels of digital cinema, the typical home viewing situation and the bright levels of mobile displays outdoors.

2.2. Key Human Vision Stimulus-Response Data Sets

Human vision contrast threshold response (contrast sensitivity) has been tested as a function of spatial frequency and mean luminance [31] [34] [35] [40] [42] [59] [60] spatiotemporal frequency and mean luminance [31] [42] [59] [61] [62], spatial frequency and area [31] [35] [59], surround [63] [32], duration [12] [63], orientation [34] [10], spatial pedestal (masker) contrast and frequency [34], and temporal pedestal contrast and frequency [61].

Human vision supra-threshold response to contrast has been tested as a function of luminance [64], spatial frequency and luminance [40], area and contrast [14], area, contrast and spatial frequency [21], luminance, area and temporal contrast and temporal frequency [53].

Perception of mean luminance has been tested as a function of temporal frequency [42], vs. luminance adaptation [43],

Perception of spatial frequency has been tested as a function of spatial and temporal frequency [39] [65] [66] [50] [51] [67].

Perception of pulses per time interval has been tested as a function of conditions required for the three-flash illusion [44] [52] [68] [69].

These stimulus-response pair sets collectively have been used to specify the response of the human vision model for general video quality. They represent a sampling across the gamut of possible video stimuli and conditions such as the low light levels of digital cinema (front row to back row) to typical home viewing to the bright levels of mobile displays outdoors.

Recently there has been growing acknowledgment of the importance of adaptation for video quality assessment. For example, Johnson [57] describes how color appearance models, which have included adaptation as a key behavior for predicting differences in simple uniform patches of color, could be extended to measure image quality by adding a spatiotemporal filter. However, in this example, filters must be selected based on frame rates or signal content, while calibration, verification and some model component details were not included. Still, it is worth noting the importance of adaptation cited in both [57] and in color appearance models including luminance appearance [58].

3. MODEL COMPONENTS

3.1. Video Quality Rating Prediction System Components

Simulation of display, environment, human vision system and contextual training as in ITU-R BT.500 are performed in succession by respective processing nodes shown in Figure 1. The display model simulates the digital video signal to light conversion. The environment, representing viewing distance, ambient light and so forth, is simulated via a view model between the display and human vision model. The vision model responds to this light as specified in section 2. The objective measures node classifies and measures visible impairments objectively, with the ability to then sum each with corresponding relative annoyance (or relative preference). The summary node extracts single summary measures per frame and/or video sequence. Also, the ITU-R BT.500 training equivalent, which maps the response summary measures to DMOS according to [29] is included in the summary node.

3.2. The Human Vision Model

The human vision model takes into account behavior, including adaptation, represented by the stimulus-response data sets of section 2.2. As described in detail in [27], the adaptive integrator of Figure 2 is used to construct an adaptive spatiotemporal filter of Figure 3, in turn used combined with resolution, viewing distance and frame rate adaptation of [28] shown in figure 4. For full reference measurements, the perceptual difference prediction system of Figure 5 is used, incorporating the adaptive spatiotemporal filter and adaptively weighted difference mechanisms of [26].

3.2.1 The Adaptive Integrator

The primary building block within the adaptive spatiotemporal filter [27] is the adaptive integrator, also known as an adaptive recursive (infinite-impulse response, IIR) low-pass filter shown in Figure 2. The integration time (or area) and corresponding frequency cut-off are controllable, while the mean of the output follows the mean of the input (unity DC gain).

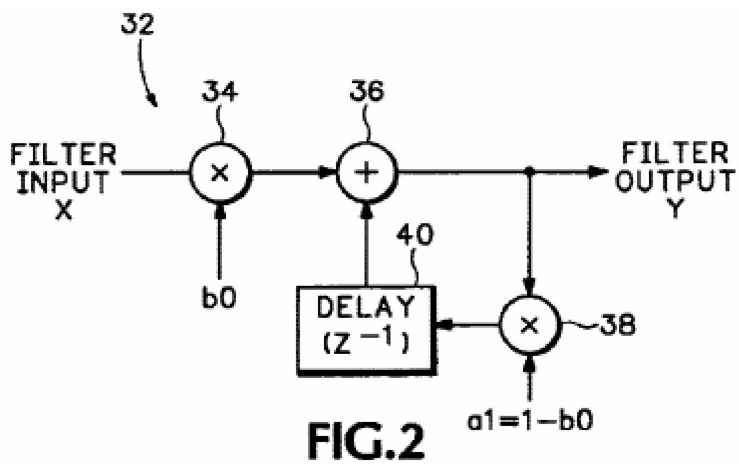

FIG.2

Fig. 2. Adaptive Integrator

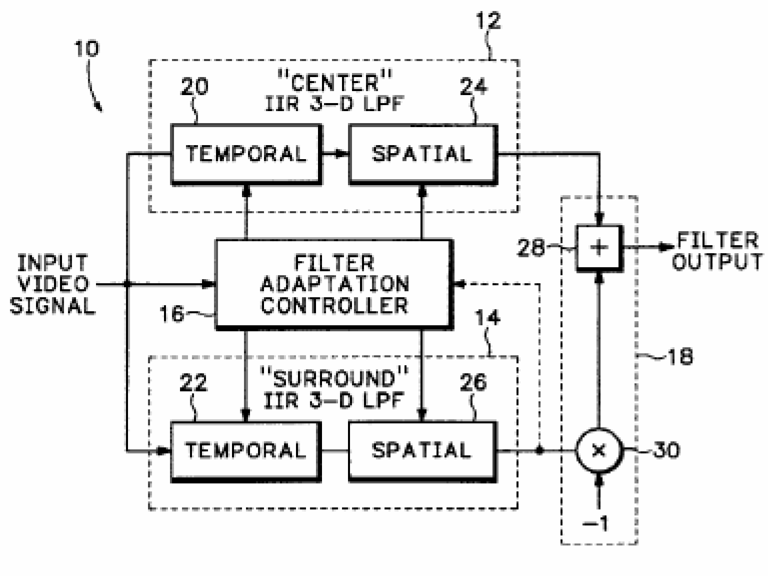

Fig. 3. Adaptive Spatiotemporal filter

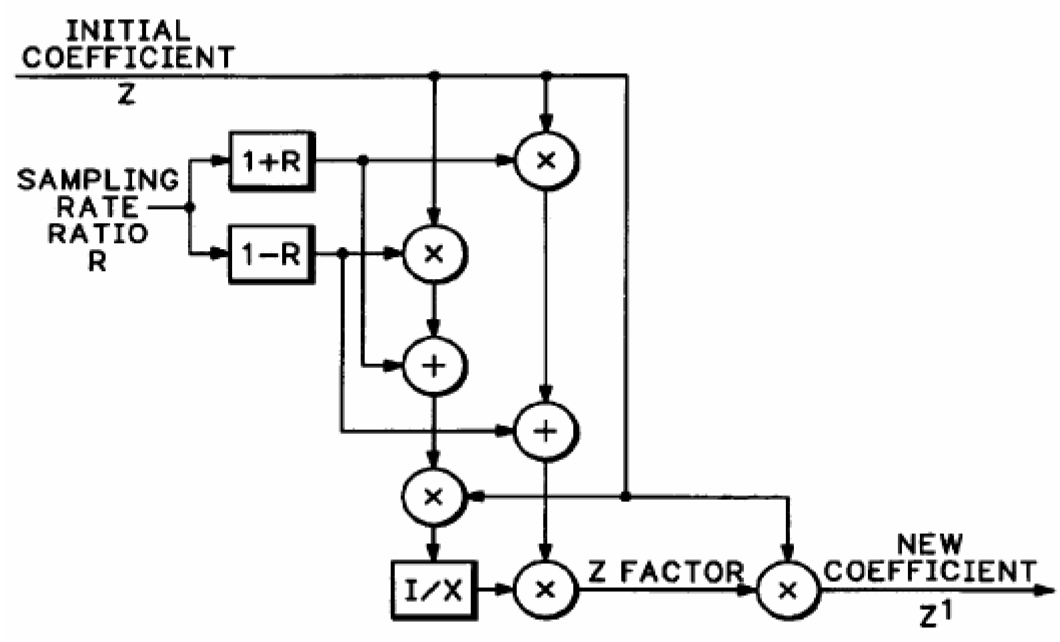

Fig. 4. Mechanism for frame rate, resolution and viewing distance adaptation

3.2.2 The Adaptive Spatiotemporal Filter

This adaptive integrator is used to integrate (and filter) in 4 spatial directions (right, left, up, down) and temporally. The result is a spatiotemporal filter that is tunable in each dimension. Consistent with previous models taking into account center and surround interaction, for example as described in [47] [90] [26], two spatiotemporal filters are used: one for the center and one for the surround. The surround spatiotemporal response is used to both subtract from the center and tune the center spatiotemporal response. In addition, the surround spatiotemporal response also alters its own response via feedback to the frequency controls, but much more slowly than for the center, consistent with longer term adaptation such as long term light and dark adaptations [43], after-images [42], and other long-term effects.

Nominally the difference of the two three-dimensional low-pass filter responses results in a band-pass response.

To allow for calibration, there are 10 controls. Eight are used for direct threshold spatiotemporal response control. Horizontal and vertical dimensions use the same controls. The spatial and temporal dimensions of both center and surround each have two filter controls. One control is for both baseline frequency cut-off (corresponding to integration time or area). Another control is for the frequency response adaptation sensitivity control (how much does the integration time or area change with the adaptation input such as average surround luminance). Two are for control of the transition between threshold and supra-threshold response (one for spatial and one temporal).

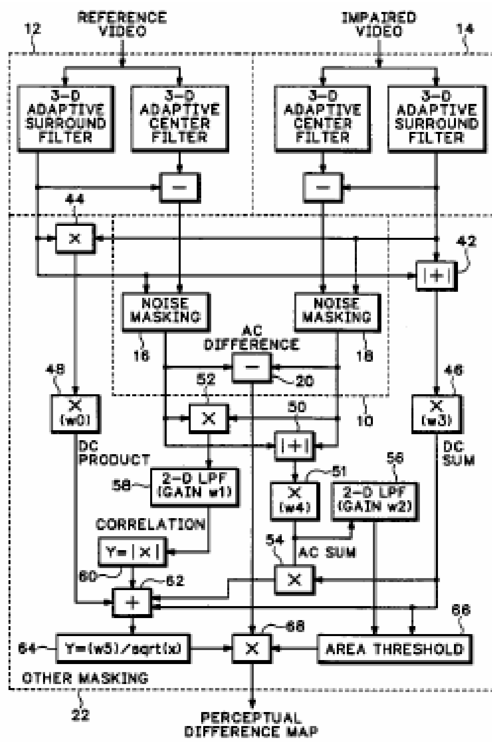

Fig. 5. Perceptual difference prediction system

3.2.3 Adapting To New Frame Rates, Resolutions and View Distances

After calibration, these adaptive spatiotemporal filters can be automatically reconfigured to account for different viewing distances and frame rates. The mechanism for this is filter coefficient recalculation for the corresponding sample rate ratio as described in [28].

3.2.4 Remaining HVS Model Structure

In addition to the adaptive spatiotemporal filter, other model components are used to take into account Weber's law, perceptual differences between correlated vs. uncorrelated images and other behavior including types of masking. Details are given in [26]. There are nine associated calibration parameters. Thus, for the human vision perception model, there are 19 calibration parameters.

4. MODEL CALIBRATION AND VALIDATION

4.1. Human Vision Model Calibration and Validation

For model calibration of the 19 parameters, light stimulus is simulated and model response is compared with expected responses from vision science literature. Calibration has used approximately 1400 light stimuli pairs (test and reference) of which 968 pairs of stimuli cover threshold contrast sensitivity using 176 from [34], 132 from [35], 33 from [10], 176 from [12], 56 from [31], 80 from [59], 78 from [61], 40 from [62], 162 from [45] and 35 from [40]. Supra threshold response was checked against 413 sequence pairs consisting of 39 from [14], 1 from [44], 328 from [53] and 35 from [40]. Care was taken to ensure units were normalized to nits (cd/m^2) consistent with natural (non-stabilized, binocular) vision.

The comparison of model response to human vision response includes the calculation of error standard deviation within specific categories (for example threshold or supra-threshold) and overall (all categories included), error histograms and fitted Gaussian curves, and visual inspections of plotted curves along with the original data sets (See Figures 6-8).

Since the data is gathered from different people and has inherent variance, it is not expected that one model would identically fit all people simultaneously. Instead, calibration has nearly minimum standard deviation error and the error distribution is approximately Gaussian (Figure 9).

Based on errors of responses, for example, spatial frequency response vs. luminance, the corresponding calibration parameter is modified to reduce the error. For each parameter changed, the corresponding effected responses are iteratively measured and analyzed for directing further parameter changes for reduced error.

Validation is achieved if calibration achieves error standard deviation within the reference standard deviation expected from the data sets taken from the literature. Different data sets have different standard deviations, and many do not have enough samples for each stimulus to be able to compute standard deviation. However, low sensitivity responses generally have higher standard deviations. Therefore, error standard deviations can be classified roughly by nominal response sensitivities, even if the corresponding standard deviations are not known.

4.2. Summary Node Calibration and Validation

For model calibration of DMOS prediction via ITU-R BT.500 training, the training sequence pairs are used to measure the corresponding perceptual best case and worst case differences. These are then used to map to DMOS using the method described in [29].

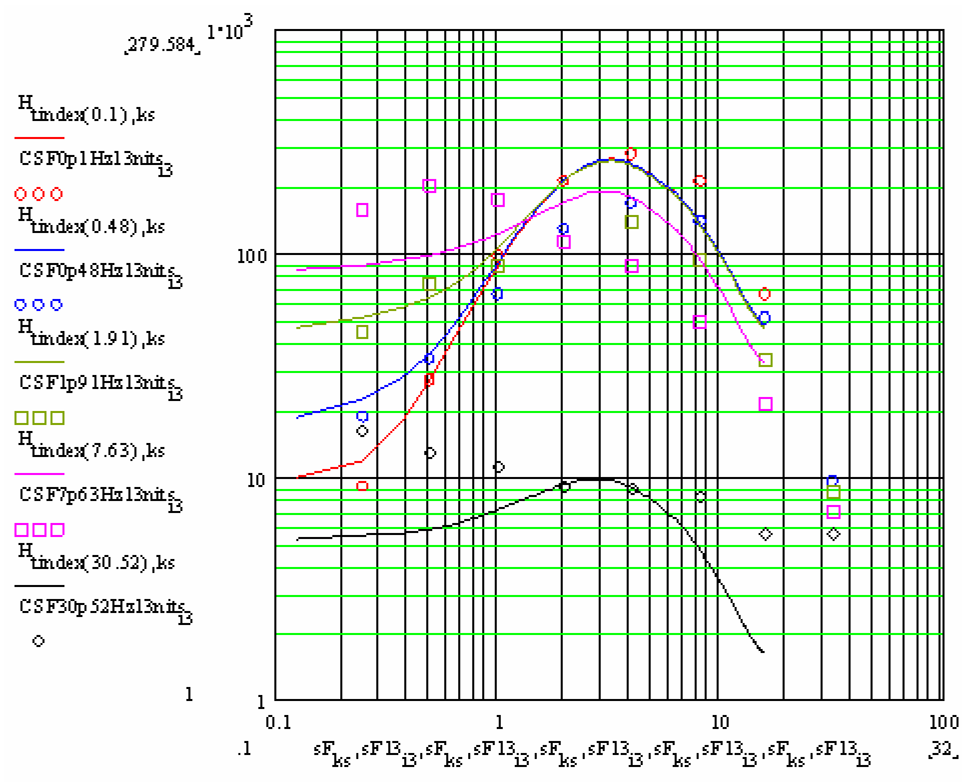

Fig 6. An example of model (solid) vs. human vision response (points) [31]: Pedestal luminance is 13 nits. Horizontal: spatial frequency, vertical: contrast sensitivity, curves: temporal frequencies.

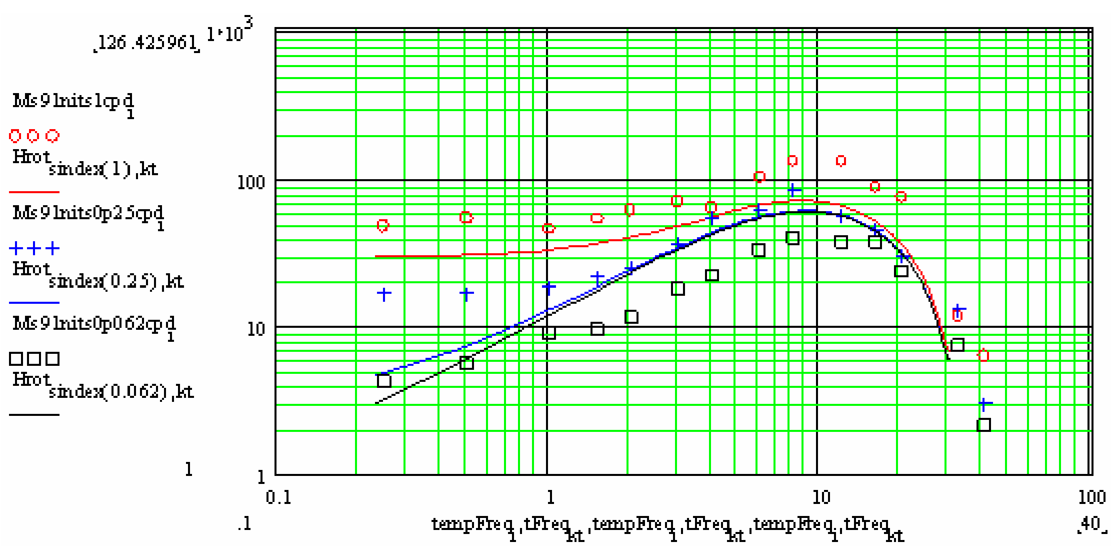

Fig 7. The same as Fig. 6 except for 91 nits [45] and the temporal frequency varies along the horizontal while each curve represents a different spatial frequency.

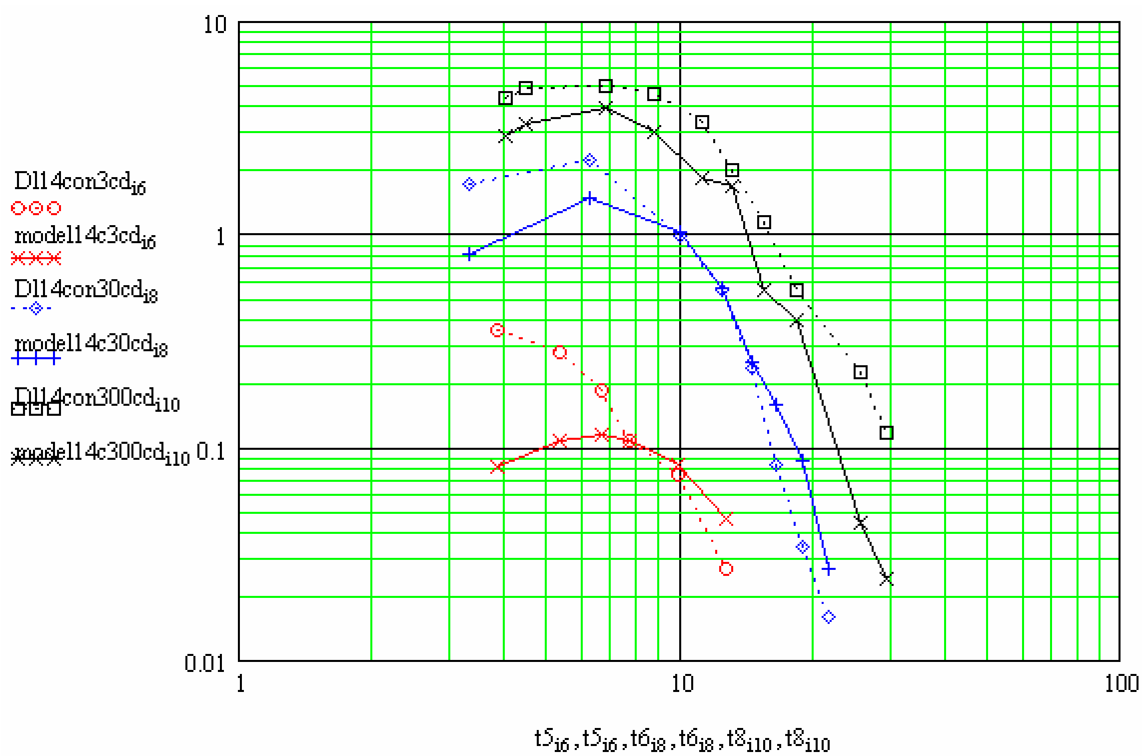

Fig 8. Relative perceived temporal contrast for constant supra-threshold (14%) contrast vs. temporal frequency at different mean luminance levels. Dashed lines connect HVS data from [53], solid for HVS model.

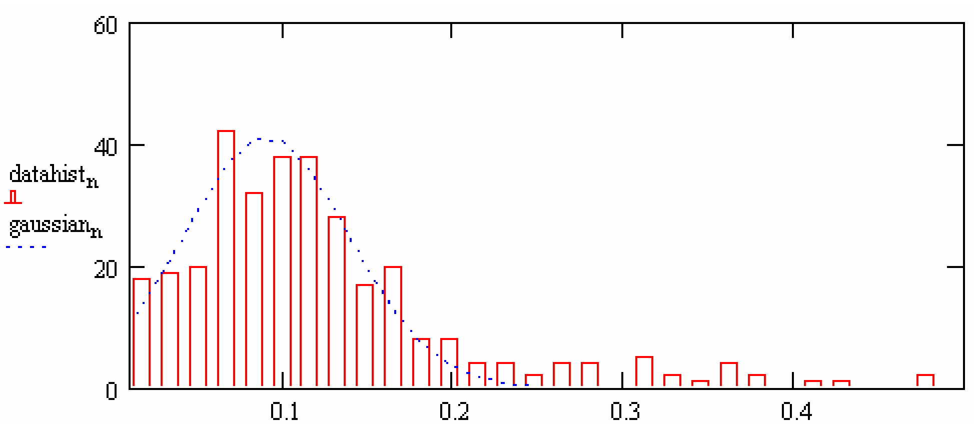

Fig 9. Example of Gaussian distribution of model vs. HVS measured response for threshold data (corresponding to 0.1 reference as per Cannon [14]).

Validation is achieved if subsequent DMOS prediction has error standard deviation within the expected DMOS standard deviation. Expected DMOS standard deviation is calculated as the standard deviation of the opinion score differences divided by square root of the number of subjects.

5. SUMMARY

A highly adaptable model is required for an improved method of predicting subjective video quality allowing the comparison of dissimilar displays, image sizes, viewing environments, frame rates and video quality classes. The test results of such a model have been compared with human vision perceptual response, including adaptation behavior exemplified by nonlinear response responsible for significant sensitivity changes, masking and visual illusions. Simulation of contextual adaptation of the training portion of ITU-R BT.500 was also discussed.

6. REFERENCES

[1] ITU-T J.144 (& ITU-R BT.1683), “Recommendation J.144 (Rev.1) - Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference,” 2004.

[2] T1A1.3 Working Group, “Digital Transport of One-Way Video Signals – Parameters for Objective Performance Assessment,” ANSI, T1.801.03-2003.

[3] S. Daly, “The visible differences predictor: an algorithm for the assessment of image fidelity,” in Digital Images and Human Vision, ed. Andrew B. Watson (MIT Press, Cambridge, MA 1993), pp. 162-206.

[4] J. Lubin, “A Visual Discrimination Model for Imaging System Design and Evaluation,” in Vision Models for Target Detection and Recognition, ed. Eli Peli (World Scientific Publishing, River Edge, NJ, 1995), pp. 245-283.

[5] R. Eriksson, B. Andren, K. Brunnstrom, “Modelling the perception of digital images: A performance study,” SPIE Conference on Human Vision and Electronic Imaging, Vol. 3299, Jan. 1998, pp. 88-97.

[6] S.J.P. Westen, R.L. Lagendijk, J. Biemond, “Perceptual Image Quality Based on a Multiple Channel HVS Model,” IEEE ICASSP-95 pp. 2351-2354.

[7] C.J. van den Branden Lambrecht, “A Working Spatio-Temporal Model of the Human Visual System for Image Restoration and Quality Assessment Applications,” IEEE ICASSP-96, May 7-10, Atlanta, GA; pp. 2291-2294.

[8] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, M. Carli, “Two New Full-Reference Quality Metrics Based on HVS,” http://enpub.eas.asu.edu/resp/vpqm2007, VPQM-06, pp.1-4, 2006.

[9] R. Ferzli, L. J. Karam, “A Human Visual System-Based Model For Blur/Sharpness Perception,” http://enpub.eas.asu.edu/resp/vpqm2007, VPQM, pp. 1-4, 2006.

[10] J. Foley, “Human luminance pattern-vision mechanisms: masking experiments require a new model,” Jounal of The Optical Society of America, Vol. 11, No. 6, pp. 1710-1719, June, 1994.

[11] J. Rovamo, O. Luntinen, R. Nasanen, “Modelling the Dependence of Contrast Sensitivity on Grating Area and Spatial Frequency,” Vision Research, Vol. 33, No. 18, pp. 2773-2788, 1993.

[12] O. Luntinen, J. Rovamo, R. Nasanen, “Modelling the Increase of Contrast Sensitivity with Grating Area and Exposure Time,”” Vision Research, Vol. 35, No. 16, pp. 2339-2346, 1995.

[13] P. Barten, “Physical Model For The Contrast Sensitivity of the Human Eye,” Barten Consultancy, 5511 KC Knegsel, The Netherlands.

[14] M. Cannon, “A Multiple Spatial Filter Model for Supra-threshold Contrast Perception,” in Vision Models for Target Detection and Recognition, ed. Eli Peli (World Scientific Publishing, River Edge, NJ, 1995), pp. 88-117.

[15] T. Frese, C. A. Bouman, J. P. Allebach, “A Methodology for Designing Image Similarity Metrics Based on Human Visual System Models,” SPIE Conference on Human Vision and Electronic Imaging, Vol. 3016, Feb. 10-13, 1997, pp. 472-483.

[16] A. Ahumada, Jr., B. Beard, “Image discrimination models predict detection in fixed but not random noise,” Journal of the Optical Society of America, Vol. 14, No. 9, pp. 2471-2476, Sep. 1997.

[17] A. B. Watson, J. A. Solomon, “Model of visual contrast gain control and pattern masking,” Journal of the Optical Society of America, Vol. 14, No. 9, pp. 2379-2391, Sep. 1997.

[18] A. M. Rohaly, G. Buchsbaum, “Global spatiochromatic mechanism accounting for luminance variations in contrast sensitivity functions,” Journal of the Optical Society of America, Vol. 6, pp. 312-317, Feb. 1989.

[19] B. Levitan and G. Buchsbaum, “Signal sampling and propagation through multiple cell layers in the retina: modeling and analysis with multirate filtering,“ J. Opt. Soc. Am., July 1993, Vol. 10, No. 7, pp. 1463-1480.

[20] J. M. Foley and G. M. Boynton, “A New Model of Human Luminance Pattern Vision Mechanisms: Analysis of the Effects of Pattern Orientation, Spatial Phase and Temporal Frequency,“ SPIE 1994, Vol. 2054, pp. 32-42.

[21] M. W. Cannon, “A Transducer Model for Contrast Perception,” Vision Research, 1991, Vol. 31, No. 6, pp. 983-998.

[22] J. M. Foley, C. Chen, “Pattern detection in the presence of maskers that differ in spatial phase and temporal offset: threshold measurements and a model,” Vision Res. 39, 1999, pp. 3855-3872.

[23] S. Pechard, S. Tourancheau, P. Le Callet, M. Carnec, D. Barba, “Towards Video Quality Metrics For HDTV,” http://enpub.eas.asu.edu/resp/vpqm2007, VPQM-06, pp. 1-6, 2006.

[24] RECOMMENDATION ITU-R BT.500-11, “Methodology for the subjective assessment of the quality of television picture”, ITU-R, 2002.

[25] VQEG, “Final Report From The Video Quality Experts Group on The Validation Of Objective Models of Video Quality Assessment, Phase II,” Tables 2-6, Aug. 25, 2003, pp. 11-14.

[26] Ferguson, Kevin, “Predicting Human Vision Perception and Perceptual Difference,” US Patent No. 6975776, Issued Dec. 13, 2005.

[27] Ferguson, Kevin, “Adaptive Spatio-Temporal Filter for Human Vision System Models,” US Patent No. 6907143, Issued June 14, 2005.

[28] Ferguson, Kevin, “Variable Sample Rate Recursive Digital Filter,” US Patent No. 6976045, Issued Dec. 13, 2005.

[29] Ferguson, Kevin, “Predicting Subjective Quality Ratings of Video,” US Patent No. 6829005, Issued Dec. 7, 2004.

[30] Graham, N., Visual Pattern Analyzers, Oxford U. Press, 1989.

[31] J. J. Koenderink, A. J. van Doorn, “Spatiotemporal contrast detection threshold surface is bimodal,” Optics Letters (Opt. Soc. Am.), Jan. 1979, Vol. 4, No. 1, pp. 32-34.

[32] B. Spehar, Q. Zaidi, “Surround effects on the shape of the temporal contrast-sensitivity function,” Journal of the Optical Society of America, Vol. 14, No. 9, pp. 2517-2525, Sep. 1997.

[33] C. A. Burbek, D. H. Kelly, “Spatiotemporal Characteristics of Visual Mechanisms: Excitatory-Inhibitory Model,” Journal of the Optical Society of America, Vol. 70, No. 9, pp. 1121-1126.

[34] C. Taylor, Z.Pizlo, J. P. Allebach, “Contrast Detection and Discrimination Thresholds for Visual Models,” unpublished ms submitted to Spatial Vision, circa 1998.

[35] J. Rovamo, J. Mustonen, R. Nasanen, “Modelling Contrast Sensitivity as a Function of Retinal Illuminance and Grating Area,” Vision Research, Vol. 34, No. 10, pp. 1301-1314, 1994.
[36] G. Legge, J. Foley, “Contrast masking in human vision,” Journal of Optical Society of America, Vol. 70, No. 12, pp. 1458-1471, Dec.1980.

[37] T. Cornsweet, J. Yellott, “Intensity-dependent spatial summation,” Journal of the Optical Society of America, Vol. 2, No. 10, pp. 1769-1786, 1985.

[38] J. Mustonen, J. Rovamo, R. Nasanen, “The Effects of Grating Area and Spatial Frequency on Contrast Sensitivity as a Function of Light Level,” Vision Research, Vol. 33, No. 15, pp. 2065-2072, 1993.

[39] N. J. Coletta, D. R. Williams and C. L. M. Tiana, “Consequences of Spatial Sampling for Human Motion Perception,” Vision Res., 1990, Vol. 30, No. 11, pp. 1631-1648.

[40] E. Peli, L. Arend, A. T. Labianca, “Contrast perception across changes in luminance and spatial frequency,” J. Opt. Soc. Am., October, 1996, Vol. 13, No. 10, pp. 1953-1959.

[41] Wandell, B. A., Foundations of Vision, Sinauer Associates, 1995.

[42] G. S. Wasserman, “Brightness Enhancement in Intermittent Light: Variation of Luminance and Light-Dark Ratio,” J. Opt. Soc. Am., February 1966, Vol. 56, pp. 242-250.

[43] J. C. Stevens and S. S. Stevens, “Brightness Function: Effects of Adaptation,” J. Opt. Soc. Am., March, 1963, Vol. 53, No. 3, pp 375-385.

[44] R. W. Bowen, “Two Pulses Seen As Three Flashes: A Superposition Analysis,” Vision Research, 1989, Vol. 29, No. 4, pp. 409-417.

[45] W. H. Swanson, T. Ueno, V. C. Smith, J. Pokorny, “Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations,” J. Opt. Soc. Am., Oct. 1987, Vol. 4, No. 10, pp. 1992-2005.

[46] A. Rose, Vision: Human and Electronic, David Sarnoff Research Center, RCA, Plenum Press, NY/London, 1973, pp. 30.

[47] D. Hubel, Eye, Brian, and Vision, Scientific American Library, NY, NY, 1995, pp. 33-136.

[48] Enroth-Cugell, “The World of Retinal Ganglion Cells,” from Shapley, R., Man-Kit Lam, D., ed., Contrast Sensitivity, MIT Press, 1993, pp. 155,159.

[49] J.J. Koenderink, “Contrast Enhancement and the Negative Afterimage,” J. Opt. Soc. Am., May 1972, Vol. 62, No. 5, pp. 685-689.

[50] D.H. Kelly, “Nonlinear visual responses to flickering sinusoidal gratings,” J. Opt. Soc. Am., Sep. 1981, Vol. 71, No. 9, pp. 1051-1055.

[51] D.H. Kelly, “Frequency Doubling in Visual Responses,” J. Opt. Soc. Am., 1966, Vol. 56, No. 11, pp. 1628-1633.

[52] R. W. Bowen, J. V. Mallow, P. J. Harder, “Some properties of the double-flash illusion,” J. Opt. Soc. Am., Apr. 1987, Vol. 4, No. 4, pp. 746-755.

[53] L. E. Marks, “Apparent Depth of Modulation as a Function of Frequency and Amplitude of Temporal Modulations of Luminance,” J. Opt. Soc. Am., July 1970, Vol. 60, No. 7, pp. 970-977.

[54] Wilkins, A., Visual Stress, Oxford Science Publications, 1995, pp. 7-23

[55] VQEG, “HDTV Group TEST PLAN,” http://www.its.bldrdoc.gov/vqeg/projects/hdtv/index.php, Draft Version 1.3, September 29, 2006.

[56] VQEG, “Multimedia Group TEST PLAN,” http://www.its.bldrdoc.gov/vqeg/projects/multimedia/index.php, Draft Version 1.15, September 28, 2006.

[57] G. M. Johnson, “Using Color Appearance in Image Quality Metrics,”http://enpub.eas.asu.edu/resp/vpqm2007, VPQM-06, pp. 1-4, 2006.

[58] Fairchild, M. D., Color Appearance Models, Addison-Wesley, Reading MA, 1998, pp. 80-81, 217-393.

[59] J. G. Robson, “Spatial and Temporal Contrast-Sensitivity Functions of the Visual System,” J. Opt. Soc. Am. (Letters to the editor), August 1966, pp. 1141-1142.

[60] F.L. Van Nes, J.J. Koenderink, H. Nas and M.A. Bouman, “Spatiotemporal Modulation Transfer in the Human Eye,” J. Opt. Soc. Am., Sep. 1967, Vol. 57, No. 9, pp. 1082-1088.

[61] D.H. Kelly, “Visual Responses to Time-Dependent Stimuli. I. Amplitude Sensitivity Measurements,” J. Opt. Soc. Am., April 1961, Vol. 51, No. 4, pp. 422-429.

[62] F.L. Van Nes, J.J. Koenderink, H. Nas and M.A. Bouman, “Spatiotemporal Modulation Transfer in the Human Eye,” J. Opt. Soc. Am., Sep. 1967, Vol. 57, No. 9., pp. 1082-1088.

[63] A. Vassilev, “Contrast Sensitivity Near Borders: Significance of Test Stimulus Form, Size and Duration,” Vision Research, Vol. 13, pp. 719-730, 1993.

[64] E. Peli, J. Yang, R. Goldstein,. Reeves, “Effect of luminance on suprathreshold contrast perception,” J. Opt. Soc. Am., August 1991, Vol. 8, No. 8, pp. 1352-1359.

[65] C. W. Tyler, “Observations on spatial-frequency doubling,” Perception, 1974, Vol. 3, pages 81-86.

[66] T. Maddess, J.M. Hemmi, A.C. James, “Evidence for spatial aliasing effects in the Y-like cells of the magnocellular visual pathway,” Vision Research, 1998, Vol. 38, pp. 1843-1859.

[67] S. Bedford, T. Maddess, K. A. Rose, AC James, “Correlations between observability of the spatial frequency doubled illusion and a multi-region pattern electroretinogram,” Australian and New Zealand Journal of Ophthamology 1997; 25 Sup. 1, pp. S91-S93.

[68] M. Ikeda, “Temporal Summation of Positive and Negative Flashes in the Visual System,” J. Opt. Soc. Am., Nov. 1965, Vol. 55, No. 11, pp. 1527-1534.

[69] Barten, P., Contrast Sensitivity of the Human Eye, SPIE Press 1999.

[70] B. Levitan and G. Buchsbaum, “Signal sampling and propagation through multiple cell layers in the retina: modeling and analysis with multirate filtering,” J. Opt. Soc. Am., July 1993, Vol. 10, No. 7, pp. 1463-1480.

06/2010 Internal/WWW 2BW-25454-0

Tektronix