Evaluation of the Color Image and Video Processing Chain and Visual Quality Management for Consumer Systems

Abhijit Sarkar

B.E. Jadavpur University, Kolkata, India (2000)

M.S. Pennsylvania State University, Pennsylvania, USA (2005)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Color Science in the Chester F. Carlson Center for Imaging Science of the College of Science, Rochester Institute of Technology

May 2008

_____

Signature of the Author

_____

Accepted by Dr. Roy S. Berns,
Coordinator, M.S. Degree Program

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NY

CERTIFICATE OF APPROVAL

M.S. DEGREE THESIS

The M.S. Degree Thesis of Abhijit Sarkar has been examined and approved by two members of the Color Science faculty as satisfactory for the thesis requirement for the Master of Science degree

_____
Dr. Mark D. Fairchild, Thesis Advisor

_____
Dr. Roy S. Berns

THESIS RELEASE PERMISSION FORM

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

Title of Thesis: **Evaluation of the Color Image and Video Processing Chain and Visual Quality Management for Consumer Systems**

I, **Abhijit Sarkar**, hereby grant permission to the Wallace Memorial Library of Rochester Institute of Technology to reproduce my thesis in whole or part. Any reproduction will not be for commercial use or profit.

I additionally grant to the Rochester Institute of Technology Digital Media Library (RIT DML) the non-exclusive license to archive and provide electronic access to my thesis or dissertation in whole or in part in all forms of media in perpetuity. I retain all other ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

_____
Signature of the Author

_____
Date

Evaluation of the Color Image and Video Processing Chain and Visual Quality Management for Consumer Systems

Abhijit Sarkar

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Color Science in the Chester F. Carlson Center for
Imaging Science of the College of Science, Rochester Institute of Technology

ABSTRACT

With the advent of novel digital display technologies, color processing is increasingly becoming a key aspect in consumer video applications. Today's state-of-the-art displays require sophisticated color and image reproduction techniques in order to achieve larger screen size, higher luminance and higher resolution than ever before. However, from color science perspective, there are clearly opportunities for improvement in the color reproduction capabilities of various emerging and conventional display technologies. This research seeks to identify potential areas for improvement in color processing in a video processing chain. As part of this research, various processes involved in a typical video processing chain in consumer video applications were reviewed. Several published color and contrast enhancement algorithms were evaluated, and a novel algorithm was developed to enhance color and contrast in images and videos in an effective and coordinated manner. Further, a psychophysical technique was developed and implemented for performing visual evaluation of color image and consumer video quality. Based on the performance analysis and visual experiments involving various algorithms, guidelines were proposed for the development of an effective color and contrast enhancement method for images and video applications. It is hoped that the knowledge gained from this research will help build a better understanding of color processing and color quality management methods in consumer video.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my advisor Dr. Mark D. Fairchild for giving me the opportunity to work with him on this very interesting research project. It was a great learning experience for me.

I would like to thank Dr. Fairchild, Dr. Roy S. Berns and other faculty members in the Munsell Color Science Laboratory and the Center for Imaging Science for imparting to me the knowledge that builds the foundation of my professional career for the rest of my life.

I owe special thanks to Dr. Carl Salvaggio, for allowing me the opportunity to work with him on an independent research project and for his tremendous support. On the same note, I am grateful to many individuals at MCSL who formed an invaluable support system for me during my graduate studies at RIT. Dr. Mitch Rosen, Lawrence, Garrett, Ken, Mark (Updegraff), Val, Li, Ying, Mahdi, Mahnaz, Sunghyun, Shen, Jonathon, Philipp, Stacey and Erin, to name a few!

This research was made possible by a generous support of Intel Corporation. The test images, image sequences as well as outputs of the proprietary algorithms were provided by the sponsor. I am particularly indebted to Dr. Jorge E. Caviedes and Mahesh Subedar of Intel Corporation for their relentless guidance, valuable inputs and cooperation all along this collaborative research.

This section will not be complete if I did not acknowledge my parents and my elder sister for their forbearance, appreciation and encouragement throughout this long journey in the pursuit of my second master's degree!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

"Research is to see what everybody else has seen, and to think what nobody else has thought" ~ Albert Szent-Gyorgyi (Hungarian Biochemist, 1937 Nobel Prize for Medicine, 1893-1986)

Chapter 1

INTRODUCTION

This is an exciting time for color scientists and engineers engaged in display engineering and consumer video applications. We are witnessing a rapid and significant development in terms of state-of-the-art display technologies. This becomes more evident if we compare activities within the other segments of the imaging industry, for example printing, camera manufacturing etc, with those within the display industry. In terms of emerging technologies today, video and display industry is well ahead of others, and is likely to outpace the growth in other industries in the next several years to come.

Video processing has come a long way since the time of analog video, achieving picture quality that was unthinkable even in the early nineties. Color plays a vital role in achieving such quality. With the advent of novel digital display technologies, color processing is increasingly becoming a key aspect in consumer video applications that demand vivid, crisp and natural pictures without any visual artifacts. Larger screen size, higher luminance and higher resolution of today's displays require sophisticated color and image reproduction techniques. For any color scientist or engineer responsible for developing high-end color processing methods and algorithms suitable for various consumer video applications, it is important to keep abreast with all these latest development in display technologies.

At the same time, any future development efforts in color video processing must recognize the potential for improvement in color reproduction capabilities of various emerging display technologies as well as the more conventional ones. To this end, it will be important to bring fresh perspectives to the concept of color processing in a typical video processing chain in consumer video systems. Knowledge earned from years of color science and vision research must be applied in an effective manner while taking into account practical limitations in specific application context. More than ever before, there is a need today for concerted developmental efforts, and a close collaboration between the video researchers and color scientists.

1.1 Thesis Objective

The principal goal of this thesis research was twofold. The first goal was to evaluate various published color and contrast enhancement algorithms, and develop a novel algorithm that meets the objective of enhancing color and contrast in images and videos in an effective and coordinated manner. The second goal was the development and implementation of a psychophysical technique for the visual evaluation of color image and consumer video quality. Attaining these goals, it was assumed, would help build new knowledge specific to color processing and color quality management methods in consumer video.

1.2 Research Hypothesis

The main objective in color and contrast enhancement in video processing is to achieve the best possible combination of colorfulness and contrast in an efficient manner. Typically, the use of independent algorithms for color and contrast enhancement results in sub-optimal enhancement, and cumbersome tune-up. It was hypothesized that an integrated algorithm could help enhance

color and contrast in a more effective and coordinated manner. However, the fundamental hypothesis of this research was that color processing in a perceptually meaningful way would lead to superior image and video quality and this could be demonstrated through a properly designed visual experiment comparing results from various methods.

1.3 Thesis Organization

This thesis starts with a discussion on various aspects of color video processing in Chapter 2. Color specifications in various video standards, including color primaries and color coding standards are discussed. Various processes involved in a typical video processing chain in consumer video applications are reviewed. To appreciate the state of the art, the working principles for various modern display devices, as well as special video processing techniques employed in some of these devices are described.

In Chapter 3, various methods for determining and assessing the video quality are reviewed. These are categorized as engineering and psychophysical, based on their methodology. Several publications on subjective assessment of video quality are also reviewed.

Several previously published methods for color and contrast enhancement are discussed in Chapter 4. The development of the new algorithm is discussed, starting with the working requirement, the color space chosen for the development, and a description of the three key components of the algorithm.

Four published methods for color and contrast enhancement were implemented to evaluate their performance as well as to determine the most appropriate enhancement strategy for the new algorithm. Further, two existing color and contrast enhancement algorithms typical of consumer video applications were provided by the research sponsor. A quantitative performance analysis of these six algorithms and the newly developed algorithm was conducted. Chapter 5 contains the details of this analysis.

Chapter 6 contains a detailed description of the psychophysical experiments performed on the outputs of the two proprietary algorithms and the new algorithm, and an analysis of the results. Both still images and image sequences were used in the experiments. All aspects of the design of the visual experiments are discussed in this chapter, including display characterization, experimental setup, test images and video clips, viewing conditions and experimental method.

Finally, Chapter 7 contains concluding remarks and summarizes the research findings by outlining some key aspects relevant for the development of an effective color and contrast enhancement method for images and video applications.

"Knowledge is of two kinds: we know a subject ourselves, or we know where we can find information upon it." ~ Samuel Johnson (English Poet, Critic and Writer. 1709-1784)

Chapter 2

COLOR VIDEO PROCESSING

With the advent of novel digital display technologies, color processing is increasingly becoming a key aspect in consumer video applications. Consumers demand more vivid, crisp and natural pictures without any visual artifacts. Modern digital display systems require significantly more advanced color imaging techniques than what was adequate ten or fifteen years ago. Larger screen size, higher luminance and higher resolution of today's displays require sophisticated color and image reproduction techniques. While the display end of the video processing chain has undergone revolutionary change over the past decade, image capture capability of video cameras has also improved, along with the signal rate in the rest of the video processing pipeline, making it feasible to achieve the high quality color and image reproduction as we see today [**Kim 2005**]. Such development has not been possible in all aspects of video processing. The requirement of backward compatibility of the receivers imposed serious restriction on adopting new scanning formats that could lead to higher resolution, wider gamut and superior depth perception. However, at the same time, this constraint has stimulated innovative solutions without sacrificing compatibility. One compelling example is the introduction of color television in the sixties [**de Haan 2007**]. This compatibility issue essentially underscores the significance of various video standards in video processing.

This chapter starts with a description of the color specifications in various video standards, including color primaries and color coding standards. While the focus of this research is color and contrast enhancement in video (and also still images), various other processing not directly related to color or contrast significantly affect the overall picture quality. Thus, this chapter includes a review of various processes typical of a video processing chain in consumer video applications. These processes are independent of any specific display technology, and so have been classified as display-independent video processing. Various modern display technologies necessitate additional color and image processing, which are essentially display-dependent processing. To appreciate the state of the art, the working principles for various modern display devices, as well as special video processing techniques employed in some of these devices, have been briefly reviewed along with appropriate references for a detailed discussion. The concluding section takes a fresh look at the way color is handled in video processing, and how color science can be used to improve color quality in consumer video applications.

2.1 Color Specifications in Video Standards

In one of the early publications on HDTV colorimetry, DeMarsh [**DeMarsh 1990**] pointed out that the emerging display technologies presented an opportunity to improve the color quality of television images. Three potential areas were identified for colorimetric improvement in television systems, i) defining color characteristics, ii) extending color gamuts to take advantage of the newer display technologies, and iii) inclusion of constant luminance operation, which ensures conveying luminance information to the fullest extent in a television system. Accordingly, color specifications in video standards can be classified into three parts, namely, specification of color primaries and the white point, specification of the Opto-Electronic Transfer

Functions (OETFs), and specification of color coding for the compression and transmission of color information.

Before describing these color specifications, it is important to differentiate between standard-definition television (SDTV) and high-definition television (HDTV). This classification is based on resolution and scanning format.

There are two main scanning formats in SDTV, formats with 480 active pixel lines with interlacing and a frame rate of 29.97 Hz (denoted as 480i or 480i29.97) and 576 active pixel lines with interlacing and a frame rate of 25 Hz (denoted as 576i or 576i25). 480i systems with 4:3 aspect ratio (ratio of width to height of the displayed image) can have a resolution of 640x480, 704x480 or 720x480. 576i systems with 4:3 aspect ratio can have a resolution of 768x576 or 720x576 or 948x576. Widescreen 16:9 format supports resolutions 720x483 and 720x576.

On the other hand, HDTV has higher resolution, typically 0.75 million pixels or more [**Poynton 2003**]. Most common HDTV formats are 1280x720 with progressive scanning at a frame rate of 60 Hz (denoted as 720p or 720p60) and 1920x1080 with progressive or interlaced scanning at a frame rate of 24 (only progressive) or 30 Hz (1080p30/1080i30, 1080p24/1080p30). 1080p60 and 1080p120 formats are also possible. The aspect ratio in HDTV is 16:9.

2.1.1 Color Primaries

The set of colorants used for a particular coloration process is referred to as a primary set [**Berns 2000**]. Except for some state-of-the-art technologies, most displays use three primaries, namely red, green and blue. All video standards describe color primaries in terms of the chromaticites of RGB and the white point (an achromatic color with highest luminance achievable in a given system), and thereby the color gamut of any device that complies with a given standard. Color gamut describes the range of colors produced by a coloration system [**Berns 2000**], including displays. Color primaries defined in various widely known standards are discussed below [**Poynton 2003**]:

CIE: CIE color primaries were defined in CIE 1931 standard observer. CIE Illuminant B was defined as the white point. CIE primaries are no longer used in video coding or reproduction.

NTSC: In 1958, National Television System Committee (NTSC) standardized color primaries, primarily to be used for color Cathode-Ray Tube (CRT) displays. The white point has the chromaticities of CIE Illuminant C. The NTSC primaries were chosen such that largest color gamut could be achieved with the commercially available phosphors for CRT monitors. These primaries and the white point are no longer used in displays, but NTSC specification is still used as the industry benchmark. Compared to NTSC compliant displays of the past, modern CRTs have brighter and more efficient phosphors, even though NTSC displays produced more saturated red and yellow [**Susstrunk 1999**].

EBU Tech 3213: In 1975, European Broadcasting Union (EBU) published a standard for chromaticity tolerances for studio monitors conforming to 576i SDTV systems (standard-definition televisions with 576 active picture lines and interlaced scanning), known as EBU Tech. 3213. D65 was used as the white point.

SMPTE RP 145: Society of Motion Pictures and Television Engineer (SMPTE) set color standards for 480i SDTV systems (standard-definition televisions with 480 active picture lines and interlaced scanning) and early 1035i30 HDTV systems (high-definition televisions with 1035 active picture lines, 29.97 frame rate and interlaced scanning). This standard also uses D65 as white point.

ITU-R BT 709/sRGB: In 1990, International Telecommunication Union's Radiocommunication Sector (ITU-R) recommended standard primaries for high-definition television (HDTV), formally known as ITU-R BT 709, or simply Rec. 709. The Rec. 709 primaries are incorporated into the sRGB specifications widely used in the computing and computer graphics community, but sRGB uses D50 white point, while Rec. 709 uses D65. These primaries are the most widely used color primaries for studio video and modern display systems. Note that displays using Rec. 709 primaries have a color gamut that is 71% of the standard NTSC color gamut obtained from conventional CRT displays [**Ok 2005**].

Adobe RGB: These primaries were designed to provide large color gamut with RGB as the working space, and were based on SMPTE-240M standard and later renamed as Adobe RGB 98

[**Susstrunk 1999**]. The primaries have been adopted in some of the modern wide gamut CRT and LCD displays [**Kwak 2005**].

Table 2.1 lists the chromaticities of the color primaries defined by various standards. Note that x+y+z = 1, thus z can be easily calculated from the given data. Figure 2.1 plots some of these primaries on the chromaticity diagram.

Table 2.1 Color primaries used in video processing

	Red		Green		Blue		White Point		
	x	y	x	y	x	y	Illuminant	x	y
CIE	0.7347	0.2653	0.2737	0.7174	0.1665	0.0089	B	0.3484	0.3516
NTSC	0.67	0.33	0.21	0.71	0.14	0.08	C	0.31	0.316
EBU Tech. 3213	0.64	0.33	0.29	0.6	0.15	0.06	D65	0.3127	0.329
SMPTE RP 145	0.63	0.34	0.31	0.595	0.155	0.07	D65	0.3127	0.329
ITU-R BT 709	0.64	0.33	0.3	0.6	0.15	0.06	D65	0.3127	0.329
sRGB	0.64	0.33	0.3	0.6	0.15	0.06	D50	0.3457	0.3585
Adobe RGB	0.64	0.34	0.21	0.71	0.15	0.06	D65	0.3127	0.329

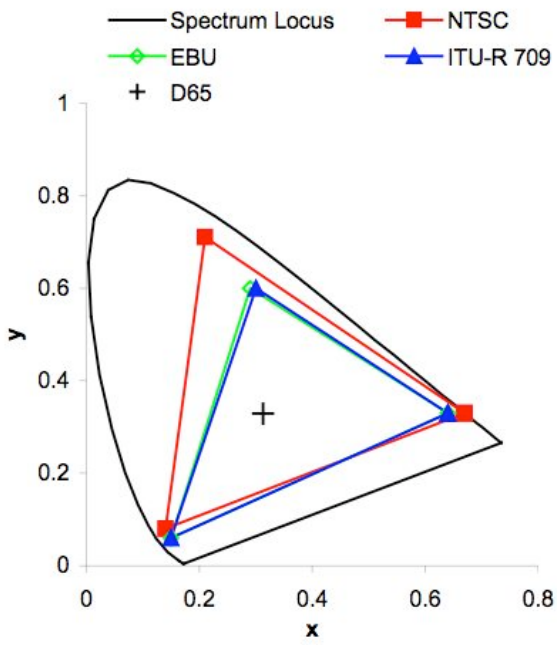


Fig. 2.1 Color primaries defined in various video standards [Source: **Susstrunk 1999**, Fig. 7]

A 3x3 matrix transformation can be used to convert from one set of primaries to another. For example, the following equation can be used to convert CIE XYZ to Rec. 709 primaries [**Poynton 2003**]:

$$\begin{bmatrix} R_{709} \\ G_{709} \\ B_{709} \end{bmatrix} = \begin{bmatrix} 3.2405 & -1.5372 & -0.4985 \\ -0.9693 & 1.876 & 0.042 \\ 0.0556 & -0.2040 & 1.0573 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2.1}$$

Other primaries can similarly be obtained.

2.1.2. Opto-Electronic Transfer Functions (OETFs)

Opto-Electronic Transfer Function (OETF) refers to the intrinsic nonlinear function in a CRT display that converts the input voltage to luminance. Gamma correction in video cameras essentially pre-compensates for this nonlinearity and achieves perceptual coding at the same time

[**Poynton 2003**]. Various video standards have defined the transfer function differently. As mentioned earlier, Rec. 709 is the international standard for HDTV. The transfer function specified by Rec. 709 is also used as SDTV studio standard. The transfer function specified for encoding is linear with a slope of 4.5 below a linearized signal value of 0.018, and follows an exponential curve above between 0.018 and 1, as shown below:

$$V' = \begin{cases} 4.5L & 0 \le L \le 0.018 \\ 1.099L^{0.45} - 0.099 & 0.018 \le L \le 1 \end{cases} \tag{2.2}$$

Here, V′ is nonlinear gamma corrected signal, R′, G′ or B′. Equation (2.3) can be inverted to decode the signal and retrieve original RGB values as follows:

$$L = \begin{cases} \dfrac{V'}{4.5} & 0 \le V' \le 0.081 \\ \left(\dfrac{V' + 0.099}{1.099}\right)^{\frac{1}{0.45}} & 0.081 \le V' \le 1 \end{cases} \tag{2.3}$$

These definitions do not take into account any display specific tone scale alterations. L and V' in above equations are normalized to unity. However, when represented in 8-bits, the values are scaled between 16 and 235 to allow for headroom above reference white and footroom below reference black, which are necessary to accommodate filter overshoot and undershoot respectively [**Poynton 2003**].

This and other transfer functions defined by various standards are discussed in detail in [**Poynton 2003**].

2.1.3. Color Coding Standards

This section outlines various color coding standards followed in the industry, details of which are beyond the scope of this thesis, but are available in [**Poynton 2003**].

Signal coding in video systems involve three steps as described below:

Step 1 – Gamma correction: A nonlinear transfer function is applied to each of the linear R, G and B signals. This function, often called gamma correction, is comparable to a square root and takes care of the nonlinearity in the conventional CRT display. Gamma correction results in nonlinear signals denoted as R′, G′ and B′. Since human visual system is sensitive to luminance changes over a wide range of luminance values, nonlinear image coding needs to be used to achieve perceptual uniformity. The nonlinear transfer function gamma is used to approximate our lightness perception. Note that in encoding, gamma correction is used before converting RGB to an opponent based color space. This is important from engineering standpoint, to reduce computational complexity in the decoding stage.

Step 2 – Formation of luma and chroma signals: From the nonlinear signals R′, G′ and B′, luma component Y′ and color difference components (B′- Y′) and (G′- Y′) are formed. Note that the term luma is used to differentiate this component from luminance. Y′ computations for SDTV (as per Rec. 601) and HDTV (as per Rec. 709) are as follows:

$$^{601}Y' = 0.299R' + 0.587G' + 0.114B' \tag{2.4}$$

$$^{709}Y' = 0.2126R' + 0.7152G' + 0.114B' \tag{2.5}$$

In component digital video, MPEG and Motion-JPEG, the color difference components are scaled to form C_B and C_R respectively.

Step 3 – Chroma subsampling: The color difference components are subsampled. Chroma subsampling is the process of reducing data capacity needed to transmit color information, while maintaining full luma information. This takes advantage of the relatively low sensitivity of our visual system to detect color differences compared to luminance. Chroma subsampling does not typically result in a perceptual loss of chromatic details in video, but is the key source of artifacts resulting from color processing in video.

Different schemes are available for chroma subsampling. If we consider a 2x2 pixel array of R′G′B′ components, converting the nonlinear RGB to Y′$C_B C_R$ will result in 12 bytes of data in 8-bit systems. This is denoted as 4:4:4 Y′$C_B C_R$. In 4:2:2 sampling included in Rec. 601 for studio digital video, the color difference components are subsampled horizontally by a factor of 2, with C_B and C_R being coincident with even numbered Y′ samples. This consumes 8 bytes, instead of 12. In 4:1:1 scheme, C_B and C_R are subsampled horizontally by a factor of 4, and coincide with every fourth Y′ sample. This scheme requires 6 bytes only. In 4:2:0 sampling scheme used in JPEG, H-261, MPEG-1, MPEG-2 etc, C_B and C_R are subsampled both horizontally and vertically each by a factor of 2. Thus, there is one set of C_B and C_R components for the four Y′ samples. The number of bytes used in this case is also 6.

SDTV color coding: Studio applications require that full 8-bit range (between 0-255) not be used for luma scaling so as to leave a headroom and footroom to accommodate higher output

resulting from filter operation and misadjusted equipment. In an 8-bit system, offsets of 16 and 128 are added to the luma and chroma signals respectively. Luma reference levels are 16 and 235, and chroma reference levels are 16 and 240. Digital values 0 and 255 are used in the video data only for synchronization purposes. Following equation is used for computing 8-bit Rec. 601 $Y'C_BC_R$ from gamma corrected digital counts normalized between 0 and 1:

$$\begin{bmatrix} Y' \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{2.6}$$

HDTV color coding: As mentioned before, ITU-R Rec. BT.709 standard is the most commonly used standard for HDTV. Rec. 709 $Y'C_BC_R$ can be computed from $R'G'B'$ using the following equation:

$$\begin{bmatrix} Y' \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 46.559 & 156.629 & 15.812 \\ -25.664 & -86.336 & 112 \\ 112 & -101.730 & -10.270 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{2.7}$$

NTSC and PAL color coding: NTSC (acronym for National Television System Committee) and PAL (acronym for Phase Alternating Line) coding are also known as composite coding, where quadrature modulation is applied to combine two color difference components into a modulated chroma signal, which is then added to the luma signal through frequency interleaving to form a composite signal. Composite decoding breaks the composite signal into constituent luma and chroma signals. A composite signal enabled the use of new color receivers in early sixties to receive black and white broadcast, and thus providing required backward compatibility to the color television sets. However, these coding schemes are not generally used anymore due to the

resulting artifacts, and also because of the availability of adequate bandwidth to carry component signals as in Rec. 601 and Rec. 709.

2.2 Display-Independent Video Processing

The components of a typical video processing pipeline in consumer video applications are shown in Figure 2.2 [**Caviedes 2008, Klompenhouwer 2004**]. Encoded signal is transmitted from a broadcasting station and is received by the signal receiver, which then passes it to the decoding module. The video stream then passes through various post processing routines for artifact removal, format conversion and enhancement. Next, a color space transformation is applied in case of different primaries for source and display formats. A gamma correction ensures correct tone reproduction on the display, while quantization is required to obtain discreet digitized values for each display channel signal. The processed video is then ready to be displayed on a designated display device.

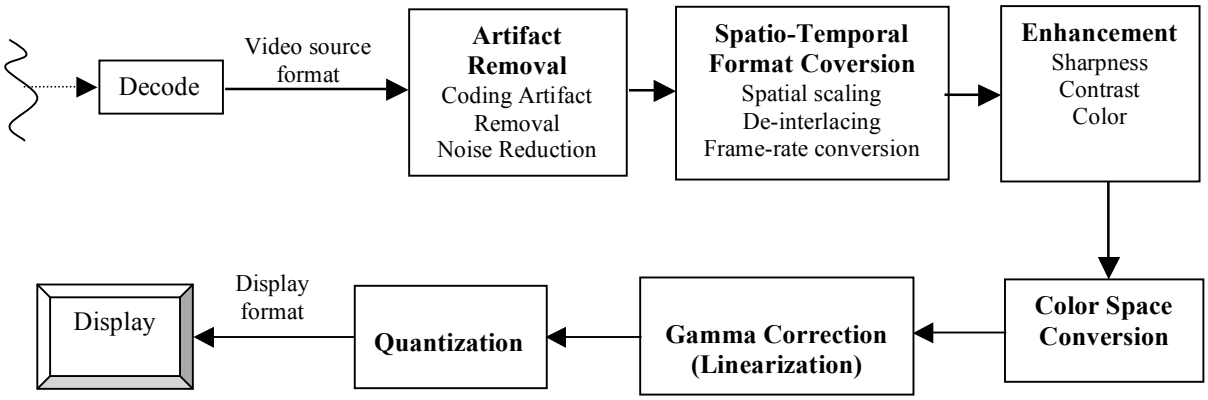


Fig. 2.2 A typical video processing pipeline in consumer video systems

What follows is a brief description of various processes mentioned above. A detailed treatise of these processes is available elsewhere [**de Haan 2003**].

2.2.1 Artifact Removal

Noise and other degradations in video must be removed during the post-processing stage. These degradations occur throughout the operating broadcasting chain, mainly during encoding and transmission. Many of these degradations do not directly relate to colors in video, but it is important to be aware of these issues, as the artifact removal is an important step in the video processing chain. This has a significant impact on the subsequent enhancement process as well as ultimate picture quality. The process of artifact removal can be broadly classified into two categories, namely, coding artifact removal and noise reduction.

2.2.1.1 Coding Artifact Removal: Most of the artifacts commonly encountered in video result from video compression, or encoding. Several encoding standards are available for video compression. The artifacts resulting from compression depends on three factors, data source, coding bit rates and compression algorithms. The coding artifacts can be dealt with at the encoding end (preprocessing) or after decoding in the post processing stage, as shown in Figure 2.2. Various types of video compression artifacts are briefly described below. Detailed reviews are available in [**Shen 1997**], [**Yuen 1998**] and [**Wu 2006**].

i) Blocking: Blocking effects are the discontinuities found at the boundaries of adjacent blocks in a reconstructed frame, resulting from each block being encoded without consideration of the correlation between adjacent blocks. This is the most common artifact that results from both

JPEG and MPEG compression standards and is more visible in smooth areas with low lightness, as shown in Figure 2.3. Note that these artifacts will be further enhanced during color and contrast enhancement.

Fig. 2.3 Blocking artifact [Source: **Yuen 1998,** Fig. 1]

ii) Blurring/ Color Bleeding: Blurring is the effect of a loss or smoothing of spatial details in areas with moderate to high spatial frequencies. The perception of lack of contrast weakens as the viewing distance increases. Blurring is associated with luminance channel, while the corresponding effect on chrominance channels cause a smearing in the areas with a drastic variation in the chrominance values. This is called *color bleeding* and is visible in Figure 2.4 around the arm.

iii) Ringing: The ringing effect is caused by a coarse quantization of high frequency components in the frequency domain. The effect is most apparent along the high contrast edges in otherwise smooth areas, and appears as shimmering or rippling outwards from the edge extending up to the adjacent block boundary. Ringing effect occurs not only in the luminance channel, but also in the

chrominance components. In this case, the ringing appears as wave-like transitions of color because of which colors in the affected areas do not correspond to the colors of the surrounding areas. Chrominance ringing is coarser than luminance ringing due to the chroma subsampling (discussed later under color coding). Figure 2.4 shows an example of color ringing, in the form of high frequency changes around the edge of the table. The ringing effect is the most noticeable artifact in the subband/wavelet coding schemes at low bit rates.

Fig. 2.4 Ringing and color bleeding effect [Source: **Yuen 1998,** Fig. 8]

iv) Staircase Effect: In block based transform coding (e.g. Discreet Cosine Transform), when a diagonal edge extends over several consecutive blocks, coarse quantization results in a series of horizontal or vertical steps. This is known as staircase effect. An example is shown in Figure 2.5. It is mainly noticeable in case of small block sizes (e.g. 6x6), but for larger blocks, it appears as occasional misalignment in an otherwise smooth edge [**Yuen 1998**].

v) Mosaic Patterns: This artifact also results from block based transform coding, due to the apparent mismatch between the contents of adjacent blocks. It manifests as a block having a

contour or texture dissimilar with those of the adjacent blocks. Mosaic pattern generally appears along with the blocking effect. An example is shown in Figure 2.6.

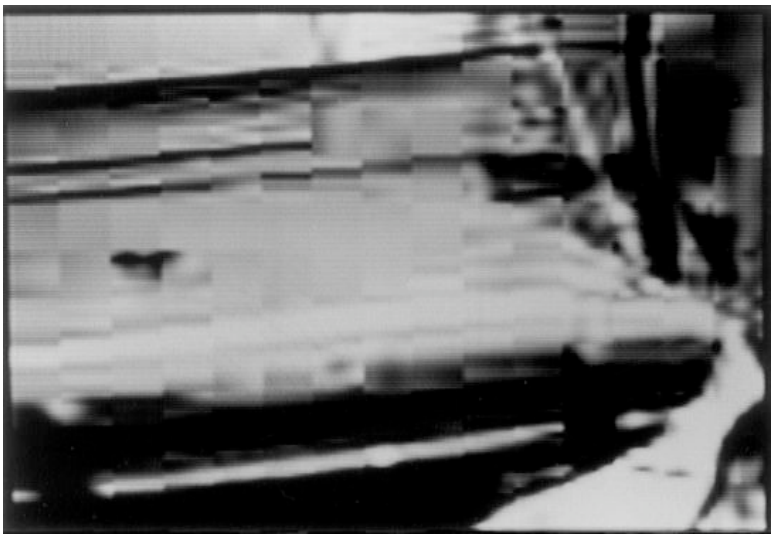


Fig. 2.5 Staircase effect [Source: **Yuen 1998,** Fig. 9]

Fig. 2.6 Mosaic patterns visible on the character's face [Source: **Yuen 1998,** Fig. 11]

vi) False contouring: This artifact results from direct quantization of pixel values. It typically occurs where the image has a uniform background with a color gradient, and shows up like a series of step-like gradations in a smoothly textured area (Figure 2.7).

Fig. 2.7 False contouring [Source: **Yuen 1998,** Fig. 17]

vii) Motion Compensation Mismatch: Motion-compensated coding is a predictive coding technique applied to video sequences involving motion. In other words, motion compensation (MC) is a technique for describing a video frame in terms of transformation from a reference frame to the current frame. Block processing assumes identical motion of all the pixels within a given block. However, this leads to problems around the boundaries of the moving objects in a video, as the blocks in these areas encompass both the moving objects and the stationary background. Figure 2.8 shows an example of this artifact in the form of high frequency spatial noise around the boundary between the arm and the background.

Fig. 2.8 Motion-compensated mismatch effect around the boundaries of moving objects [Source: **Yuen 1998,** Fig. 20]

viii) Mosquito Effect: It is a temporal artifact typically visible in areas with uniform background as variation in luminance or chrominance levels around sharp edges or moving objects in a sequence. This is caused by different coding used for the same area in consecutive frames in a sequence.

ix) Static Area Fluctuation: In an image sequence, areas with no motion but high spatial frequency can have fluctuations similar to mosquito effect. However, these fluctuations are not perceivable when there is motion. As in case of mosquito effect, varied coding of the same area leads to static area fluctuation artifacts.

x) Flickering: Intensity flicker is defined as unnatural temporal fluctuations of frame intensities that do not originate from the original scene. The flicker is a spatially localized effect that occurs in regions of substantial size [**Bovik 2005**]. This is a video compression artifact and can be caused by various factors, for example, a random noise due to digitization of the original content, a quantization noise caused by compression, unequal quantization levels between adjacent frames, variation in the bit rates assigned to different frames or to different areas within the same frame etc [**Shen 1997**].

xi) Chrominance Mismatch: In block based coding, only luminance information is used, but luminance correlation between blocks of pixel may not apply to chrominance information. This may result in chrominance mismatch, manifested as a misplacement of a block in comparison with its own color and the color of the surrounding areas in the video frame.

2.2.1.2 Noise Reduction: Noise is a form of *point degradation* in video. Point degradation affects the gray level of pixels without impacting spatial details or motion. While noise generally refers to a random process, for example, thermal noise, lightning etc, noise can also be coherent, for example, noise due to channel cross-talk [**de Haan 2003**]. De Haan classifies noise into four major categories:

Amplitude distribution: The amplitude distribution can be uniform (e.g. quantization noise), Gaussian (e.g. thermal noise), or Poisson (quantum noise).

Signal dependency: Noise can be independent of the video signal (additive noise), or can be a function of it (multiplicative noise).

Domain: Noise properties may vary along horizontal, vertical or temporal dimensions

Frequency characteristics: In any given dimension, noise may have different frequency characteristics, for example, white noise, 1/f noise and triangular noise.

Noise reduction involves determining the correlation among pixels in an image or image sequence, either in the spatial or in the temporal domain. It is essentially realized through noise filtering [**de Haan 2003**].

2.2.2 Spatio-Temporal Format Conversion

Format conversion is essential in consumer video applications to distribute the same video content using various video broadcasting formats, and present them simultaneously on various modern display devices with different characteristics. Format conversion can be either spatial or temporal, or both. The video signal represents a sequence of images, each described by the spatial addressing format depicting the number of pixels in horizontal and vertical directions. The number of images per second determines the temporal addressing format [**Klompenhouwer 2004**].

2.2.2.1 Spatial Scaling: If the number of pixels on a video line or the number of lines in an input image does not match those of the display, spatial scaling is used to resolve this issue. An

example is the aspect ratio conversion employed in wide screen displays. Earlier method of spatial scaling involved simple pixel repetition and pixel dropping, for up scaling and downscaling respectively. While these methods are still used for format conversion in the vertical and temporal domains, they are not preferred solutions because of the loss in the details in the image and resulting artifacts like jagged edges. A common artifact resulting from wrong up sampling is called “chroma bug”, appearing in the form of tiny horizontal streaks of incorrect coloration. Contemporary methods employ poly-phase filtering for sample rate conversion, which can be either integer scaling or fractional scaling. Integer scaling is achieved through decimating and interpolating low-pass filters for downscaling and up scaling respectively. Fractional or non-integer scaling applies a combined method of up sampling (by adding zero valued samples and interpolating) and down sampling (by dropping samples from the input signal) [**de Haan 1999**]. Polyphase filtering yields higher quality scaling than simple pixel dropping, as shown in Figure 2.9. The underlying assumption in employing poly-phase filter based sample rate conversion stems from the sampling theorem, i.e. the sampling frequency must be more than twice the bandwidth of the input signal.

Fig. 2.9 An example of spatial sampling: down sampling by pixel dropping (left) and polyphase filtering (right) [Source: **Klompenhouwer 2004** Fig. 4]

2.2.2.2 De-Interlacing: Video scanning represents a sequential array of pixels used to capture, transmit and display the content at a given frame rate. Interlacing, a scanning scheme fairly common in conventional broadcast television and HDTV, helps reduce transmission bandwidth without compromising resolution. In this scheme, the complete frame is broken down into two fields, each with half of the total number of scanning lines (odd and even lines). The second field is delayed by half the frame time from the first [**Poynton 2003**]. This helps achieve improved picture quality in CRT based displays with sufficient refresh rate (to avoid perceivable flicker), without consuming extra bandwidth. However, many modern displays like Liquid Crystal Displays (LCD), micromirror or Digital Light Projector (DLP) and Plasma Display Panels (PDP) use progressive scanning instead of interlacing (video processing in these devices are described later). In this scheme, the entire frame is scanned, transmitted and displayed line by line, from top to bottom into a single field. Thus, depending on the application, de-interlacing might be needed for high quality video scanning format conversion.

De-interlacing can be of two types, field repetition and line repetition. Field repetition is a temporal process where one or more fields are buffered in memory and consecutive fields are added (weaving) or averaged (blending) to form a single frame. Line repetition is spatial and achieved by extending each field to form the entire frame. However, de-interlacing is typically a lossy process and leads to various artifacts. Field repetition typically leads to motion artifacts while line repetition causes jagged edges [**de Haan 2003**]. Figure 2.10 shows an example of artifacts resulting from de-interlacing on moving objects.

Original

Staircase artifacts resulting from de-interlacing using line repetition

Feathering artifacts resulting from de-interlacing using field repetition

Fig. 2.10 An example of artifacts resulting from de-interlacing [Source: **Klompenhouwer 2004** Fig. 5]

Advanced de-interlacing technique employs motion estimation and compensation to avoid motion related artifacts.

2.2.2.3 Frame-Rate Conversion: Various video sources use various frame rates (also referred to as picture rates). Video cameras generally use 50-60 Hz, motion picture films are recorded at 24, 25 or 30 Hz, and the TV or PC displays have a frame rate between 50 and 120 Hz [**de Haan 2003**]. Thus, frame rate conversion may be needed before some video content can be displayed on a given device.

One of the common methods for frame rate conversion is 2-3 pulldown, which transfers film at 24 Hz to video at 60 Hz. The first film frame is transferred to two video fields, while the second is transferred to three, resulting in five video fields. Following frames follow the same order, with the role of first and second fields reversed [**Poynton 2003**]. However, this process, or any frame rate conversion where the difference between input and output rates is less than 30 Hz,

typically results in jerky motion, or motion judder as they are called, shown in Figure 2.11 [**de Haan 2003**].

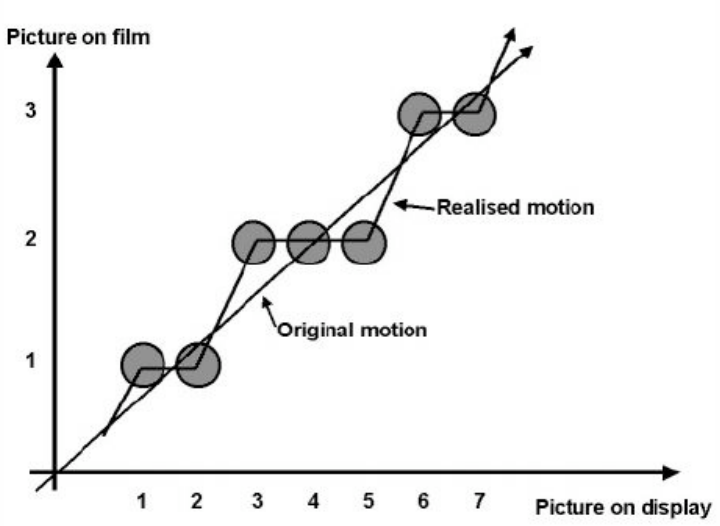


Fig. 2.11 Original and perceived motion in 2-3 pulldown [Source: **de Haan 1999,** Fig. 5]

In other cases where the difference between input and output frame rate is more than 30 Hz, picture repetition results in motion blur. In this case, we perceive an object simultaneously at two locations in the output video, at intermediate locations along the motion trajectory where we expect the object to appear, and at locations where the object is repeatedly shown by the display. This is illustrated in Figure 2.12.

As in case of de-interlacing, above problems can be remedied through motion estimation and compensation techniques. These techniques can be implemented in a single integrated circuit [**de Haan 1999**].

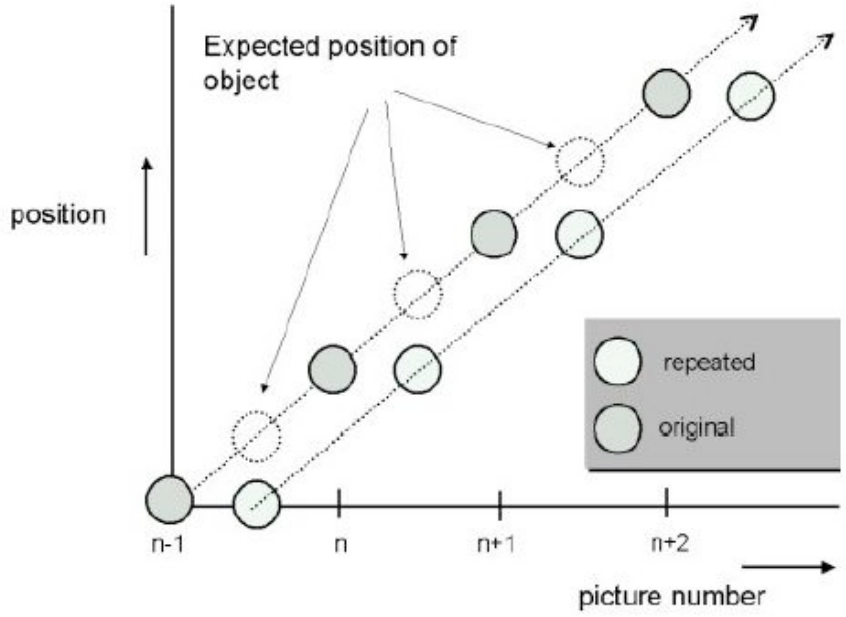


Fig. 2.12 Original and perceived motion when difference between the input and output frequency is more than 30 Hz [Source: **de Haan 1999,** Fig. 6]

2.2.3 Enhancement

The quality of color reproduction in displays is usually based on the preferences of individual viewers. Thus, the purpose of image enhancement in video is to improve the subjective picture quality, and not a reliable reproduction of the original video content. While a systematic improvement of the picture quality requires the knowledge of the process or processes that degraded it in the first place, in practice, information on the prior processes the original data was subjected to is seldom known [**Poynton 2003**]. Even though image enhancement processing can be a "matter of taste", there are certain commonly agreed upon image features that generally enhance the image quality [**de Haan 2003**]. For example, a bright and colorful image with high contrast is preferred by most observers [**Kim 2005**]. Image enhancement in consumer video applications primarily aims at improving three perceptual image attributes, namely, sharpness,

contrast and color. The following subsections discuss common enhancement strategies for these attributes.

2.2.3.1 Sharpness: Sharpness is the image attribute that indicates the extent to which edges and other fine details in an image are visible. Sharpness can be enhanced by boosting the high and/or middle frequency components in an image using linear filtering methods, also known as *linear peaking*, or by edge enhancement or transient improvement. Peaking filters can be horizontal or vertical, or both (two-dimensional). However, enhancing the diagonal frequencies in an image is not preferable as it amplifies the noise, but does not improve perceived sharpness. Linear peaking in general enhances the noise, which is most visible against uniform background. The method commonly applied to prevent this noise enhancement is called *coring*, which introduces an amplitude threshold below which linear peaking does not take place. The other alternative is an adaptive peaking signal that depends on the amplitude of the high frequency content. This is termed as *dynamic peaking*. In more advanced solutions, peaking is adaptive to local chroma, in particular the skin tone. In case of skin tone, the amount of peaking is controlled to avoid enhancing skin imperfections or introduce wrinkles.

The other method for sharpness improvement, i.e. edge enhancement, is a nonlinear process, which involves detecting the edges and compressing their extent. When applied to the luminance signal, this is called Luminance Transient Improvement (LTI), and when applied to chrominance signal, this is termed as Color Transient Improvement (CTI) [**de Haan 2003**]. When sharpness enhancement is conducted on chroma-subsampled content, or content with poor chroma upscaling, it may lead to artifacts.

2.2.3.2 Contrast: Contrast is defined as the difference between the maximum and minimum luminance in a stimulus divided by the sum of those two luminances [**Fairchild 2005**]. This is true for images and video as well. Contrast enhancement can be global, or local. Global contrast enhancement typically involves a nonlinear transfer function applied to the luminance channel. Local contrast enhancement mostly involves filtering, either in the spatial domain, or in the frequency domain.

Applying the nonlinear transfer function is generally referred to as gamma correction, the main purpose of which is to correct for the display nonlinearity at the capture end. However, changing the nonlinear shape also helps achieve different perceptual effects, and is typically provided as a user control. There are various forms of black level correction applied to the transfer function, for example, by setting the darkest part of an image to zero luminance (auto pedestal), or by applying an additional linear transfer function with an offset and a modified gain (black restore), or by using an offset coupled with different gains in the darker and lighter luminance ranges (black stretch) [**de Haan 2003**].

The other common method for contrast improvement is histogram modification, where the luminance values in an image are remapped to achieve a desirable distribution (the histogram). Sometimes the histogram modification incorporates amplitude stretch, where the desirable luminance levels in the scene are expanded to optimally use the available signal range.

Above methods mainly rely on signal processing and are not perceptual based. Various published image processing based methods for contrast enhancements are discussed in Chapter 4.

2.2.3.3 Color: Ideally, the goal of color reproduction improvement in the video processing chain is to match the primaries of the modern display system with those of the video camera used in the capture process. Different color primaries used in video processing are discussed later in this chapter. The problem with such color space conversion is in the application of gamma correction during the capture process. Because of the high gamma values in modern displays (ranging between 2.6 and 3, while cameras typically use a gamma of 2.2), red and magenta become more saturated, while green and cyan are de-saturated, and white and the skin tone remain unaltered. This makes it virtually impossible to go back to the original color space by a simple 3x3 matrix transformation. What is done instead is to apply a linear conversion to correct three key colors, for example, skin-tone, white and a natural green [**de Haan 2003**].

Skin tone correction involves correcting for the hue shift that sometimes occurs during the transmission process (mainly in NTSC coding). Skin tone detection essentially looks for colors with the hue angles falling within a predefined range, which may lead to correction of object colors resembling skin tone (false positives), correction of various skin tones to an ideal value, which may not be appropriate, and may even give rise to artifacts.

White correction is nothing but a modification of the white point of the color space, in color science terminology, chromatic adaptation transform. Many modern display devices have a white point correlated color temperature of 10000 K or higher, and thus the white point of the source video color space needs to be appropriately converted.

Human visual system is quite sensitive to a saturation change of green colors. We tend to prefer vivid, saturated green, as in grass and leaves (one of the memory colors). During the color space conversion from the camera to the display, the green ends up being significantly de-saturated, and sometimes undergoes a hue shift. The green enhancement involves increasing the saturation of natural green, coupled with a hue correction.

2.3 Display-Dependent Video Processing

Since the late nineties, display processors have seen a tremendous amount of developmental efforts. State-of-the-art processors include specialized algorithms for removing visual artifacts, increasing the resolution by up-conversion, for color correction, enhancement etc. CRTs of yesteryears are being replaced by flat panel displays based on various modern technologies, with a common goal of achieving higher luminance and larger color gamut. While one of the means to accomplish larger color gamut has been to add more primaries (Multi-Primary Display or MPD), the other strategy involves development of new light sources, including Light Emitting Diode (LED) and laser, with highly saturated colors. [**Kim 2005**].

2.3.1 Working Principles of Modern Digital Display Devices

A brief description of the working principles of various modern displays is provided below.

2.3.1.1 Liquid Crystal Display (LCD)

LCD devices take advantage of the fact that the structure of the liquid crystals can be changed by the application of an electric field so that they transmit different amounts of light. Many LCD devices are based on twisted nematic type of liquid crystals. Such a display consists of a layer of

liquid crystal molecules sandwiched between two transparent electrodes and two polarizing filters oriented perpendicular to each other. At the end, there is a reflecting surface to send light back to the viewer, and is replaced by a light source in case of a backlit LCD. Under normal condition, the liquid crystals form a helical structure between the two electrodes. When a voltage is applied across these electrodes, liquid crystal molecules are aligned along the field, with the amount of twisting reducing proportionally with applied voltage. With sufficient voltage, the crystals are completely untwisted and light passing through the first polarizer is blocked by the second polarizer with crossed orientation. A zero voltage allows most amount of light to be transmitted through the helical structure and results in white. Thus, different amounts of voltage lead to different gray levels. Figure 2.13 shows this mechanism graphically.

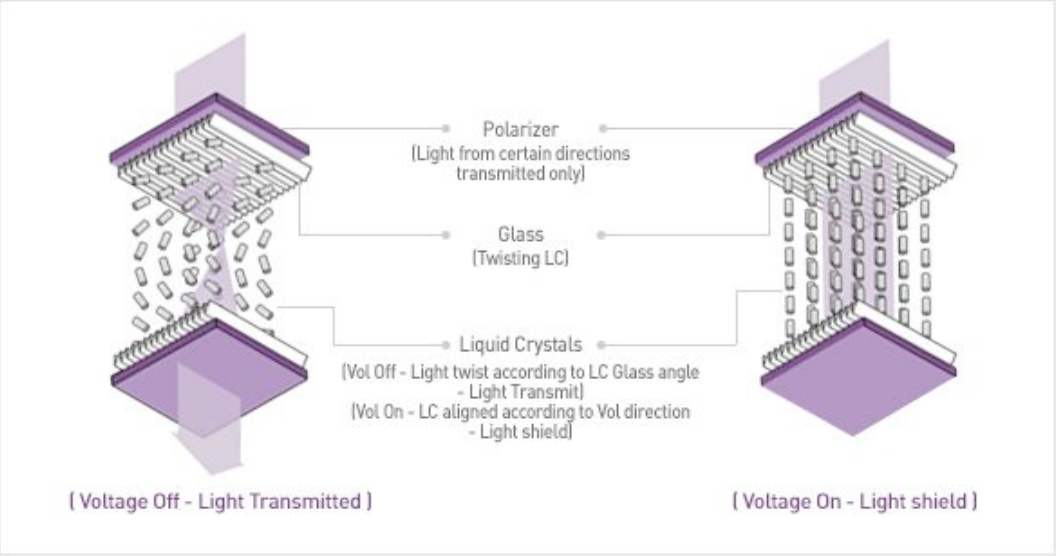


Fig. 2.13 Mechanism of operation in a Liquid Crystal Display [Source: http://www.samsungsdi.com/contents/en/tech/disClass_02_01.html]

A color LCD has each pixel divided into three subpixels with RGB color filters and arranged in various geometries. Each subpixel is individually addressable through appropriate electronics. When each row/column of the display has its own electrical circuit and the pixels are addressed one at a time, the display is called a passive-matrix display. Displays equipped with Thin Film Transistors (TFT) allow row addressing with faster response and brighter pixels, and are called active-matrix displays.

2.3.1.2 Plasma Display Panel (PDP)

PDP displays consist of small cells filled with neon or xenon gas and contained within a chamber with a front and a rear glass plate (Figure 2.14). Strong voltage is applied across the electrodes positioned between the glass plates, causing the gas to ionize and discharge ultraviolet rays. The ionized state during gas discharge is called plasma. In the color panels, the back of each cell is coated with red, green or blue phosphors. When ultraviolet rays hit the phosphors, photons with corresponding wavelengths are emitted, which we see as colored light.

PDP displays can have large size and thin width, higher brightness (1000 lux or more) with contrast ratio as high as 30,000:1, and greater viewing angle than many other competing technologies. The displays have wide color gamuts.

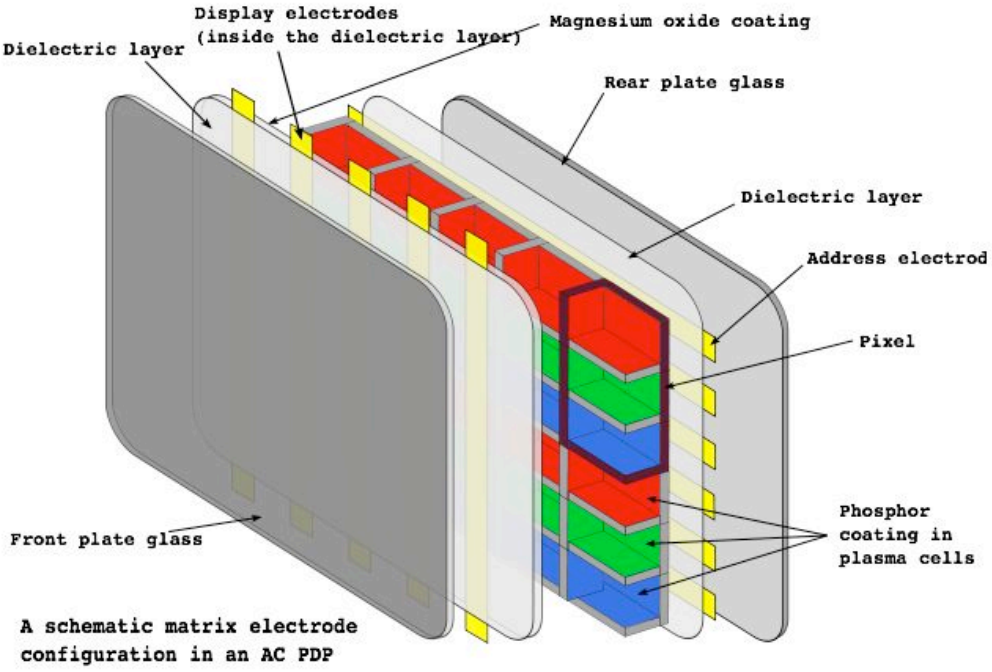


Fig. 2.14 Structure of a Plasma Display Device [Source:

http://en.wikipedia.org/wiki/Image:Plasma-display-composition.svg]

However, the three phosphors used in PDPs have different response time, leading to what is known as the phosphor lag effect, where a colored trail and edge accompany a bright moving object against a dark background. Further, when a video sequence with a moving object is displayed on a PDP, the viewer's eye tracks the motion and in the process integrates subfields along the trajectory that might belong to different pixel positions. This introduces motion blur.

Display processing has to compensate for the phosphor lag through motion vector estimation. However, this further enhances the motion blur typical of PDPs resulting from their subfield driving method. Thus, subfield motion compensation has to follow phosphor lag compensation [**Kim 2005**]. One of the proposed methods for phosphor lag compensation discolors the trails

and edges of moving objects and converts their colors into gray. This helps reduce the color artifact since human eye is less sensitive to variations in gray than variations in chromatic colors [**Oh 2006**].

2.3.1.3 Digital Light Projector (DLP™)

The core of the DLP™ projection system is an optical semiconductor called Digital Micromirror Device, or DMD, invented by Larry Hornbeck of Texas Instruments in 1987 [**DLP 2008**]. The DMD is a highly sophisticated light switch capable of modulating light with high precision. Over two million hinge-mounted microscopic mirrors arranged in a rectangular array are housed in the DMD chip. A digital signal activates a tiny electrode beneath each mirror, causing that mirror to tilt toward (ON) or away (OFF) from the light source. When the DMD is coordinated with digital video, a light source and a projection lens, a grayscale digital image can be reflected onto a screen or other surfaces, as shown in Figure 2.15, by creating a light or dark pixel through light modulation using pulse width modulation technique (PWM). The mirrors can be switched on and off several thousand times per second, allowing the display of bit-streamed video. The use of PWM techniques results in a linear device transfer function from input signal to light output [**Pettitt 2001**].

To display color, the light from the projection system is passed through a color wheel fitted with red, green and blue filters, shown in Figure 2.16. Colors are shown in sequence on to the surface of the DLP™ chip. The switching of a given mirror and the duration of its on or off time are determined based on what color the mirror is supposed to project. As an example, to generate a yellow pixel, a mirror alternately flashes red and green colors with high frequency such that a

temporal integration takes place in observer's eye, creating a sensation of yellow hue. A single-chip DLP™ projection system, used in televisions, home theater systems and business projectors, can produce 16.7 million colors.

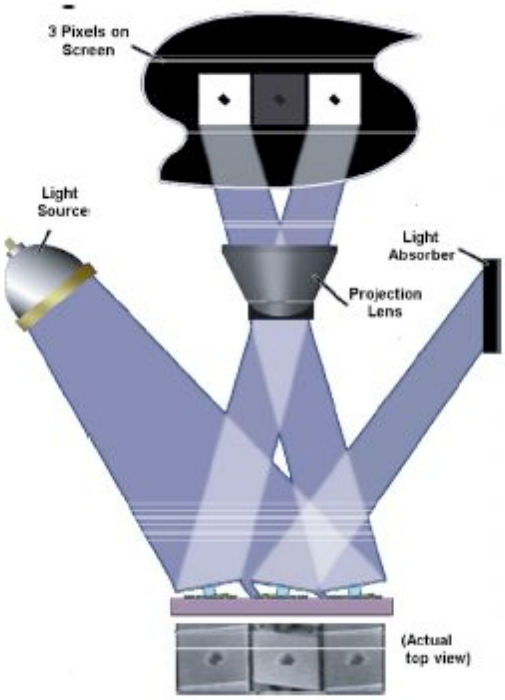


Fig. 2.15 Optical switching through DMD [Source: **Pettitt 2001**, Fig. 2]

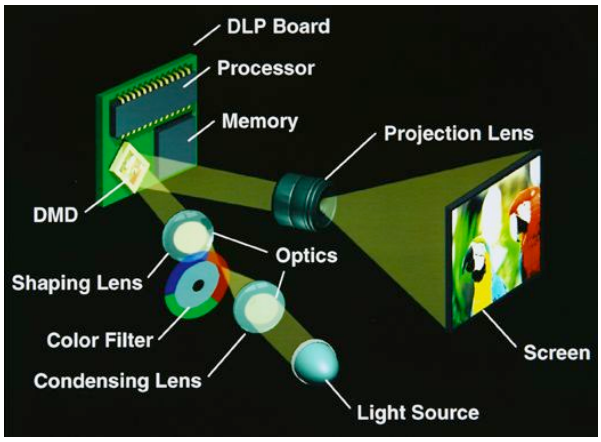


Fig. 2.16 Schematic of a DLP system [Source: **DLP 2008**]

2.3.1.4 Organic Light Emitting Diode (OLED) Displays

OLED is a form of light-emitting diode composed of phosphoric or fluorescent organic thin film (Figure 2.17). OLED structure consists of an emissive layer, a conductive layer, a substrate and the electrodes. The layers are made of organic materials with variable conduction. The electroluminescent emission is caused by electric voltage applied across the electrodes, which results in light of different colors.

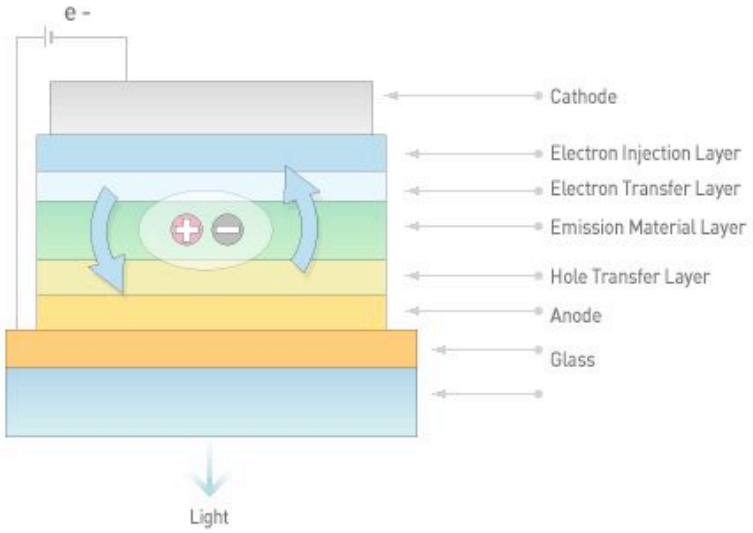


Fig. 2.17 Structure of an OLED device [Source:

http://www.samsungsdi.com/contents/en/tech/disClass_03_01.html]

OLED displays consume significantly low power and can be brighter and considerably thinner than LCDs. They do not require backlighting either. However, the organic materials deteriorate over time, which reduces the lifetime for these displays.

2.3.1.5 Laser Displays

Laser TV is an HDTV with solid state lasers providing three primary colors, red, green and blue. In a recent development, a display based on laser light sources was reported to achieve a very high luminance (500 cd/m^2) and significantly wider color gamut, namely 190% of the gamut defined by Rec. 709 [**Sugiura 2007**]. Figure 2.18 compares the color gamut of this display with that of Rec. 709 and an LED backlit LCD (discussed later in this chapter). Note that the primaries in the laser TV are almost monochromatic. The display developed by Sugiura et al eliminated color wheel typically used in the projection display to achieve field sequential color, and instead directly controlled the illumination timing of the three-primary color laser beams.

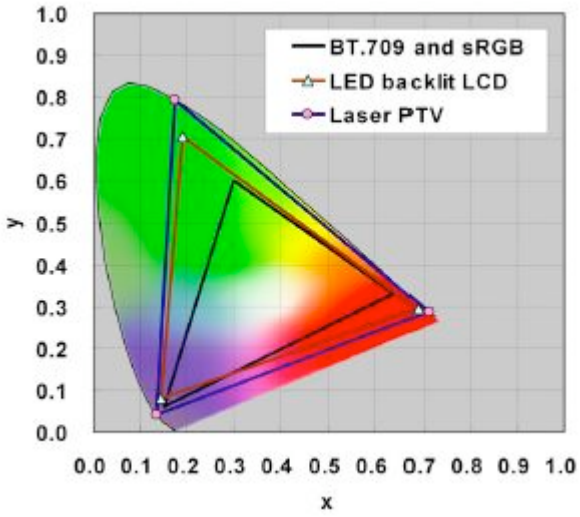


Fig. 2.18 Color gamut of laser projection TV in comparison with that of Rec. 709 and LED backlit LCD [Source: **Sugiura 2007**, Fig.3]

An additional feature of this display, and possibly of other laser displays, is its compliance with xvYCC, an extended color space for moving pictures proposed by International Electrotechnical Commission (IEC 61966-2-4) [**Sugiura 2007**]. Figure 2.19 shows the extended region available

in xvYCC color space. This color space allows red, green and blue pixel values to go below 0 (darker colors) and above 1 (very bright colors).

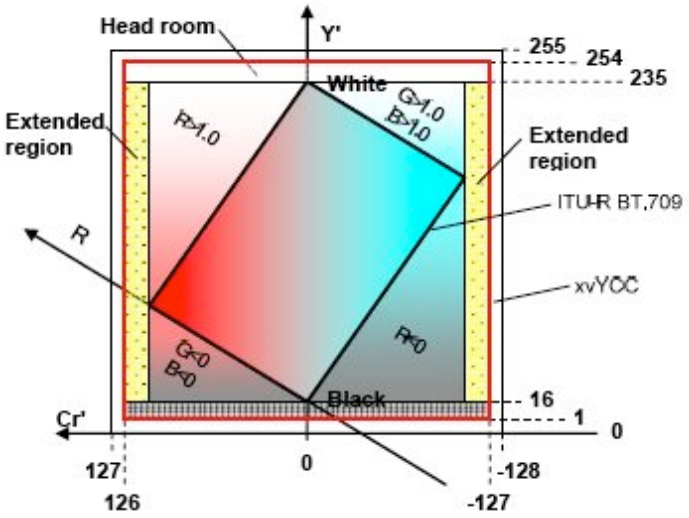


Fig. 2.19 Extended region in xvYCC color space [Source: **Sugiura 2007**, Fig.2]

2.3.1.6 Field Emission Displays (FED)

A field emission display or FED is one of the most recent developments among the display technologies. An FED is a type of flat panel display that employs field-emitting cathodes coupled with phosphor coatings to generate light. An FED is similar to a CRT in operation, but is much thinner as it does not require an electron gun. Instead, it contains a large emitter array, which, loosely speaking, is placed closely behind a phosphor layer that emits light when subjected to field emission. Figure 2.20 shows the structure of a Spindt-type color FED, recently developed by Itoh et al [**Itoh 2007**]. The focus electrode layer is designed to reduce color crosstalk. The total thickness of the panel is only 2.8 mm. FEDs are low power consuming devices particularly suited for high-quality motion pictures because of their quick response.

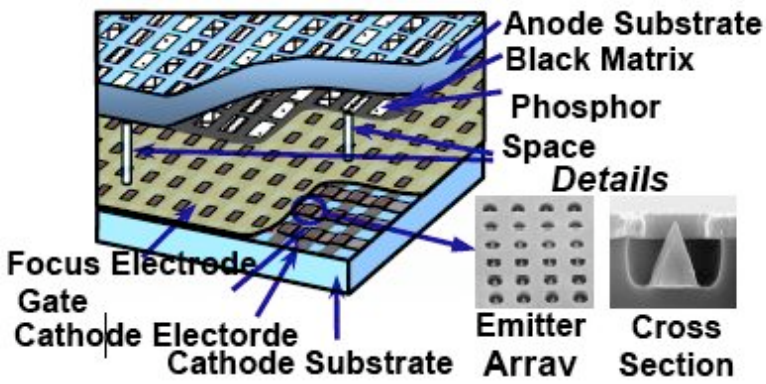


Fig. 2.20 Structure of a Spindt-type color FED [Source: **Itoh 2007**, Fig.3]

2.3.2 Color Processing in Wide Gamut and Multi-Primary Displays

Displays having larger gamuts than the gamut of a conventional display with standard primaries are termed as wide gamut (WG) displays. In one of the earliest publication on HDTV colorimetry, DeMarsh [**DeMarsh 1991**] referred to wide gamut primaries as those primaries that are capable of displaying a larger gamut than the SMPTE "C" primaries (defined in SMPTE RP-145 standard). However, there is no specific set of primaries termed as wide gamut primaries [**Kwak 2005**]. It is reasonable to make such determination based on contemporary ITU-R BT Rec. 709 primaries defined for HDTV. However, it is a common practice to represent gamuts of modern displays in terms of percent of NTSC color gamut, which is still an industry benchmark.

Multi-primary displays (MPD) are those that have more than three primary colors. These are essentially wide gamut displays as they can produce more natural colors than the conventional displays can, particularly in the cyan and yellow regions in the color space. These displays generate multi-primary signals from the incoming three standard primary signals. Note that a

wide gamut display does not necessarily have to be multi-primary, examples being 3-primary LCD with LED backlight and RGB laser displays.

Among the problems associated with color processing in an MPD are those of an over-determined system with multiple gamut mapping solutions to convert 3-primary inputs to multi-primary outputs, and of metamerism, a many-to-one mapping of colors that may result in discontinuities. Solutions to these problems have been suggested with 5-primary DLP™ projection television developed by Samsung [**Ok 2005**]. The display has yellow and cyan primaries, in addition to red, green and blue. Figure 2.21 compares the color gamut of this display to that defined by Rec. 709 primaries. Decomposition of five primaries into two color components was proposed, one consisting of three primaries and one consisting of two primaries. Two arbitrary color components from the five primaries were selected and their values at the gamut boundary were used to construct an LUT. The remaining three components were calculated using inverse matrix. The paper also suggested a method for gamut mapping so that natural images conforming to Rec. 709 can be displayed on 5-primary display without causing over-saturation or under-saturation of chroma in high lightness or low lightness regions. This involved lightness mapping of the gamut vertices (or cusps), followed by a chroma mapping.

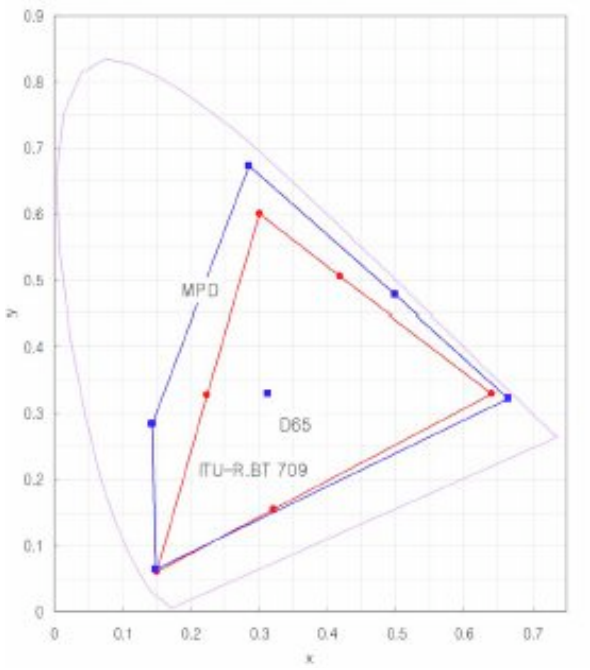


Fig. 2.21 Color gamuts of 5-primary DLP™ projection TV and that defined by Rec. 709 primaries [Source: **Ok 2005**, Fig.1]

Apart from the problems mentioned above, the other difficulty with MPDs is associated with the area-based pixel structure. As the pixel resolution in a flat panel display increases, the display luminance decreases. So, MPDs for area-based displays, e.g. LCD or PDP, have lower luminance than sequential time-based displays, e.g. projection type display, at the same resolution. The solution may require redesigning of the color filters in such a way that luminance as well as chrominance ratios of sub-pixels do not affect the color balance. Choe et al [**Choe 2005**] proposed a new pixel structure for six-primary flat-panel wide gamut displays, shown in Figure 2.22, and also developed a gamut mapping and color decomposition algorithm to convert input RGB signals to 6-channel signals (RGBCMY), with the goal to preserve color image quality of RGB displays and maximize color gamut usage. The color decomposition is essentially a gamut-mapping algorithm, functioning in the RGB space. The algorithm involved

linearly combining the input RGB to form first approximation of RGBCMY data, and then compensate for insufficient chrominance and luminance as compared to the input color.

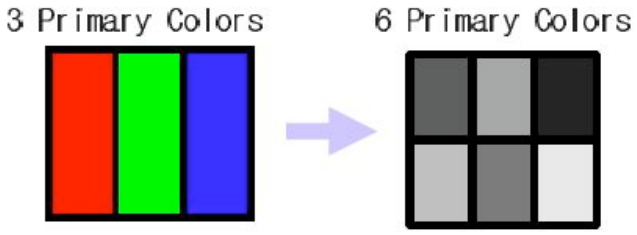


Fig. 2.22 Pixel structure for 6-primary LCD [Source: **Choe 2005**, Fig.3]

In a practical implementation of MPD technology, a prototype five-primary MPD was developed by Samsung using a color wheel with RGB DLP projection TV with 50/60" screen size [**Kim 2004**]. A rotating color wheel with five interference filter segments was placed in front of the projection lamp to generate five primary colors, namely red, green, blue, cyan and orange. Figure 2.23 compares the color gamut of this MPD to that defined by Rec. 709 primaries as well as shows the three-dimensional gamuts in CIELAB space. Although the MPD has a larger gamut volume, in some parts Rec. 709 has more saturated colors. A display model and a gamut mapping algorithm were proposed and implemented on hardware. The research found an increase in the perceived saturation compared to conventional systems because of an increase in chroma and lightness.

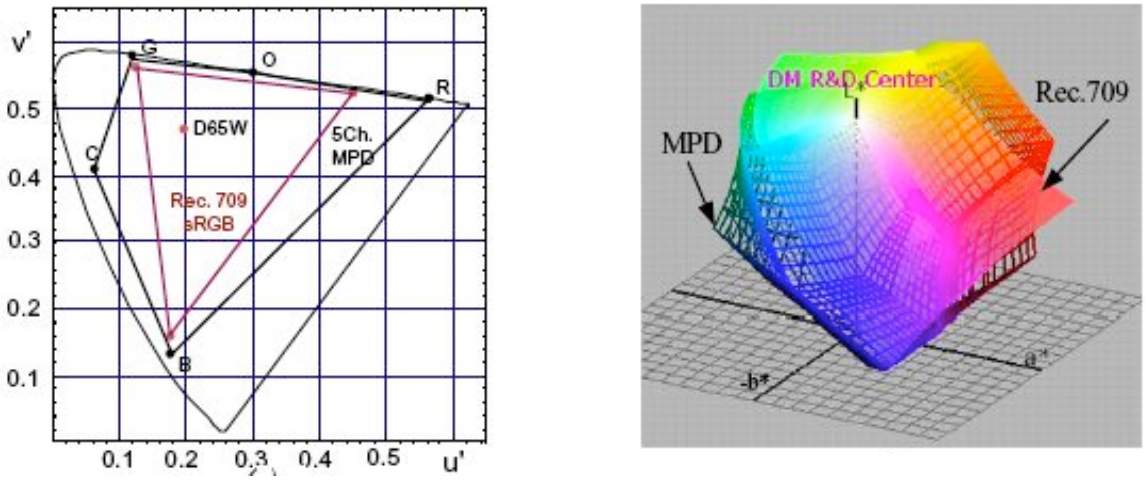


Fig. 2.23 Comparison of color gamuts of the five-primary MPD and Rec. 709 in u'-v' diagram and in CIELAB space [Source: **Kim 2004**, Fig. 2]

Multi-primary technology has also been applied to develop wide gamut high efficiency LCD panels with four (RGBY) and five (RGBCY) primaries coupled with Cold Cathode Fluorescent Lamp (CCFL) backlight [**Roth 2007**]. The backlight and color filter properties play an important role in determining the color and intensities of the individual primaries of such displays. The four-primary display was equipped with normal color gamut CCFL backlight, while the five-primary had wide gamut CCFL backlight. Identical sub-pixel sizes were used for manufacturing convenience. An optimization of the color filters and fine-tuning of the phosphor mixture in the backlight helped achieve an appropriate white point. Like many other MPDs, these displays have yellow and cyan filters aimed at increasing the color gamut area coverage in terms of % NTSC, increasing the white point luminance and getting a good color coverage and color intensity distribution. A multi-primary scaling algorithm was proposed that allowed the use of same TFT backpanel and driver electronics as in regular RGB panels. The color filters were applied to the original sub-pixels, increasing the effective pixel size. However, the perceived resolution was not

greatly affected since the bright sub-pixels increased perceived spatial resolution. A spatial scaling algorithm took into account the location, color and intensity of each sub-pixel to generate a smooth image. Figure 2.24 shows color gamuts of different displays, including regular RGB display used as a reference.

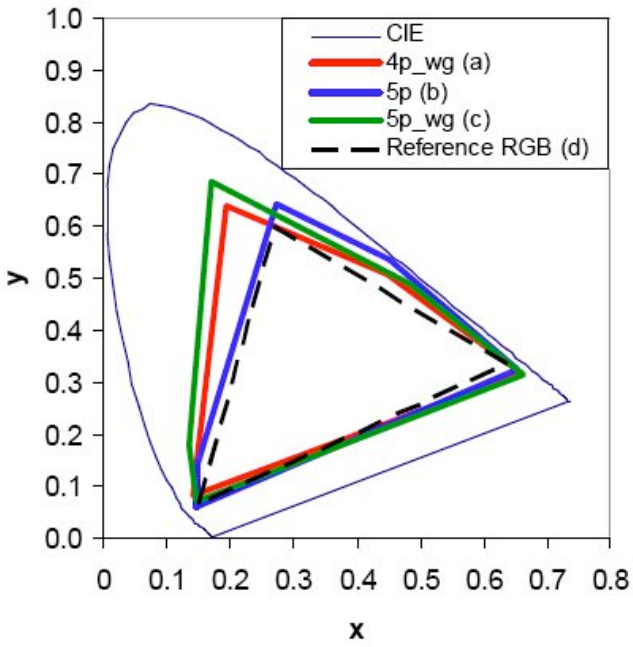


Fig. 2.24 Color gamuts of various displays: a) four-primary wide gamut CCFL, b) five-primary normal gamut CCFL, c) five- primaries display with wide gamut CCFL, d) reference RGB display [Source: **Roth 2007**, Fig. 6]

Besides increasing display gamuts, it is often desirable to increase the light intensity of the display devices. One of the methods to achieve that is to add a white channel, preserving a high rate of light transmission to red, green and blue filters. This, although increases the light output, results in reduced saturation, and thus inferior image quality. Lee et al [**Lee 2002**] proposed a method to increase the light intensity in a color sequential projection while preserving hue and

saturation of the original RGB input color. Figure 2.25 shows the concept. If T_r, T_g, T_b and T_w are filter transmittances and E_r, E_g, E_b and E_w are exposure times per frame for the four filters, the output light intensity Y_{out} corresponding to an input light intensity Yin is given by:

$$Y_{out} = Y_{in} * (T_r E_r + T_g E_g + T_b E_b + T_w E_w) \tag{2.1}$$

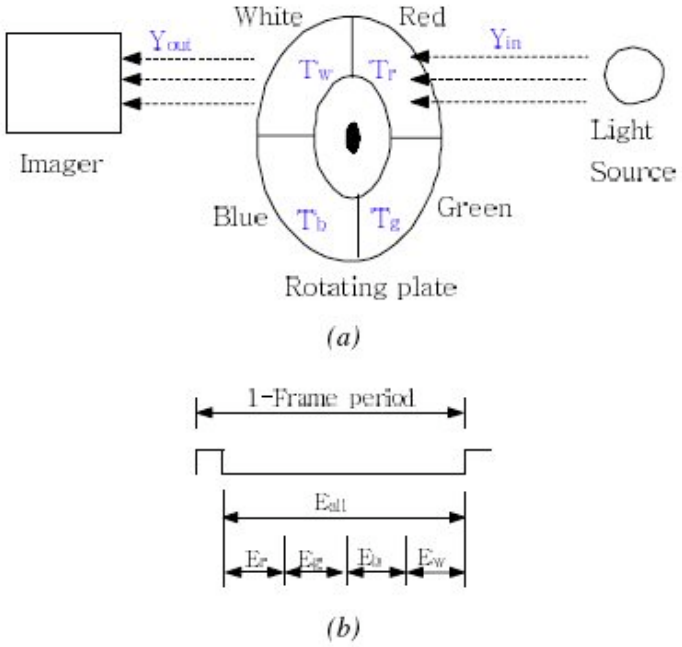


Fig. 2.25 Single panel display with four color filters a) schema, b) timing diagram [Source: **Lee 2002**, Fig. 1]

Figure 2.26 demonstrates the problem associated with adding a fourth channel in 3-color vector space. Adding a color C_2 to the input color C_1 leads to color C_3 that has a different hue (given by angle) and saturation. Note that C_2 has the same amount of red, green and blue, resulting in a white of given lightness. A compensation vector C_4 is needed to change the direction of output color vector to the direction of original input color vector. However, since RGB is not a

perceptually linear space, the compensation vector is not sufficient to preserve hue and saturation of the input color. In the proposed method for RGB to RGBW conversion, RGBW color space was divided into constant scaling space and gamut scaling space, so that different scaling could be applied as appropriate. The transformation involved three steps, namely, determination of scale space, scaling the input signal and finally, separation of RGB and white signal.

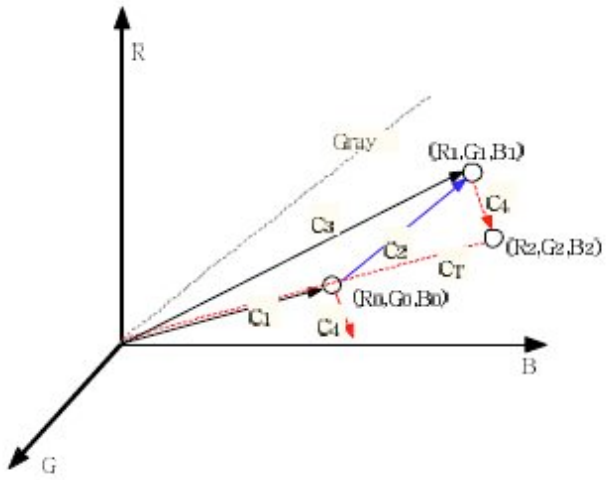


Fig. 2.26 Vector representation of RGBW processing [Source: **Lee 2002**, Fig. 3]

It seems more appropriate to use an opponent color space like CIELAB for this computation, but a solution in the RGB space was probably preferred due to the implementation issues.

Displays using RGBW primaries can achieve higher luminance, and often an improved efficiency through lower power consumption. Displays such as LCD and DMD typically employ a spatial light modulator to add an amount of luminance from the white primary that is correlated with the amount of input RGB. They typically resort to a tradeoff of color reproduction for higher efficiency, resulting in less saturated and/or less bright colors in certain cases. Emissive

displays like OLED uses an array of light-emitting subpixels as described before, thus the efficiency of the display is determined by the efficiency of the subpixels. A filtered white RGBW (W-RGBW) OLED is a multi-primary display that uses an independently controlled white emitter for each subpixel. RGB subpixels are provided with color filters while white subpixels have none and thus, are more efficient than the color subpixels. Murdoch et al [**Murdoch 2006**] presented a color-processing algorithm that exploited this feature in W-RGBW OLED displays to accomplish color accuracy and power savings at the same time. The method involved transferring the neutral luminance, given by the minimum of RGB digital counts or a fraction of it, from less efficient RGB subpixels to more efficient white subpixel to generate a metameric color with high colorimetric accuracy. If the white primary has the same chromaticities as the display primaries, RGB subpixel values can be used directly to form RGBW signal. However, the white replacement algorithm may introduce color reproduction error if broadband source used in the display does not correspond to the intended display white point. In such cases, RGB chromaticities must first be transformed to white normalized RGB, or $R_nG_nB_n$, such that the equal $R_nG_nB_n$ triad produces the same chromaticity coordinates as the white primary. Then in $R_nG_nB_n$ space, a certain fraction (white mixing ratio) of the minimum of RGB is subtracted from $R_nG_nB_n$, and the subtracted result is then normalized to return to the white-point normalized RGB space. Note that filtered white RGBW OLED displays use white to efficiently replace neutral luminance, while RGBW displays with additional white channel use white to augment luminance. The efficiency of W-RGBW OLED displays depends largely on how often the white subpixel is used to replace RGB subpixels, and thus on the picture content.

As mentioned before, using multi-primaries is not the only way to achieve wider gamuts. Recently, a prototype of a high resolution (WUXGA: 1920x1200) LCD with LED backlights has been developed, offering a wide color gamut and high luminance [**Sugiura 2006**]. In LED backlit displays, RGB color LEDs are used as point sources for backlighting, with white color being produced by mixing the three colors using an optical waveguide. However, several problems are encountered in this technology. The white point chromaticities shift over time due to the changes in the wavelength of the emitted beam as well as the light output of the LEDs. Color variation across the display can be observed as the uniformity in the luminance output of the LEDs deteriorates over time. Sugiura et al's prototype incorporated a feedback control circuit with optical sensor for the stabilization of luminance and chromaticity of the backlight, a color non-uniformity correction circuit as well as a gamma correction function.

The authors also reported the development of a six-primary LCD with LED backlights having two sets of RGB power LEDs that achieved a color gamut 175% wider than that of sRGB gamut used in conventional displays (Figure 2.27). However, high cost and low efficiency are considered technological barriers in bringing LED backlight technology to mainstream consumer display market [**Roth 2007**].

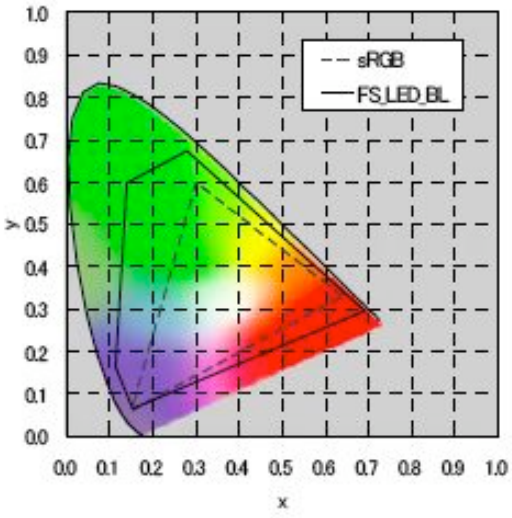


Fig. 2.27 Color gamut of six-primary LCD with LED backlight [Source: **Sugiura 2006**, Fig. 13]

2.4 Challenges and Opportunities in Color Video Processing

In the world of consumer video, research and development activities are expected to continue both in the realms of display-dependent processing and display-independent processing, with overlapping activities being a likely scenario. Many of the ongoing development efforts in the display industry focus on flat-panel displays. While we are not likely to see much increase in the display resolution from where we already are, there are several technical challenges posed by modern displays that remain to be addressed. For example, the solution to the problem of loss of perceived resolution with increasing object speed on many contemporary flat-panel displays needs motion compensated picture rate conversion as well as high-speed cameras and transmission [**de Haan 2007**]. Many of the emerging display technologies suffer from various temporal artifacts, which must be remedied at the hardware and/or software level [**de Haan 2001**]. While advanced display technologies like wide gamut, multi-primary displays pave the path for widespread adoption of HDTV standard and systems, one of the challenges will be to

optimally show legacy video with standard-definition color content on modern displays, requiring advanced resolution up-conversion [**de Haan 2007**]. In addition to the display technologies discussed in this chapter, there are other ambitious emerging technologies like high dynamic range display, 3D display and fine grey-scale quantization display (10bitTV) that can open new frontiers for color video processing.

As we look into the state-of-the-art of color video processing, it is worthwhile for us to step back and take an objective view from the perspective of color science. To this end, Fairchild presented a critical review of some of the color processing in a typical video processing chain [**Fairchild 2007**]. One of the key aspects of color processing in consumer video applications is that it is essentially display-centric. Capture-end of a typical video system workflow does not involve a colorimetrically accurate method, resulting in color information about the original scene being discarded at the outset of video processing. Color processing for video attempts to achieve high picture quality for a standard display, which may or may not be desirable in a given application context.

As outlined in this chapter, several steps in the video processing chain that are not directly related to color can affect the picture as well as color quality. Nevertheless, these steps do not appropriately address the color issues. For example, a nonlinear gamma correction followed by a subsampling of color difference signals in the encoding stage results in luminance and chrominance errors during any differential processing of luminance and chrominance data, which are ignored in video processing. Further, video processing typically avoids using a perceptual space in which luminance and chrominance channel data are orthogonal for practical purposes. A

color space like YC_BC_R introduces crosstalk between these channels leading to color error at the display end. Video processing algorithms do not take into account the viewing conditions as is done in color and image appearance modeling. These problems are handled on an ad hoc basis by the end-user to achieve a pleasing picture quality for a given display. The solution is far from desirable from accurate color reproduction perspective.

Several practical improvements in television video processing can help achieve superior color reproduction [**Fairchild 2007**]. Using characterized video cameras with recorded camera settings and video information will provide an option to retrieve colorimetric information of the scene. Video processing in a perceptually linear space can help reduce artifacts and produce better color quality. At the output end of the processing chain, displays with proper colorimetric characterization will ensure consistent color reproduction across different display technologies. Integrating color appearance issues into the video processing algorithms will make a display responsive to changes in the ambient viewing conditions.

As the capabilities of today's display technologies keep advancing at a remarkable pace, consumer demand for superior picture quality is stronger than ever before. Color plays a vital role in determining the quality of a display, and any future development efforts in color video processing must recognize the potential for improvement in color reproduction capabilities of various emerging display technologies. Video researchers and color scientists must work together to transform this potential into reality.

"To suppose that the eye with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selections, seems, I confess, absurd in the highest degree." ~ Charles Darwin (English Naturalist and Author of the theory of evolution by natural selection. 1809-1882)

Chapter 3

VIDEO QUALITY AND ITS ASSESSMENT

Work on picture quality metric goes back almost 50 years. Most of the quality metrics proposed over time are quality metrics for still images [**Wu 2006**]. First models of human vision were based on single-channel approach, in which the human visual system is represented by a single spatial filter characterized by the contrast sensitivity function. Mannos and Sakrison [**Mannos 1974**] developed the first image quality metric for encoded monochrome images. They considered integral square type of distortion measures, calculated the rate-distortion function and simulated the optimum encoding of a given image at a given bit-rate by varying different coding parameters. They also took into account some of the well-known characteristics of spatial vision and contrast sensitivity and developed a mathematical model for the human visual system, which was a closed-form expression for the contrast-sensitivity as a function of spatial frequency. The input images were filtered with this function after applying amplitude nonlinearity. Squared difference between the two images was the distortion measure. This was one of the first works in engineering that applied vision science in image processing.

The first color image quality metric was proposed by Faugeras [**Faugeras 1979**]. He presented a simple model of human color vision that quantitatively described different perceptual parameters such as brightness, hue, saturation etc. The idea was to structure the perceptual space as a vector

space with spatial filtering properties, and introduced a norm on the vector space that would allow us to measure distances and to define a distortion measure in agreement with perceptual evaluation.

Soon after the introduction of a color image quality metric, Lukas and Budrikis [**Lukas 1982**] developed the first video quality metric based on a vision model. The first stage of the model constituted a nonlinear spatiotemporal model of a visual filter describing threshold characteristics on uniform background fields. The second stage incorporated a masking function to account for the non-uniform background fields. The model attempted to predict the subjective quality of moving monochrome television pictures containing arbitrary impairments.

In this chapter, various methods for determining and assessing the video quality are reviewed. These methods are broadly categorized into two groups, one that is based on engineering approach, relying on identifying and quantizing specific features or artifacts, and one that is based on psychophysical approach, involving the modeling of human visual system. Various video quality metrics based on color difference, image difference and image appearance modeling are also categorized under psychophysical approach, since they invariably take into account various mechanisms associated with the human visual system. Over the years, various international organizations have attempted to standardize video quality assessment methods and metrics. Some were more successful than the others. These standards are briefly reviewed in this chapter. The concluding section reviews several publications on subjective assessment of video quality.

3.1 Engineering Approach

In this approach, metric design is primarily based on extracting and analyzing some specific features or artifacts in the video. These are also called objective image quality metrics. The quality metrics developed using this approach can be further classified into three groups, full-reference, reduced-reference and no-reference [**Wu 2006**]. Full-reference metrics require the entire reference video to be available. Reduced-reference metrics require information on several features from the reference video. No-reference metrics do not use any information about the reference video.

As described in Chapter 2, most artifacts encountered in video are a direct result of compression. Miyahara et al [**Miyahara 1998**] presented a new methodology to obtain a picture quality scale (PQS) for coded achromatic still images. The authors used Weber-Fechner's law and the contrast sensitivity for achromatic images over the full range of image quality defined by the subjective mean opinion score. Some of the visual perception properties related to global image impairment were used while weighting the perceptually important structured and localized errors. The resulting PQS had high correlation with subjective mean opinion scores.

The work of Wu and Yuen [**Wu 1997**] focused on a specific compression artifact, the blocking artifact, which is quite common in compressed video. The authors presented a no reference, generalized block-edge impairment metric that enabled evaluation of reconstructed picture with blocking artifacts. The metric took into account luminance masking effects in extremely bright or extremely dark areas in a reconstructed image. The performance of this metric was found to be consistent with subjective evaluation.

Tan et al [**Tan 1998**] proposed a two-stage full-reference objective measurement model for MPEG-coded video. The first stage applies low-pass spatial filtering and spatial masking to each frame in the decoded picture and then computes perceptual impairment, taking into account the reference picture. The second stage simulates human visual system's processing of visual information, and thus acts as a cognitive emulator. The model was tested with several coded video sequences of 2-3 min duration. The model, when compared with PSNR and SSCQE, provided a closer approximation to the latter. This work put an emphasis on the need of a cognitive emulation stage in the objective measurement model.

Caviedes et al [**Caviedes 2000**] used three impairment metrics for MPEG video quality assessment, namely blocking artifact level, ringing artifact level and corner outlier artifact level, and created a combined impairment metric. The metric was used for subjective and objective quality assessment. The authors concluded that before we could map such a deterministic metric onto a probabilistic perceptual space and thus use it in closed-loop quality control system, subjective test methods needed to be improved in order to increase resolution and certainty of quality prediction.

Caviedes and Oberti [**Caviedes 2003**] developed a no-reference objective quality metric for measuring improved and degraded video. The authors followed a heuristic, incremental approach to modeling quality and training the model using a variety of video sequences. The method involved dividing the training sequences into impaired, enhanced and impaired-enhanced sets in order to deal with individual impairments and enhancements.

3.2 Psychophysical Approach

In this approach, metric design is primarily based on modeling of human visual system. These metrics are essentially full-reference metrics. Ahumada [**Ahumada 1993**] presented a comprehensive overview of different vision model based quality metrics for monochrome still images. The review was however limited to image quality metrics relying on image difference between the original and a corrupted version of it. Further, several more vision model based metrics have been proposed since the time of this publication.

Lindh and van den Branden Lambrecht [**Lindh 1996**] presented a vision model for moving pictures. The model was a more advanced version of a simple spatiotemporal model proposed earlier [**van den Branden 1996**]. It accounted for the normalization of cortical receptive field responses and inter-channel masking. Two quality metrics for video were derived from the new and old model and were used to assess the coding quality of MEPEG-2 video streams. The new model yielded a better quality rating.

One of the earliest video quality metric was offered by Sarnoff Just-Noticeable Difference (JND) model. The model aims to provide accurate estimate of the visibility of differences between original and distorted image sequences without requiring direct measurement using human observers [**Lubin 1997**]. The model was based on physiological and psychophysical principles of human visual discrimination performance, and thus was applicable to a varied range of distortions. The paper discussed model performance in a range of video applications involving discrimination and fidelity assessment.

One of the most well known models for video quality was DVQ, an acronym for Digital Video Quality, described by Watson et al in 1999 [**Watson 1999**]. It was similar to Sarnoff JND model, but significantly different in implementation. DVQ metric attempted to incorporate several aspects of early visual processing into a simple image processing algorithm, including light adaptation, luminance and chrominance channels, spatial and temporal filtering, spatial frequency channels, contrast masking and probability summation. It accelerated the spatial filtering operation by using Discrete Cosine Transform (DCT), making use of already available efficient hardware and software.

In the subsequent publication following the proposal of DVQ metric, Watson and others [**Watson 2001**] reported new visual data on the visibility of dynamic DCT quantization noise. This was obtained by using an image composed of a square array of 8x8 pixel blocks. A DCT basis function of the same frequency was placed within each block, and over a sequence of frames, each basis function was temporally modulated by a Gabor function of a particular temporal frequency and phase. Human visual threshold for the DCT noise was measured through psychophysical experiments. The data were then fitted with a mathematical model and incorporated it into the DVQ metric. The DVQ metric was tested comparing its predictions to judgments of impairment for a video stream of 65 sequences to the quality estimates provided by 25 human subjects. DVQ model performed similarly to the Sarnoff model. However, a systematic failure of prediction was also observed. Like the Sarnoff model, the DVQ metric is aimed at predicting the probability of detection of threshold image differences. It does not include appearance modeling through spatial or temporal adaptation, or correlates of appearance attributes and thus, cannot be used for video rendering as well [**Fairchild 2005**].

Winkler presented a perceptual distortion metric for color video sequences [**Winkler-1 1999**]. It was based on a contrast gain control model of the human visual system incorporating spatial and temporal aspects of vision as well as color perception. After conversion of the input video sequence to the opponent color space, each of the resulting three components was subjected to a spatio-temporal perceptual decomposition, yielding a number of perceptual channels. They were weighted according to contrast sensitivity data and sent through a contrast gain control stage. Both the reference and the processed sequence were input to the model. Finally, all the sensor stages were combined into a distortion metric. The metric was used to assess the quality of MPEG coded sequence. The model achieved a close fit to contrast sensitivity and contrast masking data from several different psychophysical experiments for both luminance and color stimuli.

In an important contribution to the volume of literature on perceptual video quality, Winkler summarized the issues in vision modeling for perceptual video quality assessment (PVQA) [**Winkler-2 1999**]. The author described quality factors in relation to a human observer, and how to measure them. Then the human visual system was described, along with the process to incorporate its component or phenomenon in a vision model for PVQA. The validation and evaluation of PVQA systems were also discussed by the author.

Traditional visual quality metrics measure image fidelity, i.e. the accuracy of the reproduction of the original image, instead of perceived quality. Winkler [**Winkler-3 2001**] investigated the addition of image appeal attributes to the metric in order to account for the perceived video

quality. Sharpness and colorfulness were identified as important subjective attributes of quality and were integrated into the perceptual distortion metric [**Winkler-1 1999**]. These attributes were quantified by a sharpness rating based on local contrast and a colorfulness rating based on the distribution of chroma and saturation in the video sequence. Several subjective tests were performed to determine the relationship between these ratings and perceived visual quality. The results showed an improvement in the prediction of perceived quality by including sharpness and colorfulness ratings.

Yu et al published a review on human visual system (HVS) based digital video quality metrics [**Yu 2000**]. Three objective video quality metrics that represented the state-of-the-art of HVS based quality metric research were chosen for review, namely Sarnoff JND, Watson's DVQ and Perceptual Distortion Metric (PDM) proposed by Winkler. These models were also tested and verified by VEQG. Watson emphasized on easy implementation while formulating DVQ metric. As a result it was found to minimize the calculation and memory requirements, but at the cost of performance. Both Sarnoff JND and PDM incorporated recent results in vision research and showed similar overall performance. However, several critical aspects of HVS were missing from these models. These included i) temporal mechanism such as motion and light adaptation, spatial frequency adaptation, backward masking etc, ii) a thorough analysis of the opponent color space that models the perceptual pathways, iii) a working contrast gain control model which explains inter and intra-channel masking over temporal, spatial frequency and orientation bands and iv) supra-threshold quality metric.

Brill et al [**Brill 2004**] recognized the variability implicit in the psychophysical subjects and came up with techniques for determining the statistical resolving power of a video quality metric (VQM), defined as the minimum change in the value of the metric for which subjective test scores show a significant change. The primary data used in the analysis were subjective scores of various video-source materials subject to various kinds of digital-processing distortion. Original subjective mean opinion scores were converted to a common interval scale, and then the VQM scores were transformed to this common scale through statistical analysis. Fitting all VQMs to one scale provided a way for cross-calibration of those VQMs, in other words transformation of one VQM to another. Statistical probability was used to assess the resolving power of VQM. These new methods for assessing VQM accuracy and cross-calibrating VQMs were incorporated into the ATIS series of Technical Reports, which provided a comprehensive framework for characterizing and validating full-reference VQM.

3.2.1 Image Quality Metric Based on Image Difference

In an attempt to offer a comprehensive model for the complex process of image contrast judgment, Daly [**Daly 1993**] introduced an algorithm for the prediction of visual differences between two digital images based on a model of human visual system. The goal was to assess the image fidelity and develop an algorithm that could be used for the design and analysis of image processing algorithms, imaging systems etc. It consisted of three parts, amplitude nonlinearity, contrast sensitivity function and a hierarchy of detection mechanism. The algorithm was tested for a wide variety of image distortions including synthetic images designed for the purpose of psychophysical experiments and natural images with practical distortions. The sources of image

distortion included blur, noise, data compression artifacts, banding, blocking, contouring, low frequency non-uniformities, hyper-acuity and tone-scale changes.

3.2.2 Image Quality Metric Based on Color Difference

Research on color difference equations had an important effect on the conceptualization of image quality metrics. Color difference research culminated with the introduction of CIEDE2000, but long before that, S-CIELAB presented the first incarnation of an image difference model based on CIELAB color space and color difference equations. Zhang et al [**Zhang 1996**] introduced this as a spatial extension to the CIELAB color metric for measuring color reproduction errors of digital images. The goal was to apply a spatial-filtering operation to the color-image data in order to simulate the spatial blurring by the human visual system. The model was essentially a spatial pre-processor to the standard CIE color difference equations to account for complex color stimuli such as halftone patterns. In case of large uniform areas, results had to be consistent with CIELAB calculations. To achieve this, the image data were first transformed into the opponent space, representing a luminance channel and two chrominance channels, red-green and yellow-blue. Then each opponent color images was convolved with a one-dimensional kernel to that color dimension. The shape of each kernel was determined by the visual spatial sensitivity and was such that the area under each of those kernels integrated to unity. Finally, the filtered data were transformed into a CIEXYZ representation, followed by a conversion to CIELAB. The calculation effectively segregated patterns and colors because spatial transformation was independent of color transformation, as also suggested by psychophysical experiments on human visual perception with regard to simple colored patterns. Because of this feature, S-CIELAB could be implemented as a pre-processor to existing CIELAB based systems. The results

reported by Zhang et all showed better consistence with visual evaluation of color image difference than that predicted by general CIELAB equation.

Tong et al [**Tong 1999**] proposed a video quality metric based on a single-resolution spatial, temporal and chromatic model of human contrast sensitivity. This metric was an extension of the S-CIELAB and was termed spatio-temporal CIELAB (ST-CIELAB). The metric was designed such that it fitted published contrast sensitivity data and reduced to CIELAB value for uniform color field. ST-CIELAB was tested by conducting psychophysical experiments with MPEG video sequences. The metric ratings were consistent with subjective assessments, but overestimated the visibility of blocking artifacts in color MPEG video.

3.2.3 Image Quality Metric Based on Image Appearance Modeling

Formulation and application of **image appearance models** essentially began with the image measurement. As imaging systems became more complex, there was a need for device-independent images measurements, which started with CIE colorimetry and evolved into cross-media image reproduction, involving device-independent color imaging, gamut mapping and color-accurate computer graphics rendering with spectral imaging. However, CIE colorimetry did not provide a complete solution for image specification under widely disparate viewing conditions. Thus, color appearance models were developed to extend CIE colorimetry to the prediction of color appearance. CIECAM97s and CIECAM02 are the most widely studied color appearance models. While these models saw some successful applications in image reproduction, they did not adequately address spatially complex image appearance and image quality problems. Color appearance models do not directly incorporate any of the spatial and temporal

properties of human vision and the perception of images. They treat each pixel as independent stimuli. To address the issues of device independent color imaging and modern color management systems with regard to spatial properties of vision, image perception and image quality, the concept of image appearance models was introduced [**Fairchild 2003**]. Apart from the attributes such as lightness, brightness, colorfulness, chroma and hue, which are adequately predicted by color appearance models, image appearance models also encompass different image attributes like sharpness, graininess, contrast and resolution. Image appearance models are essentially based on uniform color space. One of the most well-known image appearance model is iCAM [**Fairchild 2004**]. This model was also extended to predict the appearance of digital video sequences and high dynamic range scenes. Implementation of a temporal low-pass filter was proposed to model the time-course of chromatic and light adaptation of rendering applications. Conversion of spatial filter to spatio-temporal filters for image difference and quality applications was also proposed. Image appearance models like iCAM employ image-wise predictors of lightness, chroma and hue, from which an image quality metric can be derived.

Table 3.1 summarizes various video quality metrics discussed in this chapter. Some of these metrics were based on psychophysical approach relying on vision modeling, and others were based on engineering approach. Assessment of the developed metrics either was subjective, based on psychophysical experiments, or objective, based on various statistical metrics, or a combination of the two. In some cases, previously collected visual data were used to test the model, which has also been classified as a subjective assessment in this table.

Table 3.1 A summary of various proposed video quality metrics

Authors	Approach (Psychophysical/ Engineering)	Assessment (Subjective/ Objective/ Both)	Description
Brill et al [**Brill 2004**]	Psychophysical	Subjective	Proposed techniques for determining the statistical resolving power and accuracy of a video quality metric (VQM) and thus for cross-calibrating various VQMs
Caviedes et al [**Caviedes 2000**]	Engineering	Both	Created a combined impairment metric for MPEG video quality assessment
Caviedes and Oberti [**Caviedes 2003**]	Engineering	Objective	Developed a no-reference objective quality metric for measuring improved and degraded video
Daly [**Daly 1993**]	Psychophysical	Objective	Introduced an algorithm for the prediction of visual differences between two digital images based on a vision model
Lindh et al [**Lindh 1996**]	Psychophysical	Subjective (visual data fitting)	Presented a vision model for moving pictures
Lubin [**Lubin 1997**]	Psychophysical	Subjective (visual data fitting)	Operation and general structure of Sarnoff JND model was described
Miyahara et al [**Miyahara 1998**]	Engineering	Both	Presented a new methodology to obtain a picture quality scale (PQS) for coded achromatic still images
Tan et al [**tan 1998**]	Engineering	Objective	Proposed a two-stage full-reference objective measurement model for MPEG-coded video
Tong et al [**Tong 1999**]	Psychophysical	Subjective (visual data fitting)	Proposed a video quality metric based on various models of human contrast sensitivity
Watson et al [**Watson 1999**]	Psychophysical	Subjective (visual data fitting)	Described Digital Video Quality (DVQ) metric
Winkler [**Winkler-1 1999**]	Psychophysical	Subjective (visual data fitting)	Presented a perceptual distortion metric for color video sequences
Winkler [**Winkler-3 2001**]	Psychophysical	Subjective (visual data fitting)	Integrated image appeal attributes like sharpness and colorfulness into the perceptual distortion metric proposed in [Winkler-1 1999]
Wu and Yuen [**Wu 1997**]	Engineering	Objective	Presented a no reference, generalized block-edge impairment metric

3.3 Standardization of Video Quality Assessment and Metrics

Since the early nineties, the rapid evolution of digital video technologies posed a significant challenge to the performance measurement task, which necessitated the development of a new measurement methodology for testing the performance of digital video systems. With this goal, American National Standard (ANSI) approved a standard in 1996, called ANSI T1.801.03. This standard provided much needed set of objective quality metrics that showed high correlation with subjective evaluations of digital video impairments. Wolf [**Wolf-1 1997**] presented an overview of different parameters that were a part of the American National Standard (ANSI) T1.801.03. These parameters were technology-independent and were relevant for a wide range of digital video compression, storage and transmission.

ANSI T1.801.03 did not include MPEG video systems and did not cover bit rates between 1.6 and 10 mbps. Wolf et al [**Wolf-2 1997**] presented the results from two MPEG studies, MPEG 1 and MPEG 2 codecs, whose bit-rates ranged from 1.5 to 8.3 mbps. An analysis of the results revealed that the objective video quality metrics primarily measured four features, addition of false edges, lost sharpness of edges, added motion and lost motion. A set of three or four of these measures achieved a high correlation with subjective responses. The authors concluded these could be used as effective predictors of subjective quality ratings for entertainment video systems.

Since digital video quality depends upon the dynamic characteristics of the input video and the digital transmission system, an accurate perception-based measurement must be performed in-service. The Institute of Telecommunication Sciences (ITS) developed spatial-temporal

distortion metrics that were valid over a wide range of quality and that could be used for in-service quality monitoring [**Wolf-3 1999**]. It was essentially a reduced reference quality metric, for which reference information was extracted from the spatial-temporal region. Horizontal and vertical edge enhancement filters to estimate spatial gradient emphasized edges as long as 10 arc min, and suppressed a large amount of noise. Spatial-temporal feature compression factor of 384 was achieved. Two separate visual masking functions emulated human perception. Seven subjective data sets spanning a wide range of bit rates, test scenes and digital video systems were used to evaluate the metrics. The metrics accounted for a large percentage of variance of the mean opinion scores. The size of the spatial-temporal region could be adjusted to match the bandwidth of the in-service data channel.

IEEE Broadcast Technology Society Subcommittee on Video Compression Measurements initiated an approach to the issue of video quality assessment with the aim of developing a scale of video impairment and unit of measure to describe video distortion from both perceptual and engineering standpoint [**Libert 2000**]. The aim was to set a standard for specifying and testing new equipment used in television production and broadcasting. It was proposed that the IEEE study would attempt to define a scale of video impairment in terms of multiple measurements of the just-noticeable difference (JND) of compression-induced video impairments. The subcommittee agreed on using actual video clips with a single type of impairments for the subjective test. The results of this proposed study was never reported. However, in response to the perceived urgent need in the industry for the sanctioned guidance on video quality, the Telecommunications committee of Alliance for Telecommunications Industry Solutions (ATIS)

released a series of technical reports. One of those reports provided full implementation details for Sarnoff's JND model [**ATIS 2001**].

The Video Quality Experts Group or VQEG was formed in October 1997 in Turin, Italy to address video quality issues. The first task undertaken by VQEG was to provide a validation of objective video quality measurements methods leading to recommendations in both the Telecommunication (ITU-T) and Radiocommunication (ITU-R) sectors of the International Telecommunications Union (ITU). VQEG outlined, designed and executed a test program to compare subjective video quality evaluations to the predictions of a number of proposed objective measurement methods for video quality in the bit rate range of 768 kbps to 50 mbps [**VQEG 2000**]. VQEG solicited submission of objective models to be included in an ITU verification process leading to one or more ITU recommendations. It required all models to receive as input a processed sequence and its corresponding source sequence. Based on this input, the model was supposed to provide one unique figure of merit that correlated with the value obtained from subjective assessment of the proposed material. A set of test sequence was selected and two test sequences for subjective and objective evaluations were executed in parallel. Psychophysical experiment was performed on a total of 287 viewers to collect the subjective data, while ten objective models were evaluated through statistical analysis with respect to three aspects of their ability to estimate subjective assessment of video quality, namely prediction accuracy, prediction monotonicity and prediction consistency. The result of the test did not find an objective measurement system that was able to replace subjective testing. Depending on the metric used for the evaluation, the performance of eight or nine out of ten models was found to be statistically equivalent, leading to the conclusion that no single model

outperformed the others in all cases. One of the major achievements of the first validation effort by VQEG was the unique data set assembled to help future development of objective models.

3.4 Subjective Assessment of Video Quality

Subjective assessment involves human observers who evaluate, compare or assess the quality of a given video. Although more time and resource intensive than an objective assessment, it is the most reliable way to determine perceived picture quality. Various aspects of video quality testing have been dealt in detail in [**Wu 2006**].

Corriveau et al [**Corriveau 1999**] used different test methods to evaluate the same video materials, and compared the stability of measurements in presence of contextual effect (due to varied impairments). Double Stimulus Continuous Quality Scale (DSCQS), Double Stimulus Impairment Scale method variant II (DSIS II) and comparison method were used in this study. DSCQS was found to be free from contextual effect, while comparison method had moderate and DSIS II had large contextual effect.

Pinson and Wolf [**Pinson 2003**] used data sets from six different subjective video quality experiments performed with single stimulus continuous scale evaluation (SSCQE), double stimulus continuous quality scale (DSCQS) and double stimulus comparison scale (DSCS) methods. A subset of video clips from each of these six experiments were combined and rated in a secondary SSCQE subjective video quality test. It was found that SSCQE with hidden surface removal and multiple randomized viewer orderings produced quality estimates comparable to DSCQS or DSCS value, which indicated that viewers performed the same error pulling function

in all methods. The study also showed that the SSCQE test subjects utilized at most 9 to 15 seconds of video. Based on their findings, the authors concluded that properly designed SSCQE testing might be an effective substitute for more complicated DSCQS method.

Pearson [**Pearson 1998**] reported a three-stage method of measuring time-varying video quality. The first stage was an SSCQE method of instantaneous quality, the second stage was a calibration process to convert SSCQE to DSCQS metric, and the last stage was a numerical procedure for relating continuous and overall quality. It was found in this study that three effects played a role in subjective tests. These are: i) a forgiveness effect (momentary occurrence of poor quality ignored after a period of time, ii) a recency effect (quality in last 10-20 seconds of the presentation has a significant influence on overall perceived quality), and iii) a negative-peak effect (more influence of depth rather than the duration of the negative peaks in quality).

Zhao and de Haan presented a subjective assessment of various de-interlacing techniques [**Zhao 2005**]. Five typical algorithms were used in split screen paired comparison experiments. The test sequences included stationary, horizontal and vertical moving sequences, and object with complex motion. Rankings of different methods were derived from the experiment and were compared with objective scores using peak signal-to-noise ratio (PSNR). Since the subjective and objective scores were highly correlated, the authors concluded that objective performance criteria like PSNR are good predictors of quality of reconstructed video resulting from various de-interlacing techniques.

3.5 Conclusions

As this chapter illustrates, a vast body of literature is available on objective video quality metrics. However, the development of reliable metrics is still in a nascent stage, with many challenging issues remaining to be resolved [**Wu 2006**]. More comparative analysis is needed to evaluate the prediction performance of metrics. More experiments with natural images need to be conducted so that more visual data are available for vision modeling purposes. The focus of such experiments should be on supra-threshold conditions, rather than on the threshold of visibility. A reliable perceptual video quality metric will eventually help in benchmarking various video processing techniques. This will require coordinated research efforts in the areas of human vision, color science and video processing.

"A Picture of many colors proclaims images of many Thoughts." ~ Donna A. Favors (Member of the Board of Directors of the Montgomery Institute, 1955)

Chapter 4

METHODS FOR COLOR AND CONTRAST ENHANCEMENT IN IMAGES AND VIDEO

The goal of color and contrast enhancement in general is to provide a more appealing image or video with vivid colors and clarity of details. These enhancements are intimately related to different attributes of visual sensation. It is important to define these attributes before discussing the objectives of color and contrast enhancement, or the various methods to achieve them.

Perceptual Attribute	Definition
Brightness	Attribute of a visual sensation according to which an area appears to emit more or less light. [**Fairchild 2005**]
Lightness	Attribute of a visual perception by which a perceived color is judged to be equivalent to one of a series of grays ranging from black to white. [**Berns 2000**] In other words, it is the brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting. [**Fairchild 2005**]
Hue	Attribute of a visual perception according to which an area appears to be similar to one of the colors, red, yellow, green, and blue, or to a combination of adjacent pairs of these colors considered in a closed ring. [**Berns 2000**]
Colorfulness	Attribute of a visual perception according to which an area appears to exhibit more of less of its hue. [**Hunt 2001**]
Chroma	Colorfulness of an area judged in proportion to the brightness of a similarly illuminated area that appears to be white or highly transmitting. [**Hunt 2001**]
Saturation	Colorfulness of an area judged in proportion to its brightness. [**Hunt 2001**] By definition, saturation = chroma/lightness.

The objective of contrast enhancement is to increase the visibility of details that may be obscured by deficient global and local lightness. The goal of color enhancement can be either to increase the colorfulness, or to increase the saturation. Increasing the lightness can give a perception of increased colorfulness, however in this case perceived saturation reduces for a given chroma. On the other hand, perceived saturation can be increased by increasing chroma or reducing lightness, or both. If chroma is increased moderately while slightly reducing the lightness, both saturation and colorfulness in an image can be enhanced. This method is also likely to avoid out-of-gamut or unrealizable colors.

The next section discusses several previously published methods for color and contrast enhancement that served as a preamble for the current research. Then, the development of the new algorithm is discussed in detail, starting with the working requirement, the color space chosen for the development, a detailed description of the three key components of the algorithm, and finally the innovation that was achieved in this work.

4.1 Color and Contrast Enhancement in Digital Images: A Review of Past Research

Most of the published works to date focus on color enhancement in digital color images. Many of these techniques can theoretically be implemented for video as well. However, hardware implementation issues can impose serious restrictions for many of these techniques. These issues have not been considered in this discussion. Also note that the methods discussed here do not involve signal processing as much as image processing. Image enhancement methods relying on signal processing do not involve a great deal of perceptual processing that will be considered

appropriate from color science standpoint. [**de Haan 2003**] A brief overview of such methods is included in Chapter 2.

4.1.1 Color Processing in LHS Space

In the early '80s, as color images and video media started getting increasingly commonplace, researchers soon realized that most of the enhancement techniques developed for monochrome images led to artifacts when applied to color images. In one of the earliest papers on color enhancement, Strickland, Kim and Mcdonnell [**Strickland 1986, Strickland 1987**] recognized that RGB color space did not correspond with the human color perception and so, image enhancement algorithms applied directly to RGB images could lead to color artifacts. They suggested performing enhancement operations in a color space whose dimension corresponded to luminance, hue and saturation. The authors also pointed out that enhancing luminance alone could lead to color artifacts in low luminance regions, and thus simultaneous saturation processing was required for proper enhancement. They presented the derivation of LHS coordinates from RGB. Because of the nonlinear transformation between the two color spaces, some processed colors were at risk of being out-of-gamut when converted back to RGB. Strickland, Kim and Mcdonnell proposed to clip the RGB pixel vector at the color cube boundary to prevent color shift during the clipping operation.

4.1.2 Histogram Based Methods

Histogram equalization is a common approach for enhancing contrast and brightness in grayscale images. Extending this tool to color images is not straightforward. Color histogram equalization is a three-dimensional problem. Moreover, RGB is not a suitable color space because of its poor

correlation with human visual system, requiring a color space transformation. Histogram equalization on the intensity component can improve the contrast, but can cause de-saturation in areas where histogram equalization results in a large reduction in intensity. Similarly, equalization of the saturation component alone can lead to color artifacts. Independent equalization of RGB components is also not advisable as it can lead to a hue shift. In one of the earliest papers on color processing in color difference space (YCC), Hague, Weeks and Myler [**Hague 1994**] presented an approach where histogram equalization was performed on saturation for each hue in the image, taking into account the maximum allowable saturation for a given luminance. The color space was segmented into various pie-shaped hue regions, each of which was further divided into several luminance regions, as shown in Figure 4.1. For each of these regions, the minimum possible saturation value was zero, while the maximum possible saturation value was a function of both hue and luminance. Maximum allowable saturation for each region was determined by computing the saturation for every RGB combination and retaining the largest value computed within each hue region for all different luminance regions. Once the saturation limits were determined, histogram equalization was applied to each of the luminance regions within each hue region. Saturation equalization was followed by luminance equalization over the entire luminance image. This method helped reduce the number of out-of-gamut colors as well as color artifacts.

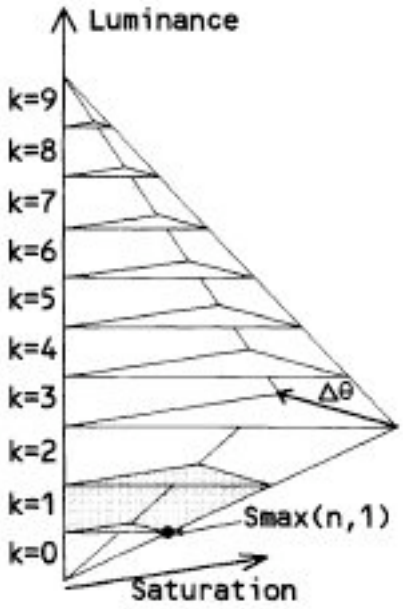


Fig. 4.1 A single C-Y hue region, divided into different luminance regions [**Hague 1994**, Fig. 4]

The main limitation of the above method, or of any histogram dependent processing for that matter, is that it is a global method and has little control over local contrast. It does not take into account the content of input images. The other critical disadvantage is that this and other histogram based methods do not consider the perceptual aspects of human visual system. As a result, a change in the lightness and saturation of a given pixel or region may or may not be perceived as a desirable effect.

Weeks, Sartor and Myler [**Weeks 1999**] extended Hague, Weeks and Myler's histogram equalization approach by using histogram specification for color image enhancement in color difference space. In this method desired hue, saturation and luminance histograms were specified separately, but the correlation between different components were also taken into account. An advantage to this approach was that the saturation component could be enhanced while leaving

hue and luminance unchanged. Saturation distribution nevertheless was a function of luminance and hue, since maximum saturation was a function of these two components. Sixty-four luminance regions and sixteen hue regions were considered. A 64x16 array stored the maximum saturation value for each region. Sixteen saturation histograms were specified for each hue region, and within each hue region 64 specified histograms were generated, scaling each of those by the maximum saturation for the given luminance/hue region. In other words, saturation histograms for different luminance regions under a given hue region had the same shape, but different widths due to varying maximum saturation as a function of luminance. Best performance was achieved by first equalizing the luminance component and then redistributing the saturation component across complimentary hue regions according to a Gaussian distribution. Figure 4.2 shows the input specified histogram for one of the test images used in Hague, Weeks and Myler's work [**Hague 1994**] and the resulting saturation histogram for a single intensity/hue region in the enhanced image.

Although this method performed better than histogram equalization in enhancing the image and reducing color artifacts and hue noise, this method was not fully automatic since different histogram specifications had to be adopted for different images.

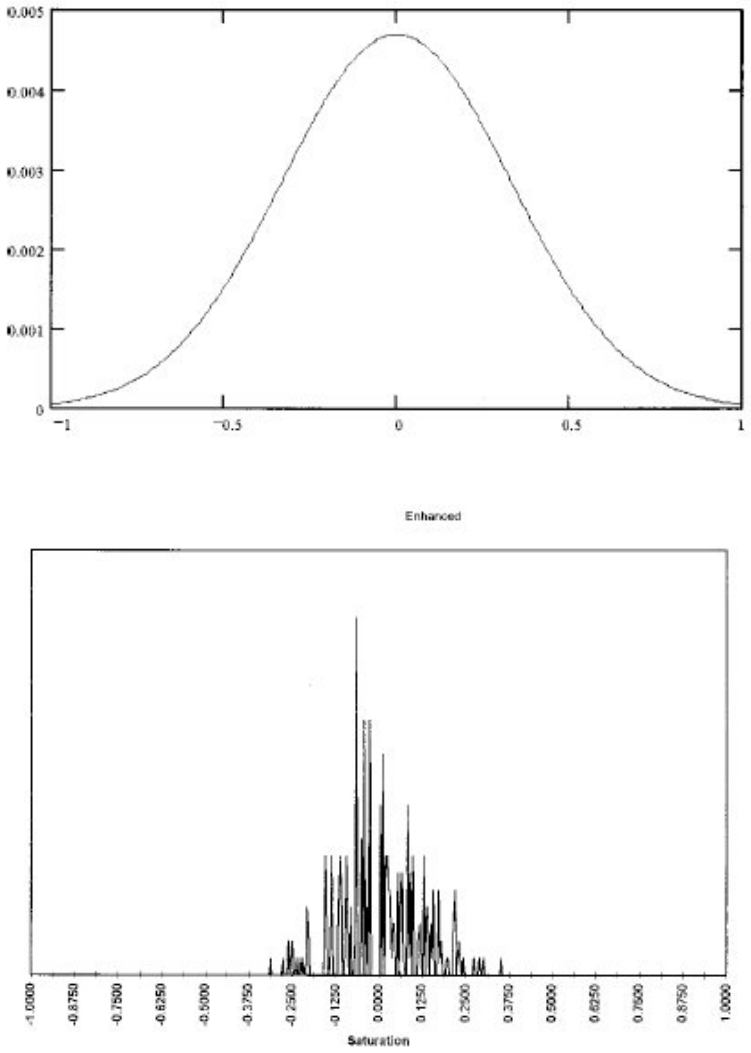


Fig. 4.2 Specified histogram saturation for one of the test images (top) and the saturation histogram for a single intensity/hue region in the saturation enhanced image [**Weeks 1999**, Fig. 9 and 12]

4.1.3 Color/Contrast Enhancement Method Based on the Chromaticity Diagram

In a different approach, Lucchese, Mitra and Mukherjee [**Lucchese 2001**] presented a two-stage method for color contrast enhancement based on xy chromaticity diagram. All colors with positive chroma values were maximally saturated through shifting to the borders of a given color gamut. In the next stage, the colors were desaturated toward a new white point by an appropriate

color-mixing rule. Lucchese et al's method is shown in Figure 4.3. RGB coordinates define the color gamut. W is the white point. Saturation of any color C_1 is enhanced by moving the point along the straight line joining W and C_1 to the point S on the spectrum locus. Next, a color mixing law is used to desaturate S and compute coordinates for the final color C_2.

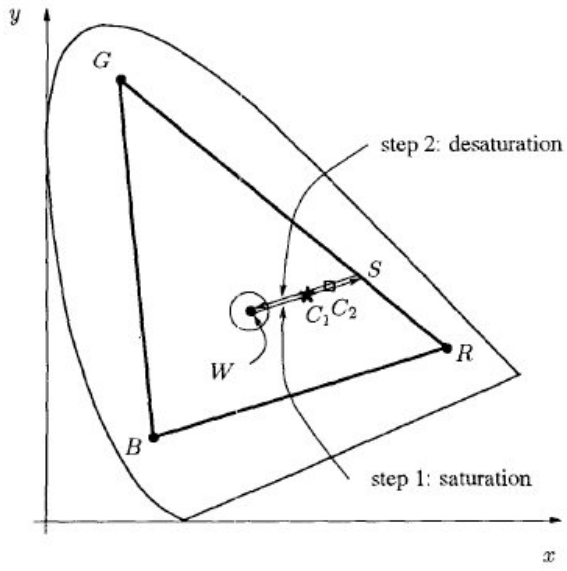


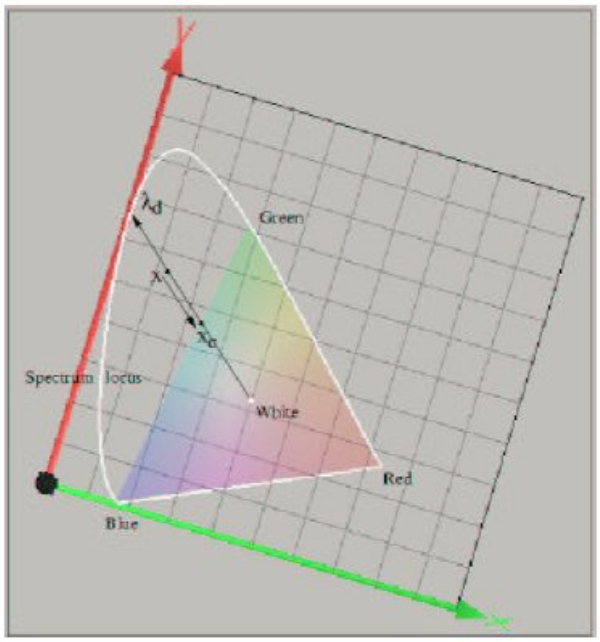
Fig. 4.3 Color enhancement using chromaticity diagram (C_1 and C_2 are color coordinates before and after enhancement) [**Lucchese 2001**, Fig. 1]

The method described above suffers from a serious flaw. Chromaticity diagram does not represent a perceptual space. Thus, a straight line joining the white point and a given point in the color gamut on the chromaticity diagram is unlikely to preserve constancy of perceived hue.

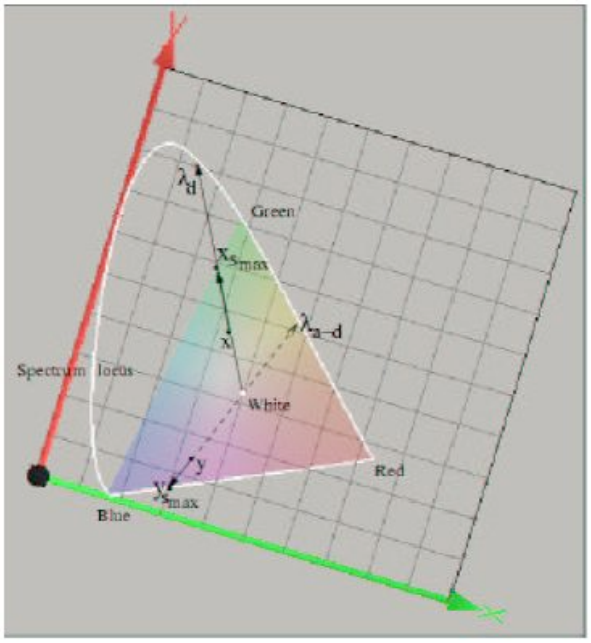
In yet another attempt to develop an image enhancement method based on the chromaticity diagram, Colantoni, Bost and Trémeau [**Colantoni 2004**] used λSY color space for colorfulness enhancement. This space is derived from xyY color space and is based on the dominant

wavelength (λ), saturation (S) and intensity (Y). Figure 4.4 describes the idea. Any real color x lying within the region enclosed by the spectrum locus and the lines BW and RW can be considered to be a mixture of the white point chromaticities and those of the spectrum light of the dominant wavelength (λ_d). The dominant wavelength is obtained by extending the line WX until it intersects the spectrum locus. The color triangle defines the device gamut.

Four methods were developed to increase the saturation of the color of each pixel in the direction of dominant wavelength, thereby enhancing the colorfulness of images. In the first method, the original saturation was increased by different fractions. In the second method, the saturation component was increased to the maximum saturation, which was a function of both hue and luminance. This method was found to greatly increase color contrast at the cost of perceptual quality. The third method reduced the luminance component by a given fraction and then increased the saturation component in the same way as the first method. In the last method, fractional luminance reduction was followed by increasing the saturation component to the maximum saturation corresponding to the adjusted luminance.



(a) Clipping of an out-of-gamut color (from x to x_c)



(b) Extending a color in the direction of dominant wavelength (x to x_s) and anti-dominant wavelength (y to y_s)

Fig. 4.4 Color enhancement in λSY color space [**Colantoni 2004**, Fig. 1]

The authors found a weak dependence between image content and chroma and lightness changes. Very strong colorfulness enhancement resulted in poor image quality. Also, increased saturation was found to introduce hue noise in uniform areas and background.

While this method is fast and inexpensive, λSY is by no means a perceptual color space. The above method is likely to result in a hue shift since it ignores the fact that constant hue lines are curvilinear in a chromaticity diagram. The perceived hue shift will depend on the amount of saturation enhancement. Further, an equal amount of change in the chromaticities in different regions of the diagram will lead to a different amount of perceptual color difference.

4.1.4 Saturation Clipping in LHS and YIQ Color Space

As already discussed, since RGB color space does not conform to human perception of color, many a time color processing is done by first transforming RGB image to a new color space, and converting it back to RGB once the processing is complete. However, due to the fact that the useful range of saturation decreases as one moves away from the medium luminance values, upon conversion back to RGB, it is possible to end up with illegal (out of gamut) colors. One solution is to clip the processed luminance value before transforming back to RGB, but this can lead to artifacts such as bright spots, washout regions, and loss of local contrast at the end regions of the range since many pixels can be clipped to the same luminance.

Yang and Rodriguez proposed a hue-preserving graphical approach involving scaling and shifting that bypassed computationally intensive coordinate transformation during color image processing [**Yang 1995**]. This method was intended for cases where only the luminance or only the saturation component needed to be modified. Later, the same authors proposed a method where the saturation of an out-of-gamut color was clipped, instead of luminance [**Yang 1996**]. This method was implemented in LHS and YIQ (formulated by NTSC) color spaces. The method is depicted in Figure 4 (a). The luminance of a pixel with saturation S changes from L to L' after processing. The point (L´S) will map to outside the RGB gamut, so it needs clipping. The point can be clipped to (L´´,S) by reducing the luminance, but may lead to reduced perceived contrast. Alternately, the saturation of the point (L´S) can be reduced to obtain the point (L´,S´), which will also map to an in-gamut RGB value. Similar processing can be performed in YIQ space, as shown in Figure 4.5 (b).

Saturation clipping resulted in improved contrast compared to the results obtained from luminance clipping. However, this method is intended for applications where only the lightness is being enhanced.

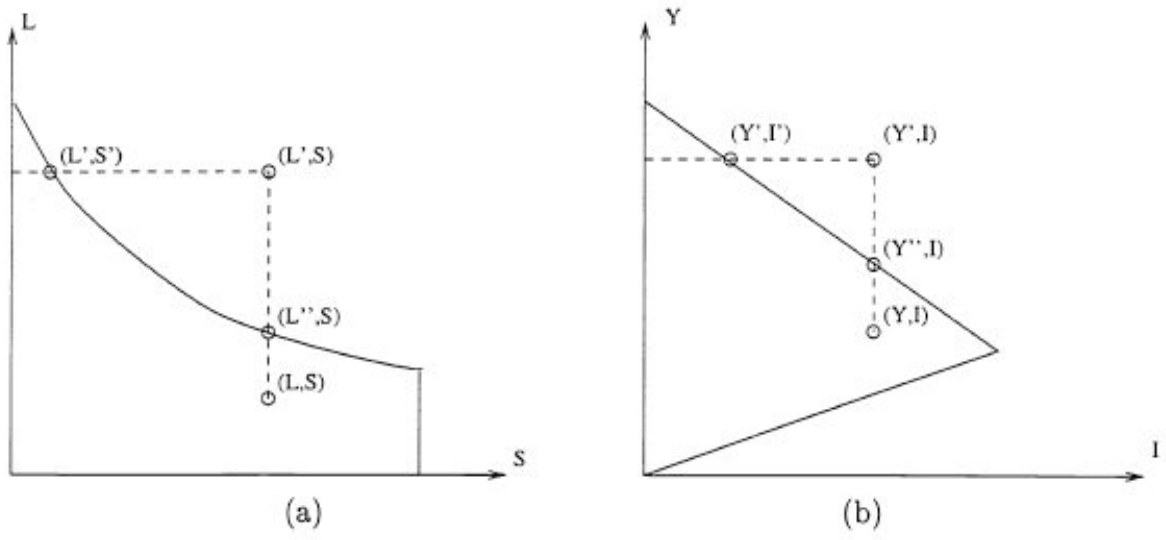


Fig. 4.5 Saturation clipping for red hue plane in (a) LHS and (b) YIQ [**Yang 2006**, Fig. 4]

4.1.5 Retinex-Based Image Enhancement Methods

In one of the significant contributions, Edwin Land conceived the retinex theory to model visual perception of lightness and human vision color constancy [**Land 1983**, **Land 1986**]. The theory was formulated based on the experimental demonstrations that color appearance was controlled by surface reflectances and spatial distribution of colors in a scene rather than the spectral perperties of reflected light. In its most recent form, the concept evolved into a center/surround spatially opponent operation related to neurophysiological functions. Proposed mechanisms were thought to be a combination of retinal and cortical mechanisms. The color of a unit area was determined by a trio of lightness numbers corresponding to three single wavebands, long, middle

and short. The numbers together represented the relationship between the unit area and the rest of the unit areas in the scene. The output for a given waveband was determined by taking the ratio of the signal at any given point in the scene and normalizing it with an average of the signals for that waveband throughout the scene. Thus, retinex theory acknowledged the influence of background in determining the color of the stimulus [**Fairchild 2005**]. Even though the retinex model has some weaknesses in physiological modeling, retinex theory has been widely used in various application areas, including image enhancement.

Meylan and Süsstrunk [**Meylan 2004**] proposed a retinex-based method for high dynamic range rendering of natural images. At first, a global tone mapping was applied on the linear image. Luminance component was then computed from the non-linear RGB image. An adaptive filter algorithm based on retinex was applied to luminance data. Then, the modified luminance and original chrominance components were transformed back to RGB. Finally, the RGB image was scaled to the output device dynamic range using hisotgram scaling. The key feature of the low-pass retinex-based filter was that the surround function was not circularly symmetric, but followed the image's high contrast edges. The filter coefficients were determined by traversing the surround radially for each rotation angle. Further, the radial one-dimensional function was a Gaussian curve with spatial constant varying with the local contrast. The method reduced artifacts like black halos around light sources, but the processing time increased significantly.

Hsu and others [**Hsu 2006**] used non-isotropic Gaussian kernel filters in a multiple-scale retinex method. The logarithmic intensity of the image was subjected to a low-pass Gaussian filter with adaptive width, following which weighted multi-scale retinex ratios were computed. Chromatic

information from the input image was used to restore color, and then luminance histogram equalization was performed to enhance image contrast. Retinex ratios and equalized intensities were multiplied to get the final output. The authors reported better local contrast and rendering effects than previously published methods.

Rahman, Jobson and Woodell focused on multi-scale retinex based approach for color image enhancement [**Rahman 1996**, **Rahman 2004**]. Main goals were to achieve an image rendering close to the original scene, and to increase the local contrast in dark regions of high dynamic range scenes. In order to achieve a balance between the dynamic range compression and tone mapping, multiple-scale surround with Gaussian distribution was used. A color restoration method was proposed to compensate for the desaturation effect inherent in retinex-based methods due to non-conformity to gray world assumption both globally and locally. Color restoration took the form of a logarithmic spectral computation. However, color restoration was found to be inadequate for preserving the saturation of the lighter colors, and thus a white balance process was introduced to address this issue.

In a different approach, Choi et al [**Choi 2007**] proposed a color image enhancement method based on the single-scale retinex with a just noticeable difference (JND)-based nonlinear filter. The processing was done in the HSV color space. Only S and V components were enhanced, and original hue was maintained. To enhance the V component, the illumination was first estimated using the JND-based filter. A fraction of the logarithm of the estimated illumination was subtracted from the logarithm of the input V component, which was followed by histogram modeling to obtain output V component. The S component of the image was enhanced in equal

proportion as the V component. Finally, RGB output image was computed from the output HSV image. The results were found to be superior than conventional histogram equalization and standard single-scale retinex methods.

For any color enhancement method based on retinex theory, the main weakness lies in the fact that no direct interdependence is assumed between the luminance and chrominance data. Even though methods like color restoration described above were proposed as an extension, it is very difficult to maintain the relationship between lightness and chroma. Note that nonlinear processing is performed on the luminance data in a color space where luminance and chrominance data are not necessarily decoupled. Algorithms based on retinex theory are in many ways simply lightness adjustment and/or local contrast enhancement algorithms. Further, retinex-based methods are computationally expensive, making them difficult to implement in commercial imaging devices.

4.1.6 Geometrical Method for Lightness Adjustment

Samadani and Li [**Samadani 2006**] proposed a method for lightness adjustment that was motivated by the fact that the color of an object of uniform material varies along a luminance-saturation curve as the intensity and the angle of the light source changes. Their approach focuses on lightening or darkening of an image where colors are directly adjusted by moving them along lightness-saturation curves (Figure 4.6) while leaving the hue unchanged. In the simplified version of the method, for each hue, the saturation is assumed a separable function of luminance and a scale parameter (Figure 4.7). The maximum saturation point determines the shape of the curve, which is a function of luminance, and implicitly, of hue. Curves can be stored

as Lookup Tables for three primary and three secondary hues, and saturation value corresponding to an intermediate hue can be obtained through interpolation. All processing is done in the YCC space. The authors compared the results with that of a standard algorithm where only the luminance was processed and out-of-gamut colors were clipped to the boundary. Samadani and Li's method offered improved clipping and avoided excessive or insufficient saturation of pixels.

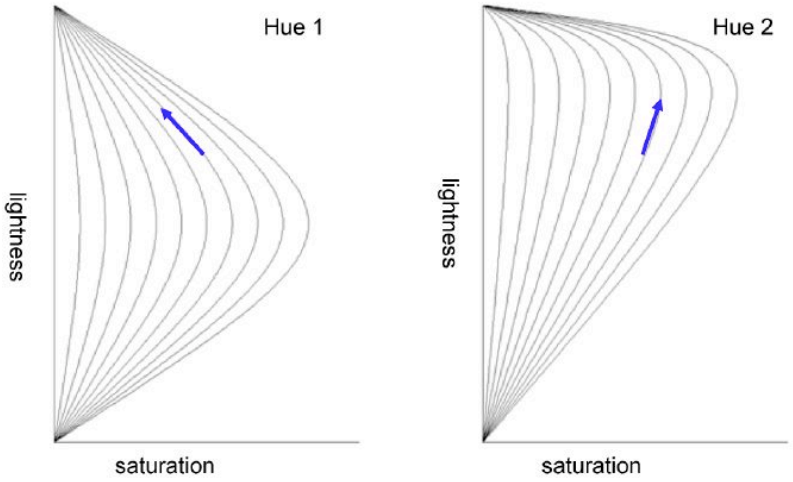


Fig. 4.6 Saturation-lightness curve families for two different hues [**Samadani 2006**, Fig. 2]

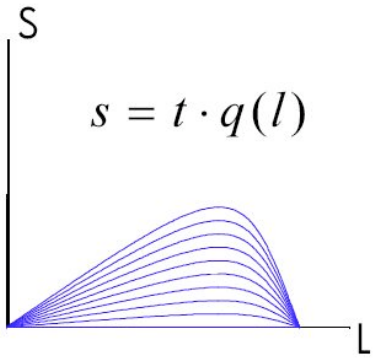
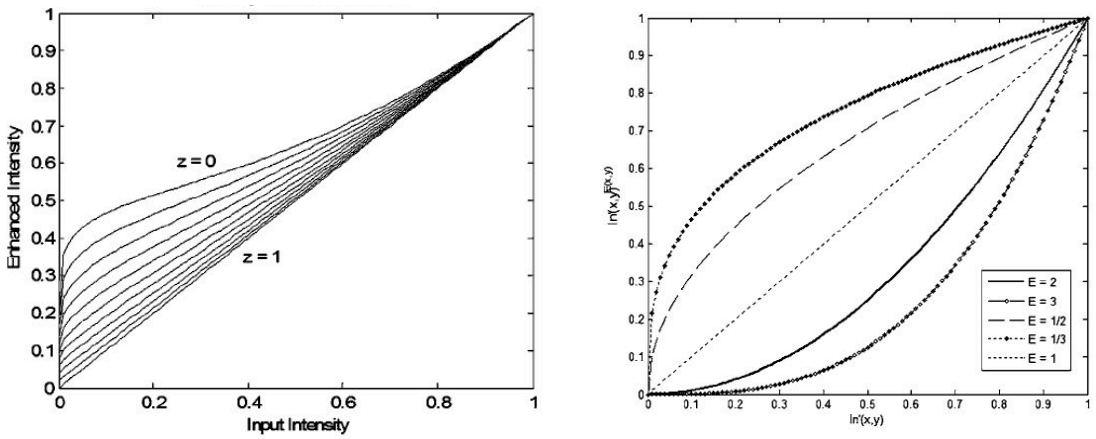


Fig. 4.7 Saturation as a separable function of luminance [**Samadani 2006**, Fig. 4]

4.1.7 AINDANE: Locally Adaptive Image Enhancement

Tao and Asari **[Tao 2004]** proposed a nonlinear image enhancement method called AINDANE (adaptive and integrated neighborhood dependent approach for nonlinear enhancement). This algorithm specifically addresses visual quality issues of digital images captured under low or non-uniform illumination conditions. The algorithm involves two independent processes, adaptive luminance enhancement for dynamic range compression and adaptive contrast enhancement to preserve visual details, followed by the tone reproduction of the original image. In the adaptive luminance enhancement, the intensity values are subject to a nonlinear transfer function, shown in Fig. 4.8 (a), which also serves as dynamic range compression as the intensities of darker pixels are increased. One of the parameters in the transfer function is the intensity corresponding to the cumulative distribution function (CDF) of 0.1. As a result, the luminance enhancement depends on the lightness of the original image. Intensity of darker pixels are increased significantly, while pixels with sufficient intensity remain unchanged.

In the following stage, the local contrast enhancement is achieved by taking into account the intensities of surrounding pixels by using a 2D discrete spatial convolution with a Gaussian kernel. The luminance-enhanced image is subject to an exponential, which is an expression of the ratio of convolved intensity and the original intensity. This exponential in turn is raised to the power of a function of the global standard deviation. The transfer function is shown in Fig. 4.8(b). Once the adaptive contrast enhancement is complete, a linear color restoration process is performed based on the chromatic information contained in the input image.



(a) (b)

Fig. 4.8 Nonlinear transfer functions for (a) adaptive luminance enhancement and (b) adaptive contrast enhancement [**Tao 2004**, Fig. 3 and 4]

Tao and Asari's method suffers from a major technical flaw in the final stage. The assumption that the relationship between the RGB channel values will be maintained even after several nonlinear processing is fundamentally wrong. This also requires a manual adjustment of the color correction operation. However, the adaptive luminance enhancement process was reported to produce superior results compared to other published methods. The global lightness adjustment method proposed in the present research has been partly inspired by Tao and Asari's method.

4.1.8 Sigmoidal Lightness Rescaling Function

Braun and Fairchild **[Braun 1999]** proposed sigmoidal lightness rescaling function to maintain the perceived lightness contrast that selectively rescaled images from a source device with a full dynamic range into a destination device with a limited dynamic range. This achieved the effect of enhancing the overall contrast by increasing the lightness difference between the highlight and

shadow regions. This method, essentially developed for color gamut mapping of pictorial images, uses lightness remapping based on the phenomenon of simultaneous lightness contrast. The highlights are lightened and the shadowed regions are darkened, thus increasing the image contrast. However, this strategy was not found to be suitable for the purpose of color enhancement. Increasing the lightness of lighter colors did not leave much room for increasing the saturation in the following step, resulting in increased clipping. Further, reducing intensity of the low-intensity pixels resulted in loss of details in darker parts of an input image.

4.1.9 Local Color Correction Using Nonlinear Masking

When a high dynamic range image has to be reproduced in lower dynamic range displays or other output devices, it becomes difficult to design a global tone reproduction that will accommodate both shadow and highlight detail. Moroney **[Moroney 2000]** presented a local color correction operation based on nonlinear masking. The method is equivalent to deriving a specific tone reproduction curve for each pixel in the image. In this method, one input value can be mapped to many output values, depending on the values of the neighboring pixels. At first, an image mask is derived from an inverted low-pass filtered image, which is then combined with the input image. The combination operator consists of a simple power function.

Moroney's method did not produce desirable results in our case, understandably because of different working requirements. Lightening the dark regions and darkening the light regions by themselves are not appropriate strategies for color and contrast enhancement. The results were not as satisfactory as with the new method described in Section 4.2.

4.1.10 Patented Methods for Color Processing in Images and Video

Many color and contrast enhancement methods developed in the industry can be found not in the publications, but in various patents. A preliminary search of US patent database was conducted as part of this research. This section discusses several approved or pending US patents that are relevant to color image or video enhancement.

Lightness increase in an image often leads to a reduction in the saturation, and lightness reduction often results in saturation increase. A patent filed by Hermann Fuchsberger of AGFA [**US Patent 4,831,434**] offered a simple method for the correction of color saturation by taking the ratio of the luminance signals before and after luminance processing and multiplying their quotient by each of the two chrominance signals. Mainly YCC color space, quite common in the video processing area, was considered in this method. This color saturation correction method was suitable for photographic applications, where maintaining the original color saturation during contrast enhancement was of interest. The fundamental drawback of this method was that it did not take into account the fact that lightness and chroma were not linearly related in a perceptual space. A 10% increase in the luminance signal does not translate to a 10% increase in the color saturation. The respective change in the two chrominance channel signals are also likely to be different.

Bachmann et al [**US Patent 5282021**] invented a method for color correction of a video signal and implemented in the hardware. It is assumed that the video signal is separated into luminance and chrominance information, as in YCC color space. Hue and saturation are computed from the

chrominance or color difference data and converted into polar coordinates. The method involves specifying six overlapping triangular color sectors, where any hue lying in the overlapped regions of two sectors has specific control values corresponding to the two overlapping sectors. These sectors overlap by half of their angular magnitude of the sectors. Thus, the degree of the correction is stored as a function of hue. Based on the control values, corrections for the input hue, saturation and luminance are determined. A color sector manual selection unit enables manual correction of hue, saturation and luminance through a common keyboard interface. An additional signal correlated to the color hue sectors is generated. The two control values and the additional signal value are used to compute the corrected hue, saturation and luminance values. Hue correction is additive while saturation and luminance corrections are multiplicative. In the final step, the corrected hue, saturation and luminance were converted into the color difference data. The main idea behind the concept was to prevent the saturation correction leading to an illegal hue, as well as to prevent a hue correction from changing the saturation of those colors that do not correspond to the given input hue.

Jeong et al implemented a color image enhancement method for video display where sharpness was improved using the saturation component obtained by converting RGB image data into LHS values [**US Patent 6028646**]. Figure 4.9 illustrates the method in a block diagram. The RGB/saturation conversion section normalizes the input RGB image and computes the saturation. The memory section 102 stores the negative image of the saturation. The saturation enhancement section receives this negative image and performs a high-frequency-emphasized filtering with a two-dimensional convolution. The minimum color signal determining section detects the minimum value among the normalized RGB primary color components and

accordingly provides a switching control signal to the switching section 105, which then routes the minimum enhanced primary color component to one of the three different paths in the saturation/RGB conversion section 107. In this section, the negative saturation image is first converted into a positive one before performing enhancement. For example, if the minimum primary color signal is R, the enhanced R component is achieved by multiplying the sum of the original RGB components by the enhanced saturation component, while keeping G and B constant. The other two cases are similar. Delay 107 is to delay the operation of the two conversion sections. The block diagram also contains an optional luminance enhancement section that enhances the Y component in the LHS space obtained from the saturation enhanced image, using the high-frequency-emphasis filter described before and then converts the enhanced data back to RGB.

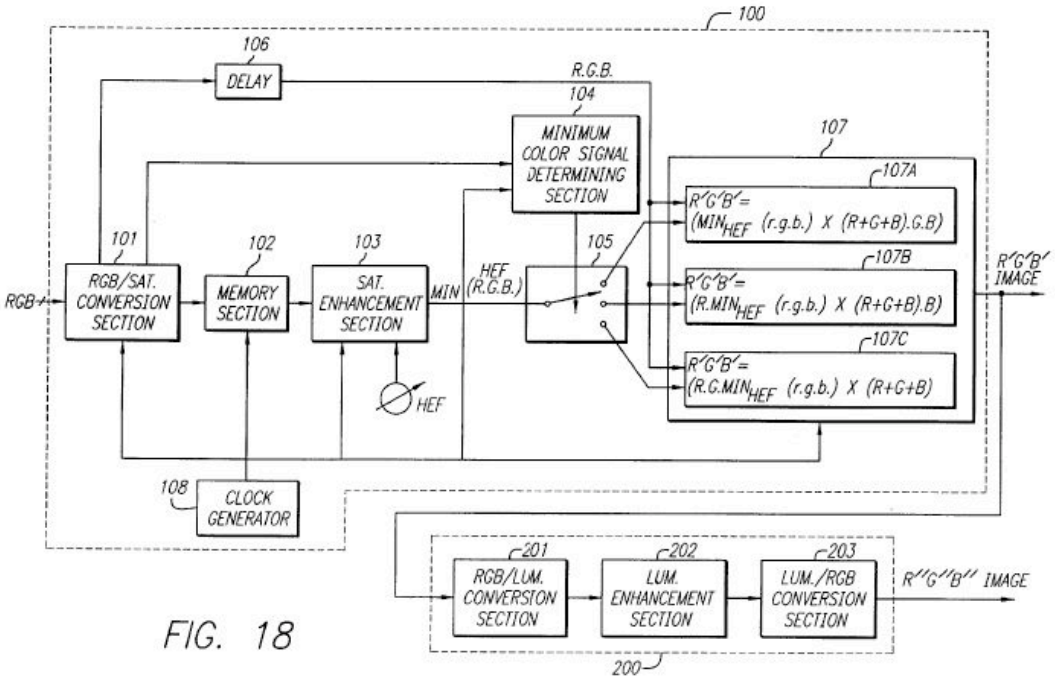


FIG. 18

Fig. 4.9 Color image enhancement device patented by Jeong et al.

Chou developed a method for automatically adjusting, correcting and enhancing the color hue and saturation of each pixel from a color video source [**US Patent 2006/0238655**]. At first, hue and saturation values are computed from input color difference values in the YCC format. Hue and saturation correction values for a given pixel are determined using lookup tables predetermined for various color tone regions. The hue and saturation adjustments are applied directly to the original chrominance data of the pixel. A separate region control lookup table is used to determine the color tone region in which an input hue falls. The region control output determines the saturation lookup table to be used in the enhancement. A fade module also receives the region control output and fades out the saturation adjustment value obtained from the given saturation lookup table to smoothen the effect of color enhancement for input colors falling at the boundaries of color tone regions, and thus avoid contouring effects.

Wang invented a method to perform color enhancement of an image in a specific color space region [**US Patent 2007/0070369 A1**]. The block diagram is shown in Figure 4.10. The image in original color space (RGB, CMY, HSI, YIQ etc) is first converted to YUV color space. The color space is partitioned into five regions corresponding to blue, green, red, yellow and skin color. The color space regional decision and enhancement attenuation calculation modules corresponding to five regions calculate the enhancement region for each pixel, the enhancement amplitude and flat transition around the edge of the enhancement regions to reduce contouring effects. These are determined based on the region parameters stored in the color space regional parameter storage module. The color space component enhance module determines the adjustment amplitude of components in all directions based on the enhancement amplitudes calculated earlier. The regional adjustment module performs a part of the adjustment such as

compensating the luminance and limiting the data boundary, and then sends the adjusted data to RGB component adaptive regional enhancement module. This module computes each enhancement component of the pixel color depending on the color space region it is in, and adds the enhancement components with the corresponding original values and outputs the result.

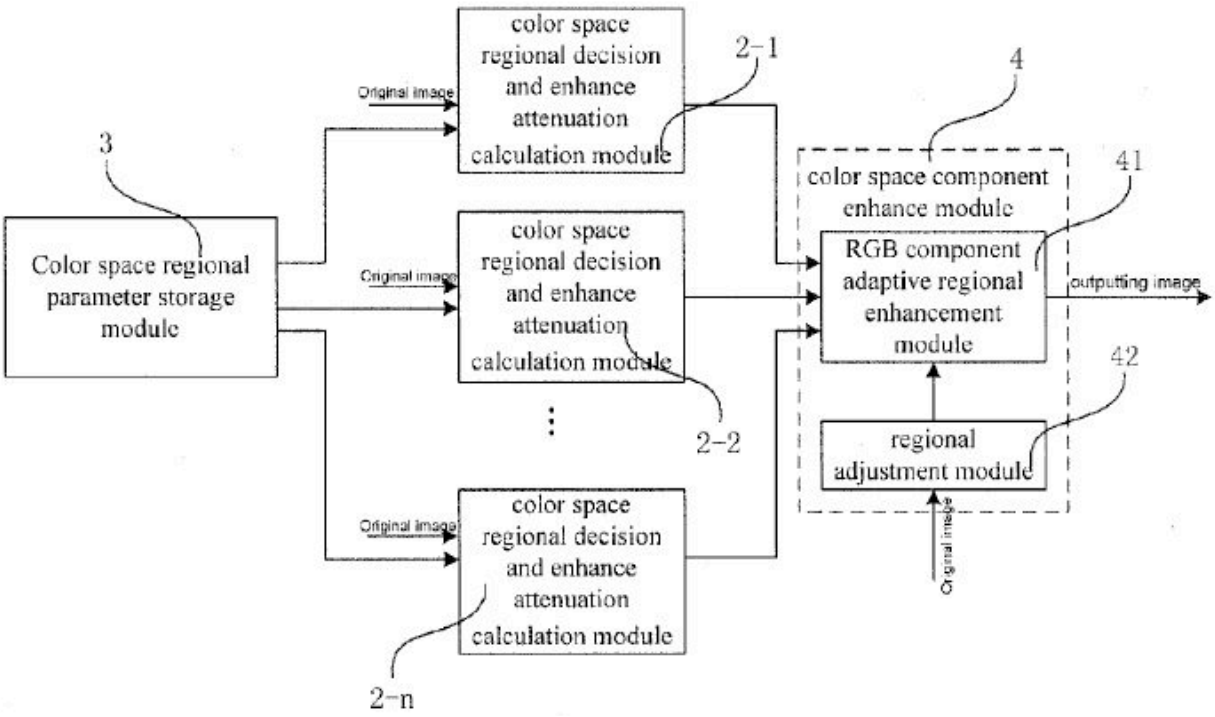


Fig. 4.10 Block diagram of the method discussed in Wang's pending patent

The above discussion on various existing and pending patents on color enhancement for images/video shows that the industry approach on color processing is largely ad hoc. While various smart implementations of color computation have been invented over the years, color processing is still performed in non-perceptual color spaces like RGB, LHS, YCC or YUV. Even though these spaces have their own computational advantages, none of them is a truly perceptual color space. Thus, an enhancement in a given region of such color space can produce a very

different result from that in another region within the same space. This happens because the relationships between the lightness, hue and chroma does not change in a linear fashion in such color spaces. To counter this problem, a color space is generally partitioned into several regions, and different enhancement strategies are employed for different regions. This often leads to contouring issues, which is solved by smoothening the enhancement in the transitional areas of various regions, as seen in case of some of the patented methods discussed in this section. Color processing in a non-perceptual color space has another important disadvantage. Luminance or saturation enhancement can frequently lead to a hue shift. Additional color correction modules are often implemented to resolve such issues.

The choice of color space is fundamentally important in digital color processing. A truly perceptual color space ensures the enhancement process is universal and more reliable, minimizing many artifacts and other issues commonly encountered in existing methods. Proposed algorithm, described in the next section, considers these aspects while employing methods that are image content independent and easy to implement since the processing is not ad hoc.

4.2 New Algorithm: Working Requirements

The main objective in color and contrast enhancement in video processing is to achieve the best possible combination of colorfulness and contrast enhancement in an efficient manner. Typically, the use of independent algorithms for color and contrast enhancement results in sub-optimal enhancement, and unwieldy combined tune-up. Following are the requirements of an image enhancement algorithm to be suitable for implementation in a video processing chain:

1. The algorithm must be automatic, i.e. deal with all content types without external intervention
2. It must integrate color and contrast enhancement functionalities
3. Enhancement must be adaptive to the overall image or video content
4. The overall lightness and saturation should be improved while maintaining the original hue
5. The achromatic colors must remain achromatic
6. Pixels with high saturation should be left unaltered to avoid any perceived loss in sharpness and undesirable color enhancement
7. The algorithm should not produce out-of-gamut colors or color artifacts (e.g. blotchiness)
8. The algorithm must be suitable for implementation in real time on the required target platform (software or hardware)

Temporal processing issues were not considered in this development.

4.3 Color Space

Following are the criteria for the suitability of a color space to be used in the implementation of a color enhancement algorithm:

1. Should be uniform in terms of perceived hue so that enhanced colors do not change in hues
2. Lightness prediction of chromatic colors must not have hue or lightness modulation dependency
3. Should be invertible functional mapping between XYZ (or some other fundamental color description) and the color space dimensions
4. Should be an opponent (has a neutral axis along one dimension) three-dimensional space
5. Transformation to and from a linear color space should not be computation expensive

Previous research showed that IPT color space satisfies all the above criteria [**Ebner 1998**].

The input image is assumed to be standard RGB (or sRGB). sRGB data are converted to CIEXYZ for the 1931 2° standard observer with an illuminant of D65. Eq. 4.1 shows the conversion from XYZ to IPT.

$$
\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0 & 0 & 0.9184 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix}
$$

$$
\begin{aligned}
L' &= L^{0.43}; \quad L \geq 0 \\
L' &= -(-L)^{0.43}; \quad L < 0 \\
M' &= M^{0.43}; \quad M \geq 0 \\
M' &= -(-M)^{0.43}; \quad M < 0 \\
S' &= S^{0.43}; \quad S \geq 0 \\
S' &= -(-S)^{0.43}; \quad S < 0
\end{aligned}
\tag{4.1}
$$

$$
\begin{bmatrix} I \\ P \\ T \end{bmatrix} = \begin{bmatrix} 0.400 & 0.4000 & 0.2000 \\ 4.4550 & -4.8510 & 0.3960 \\ 0.8056 & 0.3572 & -1.1628 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix}
$$

4.4 Details of the Algorithm

The algorithm has three components, global lightness adjustment, local contrast enhancement and finally, saturation enhancement. Following sections describe these components. Details specific to the implementation have been excluded because of the proprietary nature of the work.

4.4.1 Global Lightness Adjustment

In a typical image enhancement operation, increasing the saturation alone is not a good idea, as this can often lead to out-of-gamut or illegal colors. Increasing both saturation and intensity is even more likely to result in out-of-gamut colors. Out-of-gamut colors will necessitate clipping, which essentially leads to loss in contrast and details. Thus, in a typical image with normal exposure, the overall lightness needs to be reduced before increasing the saturation.

However, if an image is mostly dark to begin with, either due to under-saturation or due to high dynamic range of the image (which means a part of the image is quite bright at the same time),

the lightness may need to be increased before any perceived color enhancement can be attained. An example is shown in Figure 4.11. The horizon visible through the opening in the stone-wall is quite bright, while the other parts of the image are fairly dark. In this case, saturation enhancement needs to be higher than the normal to counteract the effect of a perceived loss of saturation due to an increase in the lightness. Thus, we must be able to determine whether an image needs to be darkened or lightened, and to what extent. This is achieved by the global lightness adjustment.

Fig. 4.11 “Cave”: an example image with a high dynamic range

Global lightness adjustment is done essentially through a Lookup Table (LUT). The curve is a straight line below lower threshold and above upper threshold. Thus pixel intensities less than the lower threshold and more than the upper threshold are not altered. These thresholds are parameterized, and can be used as user controls. Between these thresholds, the intensities are adjusted nonlinearly.

The intensity corresponding to a given Cumulative Distribution Function (CDF) is used to determine whether the image should be darkened or lightened. If this intensity lies below a pre-defined low CDF threshold, the image is lightened to the maximum level. If this intensity lies above a pre-defined high CDF threshold, the image is darkened to the maximum level. When the intensity lies in between, the amount of lightening or darkening is computed by linear interpolation.

Figure 4.12 shows the CDF of the "Cave" image shown in Figure 4.11 that needed the maximum amount of lightening. For this image, the intensity corresponding to a CDF of 0.5 is only 0.107, which means 50% of the image pixels have intensities below 0.107.

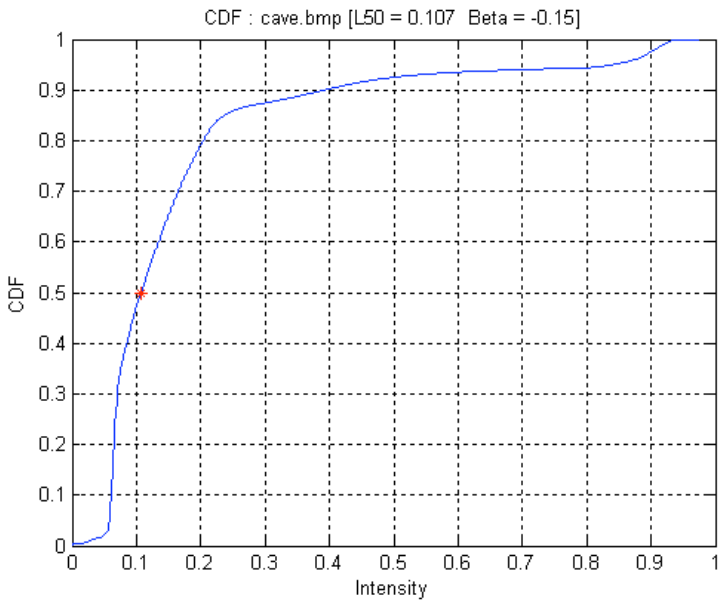


Fig. 4.12 Cumulative Distribution Function of the image "Cave"

Figure 4.13 shows an example image with normal dynamic range. Compared to the previous example, lightness variation in this image is more even. Figure 4.14 shows the CDF for this image. The intensity corresponding to a CDF of 0.5 is 0.56.

Fig. 4.13 “Faces”: an example image with normal dynamic range

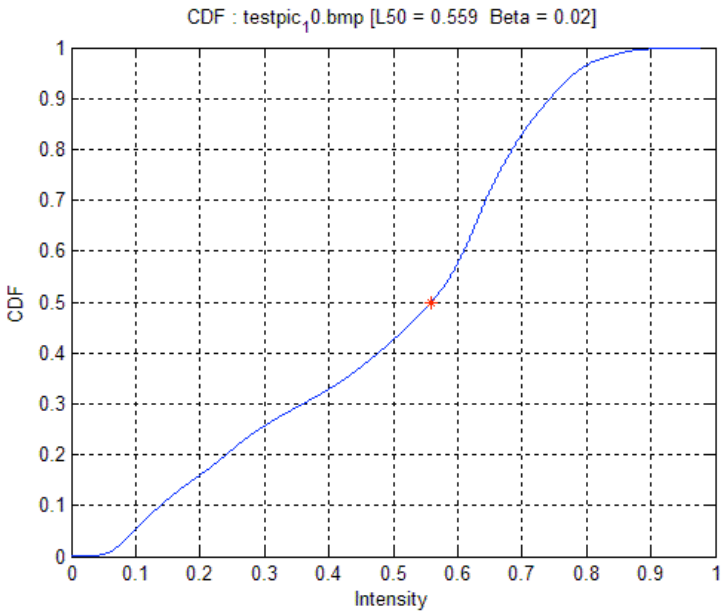


Fig. 4.14 Cumulative Distribution Function for the image “Faces”

An additional condition is imposed to ensure darkening or lightening does not cause some part of the image to be clipped. If a certain percent of pixel intensity in the CDF lies above a given value, it means a part of the image already has high intensity. In this case the image is not lightened. Similarly, if a certain percent of pixel intensity in the CDF lies below a given value, the image is not darkened.

4.4.2 Local Contrast Enhancement

A Low-Pass averaging filter is used to get a filtered version of the original intensity image. An exponent is applied to the difference of the original and the filtered intensity data while preserving the sign of the difference. The output intensity resulting from global lightness adjustment and local contrast enhancement is obtained from summing the two effects.

4.4.3 Saturation Enhancement

Similar to the global lightness adjustment, saturation enhancement is achieved through a Lookup Table. This avoids hard thresholds and the problems of artifacts associated with them. If the chroma of a pixel is below the low threshold, the color is likely to be achromatic and is left unaltered. Also, if the chroma of a pixel is above the high threshold, the color is already quite saturated and so is left unchanged. Colors in between these thresholds are increased in a nonlinear fashion.

If an image is significantly lightened, the perceived saturation reduces. Actual saturation enhancement in this case needs to be higher to counteract this effect. Whether an image is

significantly lightened was determined by a metric based on the CDF, and accordingly a higher saturation enhancement factor was used.

4.5 Novelty of the proposed method

Various published and patented methods reviewed in this chapter do not meet the objective of enhancing color and contrast in an effective and coordinated manner. The proposed algorithm addresses the need for a more complete color and contrast enhancement algorithm suitable for the video processing chain of consumer video systems. The new method is guided by the principles and foundations of color appearance and perception, and not by the statistical image processing. It takes advantage of a color space that is uniform in terms of perceived hue, thus ensuring the hue does not change during the processing, one of the main problems in color enhancement. A novel method for local contrast enhancement is also implemented in the algorithm.

"I have been impressed with the urgency of doing. Knowing is not enough; we must apply. Being willing is not enough; we must do." ~ Leonardo Da Vinci (Italian artist, sculptor, architect and engineer, 1452–1519)

Chapter 5

IMPLEMENTATION AND PERFORMANCE ANALYSIS OF SEVERAL COLOR/CONTRAST ENHANCEMENT ALGORITHMS

5.1. Algorithms Implemented

Four published methods for color and contrast enhancement were implemented to evaluate their performance as well as to determine the most appropriate enhancement strategy for the new algorithm. Further, two existing color and contrast enhancement algorithms were provided by the research sponsor, Intel Corp, to be used as benchmarks in the development of the new algorithm. Accordingly, this chapter contains a performance analysis of seven algorithms as listed below:

1. **Proposed:** New algorithm proposed in this research
2. **CH:** Algorithm provided by the sponsor
3. **Colantoni:** Colantoni, Bost and Trémeau [**Colantoni 2004**]
4. **Samadani:** Samadani and Li [**Samadani 2006**]
5. **Tao:** Tao and Asari [**Tao 2004**]
6. **Yang:** Yang and Rodriguez [**Yang 1996**]
7. **YO:** Algorithm provided by the sponsor

All implementations were carried out in Matlab.

5.1.1 Implementation of Proposed Algorithm

Implementation of the proposed algorithm has been discussed in detail in Section 4.4.

5.1.2 Algorithm CH

This was an Intel Proprietary algorithm, based on traditional multi-module approaches (i.e. cascade of methods to deal with color, contrast, and skin color).

5.1.3 Implementation of Colantoni's Algorithm

In this implementation, D65 white point was assumed. xyY image data were converted to polar coordinates for the ease of geometric computation. The distance of a given pixel (P) from the white point (W) was increased by 50%. If the new point fell outside the triangle defining the RGB gamut, it was clipped to the gamut triangle. The slope of the line PW was kept unchanged. After transforming polar coordinates to xyY, pixel Y values were reduced by 2.5% to avoid out-of-gamut colors. The computations were performed on a pixel-by-pixel basis. In the end, the modified xyY values were transformed to RGB.

The dominant wavelength in the above implementation was not directly used or computed. The implementation followed the final solution reported in the paper. Authors could not be contacted to verify if the present implementation correctly replicated their work.

5.1.4 Implementation of Samadani's Algorithm

To begin with, two tables are formed, one with the hue values, and the other with the luminance values, for the three primaries and three secondaries in the following order: red, magenta, blue, cyan, green, yellow, red. These are referred hereafter as principal colors. The hue and luminance values were generated using Eq 5.1, as described in the paper, and noting that hue $h = \tan^{-1}(c_2/c_1)$.

$$\begin{bmatrix} l \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0.3008 & 0.5859 & 0.1133 \\ 0.5117 & -0.4297 & -0.0820 \\ -0.1719 & -0.3398 & 0.5117 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix} \tag{5.1}$$

Next, the function α shown in Eq (5.2) is pre-computed for hue values of 0 through 360 degrees in steps of 0.1 degrees and stored for real-time processing. For each hue, α is computed for luminance values 0 through 1 in steps of 0.001.

$$s_0 = \frac{q(l_0,h_0)}{q(l_i,h_i)} s_i = \frac{q(f(l_i),h_i)}{q(l_i,h_i)} s_i \equiv \alpha(l_i,h_i)s_i \tag{5.2}$$

At first, the luminance l_a corresponding to maximum saturation for a give hue must be computed. For the hue (h_{in}) of any input color, the principal hues it is between are determined. Assuming ix and ix+1 are the corresponding indices in the hue table TH (as well as the luminance table TL), between which the input color lies, Eq (5.3) shows the equation to compute l_a.

$$l_a = (h_{in} - TH(ix)) * \frac{TL(ix+1) - TL(ix)}{TH(ix+1) - TH(ix)} + TL(ix) \tag{5.3}$$

Next, the luminance-saturation curves q(l) are computed for different l_a values. Since these curves are normalized with respect to saturation, s_a is always unity. In this implementation, the curves are formed by cubic interpolation passing through (0, 0), ($l_a/2$, $s_a/2$), (l_a, s_a), (l_a*1.5, s_a*0.75) and (1,0). The function α represents $q(l_0)/q(l_i)$, where l_i is the input luminance and l_0 is the output luminance, represented by the following equation:

$$l_0 = l_i + \beta \sin(\pi * l_i) \tag{5.4}$$

Where β is a user-controlled parameter (-0.15 in this implementation). A negative value resulted in reduction in the luminance. Output luminance values were also pre-computed. Function α is essentially a two-dimensional look-up table of the size (1001x3601).

For an input image, luminance and hue for each pixel are computed using Eq. (5.1). Output luminance and saturation are computed using Eq. (5.4) and Eq. (5.2) respectively. Note that input luminance and hue values were appropriately converted into integers to use as indices in the look-up table. Using sufficient resolution in the look-up table is important in obtaining proper enhancement. Finally, the output RGB values were computed using the inverse transformation of Eq. (5.1) and clipped if necessary.

5.1.5 Implementation of Tao's Algorithm

In Tao's algorithm, adaptive luminance enhancement results in dynamic range compression, which is not suitable for our application. So, adaptive luminance enhancement was excluded from the performance analysis described in Section 5.3.

For the adaptive contrast enhancement, a 5x5 gaussian filter was used for 2D convolution and then, Eq. (5.5) through Eq. (5.7) were used to computed the new intensities, as described in the paper. In Eq. (5.5), I_n' is the intensity resulting from adaptive luminance enhancement (original intensity in this implementation). E(x,y) is given by Eq. (5.6), where I_{conv} and I are filtered intensity and original intensity respectively. The standard deviation σ in Eq. (5.7) is obtained by multiplying the standard deviation of intensity by 255.

$$S(x,y) = 255 I'_n(x,y)^{E(x,y)} \tag{5.5}$$

$$E(x,y) = r(x,y)^P = \left[\frac{I_{CONV}(x,y)}{I(x,y)} \right]^P \tag{5.6}$$

$$P = \begin{cases} 3 & for \quad \sigma \le 3 \\ \dfrac{27 - 2\sigma}{7} & for \quad 3 < \sigma < 3 \\ 1 & for \quad \sigma \ge 10 \end{cases} \tag{5.7}$$

Final RGB values were obtained by multiplying original RGB values by the ratio of enhanced and original intensity computed on a pixel-by-pixel basis.

5.1.6 Implementation of Yang's Algorithm

First, luminance was computed from RGB values (Eq. 5.8). The parameter β was computed using Eq. (5.9), where k is a scale factor by which the saturation is enhanced. In this implementation, the value of k was set to 1.35.

$$L = 0.299R + 0.587G + 0.114B \tag{5.8}$$

$$\beta = \frac{3Lk}{(R+G+B)(1-k) + 3Lk} \tag{5.9}$$

Finally, output RGB values were computed using Eq. (5.10).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \beta \begin{bmatrix} R \\ G \\ B \end{bmatrix} + (1-\beta) \begin{bmatrix} L \\ L \\ L \end{bmatrix} \tag{5.10}$$

So, in fact, H and S values were not computed at all. However, transformation from RGB to LHS has been explained in the paper in detail. The enhancement process described above is in accordance with the paper.

5.1.7 Algorithm YO

This was an Intel Proprietary algorithm, based on traditional multi-module approaches (i.e. cascade of methods to deal with color, contrast, and skin color).

5.2. Images Used in the Analysis

Seven images were chosen for this analysis. Each of these images was processed by the seven algorithms. Input and output images were then further analyzed. The seven images and their various features are described in Table 5.1.

Table 5.1 Seven images used in the performance analysis

Avia

This is one frame from the Avia video test sequence. The sequence was used in a psychophysical test as described in the next chapter. It contains a close up shot of a woman's face, which is the main element in this image. A color checker and various grey patches are also present. The background, which constitutes a major part of the image, is completely black.

Carnival

This is a night scene with considerable noise in the original image. The scene has a high dynamic range because of the presence of some bright spotlights. The red and yellow light emitting diodes (LED) give rise to saturated colors. Overall, this is a challenging image.

Chinatown

This is a scene from an old Chinatown area, showing brick houses with worn-out colors, but also with some colorful signboards containing text. The brick pattern, the signs, the dark blue sky are all critical elements in the image. The interior of the buildings is in shadow, and it is important to preserve the shadow details, and even to enhance them.

Table 5.1 (contd.)

Couple

This is the last frame in the couple sequence described in the next chapter. The bright yellow shirt, the tie, the bright pink skirt, and the design on the woman's shirt are important features in this image. Also, there is green foliage in the background that need proper enhancement. Despite above features being important in this image, the most important element seems to be the faces and the skin tone. The hands and the faces look a little blurred due to motion.

Dome

The most prominent feature is the architectural details and the intricate sculpture on the walls. The other critical elements are the bright light spots above the torchieres, light and shadow on the floor and the dresses of the people at distance. It is a very good test image for contrast enhancement.

Faces

The faces of the man and the woman are the most important aspects here. Skin tone, the teeth and the hair on woman's forehead are critical in terms of color/contrast enhancement. Enhancing any imperfections in the skin may be judged negatively in terms of image quality.

Veggies

This image has several variations of color and contrast. While the red color of the tomatoes is the most prominent in the image, the broccoli, the corn and the cauliflower have many contrast details. There is a slight yellow tinge on the cauliflower, which must not be enhanced to the extent that the cauliflower looks objectionably yellow. The grey background in the image must also remain achromatic after the enhancement.

5.3. Performance Analysis

As described earlier, each of these images was processed by the seven algorithms. The input and output image data were converted to IPT color space to derive various image metrics. These metrics were:

i) Cumulative Distribution Function (CDF),

ii) Histogram of the intensity image (in IPT space)

iii) Image difference in all three dimensions, lightness (ΔJ), chroma (ΔC) and hue (Δh)

IPT color space was chosen since it is perceptually more uniform in terms of hue than other color spaces like CIELAB. Any color/contrast enhancement algorithm should not change the original hue to the extent that it is perceivable. Thus, hue difference map between the input and output image is a useful performance evaluation metric for these algorithms. The lightness difference map shows the effectiveness of contrast enhancement while the chroma difference map can be used to evaluate the color enhancement functionality.

In the following subsections, performance of all seven algorithms has been discussed on a case-by-case basis for each individual image. The lightness difference maps as well as the contour maps for chroma and hue differences between input and output are included in this chapter. To further aid the discussion, image gamut plots in CIELAB color space, cumulative distribution function plots, histogram plots and pseudo color plots of chroma and hue differences have been included in Appendix A. For chroma difference, absolute values are normalized to unity, while the hue differences are computed in degrees (between 0° and 360°).

All image difference maps include the 50^{th} and the 90^{th} percentile values of the image difference data, as well as the minimum and the maximum values. The hue difference map was obtained from a 5x5 low-pass filtered version of original hue difference data. The reason was to identify regions of significant hue difference and ignore individual pixel differences. For the same reason, the percentiles, the minimum and the maximum of the hue difference values are also computed from the filtered data. It should be noted that the perceptibility of hue difference largely depends on the corresponding lightness. The same hue difference might be imperceptible for darker pixels, yet very noticeable for lighter pixels.

It must be emphasized that none of the metrics or the image difference maps is a measure of the perceived image quality as determined by human observers. Only a properly designed psychophysical experiment can provide such information. The performance analysis presented in this chapter serves the sole purpose of demonstrating different functionality aspects of various algorithms. Results obtained from psychophysical experiments have been described in detail in the next chapter.

5.3.1 Test Image "Avia"

The most important color in this image is the skin tone, and to lesser extent the colorchecker patches. If we compare the input and output image gamuts in L*a*b* space, shown in Figure A.1, the gamut expansion is less apparent in case of the Samadani and Tao, but more obvious in case of others. Pure colors like green, blue and yellow, belonging to the colorchecker, mainly get enhanced (more apparent in the new algorithm). There is no significant change in the Cumulative Distribution Function (CDF) for the lightness (Figure A.2) and the histograms (Figure A.3), except in case of Colantoni, which shows significant clipping of darker pixels.

To analyze the lightness difference map shown in Figure 5.1, it is helpful to consider the 50^{th} and 90^{th} percentile values of the image difference data. Evidently, only algorithm CH increases the lightness significantly. As explained earlier, current implementation of Samadani's algorithm reduces the lightness in the image. Same effect can be observed in case of Colantoni's algorithm. While average lightness does not change appreciably, edge enhancement is clearly noticeable in case of the proposed algorithm, and to a lesser extent, in Tao's algorithm. Interestingly, for Yang and YO, the noise in the background, probably present in the original image itself due to compression, is enhanced. However, because of very low lightness, they cannot actually be perceived in the enhanced image.

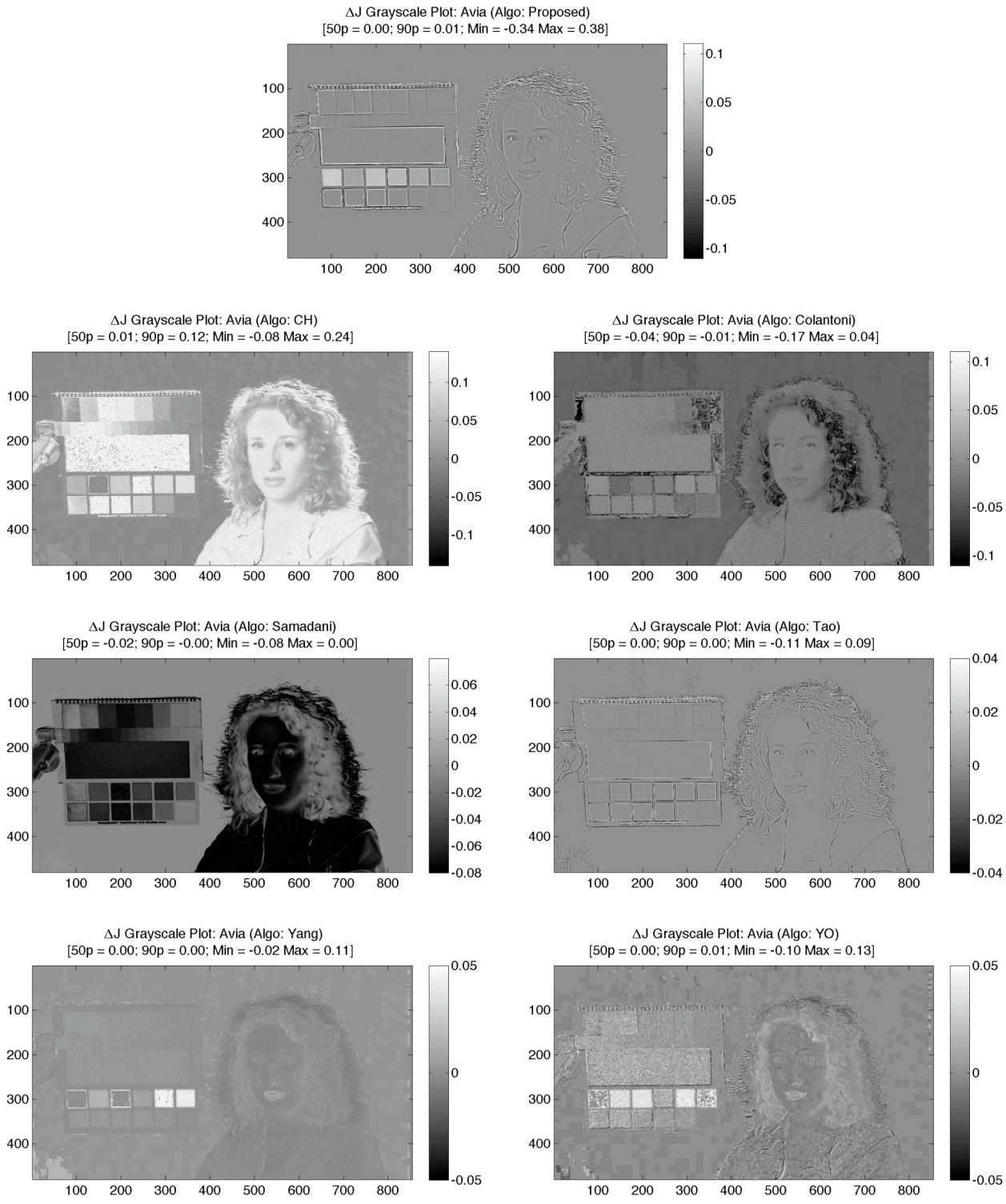


Fig. 5.1 ΔJ Image Difference Maps: Avia

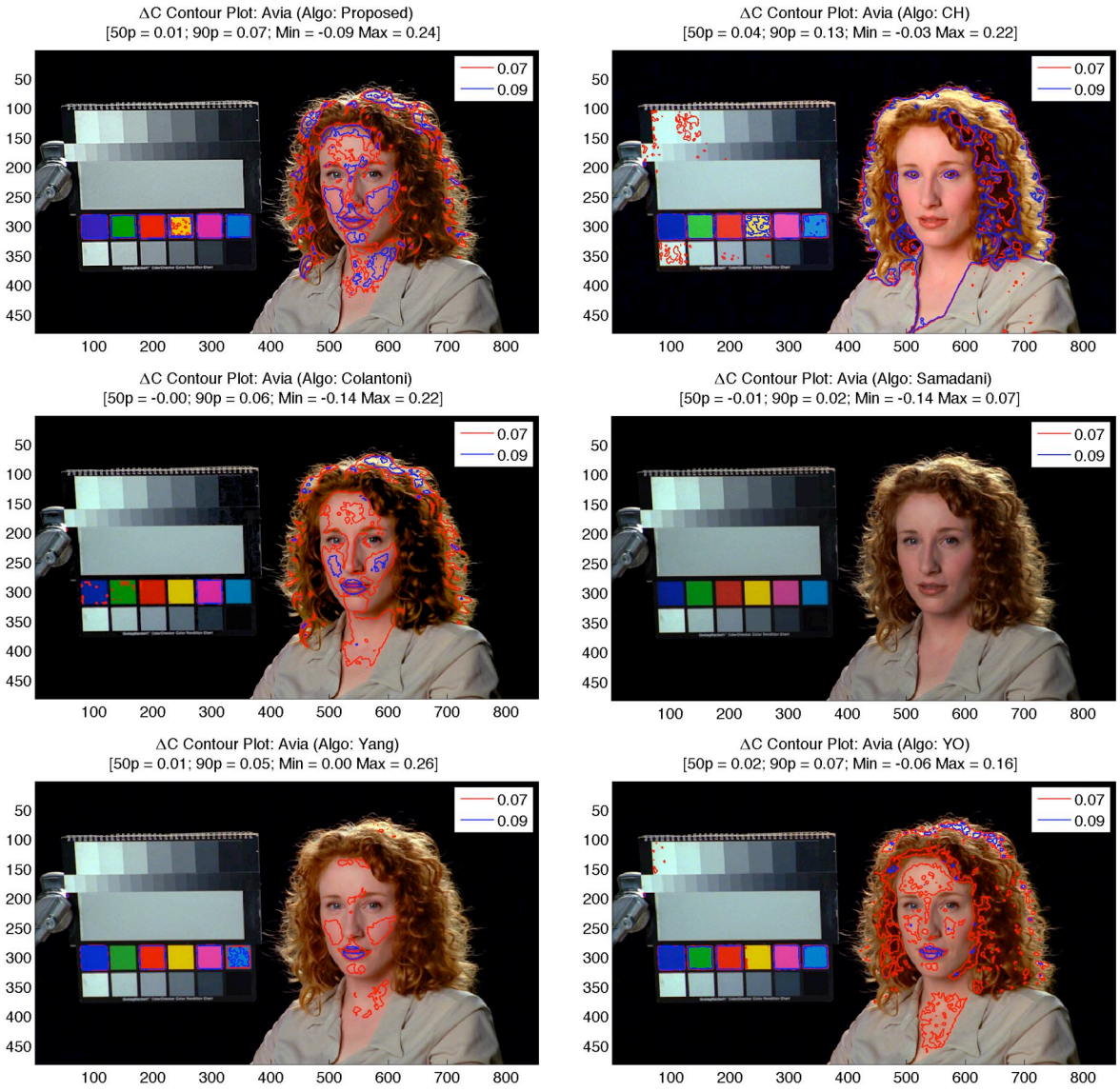


Fig. 5.2 ΔC Contour Maps: Avia

The contours for the chroma difference, shown in Figure 5.2, have been plotted for the same two levels for each algorithm. To make comparison easier, first the appropriate levels were determined for the new algorithm, and then the same levels were used for the other algorithms. Relative degrees of chroma enhancement can be judged by the percentile as well as the minimum and the maximum values in each case. As expected, in most cases, the main chroma

enhancement occurs in the face and the colorchecker patches. Samadani's algorithm does not make a significant change in the chroma. While the other algorithms selectively enhance the skin tone, the algorithm CH enhances the chroma over the whole face and neck. Further, the enhancement is significantly stronger than the other cases, and even changes the chroma for some gray patches. Such strong enhancement may not always be desirable. The chroma difference pseudo-color plot in Figure A.4 shows that the new algorithm does a good job in leaving the achromatic colors, including the gray patches and the black background, unaltered.

Hue difference contour plots, shown in Figure 5.3, are plotted for the same two levels for each algorithm. In case of the new algorithm, most of the hue change occurs at the edges where colors are changing. This is likely to happen whenever an algorithm employs low-pass filtering for enhancement purposes, since neighboring pixels to some extent affect the color of a given pixel. This hue change is unlikely to be noticed in most situations, as is the case here. Some of the colorchecker patches have undergone slight hue shift, but again, for low lightness patches they are hardly noticeable. Both CH and YO show significant hue shift on the woman's face and on several gray patches. In most of the cases, the maximum hue shift likely occurs in the black background region and thus is not critical. 90^{th} percentile value is more informative in this regard. Colantoni's algorithm output shows a serious hue shift problem, which is not surprising (explained in Chapter 4). Pseudo-color plots of hue difference (Figure A.5) support the above observations.

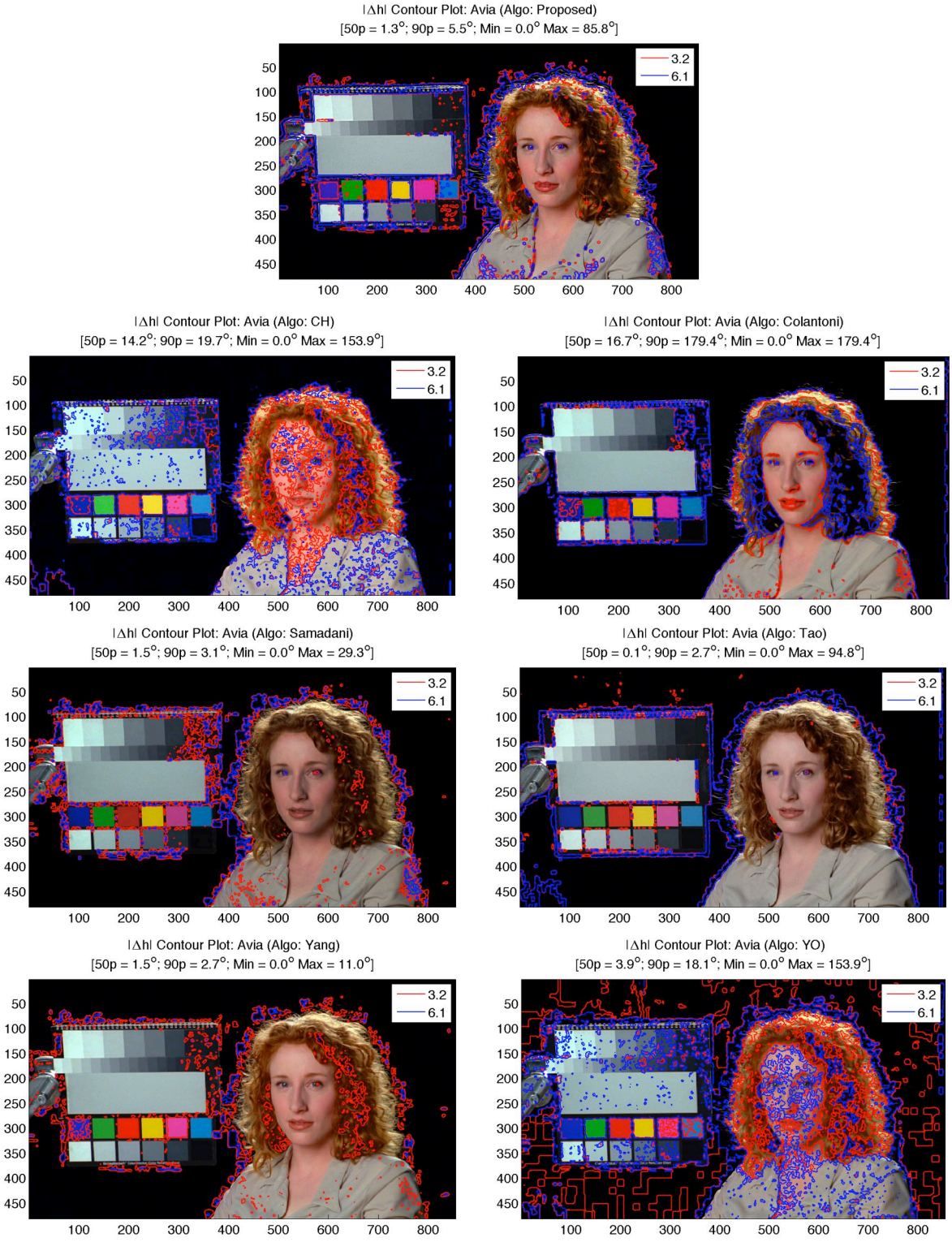


Fig. 5.3 Δh Contour Maps: Avia

5.3.2 Test Image "Carnival"

This is a night scene, and thus is mostly dark. Most of the colors in the image are due to the neon lighting. The input image contains significant amount of noise, and thus poses a challenge to the contrast enhancement algorithm in absence of a noise suppression module in the processing pipeline. The image gamut plots in Figure A.6 show that the gamut expansion is more skewed toward green and yellow. Only algorithms that show some modification of the intensity CDF (Figure A.7) are CH, Colantoni and Samadani. As for the intensity histogram, there is no significant change, except for a modest difference in its shape in case of Colantoni's output. The lightness difference map in Figure 5.4 shows the enhancement is the most significant in case of CH, the lighter regions in the image are nicely enhanced while the darker areas are further darkened, effectively enhancing the dynamic range. Edge enhancement is more apparent in case of the new algorithm and Tao's algorithm.

The contour plots for the chroma difference in Figure 5.5 demonstrate that chroma enhancement is not significant, and mainly occurs around the light sources (lighter pixels). For YO, however, the enhancement occur elsewhere in the image as well. These are also apparent in the chroma difference pseudocolor plots (Figure A.9).

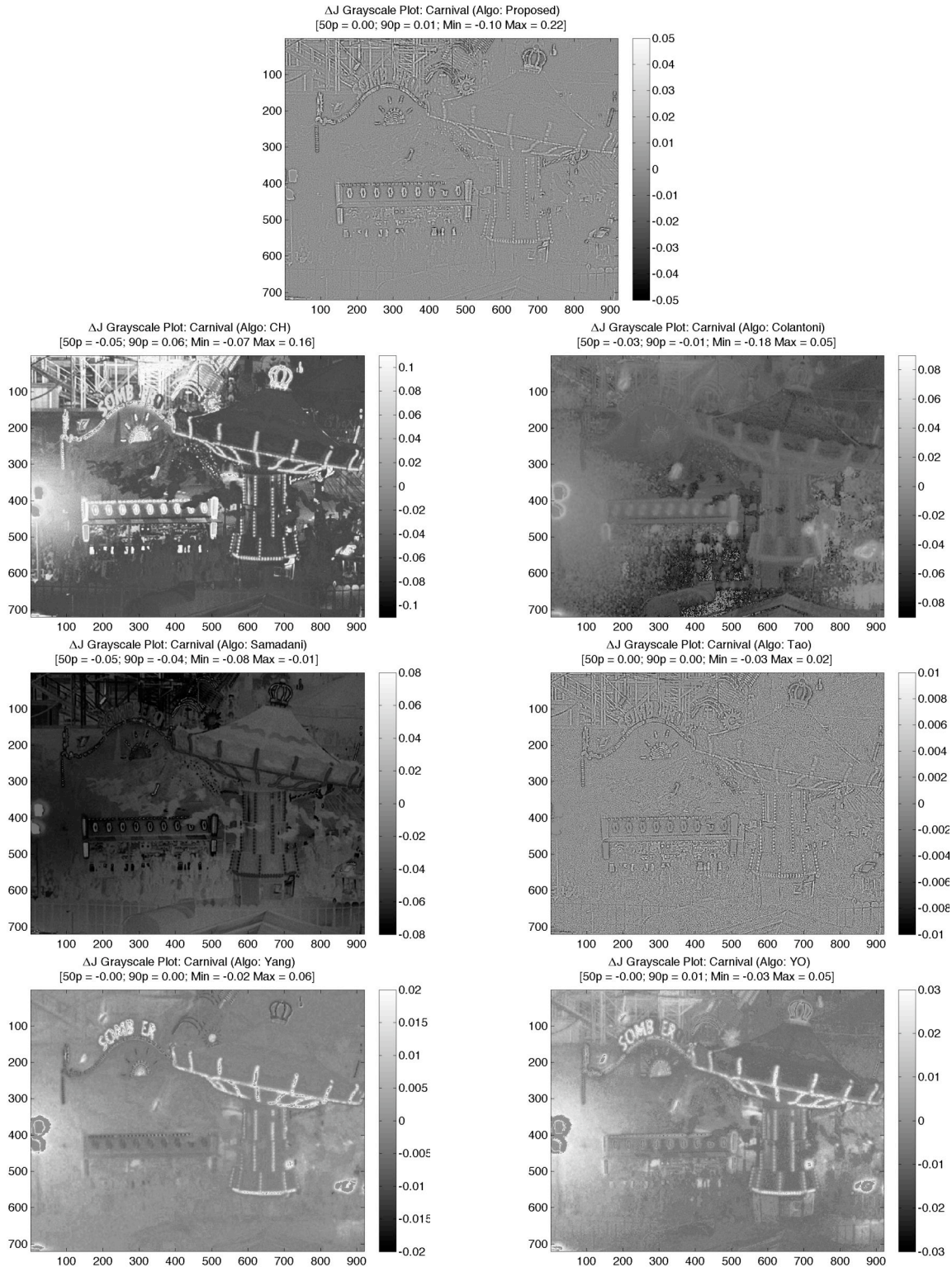


Fig. 5.4 ΔJ Image Difference Maps: Carnival

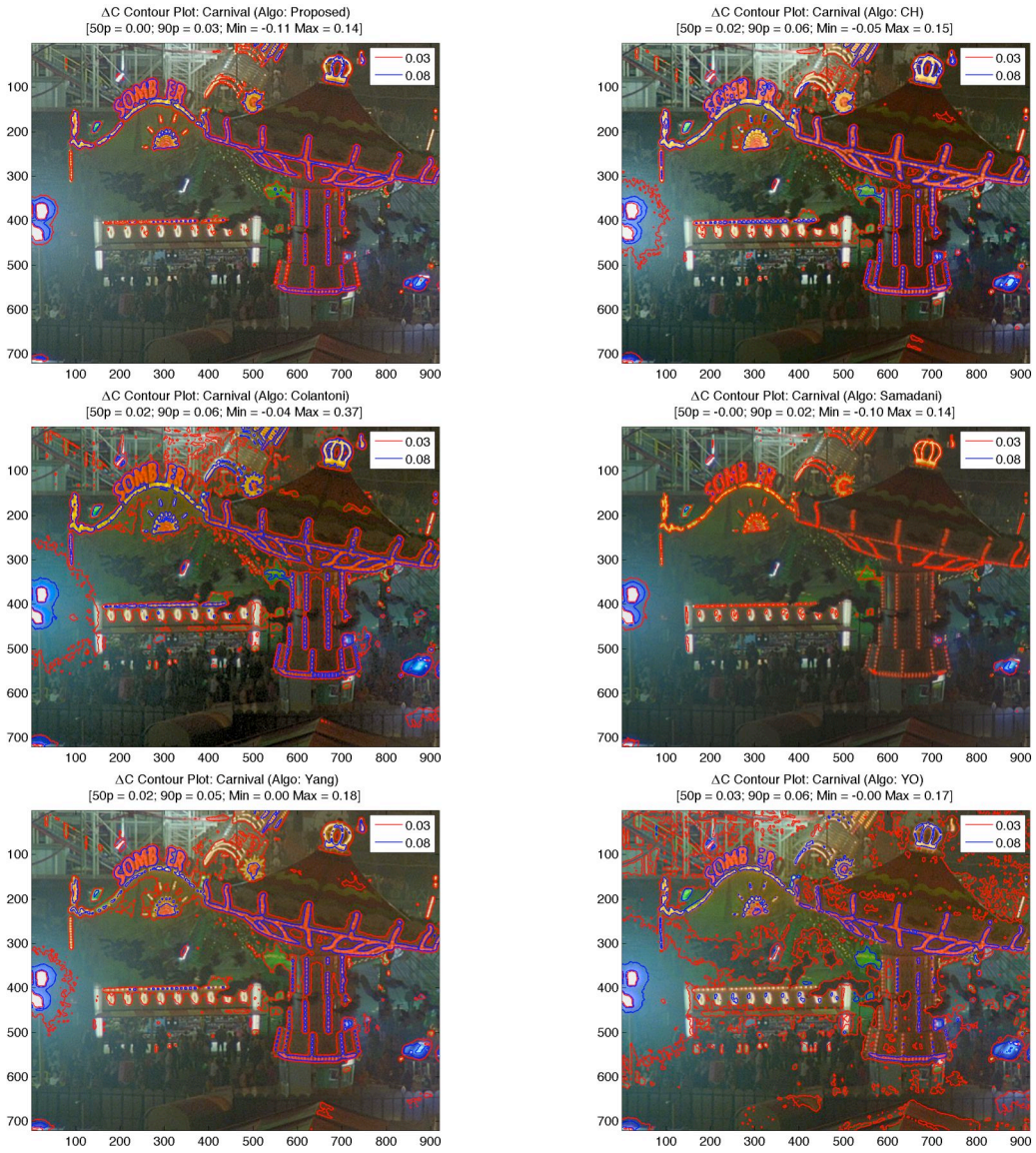


Fig. 5.5 ΔC Contour Maps: Carnival

The hue shift is predominant in YO, as shown in the hue difference contour plots in Figure 5.6, with the 90th percentile value of 6.2°. However, since the majority of the pixels are dark, the hue shift may not be very perceivable. Note that Tao's algorithm implementation does not have a color enhancement component, so the output does not show any hue difference. Hue Difference pseudocolor plots in Figure A.10 illustrate that the hue shift present in almost all outputs is high frequency and of low magnitude, and thus cannot be perceived.

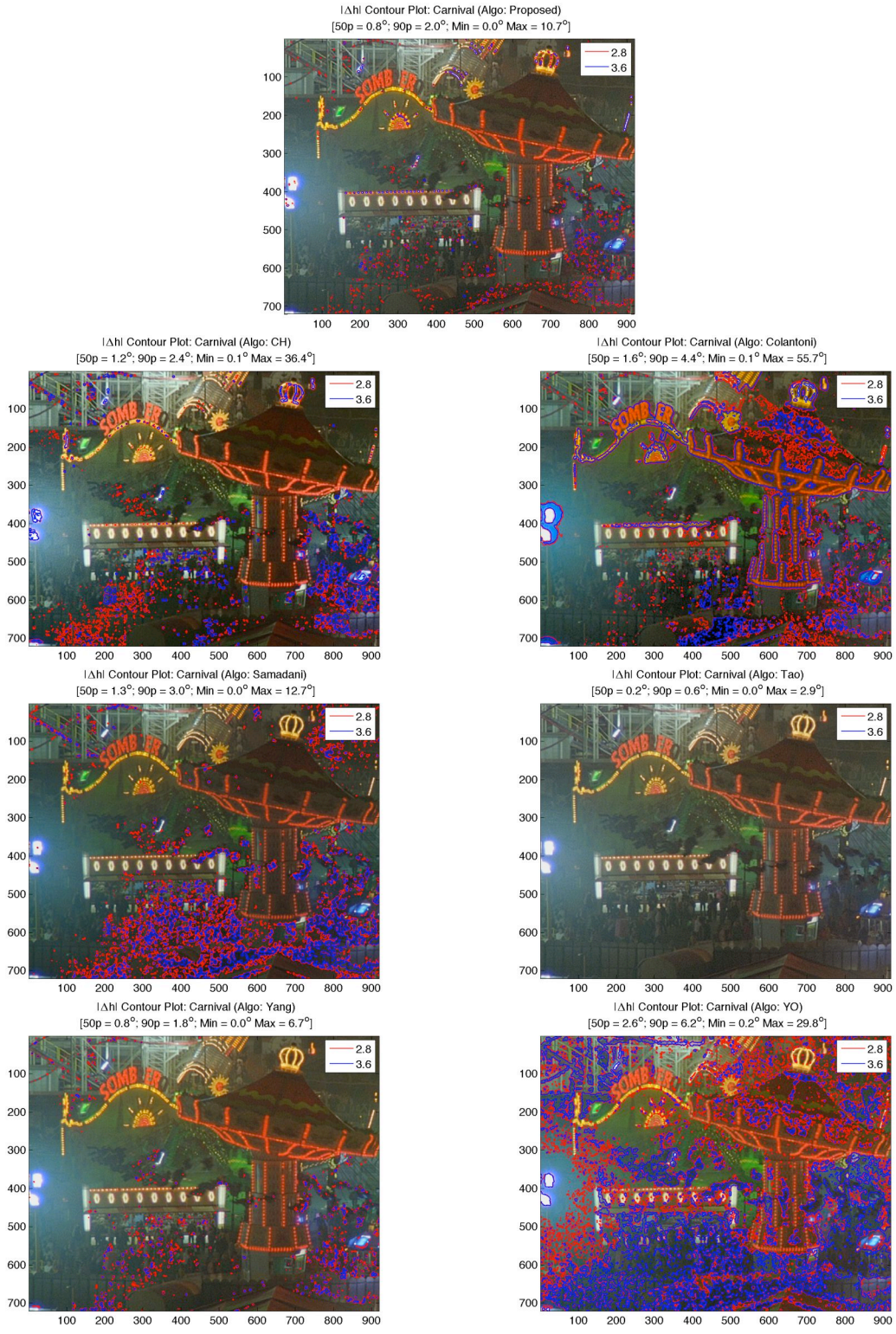


Fig. 5.6 Δh Contour Maps: Carnival

5.3.3 Test Image "Chinatown"

This image does not contain many colors, as apparent from the small image gamut (Figure A.11). The expansion in the image gamut is more apparent in case of CH algorithm. From the CDF plots in Figure A.12, Colantoni's algorithm causes 20% of the total pixels clip to zero intensity. This happens when the saturation of the darker pixels is increased while keeping the lightness same. The CDF of the proposed algorithms does not reach to the maximum level until the intensity reaches the maximum, which indicates there is a possibility of clipping for some high intensity pixels. In the intensity range of 0.2 to 0.6, Samadani's algorithm increases the intensity levels more than the others, while CH reduces the intensity in the 0.4 to 0.95 range.

The histograms in Figure A.13 further illustrate how different algorithms are achieving different lightness adjustments. Relatively high value at the upper extreme indicates the proposed algorithm is causing some clipping as mentioned before. The algorithm also reduces the peak at the intensity of 0.45 and increases the number of low-intensity pixels. Colantoni's algorithm is markedly different from the others as it completely shaves off the peak at low intensity level, but the rest of the histogram does not change significantly. This is because many pixels are clipped to the black level.

Figure 5.7 shows the lightness difference grayscale plots. For the proposed algorithm and Tao's, most of the lightness enhancement occurs along the edges. Colantoni and Samadani's algorithm mostly reduce the lightness. For YO, the lightness enhancement is quite controlled, and the local contrast enhancement has performed quite well.

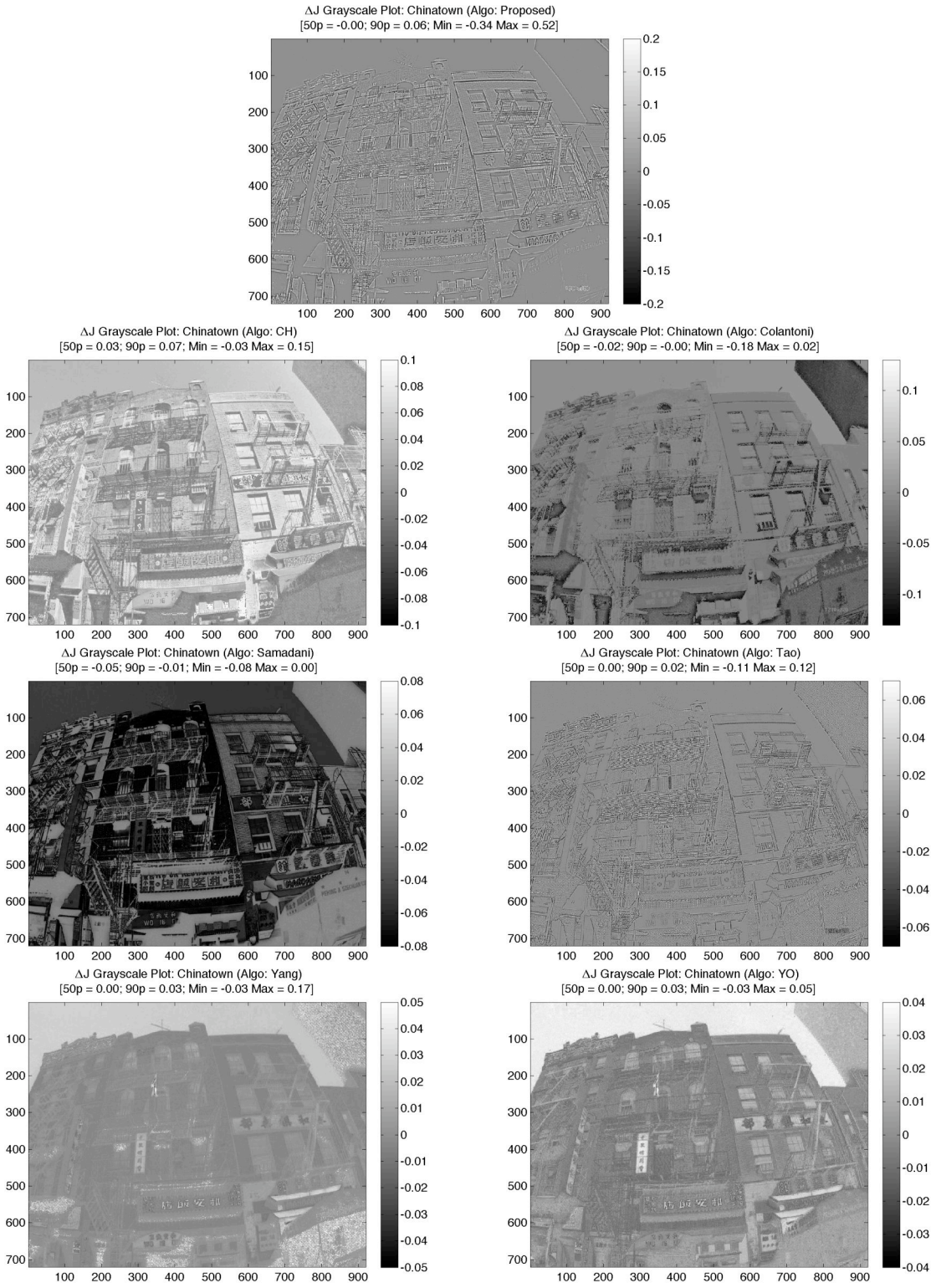


Fig. 5.7 ΔJ Image Difference Maps: Chinatown

Fig. 5.8 ΔC Contour Maps: Chinatown

As apparent in the chroma difference contour plots (Figure 5.8) and pseudocolor plots, both CH and YO enhance chroma significantly more than the other algorithms, with 90 percentile values of 0.15 and 0.11 respectively. Comparatively, the proposed algorithm enhances mainly the sky and the colored signs and hoardings. Chroma enhancement is not significant in case of Samadani. In case of Yang, selective colors are strongly enhanced.

Even though it appears from the 90th percentile values in Figure 5.9 and Figure A.15 that the proposed algorithm is causing significant hue shift, the 50th percentile value is lower than those in case of CH, Colantoni and YO and comparable to Samadani and Yang. The hue shift in this case is mostly concentrated in the dark areas of the image, and thus cannot be perceived. Interestingly, in case of Samadani and Yang, the sky undergoes some amount of hue shift. Tao's algorithm implementation does not involve color enhancement, yet there is a considerable amount of hue shift. But as in case of proposed algorithm, the hue shift cannot be perceived. From the hue difference contour plots (Figure 5.9), YO seems to cause more noticeable hue shift than the other algorithms.

Fig. 5.9 Δh Contour Maps: Chinatown

5.3.4 Test Image "Couple"

The image gamuts of input and outputs in L*a*b* space are shown in Figure A.16. The gamut expansion is not so apparent in case of Samadani and Tao, but observable in case of other algorithms. If we consider the intensity CDFs (Figure A.17), the proposed algorithm leaves it unaltered, while CH increases the intensity of pixels in the range 0.15 to 0.9, thereby significantly shifting the CDF to the right. As in previous examples, Colantoni's algorithm causes significant clipping at the lower end, 38% of the pixels getting clipped to black. Samadani's algorithm reduces the intensity for the most part, while YO slightly increases the intensity values between 0.2 and 0.8. In this example, the algorithm CH modifies the intensity histogram (Figure A.18) in a way similar to histogram equalization. Like the previous example, Colantoni's algorithm results in significant clipping and a complete shaving off of the peak.

The lightness difference maps in Figure 5.10 shows superior edge enhancement achieved by the proposed algorithm (to some extent by Tao's method as well), but otherwise the lightness enhancement seems modest. CH shows significant increase in the lightness in parts of the image, much more than the others. On the other hand, Samadani's algorithm implementation reduces the lightness in some key areas of the image. Yang's enhancement appears more selective than the others.

Fig. 5.10 ΔJ Image Difference Maps: Couple

Fig. 5.11 ΔC Contour Maps: Couple

The chroma difference contour plots above show that the proposed algorithm, YO and Colantoni's algorithms mainly enhance the colors on the dresses. However, chroma enhancement is the strongest in case of CH, and more widespread than in other cases. Even the grass and skin tone undergo strong enhancement, unlike the others where the grass and the skin tone on the faces are left unaltered. Yang's algorithm enhances chroma only on the colorful parts of the skirt

and tie. The corresponding pseudocolor plots in Figure A.19 highlight the same enhancement characteristics.

In the hue difference contour plots shown in Figure 5.12, the hue shift is least in case of Tao and Yang's algorithm. Note that Tao's algorithm implementation has no color enhancement, so hue shift problem is not expected as well. In case of CH and YO, hue shift is quite apparent. However, as shown in Figure A.20, the hue difference in case of most of the algorithms, including the proposed algorithm, is high frequency, with a majority of the change concentrated around dark pixels. As a result, the hue shift is not perceptually significant in general. Nevertheless, the plots identify the problems associated with color processing in various methods. As indicated by the 90^{th} percentile value of 1.6, the hue shift is insignificant in the proposed algorithm. Note that the maximum difference is insignificant as it most likely is located in the dark area around the edges of the tree trunks (Figure A.20). This shift arises from the fact that in low pass filtering, neighboring pixels affect the color of a given pixel.

Fig. 5.12 Δh Contour Maps: Couple

5.3.5 Test Image "Dome"

This image has limited variation in colors, mostly red and green with different lightness values, as shown in the image gamut plots in Figure A.21. As explained in previous examples, Samadani and Tao's method does not significantly increase the gamut volume. For most part of the intensity range, CH increases the intensity, as evident in the CDF plots (Figure A.22), while Samadani's method makes the image slightly darker. Once again, Colnatoni's method causes clipping of dark pixels. CDF remains mostly unaltered in all other cases.

The fact that CH increases the overall image contrast can be verified by examining the intensity histograms (Figure A.23). The histogram is moderately stretched in the upper range, signifying more high intensity pixels in the output than in the input. As in previous examples, Colantoni's method drastically reduces the peak, which is due to clipping. Even though the local contrast enhancement by the proposed algorithm is quite significant, histogram shape is more or less the same, implying that overall lightness has not changed very significantly.

The lightness difference plot in Figure 5.13 once again shows that the proposed algorithm is enhancing the edges more than the overall contrast. In other words, local contrast enhancement has a stronger influence than the global lightness adjustment. However, these effects can be adjusted by modifying appropriate parameters. Overall, increase in the lightness level is the highest for CH.

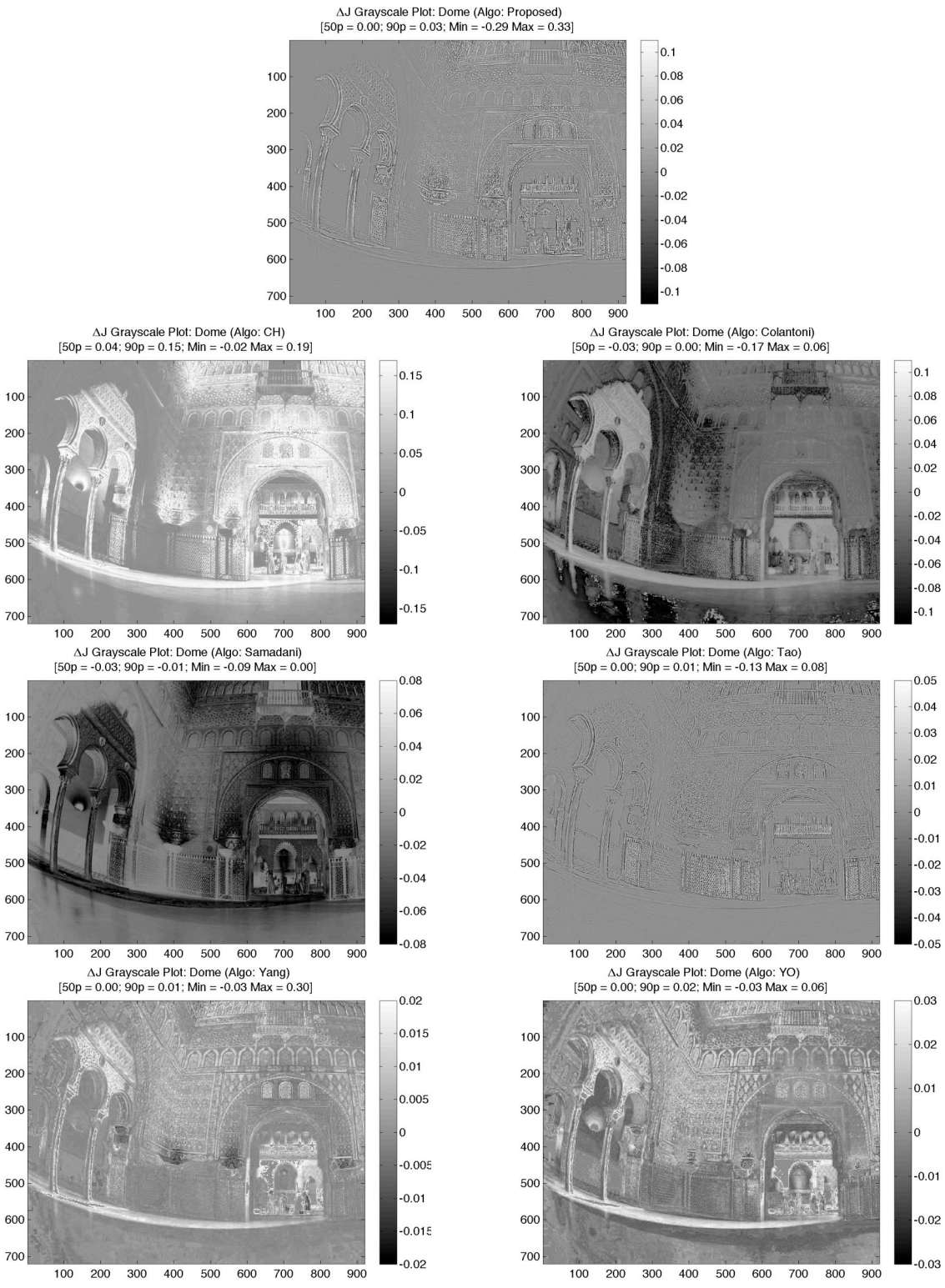


Fig. 5.13 ΔJ Image Difference Maps: Dome

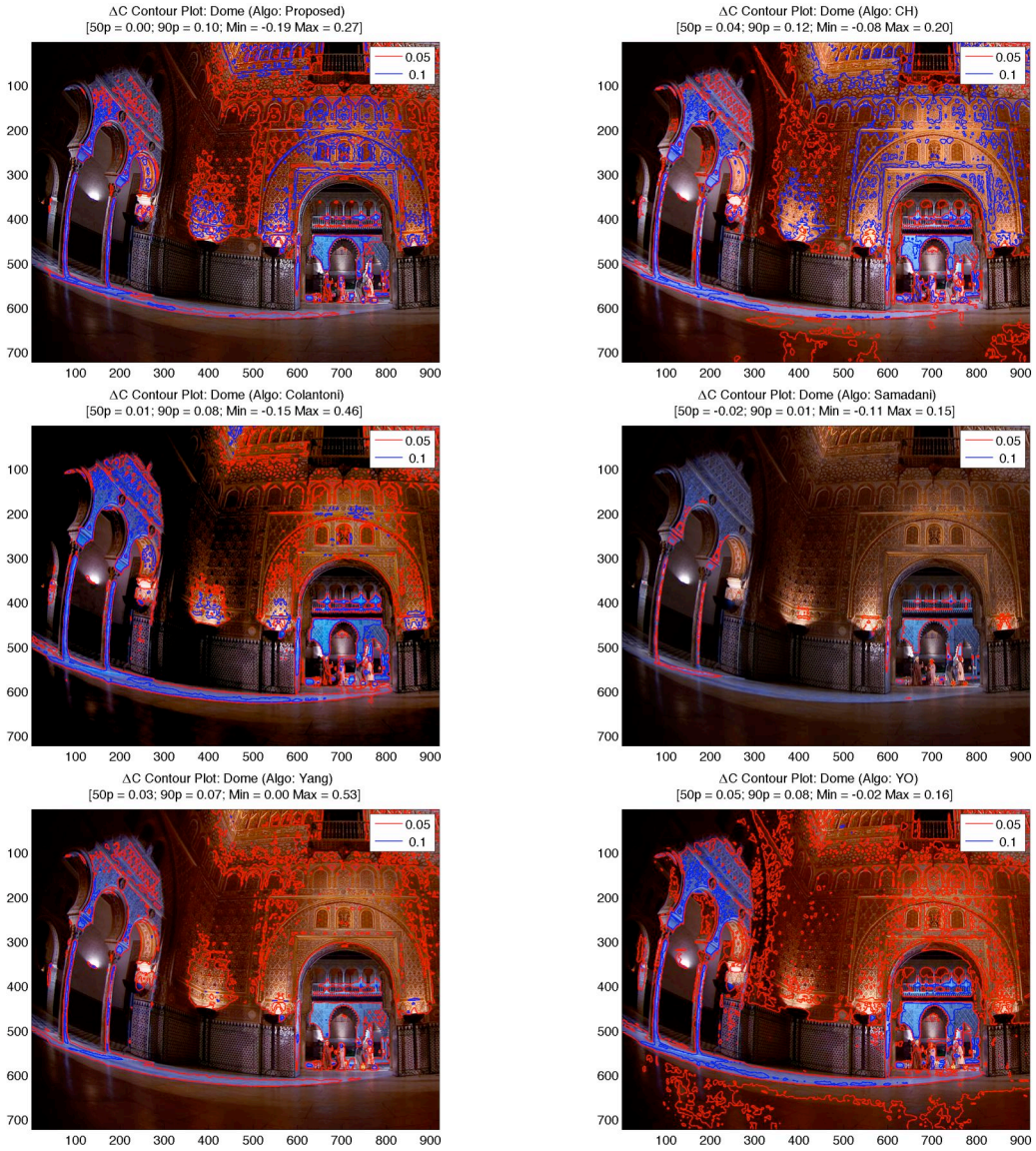


Fig. 5.14 ΔC Contour Maps: Dome

Chroma contour plots (Figure 5.14) and difference maps (Figure A.24) show that the enhancement is the strongest for CH and the proposed algorithm, and the weakest for Samadani. Interestingly, in case of Samadani's algorithm, chroma enhancement is mainly concentrated in the hotspot areas right above the uplights, which shows the chroma enhancement primarily results from the lightness adjustment.

Colantoni's method results in significant hue shift, with a 90th percentile value of 46.6° as shown in the hue difference contour plots (Figure 5.15) and the pseudocolor maps (Figure A.25). However, higher hue difference values are concentrated around the left arch that is under dark, and thus is not very perceivable. The same is true for most of the other cases. As we can see in Figure 5.15, most of the contours are located in the dark areas of the image. Compared to other algorithms, hue differences for YO are more spread out across the image. Color enhancement in new algorithm is unlikely to introduce perceivable hue shift as all processing is done in a color space uniform in terms of perceived hue. A 90th percentile value of 6.9° indicates a good performance for this image, as most of the colors have low lightness. In summary, lightness must be taken into account while evaluating perceived hue shift in a processed image.

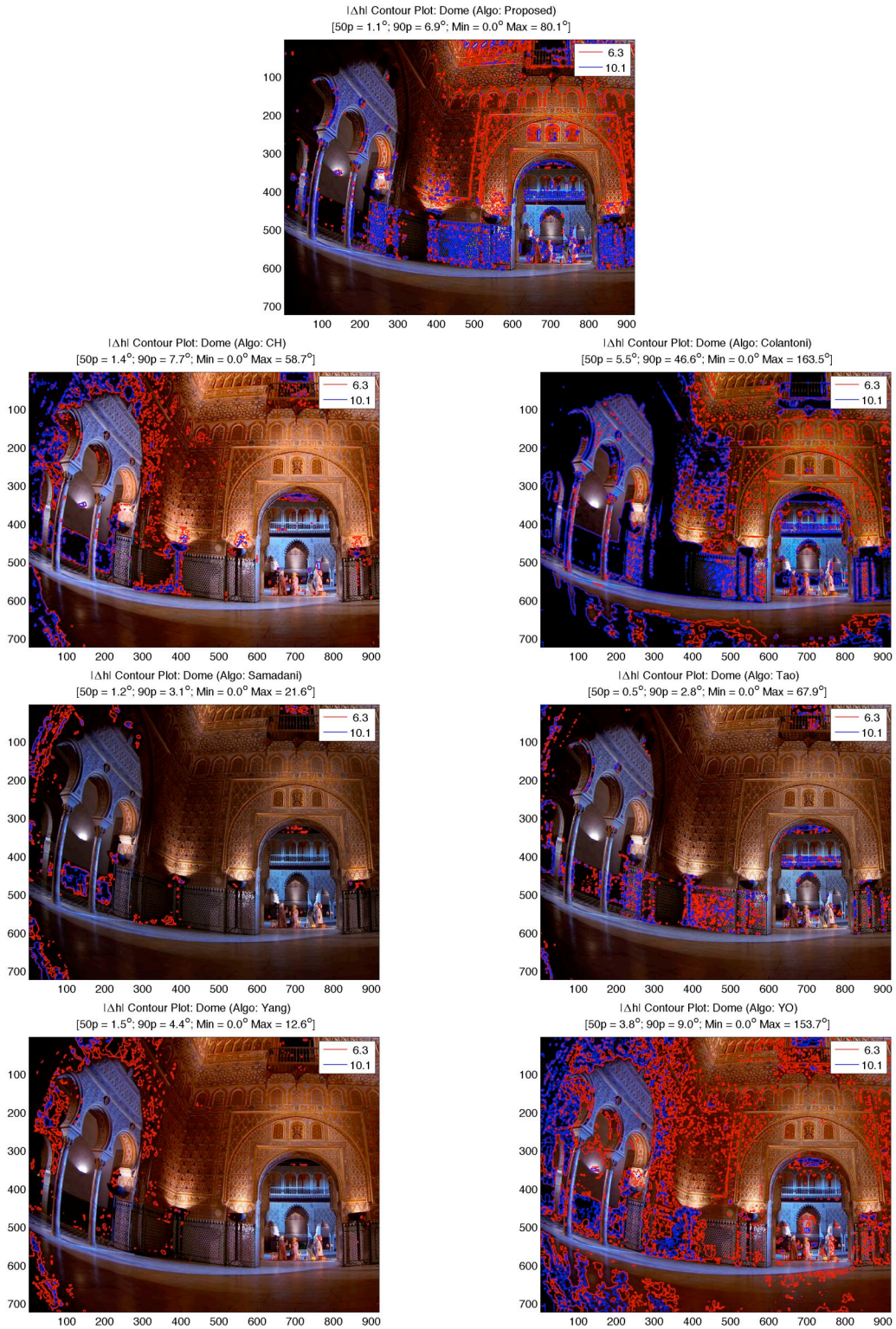


Fig. 5.15 Δh Contour Maps: Dome

5.3.6 Test Image "Faces"

Skin tone is the predominant in this image, and to some extent, blue color of the denim shirt (Figure A.26). The CDF plot (Figure A.27) shows that CH mostly increases the intensity while Samadani's method does the opposite. As for Colantoni's method, the problem of clipping is less in this image than in previous examples. For other methods, CDF remains mostly unaltered. A well spread-out input histogram (Figure A.28) indicates good overall contrast in this image, so there is not much room to enhance the intensity. Accordingly, the output histograms preserve the shape of the input.

The lightness grayscale plots in Figure 5.16 show similar trend as described in the previous examples. Chroma enhancement is more spread out in case of CH, almost covering full faces. 90^{th} percentile value is 0.14, strongest of all methods. For skin tone, this seems to be too strong an enhancement. Chroma enhancement is rather controlled in case of YO, and to a lesser extent, in case of the proposed algorithm Ho. In case of Samadani, only the lips of the woman undergoes noticeable enhancement.

Even though the magnitude of hue difference in this image is less than that in the other example images (Figure 5.18 and Figure A.30), a small hue shift will be more noticeable here because of higher overall lightness, and also because the observers are generally less tolerant about hue shift in skin tones. It is most objectionable in case of Colantoni's method. Even CH and YO suffers from a hue shift in critical areas. This underlines the importance of using an appropriate color space for image enhancement. By far, the new algorithm and Samadani's method do a superior job of preserving hue. The maximum values occur in dark areas and so are not perceivable.

Fig. 5.16 ΔJ Image Difference Maps: Faces

Fig. 5.17 ΔC Contour Maps: Faces

Tao's algorithm implementation does not involve color enhancement, so it is expected to have minimal hue shift problem.

Fig. 5.18 Δh Contour Maps: Faces

5.3.7 Test Image "Veggies"

The input image has a good overall contrast as evident in the histograms (Figure A.33), thus it is mostly local contrast, and not the global lightness, that needs to be enhanced. Lightness difference map in Figure 5.19 shows little increase in the overall lightness. As seen in case of other test images, CH boosts the chroma quite significantly (Figure 5.20 and Figure A.34). For example, the cauliflower appears more yellowish than in other cases. For Samadani's method, the chroma enhancement is rather modest and selective, as expected.

As in the previous examples, CH, YO and Colantoni's method show high hue differences with 90^{th} percentile values of 9.5°, 8.3° and 6.3° respectively. All three methods result in a perceivable hue shift on the cauliflower (Figure 5.21 and Figure A.35). All methods lead to a hue shift in the gray background, to one extent or the other. As explained before, implementation of Tao's method does not include color enhancement, so it can be expected to be mostly free of any hue shift. Hue shift problem is noticeably less in case of Samadani's method. For the proposed algorithm and Yang's method, only the achromatic background undergoes a hue shift that may be somewhat perceivable.

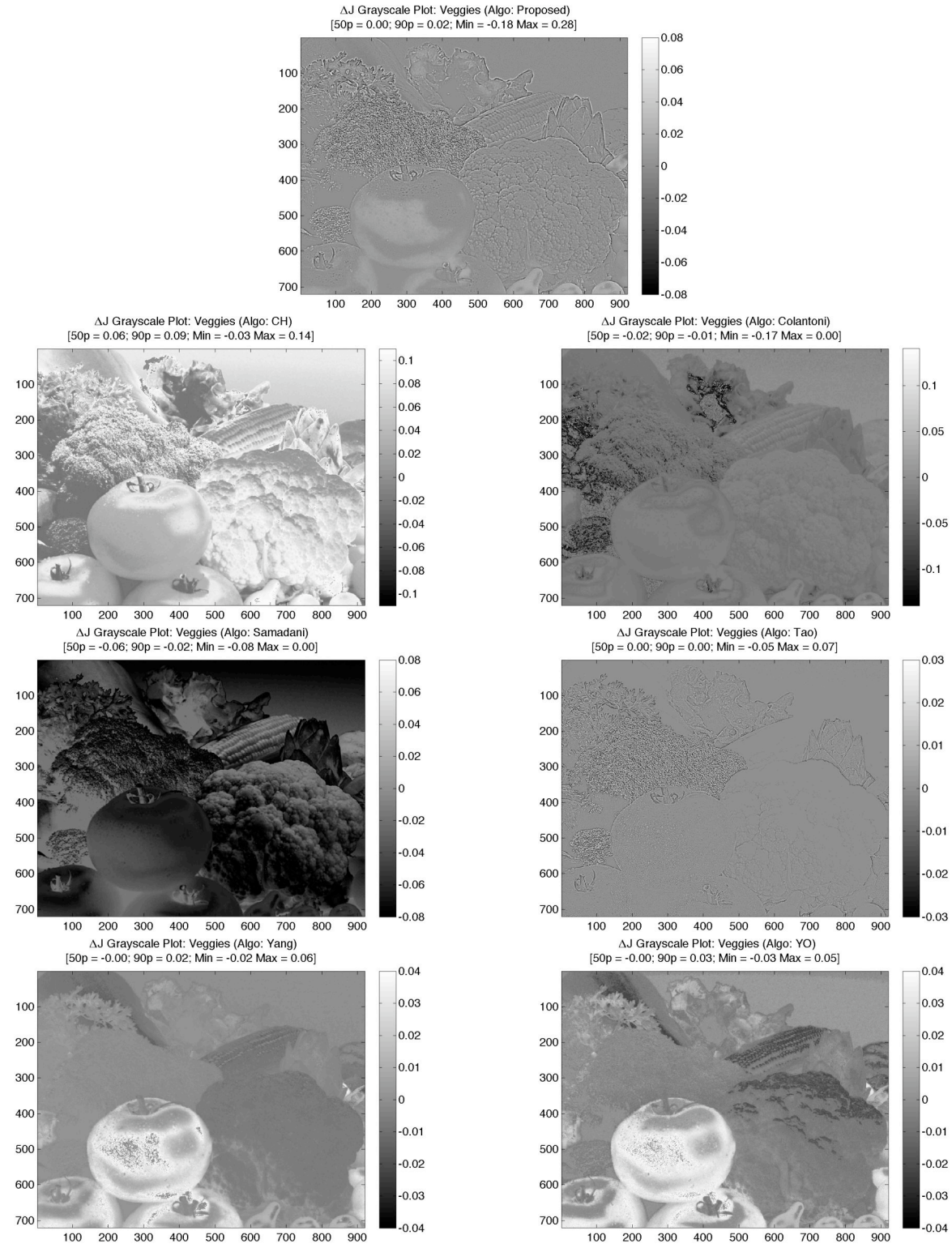


Fig. 5.19 ΔJ Image Difference Maps: Veggies

Fig. 5.20 ΔC Contour Maps: Veggies

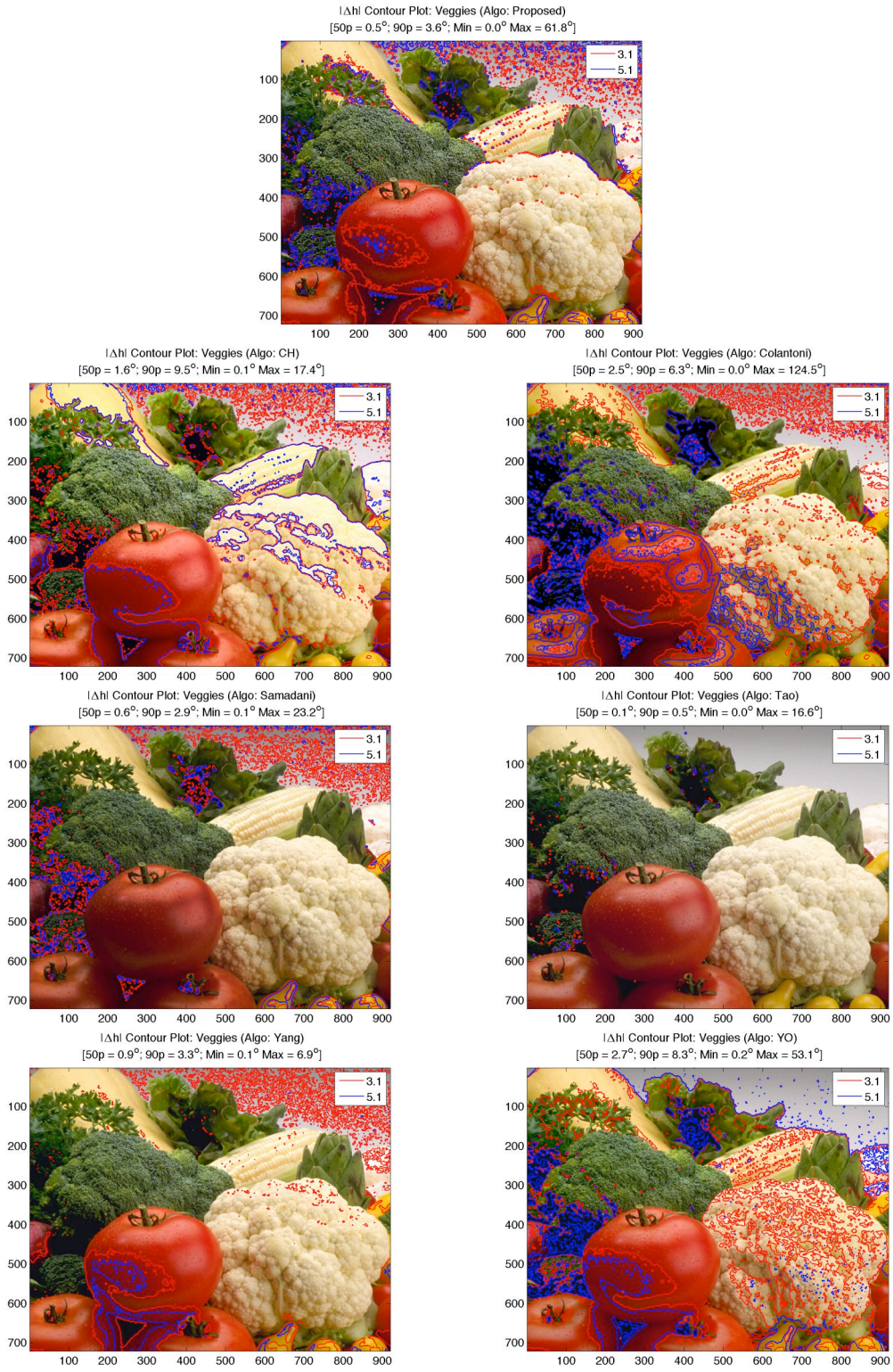


Fig. 5.21 Δh Contour Maps: Veggies

5.4. Conclusions

In this chapter, the performance of seven color/contrast algorithms was discussed. Since these algorithms are developed with different goals and applications, it is difficult to directly compare the results and determine if one algorithm is performing better than the others. Some of these algorithms are designed to enhance lightness, some are designed to enhance local contrast and some are meant to enhance overall image appearance by modifying lightness, color and contrast. Another difficulty experienced in this performance evaluation was the lack of information on specific aspects of implementations not available in the publication. In case of Samadani's algorithm, the author of the paper could be contacted to clarify some implementation issues [**Samadani 2007**]. However, the other algorithms were implemented based on personal understanding of the published paper. In some cases, the implementation was suited to serve the purpose of this performance analysis.

The performance analysis discussed in this chapter demonstrates that the choice of the color space is very important in achieving superior results. A color space that is perceptually uniform in terms of hue, like IPT, helps in preserving the original hue of the input color. When a color space is perceptually non-uniform, the result of color enhancement essentially depends on the image content. It is practically impossible to ensure that color enhancement in spaces such as RGB, YCC and LHS does not result in an unacceptable hue shift in the output image. Any perceptually uniform color space is bound to be nonlinear in nature, and when it comes to it is really a matter of optimization between superior performance and computational complexity.

Local contrast enhancement can increase noise present in the original image, but its perception depends on the lightness level. Enhanced noise cannot be perceived when lightness in a given area in the image is significantly low. The proposed algorithm as well as Tao's method cause significant edge enhancement, while the effect is more subtle in case of CH and YO. Other algorithms are not designed for local contrast enhancement.

The performance analysis also showed that the chroma enhancement strategy varies from one algorithm to the other. Algorithm CH causes very strong enhancement almost throughout the image. The proposed algorithm and YO are more selective and controlled in terms of chroma enhancement. This is particularly true in case of skin tone enhancement. YO is more selective than any other algorithms in this case. Even in the case of Colantoni and Yang, the chroma enhancement mainly took place in the colorful areas in the image. Even though Samadani's method does not include chroma enhancement feature, lightness reduction may result in chroma enhancement in parts of an image where lightness adjustment is significant. This can probably be attributed to the fact that the luma channel cannot be fully decoupled from the chroma channels in a typical opponent color space. The present implementation of Tao's method do not involve chroma enhancement.

In the concluding remark, it should be emphasized again that the main purpose of the performance analysis in this chapter was to quantitatively compare different functionality aspects of various algorithms. The analysis did not specifically take into account the perceptual effects of color and contrast enhancement.

"There are three principal means of acquiring knowledge . . . observation of nature, reflection, and experimentation. Observation collects facts; reflection combines them; experimentation verifies the result of that combination." ~ Denis Diderot (French man of letters and philosopher, 1713-1784)

Chapter 6

PSYCHOPHYSICAL EVALUATION OF THREE ALGORITHMS

In the previous chapter, a quantitative performance analysis of various algorithms was presented. This chapter contains a detailed description of the psychophysical experiments performed on outputs of several algorithms and an analysis of the results. Experiments were performed on still images as well as on video test sequences. Many of the algorithm implementations discussed in the pervious chapter were not integrated algorithms, focusing either on lightness adjustment, or on color enhancement, or simply contrast enhancement, but not all at the same time. Thus, it was not appropriate to include these algorithms in a single psychophysical experiment, as the end-results were very different. The experiments discussed in this chapter involves only three of the seven algorithms discussed, two Intel-proprietary algorithms CH and YO, and the proposed algorithm. All three algorithms attempt to enhance both color and contrast of the input images (or motion pictures).

This chapter starts with a discussion on display characterization process, which is essential for any psychophysical experiment involving display devices. Next, a paired comparison experiment conducted on several still images is described, followed by a similar experiment performed on video test sequences. Note that the experiments involved a Liquid Crystal Display (LCD) device, and so in the context of this chapter, a display mainly refers to an LCD.

6.1 Color Modeling of the LCD

For any display-based visual experiment, it is critical that the psychophysical images be transformed to colorimetric definitions accurately. This is only possible by accurately modeling, or characterizing, the inherent nonlinear nature of a computer-controlled display device. This nonlinearity is described by the optoelectronic transfer function (OETF), the relationship between the signals used to drive the display and the radiant outputs produced by the corresponding input. Determining this relationship is essentially a two-step process, namely, display calibration followed by a characterization.

6.1.1 Display Calibration

Device calibration is the process of maintaining the device with a fixed known characteristic of color response [**Sharma 2003**]. Many LCD display manufacturers build correction tables into the video card to convert the native luminance response curve (or gamma) into a desired response curve. It is advisable to turn off the built-in corrections if possible, and use the native gamma instead. As an example, the Display Calibrator application available in Mac OS goes through several steps in order to calibrate a display. These are: i) determining the display's native response, ii) selecting a target gamma (e.g. a linear gamma of 1, Mac standard of 1.8, PC standard of 2.2 or the native gamma), and iii) selecting the target white point (for example, native white point or D50 or D65). In the end, the calibration defines a display color profile, which includes the display gamut, the white point as well as the gamma. Figure 6.1 shows the report generated by Display Calibrator in Mac OS.

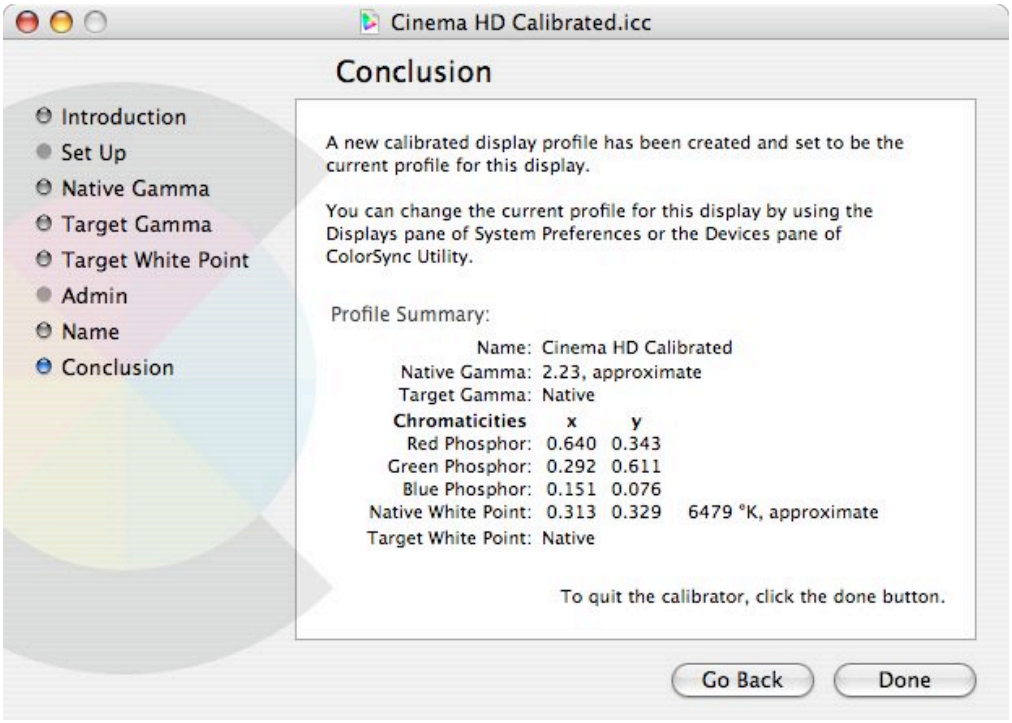


Fig. 6.1. Results of calibration using Display Calibrator in Mac OS

6.1.2 Display Characterization

Next, characterization is performed on the calibrated display. Characterization is a process that derives the relationship between device-dependent and device-independent color representations for a calibrated device. It is assumed that a calibrated device maintains the validity of the function, but the calibration process may need to be repeated from time to time to compensate for the temporal changes in the device's response and maintain it in a fixed known state [**Sharma 2003**]. Mathematical steps involved in the characterization process are described below. Discussion in this section is based on Day et al's work [**Day 2004**].

An effective way to represent the nonlinear characteristic in a computer-controlled display is to build one-dimensional look-up tables (LUTs) shown in Eq 6.1. These LUTs convert the original 0-255 RGB digital counts to linearized RGB values, thus defining the optoelectronic transfer functions for the three channels.

$$\begin{aligned} R &= LUT(d_r) \\ G &= LUT(d_g) \\ B &= LUT(d_b) \\ 0 &\le R,G,B \le 1 \end{aligned} \tag{6.1}$$

where d represents the digital counts and R, G and B are the radiometric scalars for the three channels, with values ranging between zero and unity.

The relationship between radiometric scalars and CIE tristimulus values is expressed by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_{r,\max} - X_{k,\min} & X_{g,\max} - X_{k,\min} & X_{b,\max} - X_{k,\min} & X_{k,\min} \\ Y_{r,\max} - Y_{k,\min} & Y_{g,\max} - Y_{k,\min} & Y_{b,\max} - Y_{k,\min} & Y_{k,\min} \\ Z_{r,\max} - Z_{k,\min} & Z_{g,\max} - Z_{k,\min} & Z_{b,\max} - Z_{k,\min} & Z_{k,\min} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix} \tag{6.2}$$

where $X_{r,max}$, $Y_{r,max}$ and $Z_{r,max}$ are the maximum tristimulus values obtained from the r channel. Tristimulus values corresponding to other channels are expressed similarly. $X_{k,min}$, $Y_{k,min}$ and $Z_{k,min}$ are the black-level flare. The black-level flare terms are separated into a single column to form the 3x4 transformation matrix. The above primary transformation matrix can be optimized by minimizing CIE ΔE_{00} color difference for a dataset sampling the display's colorimetric gamut.

6.1.3 Experimental Setup

A 22" flat-panel Apple Cinema LCD was characterized and subsequently used in all psychophysical experiments. The display was controlled by a 4x2.5 GHz PowerPC G5 computer running Mac OS X 10.4, with a 2 GB DDR 2 SDRAM memory. The display had a maximum resolution of 2560x1600 pixels. As explained previously, the display white point and gamma were set to native values.

A Graphical User Interface (GUI) was designed to display a round patch in the middle of the screen. Red, green and blue color samples were generated each as 11 step ramp at equal interval. Uniform grey background was used throughout the experiment. White color was displayed and measured in order to determine the display white point. Display white point and CIE 10° observer data were used in all calculations. The tristimulus values of the colors were measured with an illuminance-type LMT colorimeter, which was interfaced to the computer. Measurements were taken in a completely darkened room. The tristimulus values were saved in a file.

Three one-dimensional Look-Up Tables (LUT) corresponding to 256 digital counts, describing optoelectronic transfer functions of each of the three channels were created from the colorimetric data, as shown in Figure 6.2. An optimized 3x4 transformation matrix (Eq 6.2) to convert the digital count to tristimulus values was also created. Nonlinear optimization was used to minimize ΔE_{00} between the estimated and measured color patches. All the matrix coefficients were estimated simultaneously during the optimization process, changing the LUTs dynamically during each iteration.

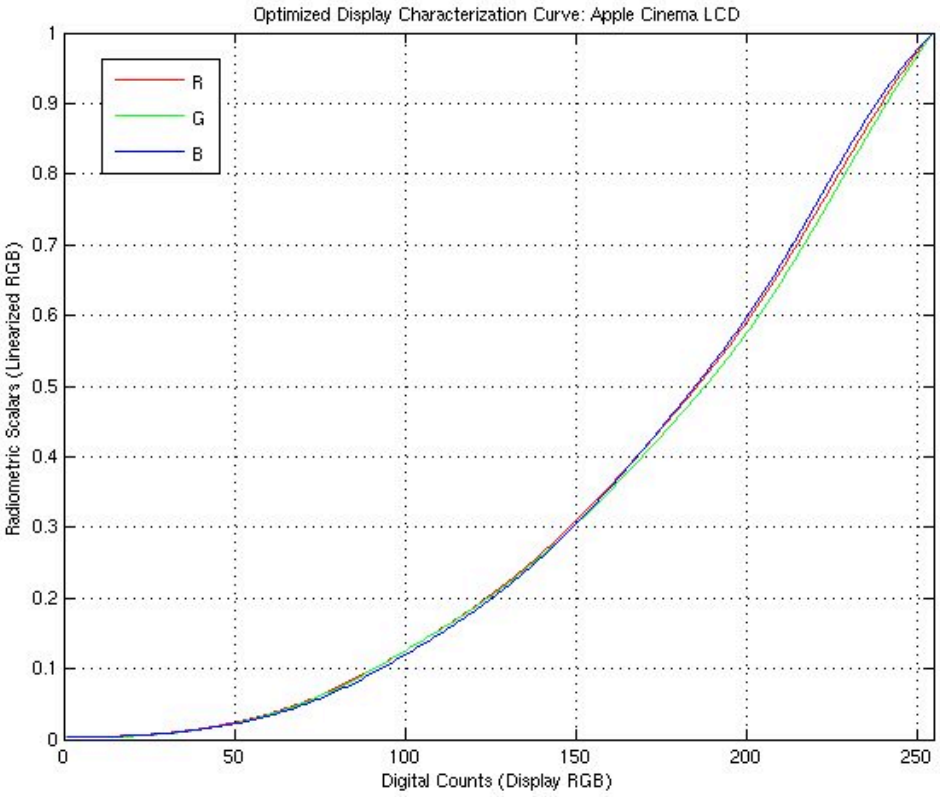


Fig. 6.2 Optimized Display Characterization Curve: Apple Cinema LCD

6.2 Psychophysical Experiments

Image quality is an integrated perception of the overall degree of excellence of an image [**Engeldrum 2004**]. Psychometric scaling is widely used in the imaging field for obtaining scale values of image quality and the "nesses", or its attributes. The process of scaling is not always straightforward. Several factors need to be taken into account during sample selection, choosing observers, formulating task instructions and finally, presenting and viewing the samples. The following subsections describe various aspects of the experimental design as well as an analysis of the results.

6.2.1 Experimental Goal

A key aspect of this research was to conduct subjective tests on the outputs of different algorithms. This was considered a potentially important contribution to the field since most of the development efforts published so far lack a systematic subjective assessment of the image enhancement algorithms. The main objective of conducting these psychophysical image quality experiments was to compare the performance of the new algorithm against some benchmark. In this case, the benchmarks were two image enhancement algorithms developed within Intel.

6.2.2 Software for Psychophysical Experiments

SiQ (pronounced sai-que), a software tool previously developed by the author was used for designing and executing the experiments as well as for analyzing the results. The software was developed in Matlab environment and had a graphical user interface. While the presentation of the trials for the still image experiment was handled by SiQ itself, some additional processing was required for the video experiment as described later.

6.2.3 Algorithms Evaluated

Apart from the new image enhancement algorithm, two Intel-proprietary algorithms were included in this test. Throughout this report, these algorithms are referred to as **CH** and **YO**. The new algorithm has been referred to as **NA**. The original images were also included in this experiment. These are referred to as **OR** in the figures. So there are four versions of each test image.

6.2.4 Test Images

Fifteen still images were included in the first psychophysical experiment. Each image had a size of 920x720 pixels.

Selection of the test images is extremely important for an unbiased evaluation of image enhancement algorithms. The following few pages contain the details of the test images used in this experiment. Note that for various color management and color appearance issues, images reproduced in this document will not have the same appearance as the actual images viewed on a characterized display. The contrast of the images in this document will generally appear a lot higher because of the reduced size. For the same reason, the output images are not included here. Various characteristics of output images generated by different algorithms have been discussed in the previous chapter.

Test Image 1

Critical aspects: Closely knit shiny beads of different colors.

Goal: To enhance color and contrast in such a way that individual beads as well as the specular reflections on them are distinctly visible

Test Image 2

Critical aspects: People's faces

Goal: This image is a little out-of-focus, and needs contrast enhancement. Because of the image content, the color enhancement has to be relatively subtle.

Test Image 3

Critical aspects: The animals, the blue jacket

Goal: Haziness due to the dust in the air needs to be preserved while enhancing the contrast

Test Image 4

Critical aspects: The colors, creases and/or the design in the dresses, skin tone, the green forest in the background

Goal: The skin tone and the greenery in the trees must not undergo strong enhancement. The man and the woman are a little blurred due to motion

Test Image 5

Critical aspects: Green grass outside the tunnel, the joints and the graffiti on the tunnel wall, the overhead fixtures

Goal: Preserve the high dynamic range in the scene; avoid introducing noise in the dark areas of the image; avoid strong enhancement of the grass

Test Image 6

Critical aspects: Colors of different vegetables, water droplets on the tomato, image contrast, particularly on the cauliflower, the broccoli and the corn

Goal: Avoid turning achromatic colors into chromatic ones while enhancing color and contrast of the vegetables

Test Image 7

Critical aspects: Same as the previous image. This is a low-contrast version of test image 6.

Goal: Increase the overall contrast (in addition to the stated goals for test image 6)

Test Image 8

Critical aspects: The bricks on the buildings, the stairs, the texts, the objects in the shadowed areas, the sky

Goal: White letters as well as the objects in the shadowed area should be more visible, the brick pattern should texts in general should look sharper

Test Image 9

Critical aspects: Intricate artwork on the walls, creases on the dresses of the people, bright light spots above the torchieres , light and shadow on the floor

Goal: enhance color and contrast while preserving the intricacies of the sculptured walls

Test Image 10

Critical aspects: Skin tone, hair on woman's forehead, her earrings, the blue denim shirt

Goal: Enhance color and contrast while avoiding strong color enhancement of the skin tone

Test Image 11

Critical aspects: color of the flowers, green plants, stone walls, dark areas under the big trees

Goal: enhance color and contrast of the flowers and green plants while preserving the naturalness; enhance the contrast of the dark areas so that shadowed objects are more visible

Test Image 12

Critical aspects: Different skin tones, colors on the dresses, text on the green badge

Goal: Avoid strong enhancement of different skin tones while enhancing color and contrast

Test Image 13

Critical aspects: Face of the man on the left, the text

Goal: Avoid enhancing the noise in the original image while boosting the contrast; avoid strong enhancement of the skin tone

Test Image 14

Critical aspects: The colorful parts in the balloon

Goal: Avoid making the noise in the original image more apparent while enhancing the color and contrast

Test Image 15

Critical aspects: LEDs and the overall contrast in this night scene

Goal: Avoid making the noise in the original image more apparent while enhancing the color and contrast

6.2.5 Test Movie Sequences

A second psychophysical experiment was performed on four video test sequences. As in the first experiment, outputs from the proposed algorithm, CH, YO, as well as the original sequences were included. The number of sequences had to be limited to four because of various reasons. Firstly, sequences appropriate for color and contrast enhancement were not readily accessible at the time of this research. Secondly, outputs from the proprietary algorithms CH and YO had to be obtained through the sponsor since IP issues were involved. Finally, the duration of the experiment had a practical limitation on the number of test sequences. The sequences had durations between 7.5 seconds and 10 seconds.

The movie sequence Avia had a resolution of 854x480. All other movie sequences were cropped to the same resolution to maintain uniformity. Having a resolution significantly larger than this for any sequence caused a playback and synchronization problem, in spite of using one of the most powerful computing resources available in the laboratory (configuration discussed earlier under experimental setup). The file sizes of the frames in these sequences varied from 1.1 MB to 1.3 MB. Processing did not change the file size appreciably, as expected.

A description of the video sequences follows.

6.2.5.1 Movie Sequence “Avia”

Description: This was a TIFF image sequence of 224 uncompressed images. This was not a continuous sequence, but rather selected from different locations in a long sequence. The duration of the clip was approximately 7.5 sec.

Content: This was mainly a restaurant scene starting with a close-up of desserts, and then panning on to the man and woman talking, then a close-up shot of fresh salad. Next, the clip shows a man talking to another man and a woman sitting in front of him. Finally, there is a close-up shot of a woman sitting against a black background with a color chart. Several frames of the sequence is shown in Figure 6.3.

Fig. 6.3 Different clips from the sequence Avia

Critical aspects: Various colors present in the content are important from the perspective of this experiment. Examples are the strawberries in the dessert, the wine glasses and other materials on the table, the skin tone, the colors of the carrot and other vegetables, the red shirt of the woman sitting in the couch and the color charts in the last several frames.

In several frames, contouring resulting from the compression led to challenging content for image enhancement. For example, the background in the first and third example frames shown in Figure 6.3 or the wall in the 4th example frame showed contours in the input image. If an algorithm makes these contours more visible while enhancing the contrast, it may result in reduced perceived video quality.

6.2.5.2 Movie Sequence "Calendar"

Description: This was a bitmap image sequence of 300 uncompressed images. The original size of 1280x720 was appropriately cropped so that the frames contained enough color even after cropping. The duration of the clip was approximately 10 sec.

Content: The clip shows a black and white stripe pattern along with some content with colors and a part of a calendar showing dates, as shown in Figure 6.4. The whole content rotates slowly throughout the duration of the clip.

Critical aspects: There is not a lot of color in this sequence. The stripe pattern or the numerals in the calendar are important for contrast enhancement evaluation and any resulting artifacts. The image is inherently noisy, so contrast enhancement might boost the noise as well, leading to poor picture quality. Even though noise was not included as a parameter in the psychophysical experiment, it is difficult for the observers to fully discount the effect of noise enhancement.

Fig. 6.4 Clips from the sequence Calendar

6.2.5.3 Movie Sequence "Vintage Car"

Description: This was a bitmap image sequence of 300 uncompressed images. The original size of 1280x720 was appropriately cropped so that the main object of interest in the clip, the vintage car in motion, was included in all frames. The duration of the clip was approximately 10 sec.

Content: The sequence shows a blue-colored vintage car approaching from a distance through a wooded area full of green foliage. The last several frames contain a close-up shot of the car and the man driving it. Figure 6.5 shows some representative frames.

Critical aspects: The foliage and the car are the main objects of interest with regard to color. In the close-up shots, the skin tone of the man and the color of the car are of importance. As in the Calendar sequence, there is noticeable noise in the input movie sequence, which is likely to get enhanced during contrast enhancement, unless a noise reduction module is included in the algorithm.

Fig. 6.5 Clips from the sequence Vintage Car

6.2.5.4 Movie Sequence "Walking Couple"

Description: This was a bitmap image sequence of 250 uncompressed images. The original size of 1280x720 was appropriately cropped so that the man and the woman walking down the wooded path were included in all frames. The duration of the clip was approximately 8.5 sec.

Content: The sequence shows a man and a woman walking together amidst the woods in the backdrop of trees. The man is wearing a bright yellow shirt and a colorful tie, while the woman is wearing a pink shirt with patterns on it. Two frames from the sequence are shown in Figure 6.6.

Critical aspects: The colors on the clothes are the most prominent features in this sequence, apart from the thick foliage in the background. The skin tone is also important.

Fig. 6.6 Clips from the sequence Walking Couple

6.2.6 Viewing Conditions

All images and video sequences were run through the Lookup Tables obtained from display characterization before displaying on the LCD screen. The experiments were performed in a completely dark room. The observers maintained a distance of around 30 inches from the screen.

6.2.7 Observers

A total of 25 color normal observers participated in each psychophysical experiment involving still images and the video test sequences. While both naïve and experienced observers were included in the experiments, no observer was familiar with the algorithms or the technology variables. It is important to not include such observers as they might have significantly different image preferences than the average observers [**Engeldrum 2001**]. Most of the observers were students and staff at the Munsell Color Science Laboratory, Rochester Institute of Technology. It is worthwhile to note that repeated empirical observation showed that experts and non-experts

judge image quality similarly when the task is application-independent, resulting in an image quality scale that is more "absolute" [**Engeldrum 2004**].

Before the actual experiments started, color vision of the observers was tested with some of the test plates from Ishihara Pseudoisochromatic test 24-plate [http://www.toledo-bend.com/colorblind/Ishihara.html]. This screening test was done by SiQ itself.

6.2.8 Experimental Method for Still Images

Since the main objective of the experiment was to generate an interval scale of image preference for the algorithm outputs, the method of paired comparison was determined to be the best method to use. This is one of the most common experimental techniques to quantity image quality [**Wu 1998**]. Note that this method generates a one-dimensional scale. In other words, we assume that the variability in the observers' responses can be fully expressed in a single dimension.

Every pair of the images was presented to the subject in a unique random order chosen by the software. The relative position of the images on the display screen was also randomized. The same pair of samples was presented only once. There were 90 observations in all (a pair can be chosen from 4 versions of a given image in ${}^4C_2$ or 6 ways, and there were 15 test images, so 6x15 = 90). Each session was completed in approximately 25 minutes on an average.

The following instruction was given to each observer:

Thank you for participating in our study.

There are 15 images. For each image, there are 4 versions. This experiment has 90 observations in all. In each observation, two versions of the same image will be displayed on the screen. Choose the image that you prefer. If you prefer the left image, click on the box displayed below the left image. If you prefer the right image, click on the box displayed below the right image.

There is no time limit, and there is no right or wrong answer. We are seeking your opinion. If you have any questions, please ask the test administrator at this point.

6.2.9 Experimental Method for Video Test Sequences

Similar to the still image experiment, the method of paired comparison was also used in the experiment involving video test sequences. Two movies of the same content were shown simultaneously, one at the top of the window and one at the bottom, playing in a continuous loop. The observers were allowed to take as long as needed before deciding on their preference. Additional processing was needed to prepare the video for this test as described below.

Step 1: Exporting Original and Processed Image Sequences As Quicktime Movies

The original and processed image sequences were first converted into movies using QuickTime Pro for Mac. Image compression was applied while generating the movies. While this was not a preferred method, uncompressed movies could not be played in the computer without experiencing playback problems. Note that two movies had to be played simultaneously, so synchronization was also an issue with uncompressed movies. The image compression reduced

the file sizes, and thus reducing the memory requirement during playback. H.264 image compression method with medium quality setting and a frame rate of 29.97 frames/sec was used. Every 24th frame was designated as a key frame in the compression process.

Step 2: Combining Clips in Pair Using SMIL Scripts

Once the movie clips were ready, SMIL (Synchronized Multimedia Integration Language) scripts were written to generate combined movie clips with every possible combination [**SMIL 2005**]. There were four versions for each movie clip (original, new algorithm and two Intel proprietary algorithms), which could be combined in ${}^4C_2$ * 2 or 12 ways, since a given version could be played either on top, or on bottom. Thus, for four input movie clips, there were 48 combined movies. With the chosen resolution, the combined movie clip could be played without any synchronization problem.

Step 3: Loading Combined Movie Clips in A Web Browser

In the final step, HTML script was written to load a combined movie clip onto a web browser. In this case, the browser was Safari (for Mac). The browser parameters, including the background color, exact position of the movie in the browser window were set within the script. The whole background was set to gray. This script was saved as a webpage. During the experiment, SiQ edited the script to insert the name of the appropriate movie according to the predefined trial sequence.

The following instruction was given to each observer:

Thank you for participating in our study.

In each trial, two versions of a given clip will be randomly picked by the software and will be presented to you in a browser. The movies will continuously play in loop. Determine which clip has the highest OVERALL picture quality in terms of color, contrast and sharpness. IGNORE NOISE IN THE MOVIES.

Once you have decided which clip to choose, click on the appropriate button (top or bottom) in the input window based on your choice. If you accidentally close the browser, you can reload the current trial by pressing the reload button. Please DO NOT resize the browser at any time.

There is no time limit, and there is no right or wrong answer. We are seeking your opinion.

To get started, click OK and then click anywhere in the browser. If you have any questions, please ask the test administrator at this point.

6.3 Results and Discussion

Data from complete pair wise comparisons were analyzed using Thurstone's Law of Comparative Judgment Case V [**Thurstone 1927**] in order to create an interval scale of overall image preference.

6.3.1 Thurstone's Law of Comparative Judgment

Following are the hypotheses behind Thurstone's Law of Comparative Judgment (Thurstone 1927):

1. Each stimulus gives rise to a discriminal process, which has some value on the psychological continuum of interest.
2. Due to momentary fluctuations (internal fluctuations occurring within or between observers), the value of a stimulus may be higher or lower on repeated presentations. The distribution of this fluctuation can be characterized by a normal distribution.
3. The mean and standard deviation of the distribution associated with a stimulus are its internal scale values and discriminal dispersion, respectively.
4. The distribution of the difference between two stimuli is also normally distributed and is a function of the proportion that one stimulus is chosen as greater than the other.
5. The difference in scale values, R, between two stimuli, i and j, is:

$$R_i - R_j = z_{ij}\sqrt{\sigma_i^2 + \sigma_j^2 - 2r_{ij}\sigma_i\sigma_j} \tag{6.3}$$

where R_i and R_j represent the scale values of stimuli i and j, σ_i and σ_j, are the standard deviations of the respective discriminal dispersions, r_{ij} is the correlation between the two discriminal

processes, and z_{ij} is the normal deviate (the z-score) corresponding to the proportion of times stimulus *j* is judged is judged greater along the psychological continuum than stimulus *i*.

Eq (6.3) can be simplified based on certain assumptions:

1. The evaluation of one stimulus along the continuum does not influence the evaluation of the other in the paired comparison ($r_{ij} = 0$).
2. The dispersions are equal for all stimuli ($\sigma_i = \sigma_j$).

Accordingly, following equation is obtained:

$$R_i - R_j = z_{ij}\sqrt{2} \tag{6.4}$$

6.3.2 Confidence Interval

In terms of interval scale unit, standard deviation σ can be expressed as

$$\sigma = \frac{1}{\sqrt{2}} \tag{6.5}$$

Standard error of the scale value is

$$\frac{\sigma}{\sqrt{N}} = \frac{1}{\sqrt{2N}} = \frac{0.707}{\sqrt{N}} \tag{6.6}$$

Where N is the number of observations per pair, which is equal to the number of observers if each observer views a trial a single time.

The confidence interval for the paired comparison data can be expressed as

$$\mu = \pm F \frac{\sigma}{\sqrt{N}} \tag{6.7}$$

Here F is a function of the Level of Confidence (LOC). F is derived from the Normal Curve.

With 95% LOC, F = 1.96. So, the confidence interval is

$$\mu = \pm 1.96 \frac{0.707}{\sqrt{N}} = \pm \frac{1.38}{\sqrt{N}} \tag{6.8}$$

Note that the confidence interval does not depend on the number of observations – this is the weakness of this statistical metric.

Since there were 25 observers in each experiment, so the confidence interval was 0.276. Note that the confidence interval computed using above equations is the same for all images/sequences in a given experiment.

6.3.3 Interval Scale Plots: Still Image Experiment

An analysis of the experimental data leads to a separate interval scale for each test image (i.e. 15 interval scales in all). An average of these 15 interval scales is presented in Figure 6.7. Strictly speaking, we should not combine different interval scales since each image is a different psychophysical experiment and there is no common anchor point in these scales so that they can be tied together. Another issue with using this figure is that it does not reflect the fact that the observer preferences are strongly dependent on the image content, as will be clear shortly. However, this figure gives an impression on how each algorithm performed for all the images on an average. Existing algorithms CH and YO performed equally well, while the new algorithm

probably performed slightly better. All three algorithms generally performed better than the original. The error bars in this figure signify confidence intervals, or statistical uncertainty. Note that the zero does not have any specific meaning in the interval scale plot, only the relative positions are important. In other words, we can add any arbitrary constant to the interval scale values without affecting the results.

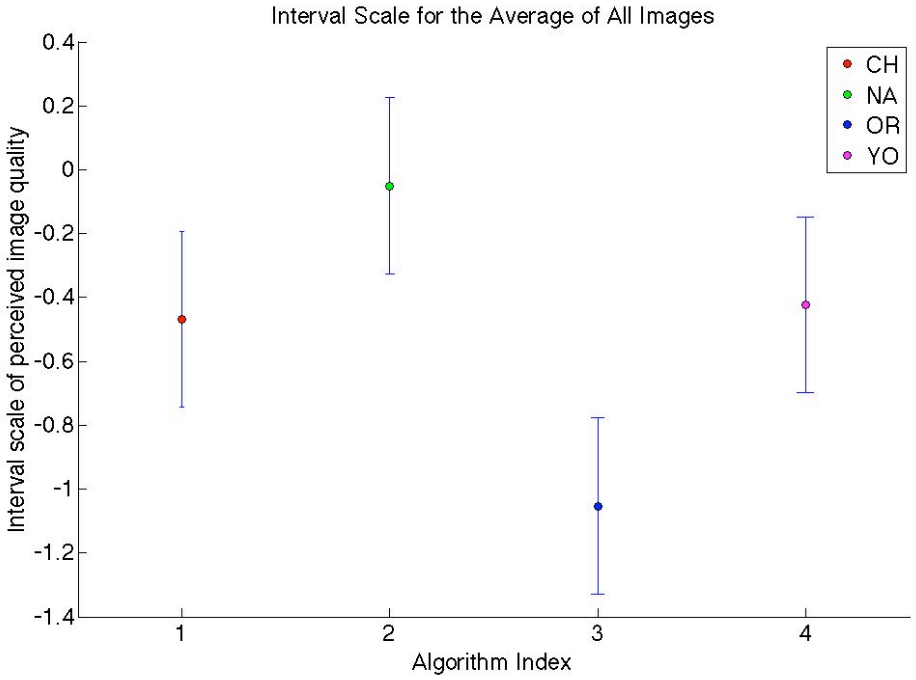


Fig. 6.7. Interval scale for the average of all images

The implication of the confidence interval is, based on this figure alone, we cannot make an inference that the new algorithm will perform better than the existing ones under all circumstances. This is more evident when we look at the result for all test images, as shown in Figure 6.8.

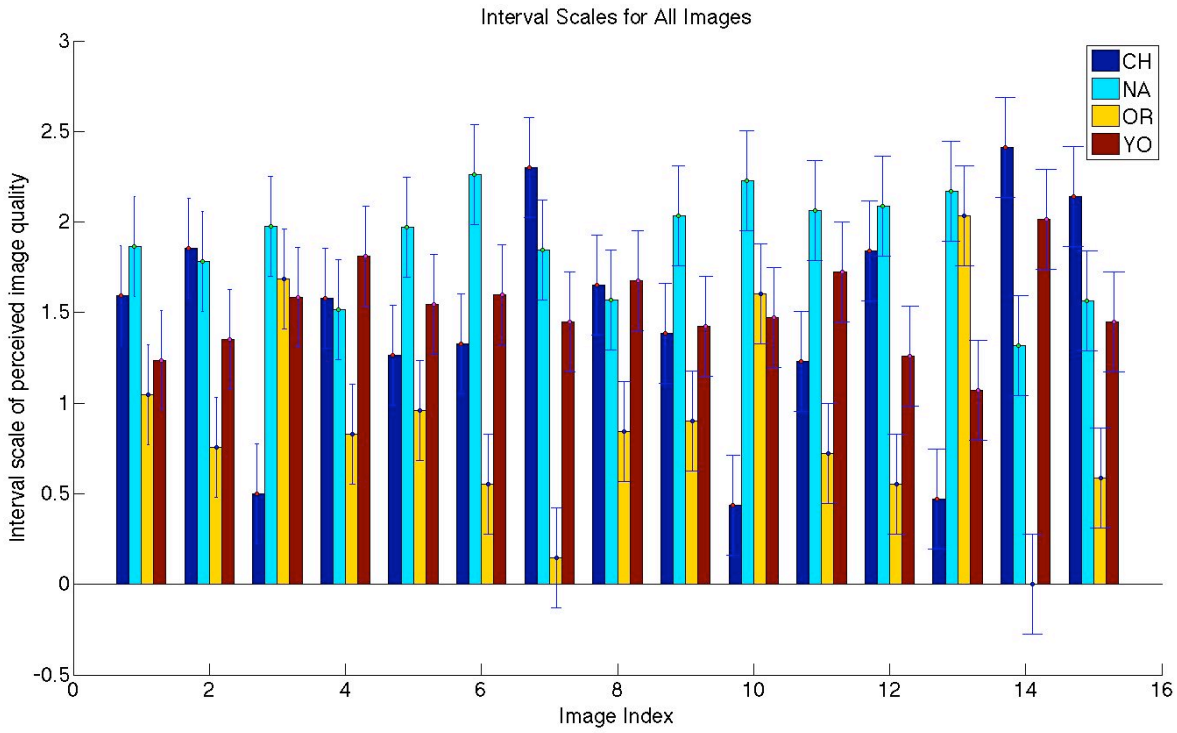


Fig. 6.8. A summary of interval scales for all test images

In the above figure, the interval scales have been shown for all 15 test images. Each bar corresponds to one version of the image (three algorithm outputs, or the original). Evidently, not a single algorithm was preferred for all these images. For many images, difference in the interval scale values for two or more algorithms is statistically not significant. Less is the overlap between two error bars, the more statistically significant is the corresponding interval scale difference. Table 6.1 summarizes observer preferences that can be considered statistically significant.

Table 6.1. Ranking Table for the performance of different algorithms in the still image experiment

Algorithm	Ranked #1		Ranked #3 or #4	
	Images	No of times	Images	No of times
CH	7, 14, 15	3	3, 10, 11, 13	4
NA	5, 6, 9, 10	4	14	1
OR	-	0	1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15	12
YO	-	0	2, 12, 13	3

Note that even though algorithm YO output was not ranked #1 for any image while CH output was ranked #1 thrice, the overall ratings for the two algorithms are comparable (from Figure 6.6), as the performance of CH was worse for several images (e.g. image 10 and 13).

Comparatively, the new algorithm has performed consistently well. Its outputs were ranked #1 for four test images, and it performed significantly worse compared to the other algorithms only once (test images 14). In comparison, CH and YO were less consistent. They were ranked among the worst two versions 4 and 3 times respectively.

As expected, for most of the test images the algorithm outputs were preferred over the originals.

Note that, this result is specific to the set of test images used in this experiment. For a different set, this result might vary. However, the set used here includes a wide variety of image content and thus, the experimental results give a fair idea about how these algorithms might perform under different circumstances.

6.3.4 Interval Scale Plots: Video Experiment

As in the still image experimental analysis, the software SiQ generated interval scale for the average of all clips (Figure 6.9) as well as for individual movie clips (Figure 6.10). There is more ambiguity in these results than in the results obtained from the still image experiment. The difference in the interval scale of perceived picture quality for the three algorithms is statistically not significant.

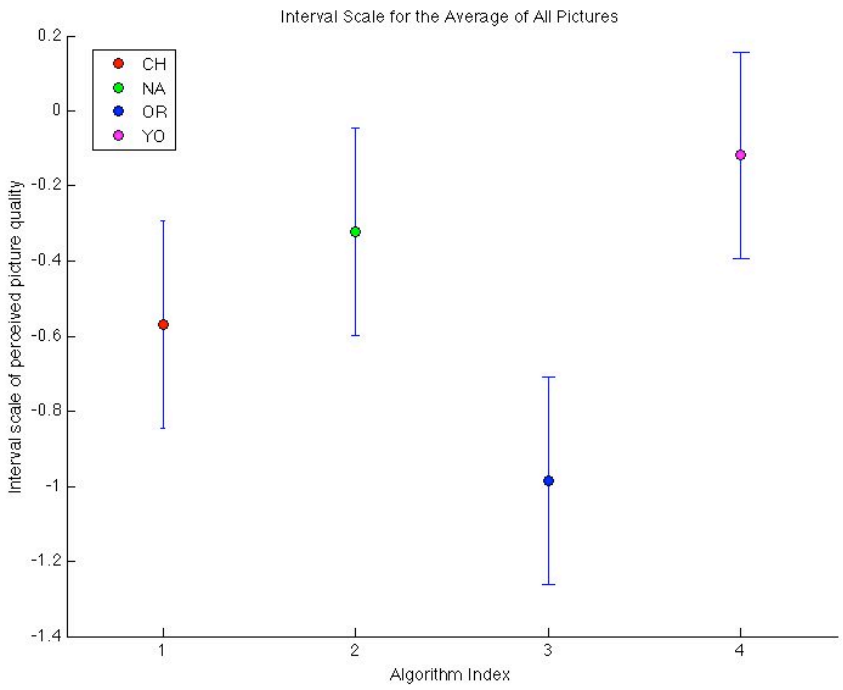


Fig. 6.9 Interval scales for the average of all clips

However, we can draw some general conclusions from the individual interval scales shown in Figure 6.10. Outputs of all three algorithms did better or similar to the original. This is more obvious for algorithms NA (proposed) and YO, whose performances were very similar for all four clips. Performance of CH was noticeably better than the other algorithms for the Calendar clip, but for other clips, it did not quite improve the perceived picture quality compared to the original.

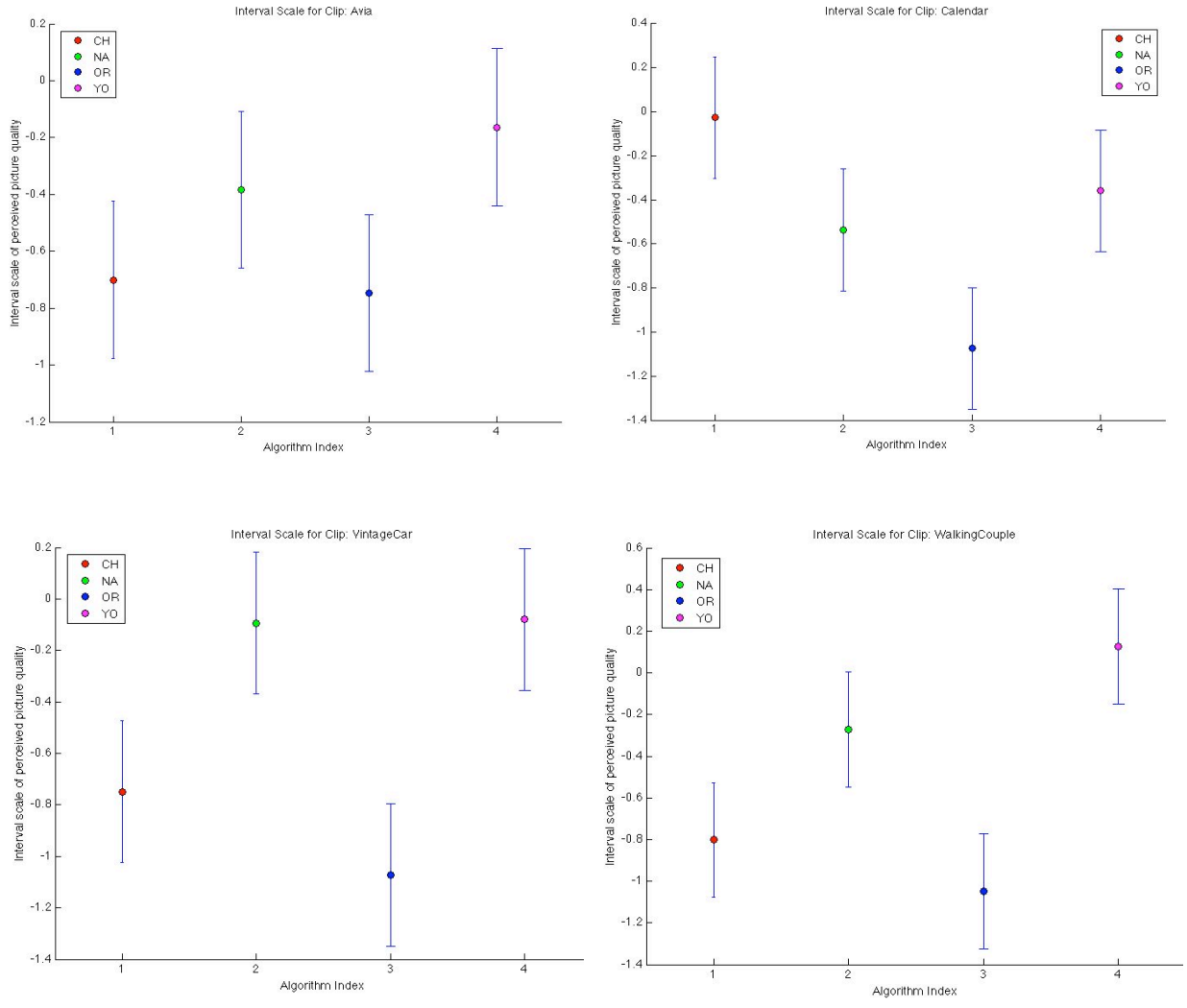


Fig 6.10 Interval scales for the four movie clips

Clip Avia shows more statistical uncertainty than the other clips. This is likely due to the image content. The change in the subject matter in this clip was rather fast, leaving observers with less time to discern perceptual difference between various outputs.

The fact that NA and YO showed a consistent performance across the clips is also evident in Figure 6.11, which shows the interval scales for all four clips and for all four versions of them.

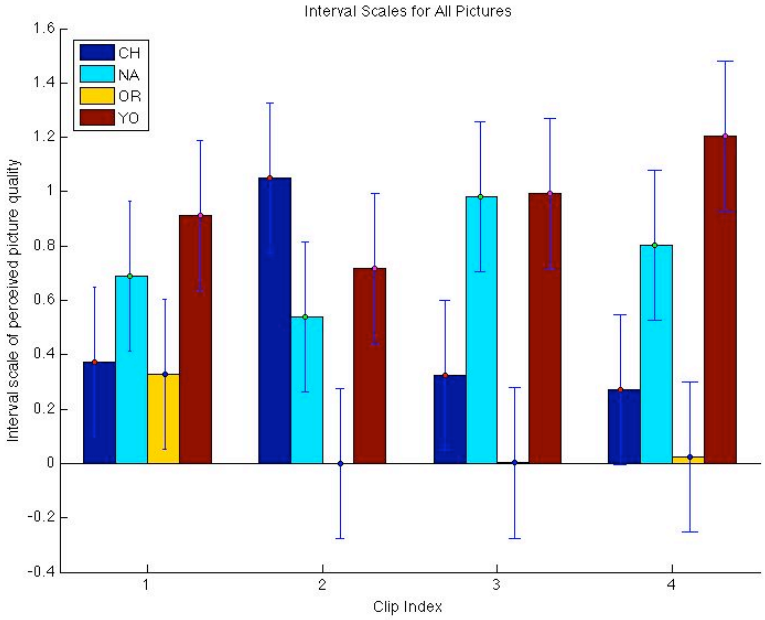


Fig. 6.11 Summary of interval scales for all movie clips

The somewhat inconclusive results from the video experiment can be attributed, at least partially, to the lack of appropriate image content. Movie clips available for this experiment were not probably very appropriate for evaluation of color and contrast enhancement aspects of different algorithms. The effect of color/contrast enhancement did not turn out to be very apparent in many cases, leading to the confusion or ambiguity in observer data. As explained earlier, higher

resolution clips could not be used in this experiment because of playback issues. Image compression had to be applied to reduce the memory requirement during playback. These imposed major limitation on the experimental setup. Reduced resolution has a direct impact on perceived contrast. The process of image compression arguably added some unknown variables into the processing chain, which might have affected the perceived picture quality.

6.3.5 Inference from the Results

The results from the psychophysical experiments, particularly the one involving still images, indicate that the new algorithm is performing well in most of the cases, although there are areas where it needs improvement. When an image is inherently noisy, the contrast enhancement in the algorithm is causing that noise to be enhanced as well, leading to a poor quality image. This is evident in cases of test image 14 and 15. This problem can be circumvented by introducing a noise detection mechanism that will prevent any contrast enhancement in case of noisy image contents. Alternately, noise detection or even noise suppression mechanism can be a part of the video processing chain, and the noise information can be provided as an algorithm input.

Even though skin tone enhancement worked reasonably well in the proposed algorithm, as evident in case of test image 10, a separate skin tone detection mechanism may need to be incorporated to improve the performance of the algorithm for this type of image contents.

With respect to processing image sequences, the results were not markedly different from the still images. No major temporal artifacts were noticed in any of the three algorithms included in the visual experiments. However, since the new algorithm relies on cumulative distribution

function to determine the intensity enhancement required in a given frame, chances are in some specific cases, there will be a noticeable shift in color from one frame to the next. Some sort of temporal processing should be included as a safeguard against that kind of situation. However, this problem was not observed in the test clips included in the visual experiment.

Overall, the performance of the new algorithm is quite promising. Further development can make this algorithm more successful as an automatic image enhancement method.

"Where is wisdom we have lost in knowledge? Where is the knowledge we have lost in information?" ~ T.S. Eliot (American born English Editor, Playwright, Poet and Critic, 1888-1965)

Chapter 7

CONCLUSIONS AND FUTURE RESEARCH

The comparative performance analysis and psychophysical experiments presented in this thesis demonstrate the challenges involved in designing an automatic color/contrast enhancement algorithm that will consistently produce pleasing results for various image/movie content. While conventional color and contrast enhancement methods in video processing typically involve signal and/or image processing, they do not generally consider the perceptual aspects of human vision. The novelty of the image enhancement approach developed as part of this research lies in the fact that it involves perceptual image processing and builds upon the knowledge acquired from past research in the field of color appearance. The new method combines color and contrast enhancement into one integrated algorithm, thereby producing optimal enhancement requiring no additional tune-up.

To summarize the research findings, following are some key aspects relevant for the development of an effective color and contrast enhancement method for images and video applications:

1. **The choice of color space is critical:** Image/video processing in a perceptually uniform color space helps in achieving visually pleasing results, while minimizing color artifacts and the need for additional color correction methods

2. **An ad hoc approach is detrimental:** It is preferable to achieve moderate enhancement for a wide variety of image content than superior enhancement in some cases and unacceptable results in others
3. **Color attributes are interdependent:** As lightness of a given color is increased, the corresponding maximum attainable saturation increases up to a certain value, then it decreases; the relationship is dependent on the hue
4. **Lightness adjustment should be globally adaptive:** An input image/video that is mostly dark should be lightened to an appropriate level, while an image/video with high lightness should be darkened
5. **Color enhancement should be content dependent:** Often times, a strong chroma enhancement can lead to a loss in detail, unrealistic colors, and in some cases, an out-of-gamut color (depending on the corresponding lightness and hue)
6. **Contrast enhancement should be locally adaptive:** A strong contrast enhancement may be objectionable in some cases (e.g. people's faces or uniform backgrounds), while in other cases it may help accentuate the details
7. **Certain colors may need special processing:** Memory colors like skin tone, natural green and blue sky may need special enhancement
8. **Noise should not be amplified:** If noise detection and suppression module does not precede color/contrast enhancement in a video processing chain, the algorithm must incorporate noise reduction filters

As the novel digital display technologies continue setting new standards for the quality of consumer video, the role of color processing becomes increasingly vital. Larger screen size, higher luminance and higher resolution of today's state-of-the-art digital display systems require significantly more sophisticated color imaging techniques than what was adequate a decade ago. Any future development efforts in color video processing must recognize the potential for improvement in color reproduction capabilities of various emerging display technologies. Using characterized video cameras with recorded camera settings at the capture end, video processing in a perceptually linear space, and displays with proper colorimetric characterization at the output end of the processing chain can all go a long way in ensuring consistent color reproduction across a myriad of display technologies. While this is easier said than done, video researchers and color scientists must work together toward the common goal of superior picture quality in consumer video applications.

At the same time, research on video quality assessment must go on in order to develop a reliable, universal perceptual video quality metric and an assessment methodology. This will be essential in benchmarking various video processing techniques, both display-specific and display-independent. Such an endeavor will require coordinated research efforts in the areas of human vision, color science and video processing.

BIBLIOGRAPHY

I. Color Science: Theory and Principles

[**Berns 2000**] Berns R.S., *Billmeyer and Saltzman's Principles of Color Technology*, 3rd Ed., John Wiley & Sons Inc, New York, USA (2000)

[**Fairchild 2005**] Fairchild M.D., *Color Appearance Models*, 2nd Ed., Wiley-IS&T Series in Imaging Science and Tehnology, Chichester, UK (2005)

[**Hunt 2001**] Hunt, R.W.G., Saturation, superfluous or superior?, *Final Program and Proceedings - IS&T/SID Color Imaging Conference*, 1-5 (2001)

[**Hunt 1998**] Hunt, R.W.G., *Measuring Colour*, Third Ed., Fountain Press, Kingston-upon-Thames, England (1998)

[**Johnson 2000**] Johnson, G.M., Fairchild, M.D., Sharpness rules, *Final Program and Proceedings - IS&T/SID Color Imaging Conference*, 24-30 (2000)

[**Susstrunk 1999**] Susstrunk, S., Buckley, R., Swen, S., Standard RGB color spaces, *Final Program and Proceedings - IS&T/SID Color Imaging Conference*, 127-134 (1999)

II. Color Video Processing

[**Bovik 2005**] Bovik, A., Editor, *Handbook of image and video processing*, 2nd Ed., Elsevier Academic Press, USA (2005)

[**Caviedes 2008**] Caviedes J., Personal Communication (March 28, 2008)

[**de Haan 1999**] de Haan, G., Video format conversion, *Society for Information Display 1999 International Symposium*, 52-57 (1999)

[**de Haan 2003**] de Haan, G., *Video processing for multimedia systems*, 3rd Ed., University Press Eindhoven (2003)

[**Fairchild 2007**] Fairchild, M.D., A color scientist looks at video, *3rd International Workshop on Video Processing and Quality Metrics (VPQM)*, Scottsdale, Invited Paper 1 (2007)

[**Kim 2005**] Kim, C.H., Lee, S., Park, D.S., Kwak, Y., DTV color and image processing: Past, present, and future, *Color Imaging XI: Processing, Hardcopy, and Applications - Proceedings of SPIE - IS&T Electronic Imaging*, **6058**, 60580E (2006)

[**Klompenhouwer 2004**] Klompenhouwer, M. A., de Haan, G., Invited Paper: Video, Display and Processing, *SID International Symposium - Digest Of Technical Papers*, **35**(2), 1466-1469 (2004)

[**Poynton 2003**] Ponton, C., *Digital video and HDTV: algorithms and interfaces*, Morgan Kaufmann, San Francisco, CA (2003)

[**Shen 1997**] Shen, M.Y., Kuo C.C.J., Review of image postprocessing techniques for compression artifact removal, *Proceedings of the SPIE - The International Society for Optical Engineering*, **3164**, 372-382 (1997)

[**Yuen 1998**] Yuen, M., Wu, H.R., A survey of hybrid MC/DPCM/DCT video coding distortions, *Signal Processing*, **70**(3), 247-278 (1998)

III. Emerging Display Technologies

[**Choe 2005**] Choe W.H., Lee, S.D., Kim, C.Y., Studying for multi-primary LCD, *Proceedings of the SPIE - The International Society for Optical Engineering*, **5667**(1), 336-43 (2005)

[**de Haan 2001**] de Haan, G.; Klompenhouwer, M.A., An overview of flaws in emerging television displays and remedial video processing, *IEEE Transactions on Consumer Electronics*, **47**(3), 326-334 (2001)

[**de Haan 2005**] de Haan, G., Zhao, M., Invited paper: Making the best of legacy video on modern displays, *SID International Symposium - Digest Of Technical Papers*, **37**(4), 1863-1866 (2006)

[**de Haan 2007**] de Haan, G., Television Display Processing: Past & Future, *International Conference on Consumer Electronics: Digest of Technical Papers*, 1-2 (2007)

[**DeMarsh 1991**] DeMarsh, L., Colorimetry for HDTV, *IEEE Transactions on Consumer Electronics*, **37**(1), 1-6 (1991)

[**DLP 2008**] DLP website, Texas Instruments, http://www.dlp.com/tech/what.aspx, Accessed May 14, 2008

[**Itoh 2007**] Itoh, S., Tanaka, M., Tonegwa, T., Taniguchi, M., Niiyama, T., Tamura, K., Namikawa, M., Naito, Y., Obara, Y., Toriumi, M., Kobayashi, H., Takeya, Y., Deguchi, K., Kawata, S., Marushima, Y., Fujimura, Y., Nawamaki, K., Kubo, Y., Sato, Y., Kataoka, F., Sakurada, K., Ishibashi, M., Otsu, K., and Tatsuo Yamaura, T., Invited Paper: Development of Field-Emission Display, *SID Symposium Digest 38*, 1297-1300 (2007)

[**Kim 2004**] Kim, M.C., Shin, Y.C., Song Y.R., Lee, S.J., Kim, I.D., Wide gamut multi-primary display for HDTV, CGIV 2004, Second European Conference on Color in Graphics, Imaging, and Vision and Sixth International Symposium on Multispectral Color Science, 248-253 (2004)

[**Kwak 2005**] Kwak, Y., Lee, S.D., Choe, W, Kim, C.Y., Optimal chromaticities of the primaries for wide gamut 3-channel display, Proceedings of SPIE - The International Society for Optical Engineering, **5667**, Proceedings of SPIE-IS&T Electronic Imaging - Color Imaging X: Processing, Hardcopy, and Applications, 319-327 (2005)

[**Lee 2002**] Lee, S.D., Kim, C.Y., Seo, Y.S., Hong, C.W., Color conversion from RGB to RGB+White while preserving hue and saturation, Final Program and Proceedings - IS&T/SID Color Imaging Conference, 287-291 (2002)

[**Murdoch 2006**] Murdoch, M.J., Miller, M.E., Kane, P.J., Perfecting the color reproduction of RGBW OLED, *Final Programs and Proceedings - International Congress of Imaging Science*, 448-51 (2006)

[**Oh 2006**] Oh, H.H., Lee, H.Y., Kim, S.S., Park, D.S. Kim, C.Y., Compensation method for color defects in PDP due to different time responses of phosphors, *Color Imaging XI: Processing, Hardcopy, and Applications - Proceedings of SPIE-IS&T Electronic Imaging*, **6058**, 60580G (2006)

[**Ok 2005**] Ok, H.W., Lee, S.D., Choe, W.H., Park, D.S., Kim, C.H., Color processing for multi-primary display devices, Proceedings - IEEE International Conference on Image Processing, **3**, 980-983 (2005)

[**Pettitt 2001**] Pettitt, G., Walker, B., DLP cinema™ technology: Color management and signal processing, *Final Program and Proceedings - IS&T/SID Color Imaging Conference*, 348-354 (2001)

[**Roth 2007**] Roth, S., Weiss, N., Chorin, M.B., David, I.B., Chen, C.H., Multi-primary LCD for TV applications, **38**(1), *SID International Symposium - Digest of Technical Papers*, 34-37 (2007)

[**Sugiura 2006**] Sugiura, H., Kagawa, S., Kaneko, H., Ozawa, M., Tanizoe, H., Kimura, T., Ueno, H., Wide color gamut displays using LED backlight - signal processing circuits, color calibration system and multi-primaries, *International Conference on Image Processing*, II-9-12 (2006)

[**Sugiura 2007**] Sugiura, H., Kuwata, M., Inoue, Y., Sasagawa, T., Nagase, A., Kagawa, S., Watanabe, N., Someya, J., Invited paper: Laser TV - Ultra wide color gamut in conformity with xvYCC, **38**(1), *SID International Symposium - Digest of Technical Papers*, 12-15 (2007)

IV. Video Quality and Video Quality Assessment

[**Ahumada 1993**] Ahumada, Jr. A.J., Computational image quality metric: a review, *SID Symposium Digest*, **24**, 305-308 (1993)

[**ATIS 2001**] ATIS Technical Report T1.TR.75-2001, Objective perceptual video quality measurement using a JND-based full reference technique (October 2001)

[**Brill 2004**] Brill, M.H., Lubin, J., Costa, P., Wolf, S., Pearson, J, Accuracy and cross-calibration of video quality metrics: new methods from ATIS/T1A1, *Signal Processing: Image Communication*, **19**(2), 101-7 (2004)

[**Caviedes 2000**] Jorge Caviedes, Antoine Drouot, Arnaud Gesnot, and Laurent Rouvellou, Impairment metrics for digital video and their role in objective quality assessment, *Proc. SPIE Int. Soc. Opt. Eng*. **4067**, pt.1-3, 791-800 (2000)

[**Caviedes 2003**] Caviedes J., Oberti F., No-reference quality metric for degraded and enhanced video. *Proc. SPIE*, **5150**, Lugano, Switzerland, 621-632 (2003)

[**Corriveau 1999**] Corriveau, P., Gojmerac, C., Hughes, B., Stelmach, L., All subjective scales are not created equal: the effects of context on different scales, *Signal Processing*, **77**(1), 1-9 (1999)

[**Daly 1993**] Daly, S., The visible differences predictor: an algorithm for the assessment of image fidelity, *Proceedings of the SPIE - The International Society for Optical Engineering*, **1666**, 2-15 (1992)

[**Day 2004**] Day, E.A., Taplin, L., Berns, R.S., Colorimetric characterization of a computer-controlled liquid crystal display, *Color Research and Application*, **29**(5), 365-37 (2004)

[Engeldrum 2001] Engeldrum, P.G., Psychometric Scaling: Avoiding the Pitfalls and Hazards, *Society for Imaging Science and Technology: Image Processing, Image Quality, Image Capture, Systems Conference*, 101-107 (2001)

[Engeldrum 2004] Engeldrum, P.G., A theory of image quality: The image quality circle, *Journal of Imaging Science and Technology*, **48**(5), 447-457 (2004)

[**Fairchild 2003**] Fairchild, M.D., Johnson, G.M., Image appearance modeling, *Proceedings of the SPIE - The International Society for Optical Engineering*, **5007**, 149-156 (2003)

[**Fairchild 2004**] Fairchild, M.D., Johnson, G.M., iCAM framework for image appearance, differences, and quality, *Journal of Electronic Imaging*, **13**(1), 126-38 (2004)

[**Faugeras 1979**] Faugeras, O.D., Digital color image processing within the framework of a human vision model, *IEEE Trans. Acoust. Speech Signal Processing*, **27**(4), 380-393 (1979)

[**ITU 2002**] ITU-R Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures, International Telecommunication Union, Geneva, Switzerland (2002)

[**Libert 2000**] Libert, J.M., Stanger, L., Watson, A.B., Rohaly, A.M., Toward developing a unit of measure and scale of digital video quality: IEEE Broadcast Technology Society Sub-committee on Video Compression Measurements, Proceedings of the SPIE - The International Society for Optical Engineering, **3959**, 160-165 (2000)

[**Lindh 1996**] Lindh, P., van den Branden Lambrecht, C.J., Efficient spatio-temporal decomposition for perceptual processing of video sequences, *Proc ICIP*, Lausanne, Switzerland, **3**, 331-334 (1996)

[**Lubin 1997**] Lubin, J., A human vision system model for objective picture quality measurements, *International Broadcasting Convention*, 498-503 (1997)

[**Lukas 1982**] Lukas, F.X.J., Budrikis, Z.L., Picture quality prediction based on a visual model, *IEEE Trans. Comm.*, **30** (7), 1679-1692, (1982)

[**Mannos 1974**] Mannos, J.L., Sakrison, D.J., The effects of a visual fidelity criterion of the encoding of images. *IEEE Transactions on Information Theory*, **IT-20**(4), 525-36 (1974)

[**Miyahara 1998**] Miyahara, M., Kotani, K., Algazi, V.R., Objective picture quality scale (PQS) for image coding, *IEEE Transactions on Communications*, **46**(9), 1215-26 (1998)

[**Pinson 2003**] Pinson, M.H., Wolf, S., Comparing subjective video quality testing methodologies, *Proceedings of the SPIE - The International Society for Optical Engineering*, **5150**(1), 573-582 (2003)

[**Pearson 1998**] Pearson, D., Viewer response to time-varying video quality, *Proceedings of the SPIE - The International Society for Optical Engineering*, **3299**, 16-25 (1998)

[**Sharma 2003**] Sharma G., *Digital Color Imaging Handbook*, Chapter 5, CRC Press. Ed. 2003a.

[**Tan 1998**] Tan, K.T., Ghanbari, M., Pearson, D.E., An objective measurement tool for MPEG video quality, *Signal Processing*, **70**(3), 279-294 (1998)

[**Thurstone 1927**] Thurstone, L.L., A law of comparative judgments, *Psychological Review*, **34**, 273-287 (1927)

[**Tong 1999**] Xin T., Heeger, D., van den Branden Lambrecht C., Video quality evaluation using ST-CIELAB, *Proceedings of the SPIE - The International Society for Optical Engineering*, **3644**, 185-196 (1999)

[**van den Branden 1996**] van den Branden Lambrecht, C., A working spatio-temporal model of the human visual system for image restoration and quality assessment applications, *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, Atlanta, GA, 2293-2296, (1996)

[**VQEG 2000**] Rohaly A.M., Corriveau P.J., Libert J., Webster A.A., Baroncini V., Beerends J., Blin J., Contin L., Hamada T., Harrison D., Hekstra A.P., Lubin J., Nishida Y., Nishihara R., Pearson J.C., Pessoa A.F., Pickford N., Schertz A., Visca M., Watson A.B., Winkler S., Video Quality Experts Group: current results and future directions, *Proceedings of the SPIE - The International Society for Optical Engineering*, **4067**(1-3), 742-53 (2000)

[**VQEG 2003**] VQEG Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment – Phase II, Aug. 2003, Available at http://www.its.bldrdoc.gov/vqeg/

[**Watson 1999**] Watson A.B., Hu J., McGowan J.F., Mulligan J.B., Design and performance of a digital video quality metric, *Proceedings of SPIE - The International Society for Optical Engineering*, **3644**, 168-174 (1999)

[**Watson 2001**] Watson A.B., Hu J., McGowan J.F., Digital video quality metric based on human vision, *J Electronic Imaging*, **10**(1), 20-29 (2001)

[Weeks 1999] Weeks, A.R., Sartor, L.J., Myler, H.R., Histogram specification of 24-bit color images in the color difference (C-Y) color space, *Proceedings of SPIE - The International Society for Optical Engineering*, **3646**, 319-329 (1999)

[**Winkler-1 1999**] Winkler, S., A perceptual distortion metric for digital color video, *Proceedings of the SPIE - The International Society for Optical Engineering*, **3644**, 175-84 (1999)

[**Winkler-2 1999**] Winkler, S., Issues in vision modeling for perceptual video quality assessment, *Signal Processing*, **78**(2), 231-252 (1999)

[**Winkler 2001**] Winkler, S., Visual fidelity and perceived quality: toward comprehensive metrics, *Proceedings of the SPIE - The International Society for Optical Engineering*, **4299**, 114-125 (2001)

[**Wolf-1 1997**] Wolf, S., Measuring the end-to-end performance of digital video systems, IEEE Transactions on Broadcasting, **43**(3), 320-328 (1997)

[**Wolf-2 1997**] Wolf, S., Pinson, M. H., Webster, A.A., Cermak, G.W., Tweedy, E.P., Objective and subjective measures of MPEG video quality, *Proc. Society of Motion Picture and Television Engineers*, 160–178 (1997)

[**Wolf-3 1999**] Wolf, S., Pinson, M.H., Spatial-temporal distortion metric for in-service quality monitoring of any digital video system, *Proceedings of the SPIE - The International Society for Optical Engineering*, **3845**, 266-277 (1999)

[**Wu 1997**] Wu, H.R., Yuen, M., A generalized block-edge impairment metric for video coding, *IEEE Signal Processing Letters*, **4**(11), 317-320 (1997)

[Wu 1998] Tuo, W., Joyce, F., Amnon, S., Qualifying Image Quality from Pairwised Comparisons, *International Conference on Digital Printing Technologies*, 564-567 (1998)

[**Wu 2006**] Wu H.R., Rao K.R., *Digital Video Image Quality and Perceptual Coding*, CRC Press, Taylor & Francis Group, Florida, USA (2006)

[**Yu 2000**] Yu, Z., Wu, H.R., Human visual system based objective digital video quality metrics, 5th International Conference on Signal Processing Proceedings and 16th World Computer Congress, **2**(2), 1088-95 (2000)

[**Zhang 1996**] Zhang, X., Wandell, B.A., A spatial extenstion of CIELAB for digital color-image reproduction, *Proc. of the SID symposiums* (1996)

[**Zhao 2005**] Zhao, M., de Haan, G., Subjective evaluation of de-interlacing techniques, *Proceedings of SPIE-IS&T Electronic Imaging - Image and Video Communications and Processing*, **5685**(2), 683-691 (2005)

[**Choi 2007**] Choi, D.H., Jang, I.H., Kim, M.H., Kim, N.C., Color image enhancement based on single-scale retinex with a JND-based nonlinear filter, *IEEE International Symposium on Circuits and Systems*, 3948-3951 (2007)

[Colantoni 2004] Colantoni, P., Bost, N., Tremeau, A., Colorfulness enhancement in λSY color space, *CGIV 2004 - Second European Conference on Color in Graphics, Imaging, and Vision and Sixth International Symposium on Multispectral Color Science*, 161-166 (2004)

[Ebner 1998] Ebner, F., Fairchild, M.D., Development and testing of a color space (IPT) with improved hue uniformity, *Proceedings of the Color Imaging Conference: Color Science, Systems, and Applications*, 8-13 (1998)

[Hague 1994] Hague, G.E., Weeks, A.R., Myler, H.R., Histogram equalization of the saturation component for true-color images using the C-Y color space, *Proceedings of SPIE - The International Society for Optical Engineering*, **2298**, 236-247 (1994)

[**Hsu 2006**] Hsu T., Liu, C.T., Kuo-Jui H., Improved Retinex approach for color image enhancement, *CGIV 2006 - 3rd European Conference on Colour in Graphics, Imaging, and Vision*, 390-393 (2006)

[**Land 1983**] Land E, Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image, *Proc. Natl. Acad. Sci. USA*, **80**, 5163-5169 (1983)

[**Land 1986**] Land, E.H., Recent advances in Retinex Theory, *Vision Research*, **26**(1), 7-21 (1986)

[Lucchese 2001] Lucchese, L., Mitra, S.K., Mukherjee, J., A new algorithm based on saturation and desaturation in the xy chromaticacity diagram for enhancement and re-rendition of color images, *IEEE International Conference on Image Processing*, **2**, 1077-1080 (2001)

[**Meylan 2004**] Meylan, L., Susstrunk, S., Color image enhancement using a Retinex-based adaptive filter, *CGIV 2004 – Second European Conference on Color in Graphics, Imaging, and Vision and Sixth International Symposium on Multispectral Color Science*, 359-363 (2004)

[Moroney 2000] Moroney, N., Local color correction using non-linear masking, *Proceedings of the Color Imaging Conference: Color Science, Systems, and Applications*, 108-111 (2000)

[**Rahman 1996**] Rahman, Z., Jobson, D.J., Woodell, G.A., Multi-scale retinex for color image enhancement, *Proceedings of International Conference on Image Processing*, **3**, 1003-1006 (1996)

[**Rahman 2002**] Rahman, Z., Jobson, D.J., Woodell, G.A., Retinex processing for automatic image enhancement, *Proceedings of SPIE - The International Society for Optical Engineering*, **4662**, 390-401 (2002)

[Samadani 2006] Samadani, R., Li, G., Geometrical methods for lightness adjustment in YCC color spaces, *Color Imaging XI: Processing, Hardcopy, and Applications - Proceedings of SPIE-IS&T Electronic Imaging*, **6058**, 605809 (2006)

[Samadani 2007] Samadani, R; Personal communication (February 22, 2007)

[**SMIL 2005**] Apple Developer Connection; Introduction to SMIL scripting guide for QuickTime; available on http://developer.apple.com/documentation/QuickTime/index.html, last updated June 4, 2005, accessed March 24, 2008

[Strickland 1986] Strickland, R.N., Kim, C.S., McDonnell, W.F., Luminance, hue, and saturation processing of digital color images, *Proceedings of SPIE - The International Society for Optical Engineering*, **697**, 286-292 (1986)

[Strickland 1987] Strickland, R.N., Kim, C.S., McDonnell, W.F., Digital color image enhancement based on the saturation component, *Optical Engineering*, **26**(7), 609-616 (1987)

[Tao 2004] Tao, L., Asari, V., An integrated neighborhood dependent approach for nonlinear enhancement of color images, *International Conference on Information Technology: Coding and Computing*, 138-139 (2004)

[**US Patent 4831434**] Fuchsberger, H., AGFA Gevaert Aktiengesellschaft, Method of correcting color saturation in electronic image processing, US Patent 4,831,434 (May 16, 1989)

[**US Patent 5282021**] Bachmann, P., Christmann, M., Poetsch, D., BTS Broadcast Television Systems, Video hue correction taking account of saturation and luminance, US Patent 5,282,021 (January 25, 1994)

[**US Patent 6028646**] Jeong, J.K., Song, W.J., LG Electronics Inc., Color image enhancement device for video display appliance, US Patent 6,028,646 (February 22, 2000)

[**US Patent 2006//0238655**] Chou, C.H., Pinnacle Patent Law Group, Method and system for automatic color hue and color saturation adjustment of a pixel from a video source, US Patent 2006//0238655 (Oct 26, 2006)

[**US Patent 2007//0070369 A1**] Wang, D., Pixelworks Inc, Processing apparatus and method for enhancing image color, US Patent 2007//0070369 A1 (Mar 29, 2007)

[Yang 1995] Yang, C.C., Rodriguez, J.J., Efficient luminance and saturation processing techniques for bypassing color coordinate transformations, *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, **1**, 667-672 (1995)

[Yang 1996] Yang, C. C., Rodriguez, J.J., Saturation clipping in the LHS and YIQ color spaces, *Proceedings of SPIE - The International Society for Optical Engineering*, **2658**, 297-307 (1996)

Appendix A: Algorithm Performance Analysis Plots

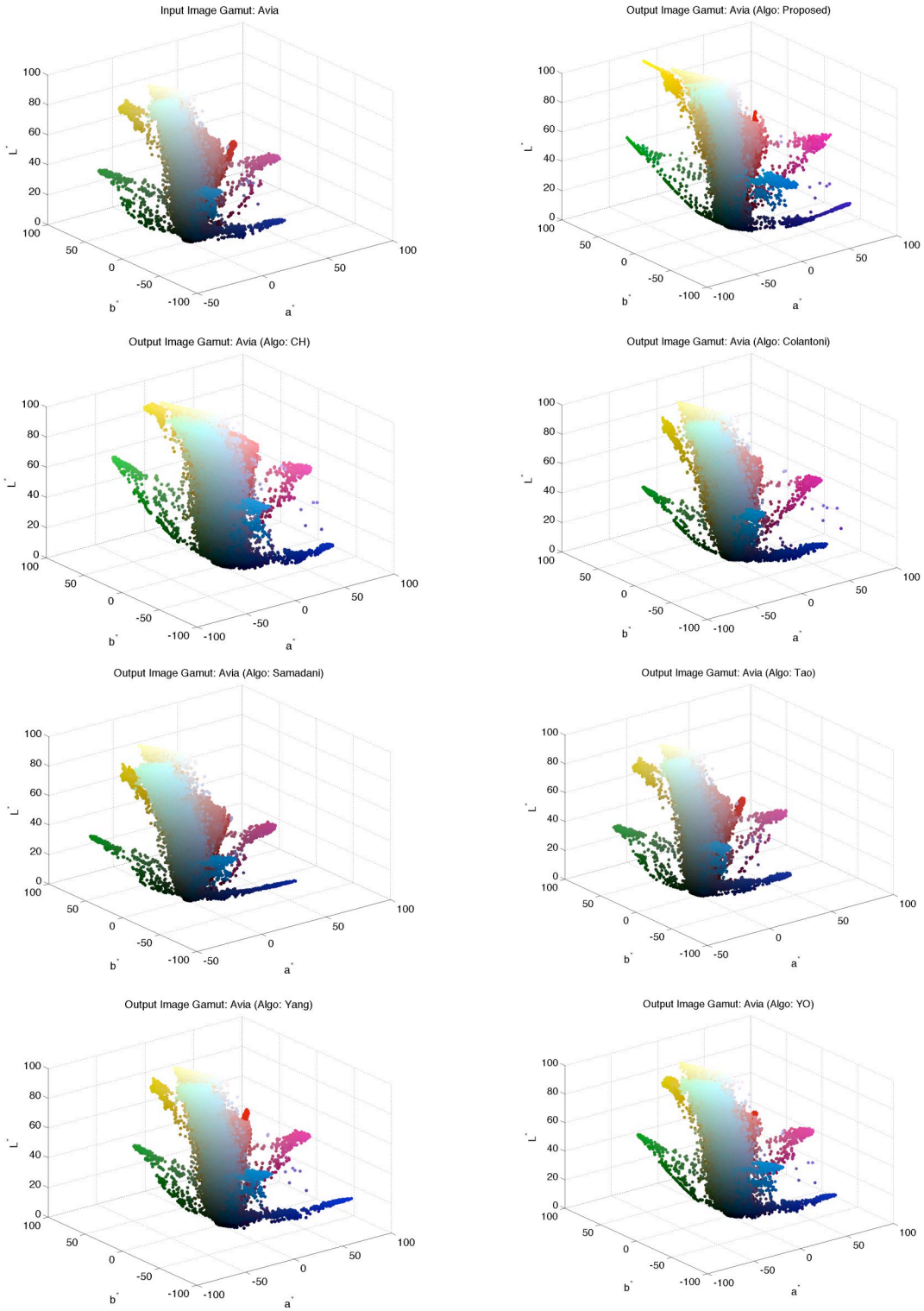


Fig. A.1 Image Gamut Maps: Avia

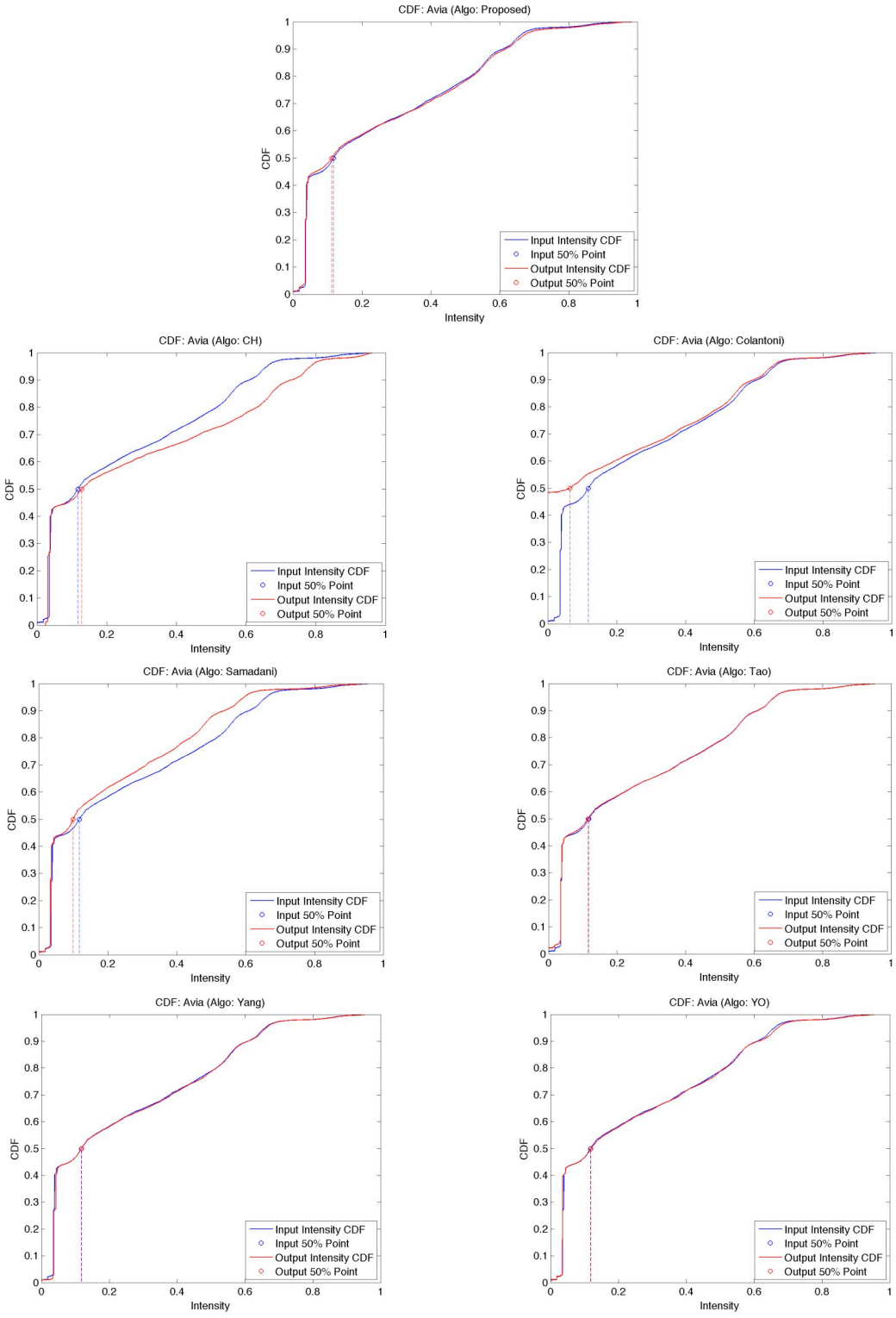


Fig. A.2 Cumulative Distribution Functions: Avia

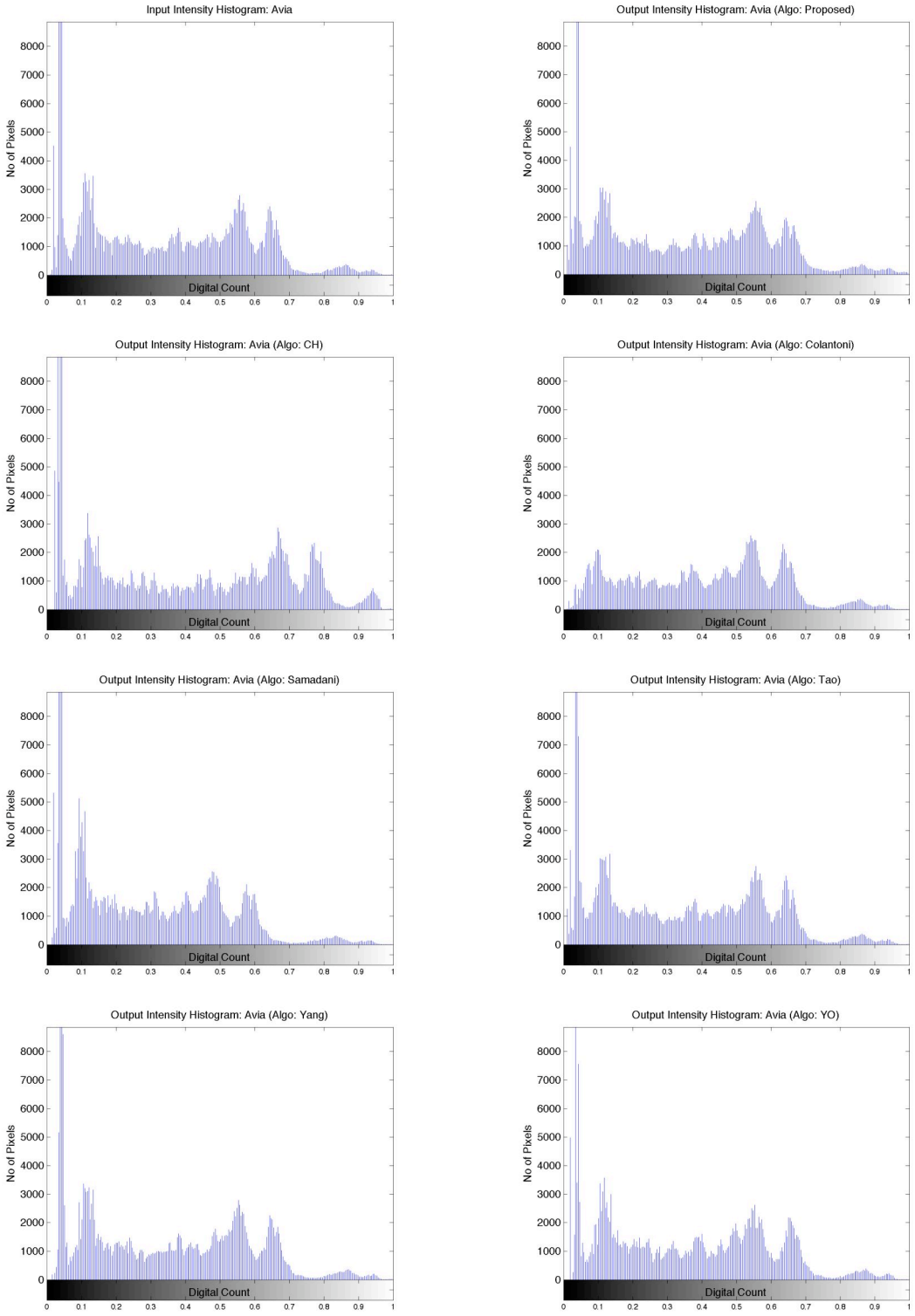


Fig. A.3 Intensity Histograms: Avia

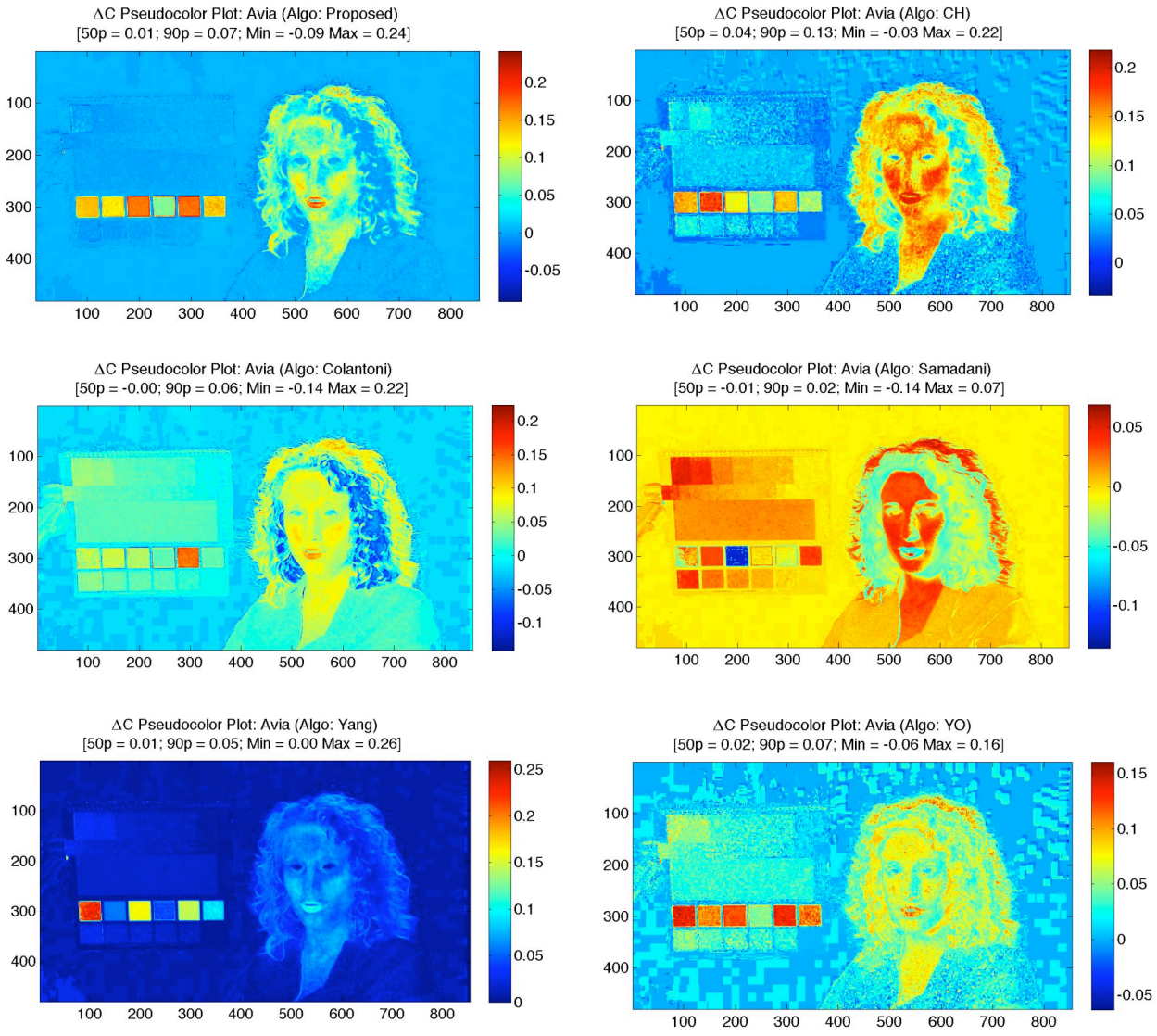


Fig. A.4 ΔC Image Difference Maps: Avia

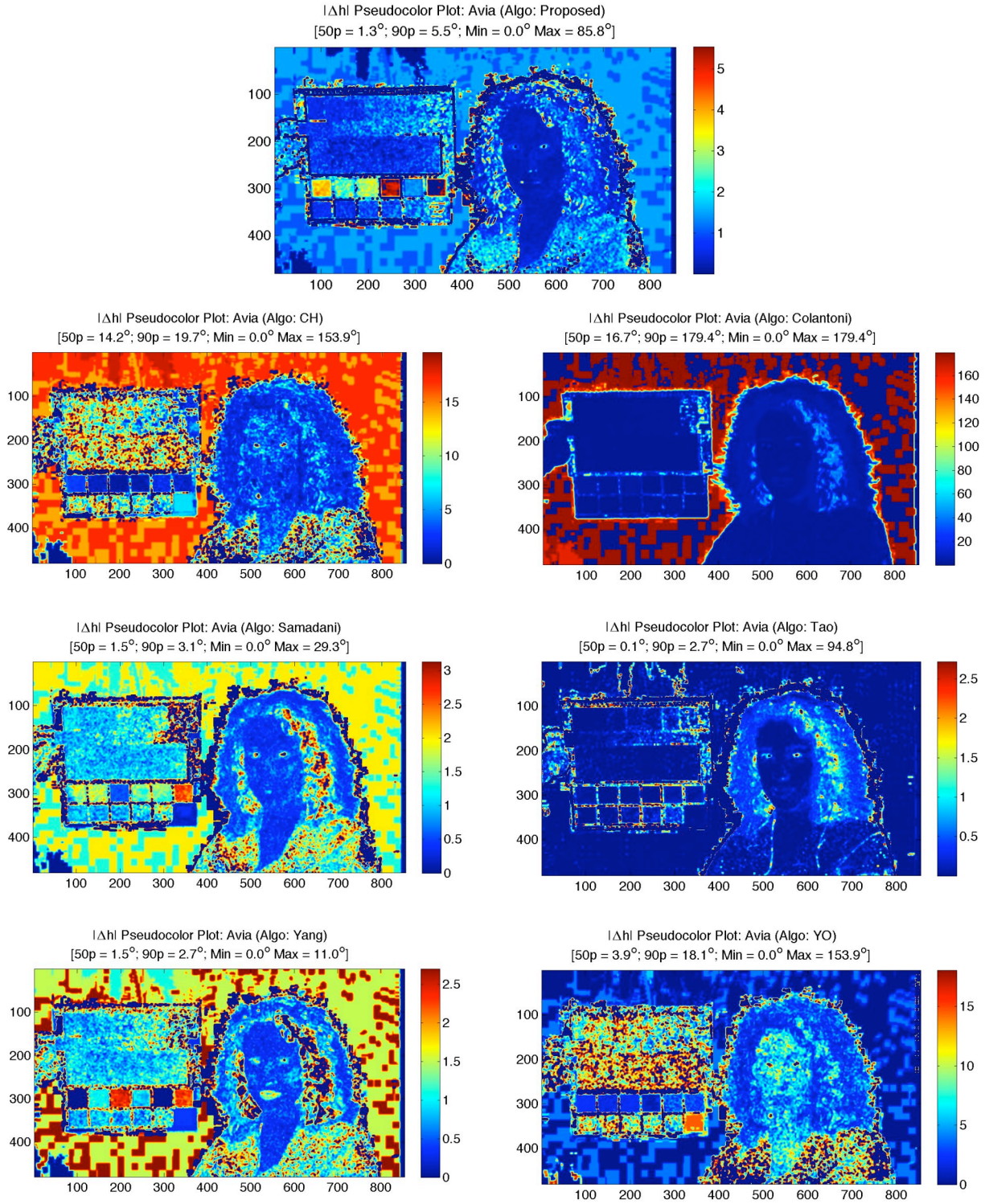


Fig. A.5 Δh Image Difference Maps: Avia

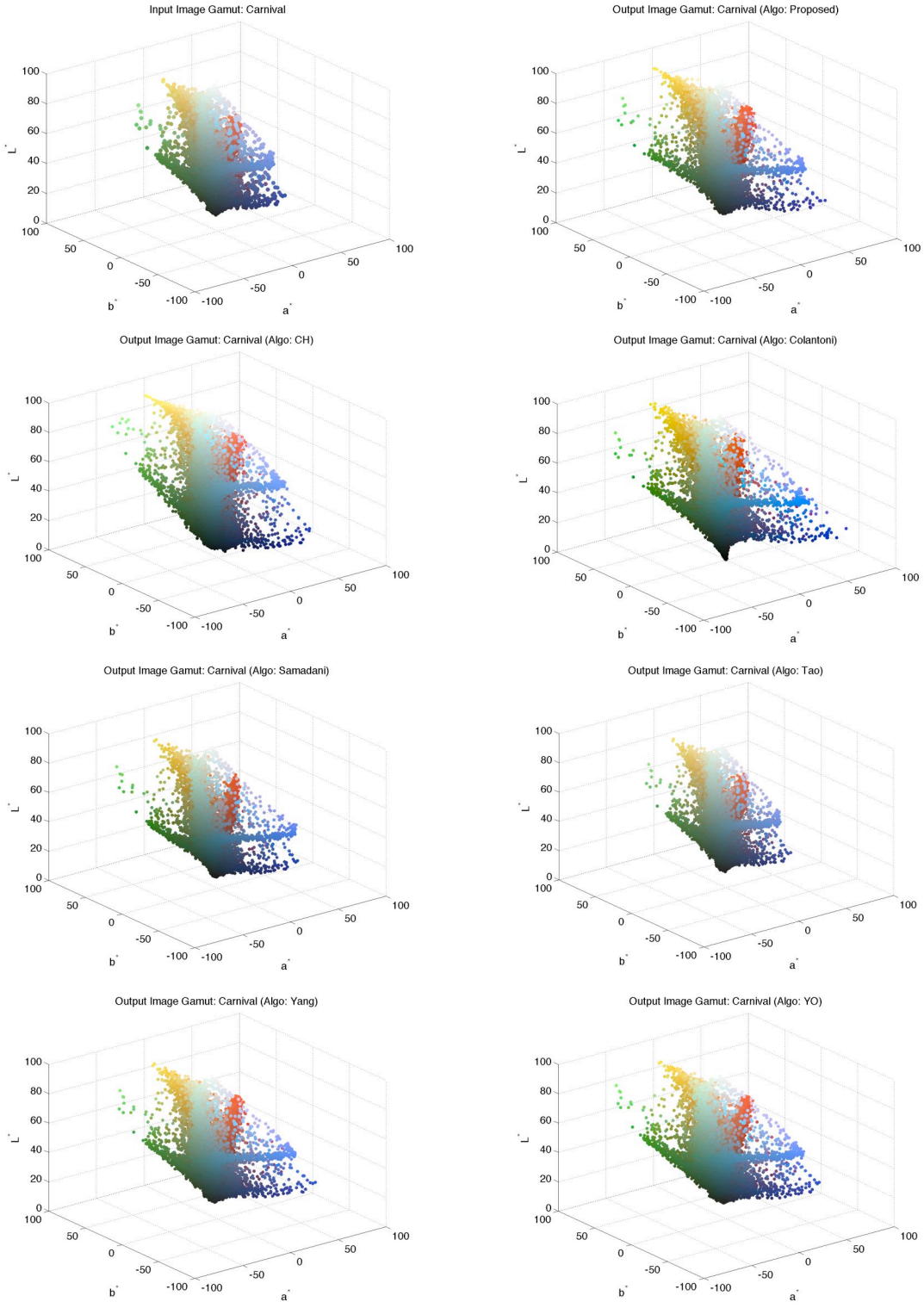


Fig. A.6 Image Gamut Maps: Carnival

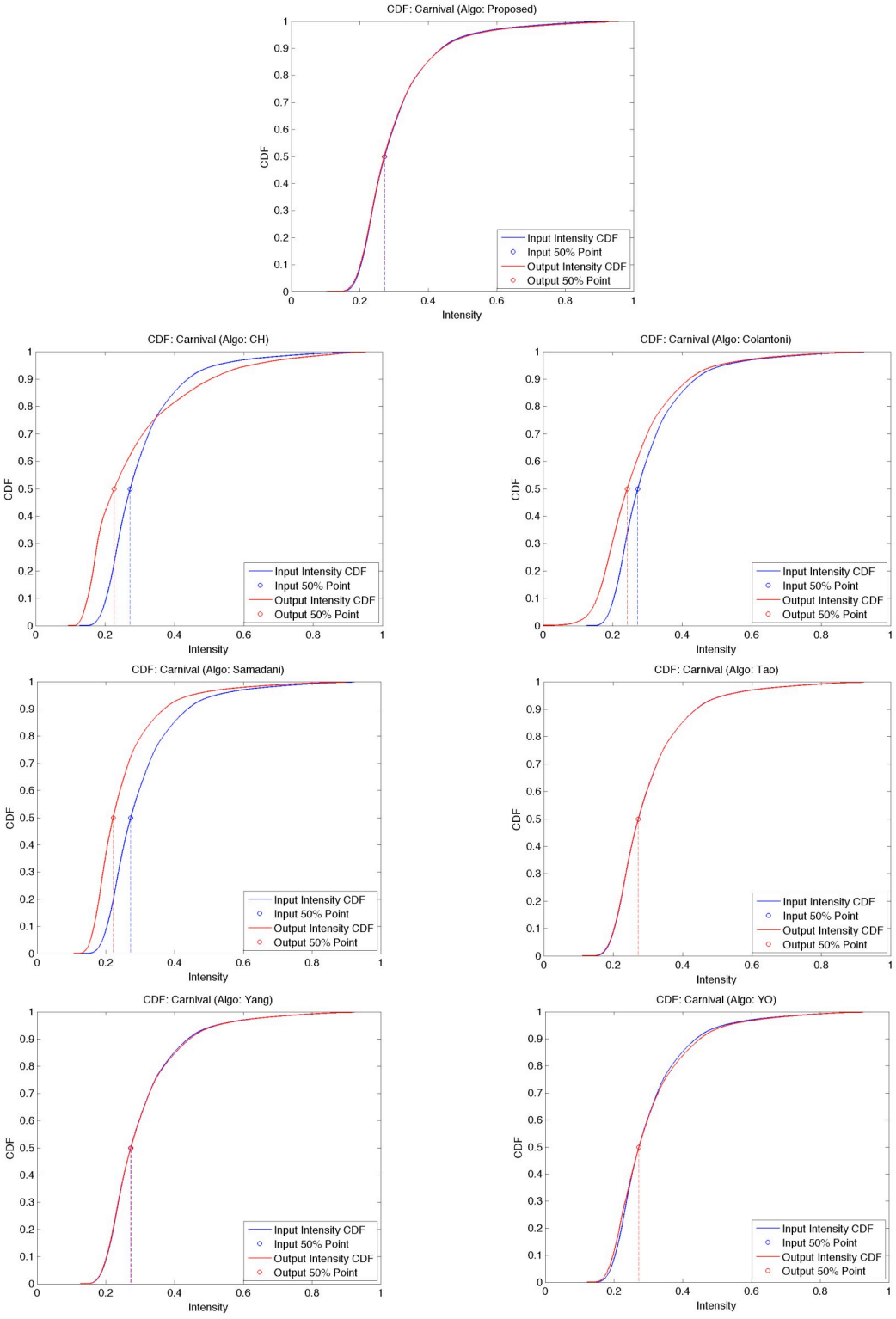


Fig. A.7 Cumulative Distribution Functions: Carnival

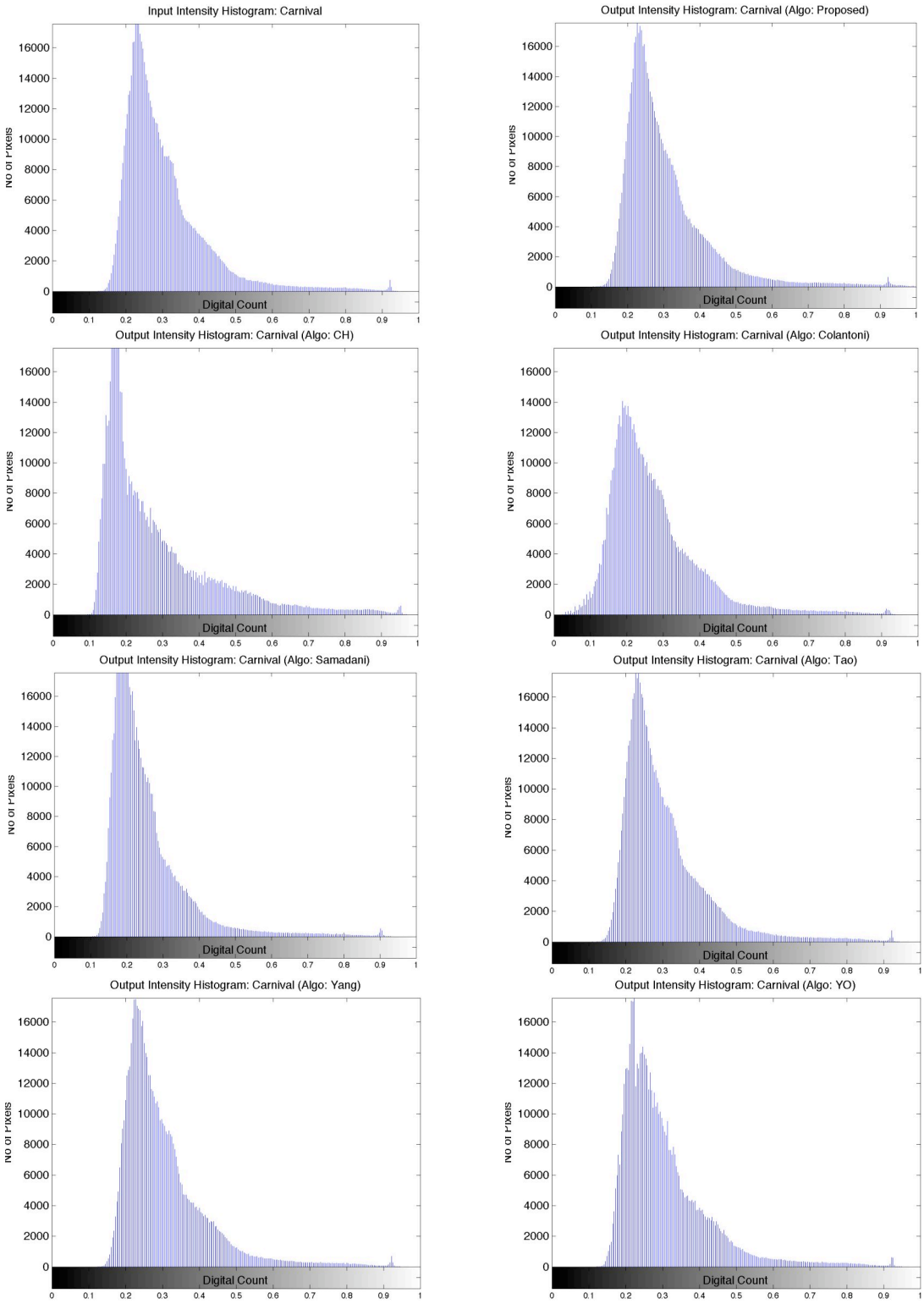


Fig. A.8 Intensity Histograms: Carnival

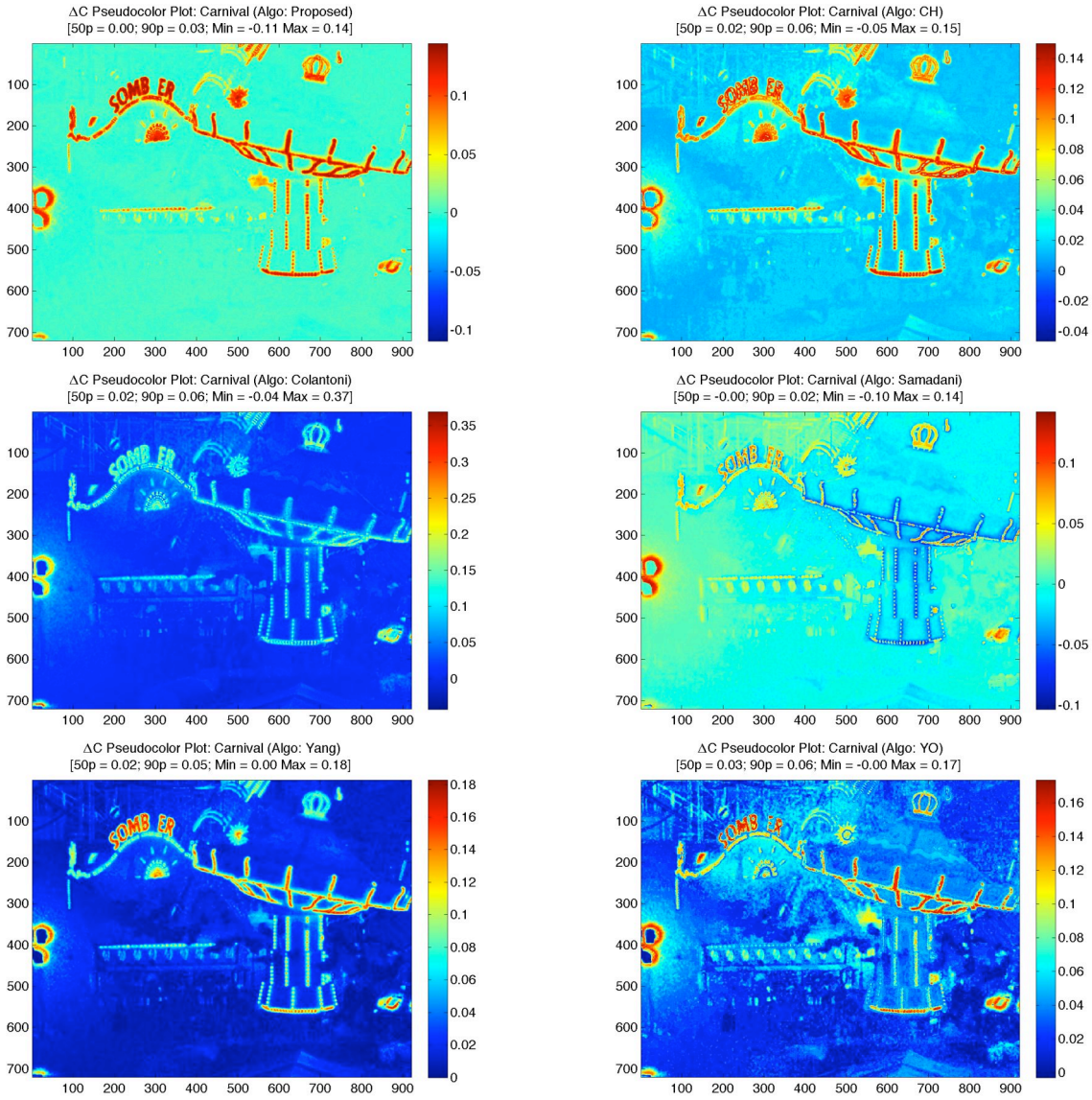


Fig. A.9 ΔC Image Difference Maps: Carnival

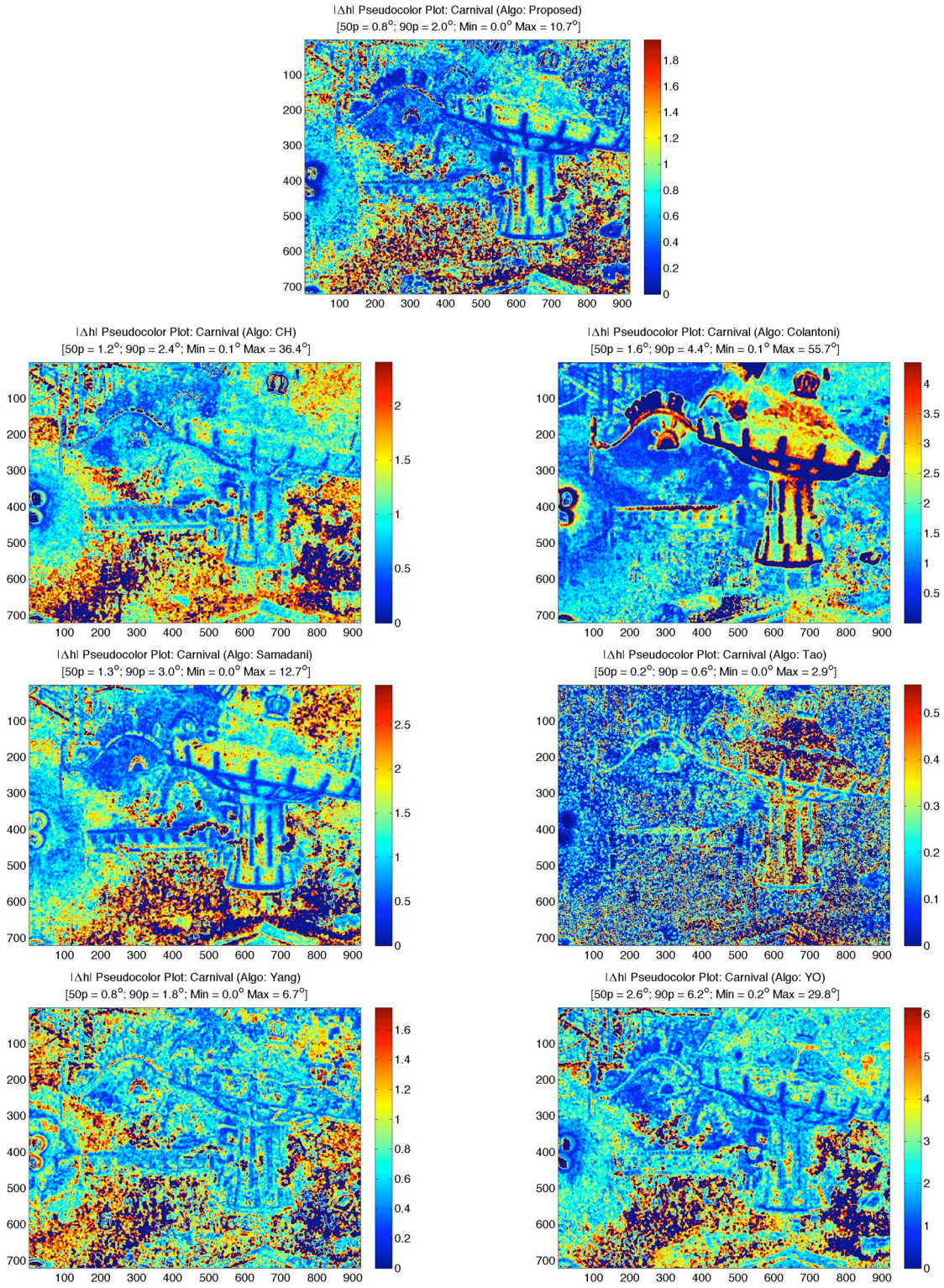


Fig. A.10 Δh Image Difference Maps: Carnival

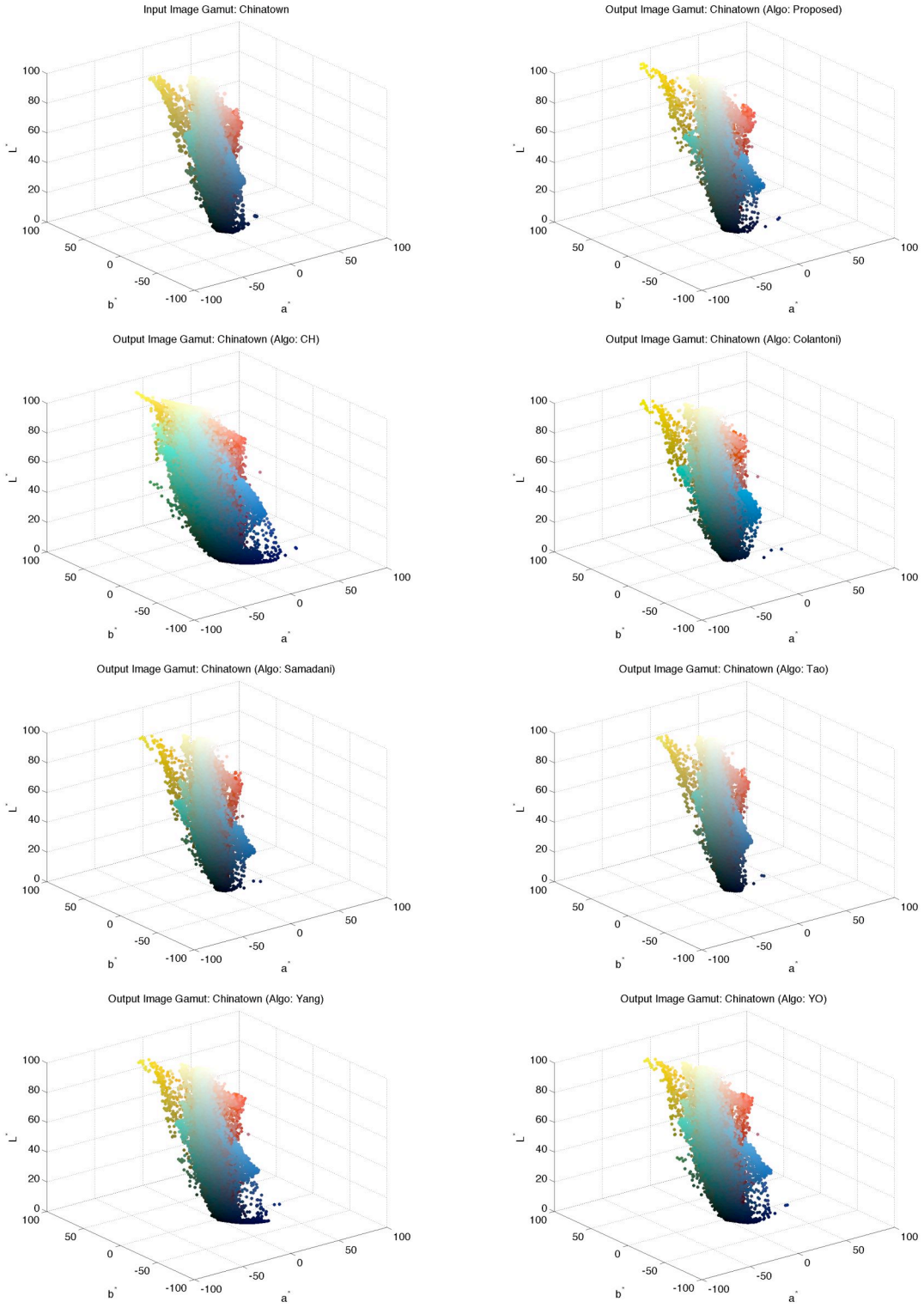


Fig. A.11 Image Gamut Maps: Chinatown

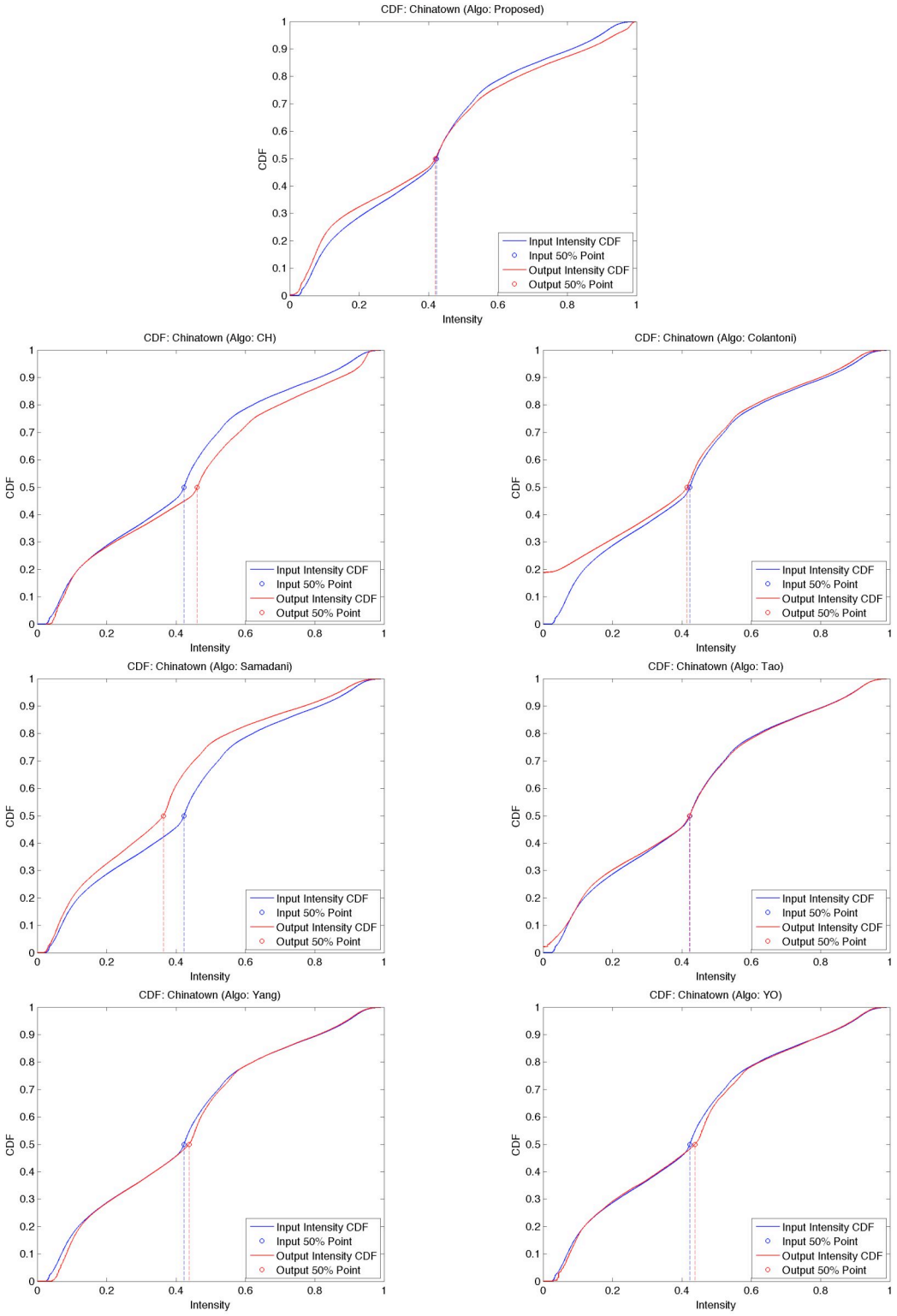


Fig. A.12 Cumulative Distribution Functions: Chinatown

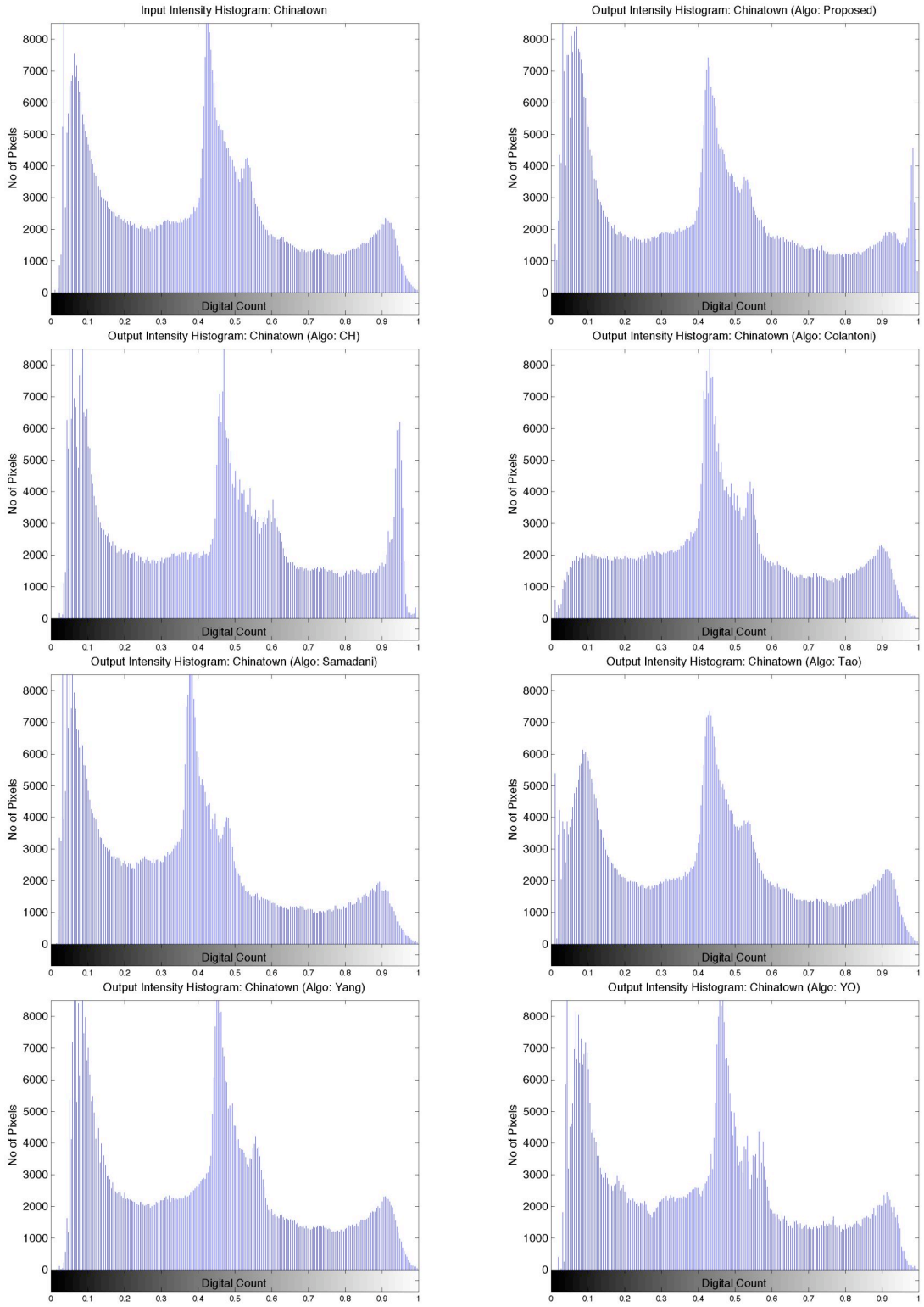


Fig. A.13 Intensity Histograms: Chinatown

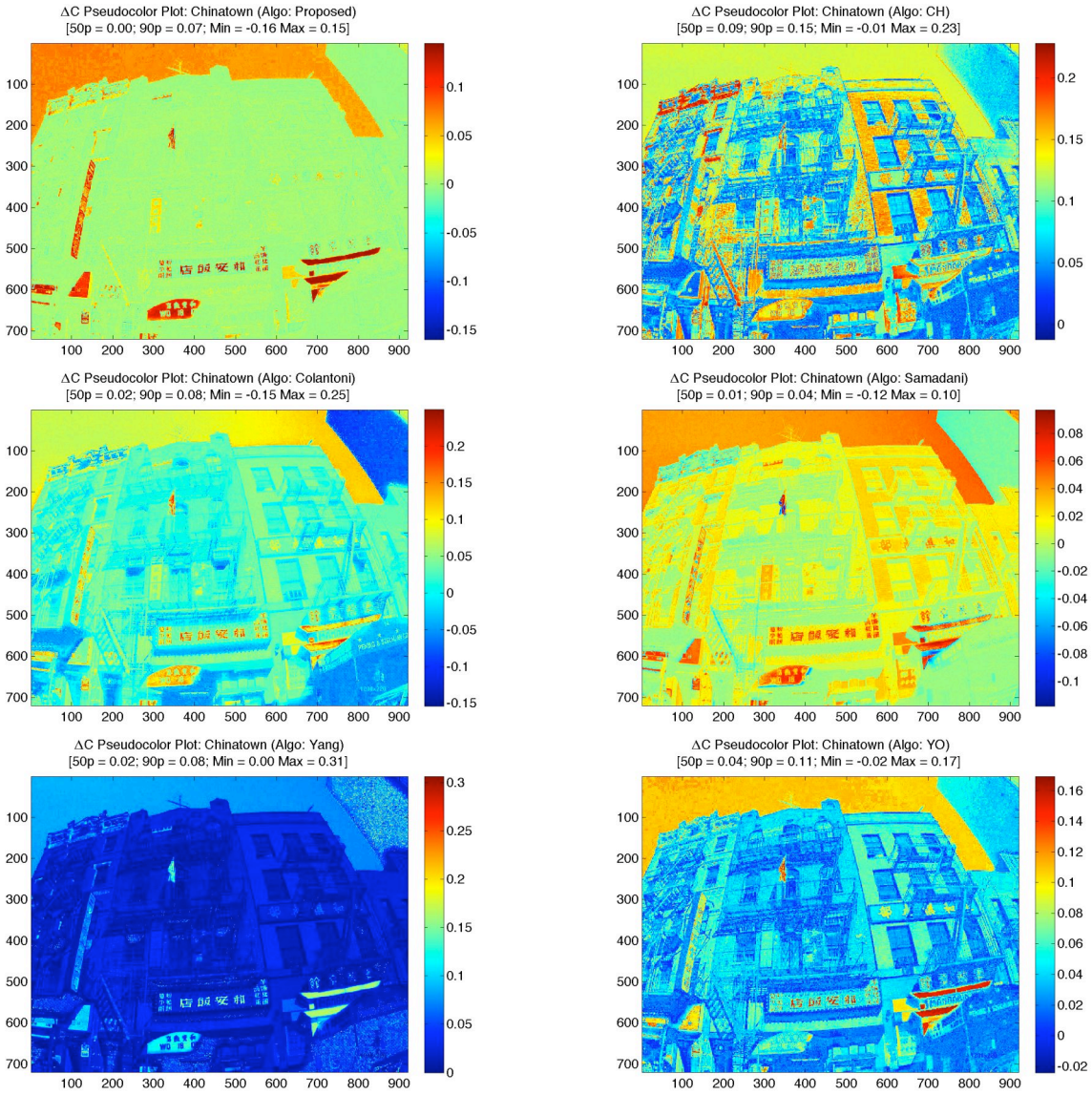


Fig. A.14 ΔC Image Difference Maps: Chinatown

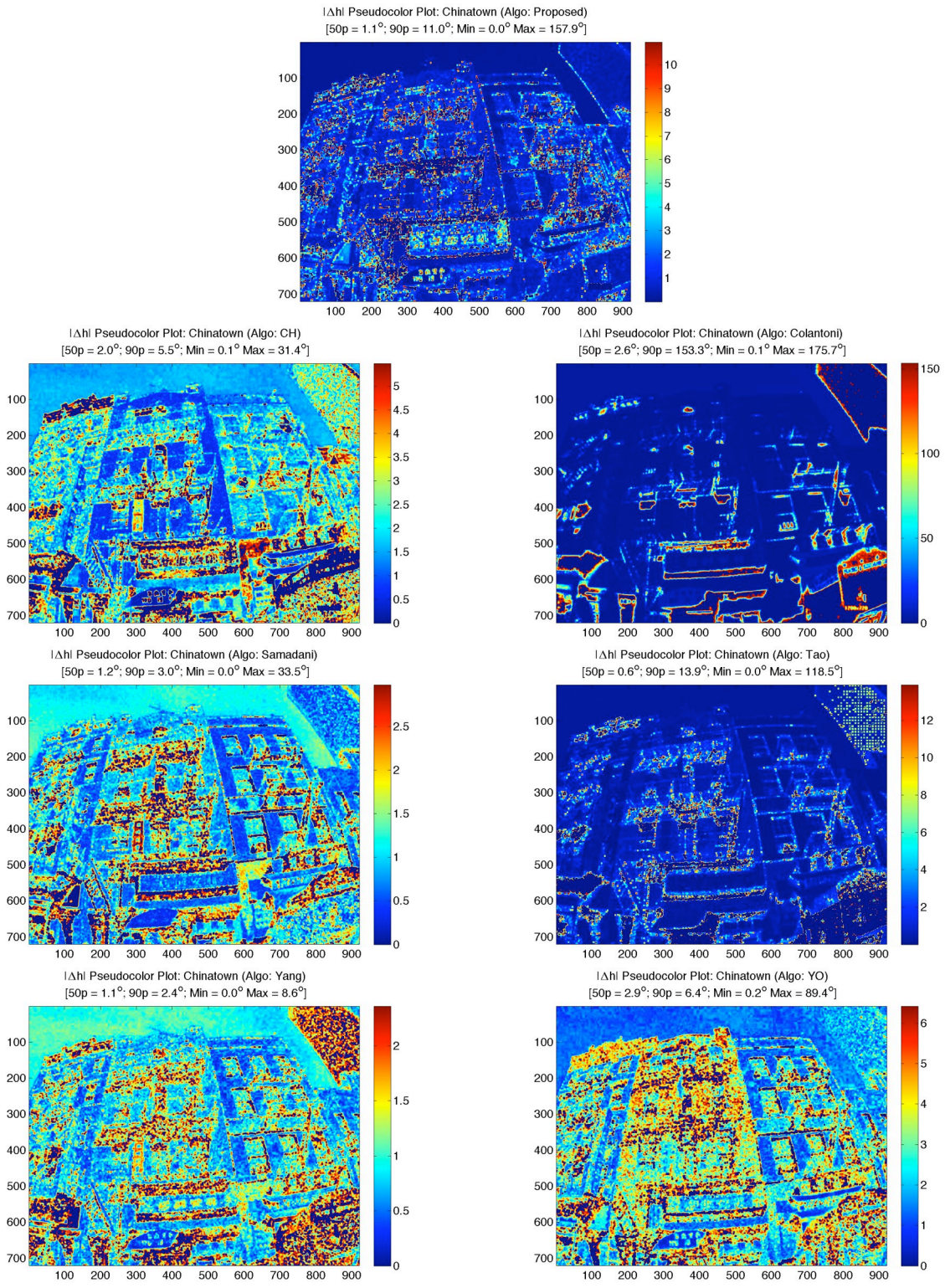


Fig. A.15 Δh Image Difference Maps: Chinatown

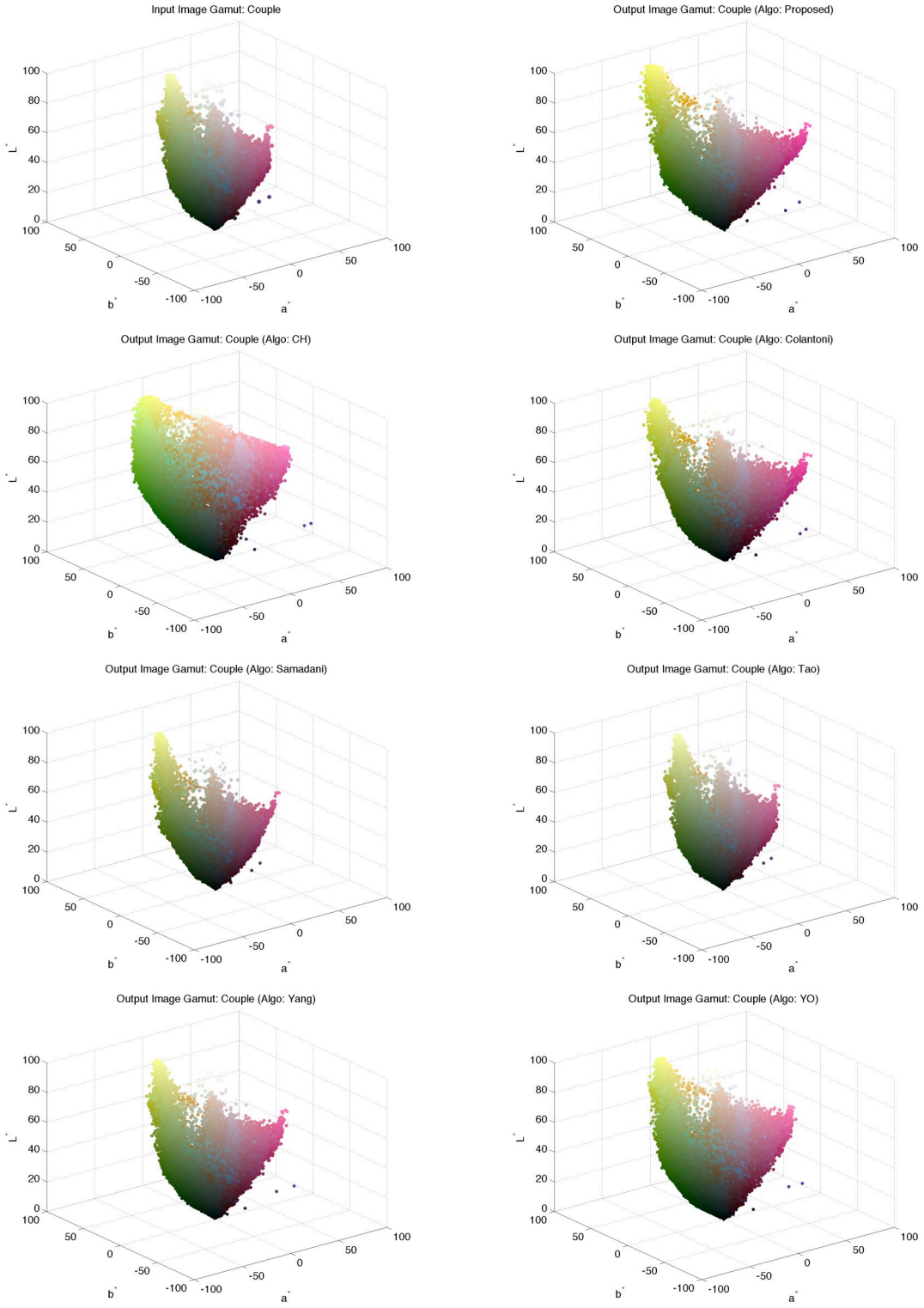


Fig. A.16 Image Gamut Maps: Couple

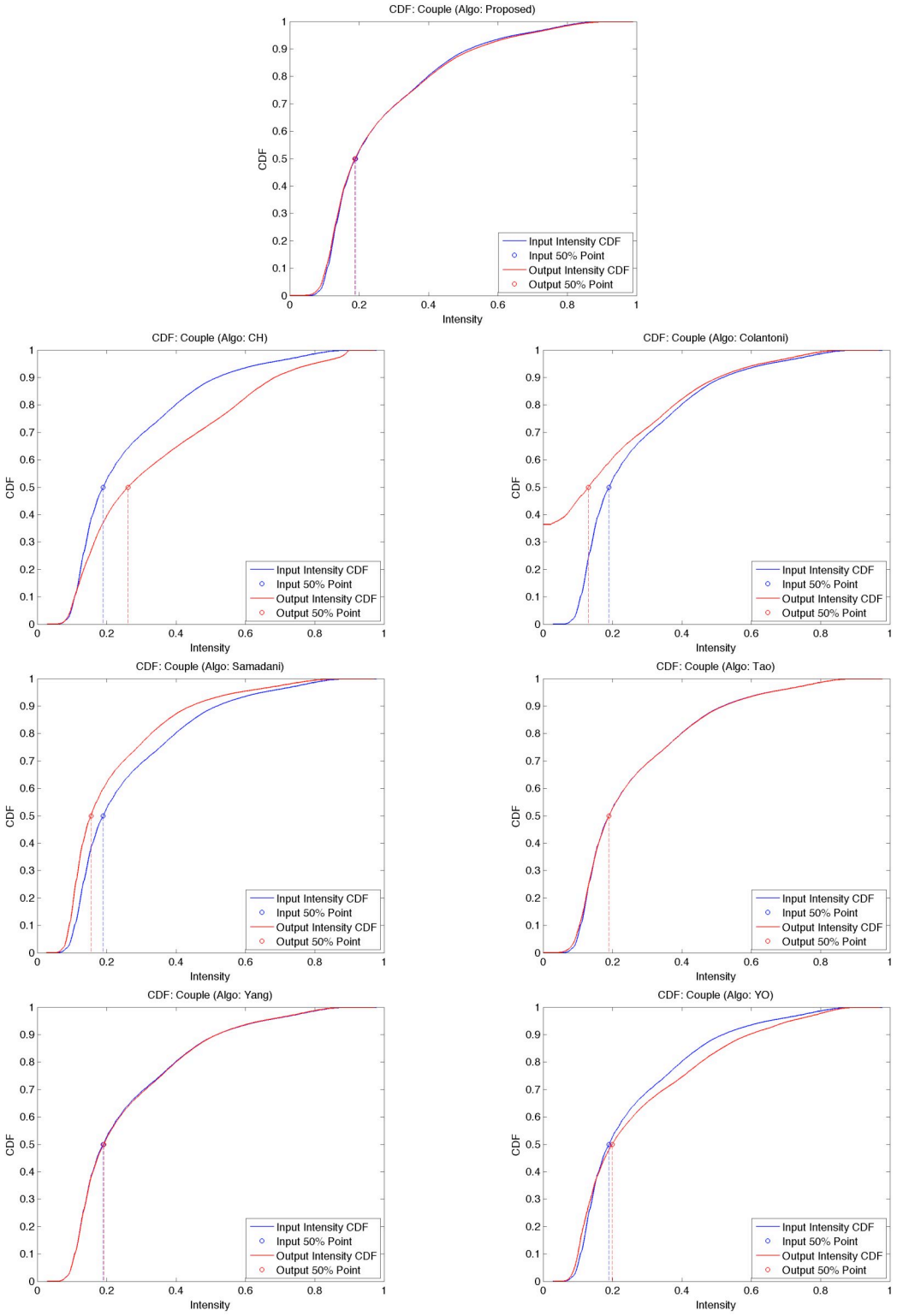


Fig. A.17 Cumulative Distribution Functions: Couple

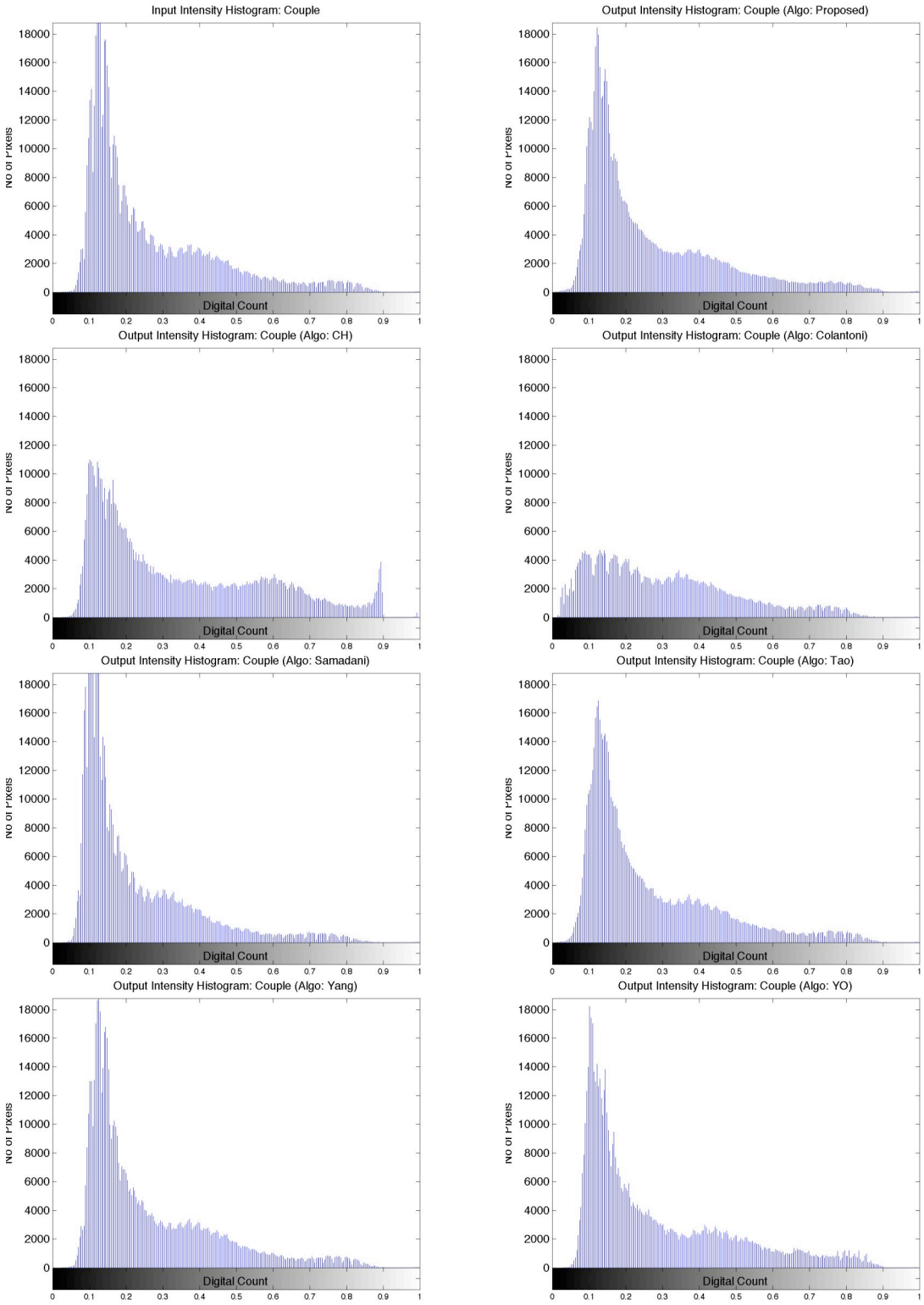


Fig. A.18 Intensity Histograms: Couple

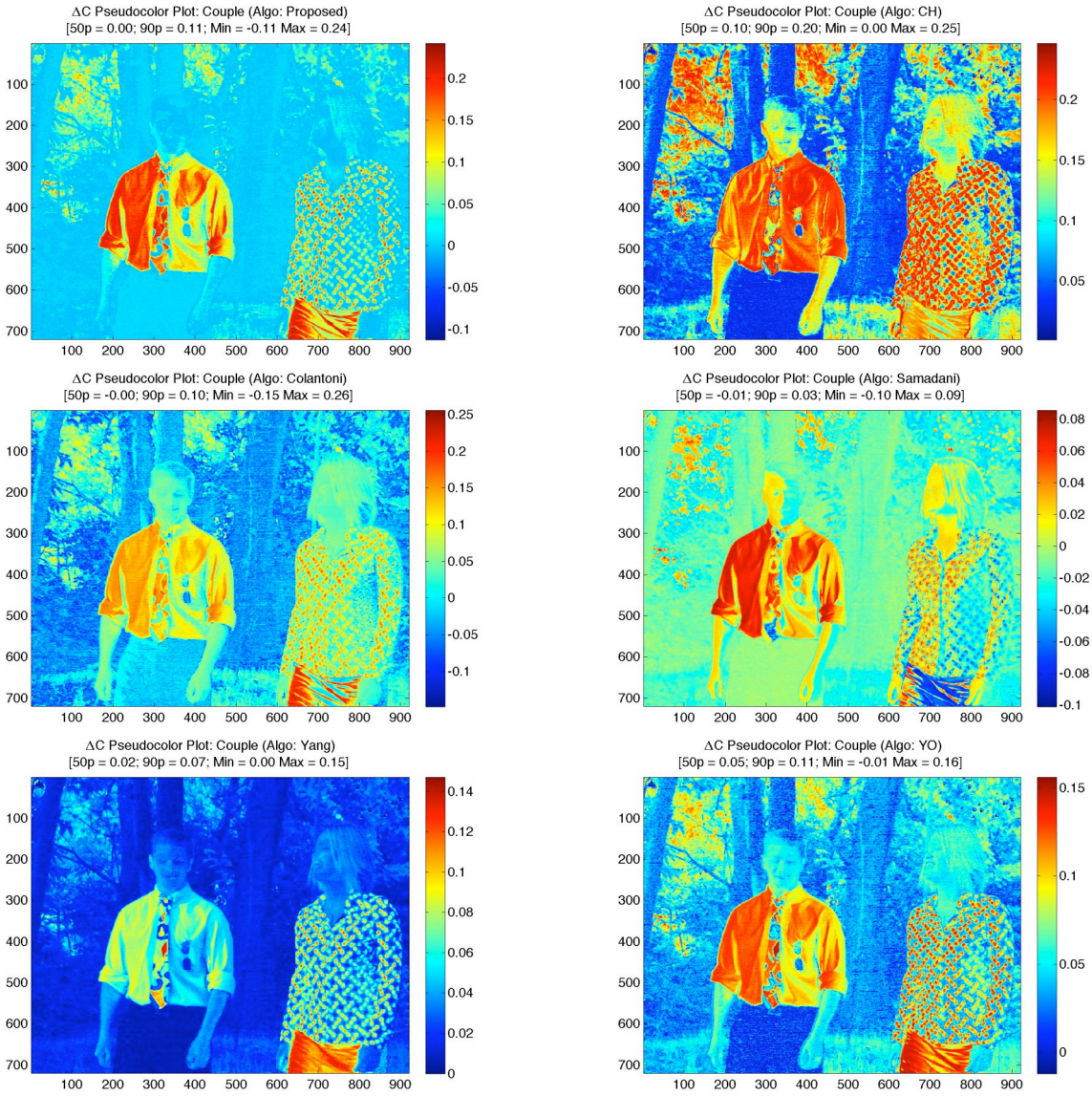


Fig. A.19 ΔC Image Difference Maps: Couple

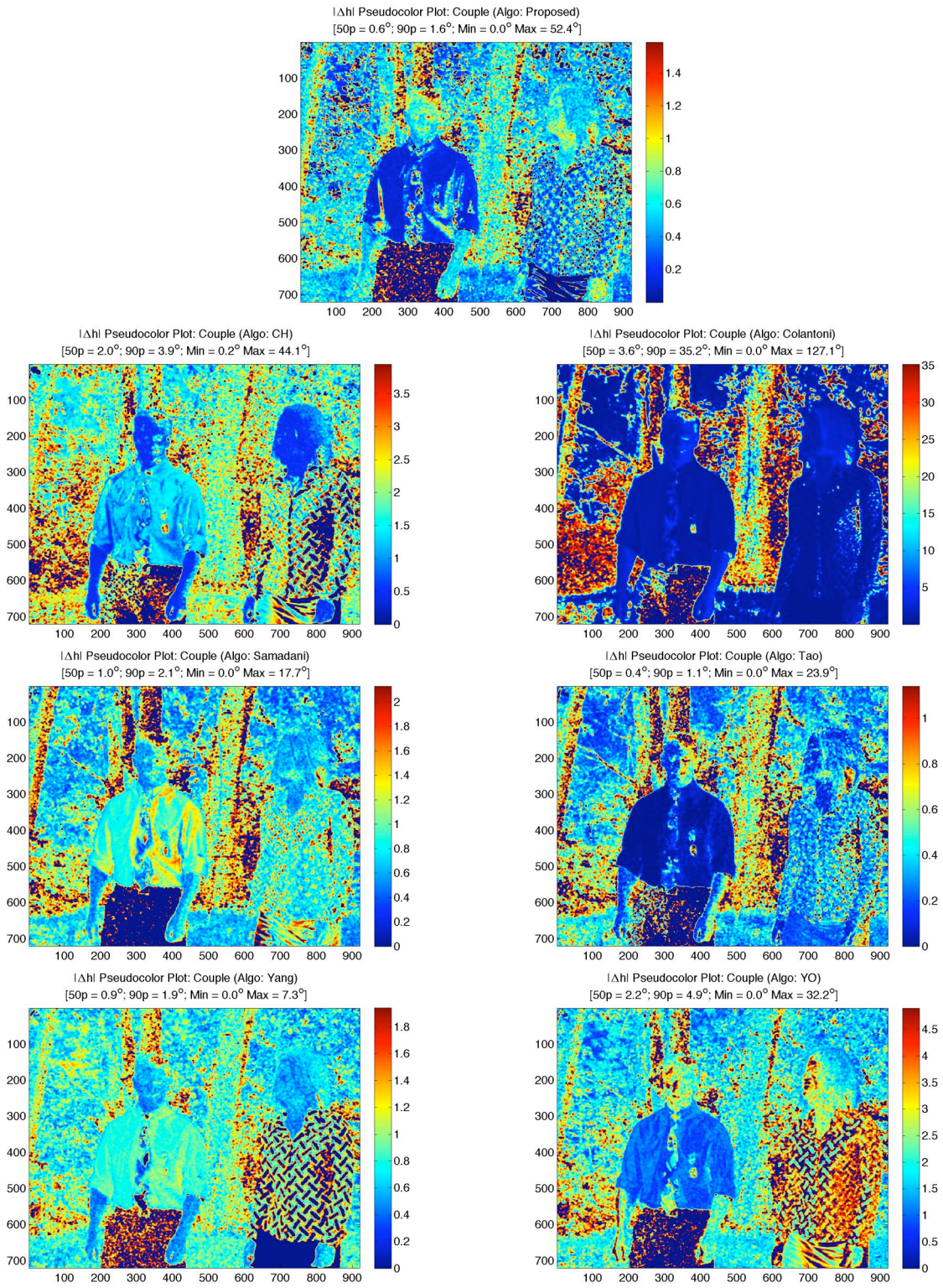


Fig. A.20 Δh Image Difference Maps: Couple

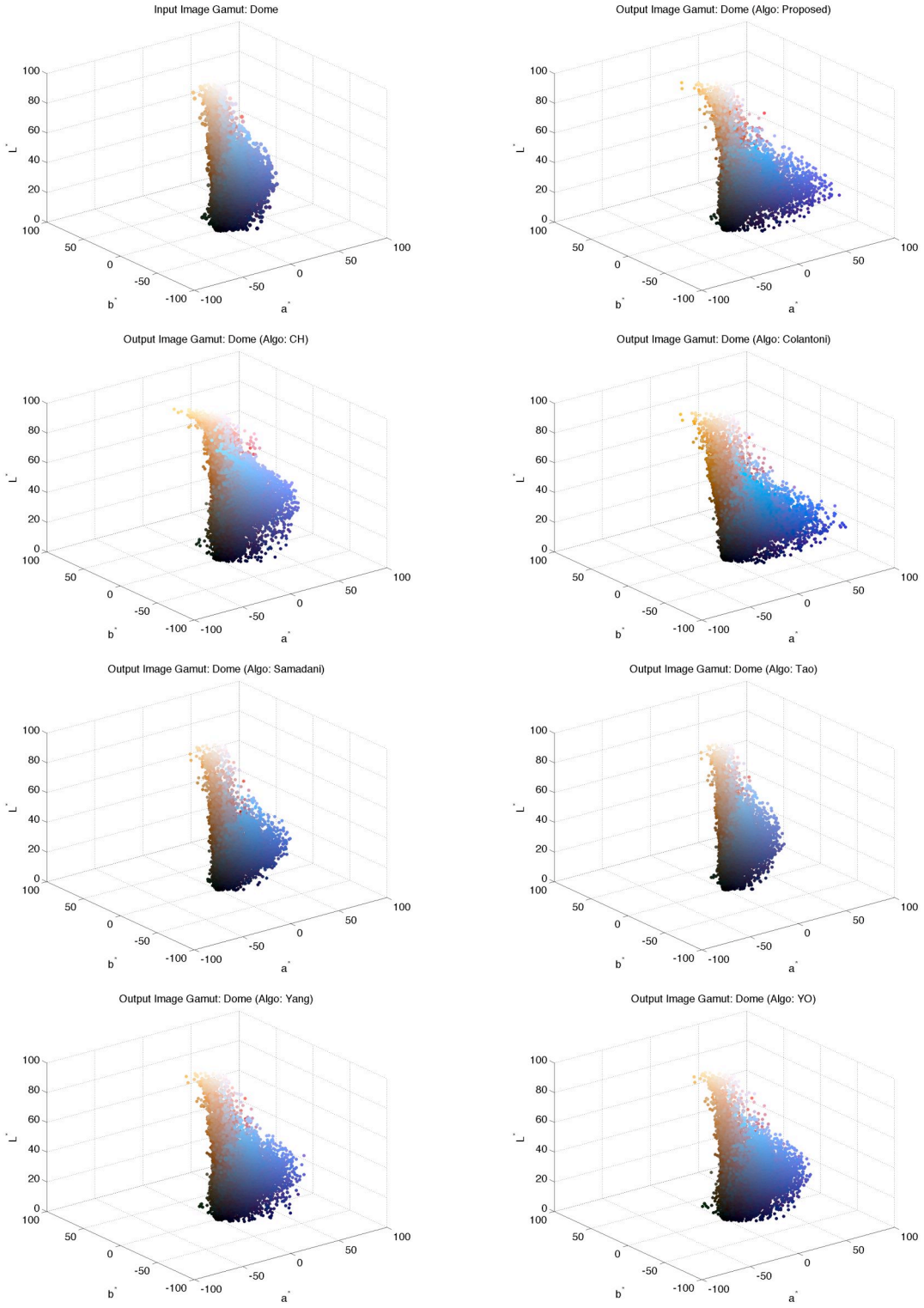


Fig. A.21 Image Gamut Maps: Dome

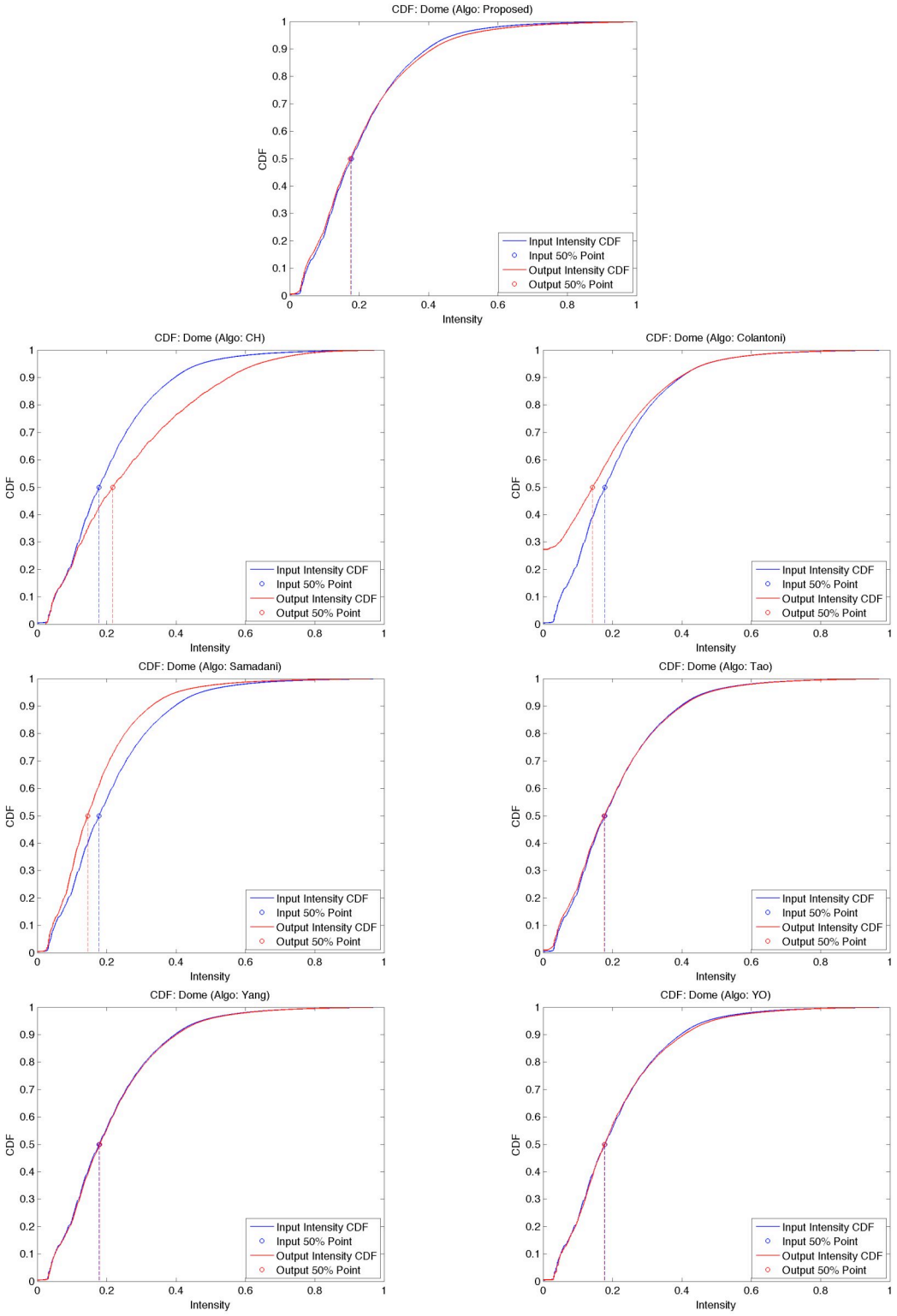


Fig. A.22 Cumulative Distribution Functions: Dome

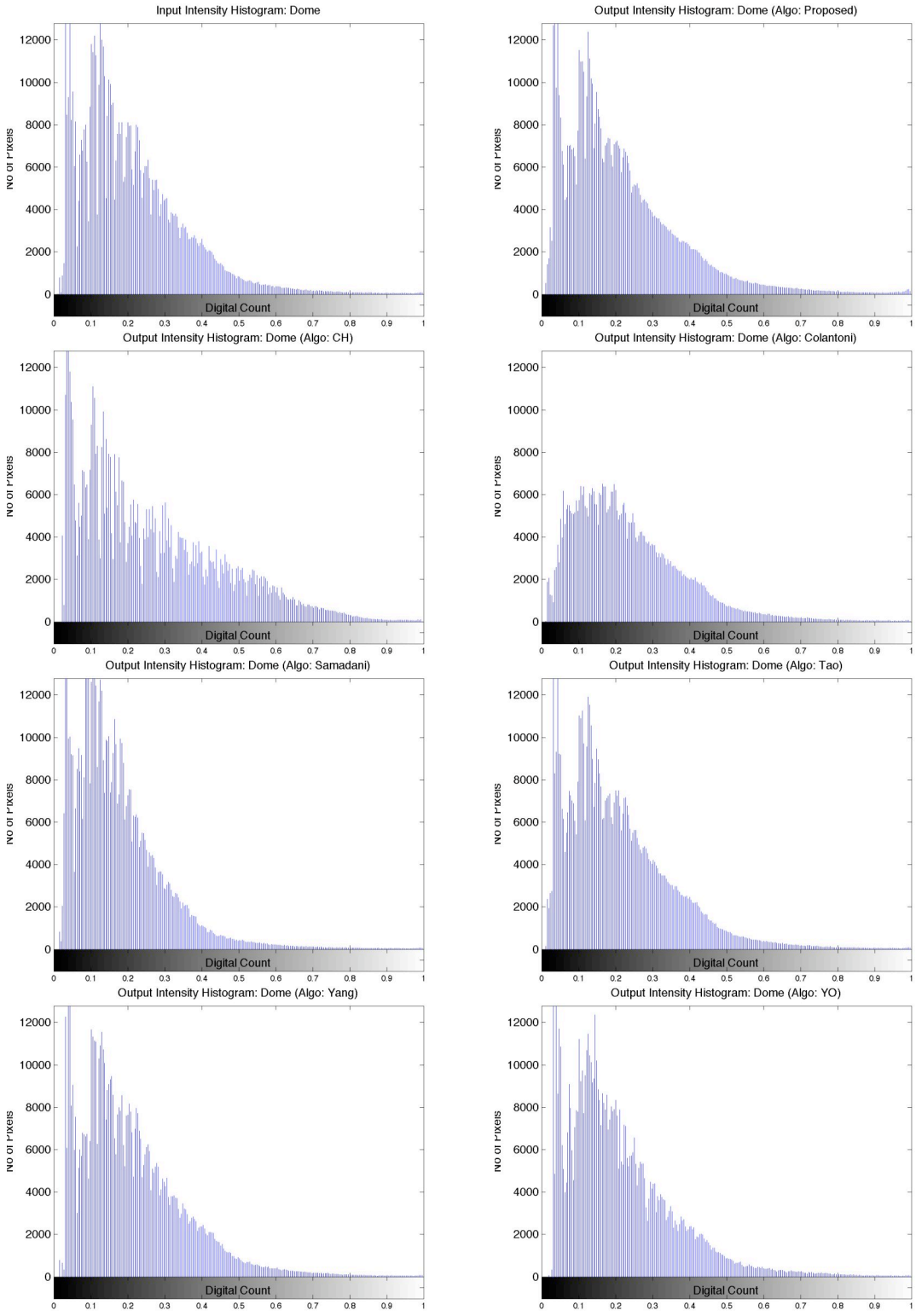


Fig. A.23 Intensity Histograms: Dome

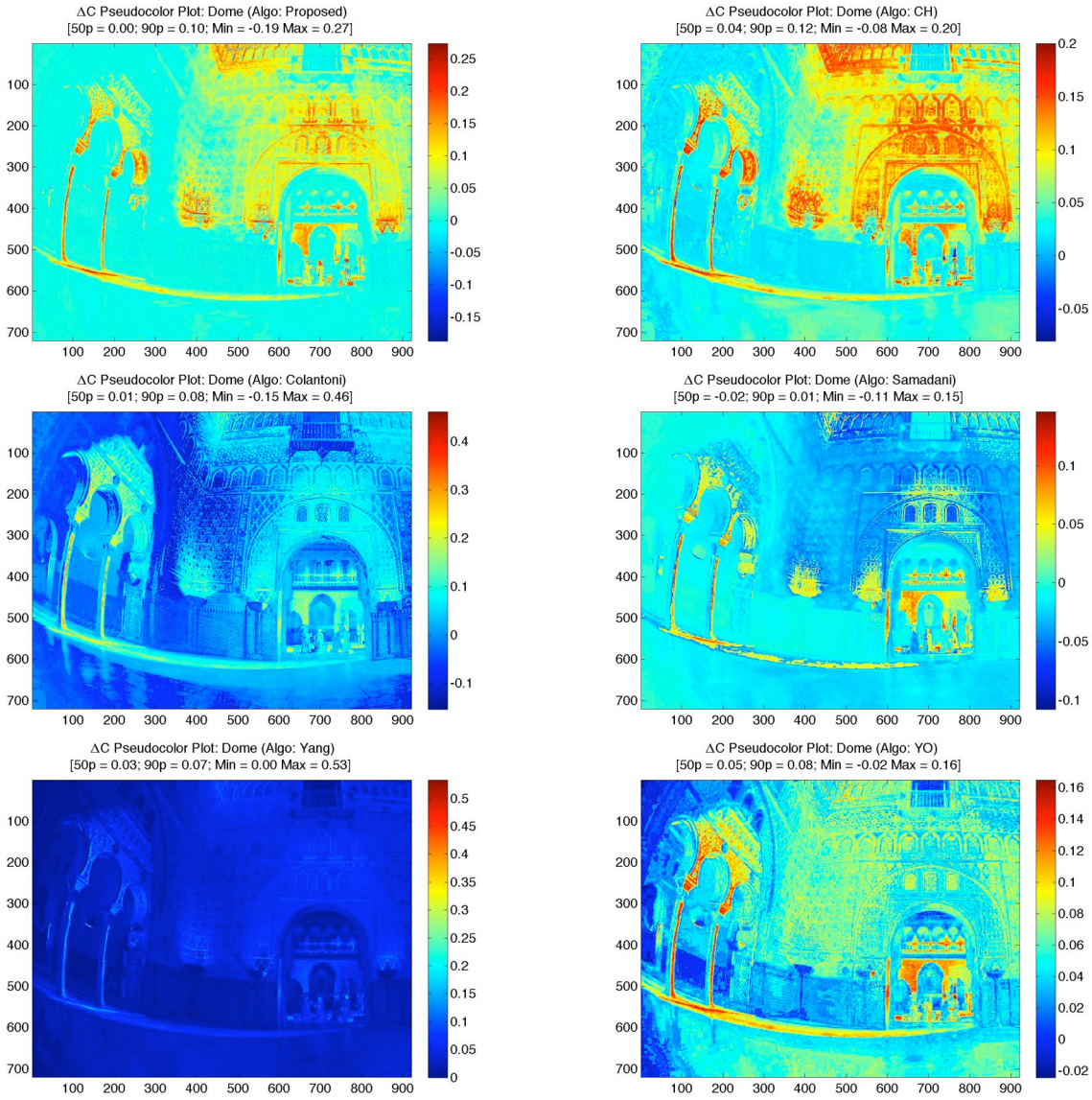


Fig. A.24 ΔC Image Difference Maps: Dome

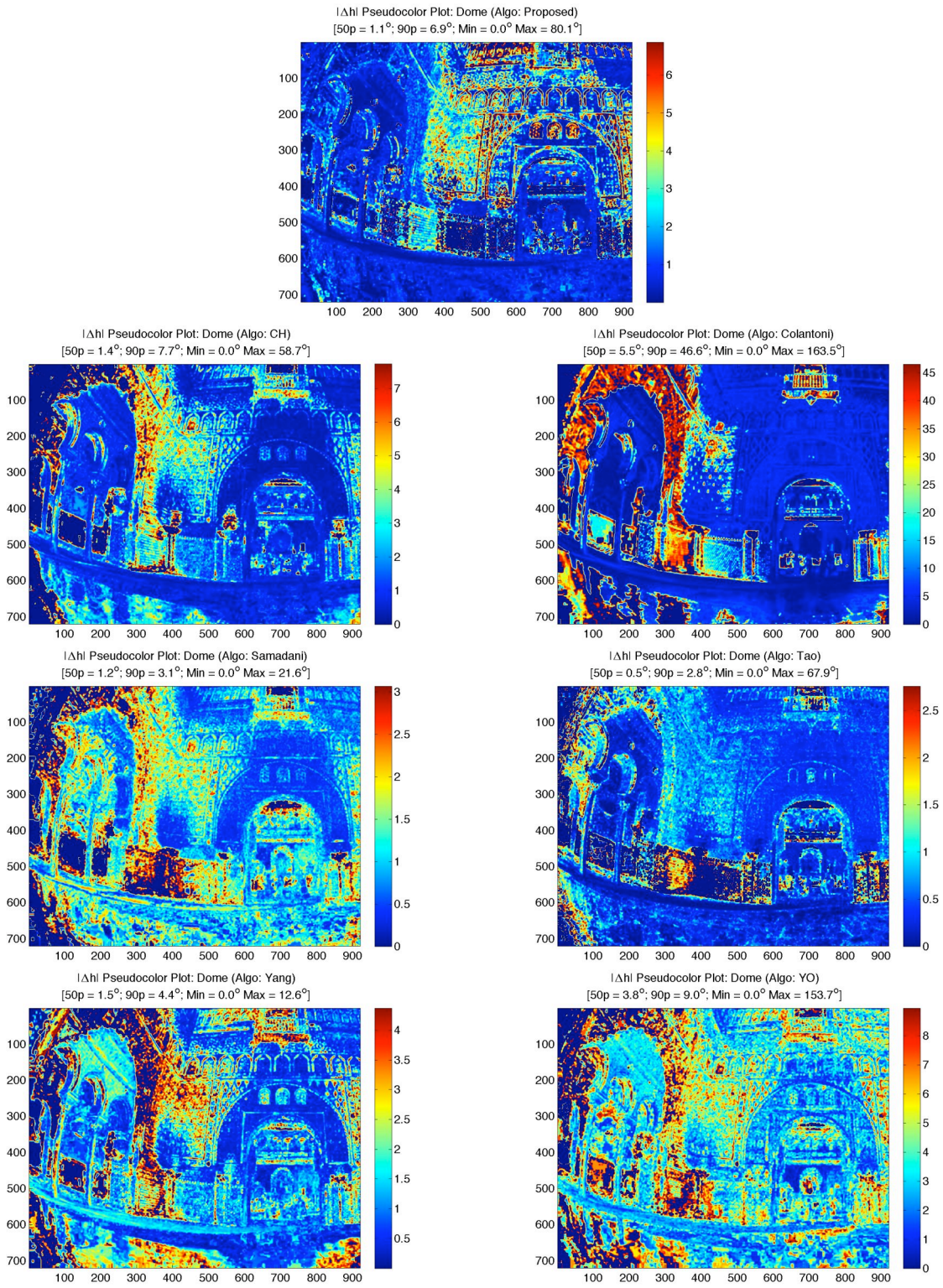


Fig. A.25 Δh Image Difference Maps: Dome

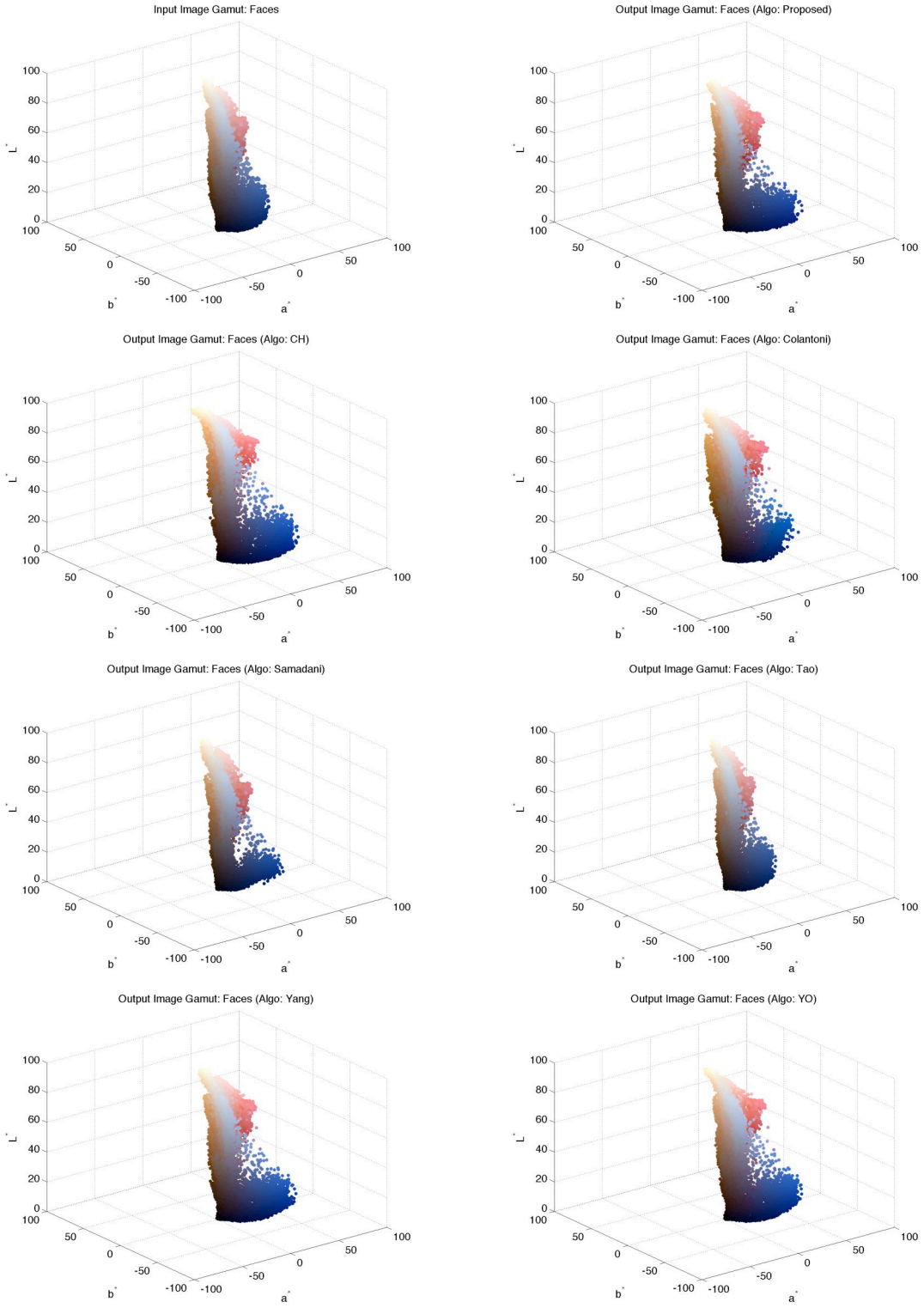


Fig. A.26 Image Gamut Maps: Faces

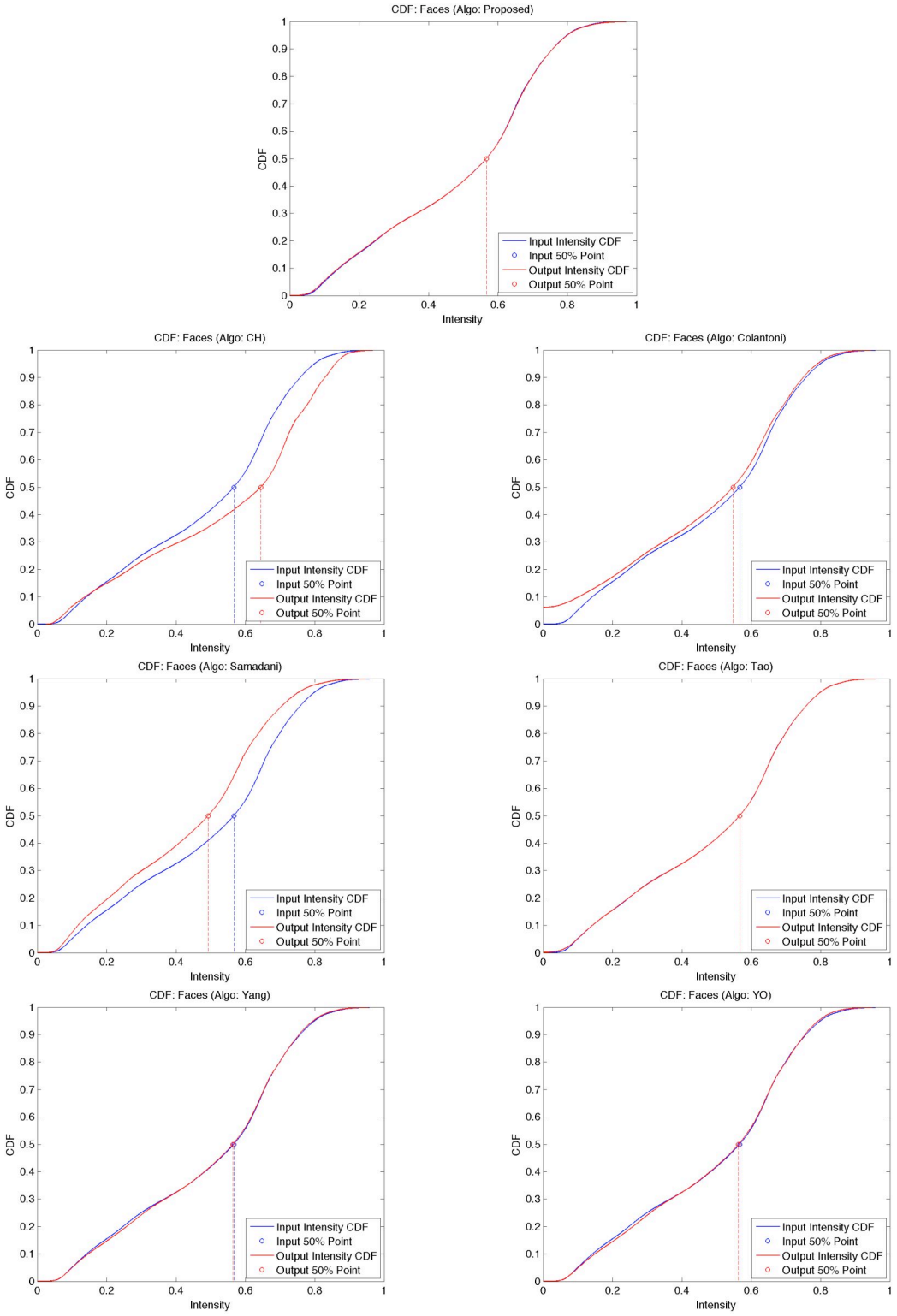


Fig. A.27 Cumulative Distribution Functions: Faces

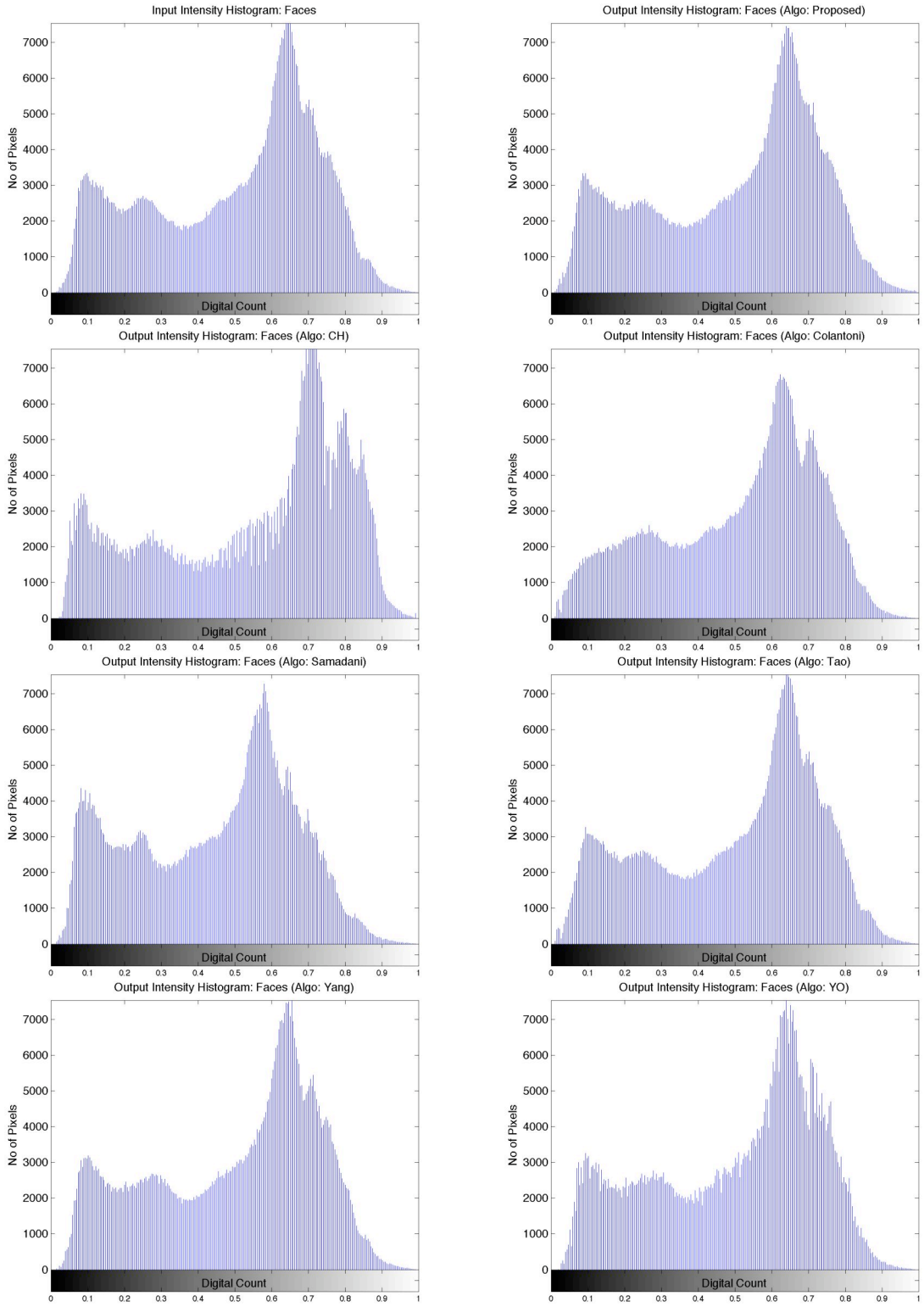


Fig. A.28 Intensity Histograms: Faces

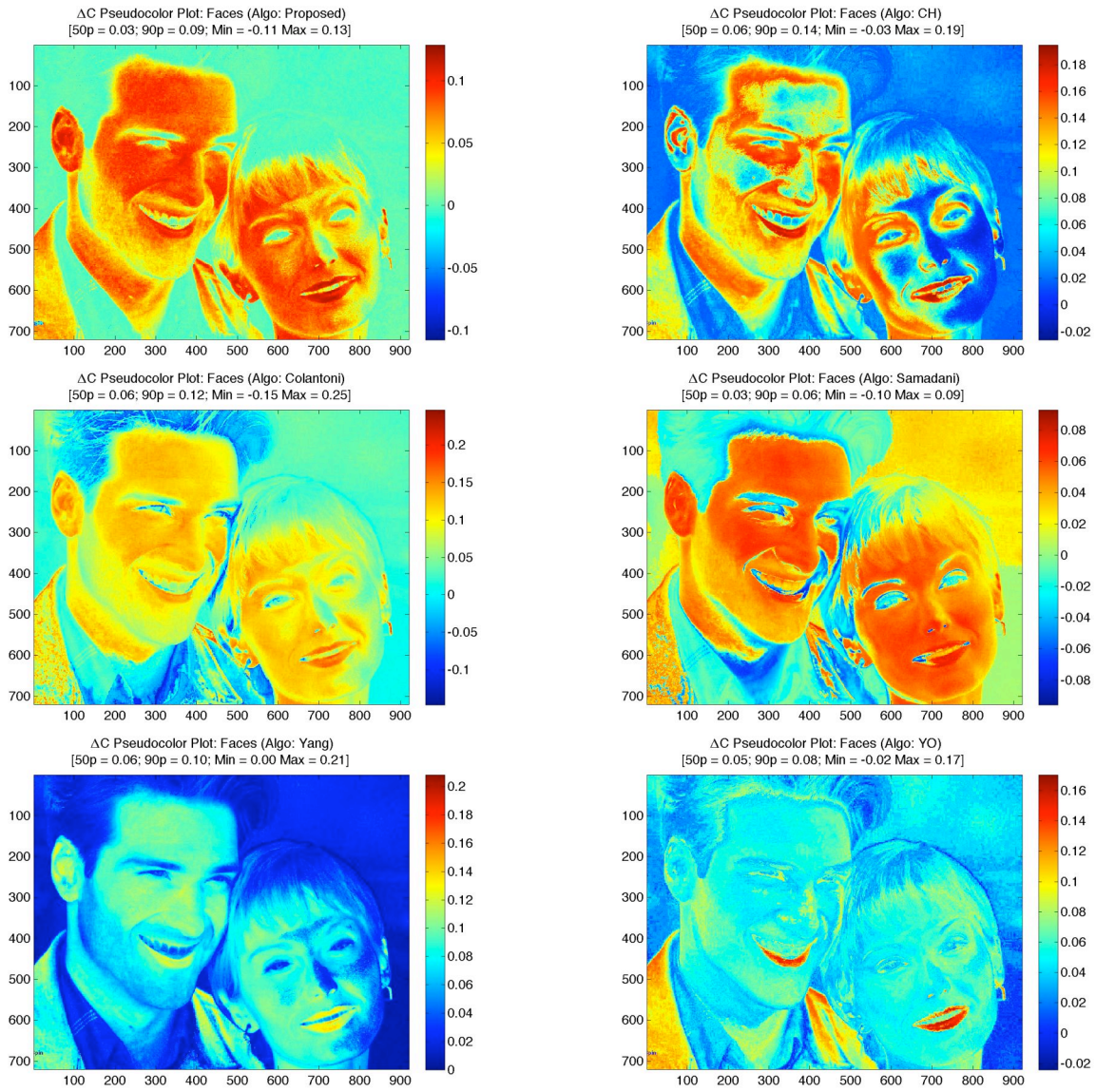


Fig. A.29 ΔC Image Difference Maps: Faces

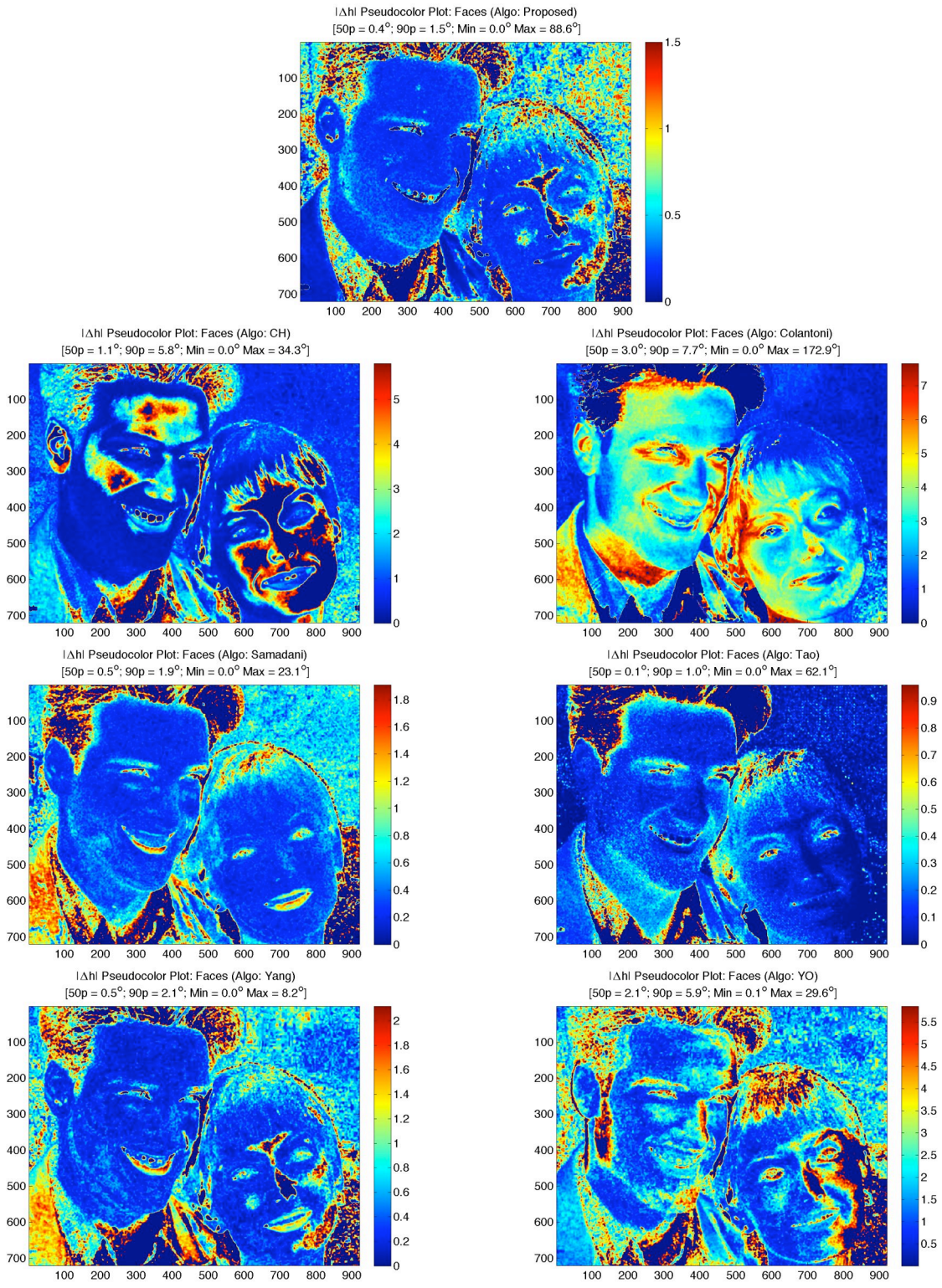


Fig. A.30 Δh Image Difference Maps: Faces

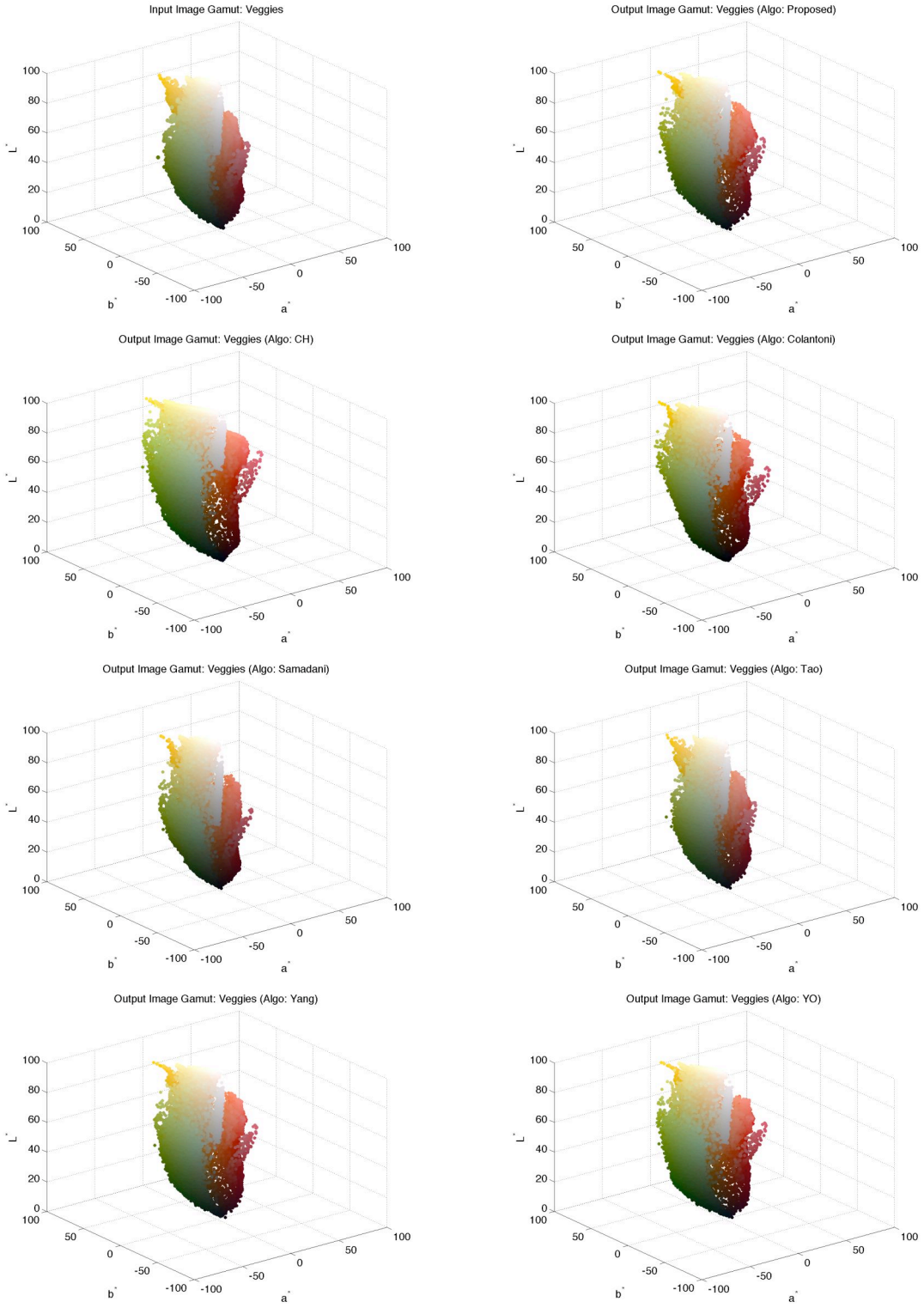


Fig. A.31 Image Gamut Maps: Veggies

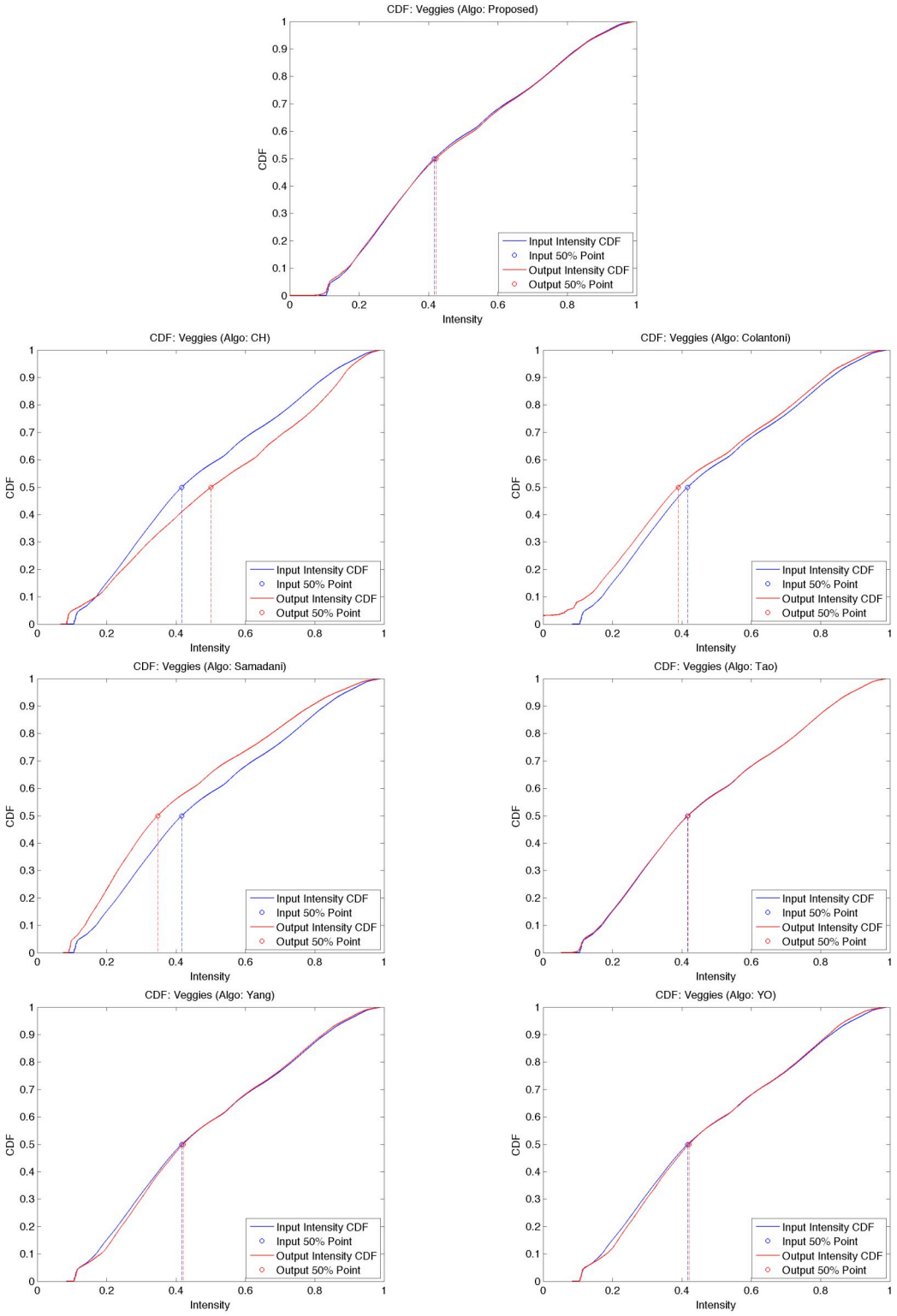


Fig. A.32 Cumulative Distribution Functions: Veggies

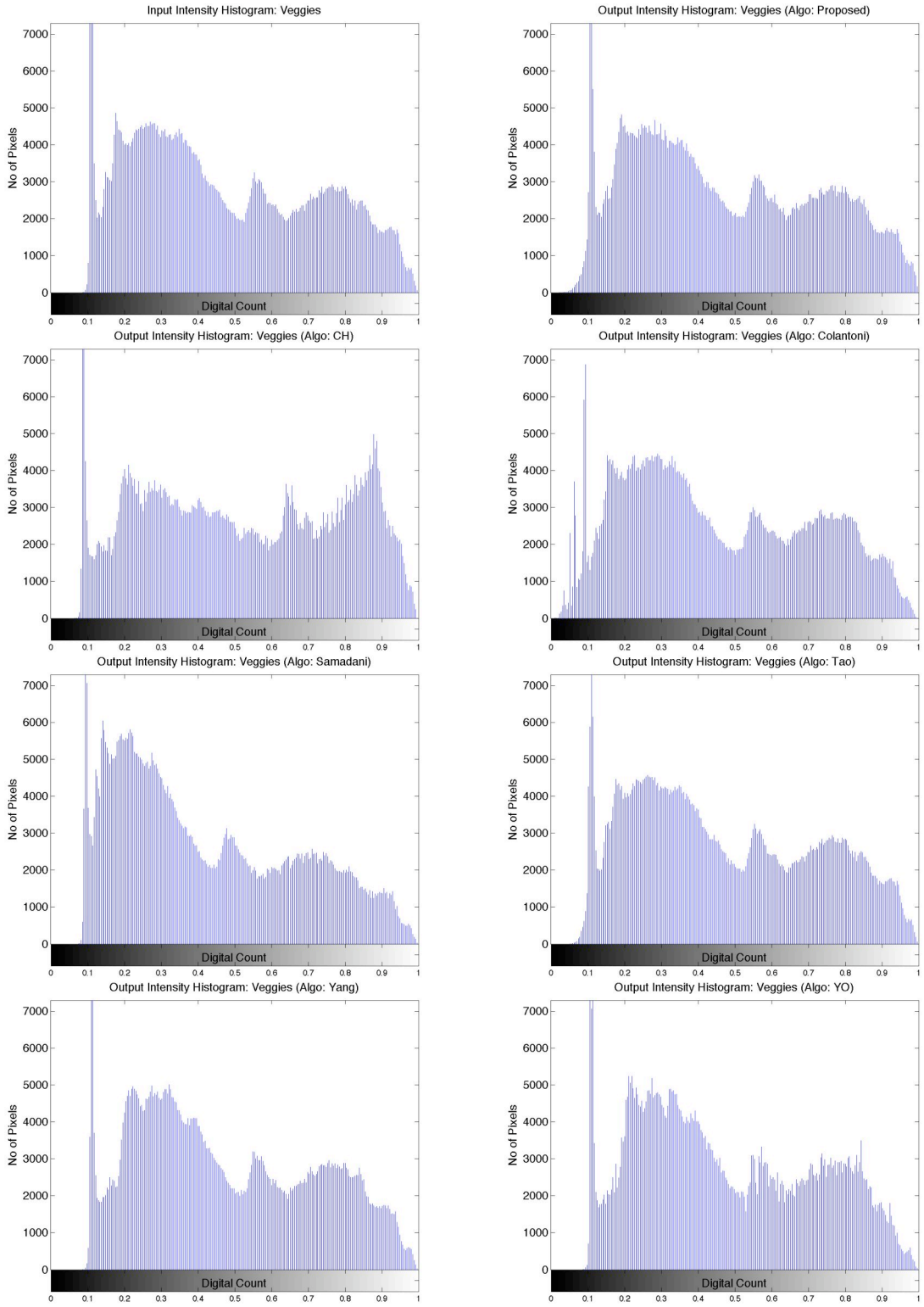


Fig. A.33 Intensity Histograms: Veggies

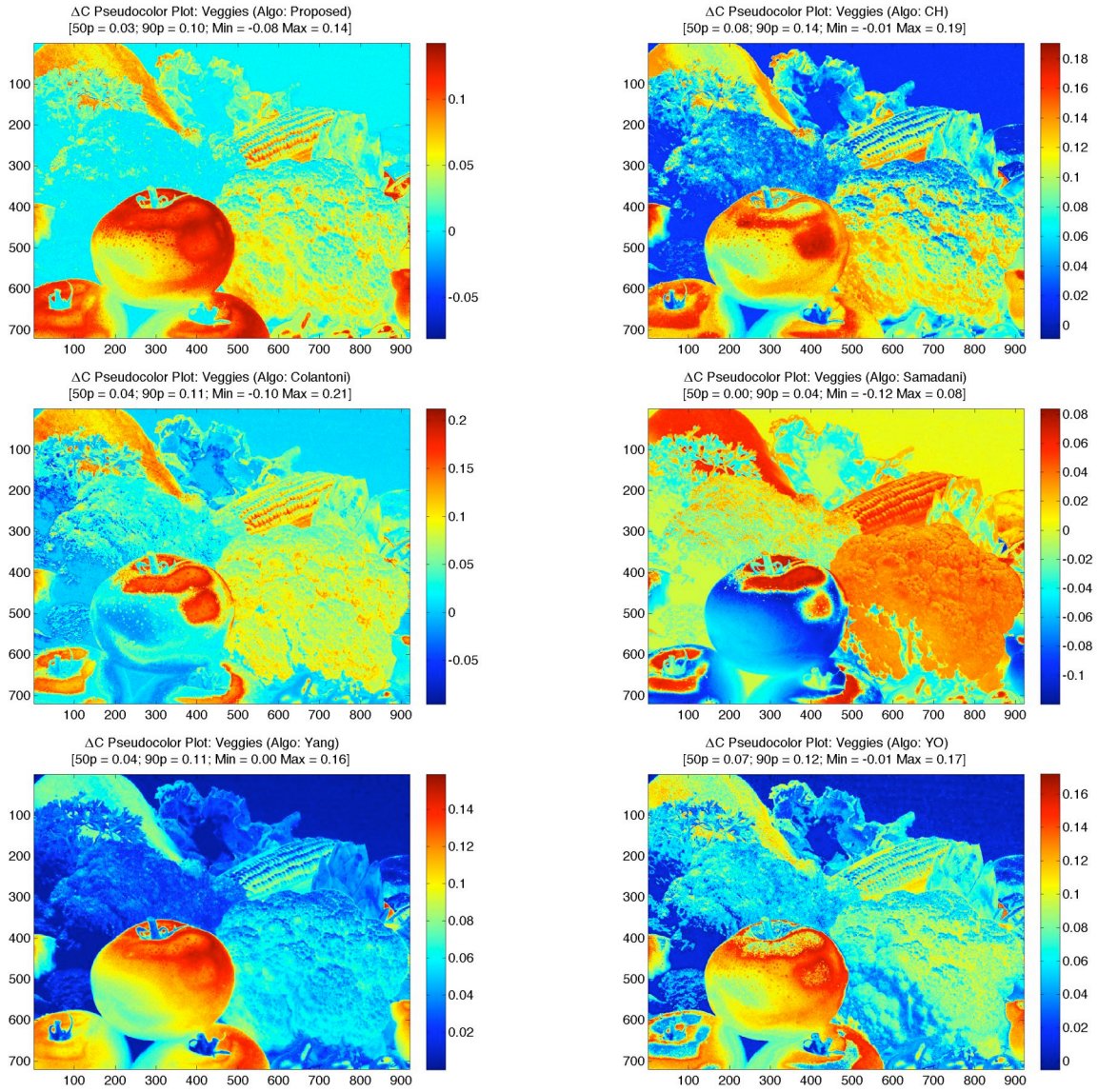


Fig. A.34 ΔC Image Difference Maps: Veggies

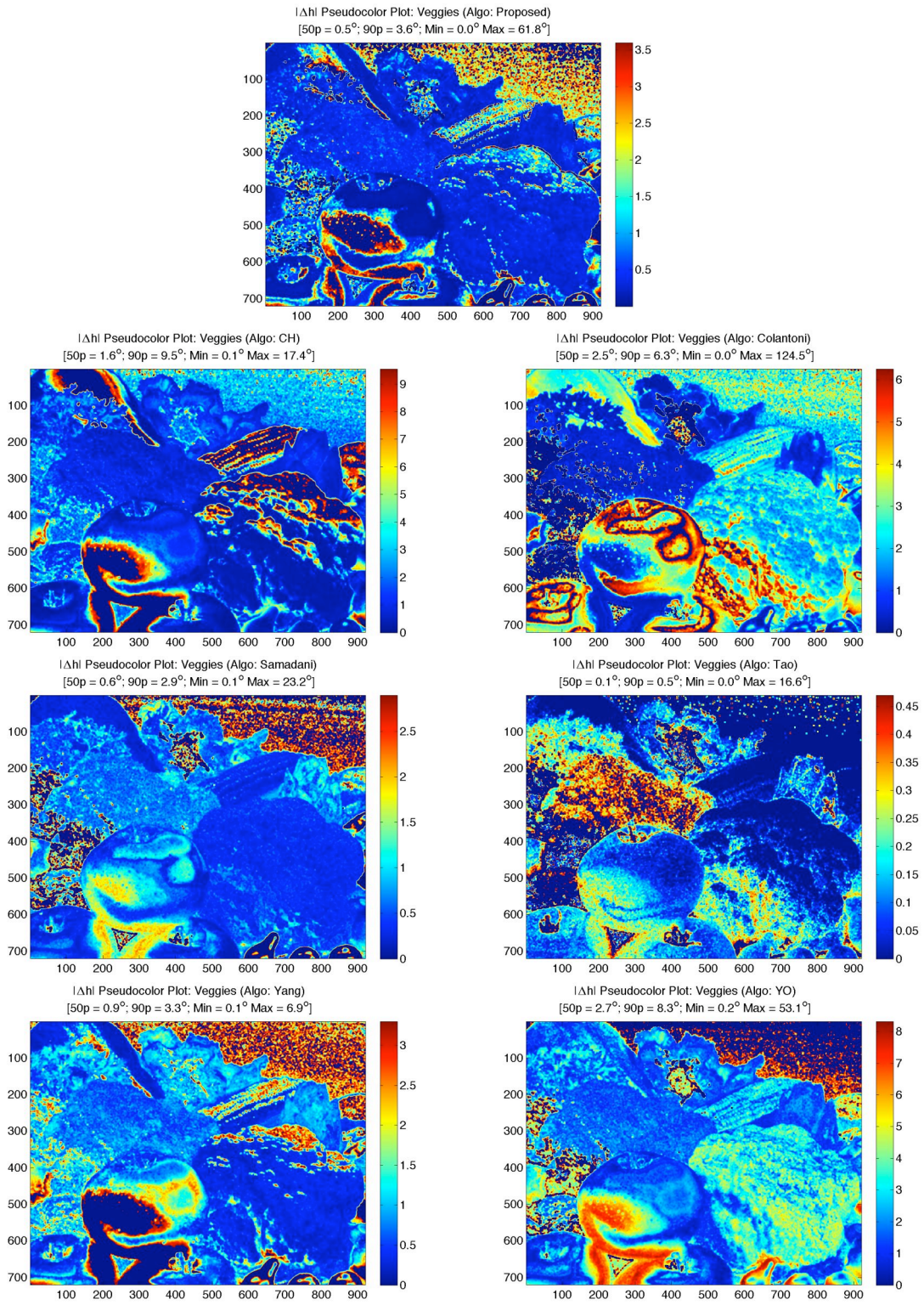


Fig. A.35 Δh Image Difference Maps: Veggies