

CHAPTER II

THEORETICAL CONSIDERATIONS AND LITERATURE REVIEW

2.1 Theoretical considerations

This chapter provides some basic theories behind the present study. First, the concept of color perception needs to be addressed (Section 2.1.1). A description of color spaces is then given in Sections 2.1.2, 2.1.3, and 2.1.4, for CIE XYZ, RGB, and sRGB color space, respectively. The main contributions of this thesis, i.e. digital image enhancement and iCAM, are described in Sections 2.1.5 and 2.1.6.

2.1.1 The perception of color

Color perception involves three basic factors: the source of light, objects under illumination, and the eyes and neural responses of observers. The visual process begins when radiant energy from the source strikes the object and some of this energy is reflected and passes through the lens to strike the retina in the eye. The retina is made up of a complex network of cells and neurons. The retina consists of a large number of cells which are sensitive to light; these receptors cells are of two kinds, rods and cones. Rods are sensitive to brightness of light only at low illuminate. Cones are cells of three different types which respond to red, green and blue regions of light, respectively, and it is through these that all colors are seen. When the three type of cones are all stimulated equally, the eye and the brain see achromatic, but if one types

of cone is stimulated more than the other two, the image appears to be tinted with the corresponding primary hue.

The most central part of the retina is called the fovea and it has the largest concentration of cells. The fovea vision is used for distinguishing very fine detail, such as reading and seeing objects at distance. Outside the fovea, the number of cones is greatly reduced and they are situated quite apart from one another. The rods are completely absent from the fovea and fall out to the extreme periphery. The signals leave the retina via the optic nerve and eventually arrive at the back of the brain. The brain signals are interpreted through mental impressions that result in perception [5].

2.1.2 The CIE color system

In the study of the perception of color, one of the first mathematically defined color spaces was the CIE XYZ color space (also known as CIE 1931 color space), created by the International Commission on Illumination (CIE) in 1931. The human eye has receptors for short (S), middle (M), and long (L) wavelengths, also known as blue, green, and red receptors. That means that one, in principle, needs three parameters to describe a color sensation. A specific method for associating three numbers (or tristimulus values) with each color is called a color space, of which the CIE XYZ color space is one of many such spaces. However, the CIE XYZ color space is special, because it is based on direct measurements of the human eye, and serves as the basis from which many other color spaces are defined. The CIE XYZ color space was derived from a series of experiments done in the late 1928's by W. David Wright [6] and John Guild [7]. Their experimental results were combined into the specification of the CIE RGB color space, from which the CIE XYZ color space

was derived. This thesis is actually concerned with both of these color spaces. In the CIE XYZ color space, the tristimulus values are not the S, M, and L stimuli of the human eye, but rather a set of tristimulus values called X, Y, and Z, which are also roughly red, green and blue, respectively. Two light sources may be made up of different mixtures of various colors, and yet have the same color (metamerism). If two light sources have the same apparent color, then they will have the same tristimulus values, no matter what different mixtures of light were used to produce them [8].

2.1.2.1 CIE illuminants

The CIE has established a number of spectral power distributions as CIE illuminants for colorimetric. These distributions are based on physical standards, such as blackbody radiators or Planckian radiator, or are based on statistical representations of measured light.

CIE illuminant A represents a Planckian radiator with a color temperature of 2856 K, as shown in Figure 2-1. It is used for colorimetric calculations when incandescent illumination is of interest.

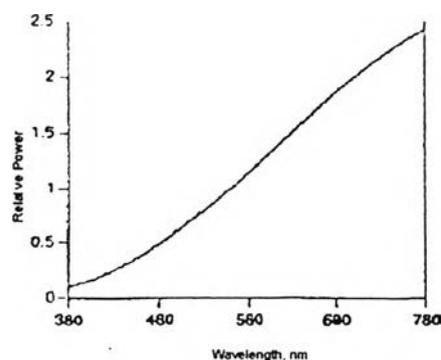


Figure 2-1: The spectral power distribution of CIE illuminant A [9].

CIE illuminant C is the spectral power distribution of illuminant A as modified by particular liquid filters defined by the CIE. It represents a daylight simulator with a correlated color temperature of 6774 K, as shown in Figure 2-2.

CIE illuminants D65 and D50 are part of the CIE D series illuminants that have been statistically defined based upon a large number of measurements of nature daylight. Illuminant D65 represents an average daylight with a correlated color temperature of 6500 K, and D50 represents an average daylight with a correlated color temperature of 5003 K, as shown in Figure 2-2. D65 is commonly used in colorimetric applications, such as paints, plastics, and textiles. D50 is often used in graphic arts and computer industries.

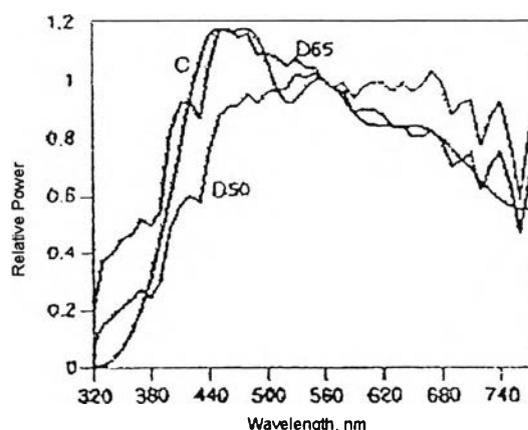


Figure 2-2: The spectral power distribution of CIE illuminants D50, D65 and C [9].

CIE F series illuminants represent typical spectral power distributions for various types of fluorescent sources including standards cool white, warm white, “full spectrum”, and tri-band, 12 in all. CIE illuminant F2 represents a fluorescent with a correlated color temperature of 4230 K. Illuminant F8 represents a fluorescent

D50 simulator with a correlated color temperature of 5000 K, and illuminant F11 represents a tri-band fluorescent source with a correlated color temperature of 4000K, as shown in Figures 2-3 and 2-4. Tri-band fluorescent sources are popular because of their efficiency, efficacy, and pleasing color-rendering properties.

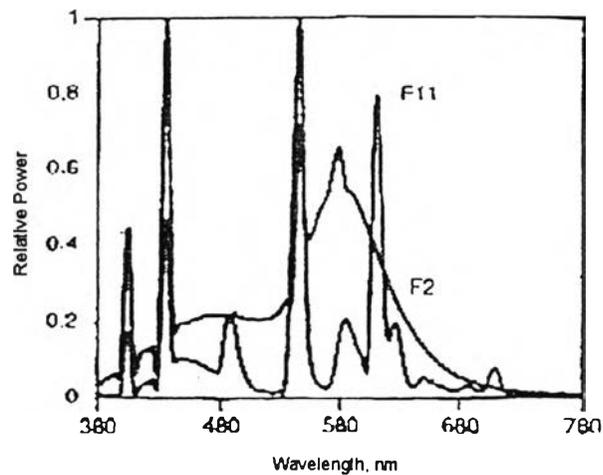


Figure 2-3: The spectral power distribution of CIE illuminants F2 and F11 [10].

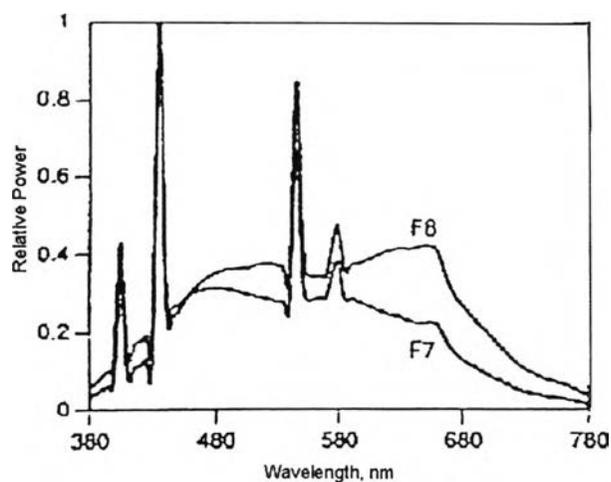


Figure 2-4: The spectral power distribution of CIE illuminants F7 and F8 [10].



2.1.2.2 Standard of reflectance factor

The CIE recommends that reflectance measurement be made relative to the perfect reflecting diffuser. There is no object surface that has the properties of the perfect reflecting diffuser, but working standards of known spectral reflectance factors are normally used. The working standards for reflectance factor measurement are called white standard. The effect of an object on light can be described by its spectral transmittance or reflectance curve. The spectral reflectance curve describes the object just as the spectral power distribution curve describes a source. Figure 2-5 shows examples of spectral reflectance factor of a perfect reflecting diffuser and a sample.

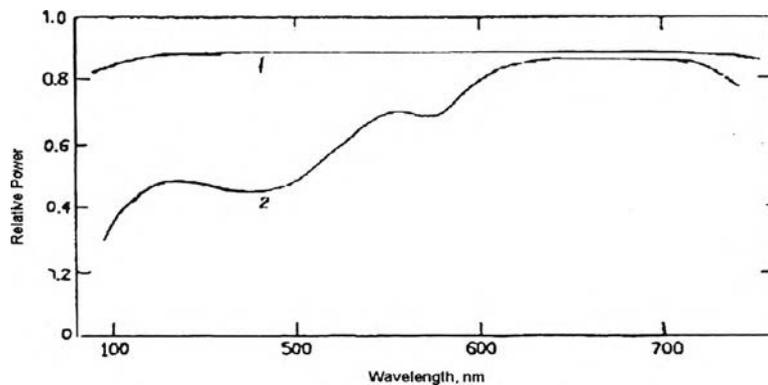


Figure 2-5: The spectral reflectance factor of a perfect reflecting diffuser and a sample [11].

2.1.2.3 CIE standard observers

The colorimetric specifications of colors are based on the spectral tristimulus values of $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ which are also called color matching functions. There are two sets of color-matching functions established by the CIE. The CIE 1931 standard colorimetric observer was determined from experiments by Guild [6] and Wright [7], using a visual field that subtended 2 degrees so that the matching stimuli were imaged onto the retina completely within the fovea. In 1964, the CIE recommended a set of color-matching functions notated as $\bar{x}_{10}(\lambda)$, $\bar{y}_{10}(\lambda)$, and $\bar{z}_{10}(\lambda)$ for the experiments using a 10° visual field that excluded the central fovea. The results for large fields were significantly different from the 2° standard, enough to warrant the establishment of the CIE 1964 supplementary standard colorimetric observer, sometimes called the 10° observer, as shown in Figure 2-6. Nowadays standards exist for field size of 2° and 10° .

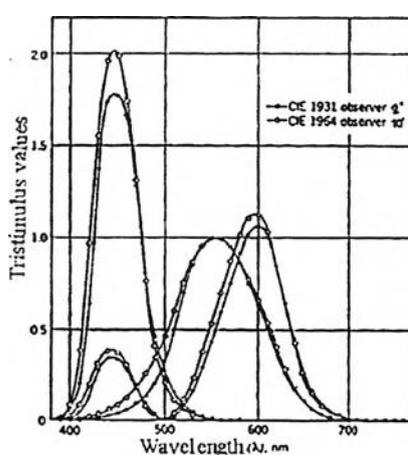


Figure 2-6: Comparison of color matching functions of the 1931 CIE standard observers and the 1964 CIE supplementary observers [12].

2.1.2.4 CIE XYZ tristimulus values

The CIE tristimulus values X, Y, and Z of color are obtained by multiplying together the relative power of a CIE standard illuminant, the reflectance or the transmittance of the object and the standard observer function (Figure 2-7).

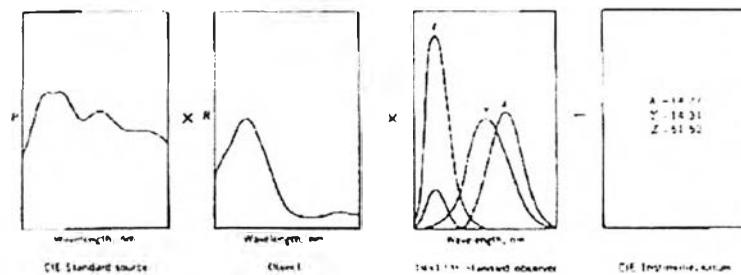


Figure 2-7: The CIE tristimulus values X, Y and Z of color [13].

The CIE XYZ tristimulus values are calculated by:

$$\begin{aligned}
 X &= k \sum S_{\lambda} R_{\lambda} \bar{x}_{\lambda} \Delta\lambda \\
 Y &= k \sum S_{\lambda} R_{\lambda} \bar{y}_{\lambda} \Delta\lambda \\
 Z &= k \sum S_{\lambda} R_{\lambda} \bar{z}_{\lambda} \Delta\lambda \\
 k &= 100 / \sum S_{\lambda} \bar{y}_{\lambda} \Delta\lambda
 \end{aligned}
 \tag{2.1}$$

Where, S_{λ} is the spectral power distribution of light source

R_{λ} is the spectral reflectance factor of object

$\bar{x}_{\lambda}, \bar{y}_{\lambda}, \bar{z}_{\lambda}$ are the color matching functions

k is a normalizing constant

$\Delta\lambda$ is the measurement wavelength interval

\sum_{λ} is summation across wavelength

By convention, the value $Y = 100$, assigned to perfect white object reflecting or transmitting 100% at all wavelengths, or to the perfect colorless sample. Thus, a value of 100 is the maximum value that Y can have for non fluorescent sample.

2.1.3 RGB color space

The RGB color space is commonly used for specifying colors based on the additive-mixing theory. In the case of CRT monitor, the three primaries are the particular color emitted by the three phosphors. It is therefore highly device dependent; the same color may be specified as two different sets of numbers on two different monitors. The RGB color space is the device dependent color space, which defines colors within a unit cube by additive-color-mixing model. Red, green and blue are additive primaries represented by the three axes of cube as shown in Figure 2-8. All colors within the cube can be represented as the triplet (R, G, B) , where values R , G and B are assigned in the range from 0 to 1. An important characteristic of the additive system is that the object itself is a light emitter such as a television. Scanners and computer monitor are also used the RGB space. RGB values in one device might not look the same as RGB values in other devices.

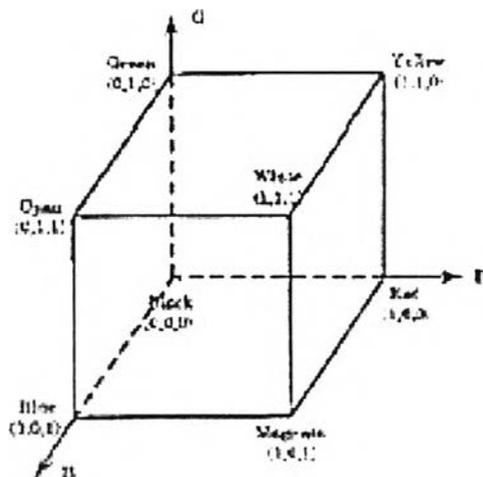


Figure 2-8: RGB color space [14].

2.1.4 sRGB color space

sRGB color space, or standard RGB (Red Green Blue), is an RGB color space created cooperatively by Hewlett-Packard and Microsoft Corporation [15]. sRGB defines the red, green, and blue primaries as color where one of the three channels is at the maximum value and the other two are at zero. In CIE xy chromaticity coordinates red is at [0.6400, 0.3300], green is at [0.3000, 0.6000] and blue is at [0.1500, 0.0600] and white point is the D65 white point at [0.3127, 0.3290]. sRGB has been criticized for poor placement of these primary colors. If you restrict the indexes to the 0 to 1 range, you are unable to address outside the gamut, which is well inside the set of visible colors to a human, as shown in Figure 2-9.

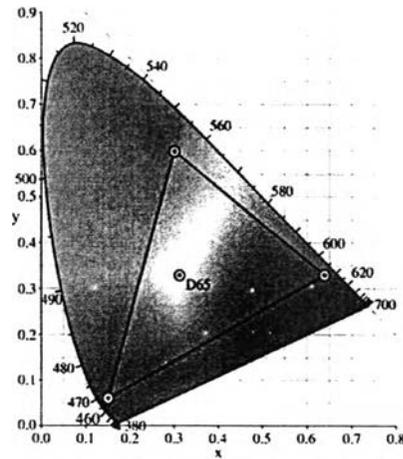


Figure 2-9: CIE 1931 xy chromatic diagram showing the gamut of the sRGB color space and location of primary. The D65 white point is shown in center [15].

2.1.4.1 The forward transformation

sRGB can be transformed to CIE XYZ using Equations 2.1-2.5.

The 8 bit integer RGB values are converted to floating point non-linear sR'G'B' values as follows:

$$R'_{sRGB} = R_{8bit} / 255.0$$

$$G'_{sRGB} = G_{8bit} / 255.0 \quad (2.2)$$

$$B'_{sRGB} = B_{8bit} / 255.0$$

The nonlinear R'G'B' values are transformed to linear R, G, and B values by:

$$\text{If } R'_{sRGB}, G'_{sRGB}, B'_{sRGB} \leq 0.04045$$

$$R = R'_{sRGB} / 12.92$$

$$G = G'_{sRGB} / 12.92 \quad (2.3)$$

$$B = B'_{sRGB} / 12.92$$

$$\text{else if } R'_{sRGB}, G'_{sRGB}, B'_{sRGB} > 0.04045$$

$$R = ((R'_{sRGB} + 0.055) / 1.055)^{2.4}$$

$$G = ((G'_{sRGB} + 0.055) / 1.055)^{2.4} \quad (2.4)$$

$$B = ((B'_{sRGB} + 0.055) / 1.055)^{2.4}$$

Then convert to XYZ by:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{D65} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.5)$$

2.1.4.2 The reverse transformation

From CIE XYZ, it is also possible to obtain sRGB. The first step is to calculate linear RGB values (Equation 2.6).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2406 & -1.5372 & -0.4986 \\ -0.9689 & 1.8758 & 0.0415 \\ 0.0557 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{D65} \quad (2.6)$$

The red, green and blue phosphor chromaticities and D65 white point result in the following relationship between D65 tristimulus values and linear RGB values. Any values greater than 1.0 or less than 0.0 are clipped to 1.0 and 0.0

If $R, G, B \leq 0.0031308$

$$R'_{sRGB} = 12.92 \times R$$

$$G'_{sRGB} = 12.92 \times G \quad (2.7)$$

$$B'_{sRGB} = 12.92 \times B$$

else if $R, G, B > 0.0031308$

$$R'_{sRGB} = 1.055 \times R^{(1.0/2.4)} - 0.055$$

$$G'_{sRGB} = 1.055 \times G^{(1.0/2.4)} - 0.055 \quad (2.8)$$

$$B'_{sRGB} = 1.055 \times B^{(1.0/2.4)} - 0.055$$

The nonlinear R'G'B' values are then converted into 8 bit integers by:

$$\begin{aligned}R_{8\text{bit}} &= 255.0 \times R'_{\text{sRGB}} \\G_{8\text{bit}} &= 255.0 \times G'_{\text{sRGB}} \\B_{8\text{bit}} &= 255.0 \times B'_{\text{sRGB}}\end{aligned}\tag{2.9}$$

2.1.5 Digital image enhancement

In computer graphics, the process of improving the quality of a digitally stored image by manipulating the image with software is quite easy. For example, it is possible to make an image lighter or darker, or to increase or decrease contrast. Advanced image enhancement software also supports many filters for altering images in various ways. Programs specialized for image enhancements are sometimes called image editors [16].

2.1.5.1 Image histogram

Understanding image histograms is probably the single most important concept to become familiar with when working with pictures from a digital camera. A histogram can tell whether or not this image has been properly exposed, whether the lighting is harsh or flat, and what adjustments will work best. It will not only improve your skills on the computer, but as a photographer as well.

Each pixel in an image has a color which has been produced by some combination of the primary colors red, green, and blue (RGB). Each of these colors can have a brightness value ranging from 0 to 255 for a digital image with a bit depth of 8 bits. A RGB histogram is obtained when the computer scans through each of these RGB brightness values and counts how many pixels are at each level from 0 through 255 [17]. Other types of histograms exist, although all will have the same basic layout as the histogram example shown in Figure 2-10.

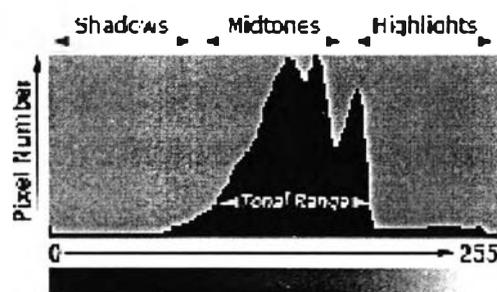


Figure 2-10: The histogram example [17].

2.1.5.1.1 Tones

The region where most of the brightness values are present is called the "tonal range." Tonal range can vary drastically from image to image, so developing an intuition for how numbers map to actual brightness values is often critical—both before and after the photo has been taken. There is no one "ideal histogram" which all images should try to mimic; histograms should merely be representative of the tonal range in the scene and what the photographer wishes to convey.

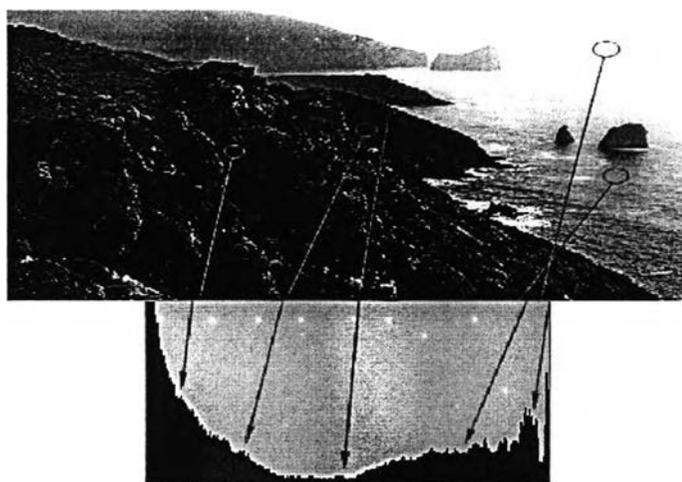


Figure 2-11: The histogram example of high contrast [17].

Figure 2-11 is an example image which contains a very broad tonal range, with markers to illustrate where regions in the scene map to brightness levels on the histogram. This coastal scene contains very few midtones, but does have plentiful shadow and highlight regions in the lower left and upper right of the image, respectively. This translates into a histogram which has a high pixel count on both the far left and right-hand sides.

Lighting is often not as extreme as the last example. Conditions of ordinary and even lighting, when combined with a properly exposed subject, will usually produce a histogram which peaks in the centre gradually taper off into the shadows and highlights. With the exception of the direct sunlight reflecting off the top of the building and off some windows, the boat scene to the right is quite evenly lit (Figure 2-12). Most cameras will have no trouble automatically reproducing an image which has a histogram similar to the one shown in Figure 2-12.



Figure 2-12: The histogram example with peaks in the center [17].

2.1.5.1.2 Contrast

A histogram can also describe the amount of contrast. Contrast is a measure of the difference in brightness between light and dark areas in a scene. Broad histograms reflect a scene with significant contrast, whereas narrow histograms reflect less contrast and may appear flat or dull. This can be caused by any combination of subject matter and lighting conditions. Photos taken in the fog will have low contrast, while those taken under strong daylight will have higher contrast [17].

Contrast can have a significant visual impact on an image by emphasizing texture, as shown in Figure 2-13. The high contrast water has deeper

shadows and more pronounced highlights, creating texture which "pops" out at the viewer.

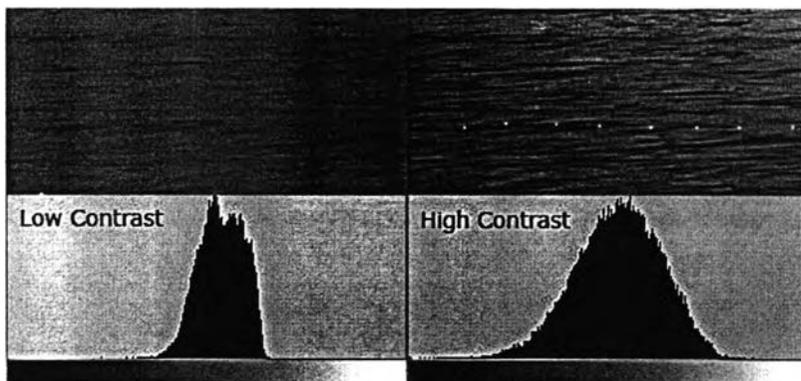


Figure 2-13: Difference between low contrast and high contrast [17].

2.1.5.2 Dynamic range

Dynamic range is the range of tonal difference between the lightest light and darkest dark of an image. The higher the dynamic range, the more potential shades can be represented, although the dynamic range does not automatically correlate to the number of tones reproduced. For instance, high-contrast microfilm exhibits a broad dynamic range, but renders few tones. Dynamic range also describes a digital system's ability to reproduce tonal information. This capability is most important for continuous-tone documents that exhibit smoothly varying tones, and for photographs it may be the single most important aspect of image quality [18].

2.1.5.3 Bit depth

Bit depth quantifies how many unique colors are available in an image's color palette in terms of the number of 0's and 1's, or "bits," which are used to specify each color. This does not mean that the image necessarily uses all of these colors, but that it can instead specify colors with that level of precision. For a grayscale image, the bit depth quantifies how many unique shades are available. Images with higher bit depths can encode more shades or colors since there are more combinations of 0's and 1's available.

Every color pixel in a digital image is created through some combination of the three primary colors: red, green, and blue. Each primary color is often referred to as a "color channel" and can have any range of intensity values specified by its bit depth. The bit depth for each primary color is termed the "bits per channel." The "bits per pixel" (bpp) refers to the sum of the bits in all three color channels and represents the total colors available at each pixel. Confusion arises frequently with color images because it may be unclear whether a posted number refers to the bits per pixel or bits per channel. Using "bpp" as a suffix helps distinguish these two terms [19].

Most color images from digital cameras have 8 bits per channel and so they can use a total of eight 0's and 1's. This allows for 2^8 or 256 different combinations—translating into 256 different intensity values for each primary color. When all three primary colors are combined at each pixel, this allows for as many as 2^{8*3} or 16,777,216 different colors, or "true color." This is referred to as 24 bits per pixel since each pixel is composed of three 8-bit color channels. The number of

colors available for any X-bit image is just 2^X , if X refers to the bits per pixel, and 2^{3x} if X refers to the bits per channel. The following Table 2-1 shows different image types in terms of bits (bit depth), total colors available, and common names.

Table 2-1: The different image types in terms of bits (bit depth), total colors available, and common names [19].

Bits Per Pixel	Number of Colors Available	Common Name(s)
1	2	Monochrome
2	4	CGA
4	16	EGA
8	256	VGA
16	65536	XGA, High Color
24	16777216	SVGA, True Color
32	16777216 + Transparency	
48	281 Trillion	

2.1.5.4 Image Filtering

Digital image can be processed in a variety of ways. The most common one is called filtering and creates a new image as a result of processing the pixels of an existing image. Each pixel in the output image is computed as a function of one or several pixels in the original image, usually located near the location of the output pixel. If the function used does some kind of interpolation (e.g. linear, cubic or Gaussian), then the result will look smoother than the original, but care needs to be taken that the output values are not computed from too many input pixels, or the resulting image may get blurred. The most common purpose for this interpolation is anti aliasing [20].

2.1.5.4.1 Gaussian filter

The Gaussian smoothing operator is a 2-D convolution operator that is used to blur images and remove detail and noise. In 2-D, an isotropic (i.e. circularly symmetric) Gaussian has the form in Equation 2-10 [21].

$$G_{(x,y)} = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2-10)$$

From Equation 2-10, the variables taking effects with the shape of filter are σ and size of the filter, x and y .

This distribution is shown in Figure 2-14.

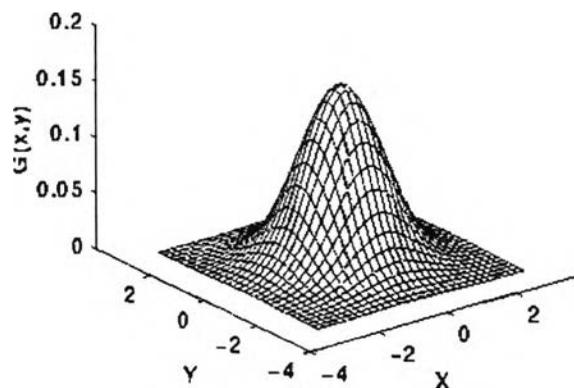


Figure 2-14: 2-D Gaussian distribution with mean $(0, 0)$ and $\sigma = 1$ [21].

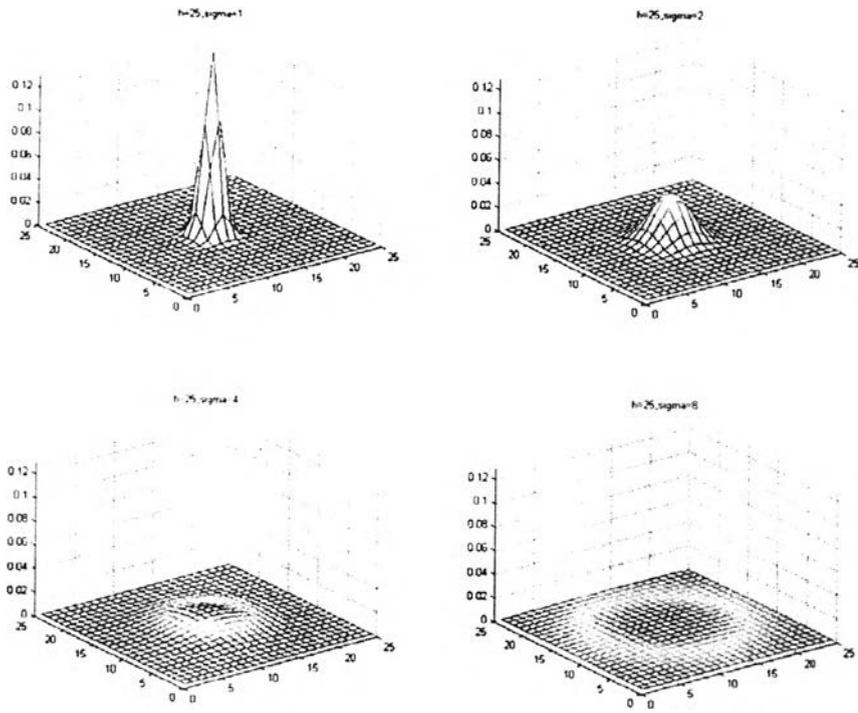


Figure 2-15: Different shapes of Gaussian filter when changing σ .

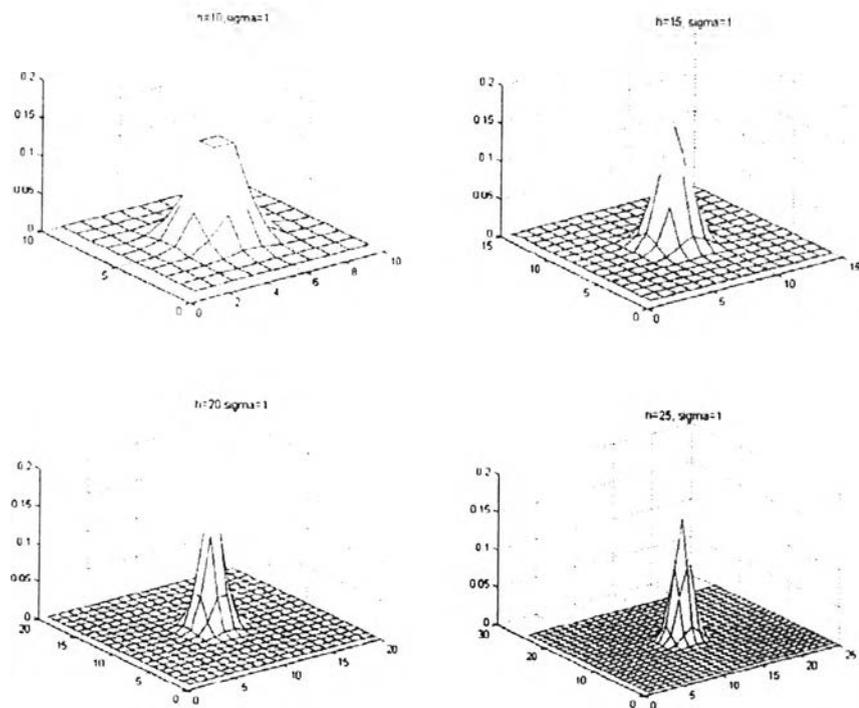


Figure 2-16: Different shapes of Gaussian filter when changing size of filter.

The ideal of Gaussian smoothing is to use this 2-D distribution as a 'point-spread' function and this is achieved by convolution. Since the image is stored as a collection of discrete pixels, we need to produce a discrete approximation to the Gaussian function before we can perform the convolution. In theory, the Gaussian distribution is non-zero everywhere, which would require an infinitely large convolution kernel, but in practice it is effectively zero more than about three standard deviations from the mean, and so we can truncate the kernel at this point.

One of the principle justifications for using the Gaussian as a smoothing filter is due to its frequency response. Most convolutions based smoothing filters act as low-pass frequency filters. This means that their effect is to remove low spatial frequency components from an image.

2.1.5.4.2 Convolution filter

Convolution is a simple mathematical operation which is fundamental to many common image processing operators. Convolution provides a way of "multiplying together" two arrays of numbers, generally of different sizes, but of the same dimensionality, to produce a third array of numbers of the same dimensionality. This can be used in image processing to implement operators whose output pixel values are simple linear combinations of certain input pixel values [22]. Convolution is the modification of a pixel's value on the basis of the value of neighboring pixels. Images are convolved by multiplying each pixel and its neighbors by a numerical matrix, called a kernel. This matrix is essentially moved over each pixel in the image, each pixel under the matrix is multiplied by the appropriate matrix value, the total is

summed and normalized, and the central pixel is replaced by the result. Figure 2-17 shows an example image and kernel that we will use to illustrate convolution.

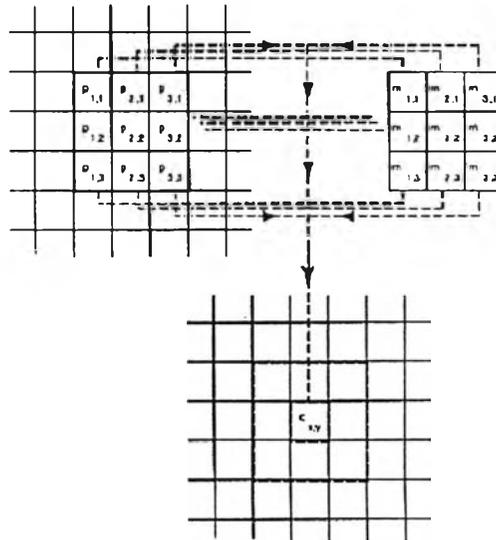


Figure 2-17: An example image (left) and kernel (right) for illustrating convolution [23].

The convolution is performed by sliding the kernel over the image, generally starting at the top left corner, so as to move the kernel through all the positions where the kernel fits entirely within the boundaries of the image. (Note that implementations differ in what they do at the edges of images as explained below.) Each kernel position corresponds to a single output pixel, the value of which is calculated by multiplying together the kernel value and the underlying image pixel value for each of the cells in the kernel, and then adding all these numbers together. So in example, the value of the bottom right pixel in the output image will be given by:

$$C_{(x,y)} = p_{11}m_{11} + p_{21}m_{21} + p_{31}m_{31} + p_{12}m_{12} + p_{22}m_{22} + p_{32}m_{32} + p_{13}m_{13} + p_{23}m_{23} + p_{33}m_{33}$$

At the edge of an image we usually want the filter to act as if the pixels beyond the edge are “just like” the ones in the image. So, for example, we can perform computations as if the (nonexistent) column of pixels just to the left of the image is identical to the left-most column of the image (Figure2-18).



Figure 2-18: An example convolution at the edge of an image [23].

2.1.6 iCAM

A next generation of image enhancement is image appearance model, namely iCAM for image color appearance model [2]. An iCAM capable of predicting perceived color difference between complex image stimuli is a useful tool, but has some limitations. Just as a color appearance model is necessary to fully describe the appearance of color stimuli, an image appearance model is necessary to describe spatially complex color stimuli. Color appearance models allow for the description of attributes such as lightness, brightness, colorfulness, chroma, and hue. Image appearance models extend upon this to also predict such attributes as sharpness, graininess, contrast, and resolution.

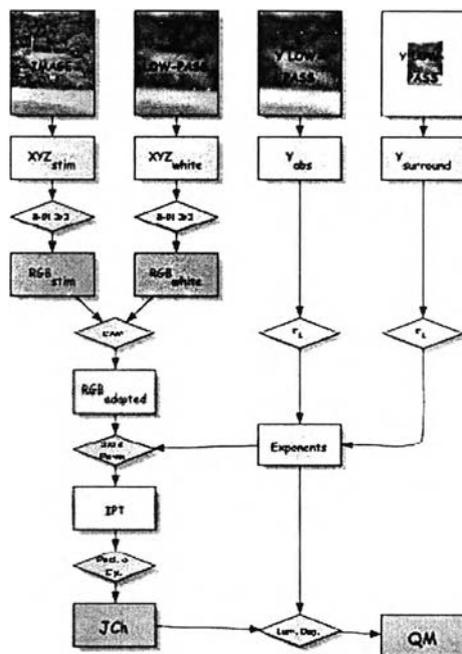


Figure 2-19: Flowchart of the iCAM image appearance model [24].

Figure 2-19 presents a flow chart of the general framework for the iCAM image appearance model as applied to complex stimuli still image, originally presented by Fairchild and Johnson [24]. This represents the traditional appearance modeling approach of treating each pixel as a stimulus in a point-wise fashion. The process is to start with tristimulus values for the stimulus and adapting white point and luminance values for the adapting level and surround. The tristimulus values are transformed to RGB values that are utilized in a linear, von Kries adaptation transform identical to the one proposed for the CIECAM02 color appearance model [25] as shown in Equations 2-11 to 2-13.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_{C4702} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2-11)$$

$$M_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \quad (2-12)$$

$$R_c = \left[\left(Y_w \frac{D}{R_w} \right) + (1-D) \right] R \quad (2-13)$$

The linear von Kries transform with an incomplete adaptation term, D , is given in Equation 2-13 for a single sensor. The primary difference between the iCAM chromatic adaptation transform and the CIECAM02 transform is in the definition of the white point, R_w in Equation 2-13. The iCAM transform uses a low-pass version of the image itself as the adapting white point to perform a localized adaptation. This adaptation can be a chromatic adaptation, as described in Equation 2-13, or can be luminance only adaptation Y_w tristimulus values of D65. This can be accomplished by replacing Equation 2-11 with Equation 2-14.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_{CAT02} \begin{bmatrix} Y \\ Y \\ Y \end{bmatrix} \quad (2-14)$$

The stimulus is replaced with an image and the adapting stimulus becomes a spatially low-pass image. The specific low-pass filters used for the adapting images depend on viewing distance and application. Additionally, in some image rendering circumstances it might be desirable to have different low-pass adapting images for luminance and chromatic information to avoid desaturation of the rendered images due to local chromatic adaptation (decrease in visual sensitivity to the color of the stimulus). The adapting luminance is also derived from a low-pass image of the

luminance channel and the surround luminance is derived from another low-pass image derived from a larger spatial extent. The surround luminance from low-pass image is used to calculate a degree-of-adaptation factor, D and a series of power function, F_L as a function of adaptation luminance, L_A , for various viewing conditions, as shown in Equations 2-15 to 2-16.

$$D = F \left[1 - \left(\frac{1}{3.6} \right) e^{\left(\frac{-L_A - 42}{92} \right)} \right] \quad (2-15)$$

$$F_L = 0.2 \left(\frac{1}{(5L_A + 1)} \right)^4 (5L_A) + 0.1 \left(1 - \left(\frac{1}{(5L_A + 1)} \right)^4 \right)^2 (5L_A)^{\frac{1}{3}} \quad (2-16)$$

The surround exponents calculated in the previous equation are actually used in the transform from XYZ tristimulus values to IPT appearance space. The first stage in this transforms is to convert the XYZ units into LMS cone responses. These cone responses are then compressed using a nonlinear power function which is modified on a per-pixel-basis by the surround map calculated (Equations 2-17 to 2-18). Typically the LMS responses are brought into the IPT appearance space for calculation of appearance correlates such as lightness, chroma, and hue (Equation 2-19).

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0.0 & 0.0 & 0.9184 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix} \quad (2-17)$$

$$\begin{aligned} L' &= L^{0.43^{*F_L}} ; L \geq 0 \\ L' &= -|L|^{0.43^{*F_L}} ; L \leq 0 \end{aligned} \quad (2-18)$$

$$\begin{bmatrix} I \\ P \\ T \end{bmatrix} = \begin{bmatrix} 0.4000 & 0.4000 & 0.2000 \\ 4.4550 & -4.8510 & 0.3960 \\ 0.8056 & 0.3572 & -1.1628 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \quad (2-19)$$

To invert the IPT image back for display we can invert Equation 2-18 for a single surround condition. This is shown in Equation 2-20. The LMS cone responses, after applying the surround tone reproduction functions, are then converted back into CIE XYZ tristimulus values.

$$\begin{aligned} L' &= L^{\frac{1}{0.43}} ; L \geq 0 \\ L' &= -|L|^{\frac{1}{0.43}} ; L \leq 0 \end{aligned} \quad (2-20)$$



2.2 Literature review

The image color appearance model, iCAM, was proposed by Fairchild and Johnson [2] as a model capable of predicting the appearance of spatially-simple color stimuli under a wide variety of viewing conditions. The model was applied to images by treating each pixel as an independent stimulus. Its revolutionary advances in color appearance modeling would require more rigorous treatment of spatial appearance phenomena. Several examples of the performance of iCAM were given in the study by Fairchild and Johnson including its prediction of simultaneous contrast, crispening, and spreading, high dynamic range tone mapping.

Johnson and Fairchild [3] described the use of iCAM in rendering high dynamic range images for display. Recent advances in color imaging have led to systems that are capable of capturing high dynamic range (HDR) scenes, as described by Debevec [26] and Xiao et al [27]. There has been research in development of tone reproduction algorithms for rendering high dynamic range images onto lower dynamic range display.

A number of algorithms related to scene rendering were investigated in a study by Moroney and Tastl [4]. They conducted an experiment to compare performance of McCann 99 Retinex and iCAM. Three high dynamic range grayscale images were rendered using the algorithms tested. Then ten observers performed a rank ordering of the printed rendering obtained. The results varied by image with iCAM being most preferred for one of the images, McCann 99 Retinex being most preferred for one of the images and a tie for all three rendering for the third image.

Wang et al. [28] described methods for assessing perceptual image quality traditionally attempted to quantify the visibility of error (differences) between a distorted image and a reference image using a variety of known properties of the human visual system. Under the assumption that human visual perception is highly adapted for extracting structural information from a scene, they introduced an alternative complementary framework for quality assessment based on the degradation of structural information. As a specific example of this concept, they developed a structural similarity index (SSIM) that compares local patterns of pixel intensities which have been normalized for luminance and contrast. The results showed that SSIM index agreed well with human perception of image quality.

In summary, the previous studies have proved the capability of iCAM to enhance image quality, but yet given a clear instruction as to what types of the specific low-pass filter should be used. In addition, in Moroney and Tastl's study, the performance of iCAM was tested in comparison with other algorithms for grayscale images, i.e. color were not considered in this study. The results from visual experiments revealed that iCAM should be implemented differently for different images, as it was found that iCAM performed best for only one out of three images. The present study thus aimed to determine the type of filter with respect to size and shape of the filter, in order to obtain an optimum image enhancement. Quantitative analysis using the SSIM index was employed to investigate the quality of images enhanced by iCAM, together with subjective analysis based on the method of rank order.