The iCAM Framework for Image Appearance, Image Differences, and Image Quality

Mark D. Fairchild and Garrett M. Johnson

Munsell Color Science Laboratory

Chester F. Carlson Center for Imaging Science

Rochester Institute of Technology

54 Lomb Memorial Drive

Rochester, NY 14623-5604

www.cis.rit.edu/mcsl

mdf@cis.rit.edu, garrett@cis.rit.edu

Traditional color appearance modeling has recently matured to the point that available, internationally-recommended models such as CIECAM02 are capable of making a wide range of predictions to within the observer variability in color matching and color scaling of stimuli in somewhat simplified viewing conditions. It is proposed that the next significant advances in the field of color appearance modeling and image quality metrics will not come from evolutionary revisions of these models. Instead, a more revolutionary approach will be required to make appearance and difference predictions for more complex stimuli in a wider array of viewing conditions. Such an approach can be considered image appearance modeling since it extends the concepts of color appearance modeling to stimuli and viewing environments that are spatially and temporally at the level of complexity of real natural and man-made scenes and extends traditional image quality metrics into the color appearance domain. This paper reviews the concepts of image appearance modeling, presents iCAM as one example of such a model, and provides a number of examples of the use of iCAM in image reproduction and image quality evaluation.

Keywords: color appearance, image appearance, image quality, vision modeling, image rendering

INTRODUCTION

The fundamental theme of this research can be considered image measurement and the application of those measurements to image rendering and image quality evaluation. Consideration of the history of image measurement helps set the context for the formulation and application of image appearance models, a somewhat natural evolution of color appearance, spatial vision, and temporal vision models. Early imaging systems were either not scientifically measured at all, or measured with systems designed to specify the variables of the imaging system itself. For example, densitometers were developed for measuring photographic materials with the intent of specifying the amounts of dye or silver produced in the film. In printing, similar measurements would be made for the printing inks as well as measures of the dot area coverage for halftone systems. In electronic systems like television, system measurements such as signal voltages were used to quantify the imaging system. As imaging systems evolved in complexity and openness, the need for device-independent image measures became clear.

Image Colorimetry

Electronic imaging systems, specifically the development of color television, prompted the first application of device-independent color measurements of images. Device-independent color measurements are based on the internationally-standardized CIE system of colorimetry first developed in 1931. CIE colorimetry specifies a color stimulus with numbers proportional to the stimulation of the human visual system independent of how the color stimulus was produced. The CIE system was used very successfully in the design and standardization of color television systems (including recent digital television systems).

Application of CIE colorimetry to imaging systems became much more prevalent with the advent of digital imaging systems and, in particular, the use of computer systems to generate and proof content ultimately destined for other media such as print. As color-capable digital imaging systems (from scanners

and cameras, through displays, to various hardcopy output technologies) became commercially available in the last two decades, it was quickly recognized that device-dependent color coordinates (such as monitor RGB and printer CMYK) could not be used to specify and reproduce color images with accuracy and precision. An additional factor was the open-systems nature of digital imaging in which the input, display, and output devices might be produced by different manufacturers and one source could not control color through the entire process. The use of CIE colorimetry to specify images across the various devices promised to solve some of the new color reproduction problems created by open, digital systems. The flexibility of digital systems also made it possible and practical to perform colorimetric transformations on image data in attempts to match the colors across disparate devices and media.

Research on imaging device calibration and characterization has spanned the range from fundamental color measurement techniques to the specification of a variety of devices including CRT, LCD, and projection displays, scanners and digital cameras, and various film recording and print media. Some of the concepts and results of this research have been summarized by Berns.[1] Such capabilities are a fundamental requirement for research and development in color and image appearance. Research on device characterization and calibration provides a means to tackle more fundamental problems in device-independent color imaging. For example, conceptual research on design and implementation of device-independent color imaging,[2] gamut mapping algorithms to deal with the reproduction of desired colors that fall outside the range that can be obtained with a given imaging device,[3] and computer graphics rendering of high-quality spectral images that significantly improve the potential for accurate color in rendered scenes.[4] This type of research built upon, and contributed to, research on the development and testing of color appearance models for cross-media image reproduction.

Color Difference Equations
Color difference research has culminated with the recently published CIEDE2000 color difference formula.[5] A color difference equation allows for the mapping of

physically measured stimuli into perceived differences. At the heart of such color difference equations lies some form of uniform color space. The CIE initially recommended two such color spaces, in 1976, CIELAB and CIELUV. Both spaces were initially described as interim color spaces, with the knowledge that they were far from complete. Over 25 years later these spaces are still the CIE recommendations, although CIELUV has fallen out of favor.

With a truly uniform color space, color differences can then be taken to be a simple measure of distance between two colors in the space, such as CIE ΔE^*_{ab}. The CIE recognized the non-uniformity of the CIELAB color space, and formulated more advanced color difference equations such as CIE DE94 and CIEDE2000. These more complicated equations are very capable of predicting perceived color differences of simple color patches.

Image Difference

The CIE color difference formula were developed using simple color patches in controlled viewing conditions. There is no reason to believe that they are adequate for predicting color difference for spatially complex image stimuli. The S-CIELAB model was designed as a spatial pre-processor to the standard CIE color difference equations, to account for complex color stimuli such as halftone patterns.[6] The spatial preprocessing uses separable convolution kernels to approximate the contrast sensitivity functions (CSF) of the human visual system. The CSF serves to remove information that is imperceptible to the visual system. For instance, when viewing halftone dots at a certain distance the dots tend to blur, and integrate into a single color. A pixel-by-pixel color difference calculation between a continuous image and a halftone image would result in very large errors, while the perceived difference might in fact be small. The spatial pre-processing would blur the halftone image so that it more closely resembles the continuous tone image.

S-CIELAB represents the first incarnation of an image difference model based upon the CIELAB color space and color difference equations. Recently this model has been refined and extended into a modular framework for image color

difference calculations.[7] This framework refines the CSF equations from the S-CIELAB model, and adds modules for spatial frequency adaptation, spatial localization, and local and global contrast detection. This framework is discussed in more detail below.

Color Appearance

Unfortunately, fundamental CIE colorimetry does not provide a complete solution for image specification. CIE colorimetry is only strictly applicable to situations in which the original and reproduction are viewed in identical conditions. By their very nature, the images produced or captured by various digital systems are examined in widely disparate viewing conditions, from the original captured scene, to a computer display in a dim room, to printed media under a variety of light sources. Thus color appearance models were developed to extend CIE colorimetry to the prediction of color appearance (not just color matches) across changes in media and viewing conditions (not just within a single condition). Color appearance modeling research applied to digital imaging systems was very active throughout the 1990s culminating with the recommendation of the CIECAM97s model in 1997[8] and its revision, CIECAM02, in 2002.[9] Details on the evolution, formulation, and application of color appearance models can be found in Fairchild.[10] The development of these models was also enabled by visual experiments performed to test the performance of published color appearance models in realistic image reproduction situations.[11] Such research on color appearance modeling in imaging applications naturally highlighted the areas that are not adequately addressed for spatially complex image appearance and image quality problems.

Image Appearance and Image Quality

Color appearance models account for many changes in viewing conditions, but are mainly focused on changes in the color of the illumination (white point), the illumination level (luminance), and surround relative luminance. Such models do not directly incorporate any of the spatial or temporal properties of human vision and the perception of images. They essentially treat each pixel of an image (and each frame of a video) as completely independent stimuli. While

color appearance modeling has been successful in facilitating device-independent color imaging and is incorporated into modern color management systems, there remains significant room for improvement. To address these issues with respect to spatial properties of vision and image perception (localized adaptation and spatial filtering) and image quality, the concept of image appearance models has been recently introduced and implemented.[12,13] These models combine attributes of color appearance models with attributes of spatial vision models that have been previously used for image quality metrics in an attempt to further extend the capabilities of color appearance models. Historically color appearance models largely ignored spatial vision (*e.g.*, CIECAM97s) while spatial vision models for image quality largely ignored color.[14,15] One notable exception, and the theme of this special issue of the *Journal of Electronic Imaging*, has been the retinex model[16,17] and its various derivatives.[18] While the retinex model was never designed as a complete model of image appearance and quality, its spatially variable mechanisms of chromatic adaptation and color constancy serve some of the same purposes in image rendering and provide some of the critical groundwork for image appearance modeling. The goal in developing an image appearance model has been to bring these research areas together to create a single model applicable to image appearance, image rendering, and image quality specifications and evaluations. One such model for still images, referred to as iCAM, has recently been published by Fairchild and Johnson[13] and is detailed in this paper. This model was built upon previous research in uniform color spaces,[19] the importance of image surround,[20] algorithms for image difference and image quality measurement,[21,22] insights into observers eye movements while performing various visual imaging tasks and adaptation to natural scenes,[23,24] and an earlier model of spatial and color vision applied to color appearance problems and high-dynamic-range (HDR) imaging.[25] The structure of the iCAM model and examples of its implementation for image appearance are presented below.

Color and Image Appearance Models

A model capable of predicting perceived color difference between complex image stimuli is a useful tool, but has some limitations. Just as a color appearance

model is necessary to fully describe the appearance of color stimuli, an image appearance model is necessary to describe spatially complex color stimuli. Color appearance models allow for the description of attributes such as lightness, brightness, colorfulness, chroma, and hue. Image appearance models extend upon this to also predict such attributes as sharpness, graininess, contrast, and resolution.

A uniform color space also lies in the heart of the of an image appearance model. The modular image difference framework allows for great flexibility in the choice of color spaces. Examples are the CIELAB color space, similar to S-CIELAB, the CIECAM02 color appearance model, or the IPT color space.[9,19]

Models of image appearance can be used to formulate multi-dimensional models of image quality. For example it is possible to take weighted sums of various appearance attributes to determine a metric of overall image quality, as described by Keelen[26] and Engledrum.[27] Essentially these models can augment or replace human observations to weight image attributes with overall appearances of quality. For instance a model of quality might involve weighted sums of tonal balance, contrast, and sharpness. A first step towards this type of model is illustrated in more detail below.

THE iCAM FRAMEWORK

Figure 1 presents a flow chart of the general framework for the iCAM image appearance model as applied to still images originally presented by Fairchild and Johnson.[13] A description of the model along with example images and code can be found at <www.cis.rit.edu/mcsl/iCAM>. For input, the model requires colorimetrically characterized data for the image (or scene) and surround in absolute luminance units. The image is specified in terms of relative CIE XYZ tristimulus values. The adapting stimulus is a low-pass filtered version of the CIE XYZ image that is also tagged with absolute luminance information necessary to predict the degree of chromatic adaptation. The absolute luminances (Y) of the image data are also used as a second low-pass image to control various luminance-dependant aspects of the model intended to predict

the Hunt effect (increase in perceived colorfulness with luminance) and the Stevens effect (increase in perceived image contrast with luminance). Lastly, a low-pass, luminance (Y) image of significantly greater spatial extent is used to control the prediction of image contrast that is well-established to be a function of the relative luminance of the surrounding conditions (Bartleson and Breneman equations). Refer to Fairchild[10] for a full discussion of the various image appearance effects mentioned above and detailed specifications of the data required. The specific low-pass filters used for the adapting images depend on viewing distance and application. Additionally, in some image rendering circumstances it might be desirable to have different low-pass adapting images for luminance and chromatic information to avoid desaturation of the rendered images due to local chromatic adaptation (decrease in visual sensitivity to the color of the stimulus). This is one example of application dependence. Local chromatic adaptation might be appropriate for image-difference or image-quality measurements, but inappropriate for image-rendering situations.

The first stage of processing in iCAM is to account for chromatic adaptation. The chromatic adaptation transform embedded in the recently-published CIECAM02 model[9] has been adopted in iCAM since it was well researched and established to have excellent performance with all available visual data. It is also a relatively simple chromatic adaptation model amenable to image-processing applications. The chromatic adaptation model, given in Eqs. 1-6, is a linear von Kries normalization of RGB image signals to the RGB adaptation signals derived from

$$
\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \mathbf{M}_{CAT02} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}
\tag{1}
$$

$$
\mathbf{M}_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix}
\tag{2}
$$

$$D = F\left[1 - \left(\frac{1}{3.6}\right)e^{\left(\frac{-L_A - 42}{92}\right)}\right]$$ (3)

$$R_c = \left[\left(Y_W \frac{D}{R_W}\right) + (1 - D)\right]R$$ (4)

$$G_c = \left[\left(Y_W \frac{D}{G_W}\right) + (1 - D)\right]G$$ (5)

$$B_c = \left[\left(Y_W \frac{D}{B_W}\right) + (1 - D)\right]B$$ (6)

the low-pass adaptation image at each pixel location ($R_W G_W B_W$). The RGB signals are computed using a linear transformation from XYZ to RGB derived by CIE TC8-01 in the formulation of CIECAM02. This matrix transformation has come to be called the M_{CAT02} matrix, where CAT stands for chromatic adaptation transform. The von Kries normalization is further modulated with a degree-of-adaptation factor, D, that can vary from 0.0 for no adaptation to 1.0 for complete chromatic adaptation. Equation 3 is provided in the CIECAM02 formulation, and used in iCAM, for computation of D as a function of adapting luminance, L_A, for various viewing conditions. Alternatively the D factor can be established manually. The chromatic adaptation model is used to compute corresponding colors for CIE Illuminant D65 that are then used in the later stages of the iCAM model. It should be noted that, while the adaptation transformation is identical to that in CIECAM02, the iCAM model is already significantly different since it uses spatially-modulated image data as input rather than single color stimuli and adaptation points. It also differs completely in the remainder of the formulation although using CIECAM02 equations where appropriate. One example of this is the modulation of the absolute-luminance image and surround luminance image using the F_L function from CIECAM02 given in Eq. 7. This function, slowly

$$F_L = 0.2\left(\frac{1}{(5L_A+1)}\right)^4 (5L_A) + 0.1\left(1-\left(\frac{1}{(5L_A+1)}\right)^4\right)^2 (5L_A)^{\frac{1}{3}} \qquad (7)$$

varying with luminance, has been established to predict a variety of luminance-dependent appearance effects in CIECAM02 and earlier models. Since the function has been established and understood, it was also adopted for the early stages of iCAM. However, the manner in which the F_L factor is used in CIECAM02 and iCAM are quite different.

The next stage of the model is to convert from RGB signals (roughly analogous to cone signals in the human visual system) to opponent-color signals (light-dark, red-green, and yellow-blue; analogous to higher-level encoding in the human visual system) that are necessary for constructing a uniform perceptual color space and correlates of various appearance attributes. In choosing this transformation, simplicity, accuracy, and applicability to image processing were the main considerations. The color space chosen was the IPT space previously published by Ebner and Fairchild.[19] The IPT space was derived specifically for image processing applications to have a relatively simple formulation and specifically to have a hue-angle component with good prediction of constant perceived hue (important in gamut-mapping applications). More recent work on perceived hue has validated the applicability of the IPT space. The transformation from RGB to the IPT opponent space is far simpler than the transformations used in CIECAM02. The process, expressed in Eqs. 9-12, involves a linear transformation to a different cone-response space (a different RGB), application of power-function nonlinearities, and then a final linear transformation to the IPT opponent space (I: light-dark; P: red-green, T: yellow-blue).

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0.0 & 0.0 & 0.9184 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix} \qquad (8)$$

$$L' = L^{0.43}; \quad L \geq 0$$
$$L' = -|L|^{0.43}; \quad L \leq 0 \tag{9}$$

$$M' = M^{0.43}; \quad M \geq 0$$
$$M' = -|M|^{0.43}; \quad M \leq 0 \tag{10}$$

$$S' = S^{0.43}; \quad S \geq 0$$
$$S' = -|S|^{0.43}; \quad S \leq 0 \tag{11}$$

$$\begin{bmatrix} I \\ P \\ T \end{bmatrix} = \begin{bmatrix} 0.4000 & 0.4000 & 0.2000 \\ 4.4550 & -4.8510 & 0.3960 \\ 0.8056 & 0.3572 & -1.1628 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \tag{12}$$

The power-function nonlinearities in the IPT transformation are a critical aspect of the iCAM model. First, they are necessary to predict response compression that is prevalent in most human sensory systems. This response compression helps to convert from signals that are linear in physical metrics (*e.g.*, luminance) to signals that are linear in perceptual dimensions (*e.g.*, lightness). The CIECAM02 model uses a hyperbolic nonlinearity for this purpose. The behavior of which is that of a power function over the practical ranges of luminance levels encountered. Secondly, and a key component of iCAM, the exponents are modulated according to the luminance of the image (low-pass filtered) and the surround. This is essentially accomplished by multiplying the base exponent in the IPT formulation by the image-wise computed F_L factors with appropriate normalization. These modulations of the IPT exponents allow the iCAM model to be used for predictions of the Hunt, Stevens, and Bartleson/Breneman effects mentioned above. They also happen to enable the tone mapping of high-dynamic-range images into low-dynamic range display systems in a visually meaningful way (see example in Fig. 7).

For image-difference and image-quality predictions, it is also necessary to apply spatial filtering to the image data to eliminate any image variations at spatial

frequencies too high to be perceived. For example, the dots in a printed half-tone image are not visible if the viewing distance is sufficiently large. This computation is dependent on viewing distance and based on filters derived from human contrast sensitivity functions. Since the human contrast-sensitivity functions vary for luminance (band-pass with sensitivity to high frequencies) and chromatic (low pass) information, it is appropriate to apply these filters in an opponent space. Thus in image-quality applications of iCAM, spatial filters are applied in the IPT space. Since it is appropriate to apply spatial filters in a linear-signal space, they are applied in a linear version of IPT prior to conversion into the non-linear version of IPT for appearance predictions. Johnson and Fairchild have recently discussed some of the important considerations for this type of filtering in image-difference applications and specified the filters used based on available visual data.[22,28] Since the spatial filtering effectively blurs the image data, it is not desirable for image rendering applications in which observers might view the images more closely than the specified viewing distance. The result would be a blurrier image than the original. It is only appropriate to apply these spatial filters when the goal is to compute perceived image differences (and ultimately image quality). This is an important distinction between spatially-localized adaptation (good for rendering and image quality metrics) and spatial filtering (good for image quality metrics, bad for rendering). In image-quality applications, the spatial filtering is typically broken down into multiple channels for various spatial frequencies and orientations. For example, Daly,[14] Lubin,[15] and Pattanaik *et al.*[25] describe such models. More recent results suggest that while such multi-scale and multi-orientation filtering might be critical for some threshold metrics, it is often not necessary for data derived from complex images and for supra-threshold predictions of perceived image differences (one of the main goals of iCAM). Thus, to preserve the simplicity and ease of use of the iCAM model, single-scale spatial filtering with anisotropic filters was adopted.

Once the IPT coordinates are computed for the image data, a simple coordinate transformation from rectangular to cylindrical coordinates is applied to obtain image-wise predictors of lightness (J), chroma (C), and hue angle (h). Differences in these dimensions can be used to compute image difference statistics and those

used to derive image quality metrics. The overall Euclidean difference in IPT is referred to as ΔIm, for image difference, to distinguish it from a traditional color difference metric, ΔE, that includes no spatial filtering. In some instances, correlates of the absolute appearance attributes of brightness (Q) and colorfulness (M) are required. These are obtained by scaling the relative attributes of lightness and chroma with the appropriate function of F_L derived from the image-wise luminance map.

$$J = I \tag{13}$$

$$C = \sqrt{P^2 + T^2} \tag{14}$$

$$h = \tan^{-1}\left(\frac{P}{T}\right) \tag{15}$$

$$Q = \sqrt[4]{F_L}J \tag{16}$$

$$M = \sqrt[4]{F_L}C \tag{17}$$

$$\Delta Im = \sqrt{\Delta I^2 + \Delta P^2 + \Delta T^2} \tag{18}$$

For image rendering applications it is necessary to take the computed appearance correlates (JCh) and then render them to the viewing conditions of a given display. The display viewing conditions set the parameters for the inversion of the IPT model and the chromatic adaptation transform (all for an assumed spatially uniform display adaptation typical of low-dynamic-range output media). This inversion allows the appearance of original scenes or images from disparate viewing conditions to be rendered for the observer viewing a given display. One important application of such rendering is the display of high-dynamic-range (HDR) image data on typical displays.

A MODULAR IMAGE DIFFERENCE MODEL

A framework for a color image difference metric has recently been described.[7,28] This framework was designed to be modular in nature, to allow for flexibility and adaptation. The framework itself is based upon the S-CIELAB spatial extension to the CIELAB color space. S-CIELAB merges traditional color difference equations with spatial properties of the human visual system. This was accomplished as a spatial filtering pre-processing, before a pixel-by-pixel color difference calculation.[6]

The modular framework further extends this idea by adding several pre-processing steps, in addition to the spatial filtering. These pre-processing steps are contained in independent modules, so they can be tested and refined. Several modules have been described, and include spatial filtering, adaptation, and localization, as well as local and global contrast detection. Figure 2 shows a general flowchart with several distinct modules. These modules are described briefly below.

Spatial Filtering

The behavior of the human visual system in regards to spatially complex stimuli has been well studied over the years.[29] The contrast sensitivity function describes this behavior in relation to spatial frequency. Essentially the CSF is described in a post-retinal opponent color space, with a band-pass nature for the luminance channel and low-pass nature for the chrominance channels. S-CIELAB uses separable convolution kernels to approximate the CSF, and modulate image details that are imperceptible. More complicated contrast sensitivity functions that include both modulation and frequency enhancement were discussed in detail by Johnson and Fairchild.[28]

Spatial Frequency Adaptation

The contrast sensitivity function in this framework serves to modulate spatial frequencies that are not perceptible, and enhance certain frequencies that are most perceptible. Generally CSFs are measured using simple grating stimuli with care taken to avoid spatial frequency adaptation. Spatial frequency adaptation

essentially decreases sensitivity to certain frequencies based upon information present in the visual field.

Since spatial frequency adaptation cannot be eliminated in real world viewing conditions, several models of spatial frequency adaptation have been described.[28] These models alter the nature of the CSF based upon either assumptions of the viewing conditions, or based upon the information contained in the images themselves.

Spatial Localization

The band-pass and low-pass contrast sensitivity serve to modulate high-frequency information, including high-frequency edges. The human visual system is generally acknowledged to be very adept at detecting edges. To accommodate this behavior, a module of spatial localization has been developed. This module can be as simple as an image processing edge-enhancing kernel, although that kernel must change as a function of viewing distance. Alternatively, the CSF can be modified to boost certain high frequency information.

Local Contrast Detection

This module serves to detect local and global contrast changes between images. It is based upon the nonlinear mask based local contrast enhancement described by Moroney.[30] Essentially a low-pass image mask is used to generate a series of tone-reproduction curves. These curves are based upon the global contrast of the image, as well as the relationship between a single pixel and its local neighborhood.

Color Difference Map

The output of the modular framework is a map of color differences, ΔIm, corresponding to the perceived magnitude of error at each pixel location. This map can be very useful for determining specific causes of error, or for detecting systematic errors in a color imaging system. Often it is useful to reduce the error map into a more manageable dataset. This can be accomplished using image

statistics, so long as care is taken. Such statistics can be image mean, max, median, or standard deviation. Different statistics might be more valuable than others depending on the application, as perhaps the mean error better describes overall difference, while the max might better describe threshold differences.

IMAGE APPEARANCE APPLICATIONS (RENDERING)

Figure 3 illustrates the extensions to the basic iCAM model required to complete an image rendering process necessary for HDR image tone mapping. The components essential in this process are the inversion of the IPT model for a single set of spatially constant viewing conditions (the display) and the establishment of spatial filters for the adapting stimuli used for local luminance adaptation and modulation of the IPT exponential nonlinearity. While the derivation of optimal model settings for HDR image rendering is still underway, quite satisfactory results have been obtained using the settings outlined in Fig. 3.

The iCAM model has been successfully applied to prediction of a variety of color appearance phenomena such as chromatic adaptation (corresponding colors), color appearance scales, constant hue perceptions, simultaneous contrast, crispening, spreading, and image rendering.[13]

Since iCAM uses the same chromatic adaptation transform as CIECAM02, it performs identically for situations in which only a change in state of chromatic adaptation is present (*i.e.*, change in white point only). CIE TC8-01 has worked very hard to arrive at this adaptation transform and it is clear that no other model currently exists with better performance (although there are several with equivalent performance). Thus the chromatic adaptation performance of iCAM is as good as possible at this juncture.

The appearance scales of iCAM are identical to the IPT scales for the reference viewing conditions. The IPT space has the best available performance for constant hue contours and thus this feature is retained in iCAM. This feature makes accurate implementation of gamut-mapping algorithms far easier in iCAM than in other appearance spaces. In addition, the predictions of lightness

and chroma in iCAM are very good and comparable with the best color appearance models in typical viewing conditions. The brightness and colorfulness scales will also perform as well as any other model for typical conditions. In more extreme viewing conditions, the performance of iCAM and other models will begin to deviate. It is in these conditions that the potential strengths of iCAM will become evident. Further visual data must be collected to evaluate the model's relative performance in such situations.

The color difference performance of iCAM will be similar to that of CIELAB since the space is very similar under the reference viewing conditions. Thus, color difference computations will be similar to those already commonly used and the space can be easily extended to have a more accurate difference equation following the successful format of the CIE94 equations. (Following the CIEDE2000 equations in iCAM is not recommended since they are extremely complex and fitted to particular discrepancies of the CIELAB space such as poor constant-hue contours.)

Simultaneous contrast (or induction) causes a stimulus to shift in appearance away from the color of the background in terms of opponent dimensions. Figure 4 illustrates a stimulus that exhibits simultaneous contrast in lightness (the gray square is physically identical on all three backgrounds) and its prediction by iCAM as represented by the iCAM lightness predictor. This prediction is facilitated by the local adaptation features of iCAM.

Crispening is the phenomenon whereby the color differences between two stimuli are perceptually larger when viewed on a background that is similar to the stimuli. Figure 5 illustrates a stimulus that exhibits chroma crispening and its prediction by the iCAM chroma predictor. This prediction is also facilitated by the local adaptation features of iCAM.

Spreading is a spatial color appearance phenomenon in which the apparent hue of spatially complex image areas appears to fill various spatially coherent regions. Figure 6 provides an example of spreading in which the red hue of the

annular region spreads significantly from the lines to the full annulus. The iCAM prediction of spreading is illustrated through reproduction of the hue prediction. The prediction of spreading in iCAM is facilitated by spatial filtering of the stimulus image.

One of the most interesting and promising applications of iCAM is to the rendering of high-dynamic-range (HDR) images to low-dynamic-range display systems. HDR image data are quickly becoming more prevalent. Historically HDR images were obtained through computer graphics simulations computed with global-illumination algorithms (*e.g.*, ray tracing or radiosity algorithms) or through the calibration and registration of images obtained through multiple exposures. Real scenes, especially those with visible light sources, often have luminance ranges of up to six orders of magnitude. More recently, industrial digital imaging systems have become commercially available that can more easily capture HDR image data. It is also apparent that consumer digital cameras will soon be capable of capturing greater dynamic ranges. Unfortunately display and use of such data are difficult and will remain so since even the highest-quality displays are generally limited in dynamic range to about two orders of magnitude. One approach is to interactively view the image and select areas of interest to be viewed optimally within the display dynamic range. This is only applicable to computer displays and not appropriate for pictorial imaging and printed output. Another limitation is the need for capability to work with greater than 24-bit (and often floating point) image data. It is desirable to render HDR pictorial images onto a display that can be viewed directly (no interactive manipulation) by the observer and appear similar to what the observer would perceive if the original scene was viewed. For printed images, it is not just desirable, but necessary. Pattanaik *et al.*[25] review several such HDR rendering algorithms and it is worth noting that several papers were presented on the topic at the 2002 SIGGRAPH meeting, illustrating continued interest in the topic.

Since iCAM includes spatially-localized adaptation and spatially-localized contrast control, it can be applied to the problem of HDR image rendering. This

is not surprising since the fundamental problem in HDR rendering is to reproduce the appearance of an HDR image or scene on a low-dynamic-range display. Since the encoding in our visual system is of a rather low dynamic range, this is essentially a replication of the image appearance processing that goes on in the human observer and is being modeled by iCAM. Figure 7 illustrates application of the iCAM model to HDR images obtained from Debevec <www.debevec.org>. The images in the left column of Fig. 7 are linear renderings of the original HDR data normalized to the maximum presented simply to illustrate how the range of the original data exceeds a typical 24-bit (8-bits per RGB channel) image display. For example, the memorial image data (top row) have a dynamic range covering about six orders of magnitude since the sun was behind one of the stained-glass windows. The middle column of images represents a typical image-processing solution to rendering the data. One might consider a logarithmic transformation of the data, but that would do little to change the rendering in the first column. Instead the middle column was generated interactively by finding the optimum power-function transformation (also sometimes referred to as gamma correction; note that the linear images in the first column are already gamma corrected). For these images, transformations with exponents, or gammas, of approximately 1/6 (as opposed to 1/1.8 to 1/2.2 for typical displays) were required to make the image data in the shadow areas visible. While these power-function transformations do make more of the image-data visible, they required user interaction, tend to wash out the images in a way not consistent with the visual impression of the scenes, and introduce potentially-severe quantization artifacts in shadow regions. The rightmost column of images shows the output of the iCAM model with spatially-localized adaptation and contrast control (as shown in Fig. 3). These images both render the dynamic range of the scene to make shadow areas visible and retain the colorfulness of the scene. The resulting iCAM images are quite acceptable as reproductions of the HDR scenes (equivalent to the result of dodging and burning historically done in photographic printing). It is also noteworthy that the iCAM-rendered images were all computed with an automated algorithm mimicking human perception with no user interaction.

IMAGE QUALITY APPLICATIONS (DIFFERENCE PERCEPTIBILITY)

A slightly different implementation of iCAM is required for image quality applications in order to produce image maps representing the magnitude of perceived differences between a pair of images. In these applications, viewing-distance-dependent spatial filtering is applied in a linear IPT space and then differences are computed in the normal nonlinear IPT space. Euclidean summations of these differences can be used as an overall image difference map and then various summary statistics can be used to predict different attributes of image difference and quality. This process is outlined in Fig. 8 and described more fully in Johnson and Fairchild.[31]

Image quality metrics can be derived from image difference metrics that are based on normal color difference formulas applied to properly spatially-filtered images. This approach has been used to successfully predict various types of image quality data.[7] Figure 9 illustrates the prediction of perceived sharpness[32] and contrast[33] differences in images through a single summary statistic (mean image difference). This performance is equivalent to, or better than, that obtained using other color spaces optimized for the task.[7]

CONCLUSIONS AND FUTURE DIRECTIONS

Advances in imaging and computing technologies along with increased knowledge of the function and performance of the human visual system have allowed for the integration of models of color, spatial, and temporal vision to create a new type of color appearance model, referred to as an image appearance model. Such models show promise in a variety of applications ranging from image difference and image quality metrics to the rendering of image data. This paper described the framework of one example of an image appearance model referred to as iCAM and illustrated its applicability to HDR image tone mapping and image quality metrics. Recently, initial efforts have been made to incorporate psychophysical data on the time-course of chromatic adaptation[34] to extend the model to video appearance and quality applications.[35] Future efforts will be directed at completion of the spatio-temporal filters required for video difference metrics, the collection of more psychophysical data on image and

video appearance and differences, and the formulation of specific iCAM algorithms for various applications.

REFERENCES

1. R.S. Berns, A generic approach to color modeling, *Color Research and Application* **22,** 318-325 (1997).

2. M.D. Fairchild, Some hidden requirements for device-independent color imaging, *SID International Symposium,* San Jose 865-868 (1994).

3. G.J. Braun and M.D. Fairchild, General-purpose gamut-mapping algorithms: Evaluation of contrast-preserving rescaling functions for color gamut mapping, *Journal of Imaging Science and Technology* **44**, 343-350 (2000).

4. G.M. Johnson and M.D. Fairchild, Full-spectral color calculations in realistic image synthesis, *IEEE Computer Graphics & Applications* **19:4**, 47-53 (1999).

5. M. R. Luo, G. Cui, and B. Rigg, The development of the CIE 2000 Colour Difference Formula, *Color Research and Application* **26**, 340-350 (2001).

6. X. M. Zhang and B. A. Wandell, A spatial extension to CIELAB for digital color image reproduction, *Proceedings of the SID Symposiums*, 731-734 (1996).

7. G.M. Johnson and M.D. Fairchild, Darwinism of Color Image Difference Models, *Proc. of IS&T/SID 9th Color Imaging Conference*, 108-112 (2001).

8. CIE, The CIE 1997 Interim Colour Appearance Model (Simple Version), CIECAM97s, *CIE Pub. 131* (1998).

9. N. Moroney, M.D. Fairchild, R.W.G. Hunt, C.J Li, M.R. Luo, and T. Newman, The CIECAM02 color appearance model, *IS&T/SID 10th Color Imaging Conference*, Scottsdale, 23-27 (2002).

10. M.D. Fairchild, *Color Appearance Models,* Addison-Wesley, Reading, Mass., (1998).

11. K.M. Braun and M.D. Fairchild, Testing five color appearance models for changes in viewing conditions, *Color Research and Application* **22,** 165-174 (1997).

12. M.D. Fairchild, Image quality measurement and modeling for digital photography, *International Congress on Imaging Science '02*, Tokyo, 318-319 (2002).

13. M.D. Fairchild and G.M. Johnson, Meet iCAM: A next-generation color appearance model, *IS&T/SID 10th Color Imaging Conference*, Scottsdale, 33-38 (2002).

14. S. Daly, The Visible Differences Predictor: An algorithm for the assessment of image fidelity, in *Digital Images and Human Vision*, A. Watson, Ed., MIT, Cambridge, 179-206 (1993).

15. J. Lubin, The use of psychophysical data and models in the analysis of display system performance, in *Digital Images and Human Vision*, A. Watson, Ed., MIT, Cambridge, 163-178 (1993).

16. E.H. Land, Recent advances in retinex theory, *Vision Res*. **26**, 7-21 (1986).

17. E.H. Land and J.J. McCann, Lightness and the retinex theory, *J. Opt. Soc. Am.* **61**, 1-11 (1971).

18. B. Funt, F. Ciurea, and J.J. McCann, Retinex in Matlab, *Proc. of IS&T/SID 8th Color Imaging Conference*, 112-121 (2000).

19. F. Ebner, and M.D. Fairchild, Development and testing of a color space (IPT) with improved hue uniformity, *IS&T/SID 6th Color Imaging Conference,* Scottsdale, 8-13 (1998).

20. M.D. Fairchild, Considering the surround in device-independent color imaging, *Color Research and Application* **20,** 352-363 (1995).

21. M.D. Fairchild, Modeling color appearance, spatial vision, and image quality, *Color Image Science: Exploiting Digital Media,* Wiley, New York, 357-370 (2002).

22. G.M. Johnson and M.D. Fairchild, A top down description of S-CIELAB and CIEDE2000, *Color Research and Application*, in press (2003).

23. J.S. Babcock, J.B. Pelz and M.D. Fairchild, Eye tracking observers during color image evaluation tasks, *SPIE/IS&T Electronic Imaging Conference*, Santa Clara, in press (2003).

24. M.A. Webster and J.D. Mollon, Adaptation and the color statistics of natural images, *Vision Res.* **37**, 3283-3298 (1997).

25. S.N. Pattanaik, J.A. Ferwerda, M.D. Fairchild, and D.P. Greenberg, A multiscale model of adaptation and spatial vision for image display, *Proceedings of SIGGRAPH 98,* 287-298 (1998).

26. B.W. Keelan, *Handbook of Image Quality: Characterization and Prediction,* Marcel Dekker, New York, NY (2002).

27. P.G.Engledrum, Extending image quality models, *Proc IS&T PICS Conference,* 65-69 (2002).

28. G.M. Johnson and M.D. Fairchild, On contrast sensitivity in an image difference model, *Proc. of IS&T PICS Conference,* 18-23 (2001).

29. B.A. Wandell, *Foundations of Vision,* Sinauer Associates Inc. Sunderland, MA (1995*).*

30. N. Moroney, Local color correction using non-linear masking, *Proc. of IS&T/SID 8th Color Imaging Conference,* 108-111 (2000).

31. G.M. Johnson and M.D. Fairchild, Measuring images: Differences, Quality, and Appearance, *SPIE/IS&T Electronic Imaging Conference*, Santa Clara, in press (2003).

32. G.M. Johnson and M.D. Fairchild, Sharpness rules, *Proc of IS&T/SID 8th Color Imaging Conference,* 24-30 (2000).

33. A.J. Calabria and M.D. Fairchild, Compare and contrast: Perceived contrast of color images, *Proc. of IS&T/SID 10th Color Imaging Conference*, 17-22 (2002).

34. M.D. Fairchild and L. Reniff, Time-course of chromatic adaptation for color-appearance judgements, *Journal of the Optical Society of America A* **12,** 824-833 (1995).

35. M.D. Fairchild and G.M Johnson, Image appearance modeling, *Proc. SPIE/IS&T Electronic Imaging Conference,* in press (2003).

FIGURE CAPTIONS

Figure 1. Flowchart of the iCAM image appearance model.

Figure 2. Flowchart of a modular image difference metric.

Figure 3. Implementation of iCAM for tone mapping of HDR images.

Figure 4. (a) Original stimulus and (b) iCAM lightness, J, image illustrating the prediction of simultaneous contrast.

Figure 5. (a) Original stimulus and (b) iCAM chroma, C, image illustrating the prediction of chroma crispening. Original image from <www.hpl.hp.com/persona./Nathan_Moroney/>.

Figure 6. (a) Original stimulus and (b) iCAM hue, h, image illustrating the prediction of spreading.

Figure 7. Three HDR images from <www.debevec.org>. The leftmost column illustrates linear rendering of the image data, the middle column illustrates manually-optimized power-function transformations, and the rightmost column represents the automated output of the iCAM model implemented for HDR rendering (see Fig. 3).

Figure 8. Implementation of iCAM for image difference and image quality metrics.

Figure 9. iCAM image differences as a function of (a) perceived image contrast and (b) perceived image sharpness for a variety of image transformations. (Note: Desired predictions are a V-shaped data distributions since the perceptual differences are signed and the calculated differences are unsigned.)
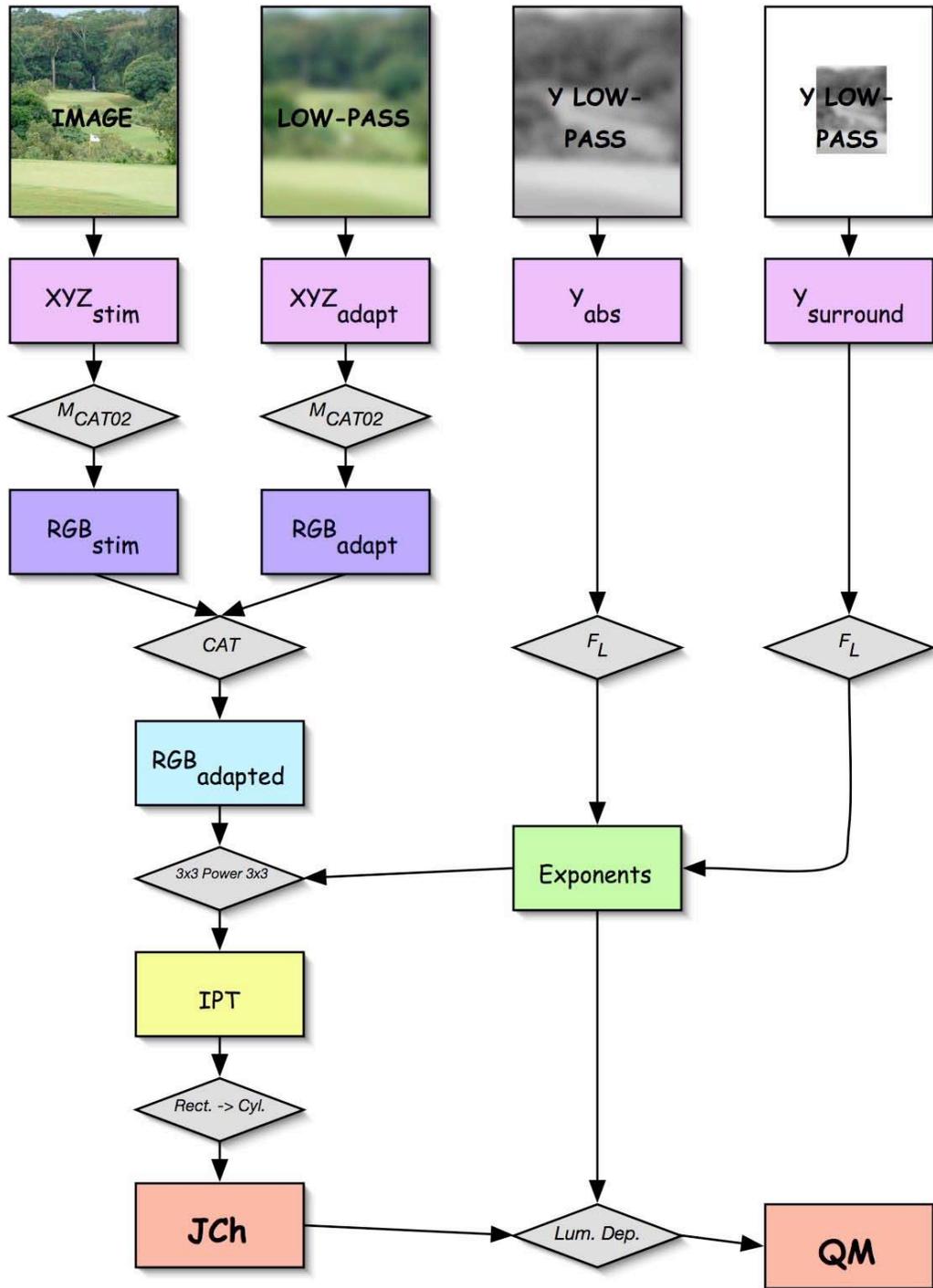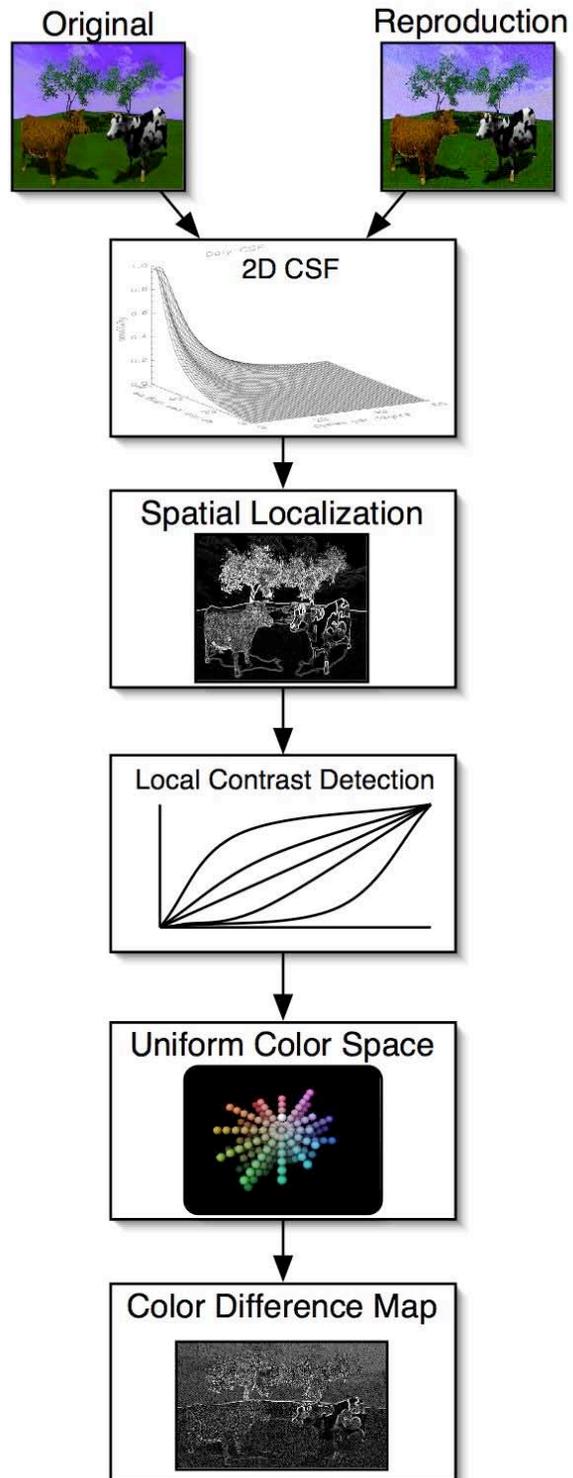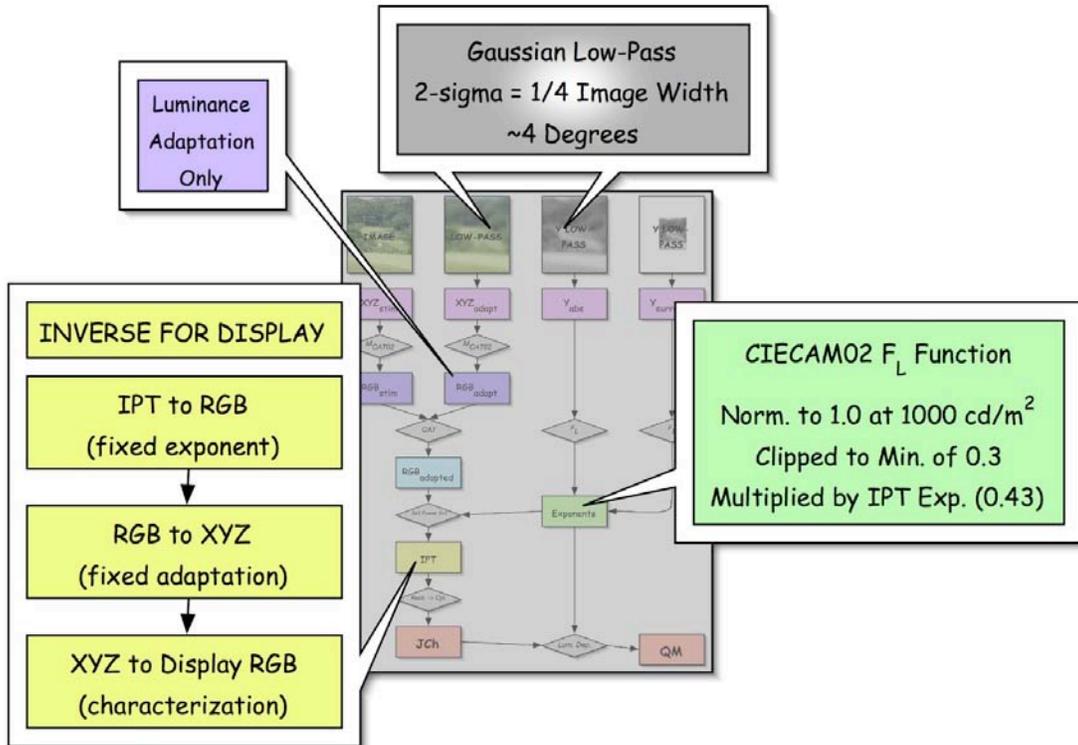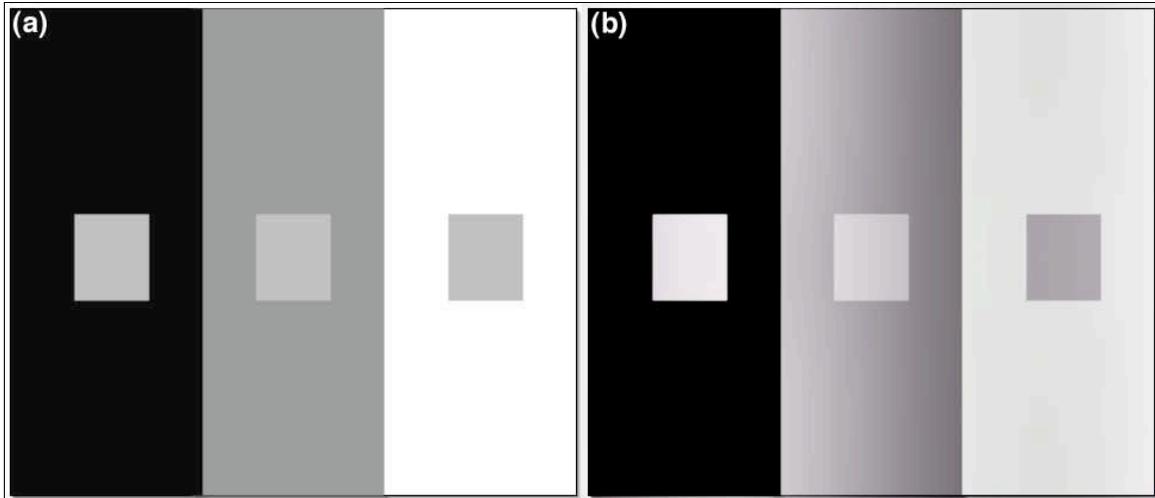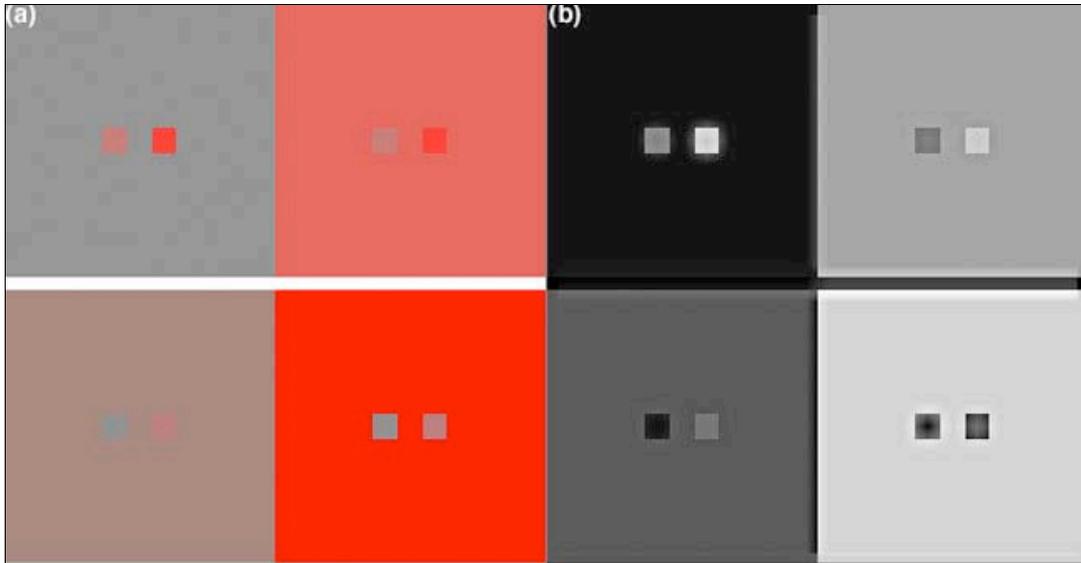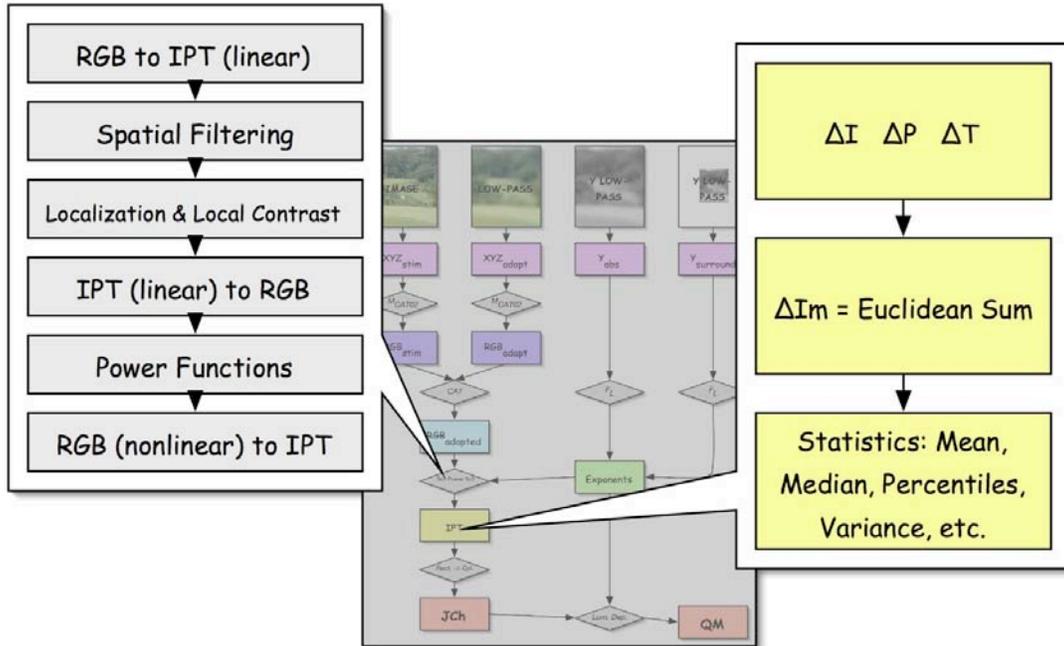
Fig. 1.

Fig. 2.

Fig. 3.

Fig. 4.

Fig. 5.

Fig. 6.

Fig. 7.

Fig. 8.

Fig. 9.