1 Introduction

Chromatic Adaptation is the ability of the human visual system to approximately preserve the appearance of an object color, no matter which light “color” illuminates a particular scene. We perceive a white piece of paper as white, regardless whether viewed under a bluish daylight or yellowish tungsten light source. Similar effects can be observed with other color hues.

Figure 1 illustrates this visual phenomenon. The left image is the same image as the middle image, except that it is overlayed by a bluish filter. While the image seems to have a color cast, i.e. we perceive all the colors to be too blue, we still can clearly identify the original hues. For example, the helmet still appears yellowish. In the right image, the filter was just overlayed on the helmet, which now appears greenish.

![Figure 1: An example of chromatic adaptation. See text for explanation.](image)

Image capturing systems, such as scanners and digital cameras, do not have the ability to adapt to an illumination source like the human visual system, as illustrated in Figure 2. The left
image shows an object, the Macbeth color rendition chart [MMD76], captured under a tungsten light source. When we view an image, we adapt to the prevalent illuminant inherent in the viewing conditions. The image will appear to have a color cast if we are not adapted to the same light source the object was illuminated with when the image was taken. The right image was transformed to the viewing illuminant we usually adapt to when we view an image on a CRT monitor and appears more natural to us.

![Macbeth Color Rendition Chart](image)

Figure 2: Left: An object imaged under a tungsten illuminant. Right: The same image transformed to appear correctly under monitor viewing conditions.

If the two helmets in the left and middle images of Figure 1 looked identical, we would have perfect color constancy. That is, we would be able to totally discount the effect of the blue filter, and the appearance of the helmet would depend only on its surface characteristics. However, as illustrated, that is not always the case. Indeed, the level of color constancy changes according to actual surface color, the viewing conditions, the scene composition, and the adaptation period. The appearance of a color under two different illuminants can therefore not simply be predicted by calculating only illumination change.

Figure 3 graphically illustrates this concept. A surface color under one adapting illuminant can be described with tristimulus coordinates $XYZ^1$. Changing the illuminant will result in coordinates $XYZ^2$. However, the appearance of the surface color under the second illuminant is best described with coordinates $XYZ^3$. The corresponding color coordinates of the surface color are $XYZ^1$ and $XYZ^3$, the distance between $XYZ^2$ and $XYZ^3$ is an indicator of the color inconstancy of the given surface color.

To faithfully reproduce the appearance of image colors, it thus follows that all image processing systems need to apply a transform that converts the input colors captured under the input illuminant to the corresponding output colors under the output illuminant. This can be achieved by using a chromatic adaptation transform (CAT). Indeed, the prevalent color management framework used in color image applications, developed by the International Color Consortium (ICC) [ICC04], contains a chromatic adaptation transform. Similarly in color science, when the appearance of a color under a different illuminant needs to be predicted, a color appearance model [MFH+02] containing
Figure 3: A graphical illustration of color constancy: If the human visual system is perfectly color constant, the effect of the illuminant could be discounted by simply modeling illumination change. However, most colors exhibit some degree of color inconstancy. The corresponding color of the sample under the reference illuminant \((XYZ)^1\), i.e. the color coordinates that best describe its appearance under the test illuminant, are given by \((XYZ)^3\).

A CAT is used. Basically, applying a chromatic adaptation transform to the color values under one adapting light source predicts the corresponding color values under another adapting light source.

2 Chromatic Adaptation Transforms (CATs)

Chromatic adaptation can be described as the mechanism of the human visual system to discount the influence of the illuminant SPD and to approximately preserve the appearance of object colors [Fai98]. Color constancy is achieved when the appearance of an object is determined only by its surface spectral reflectance [Mal01]. Thus, chromatic adaptation refers to a process of the human visual system, while color constancy refers to the physical state under which the process works. While we often assume that the visual system is “approximately color constant,” given by our daily visual experiences, there are also instances when color constancy fails. Such examples is illustrated in Figures 1 and 3.

Chromatic adaptation research has therefore been primarily concerned with finding a chromatic adaptation model, based on physiological knowledge of the human visual system and psychophysical experiments. Chromatic adaptation models (or transforms) aim to predict the appearance of a surface color under different illuminant conditions. Color constancy algorithms, on the other hand, aim to recover the physical illuminant, reflectances and/or sensor characteristics necessary to achieve color constancy.
2.1 Performance Measure

In the following sections, we introduce the chromatic adaptation models that were derived based on different experimental methods, starting with the von Kries model. The goal of a chromatic adaptation model, equivalently a chromatic adaptation transform (CAT), is to predict as well as possible the corresponding color of a stimulus $\rho^b$ under a test adaptation condition $A^b$, given the experimental measurements $\rho^a$ under a reference adaptation condition $A^a$:

$$\rho^b \approx \text{CAT}(\rho^a)$$

(1)

In other words, when applying a model or transform we obtain an estimate $\tilde{\rho}^b$ of the corresponding colors under the test conditions:

$$\tilde{\rho}^b = \text{CAT}(\rho^a)$$

(2)

The difference of $\tilde{\rho}^b$ to $\rho^b$ is an indicator of the appropriateness of a given model or transform.

In general, the quality of a chromatic adaptation model or transform, i.e. the prediction errors, is evaluated using Euclidean distances in a perceptually uniform color space, either in CIELUV, CIE $u', v'$, or in CIELAB.

2.2 Von Kries Chromatic Adaptation Model

Johannes von Kries is generally considered to be the first who attempted to describe the visual phenomenon of chromatic adaptation with a model. In [vK02], he investigates the additivity laws (see Colorimetry Lecture Notes) under different adaptation conditions and concludes that certain properties hold. He studied the effect of chromatic adaptation using an asymmetric matching experiment.

As illustrated in Figure 4, the basic concept of (simultaneous) asymmetric matching is to expose the observer to color stimuli under different adaptation conditions $A^a$ and $A^b$ [WS82]. The difference in adaptation conditions is usually limited to using a uniform, often neutral surround illuminated by two different illuminants, $E^a(\lambda)$ and $E^b(\lambda)$. The observer, using a colorimetric like device with two viewing areas matches under $A^b$ the perceived color stimulus $C_1(\lambda)$ under $A^a$. One eye is looking at the reference stimulus $C_1(\lambda)$ under $A^a$, and the other eye is looking at a viewing area under $A^b$, which contains a field with a mixture of red, green and blue primaries that are adjusted by the observer to match the reference stimulus. The resulting match is $C_2(\lambda)$. $C_1(\lambda)$ and $C_2(\lambda)$ are corresponding color signals.

Von Kries observed that if $C_1(\lambda), A^a$ under adaptation condition $A^a$ has the same appearance as $C_2(\lambda), A^b$ under adaptation condition $A^b$, and equivalently for $C_3(\lambda), A^a$ and $C_4(\lambda), A^b$, then ($C_1(\lambda) + C_3(\lambda)), A^a$ has the same appearance as ($C_2(\lambda) + C_4(\lambda)), A^b$. Equally, $\alpha C_1(\lambda), A^a$ matches $\alpha C_2(\lambda), A^b$.

Comparing his theorem of proportionality [vK02] with the trichromacy laws (see Colorimetry LN), von Kries deduced that additivity and proportionality hold for color responses independent of
adaptation, if adaptation can be modeled by modifying the sensor sensitivities. He concluded that color changes caused by adaptation can be explained as a modification of the spectral sensitivities of the cone fundamentals, i.e. $\mathcal{A}^a \mapsto \{L^a, M^a, S^a\}$ and $\mathcal{A}^b \mapsto \{L^b, M^b, S^b\}$. He further assumes that the cone responses are independent from one another, each adapting exclusively to its own function. In \cite{vK70}, he states that the change of stimuli appearances can be modeled by a multiplicative coefficient that is a function of the higher or lower excitability of the individual cones. Von Kries does not explicitly state how the coefficients should be calculated, except to say that they are dependent on the adaptation condition. He also mentions that his model is a simplification that might not hold for all conditions.

The von Kries's chromatic adaptation model can therefore be expressed as follows: the cone responses $L^a, M^a, S^a$ of a color stimulus under adaptation condition $\mathcal{A}^a$ can be mapped to its illuminant-invariant descriptors $d_1^a, d_2^a, d_3^a$ with three adaptation dependent gain factors $g_1^a, g_2^a, g_3^a$
that are independent for each channel. Mathematically, it follows that:

\[
\begin{align*}
\mathbf{{d}}_1^o &= g_1^a \mathbf{{L}}^a \\
\mathbf{{d}}_2^o &= g_2^a \mathbf{{M}}^a \\
\mathbf{{d}}_3^o &= g_3^a \mathbf{{S}}^a \\
\end{align*}
\]

or, expressed as a Diagonal Matrix Transformation (DMT):

\[
\begin{bmatrix}
\mathbf{{d}}_1^o \\
\mathbf{{d}}_2^o \\
\mathbf{{d}}_3^o \\
\end{bmatrix} = \begin{bmatrix} g_1^a & 0 & 0 \\ 0 & g_2^a & 0 \\ 0 & 0 & g_3^a \\
\end{bmatrix} \begin{bmatrix}
\mathbf{{L}}^a \\
\mathbf{{M}}^a \\
\mathbf{{S}}^a \\
\end{bmatrix}
\]

(4)

Equivalently, \( L^a, M^a, S^a \) and \( L^b, M^b, S^b \) of two color stimuli that appear identical when regarded under the two different adaptation conditions \( A^a \) and \( A^b \) are related by the ratio of their coefficients:

\[
\begin{bmatrix}
\mathbf{{L}}^b \\
\mathbf{{M}}^b \\
\mathbf{{S}}^b \\
\end{bmatrix} = \begin{bmatrix} \frac{g_1^a}{g_1^b} & 0 & 0 \\ 0 & \frac{g_2^a}{g_2^b} & 0 \\ 0 & 0 & \frac{g_3^a}{g_3^b} \\
\end{bmatrix} \begin{bmatrix}
\mathbf{{L}}^a \\
\mathbf{{M}}^a \\
\mathbf{{S}}^a \\
\end{bmatrix}
\]

(5)

Von Kries assumes that his model is valid for cone responses. Thus, to use his model with any colorimetric measurements, i.e. tristimulus values XYZ or colorimetric RGB values, eq. 4 can be extended to:

\[
\begin{bmatrix}
\mathbf{{R}}^b \\
\mathbf{{G}}^b \\
\mathbf{{B}}^b \\
\end{bmatrix} = \mathbf{{M}}^{-1} \mathbf{{D}} \begin{bmatrix} \mathbf{{R}}^a \\
\mathbf{{G}}^a \\
\mathbf{{B}}^a \\
\end{bmatrix}
\]

(6)

where \( \mathbf{{M}} \) is a nonsingular (3x3) matrix linearly transforming XYZ or RGB values to cone responses, \( \mathbf{{M}}^{-1} \) its inverse and \( \mathbf{{D}} \) the diagonal matrix containing the gain coefficients.

It is generally accepted today that a von Kries chromatic adaptation model is, at first approximation, able to model chromatic adaptation. By defining all the parameters of the model described in eq. 6, it is possible to design a chromatic adaptation transform (CAT) that predicts color appearance under different illuminants.

The parameters to define are the linear transform \( \mathbf{{M}} \) and the scaling coefficients \( g_1^a, g_2^a, g_3^a \) contained in the \( 3 \times 3 \) diagonal matrix \( \mathbf{{D}} \). According to von Kries, we would assume that the linear transform \( \mathbf{{M}} \) is used to map from tristimulus values XYZ or colorimetric RGB values to cone responses. In other words, the sensors are broad band and are equal (or at least close) to cone fundamentals. While this makes sense from the point of view of human visual processing, other sensors are more optimal, as will be discussed below.

### 2.3 Scaling Coefficients

As discussed in section 2.2, von Kries states that the scaling coefficients are dependent on the adaptation condition, but are independent for each visual channel. Since von Kries, many researchers
have studied what these adaptation conditions \( A \) are. Depending on their experimental data, they have either based the chromatic adaptation model on illuminant information alone, or taken the surround of a stimulus into consideration. It is known that both have an effect on the color appearance of a stimulus, but are both adaptation phenomena, or is just the influence of the the illuminant important?

Ives [Ive12] early on hypothesized, without mentioning von Kries, that adaptation is a function of the illuminant “color.” Long exposure to one color decreases the cone sensitivity, while the other sensitivities increase. Land and McCann and West and Brill [WB82] follow this assumption by defining the scaling coefficients to be the inverse of the color response of a white patch within the scene. Thus, the illuminant independent descriptor \( d_k^o(x) \) of a surface at position \( x \) with cone response \( \rho_k^a(x) \) in channel \( k \) under illuminant \( E^a(\lambda) \) is calculated as follows:

\[
d_k^o(x) = \frac{1}{\rho_w^k} \rho_k^a(x) \tag{7}
\]

where \( \rho_w^k \) is the cone response in channel \( k \) of the white patch. If, for that surface, \( S(\lambda) = 1 \) is valid over the whole visible spectrum, the color response is dependent only on the SPD of the illuminant. This adaptation is often referred to as the von Kries-Ives model [Mal01] and is used in modern color science and color imaging CATs.

From a human visual processing point of view, however, this way of calculating the coefficients seems too simple. It is well known that the surround of a stimulus has a strong influence on its appearance, an effect known as simultaneous contrast. Thus, it seems unreasonable to neglect the effect of the surround. Helson [Hel34, Hel38] proposed that the coefficients should be calculated as the inverse of the average cone responses, independently for each cone class, of the visual field. Thus,

\[
d_k^o(x) = \frac{1}{\bar{\rho}_k} \rho_k^a(x) \tag{8}
\]

where \( \bar{\rho}_k \) is the arithmetic mean of the color response in channel \( k \). Note that if the average of the surround reflectances is neutral, then the difference in average cone response and the cone response of the adapting illuminant is a scale factor that cancels out if responses are mapped from one illuminant to another.

Both the von Kries-Ives and von Kries-Helson model assume the coefficients are dependent only on signals from the same channel, and are not influenced by the responses of the other channels. In recent literature [BW92, DB00], this assumption is called the strong von Kries coefficient model (see Figure 5). Some studies, however, found better predictions if they relaxed the independence criterion and allowed for signals from other channels to influence the gain factors for one cone class.

Brainard and Wandell [BW92] therefore formulated a more general linear model, which they call the weak form of the von Kries coefficient model. As illustrated in Figure 5, they allow for the gain factors to be determined by signals from all three color channels. In their interpretation, the
gain factors are only dependent on the illuminant’s cone responses. More generally, the weak von Kries model can be expressed as:

\[
d^w_k(x) = \frac{1}{g^w_k} \rho^w_k(x); \quad g^w_k = f(\rho^w_r, \rho^w_g, \rho^w_b)
\] (9)

where the scaling coefficient \( g^w_k \) is a function of \( \rho^w_r, \rho^w_g, \rho^w_b \), the color signals of the illuminant.

\[
\begin{bmatrix}
X^b \\
Y^b \\
Z^b
\end{bmatrix} = M^{-1}DM
\begin{bmatrix}
X^a \\
Y^a \\
Z^a
\end{bmatrix}
\] (10)

Figure 5: Strong and week von Kries coefficient models. Left: The strong von Kries model assumes that the gain is dependent only on signals from the same cone class. Right: The weak von Kries model assumes that the gain is dependent on signals from all cone classes. The illustrations are taken from [BW92].

In general, modern chromatic adaptation models and transforms thus assume that adaptation can be modeled by illuminant-dependent coefficients alone, either independent for each channel or dependent on all three color responses. Other studies considered the surround and calculated the illuminant-independent descriptor based on a background factor, or as a spatial interaction of the stimulus with many surround surfaces.

### 2.4 Strong von Kries Coefficient Model

The chromatic adaptation transforms used in color imaging applications to map from one adapting illuminant to another, such as in the color management framework developed by the International Color Consortium (ICC) [ICC04], are based on the strong von Kries-Ives coefficient model. The same applies to today’s color appearance models [MFH+02]; when the appearance of a color under a different illuminant needs to be predicted, the chromatic adaptation transform applied is von Kries-Ives. Starting from \( XYZ \) tristimulus values:
\[
D = \begin{bmatrix}
R_a w & B_a w \\
G_a w & G_b w \\
B_a w & B_b w
\end{bmatrix}
\]

Quantities \(R_a w, G_a w, B_a w\) and \(R_b w, G_b w, B_b w\) are computed from the tristimulus values of the reference and test illuminants by multiplying the corresponding XYZ vectors by \(M\).

However, the scaling is not applied to cone responses, but to modified cone responses. In other words, the matrix \(M\) does not transform the tristimulus values \(XYZ\) to cone responses \(LMS\), but to modified cone responses, i.e., color responses that are more narrowband than cones. The following sections describe two different approaches to finding the optimal \(M\) transform.

### 2.5 Corresponding Colors

Before we can develop new models for chromatic adaptation, we must first have experimental measurements of pairs of stimuli (corresponding colors) that match in appearance with respect to different adaptation conditions. Corresponding colors can be described as a pair of tristimulus values \((XYZ)\) or cone responses \((LMS)\), based on one physical stimulus that appear to be the same color when viewed under two different illumination sources [Hun95]. Below we describe the psychophysical experiments conducted to collect such stimuli pairs.

Chromatic adaptation is a color appearance phenomenon, and it is well known that color appearance could depend on many factors, such as local surround, spatial and chromatic scene arrangement, adaptive luminance, etc. [Fai98]. It is very difficult to “distill” the effect of chromatic adaptation from all other color appearance mechanisms, and as such define an experimental protocol for it. Thus, several experimental methods have been proposed that are briefly described below.

#### 2.5.1 Asymmetric Matching

Many studies investigating chromatic adaptation models used asymmetric matching experiments. The basic concept of asymmetric matching is to expose the observer to color stimuli under different illuminants \(E^a(\lambda)\) and \(E^b(\lambda)\). The observer then matches under \(E^b(\lambda)\) the perceived color under \(E^a(\lambda)\) [WS82]. Simultaneous asymmetric matching refers to the observer being adapted to two different illuminants at the same time. Some early experiments use haploscopic (also called dichoptic) color matching where the observers use a colorimeter like device (see Figure 4) with two viewing areas. One eye is looking at a test light surrounded by an area illuminated by \(E^a(\lambda)\), and the other eye is looking at a reference light, surrounded by an area illuminated by \(E^b(\lambda)\). The viewing area under \(E^b(\lambda)\) contains a field with a mixture of red, green and blue primaries that can
be adjusted to match the test stimulus. The uniform (and often neutral) surround is adjusted for
different luminance conditions.

The advantages of this method is its simplicity and high precision. Disadvantages are that both
eyes are adapted to different illuminants. As signals of both eyes are mixed at the level of the
LGN, it is not clear what the effect of this interaction is. Thus, MacAdam [Mac56] used a single
bipartite viewing field where both halves show different adapting conditions. The observer fixates
to the partition line. In this case, the same retinal area in both eyes is viewing the same adaptation
conditions. Such a set-up is called monocular matching.

In successive asymmetric color matching, the observer views a single scene first under $E^a(\lambda)$
and then $E^b(\lambda)$. The observer adjusts a colored patch under $E^b(\lambda)$ to appear the same as under
$E^a(\lambda)$. Experiments have been done using two CRT monitors calibrated to different illuminants,
or one CRT monitor to match hardcopy. Obviously, the accuracy of these experiments depend
on the ability of the observers to remember color accurately. Some studies used a combination of
successive and haploscopic matching: one eye was adjusted to one viewing conditions, while the
other eye successively matched the stimuli.

2.5.2 Achromatic Matching

In achromatic matching, the observers adjust a specified color to appear neutral. This method is
less exact in predicting the remapping of colors due to illumination change. It provides coordinates
of points in the observer’s achromatic locus, but not exactly which point. However, Speigle and
Brainard [SB99] reported after comparing the two different methods that achromatic adjustments,
together with a gain-control model, can be used to make accurate predictions of the chromaticity
of asymmetric matches.

2.5.3 Memory Matching and Magnitude Estimation

In memory matching experiments, observers first learn to describe colors in terms of hue, lightness
and chroma of the Munsell color ordering system [WS82, Mun76]. The training is time intensive,
but a trained observer is reasonably accurate in predicting Munsell coordinates. Arend [Are93] used
a combination memory-achromatic matching experiment. He trained the observers on the Munsell
color system, and then had them create neutrals or unique hues under different illuminants on a
CRT monitor, as they remembered from the Munsell coordinates.

In magnitude estimation (or color naming) experiments, the observers are asked to rate the
lightness, colorfulness and hue of a stimulus. These values are then projected to a chromaticity
diagram, from where tristimulus values can be derived at equal luminance, i.e. equal tristimulus
value $Y$. 
2.6 CATs derived by minimizing perceptual error

Several CATs that are used in color imaging applications are based on minimizing perceptual error ($\Delta E$) over one or more corresponding color data sets. Lam [Lam85], for example, minimized $\Delta E_{Lab}$ over the experimental results he obtained from a memory experiment under illuminants D65 and A. The matrix $M$ is as follows:

$$M = \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix}$$

The corresponding sensors are illustrated in Figure 6.

The original Bradford CAT was not a linear transform according to the von Kries model, it contained a non-linearity in the blue [Lam85]. However, mostly out of computational reasons, that non-linearity has been suppressed and the linear Bradford CAT is as formulated in eq. 10.

One implication of the Bradford chromatic adaptation transform is that color correction for illumination takes place not in cone space but rather in a “narrowed” cone space. The Bradford sensors (the linear combination of XYZs defined in the Bradford transform) have their sensitivity more narrowly concentrated than the cones (see Figure 6). Additionally, the long Bradford sensor has its peak sensitivity shifted to longer wavelength.

Subsequent studies later also proposed linear CATs in a narrower cone space. Fairchild [Fai01], in an effort to simplify the color appearance model CIECAM97s and make it reversible, proposed a linear CAT that most closely performed to the original non-linear Bradford transform in terms of perceptual error. He used Munsell samples to calculate corresponding colors under illuminants A and D65 using the non-linear Bradford CAT. He then developed a linear CAT by minimizing the CIELAB differences to the predictions of the Bradford CAT on this corresponding color data set. Li et al. [LLRH02] derived a new CMCCAT2000 based on minimizing perceptual error ($\Delta E$) over a set of corresponding color data compiled by Luo and Rhodes [LR99]. A slightly different transform was derived by excluding the data set of McCann et al. [MMT76]. This latter CAT was adopted by the CIE as the chromatic adaptation transform CAT02 for a new color appearance model CIECAM02 [MFH02]. The linear transformations $M$ are as follows:

$$M_{Fair} = \begin{bmatrix} 0.8562 & 0.3372 & -0.1934 \\ -0.8360 & 1.8327 & 0.0033 \\ 0.0357 & -0.0469 & 1.0112 \end{bmatrix}$$

$$M_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6974 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix}$$
The corresponding sensors are illustrated in Figure 6. The three transforms, linear Bradford, Fairchild and CAT02 result in very similar sensors, slightly more narrowband than the cone responses, and containing some negative values. For comparison, the Hunt-Pointer-Estevez (HPE) [Hun98] cone fundamentals are also plotted. The linear transform from XYZ to LMS is as follows:

$$
M_{HPE} = \begin{bmatrix}
0.3897 & 0.6890 & -0.0787 \\
-0.2298 & 1.1834 & 0.0464 \\
0 & 0 & 1
\end{bmatrix}
$$

Figure 6: The normalized sensors resulting from different chromatic adaptation transforms: HPE (solid color), Bradford (dash), Fairchild (dotted), CAT02 (dash-dot).

3 Spectral Sharpening

The theoretical limits of the strong von Kries coefficient model has primarily been studied in the context of computational color constancy. Computational color constancy algorithms either try to estimate the adapting illuminant from a scene and/or estimate surface reflectance, to find appropriate sensors, or to render image appearance under different illuminants.

Mathematically, the assumption that the exact illuminant SPD and spectral reflectances can be recovered from color responses infers that the product of a color response of an illuminant $$\rho^E_k(x)$$ and the color response of a surface $$\rho^S_k(x)$$ (under equi-energy illuminant, i.e. $$E(x, \lambda) = 1$$) are equal
to the color response of their product:

$$\rho_k^{E,S}(x) = \rho_k^E(x)\rho_k^S(x)$$  \hfill (11)

However, this is true only when the sensor sensitivities are independent, implying very narrow-band filters that are in the extreme Delta functions responsive only at one wavelength. For example, if the sensor $k$ has non-zero response only at wavelength $\lambda_i$, the image formation equation can be written as:

$$\rho_k(x) = \int \omega S(x, \lambda)E(x, \lambda)\delta(\lambda_i)d\lambda = S(x, \lambda_i)E(\lambda_i)\delta(\lambda_i)$$  \hfill (12)

In this case, a change in illuminant changes $\rho_k(x)$ only by a scale factor.

However, such sensor sensitivities apply neither for the human visual system, nor for any camera sensors that have quantum efficiency constraints. Therefore, additional assumptions about “the world” have been investigated to model color constancy. For an extended discussion on illuminant or surface reflectance recovery, see [Hur98, Hor99, Mal01].

Finlayson et al. [FDF94b, FDF94a, FS00b, FS00a] have shown that a generalized coefficient model can be successfully applied in chromatic adaptation and color constancy. In their model, the sensors sensitivities are first transformed by a linear (3x3) matrix to a new set before the diagonal transform is applied. Equivalently, the transform can be directly applied to the color responses. With reference back to eq. 6, the generalized coefficient model can be written as

$$\begin{bmatrix} d_1^a \\ d_2^a \\ d_3^a \end{bmatrix} = \begin{bmatrix} g_1^a \\ g_2^a \\ g_3^a \end{bmatrix} \begin{bmatrix} t_{1,1} & t_{1,2} & t_{1,3} \\ t_{2,1} & t_{2,2} & t_{2,3} \\ t_{3,1} & t_{3,2} & t_{3,3} \end{bmatrix} \begin{bmatrix} R^a \\ G^a \\ B^a \end{bmatrix}$$  \hfill (13)

Equivalently:

$$d^a = DT\rho^a$$  \hfill (14)

Note that the fundamental difference between eq. 6 and eq. 13 is that the von Kries model assumes adaptation is a function of cone responses, while Finlayson et al. do not put any restrictions on the sensors. They found that the optimal transform $T$ is “sharp,” i.e. the resulting sensors are more narrowband and decorrelated than the cone fundamentals (see Figure 7).

While there are many applications of spectral sharpening in imaging, we will limit the discussion here to finding the appropriate color space, i.e. the appropriate transform matrix $M$ for a DMT-type von Kries-Ives chromatic adaptation transform.

Let $W^a 3 \times n$ be a set of tristimulus values under illuminant $E^a(\lambda)$ and $W^b$ a set of tristimulus values under illuminant $E^b(\lambda)$. According to von Kries, the two sets should be approximately equal, subject to a diagonal matrix transform (DMT):

$$W^b \approx \Lambda^{b,a}W^a$$  \hfill (15)
where \( \Lambda^{b,a} \) is a diagonal transform. The idea of spectral sharpening is that the residual error in the mapping can be reduced if both data sets are first linearly transformed by a sharpening matrix \( T \):

\[
TW^b \approx \Lambda^{b,a}TW^a
\]  

(16)

Solving for \( \Lambda^{b,a} \) requires minimizing some error criterion. Using least-squares, \( \Lambda^{b,a} \) can be solved for by the Moore-Penrose inverse:

\[
\Lambda^{b,a} = TW^b (TW^a)^+ 
\]  

(17)

where \([.]^+\) denotes the pseudoinverse [GvL96]. The solution for \( T \) must ensure that \( \Lambda^{b,a} \) is diagonal. Rearranging eq. 17, we obtain:

\[
[T]^{-1}\Lambda^{b,a}T = W^b[W^a]^+ 
\]  

(18)

The solution for \( T \) is therefore in the eigenvector decomposition of:

\[
W^b[W^a]^+ = U^{b,a} \Lambda^{b,a} [U^{b,a}]^{-1} 
\]

Thus, \( T \) is equal to \([U^{b,a}]^{-1}\).

Note that the best general transform that maps \( W^a \) to \( W^b \) in the least-squares sense is found by:

\[
B^{b,a} = W^b(W^a)^+ = W^b[(W^a)^T(W^a)]^{-1}(W^a)^T 
\]  

(19)

Thus, eq. 18 can be interpreted as simply being the eigenvector decomposition of the general transform \( B^{b,a} \). What is important to note though is that if the sharpening transformation \( T \) is first applied to the color responses \( W^b, W^a \), then the optimal transformation is the diagonal transform \( \Lambda^{b,a} \).

Thus, using corresponding color data (for this example the data set of Lam [Lam85], which contains 58 XYZ values under illuminants A and D65), the prediction of the corresponding colors under D65 using a von Kries-Ives DMT chromatic adaptation model should approximately equal

\[
TW^b \approx D^{b,a}TW^a
\]  

(20)

where \( W^b \) is a 3 x 58 matrix of corresponding color XYZs under illuminant D65, \( W^a \) is a 3 x 58 matrix of the measured XYZs under illuminant A and \( D^{b,a} \) is the diagonal matrix formed from the ratios of the two sharpened white-point vectors \([R^D_{65}, G^D_{65}, B^D_{65}]\) and \([R^A_w, G^A_w, B^A_w]\), derived by multiplying vectors \([X^D_{65}, Y^D_{65}, Z^D_{65}]\) and \([X^A_w, Y^A_w, Z^A_w]\) with \( T \).

The matrix \( T \) is derived from the matrix \( B^{b,a} \) that best maps \( W^a \) to \( W^b \) minimizing least-squares error. In the example shown above, it minimizes least-squares error in XYZ, as the Lam data set uses XYZ coordinates to describe the corresponding colors. Thus, it does not minimize a perceptual error as the other methods described above.

While \( B^{b,a} \) calculated using eq. 19 results in the smallest XYZ mapping error, it will not fulfill the requirement that particular colors are mapped without error, e.g. preserving achromaticity for
neutral colors. We are very sensitive to color casts in neutral (grey) colors, and thus need to define a transform that keeps the mapping errors of colors around the neutral axis very small.

Therefore, we can derive $B^{b,a}$ using a white point preserving least-squares regression algorithm [FD97a, FD97b]. The intent is to map the values in $W^a$ to corresponding values in $W^b$ so that the RMS error is minimized subject to the constraint that, as an artifact of the minimization, the achromatic scale is correctly mapped. In order to preserve white:

\[ B^{b,a} = D + (ZN) \]  

where $D$ is the diagonal matrix formed from the ratios of the two white point vectors $[X_{w}^{D65}, Y_{w}^{D65}, Z_{w}^{D65}]$ and $[X_{w}^{A}, Y_{w}^{A}, Z_{w}^{A}]$, respectively. $Z$ is a $3 \times 2$ matrix composed of any two vectors orthogonal to the $[X_{w}^{A}, Y_{w}^{A}, Z_{w}^{A}]$ vector. $N$ is obtained by substituting $Z$, $N$, and $D$ in eq. 19 and solving for $N$. The sharpening transform $T$ can then be derived through eigenvector decomposition of the general transform $B^{b,a}$:

\[ B^{b,a} = U^{b,a} A^{k,a}[U^{b,a}]^{-1} \]  

where $T$ is equal to $[U^{b,a}]^{-1}$.

The predicted corresponding colors under illuminant D65 of Lam’s 58 samples, using this Sharp transform, are calculated with eq. 20.

Numerically, the linear transformation $T$, which is $M_{Sharp}$ [FS00b, FS00a], is equal to:

\[
M_{Sharp} = \begin{bmatrix}
1.2694 & -0.0988 & -0.1706 \\
-0.8364 & 1.8006 & 0.0357 \\
0.0297 & -0.0315 & 1.0018
\end{bmatrix}
\]  

The corresponding sensors are illustrated in Figure 7.

4 Comparison of the Sharp CAT with other linear CATs

Applying the resulting Sharp CAT, derived via data-based sharpening of the corresponding colors of the 58 Lam samples under illuminants A and D65 minimizes the RMS error between corresponding XYZs. It also yields sensors that are visibly sharper than those implied by the Bradford transform (see Figure 7). However, what we are most interested in is to compare the perceptual error between actual appearance and predicted appearance of a color under different illuminants using the Sharp and the other linear transforms, an we should thus compare errors in a more perceptual color space encoding (see section 2.1).

We can show [FS00b, FS00a] that the different chromatic adaptation transforms (CAT02 and Sharp) are not statistically significantly different from each other. However, the color spaces in which the scaling takes place is quite different. So how can we choose the right transformation matrix $M$, i.e. which CAT should be used?
This is a question that is not answered today. Even more, Finlayson and Susstrunk [FS01] have shown that there is an indefinite number of CATs, i.e. $M$, that give the same corresponding color prediction performance when using a DMT-type von Kries-Ives chromatic adaptation model.

The reasons why so many different CATs give statistically the same results is manyfolds. As mentioned above, creating a “true” corresponding color data set through psychovisual experiments is very difficult. Also, modeling a human visual effect with a diagonal matrix transform is simplifying the biological process enormously, thus several solutions are possible. Additionally, a small deviation from the “perfect” transform is visually acceptable in color imaging, where the distribution of colors is relatively large.

Currently, CAT02 is most used, as it is part of the CIECAM02 color appearance model. However, there are some problems with boundary conditions, it is very well possible that in the new future, a new CAT will be standardized and adopted.

5 Illuminant Estimation

So far, we have assumed that we know the illuminant of a scene, and can use a chromatic adaptation transform to map from one illuminant to another. However, recall from the Multispectral Imaging lecture that the color signal $C(\lambda) = E(\lambda)S(\lambda)$ is ambiguous. In other words, it is difficult to deduce
from the color responses of a camera with three sensors what exactly the contribution from light and surface are.

Yet, as discussed above, we adapt to the prevalent illuminant of the scene and are able to discount most of its effect. To do effective rendering that result in pleasing reproductions, we thus need to estimate the illuminant in a scene and compensate for it.

Recall from previous discussions (Color Spaces and Color Encoding lecture) that neutral (white, grey) colors have equal code values: \( R = G = B \). If a scene is captured under an equi-energy illuminant \( E(\lambda) = 1 \), and the sensor sensitivities are equal (i.e. have equal quantum efficiency), then it is evident that a surface with a reflectance factor \( S(\lambda) = \text{const} \) will result in equal color responses and thus equal code values. If one (or both) of these two conditions is not met, than the camera response values for a neutral surface are not equal, i.e. \( R \neq G \neq B \). The visual result is that an image that is not viewed under the exact same illuminant conditions as the image was captured will appear to have a color cast, such as illustrated in Figure 2.

In most capturing situations, both conditions are not met. However, as the engineer of a camera, you will know the difference in average quantum efficiency of your camera sensors, either by measuring them or by taking the theoretical values given by the sensor manufacturer. However, as the illuminants in real world scenes can vary dramatically, any camera needs to contain one (or more, as some camera manufacturers do) illuminant estimation method.

5.1 Algorithms to estimate the illuminant

There are many algorithms published that try to estimate the illuminant in a scene. Covering all of them would go to far for this course. We resume a few approaches in the following paragraphs, and look at three algorithms, namely \textit{max RGB}, \textit{greyworld}, and \textit{color by correlation} in a bit more detail.

5.1.1 Lightness Algorithms

\textit{Lightness algorithms} [Lan64, Lan77, Lan86b, Hor74, Hur86, BB87] presume that adaptation behaves according to eq. 3, i.e. that each receptor response is independently scaled by some scene-dependent factor. Edwin Land [Lan64, LM71, Lan74, Lan77, Lan83, Lan86a] pioneered the work on lightness algorithms with his \textit{retinex} model of color vision. He observed during experiments with Mondrian scenes that the perceived colors of surface reflectances do not depend on the radiant energy of their stimuli, but that their \textit{relative lightness rank-ordering} remained invariant independent of the illumination uniformity and color. For a review of his experiments and early algorithm, see Appendix A.

In general, the lightness value for a sensor \( k \) is calculated by finding the average ratio between the color response of a position \( x_0 \) and many surrounding positions. Early retinex algorithms
calculated the ratios along a path [Lan64], while later implementations use a weighted average of the surround [Lan86a]. As discussed by McCann et al. [MMT76], there are some parallels to human visual processing. The retinex algorithm contains a local normalization due to the influence of the neighboring pixels, which could be related to the ganglion cell receptive fields. The spatial normalization over the whole visual field could be explained by the double-opponent fields found in the visual cortex.

Setting the paths’ lengths and the number of paths to infinity is equivalent to normalizing the surface reflectance at position \( x_0 \) by the geometric mean of all surface reflectances [BW86]. As the triplet of computed lightness values for the three spectral channels should define the illuminant-independent descriptor of a color patch, the algorithm implicitly assumes then that the mean reflectance is the same in each channel and for every scene. This constraint, called *gray-world*, assumes that the (geometric) mean surface reflectance of each scene in each spectral channel is the same: gray. This assumption also appears in other color constancy algorithms [Buc80, DL86, GJT88], and has been used for many color reproduction algorithms [Hun95]. It is, in spirit, close to Helson’s [Hel34] proposal that the von Kries coefficients \( g_1^a, g_2^a, g_3^a \) are inversely proportional to the (arithmetic) average photoreceptor excitations within a cone class under adaptation condition \( A^a \).

The lightness values are not accurate illuminant-independent descriptors if the mean scene reflectance is not gray. Land and McCann [LM71, MMT76] therefore introduced a normalization into the path equation. To recover illuminant-independent lightness descriptors, this normalization assumes that the brightest patch in each channel is a perfect reflector, i.e. it reflects all illuminant energy. This *maxRGB* assumption is close to the von Kries-Ives model, which also predicts that the coefficient is based solely on the illuminant.

### 5.1.2 Linear Models for Illuminant Reflectance Recovery

The limitations of lightness algorithms has inspired the development of other approaches. Many color constancy algorithms assume that illuminant SPDs and surface reflectance functions are “smooth” over the visible spectrum and can therefore be represented as finite-dimensional linear models.

Recall from the Multispectral Imaging lecture that if both illuminant SPD and surface reflectance can be expressed with a linear model, the color response \( \rho \) of a surface is equal to:

\[
\rho = R \text{diag}(B_c \epsilon) B_s \sigma
\]  

(24)

where \( R \) is a \( k \times n \) matrix containing the sensor sensitivities and \( \rho \) a \( k \)-dimensional vector containing the sensor responses. The \( n \times m \) matrix \( B_c \) contain the \( m \) basis function for the illuminant, with \( n \) samples over the visible spectrum. \( \epsilon \) is a \( m \times 1 \) vector containing the illuminant weights.
Equivalently, the \( n \times m \) matrix \( \mathbf{B} \) contain the \( n \) basis function for the surface reflectance and \( \sigma \) the surface weights. Eq. 24 can be rewritten as:

\[
\rho = \Lambda \epsilon \sigma \tag{25}
\]

where \( \Lambda \) is called the *lighting matrix*, dependent only on the illuminant (and the reflectance basis functions), and \( \sigma \) is the vector of surface weights for a given surface reflectance.

If the illuminant is known, \( \Lambda \) is known, and the recovery of the \( m \) surface weights can be done by solving a set of linear equations. If \( m = k \), this reduces to matrix inversion:

\[
\Lambda^{-1} \rho = \sigma \tag{26}
\]

In human vision and trichromatic imaging systems, \( k = 3 \). Three basis functions is not necessarily sufficient to find good approximation of spectral reflectances. However, if \( m > k \), eq. 26 is underdetermined and there is no unique solution.

All the linear-basis algorithms that try to estimate the illuminant have in common that they assume a uniform illumination over the scene, i.e. \( \Lambda \) is constant at each position in the scene. They do not explicitly assume a diagonal coefficient mapping of the form of eq. 4. To solve for the 9 coefficients of \( \Lambda \), they exploit the redundancy in the illuminant’s contribution to the color signals sampled from several different surfaces in the scene, and apply additional constraints. Briefly, Buchsbaum [Buc80], for example, makes the assumption that the spectral average of a scene is constant and equal to that of an achromatic gray, similar to Land’s retinex. The lightness matrix \( \Lambda \) can then be recovered by first recovering the equivalent \( \Omega \sigma \) surface weight matrix. To avoid some of the gray-world failures, Gershon et al. [GJT88] first segment the image so that each (material) surface reflectance only counts once. Thus, the average of surface reflectance is independent of the spatial distribution of the reflectances. Maloney and Wandell [MW86] restrict the number of the reflectance basis functions allowed, i.e. \( K = N + 1 \). That leaves two reflectance basis functions for a trichromatic imaging system. D’Zmura [D’Z87], using a similar model, could recover three surface weights, but required that the same image is taken under two different illuminants. For a extended discussion of these algorithms, see [Hur98, Hor99, Mal01].

Forsyth [For90] and Finlayson, Hordley, and Hubel [Fin96, FH98, Hor99, FHH01] propose a different approach to recovering scene illuminants, recognizing that finding a unique solution to the underdetermined eq. 26 puts too many constraints on the real world and/or makes the algorithms fail too often. These *gamut-mapping* color constancy algorithms do not attempt, at first instance, to find a unique solution. Rather they attempt to find all possible solutions, and from this set choose the best solution. Such an approach is described below as *color by correlation*.

Note that there are other algorithms not discussed here, some based on neural networks, some based on probabilistic frameworks, and some based on real surface characteristics (as opposed to the lambertian surfaces assumed so far).

\(^1\)Recovery is also possible when \( m < k \), a technique often used in multispectral imaging.
5.2 Max RGB, Greyworld, and Color by Correlation

5.2.1 Max RGB

The simplest algorithm (and also one of the most often used) is max RGB. The assumption here is, as already discussed above, that each scene contains a white patch (i.e. a surface whose reflectance factor $S(\lambda) \approx 1$). Thus, most of the light will be reflected, and the color response of this surface is related to the color response of the illuminant $\rho^E$. Thus:

$$\text{max}(\text{RGB}) \mapsto \rho^E$$ (27)

Note that if the quantum efficiencies of the three camera sensors are equal, the operator “$\mapsto$” can be replaced with the “$=$” sign.

Clearly, this approach fails if there is no white surface in the scene.

5.2.2 Greyworld

Another simple algorithm is derived from the idea, also mentioned above, that the average (or geometric mean) of a scene is grey. Thus, change from equal RGB values are an indicator of the illuminant’s color response:

$$\text{mean}(\text{RGB}) \mapsto \rho^E$$ (28)

Clearly, this is also not true for many scenes, especially if there is a predominant background color (green grass, blue sky, etc.). Gershon et al. [GJT88] thus modified the algorithm by segmenting the image into patches of equal surface colors before estimating the illuminant. The color response from each patch is then only counted once in the averaging, so that surfaces of different sizes are given equal weight. This can be easily implemented by creating a histogram of the image RGBs:

$$\text{mean}(\text{histRGB}) \mapsto \rho^E$$ (29)

5.2.3 Color by Correlation

Finlayson et al. [FHH01] used the property of any imaging system that if illuminant and sensor sensitivities are known, a gamut of color chromaticities can be established (see Colorimetry LN), as illustrated in Figure 8. While any chromaticity space could be used, they chose to use a simple $r/g$ space, with $r = R/(R+G+B)$ and $g = G/(R+G+B)$, respectively. The gamuts are precalculated for different illuminants, using the sensor sensitivity of the camera in question. For each considered illuminant, a probability distribution is calculated, which gives the likelihood of observing a given image color under a given light. These probability distributions are then quantized, and a correlation matrix is built that contains a column vector of probabilities for each illuminant considered (see Figure 9). Solving for the illuminant is accomplished in three steps. For each image, a chro-
Figure 8: The solid (blue) and dashed (red) polygons illustrate the boundary of the chromaticity gamuts observable under daylight and Tungsten illumination, respectively. The black dots represent colors in an image [Hor99].

The chromaticity diagram is established. The quantized chromaticity values are then captured in a image vector $\mathbf{v}$. This vector is then used to find an estimate of the unknown illuminant, for example, the illuminant which is most correlated ($\mathbf{v}^T \mathbf{M} \rightarrow \mathbf{\rho}$) with the image data (see Figure 10).

5.3 White Balancing

Once we know the illuminant, the actual algorithm to compensate for the illuminant is (very simply) implemented as a color channel multiplication. We call the color image processing step where the response of a given channel is normalized with respect to the illuminant (and the different quantum efficiencies of the real filter and sensor combination) white-balancing:

$$R^o = \frac{R^a}{g_R}, \quad G^o = \frac{G^a}{g_G}, \quad B^o = \frac{B^a}{g_B} \quad (30)$$

Here, the factors $g_R, g_G, g_B$ indicate a multiplication of illuminant response and quantum efficiency gain factors. In general, the factors are normalized such that $g_G = 1$:

$$\begin{bmatrix} R^o \\ G^o \\ B^o \end{bmatrix} = \begin{bmatrix} \frac{g_R}{g_G} & 0 & 0 \\ 0 & \frac{g_G}{g_G} & 0 \\ 0 & 0 & \frac{g_B}{g_G} \end{bmatrix} \begin{bmatrix} R^a \\ G^a \\ B^a \end{bmatrix} \quad (31)$$

21
Figure 9: How to build the correlation matrix, see text for explanation [FHH01].

Figure 10: Solving for the illuminant, see text for explanation [FHH01].
6 Appendix A: Land’s Experiments

For the derivation of his vision model, Land performed various experiments, usually with black-and-white or color Mondrians. We summarize his experiments by describing two. In one experiment [Lan64, LM71], he created a black-and-white Mondrian and illuminated it with a projector. The observers’ sensations varied from white to gray to black in various parts of the display. He then inserted a gray filter wedge into the projection path, whose transmission varied monotonically in one image direction. The observers saw very little change. All the various white, gray and black rectangles looked nearly as they did before (see Figure 11). However, measuring the energy reflected from the different areas revealed that the energy coming from a perceived black rectangle at one part of the Mondrian was equal to the energy of a perceived white rectangle in another part. Repeating the experiment with different colored filters and obtaining the same result, he concluded that the human visual system actually encodes for each channel a “lightness” map whose rank-ordering remains invariant to (slow) spatial changes in illuminant radiant energy as well as illuminant “color.”

Figure 11: Black-and-white Mondrian. Left: uniformly illuminated. Right: spatially slow varying illumination (lower right hand to upper left hand corner). The two patches indicated by the arrows have the same luminance. Both photos are scanned from film supplied in [LM71].

In another experiment [Lan74, Lan77], Land uniformly illuminated a color Mondrian with three projectors emitting narrow-band short, medium, and long wavelength radiation. He measured the color signal of the white patch with a photometer. Considering that he used narrow-band illumination, he could equate one energy measure per spectral band to one cone quantum catch, i.e. one long, one medium, and one short wave energy measure. In a second color Mondrian, he randomly selected a “colored” patch \( S(x, \lambda) \neq 1 \), and adjusted the radiant power of the three projectors so that the color signal equaled the color signal of the white patch in the first Mondrian (see Figure 12). Consequently, the quantum catch of the cones was identical for both stimuli, which should
lead to the same color appearance if the trichromatic theory of color vision is correct (see section Colorimetry LN). However, when observers looked at both Mondrians together, the did not perceive the same color. The white patch in the first Mondrian looked white, while the colored patch in the second Mondrian retained its “color,” i.e. blue remained blue, green remained green, etc. The observers were able to “extract” the original reflectances from the color signal.

Based on his experiments, Land therefore concluded that color appearance is not solely dependent on cone quantum catches, and that some cortical processing is responsible for appearance. Thus, he coined the term retinex from retina and cortex. He proposed a number of algorithms intended to model how the HVS computes these illuminant independent lightness values for each cone channel.

Figure 12: The set-up used in the color Mondrian experiment. The illustration was taken from [Lan77].

One of the early Land retinex algorithm is as follows [BW86, Hur98]. The lightness value for a sensor $k$ is calculated by finding the average ratio between the color response of a position $x_0$ and many surrounding positions. As the non-linear lightness response of the HVS is often modeled with a log function, the ratios can be expressed as subtractions. Let $x_0$ be the location where the lightness $l_k(x_0)$ is to be computed. Let $p_k(x_i), p_k(x_{i+1})$ be color responses of subsequent locations on an arbitrary path $n$ through the scene, with end point response $p_k(x_w^n)$:

$$l_k^n(x_0, x_w^n) = \sum_{i=0}^{w} T [\log p_k(x_{i+1}) - \log p_k(x_i)]$$

(32)

$l_k^n(x_0, x_w^n)$ is the lightness at $x_0$ relative to endpoint $x_w^n$. $T$ represents a thresholding operation that allows to disregard small changes due to small illuminant changes. Recall that he used Mondrians, so the response $p_k(x_i)$ varies sharply with discrete patch boundaries. However, the effective illuminant
varies smoothly across the entire scene, and so induces only small changes in \( \rho_k(x_i) \). Thus, \( T \) ensures that such small ratios are ignored:

\[
T = \begin{cases} 
0 & | \log \rho_k(x_{i+1}) - \log \rho_k(x_i) | < \text{threshold} \\
1 & \text{otherwise}
\end{cases}
\] (33)

Therefore, eq. 32 can be approximated as:

\[
l^n_k(x_0, x_w) \approx \sum_{i=0}^{w} [\log \rho^S_k(x_{i+1}) - \log \rho^S_k(x_i)] = [\log \rho^S_k(x_0) - \log \rho^S_k(x^n_w)]
\] (34)

Spectral normalization is then achieved by averaging the relative lightness \( l^n_k(x_0, x^n_w) \) over all paths \( N \):

\[
l_k(x_0) = \frac{1}{N} \sum_{n=1}^{N} l^n_k(x_0, x^n_w) \approx \log \rho^S_k(x_0) - \frac{1}{N} \sum_{n=1}^{N} \log \rho^S_k(x^n_w).
\] (35)

References


