MOVING IMAGE COLOR APPEARANCE MODEL (MICAM) FOR VIDEO QUALITY RATINGS PREDITION

Kevin Ferguson

Tektronix, Inc.
Portland, OR 97077

ABSTRACT

For predicting video quality via human vision perceptual color response and associated perceptual differences of video relative to a reference, a model of the human vision system has been developed as an extension to CIECAM02 to predict the macro-behavioral response to light stimuli with the following varying parameters: spatial frequency, temporal frequency, angular extent, temporal extent, surround and all parameters relevant to CIECAM02. This paper briefly outlines the motivation for developing this new model, reviews previously developed models, reviews reference stimulus-response data from vision science literature giving responses to stimuli over the above parameter space, and presents details of resulting new model components and example validation test results.

1. INTRODUCTION

Recent proliferation of video resolutions, frame rates, application viewing contexts and technologies for displays, codecs, distribution, transmission, storage and other processing have translated into an underlying motivation for developing highly adaptive technologies to account for the associated high degree of relative change in perceptual sensitivities. Recent advances in human vision models used in predicting perceptional response have come directly as a result of the recognition of adaptation as a key behavioral trait that was prior lacking. Models for color appearance such as CIECAM02 [1] and the luminance perception model of [2] are examples. Combined, these two technologies include mechanisms that adapt to contexts of viewing environment and video content: color, brightness, viewing distance, video frame rates, resolutions and other important factors that effect human vision perceptual response. This paper discusses a model synthesized from the combination of CIECAM02 [1] and the adaptive building blocks of [2].

1.1. Prior Models

The CIE has created several improved color appearance models (CAMs) since CIE Luv was introduced. Currently, the most accurate international standard CAM is CIECAM02. A subset of the CIECAM02 development team has developed still image color appearance models based on CIECAM02: iCAM and iCAM06, both include spatial processing for images, the latter with extra processing to handle high dynamic luminance range. However, these are not international standards and do not account for changes in viewing distance or similar factors effecting spatial response. Meanwhile, CIECAM02 has found acceptance for its original use, for reflected light applications such as matching paints in different lighting environments. The reflected light application represents a smaller gamut than direct light, excluding the extremes in saturation and luminance levels. However, for current broadcast video standards and the maximum luminance output of most commercial display technologies, the video gamut mostly overlaps the CIECAM02 target. And while CIECAM02 takes into account color adaptation, it requires specification of the level of adaptation and does not include any temporal dimension to the input or output. In other words, it is designed to work with static color patches with static surround and static lighting. Besides the addition of spatial filters for iCAMs, spatiotemporal filtering has been proposed for adapting CIECAM02 to video applications, for example in [3], though no specific design was proposed.

On the other hand, [2] presented a luminance perception model that takes into account adaptation across space and time. It mentions that color perception may similarly be modeled using the same adaptive integrator building block to process spatial and temporal color stimuli in order to predict color related adaptation effects over space and time. For both luminance and color, adaptive response is relative to an integrated adaptation point represented as the output of a "surround" spatiotemporal filter with upper resolution in time and space set by a "center" spatiotemporal filter.

2. REFERENCE STIMULUS-RESPONSE DATA

For purposes of calibrating center and surround filters as well as regression testing of the adapted response of CIECAM02 for 2 degree patches, experiments conducted in vision science research were replicated in simulation in order to check for appropriate responses. Each data set was roughly categorized by stimuli type corresponding to spatial and temporal frequencies and expected responses: near perceptual threshold or relatively constant (equal magnitude) suprathreshold responses.

2.1. "Static Patch" Threshold and Suprathreshold Tests for CIECAM02

For both CIECAM02 validation and regression testing after adding filters, the following were used. The Munsell data set was used for both absolute and relative difference validation [4], Hita's metameric data [5], MacAdam metameric [6] and JND [7] ellipses, Brown's JND color difference vs. mean luminance data [8], Newhall's "39%" threshold color difference data [9], fixed hue data of Boynton [10], Hurvich [11] and Fuld [12], Wright's suprathreshold equal changes covering most of the CIE 1931 xy color gamut in several directions [13] and OSA/MacAdam suprathreshold dataset [14].

2.2. Spatiotemporal Tests

For temporal center filter calibration, data from the following were used: [15], [16], [17], [18], [19], [20], and [21]. For spatial center filter calibration, data sets from [22], [23] & [24] were used. Note that [25] was also used to verify the spatiotemporal combination. For temporal surround, [26] and [27] were used. Spatial surround data was in part gathered from the combination of the above at lowest frequencies. In aggregate, these stimuli also sampled angular and temporal extent.

3. MODEL COMPONENTS

Model components consist of CIECAM02 [1] to obtain at least two channel {a,b} (& "orthogonal" {ac,bc}, though no net advantage was seen) color response, together with the appropriate adaptive spatial and temporal processing derived from [2] to account for color perception of moving images, and controls for levels of cognitive discounting and color adaptation localization (Figure 1).

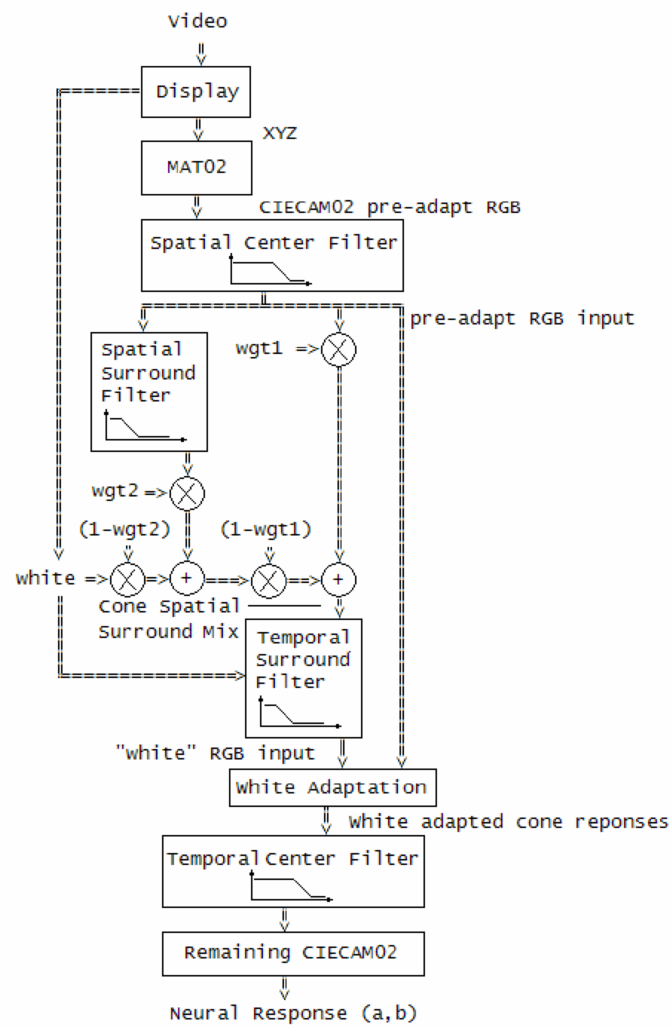

Fig. 1. MiCAM processing block diagram using CIECAM02 and adaptive spatiotemporal center and surround filtering.

The adaptive center spatiotemporal filter design and verification parallels [2] for luminance. As such, and due to space limitations, details will be omitted here as they are not so new. However, a new type of surround filter, particularly when considering the temporal processing, is required for two primary reasons. First, CIECAM02 already has a mechanism to take into account fixed adaptation relative to "surround" and/or the illuminant, given the percentage of adaptation. One input to CIECAM02 that in part controls adaptation is "adaptation white" {Rw,Gw,Bw}, used to modify the cone responses according to instantaneous level of adaptation. Likewise in both CIECAM02 and in [2], the surround channel is used to modify the center channel to produce the primary adapted spatiotemporal response. So as to leave CIECAM02 static response unchanged, the temporal processing is applied to "adaptation white" input which is treated as the overall surround channel.

More specifically, the temporal adaptation filter model (as well as all spatial and temporal filtering) is applied to CIECAM02 after image light stimulus is converted into the three pre-adapted (RGB) cone responses. CIE1931 XYZ tristimulus components of the image formed by simulated light (i.e. display simulation output, Y as per [2]) are converted to pre-adapted RGB cone responses using the MAT02 [1] conversion matrix as depicted at the top of Figure 1. These cone responses are used as inputs to a series combination of center and surround spatial and temporal filters as shown, though the spatial and temporal centers could also be combined as in [2]. The filters are applied as aggregate or "lumped" spatial and/or temporal response as an approximation of the combination of spatial and temporal response of each anatomical component of the human vision system as in [2]. If the static response of CIECAM02 is to remain unchanged, this aggregation is required in order to prevent non-linearities of subsequent CIECAM02 processing from creating unwanted artifacts such as rectification of intermediate responses.

3.1. Control of Local vs. Global Color Adaptation

The spatial center filter output connects to both the spatial surround filter and two other inputs: the pre-adapted cone response input to the "White Adaptation" portion of CIECAM02 (responsible for color adaptation), and a weighting mechanism (via wgt1) to control how much of local vs. global spatial aspect of the adaptation is used. Thus wgt1 controls localization of spatial color adaptation. This reflects effects of fixated vs. roaming gaze. For minimum (no) local color adaptation (after-images), wgt1 = 0, while for maximum wgt1 = 1.

3.2. Surround vs. White Point and Ambient

The "White" input represents the pre-adapted RGB cone responses to the combined display white and ambient illumination XYZ light combination. For pure display and ambient white point adaptation wgt2 = 0 vs. spatial surround only (each channel normalized to Yw) adaptation, wgt2=1. Conventional use of CIECAM02 corresponds to wgt1=0, wgt2=0, ambient=0, display white = illuminant.

3.3. Control of Cognitive Discounting

Cognitive discounting, the ability to compensate for to the illuminant, display white, ambient illumination, etc. when identifying colors, is controlled by wgt3 (not shown). The "White" input mentioned in 3.2 may be cross-faded to Rw'=Gw'=Bw'=Yw (equivalent to adaptation parameter D=0 in CIECAM02) by setting wgt3 = 1, where Rw'=wgt3*Yw + (1-wgt3)*Rw, and likewise for Gw and Bw.

3.4. Temporal Surround Filter

For each channel, the composite temporal surround filter is a parallel combination of the adaptive integrator based filters mentioned above. The temporal surround filter block diagram is shown in Figure 2.

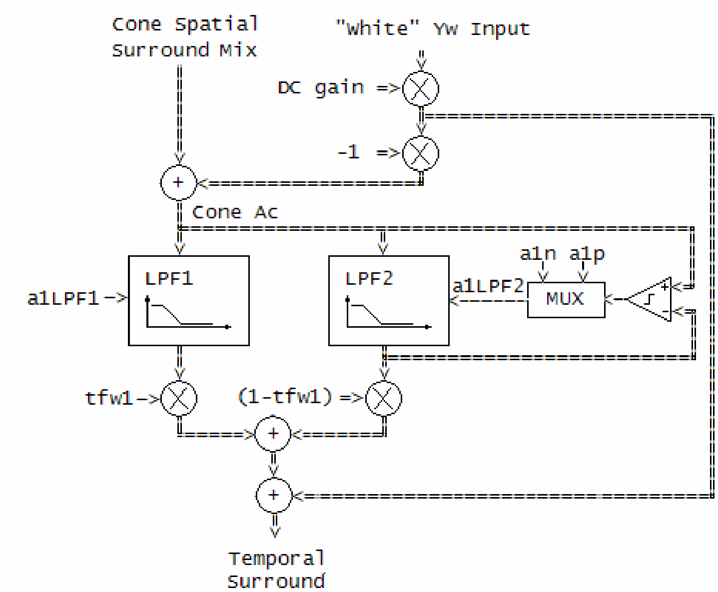

Fig. 2. Block diagram of the low frequency temporal processing : "Temporal Surround Filter" of Figure 1.

The "Cone Spatial Surround Mix" created by the sums of weighted inputs (as shown in Figure 1) to account for localization of color adaptation and cognitive discounting are a set of weighted cone responses (a channel each of R,G and B).

The "White" Yw input is the Y portion of the white input (as per CIECAM02). Yw is either directly input from the display model or converted back to Yw from the white RGB, otherwise known for example in CIECAM02 as Rw,Gw,Bw. This "White" Yw input is weighted by a "DC gain" factor = 0.48 and subtracted from each of the three "Cone Spatial Surround Mix" response inputs. The result is pre-temporal surround filtered differential or "AC" components of the "Cone Spatial Surround Mix" response. Each channel represents the difference in respective cone response from intensity scaled "adaptation white."

The two adaptive integrator based IIR low-pass filters detailed below are used in parallel to filter these differential color signals. The slow first LPF, LPF1 is an instantiation of the same adaptive temporal surround filter used in [2], only with updated parameters for zero adaptation and nominal feedback coefficient a1 set to 0.999516 for a nominal sample rate of 60 samples per second. The faster second IIR LPF is similar except that it has variable coefficients depending on whether its input is less than or greater than its output. A comparator is used to determine which coefficient to select: a1p = 0.9854 if the quantity (LPF2 input minus LPF2 output) is positive, a1n = 0.99370 if negative. Next, a weighted average (using tfw1, approximately equal to .5) of LPF1 and LPF2 outputs is calculated resulting in the composite filtered differential signal. This resulting composite filtered differential signal is restored to an absolute one by adding back the intensity scaled white luminance signal.

The Dirac delta impulse response of LPF1 is shown in Figure 3. The two curves represent 10 and 60 samples per second. This shows an example of maintaining filter response at two different frame rates, the temporal sample rate adaptation requirement met by using the adaptive integrator of [2].

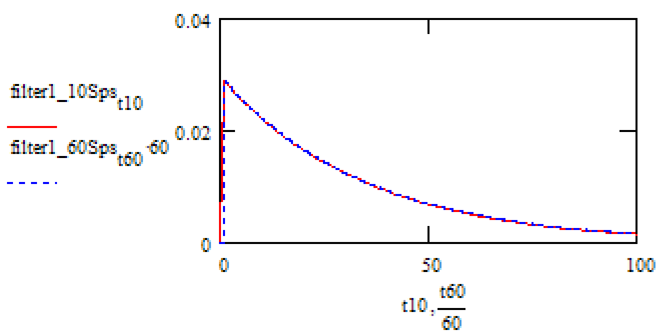

Fig. 3. Temporal Surround Filter LPF1 Dirac delta impulse response shown for a sample rate of 10 Samples/second, solid line, and 60 Samples/second, the dashed line. The horizontal axis is in units of tenths of a second.

The temporal surround filter response is nonlinear primarily due to the difference between responses to positive vs. negative changes accounted for by LPF2. This can be seen in Figure 4. An example of this difference in response involves abrupt changes in surround and tracking stimulus changes required for achromatic response, as will be explored in the next section.

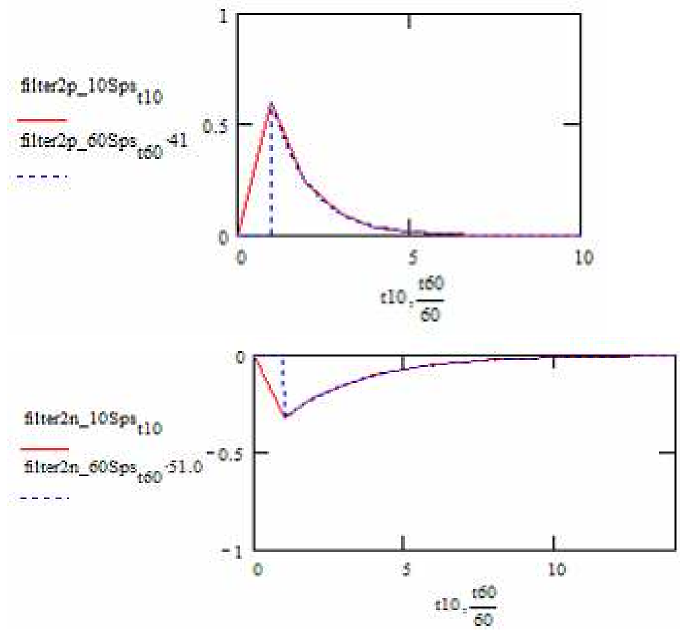

Fig. 4. Temporal Surround Filter LPF2 positive (top) and negative (bottom) Dirac delta impulse responses shown in two model simulation sample rates (10 and 60 samples/second for solid and dashed lines respectively) vs. 10ths of a second.

4. VERIFICATION OF ACCURATE MODEL RESPONSE: TEMPORAL SURROUND EXAMPLE

Of the temporal surround filter data gathered, one experiment was particularly useful for indicating details of longer term temporal response of human color perception. Fairchild and Renif [26] conducted an experiment to track temporal step responses to spatial surround changes. In this experiment, subjects were given the task of maintaining achromatic (grey) appearance of stationary target patch in the center following a temporal step change in color in the surround. Thus, simulation of this experiment should result in achromatic MiCAM response even though input target color (as defined in CIE 1931 xyY, XYZ, etc.) is changing significantly over time.

The direction in the color plane (CIE xy or CIECAM02 ab, etc.) of changes in input test target color required for maintaining achromatic response in [26] depend on the direction of the change in surround.

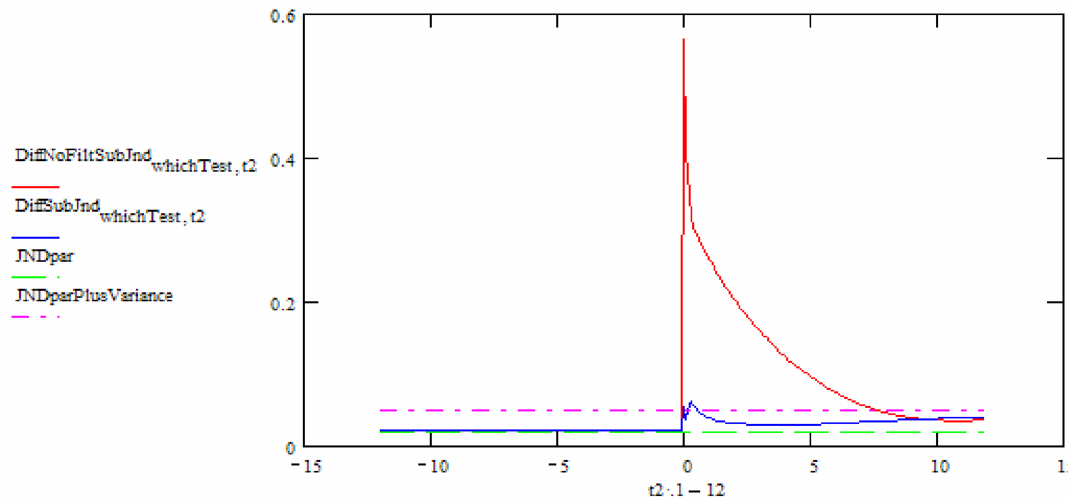

Fig. 5. Blue to yellow surround step response: Temporal surround sub-JND maximum difference with (low amplitude blue curve) and without filters (high amplitude red curve) vs. time. Nominal threshold and "guard-band markers" (threshold plus one standard deviation.

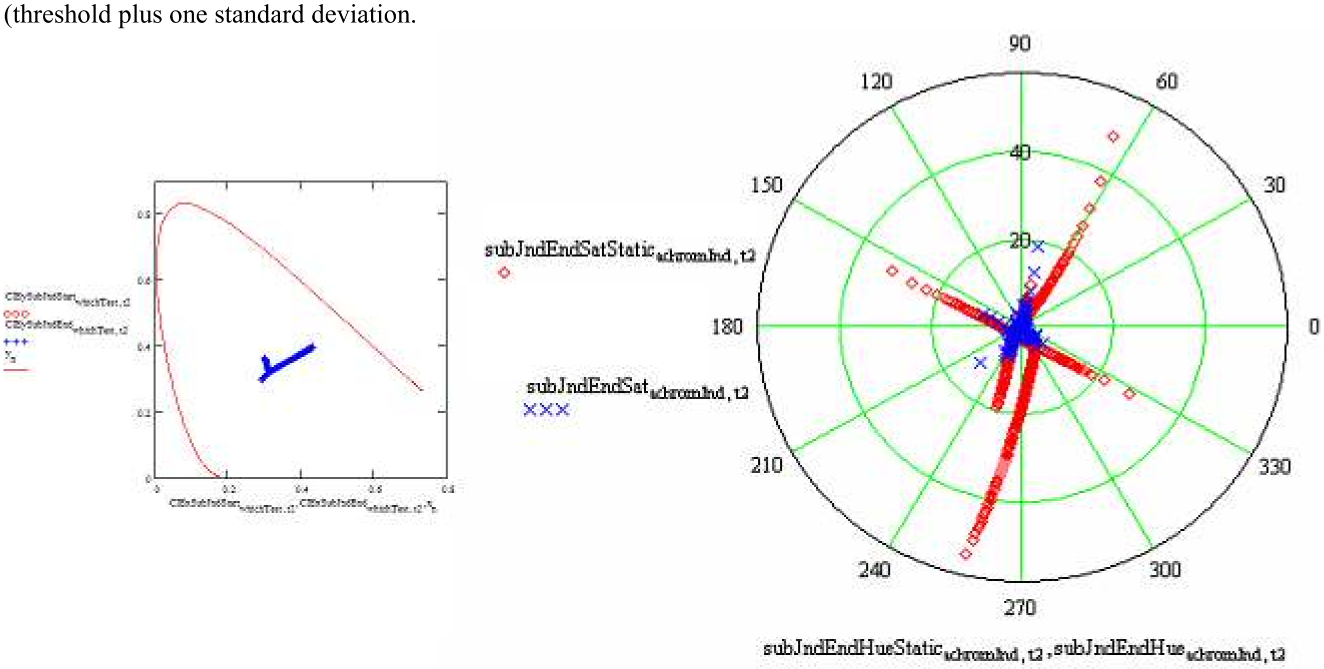

Fig. 6. Left: Step stimuli of [26] in CIE 1931 {x,y} plane. Right: Responses in CIECAM02 {Saturation, Hue} plane, where the origin represents the achromatic (grey) response expected as in [26]. Diamonds represent CIECAM02 response without any temporal processing, while X's show response once temporal surround processing is added. In this case, the temporal center filter has been omitted and therefore initial transients cause some of the X's to deviate from the origin somewhat, as can also be seen in the short peak of the response in Figure 5. Adding the temporal center filter mitigates these transients.

Three directions (red, green, blue approximating respective cone response peaks) and their CIE XYZ opposites, for a total of six surround color steps were used. Thus, six time courses (temporal trajectories) were obtained in the experiment.

The six stimuli were simulated using CIECAM02 and the temporal surround filter. An example temporal chromatic response is shown in Figure 5, depicting CIECAM02 {a,b} magnitude equivalent vs. time (with and without the temporal surround filter). Hue-chroma plots for all six surround steps are shown superimposed in Figure 6. In Figure 5, the lower dashed line represents nominal color difference threshold (~1 JND) from achromatic (no color) whereas the higher dot-dash line represents the threshold plus 1 standard deviation. The peak chromatic responses shown in Figure 5 correspond to the maximum excursions of Figure 6. Note that these maximum excursions occur at the instant of surround step transition. Therefore, the addition of the temporal center filter is expected to greatly mitigate these deviations from achromatic response. In addition, the spatial filter is expected to slightly reduce all six chromatic responses.

5. SUMMARY

A highly adaptable moving picture color appearance 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 combination of a prior adaptable spatiotemporal luminance model and the CIECAM02 "static patch" predicting color appearance model has been outlined. A new temporal surround filter is required for this combination. Example model temporal surround response test results compare well with human vision perceptual response.

6. REFERENCES

[1] Paula J. Alessi et al, "A Colour Appearance Model For Colour Management Systems: CIECAM02", CIE 159:2004.

[2] Kevin Ferguson, "An Adaptable Human Vision Model for Subjective Video Quality Rating Prediction Among CIF, SD, HD AND E-CINEMA," VPQM 2007.

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

[4] Gunter Wyszecki, W. S. Stiles, "Color Science: Concepts and Methods, Quantitative Data and Formulas, 2nd Edition," 1982, John Wiley & Sons, NY., Table I(6.6.1), pp. 840-852.

[5] E. Hita, L.Jimenez del Barco, J. Romero, "Differential color thresholds from metameric matches: experimental results concerning failures of colorimetric additively," J. Opt. Soc. Am., August 1986, Vol. 3, No. 8, pp 1203-1209.

[6] D. L. MacAdam, "Visual sensitivities to color differences in daylight," JOSA, May, 1942, Vol. 32, pp. 247-274.

[7] W.R.J. Brown, D.L. MacAdam, "Visual Sensitivities to Combined Chromaticity and Luminance Differences," J. Opt. Soc. Am., October 1949, Vol. 39, No. 10, pp 808-834.

[8] W.R.J. Brown, "The Influence of Luminance Level on Visual Sensitivity to Color Differences," J. Opt. Soc. Am., October 1951, Vol. 41, No. 10, pp 684-688.

[9] S. M. Newhall, R. W. Burnham, Joyce R. Clark, "Comparison of Successive with Simultaneous Color Matching," JOSA, January, 1957, Vol. 47, No. 1, pp. 43-56.

[10] Robert M. Boynton, James Gordon, "Bezold-Brucke Hue Shift Measured by Color-Naming Technique," JOSA, January, 1965, Vol. 55, No. 1, pp. 78-86.

[11] Leo M. Hurvich and Dorothea Jameson, "Some Quantitative Aspects of an Opponent-Colors Theory. II. Brightness, Saturation, and Hue in Normal and Dichromatic Vision", JOSA, August, 1955, Vol. 45, No. 8, pp. 602-616.

[12] Kenneth Fuld, Timothy A. Otto, "Colors of monochromatic lights that vary in contrast-induced brightness," JOSA, January, 1985, Vol. 2, No. 1, pp. 76-83.

[13] W. D. Wright, "The Sensitivity of The Eye to Small Colour Differences," Proc. Phys. Soc. (The Proceedings of The Physical Society), March, 1941, Vol. 53, Part 2, No. 296, pp. 93-112.

[14] David L. MacAdam, "Colorimetric data for samples of OSA uniform color scales," JOSA, January, 1978, Vol. 68, No. 1, pp. 121-130.

[15] C. Noorlander and J.J. Koenderink, "Spatial and temporal discrimination ellipsoids in color space," JOSA, November, 1983, Vol. 73, No. 10, pp. 1533-1543.

[16] William H. Swanson, Takehiro Ueno, Vivianne C. Smith and Joel 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.

[17] David C. Burr, M. Concetta Morronne, "Impulse-response functions for chromatic and achromatic stimuli," J. Opt. Soc. Am., August 1993, Vol. 10, No. 8, pp 1706-1713.

[18] Tatsuya Yoshizawa, Keiji Uchikawa, "Temporal integration characteristics of chromatic response as determined by use of the isoluminant double-pulse method," J. Opt. Soc. Am., September 1997, Vol. 14, No. 9, pp. 2069-2080.

[19] John J. Wisowaty, "Estimates for the temporal response characteristics of chromatic pathways," JOSA, August, 1981, Vol. 71, No. 8, pp. 970-977.

[20] Gerard J. C. Van Der Horst, "Chromatic Flicker," J. Opt. Soc. Am., 1969, Vol. 59, pp 1213-1217.

[21. K. Uchikawa and I Ikeda, "Temporal integration of chromatic double pulses for detection of equal-luminance wavelength changes," JOSA, A 3, 2109-2115 (1986)

[22] Gerard J.C. Van Der Horst, Charles M.M. De Weert, Maarten A. Bouman, "Transfer of Spatial Chromaticity-Contrast at Threshold in the Human Eye," J. Opt. Soc. Am., October 1967, Vol. 57, No. 10, pp. 1260-1266.

[23] D.H. Kelly, "Spatiotemporal variation of chromatic and achromatic contrast thresholds," J. Opt. Soc. Am., June 1983, Vol. 73, No. 6, pp 742-750.

[24] C. Noorlander and J.J. Koenderink, "Spatial and temporal discrimination ellipsoids in color space," JOSA, November, 1983, Vol. 73, No. 10, pp. 1533-1543.

[25] D.H. Kelly, "Spatiotemporal variation of chromatic and achromatic contrast thresholds," J. Opt. Soc. Am., June 1983, Vol. 73, No. 6, pp 742-750.

[26] Mark D. Fairchild and Lisa Reniff, "Time-course of chromatic adaptation for color-appearance judgments," J. Opt. Soc. Am., May, 1995, Vol. 12, No. 5, pp. 824-833.

[27] R.W.G. Hunt, "The Effects of Daylight and Tungsten Light-Adaptation on Color Perception," J. Opt. Soc. Am., June, 1950, Vol. 40, No. 6, pp. 362-371.