A Physically Based Numerical Color Calibration Algorithm for Film

Alan Edelman* Frank Wang† Arun Rao‡

Abstract

In this paper, we propose a physically based calibration algorithm for color reproduction using film. We demonstrate that color reproduction is a natural application of the Non-Negative Matrix Factorization (NNMF) algorithm. The color reproduction of film is modelled as a two-stage process: color-recording followed by color-reconstruction. To compute the input that will produce a desired output, the proposed algorithm inversely solves the two stages in sequence. The inversion of the color-reconstruction stage consists of a basis extraction routine and an application of Newton's method. Two basis extraction algorithms, the NNMF and an SVD-based alternative, are compared. For the inversion of the color-recording stage, both a linear interpolation scheme and a physical model of color film development processes are proposed. Our intended application is the matching of the color seen through a film projector with the color that an artist originally intended on a computer screen.

1 Introduction

A common problem in animation film production is the discrepancy between the color seen on a cinema screen and the original intent. To compensate, one can build a calibration model that computes the necessary input for a desired output.

*Department of Mathematics and CSAIL, room 2-343, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. Supported by NSF grant DMS-9971591. (edelman@math.mit.edu; http://math.mit.edu/~edelman)

†Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. (wangq@mit.edu)

‡Pixar Animation Studios, 1200 Park Avenue, Emeryville, CA 94608. (arun@pixar.com)

Calibration of a color system involves building a forward mapping that characterizes system input-output relations, and inversely solving this mapping numerically. In [18], sharma reviews the applications of linear algebra in the modelling of current color imaging systems. Communication of color information between different color systems necessitates the use of a device-independent color space in system characterization [17]. Overviews of color system calibration and characterization can be found in [20, Chapter 1 and Chapter 5]. Discussions of conditions for faithful color reproduction can be found in Horn [10] and Sharma [19].

There are two common approaches for system characterization: interpolation and physical modelling. In an interpolation approach, one first samples the input space of the system with a large number of samples, then chooses a proper interpolation scheme given the inputs and outputs. In a physical modelling approach, one constructs a parametric model based on knowledge of the underlying physical processes and takes sampling data to estimate the parameters. By exploiting extra information, physical models usually have fewer parameters to estimate and can be reused for different systems. For example, Hardeberg [9, Chapter 6] uses the Singular Value Decomposition (SVD) [8], also known as Principle Component Analysis (PCA) [12], to estimate the sensitivity functions of a digital camera, thus generates a physical characterization of the color system of interest.

Gennetten [7] uses a one-step interpolation approach to build a direct forward input-output mapping to calibrate an ink-jet printer. Hardeberg [9] also uses such an approach to characterize a scanner. This paper proposes a novel two-stage calibration model for color reproduction of film and demonstrates the implementations of both an interpolation approach and a physical modelling approach. In a one-step interpolation approach, one assumes no special knowledge of the system by treating it as a black box. However, we already know that the reproduction of color using film consists of two stages: First, the color information is recorded as dye densities on a frame of developed film; then, color is reconstructed by shining a white light through the frame of developed film. We refer to the two stages as a *color-recording stage* and a *color-reconstruction stage* respectively, as shown in Figure 1. Calibration is an inverse problem that requires an estimate of the input-output mapping. By taking advantage of more physical information available, i.e. treating the system as a gray box, a two-stage model narrows down possible functional mappings to a more meaningful set than that in a one-step interpolation approach.

Physical modelling of the color-reconstruction stage requires estimation of basis functions of a lower dimensional vector space from sample data of

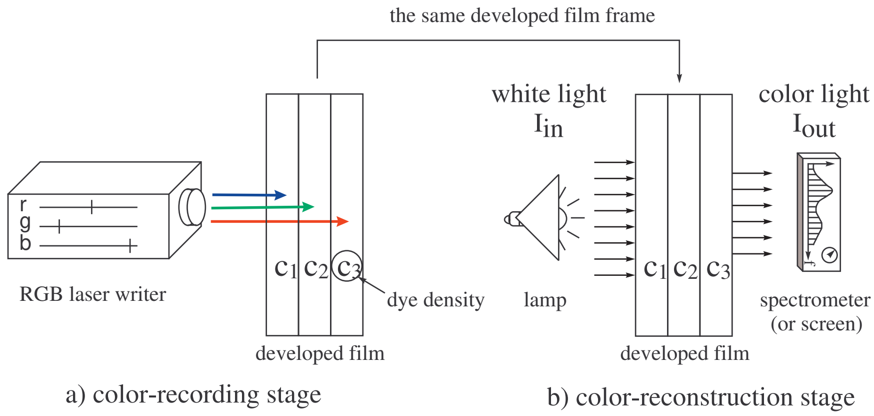


Figure 1: Two stages of the color reproduction by film: a) Color-recording stage: RGB color information of one color sample is recorded as a set of three dye densities (c_1, c_2, c_3). b) Color-reconstruction stage: the color is reconstructed by filtering white light of a projecting lamp.

much higher dimensions. With the current application in color science as an example, we argue generally that the Non-Negative Matrix Factorization (NNMF) [13] can be physically and computationally more appropriate than the SVD factorization in the estimation of underlying basis from a data set. Both algorithms reduce the dimensionality of a problem to the "right" size, but the NNMF goes further. Suppose one is given a data set in large dimensional space and one needs the best approximation in a smaller space, say three dimensional for our purpose. The SVD is well known to compute the best approximation in the sense of the 2-norm or Frobenius norms [8]. In our application, the extracted basis is a set of three vectors corresponding to the three primary color spectra for a subtractive color mixing process, which are analogous to the additive primary color spectra of red, green, and blue [11]. The output spectra of color samples are results of non-negatively adding three non-negative basis functions. In Figure 2(a) and Figure 2(b), we show the orthogonal basis extracted by the SVD and the "primary colors" produced by the NNMF. The plots illustrate that the orthogonality of the SVD is too strong. What is needed instead are the edges of a mathematical cone formed by points that are the projections of spectra onto the three SVD-generated basis vectors. The three edges are the generating basis of points inside or on the cone. They are not necessarily orthogonal.

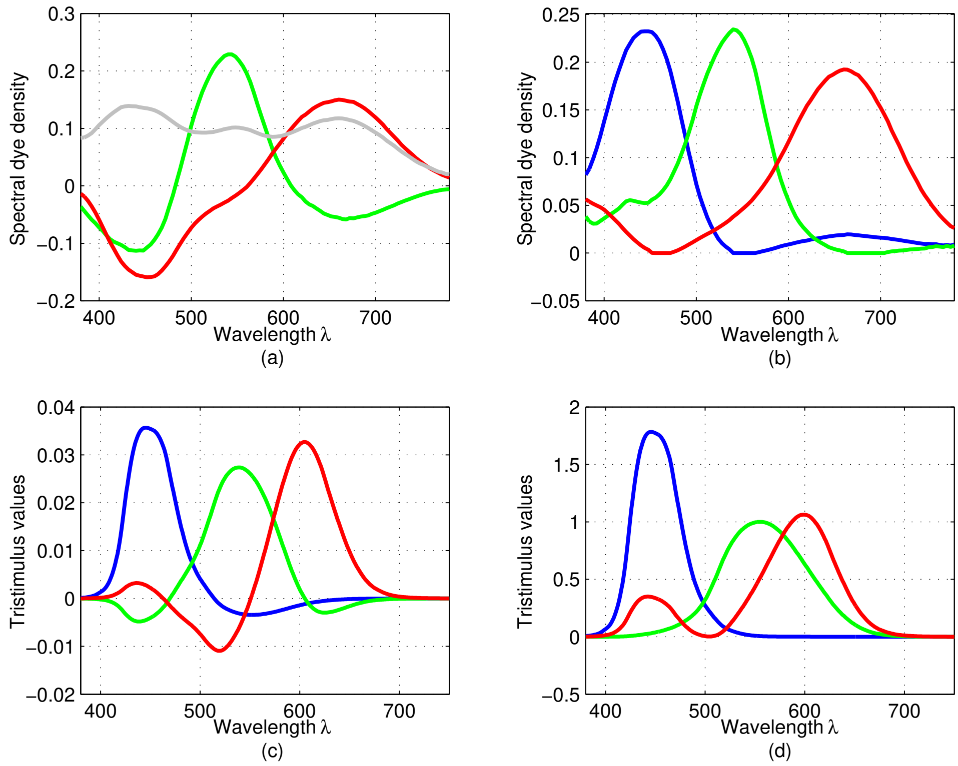


Figure 2: a) Orthogonal basis extracted by SVD b) "Primary color" basis extracted by NNMF c) CIERGB Color Matching Functions d) CIEXYZ Color Matching Functions

Respecting the non-negative constraints, the NNMF extracts such edges of the cone structure by an approximated matrix factorization. The principle of non-negatively adding basis, however, applies more generally. Other applications of the NNMF may be found in [13] [14]. For example, in face recognition, we might say one tries to identify primary facial features to combine non-negatively. In [22], we proposed an alternative non-negative matrix factorization algorithm, which involves the SVD factorization and computational processing in the light of the geometry intuitions aforementioned.

The color-recording stage transforms RGB color information of exposing light into a set of three dye densities of color film in a sequence of chemical reactions [11] [4]. To inversely solve for an input color that produces a desired set of densities, we use both an interpolation approach and a physical modelling approach. In the latter approach, we build a mapping from the input color space to the dye density space using a global least square fit of physically motivated basis functions. Specifically, basing on results from sensitometry research [21] [2], we choose linear basis functions to model the density response to light exposure of a single dye layer, and multi-linear basis functions to model possible inter-layer effects [1]. Result shows that the proposed physically based numerical model captures main features in the data.

The rest of this paper is organized into five sections: In Section 2, we give the background for color calibration and introduce the proposed two-stage calibration model for color film. In Section 3, we explain the physical model of color-reconstruction stage and present the analysis behind the inversion that utilizes Newton's method. We also compare the performance of NNMF with a SVD-based alternative as basis extraction schemes; In Section 4, we describe color film processing and argue the plausibility of a physically based numerical model for the color-recording stage. In this section, we also describe an interpolation model; In Section 5, we list the steps of the proposed two-stage calibration algorithm for color film; Section 6 is the conclusion.

2 Background

2.1 Color Science

Visible light occupies a wavelength band ranging approximately from 380nm to 780nm. Human color perception is facilitated by three types of color sensitive cones in the retina. Each cone is most sensitive to light spectra that correspond to one of the three primary colors, red, green and blue. Figure 3 shows a typical path of light that induces color perception: light $l(\lambda)$ radiated from a light source illuminates a color object; The reflected light $r(\lambda)$ impinges on the three cones in the retina; the light signal causes neurons connected to the cones to fire at a rate that is determined by both the spectrum of the reflected light from the color object and the cone sensitivity functions [23].

Industry standards assume that the *perceptual color space* is three dimensional [3]. Mathematically, it means light spectra are projected onto a three dimensional vector space spanned by the spectral sensitivity functions of the three color cones [23]. In practice, convenient basis functions, such as the CIERGB color matching functions $b_R(\lambda), b_G(\lambda), b_B(\lambda)$ as shown in Figure 2(c) [6], are used instead. With the CIERGB basis available, any

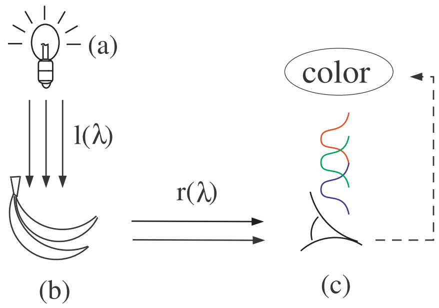


Figure 3: A typical light path that induces human color perception a) a light source b) an opaque color object c) human vision system

spectrum $r(\lambda)$ of color light can be represented by a set of RGB tristimulus values (R, G, B),

$$\begin{aligned} R &= \int r(\lambda) b_R(\lambda) d\lambda \\ G &= \int r(\lambda) b_G(\lambda) d\lambda \\ B &= \int r(\lambda) b_B(\lambda) d\lambda. \end{aligned} \tag{1}$$

However, the projection from a higher dimensional space of spectra to a lower dimensional space of tristimulus values gives rises to "metamers", which refers to sets of different light spectra that have the same tristimulus values [23], and are therefore perceived as one color by human.

Another set of commonly used basis functions is the set of CIEXYZ color matching functions, which is a linear transformation of the set of CIERGB functions [23]. All the entries of CIEXYZ functions (see Figure 2(d)) are chosen to be non-negative to avoid numerical instability due to substraction of two like numbers [20, Chapter 1]. XYZ color coordinates are the tristimulus values corresponding to CIEXYZ basis.

2.2 Calibration of Color Reproduction Systems

A color reproduction system consists of mainly three components: 1) a color sensing unit, such as human color vision system or a camera that senses color light with either photosensitive grains in film or a CCD photon detector array; 2) some color storage media, such as color film or a digital storage device; 3) a color reconstruction unit, such as a projector with slides or a CRT display. Color is captured by the color sensing unit and stored in some intermediate media, then reconstructed by the color reconstruction

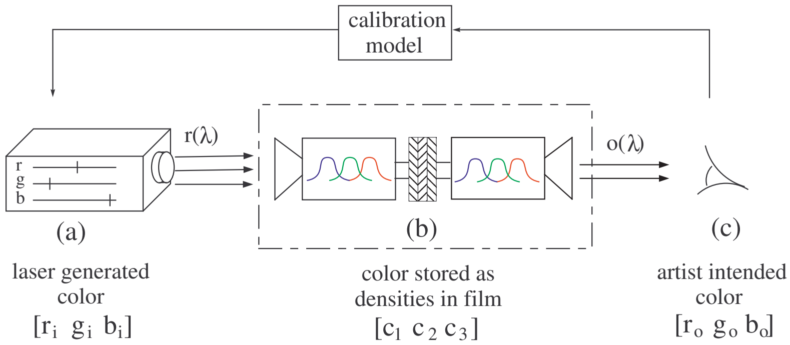


Figure 4: Color calibration of a film system a) RGB laser writer b) film as color reproduction system c) color matching by an artist

unit (with distortions), see Figure 4(b). In movie production, the color reproduction system consists of a movie camera, color film, and a projector. Similar color reproducing systems are key in animation film production, photography, etc.

Figure 4 illustrates the color calibration process of film as a color reproduction system: Artists provide intended color in RGB format (usually obtained using a calibrated computer monitor) that should be produced at the output of the color system; With the proposed calibration model, the input settings of a RGB laser is computed such that with such inputs the film would reproduce the desired color. The goal of the calibration is to solve for a set of three RGB laser settings inversely given a desired output color.

Since most currently available color reproduction systems are built with the assumption that the perceptual color space (or simply color space) is a three dimensional vector space. Ideally we can view the reproduction of color as applying a series of mappings from one vector space to another [10], e.g., from $[r_i, g_i, b_i] \in \Omega_1$ to $[c_1, c_2, c_3] \in \Omega_2$, then to $[r_o, g_o, b_o] \in \Omega_3$, where Ω_1 is the RGB color space of the input color, Ω_2 is the vector space of dye density vectors of the film, Ω_3 is the RGB color space of the output color. However, these mappings are not always linear. In this paper, we assume the color space of the output device contains that of the input device, thus avoid the complication of "out-of-gamut" color.

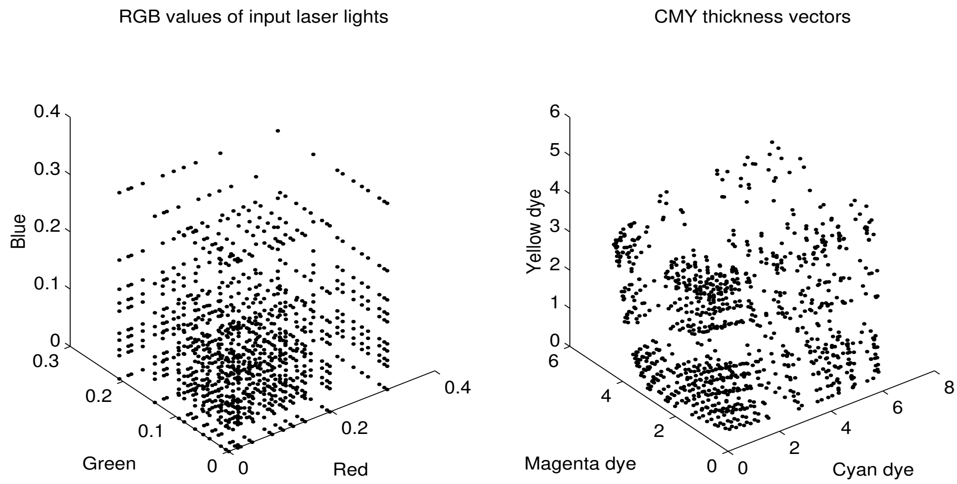


Figure 5: (a) 1000 RGB values sampled from a RGB color cube (b) 1000 dye density vectors of color negative film produced by the 1000 RGB colors

2.3 A Two-Stage Calibration Model for Color Film

In treating color film as a color reproduction system, we take as outputs: the transmitted light spectra resulted from shining white light through a frame of developed film, and as inputs: the set of (R,G,B) values of the laser light used to expose the film (see Figure 1). The objective of our calibration model is to compute a set of three input settings of RGB laser that will produce a desired output color specified in XYZ coordinates (which is equivalent to RGB coordinates by a linear transformation).

2.3.1 The data set

The data set available for calibration is collected in the following steps: 1000 colors are sampled from a RGB color cube with 10 discrete sampling points along each primary color axis; one frame of film is exposed to one *solid color* out of the 1000 color samples and then developed; Then, we shine a white light through each frame of developed film and measure the spectrum of the output light using a spectrometer with a 101-point discretization of the wavelength region of visible light. Figure 5(a) displays the 1000 inputs as the grid points in a RGB color cube.

The available data set contains the inputs–a $10 \times 10 \times 10$ sampling grid points of a regular RGB color cube, the outputs–1000 discretely sampled spectra of light outputs, plus one more spectrum of the a white light used as the light source to produce the outputs. Two data sets have been collected, one for a type of color positive film and the other for a type of color negative film respectively.

2.3.2 The calibration model

Given inputs and outputs of a system, Gennetten [7] builds a direct interpolation model by treating the film system as one black box. Hardeberg [9] also employs such an interpolation approach in the characterization of a scanner system. However, we believe more refined physical models of the system of interest would enable more accurate calibrations.

In this paper, we suggest a two-stage model for the film color reproduction process shown in Figure 1. In the first stage, a color light generated by an RGB laser exposes a frame of color film. After the exposure and development of the film, the color information of the input light is recorded uniquely as a *density vector* (c_1, c_2, c_3) where the components are the densities of primary-color-absorbing dyes, i.e. the cyan, yellow and magenta dyes [4]. In the second stage, a white light is shined through the developed film. As a component of the subtractive basis of color mixing, each primary-color-absorbing dye in a separate dye layer mainly absorbs energy of the white light at the wavelength range of its corresponding primary color. (Specifically, cyan dye mainly absorbs red light, similarly, magenta dye absorbs green light and yellow dye absorbs blue light.) The resulted color will be a reconstruction of the original color with some distortion.

With respect to color processing, we refer to the first stage as the *color-recording stage*, where the RGB values of a color input are recorded as a density vector (c_1, c_2, c_3), and the second stage as the *color-reconstruction stage*, where colors are reconstructed by shining white light through the three dye layers of the developed frames of film. The three dye layers can be thought of as a stack of three color filters (see Figure 6(b)).

The use of a set of density values (c_1, c_2, c_3) as an intermediate state decouples the two stages of color processing and allows us to inversely solve the two stages separately. In the proposed algorithm, we first inversely solve the color-reconstruction stage, i.e., solve for a set of density values (c_1, c_2, c_3) given the XYZ values of a desired output color. To perform this task, we combine Newton's method with either the well known SVD factorization or the less known NNMF factorization. Then, we invert the color-recording stage, by either an interpolation scheme or a physical modelling approach, to compute a set of RGB laser input values from the set of density values (c_1, c_2, c_3) obtained from the inversion of color-reconstruction stage.

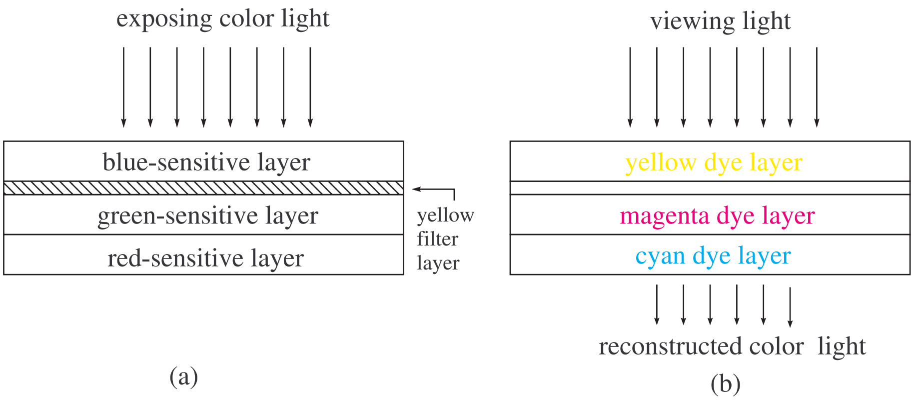


Figure 6: Color reproduction using "tri-pack" color film a) unexposed film frame b) developed film frame

3 Inverting the Color Reconstruction Step

To solve the inverse problem of the color reconstruction step, we follow three steps: 1) use the SVD to verify a hypothesis that each spectrum in the data set is a linear superposition of three basis functions; 2) extract the basis functions using either the NNMF or an alterative non-negative matrix factorization algorithm that consists of the SVD and some geometrical processing; 3) use Newton's method to compute (c_1, c_2, c_3) for a given XYZ values.

3.1 The Model of Light Absorbing Dye Layers

Beer's law [23] is valid for homogeneous isotropic absorption filters,

$$\frac{I_{\text{out}}(\lambda)}{I_{\text{in}}(\lambda)} = \phi(\lambda)^c \tag{2}$$

where λ is the wavelength, $\phi(\lambda)$ is an absorption constant of the light absorbing material and c is its concentration (or density). The three layers of primary-color-absorbing dyes, i.e., cyan, yellow and magenta dyes in a frame of developed color film can be modelled as a stack of three such filters [5] (see Figure 6(b)). The three dyes attenuate the input white light according to their corresponding absorption functions $\phi_y(\lambda), \phi_m(\lambda), \phi_c(\lambda)$ (see Figure 6). (Examples of the logarithm of these functions extracted numerically

degree n=	1	2	3	4	5
$\frac{\sum_{i=1}^{n}\sigma_i^2}{\sum_{i=1}^{101}\sigma_i^2}$	0.83602	0.95044	0.99797	0.99818	0.99834

Table 1: Ratios of sums of singular value squares

from the available data set is shown in Figure 7.) Then, the output light spectrum is determined by the densities of the dye layers (c_1, c_2, c_3), the absorption functions of the dyes, $\phi_y(\lambda), \phi_m(\lambda), \phi_c(\lambda)$ and the spectrum of the projecting lamp $I_{in}(\lambda)$ as in the following equation,

$$I_{out}(\lambda) = I_{in}(\lambda)\phi_y(\lambda)^{c_1}\phi_m(\lambda)^{c_2}\phi_c(\lambda)^{c_3}. \tag{3}$$

Rewrite this equation and define the *log-spectra-ratio* (also known as the spectral-dye-density function) $d(\lambda)$ as,

$$d(\lambda) \equiv -\log(\frac{I_{out}(\lambda)}{I_{in}(\lambda)}) = c_1 b_1(\lambda) + c_2 b_2(\lambda) + c_3 b_3(\lambda), \tag{4}$$

where $b_1(\lambda) = -\log(\phi_b(\lambda))$, $b_2(\lambda) = -\log(\phi_g(\lambda))$, $b_3(\lambda) = -\log(\phi_r(\lambda))$ are defined as the negative of logarithm of absorption functions. Thus, the log-spectra-ratios are generated by *non-negative* linear combination of a set of three *log-absorption-basis functions*, $(b_1(\lambda), b_2(\lambda), b_3(\lambda))$, which are also called the *generators*. It is important to note here that the entries of a density vector (c_1, c_2, c_3) are always non-negative numbers.

3.2 Three Generators

If the data set is indeed generated using the aforementioned typical tri-pack color film, we would expect that the spectra data has a three-generator structure. In other words, each of the 1000 log-spectra-ratios is a linear combination of the same set of three generators. In order to verify, we first collect the 1000 log-spectra-ratios $\log(\frac{I_{out}}{I_{in}})$ into a matrix $V \in \mathbb{R}^{101\times1000}$, where 101 is the number of sampling wavelengths of the spectrometer used and 1000 is the number of samples. Then, we apply the SVD to this matrix. From Table 1, we see that the sum of the largest three singular value squares captures 99.797% of the total sum, and the inclusion of the fourth largest singular value square only improves the percentage by 0.19%. This suggests strongly that a three-generator structure is embedded in the spectra data.

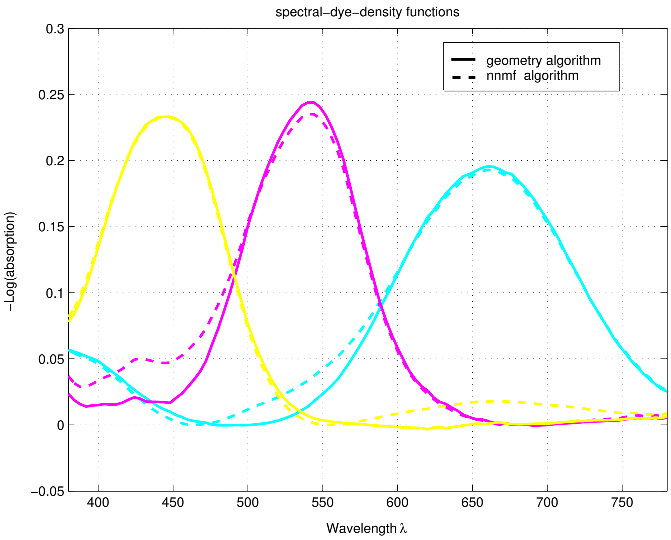


Figure 7: The negative of log-absorption-basis functions (also know as the spectral dye density functions) of cyan, magenta and yellow dye layers extracted by NNMF of 10000 iterations (dash lines) and by a geometry approach with SVD (solid lines).

3.3 Generators Extraction

Finding the three generators underlying the 1000 101-dimensional spectra, which are listed as columns of the matrix V, can be formulated as a matrix factorization of the form:

$$V \approx WH, \tag{5}$$

where $W \in \mathbb{R}^{101\times 3}$, $H \in \mathbb{R}^{3\times 1000}$ having only non-negative elements. The three columns of W are the generators, i.e., the negative of log-absorption-basis functions of the three dye layers. H is 1000 sets of three densities $(c_1, c_2, c_3)^T$. The non-negativity is a physical constraint, which is due to the nature of the color reproduction by color film. While respecting the non-negativity, the NNMF iteratively updates W and H by a gradient descent like update rule and optimizes a cost function that measures how well WH approximates V. The NNMF has guaranteed convergence [13].

As an alternative, we suggest a generator extraction algorithm that is based on the SVD factorization and geometric intuitions mentioned briefly in the introduction section: In 3D Euclidean geometry, adding three independent generators non-negatively will generate new points inside a convex

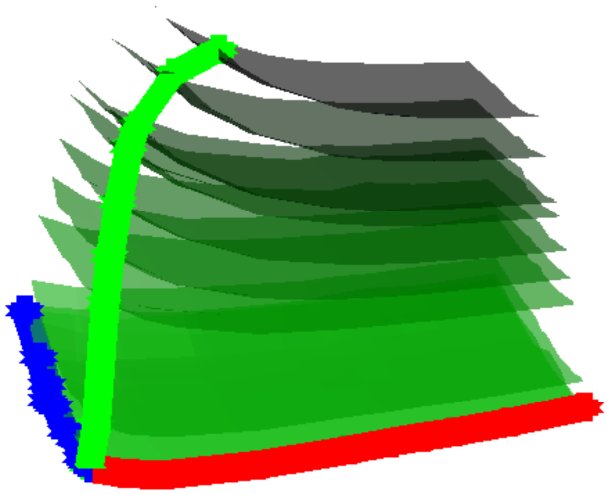


Figure 8: The 1000 density vectors form a parallelepiped structure. The surfaces are densities of constant green exposure.

cone and the generators are the three bounding edges through the origin. Generators extraction means to uncover these three bounding edges from the density vectors. To extract the generators geometrically, we followed a series of simple steps: Take SVD of the log-spectra-ratios; Project 1000 spectra onto the assumed three dimensional space spanned by the first three orthogonal u vectors (i.e., the first three columns of matrix U resulted from the SVD factorization). The resulted three dimensional density vectors have a structure of a parallelepiped. See Figure 8, since each density is non-negative and has an upper bound; Find the convex hull using existing MATLAB tool, *convexhulln* [15]; Compute the three normal vectors of the three sides of the parallelepiped that pass through the origin. Here, we use the normal vector of the facet having the largest area on that side of the convex hull; The three generator vectors are cross products of pairs of the three normal vectors. The detail of the alternative algorithm and its performance comparison with the NNMF are provided in [22].

Figure 7 compares the generators, or the log-absorption-basis functions of the three dye layers obtained by the two different non-negative factorization schemes. By considerations of color film design, the solid curves obtained by the geometry approach would be a better set of absorption functions for film, because there is less undesired cross frequency band absorption. Cross frequency band absorption is referred to as *cross-layer absorption* at later sections.

3.4 From XYZ Color Coordinates to Densities By Newton's Method

Once having the three basis functions (generators), we proceed to invert the color-reconstruction stage using Newton's method. The task is to obtain a density vector given a desired color specified in XYZ color coordinates.

The formula that turns densities and the spectrum of a projecting lamp into a set of XYZ tristimulus values (x, y, z) is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \int \begin{pmatrix} x(\lambda) \\ y(\lambda) \\ z(\lambda) \end{pmatrix} I_0(\lambda) e^{-(c_1 b_1(\lambda)+c_2 b_2(\lambda)+c_3 b_3(\lambda))} d\lambda. \tag{6}$$

where $x(\lambda), y(\lambda)$ and $z(\lambda)$ are the three CIEXYZ color matching functions as shown in Figure 2(d). The inversion of the above equation requires the solution of a set of nonlinear equations. This equation may be solved by Newton's method. Some experience shows that Newton's method works very well no matter where the starting point is.

The linearization requires the Jacobian

$$J = \begin{pmatrix} \frac{\partial}{\partial c_1} \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \frac{\partial}{\partial c_2} \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \frac{\partial}{\partial c_3} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{pmatrix}.$$

Substituting in Equation 6, we see that the ith column of the Jacobian is

$$\frac{\partial}{\partial c_i} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = -\int \begin{pmatrix} x(\lambda) \\ y(\lambda) \\ z(\lambda) \end{pmatrix} I_0(\lambda) b_i(\lambda) e^{-(c_1 b_1(\lambda)+c_2 b_2(\lambda)+c_3 b_3(\lambda))} d\lambda.$$

In other words an extra factor of -$b_i(\lambda)$ is introduced into the integrand. The same code that calculates XYZ values, may be used with this slight modification to compute the Jacobian.

Suppose $c_1^{(n)}, c_2^{(n)}$ and $c_3^{(n)}$ are the nth iterates with corresponding $x^{(n)}, y^{(n)}$ and $z^{(n)}$. Newton's method computes,

$$\begin{pmatrix} c_1^{(n+1)} \\ c_2^{(n+1)} \\ c_3^{(n+1)} \end{pmatrix} = \begin{pmatrix} c_1^{(n)} \\ c_2^{(n)} \\ c_3^{(n)} \end{pmatrix} - J^{-1} \left\{ \begin{pmatrix} x^{(n)} \\ y^{(n)} \\ z^{(n)} \end{pmatrix} - \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\},$$

where the last vector is the desired XYZ values.

It is worthwhile to note that the Jacobian is quite well conditioned. This means that the map from density space to color space is everywhere well determined. Using the data, we computed all 1000 condition numbers. There were only about 100 of them that are bigger than 200 and the largest was 6000. We recall from numerical analysis that the base-10 logarithm of condition number is the number of digits that we expect to lose in a computer. A condition number that is larger than 10^{10} is worrisome. The observed condition numbers are considered to be very good and no fear needs to occur when solving the linear equation involving J in Newton's method. Indeed we believe the negative densities are fine too. (Negative densities are simply interpreted as amplification of the projection signal rather than absorption.) Condition numbers only seem to get large when one of the density components is large in magnitude, larger than physically realistic.

4 Inverting the Color Recording Step

For inversely solving the color-recording stage, i.e., obtaining the correct input RGB values for a given density vector, we propose both an interpolation approach and a physical modelling approach.

4.1 An Interpolation Scheme

In an interpolation approach, one treats the system as an unknown black box with only the inputs and the outputs visible to an observer. Then, interpolation functions are chosen to fit the output data with input data over the domain of interest. With a linear three-dimensional interpolation function $\vec{f}$, we can inversely solve for the set of RGB values for a given set of density values $\vec{c} = (c_1, c_2, c_3)$, where

$$\vec{c} = \vec{f}(\begin{bmatrix} R \\ G \\ B \end{bmatrix}).$$

Our interpolation algorithm consists mainly the following three steps: For a given $\vec{c}$, we find, in some optimal way, a set of four neighboring points among all 1000 points shown in Figure 5(b), which forms a tetrahedron encloses $\vec{c}$ but not any other output points; Then, we choose an interpolation function from a list of well known ones such as linear, Hermite, spline, and etc.; Lastly, we solve for the RGB values inversely.

We found the interpolation tools in MATLAB convenient to use. In particular, we used the 3D Delaunay tesselation routine of MATLAB [15] to find the enclosing tetrahedron for each density vector. For the interpolation, we choose to use linear interpolation for the convenience of implementation.

4.2 A Physical Modelling Approach

4.2.1 Color film processing

The shortcomings of an interpolation approach is that the resulting interpolation functions can not be generalized to predict results in similar situations and requires a large number of fitting parameters. If there are only a few underlying physical processes at work in producing the data, then with enough knowledge of the underlying processes, one can possibly device a physical model that will explain the data with only a few parameters and capture the general physics of alike systems. From a data processing perspective, this results in efficient data compression.

In modelling the functional relationships between exposing light and dye densities of developed film, an understanding of the film development process is necessary. In the following paragraphs, we give a brief introduction of the film development process. For a detailed description, we refer readers to Hunt [11].

Figure 6(a) shows a typical "tri-pack" structure of color film. Each layer contains silver halide grains that are mainly sensitive to one corresponding primary color. Figure 9 illustrates a four-step development process for color negative film in a microscopic scale. New film contains unexposed silver halide grains (white squares) and coupler particles (circles). When exposed to light, silver ions in some unexposed silver halide grains will acquire electrons and form silver atom aggregate on the surface of the grain. Such grains become exposed silver halide grains (square with a dot), which are also called developable grains. Exposed grains in one emulsion layer form a latent image whose density is dependent on the exposure of a corresponding primary light. (Exposure is defined as the product of illuminance [16, Pages 540] of exposing light and exposure time.) After the latent image has formed, the film is treated with a color developer. Exposed silver halide grain and nearby coupler particles will react with the developer and result in completely developed grains of silver (black squares) and color dye particles (circled letters). A bleach step that follows removes silver grains and destroy the yellow filter layer. Finally, a fix step dissolves all remaining unexposed silver halide grains and renders a color negative film with color dye images

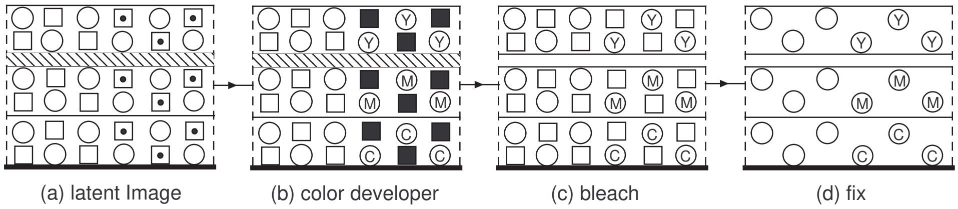


Figure 9: Sequence of chemical operations in film development (Legends: □ undeveloped silver halide grain ⊡ developable silver halide grain ■ developed silver grain ○ coupler particle Ⓨ yellow dye particle Ⓜ magenta dye particle © cyan dye particle)

in three layers as seen in Figure 6(b). By the design of color negative film, the more intense is the exposing light, the higher is the resulted dye densities, in turn, the dimmer the negative is when viewed directly against a light source. Therefore, we expect the dye densities *increase monotonically* with their primary light exposure.

It is easy to confuse the light absorption by light sensitive emulsion layers with the absorption by color dye layers in the other stage, i.e., the color-reconstruction stage. Even though at the end of the development process of color film, a light sensitive layer turns into its corresponding dye layer (ref. Figure 9), but during the development process, the color light sensitive layers are transparent. The absorption of color light during exposure is by colorless sensitized frequency selective silver halide grains rather than dye particles. Ideally, a color sensitive layer would only respond to its primary color light. However, *cross-layer-exposure* effect always exists, just as does the cross-layer absorption in color reconstruction (ref. the end of section 3.3). Cross-layer-exposure effect is especially prominent in the cases of green-sensitive and red-sensitive layers absorbing blue light. To alleviate the problem, a yellow filter layer is inserted immediately after the blue-sensitive layer at the top to absorb excess blue light, see Figure 6(a).

The quantitative study of the relationship between the density of developed silver grains and light exposure is the subject of a field called *sensitometry*. Experiments of sensitometry show that the optical density of silver image D and the logarithms of exposure $\log(E)$ follows an H-D curve first used by Hurter and Driffield [21] [2], see Figure 10. A H-D curve has three regions, a “toe”, a linear region and a “shoulder”. Most types of color film are designed to operate in the linear region. Thus, under normal conditions, we would expect a *linear* relationship between density and log-exposure.

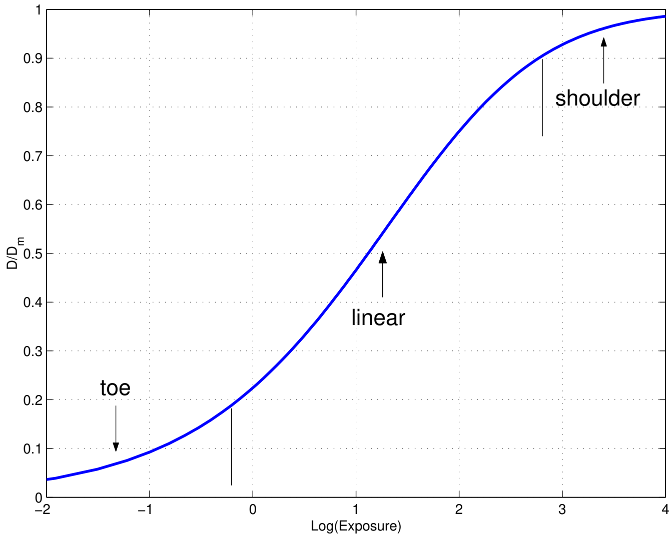


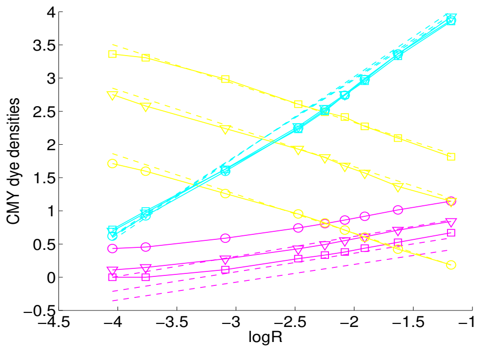
Figure 10: H-D characteristic curve of a single emulsion layer

4.2.2 A physically based numerical model

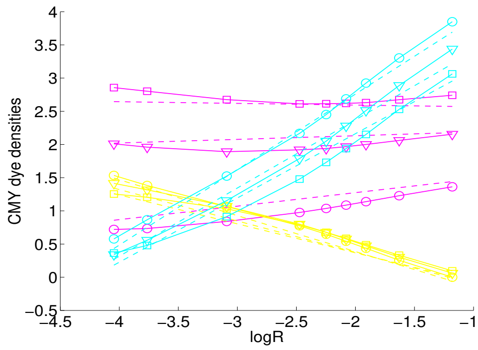
In Figure 11, we present the NNMF generated densities (i.e., the H in Equation 5) of color negative film in three sets of two plots. In the two plots of the first row, densities of cyan, magenta and yellow dyes are plotted over $\log R$ in solid color curves. The first plot shows the densities over $\log R$ at three constant B values; The increasing order of the three B values is indicated by circle, triangle and square; In this plot, the G value is held at its minimum value. The second plot of the first row is the same except that G takes on three constant values and B is held at its minimum. The other two sets are the same except the roles of the R, G, and B are permuted.

The shape of the density curves of a complementary dye responding to its primary light component does resemble the linear region of a H-D characteristic curve. However, the existence of inter-layer effects is also apparent: if such effects were absent, then as the intensity of a primary color light increases along the abscissa, it should be that only its corresponding color light sensitive layer would respond with deposited dye density increases and the dye density curves of the other two layers would have stayed flat. The pairings of a dye and the primary light that it mainly responds to are as follows: yellow dye mainly responds to blue light, magenta dye mainly responds to green light, and cyan dye mainly responds to red light.

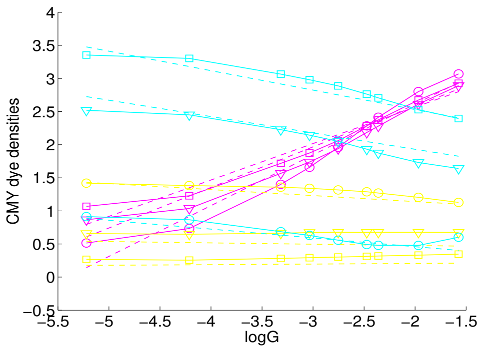
By inter-layer effect, also called inter-image effect [1], we refer to the



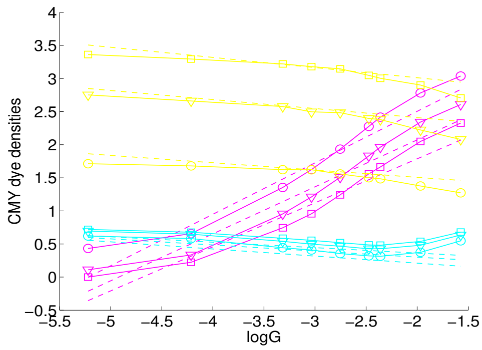
(a) G=0.005 ○:B=0.041 ∇:B=0.067 □:B=0.157
—: cyan —: magenta —: yellow
—: experiment – –: model



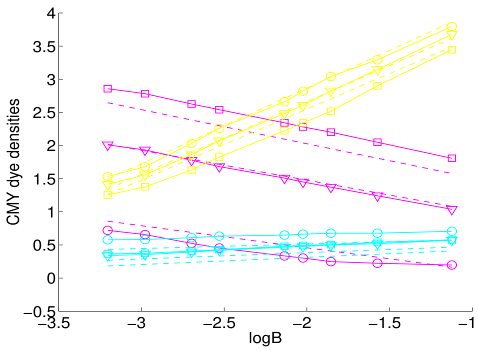
(b) B=0.041 ○:G=0.005 ∇:G=0.036 □:G=0.094
—: cyan —: magenta —: yellow
—: experiment – –: model



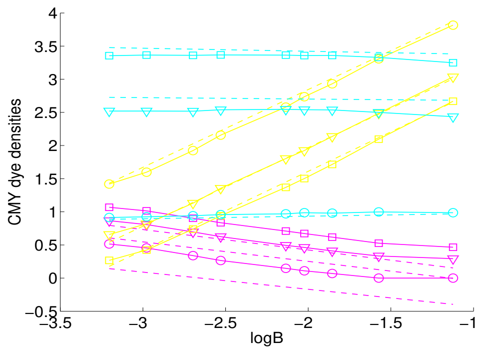
(c) B=0.041 ○:R=0.018 ∇:R=0.046 □:R=0.148
—: cyan —: magenta —: yellow
—: experiment – –: model



(d) R=0.018 ○:B=0.041 ∇:B=0.067 □:B=0.157
—: cyan —: magenta —: yellow
—: experiment – –: model



(e) R=0.018 ○:G=0.005 ∇:G=0.036 □:G=0.094
—: cyan —: magenta —: yellow
—: experiment – –: model



(f) G=0.005 ○:R=0.018 ∇:R=0.046 □:R=0.148
—: cyan —: magenta —: yellow
—: experiment – –: model

Figure 11: Experimental data and a model of the set of three tri-variate dye density functions of negative film presented in six 2D plots.

phenomena that a color light sensitive layer also responds to non-primary light components. Possible causes include cross-layer exposure, cross-layer inhibition, and etc.. Here, the most prominent inter-layer effect manifests as that most dye densities decrease monotonically as the intensity of a non-primary light component increases, e.g., yellow dye density in Figure 11(a), cyan in Figure 11(c), and magenta in Figure 11(e). Development-inhibitor-releasing (DIR) [11] couplers that are designed to retard the development of neighboring dye layers to compensate for the cross-layer absorption in color reconstruction (ref. the end of section 3.3) could be responsible for all the negative linear response of dye densities to the log of non-primary light exposure.

The unique pattern of the magenta dye in Figure 11(b) calls for special attention: the decreasing of its density as non-primary red light increases only happens at high constant $\log G$ values. Its density increases with $\log R$ at low constant $\log G$ values. The slopes of the curves of magenta dye density over $\log R$ decrease as $\log G$ increases. This phenomena also displays itself in Figure 11(c) where the magenta dye density curves of constant $\log R$ values crosses each other as $\log G$ increases. A possible explanation for this phenomena is that the increasing red exposure has two simultaneous effects on magenta layer: one effect is the cross-layer-exposure effect, i.e. green-sensitive layer also develops against red light (evident in Figure 11(a)), however, much less than green light. Such effect is strong when only red light presents and becomes less effective as the development of magenta dye to the increasing primary exposure $\log G$ dominates; The other effect is the DIR cross-layer-inhibition effect of the developing cyan layer on the magenta layer. This effect is first dominated by the previous effect and only becomes important when the first effect diminishes at high green exposure. Its existence is indicated by the negative slopes of magenta curves of high constant $\log G$ values in Figure 11(b). The diminishing of cross-layer-exposure effect motivates a bi-linear term of $\log R \log G$ in a model of the functional relationship between magenta dye density and light exposure.

By earlier analysis, we hypothesize that there are only *multi-linear* terms involved in the film development process, and each dye density function is a linear superposition of a set of multi-linear basis,

$$(1, \log R, \log G, \log B, \log R \log G, \log G \log B, \log B \log R, \log R \log G \log B).$$

Denote the density vector $\vec{c} = (C, M, Y)^T$, where $C = C(R, G, B)$, $M = M(R, G, B)$, $Y = Y(R, G, B)$ are the tri-variate density functions of cyan, magenta and yellow dyes of the negative color film used. A least-squares global fit generates the following model:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 3.4811 & 2.0233 & 2.3304 \\ 0.7549 & -0.0088 & -0.4854 \\ -0.3101 & 0.3522 & -0.1433 \\ 0.0632 & -0.3981 & 0.8333 \\ -0.0635 & -0.0460 & 0.0145 \\ 0.0333 & -0.0099 & -0.0712 \\ -0.0202 & 0.0676 & -0.0437 \\ 0.0038 & 0.0174 & -0.0100 \end{bmatrix}^T \begin{bmatrix} 1 \\ \log R \\ \log G \\ \log B \\ \log R \log G \\ \log G \log B \\ \log B \log R \\ \log R \log G \log B \end{bmatrix}. \quad (7)$$

The model are plotted in dashed color lines in Figure 11. The results show that our physically based modelling hypotheses are plausible.

5 The Calibration Algorithm

- Expose 1000 frames of color film to 1000 solid colors sampled from a RGB color cube. Measure the spectra of the output color light produced by shining white light through the developed 1000 frames of film.
- Invert the color-reconstruction stage:
  1. From the spectra data V, extract a set of three basis vectors W, i.e. the log-absorption functions of three dye layers using either the NNMF or its SVD-based alternative.
  2. From a given intended color specified in XYZ color coordinates, solve Equation 6 for a dye density vector (c_1, c_2, c_3) using Newton's method.
- Invert the color-recording stage:
  1. Obtain the 1000 density vectors H of the 1000 spectra as the other factor from the NNMF factorization.
  2. Build a functional mapping between input RGB color space and the dye density vector space via either interpolation or physical modelling. The model is obtained by a global least squares fit over a multi-linear basis.
  3. Solve this function inversely for a set of calibrated RGB inputs that will give the density vector obtained from the previous stage.

With the set of calibrated RGB values as input, the color film as a color reproduction system will produce the intended color at the output.

6 Conclusion

In this paper, we developed a two-stage calibration model for color film. We demonstrated that NNMF-like algorithms are readily applicable to the characterization of color reproduction systems due to the existence of intrinsic non-negativity. We compared a SVD-based non-negative matrix factorization algorithm with the NNMF algorithm on the task of log-absorption-basis function extraction. We showed that the NNMF-like basis extraction combined with Newton's method can be used to solve an inverse problem. We demonstrated the construction of an interpolation model and a physically based numerical model as the mapping from input color space to dye density space. The calibration procedure and the techniques used in this paper can be generalized to apply to similar color systems.

7 Acknowledgement

The authors would like to thank Dr. Victor Ostromoukhov and Dr. Frédo Durand for valuable input.

References

[1] G. G. Attridge, R. E. Jacobson, P. Newman, and J. Roychoudhury. Observations concerning inter-image effects in colour printing paper. *The Journal of Photographic Science*, 31:197–205, 1983.

[2] G. G. Attridge. The characteristic curve. *The Journal of Photographic Science*, 39:55–62, 1991.

[3] Steven M. Boker. The representation of color metrics and mappings in perceptual color space. `http://www.nd.edu/~sboker/ColorVision2/ColorVision2.html`, 1988.

[4] Eastman Kodak Company. *Exploring The Color Image*, volume H-188 of *KODAK Publication*. Rochester, New York.

[5] Eastman Kodak Company. *Student Filmmaker's Handbook*, volume H-19a of *KODAK Publication*. Rochester, New York.

[6] CVRL. CIE color matching functions data sets. `http://cvrl.ucl.ac.uk`, 1988.

[7] K. Douglas Gennetten. RGB to CMYK conversion using 3-d barycentric interpolation. In *SPIE conference on Device-Independent Color Imaging and Imaging Systems Integration*, number 1909, pages 116–126, San Jose, CA, February 1993.

[8] Gene H. Golub and Charles F. Van Loan. *Matrix Computation*. John Hopkins, Baltimore, Maryland, third edition, 1996.

[9] Jon Y. Hardeberg. *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches*. Ph.D. dissertation, Ecole Nationale Superieure des Telecommunications (ENST), Department of Signal and Image Processing, 2001.

[10] Berthold K. P. Horn. Exact reproduction of colored images. *Computer Vision, Graphics, and Image Processing*, 26:135–167, 1984.

[11] R. W. G. Hunt. *The Reproduction of Colour in Photography, Printer and Television*. Fountain Press, Tolworth, England, sixth edition, 1987.

[12] I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, New York, New York, 1986.

[13] Daniel D. Lee and H. Sebastian Seung. Algorithms for non-negative matrix factorization. In *NIPS*, pages 556–562, Denver, CO, 2000.

[14] S. Li, X. Hou, and H. Zhang. Learning spatially localized, parts-based representation. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) Volume 1*, pages 207–212, 2001.

[15] *MATLAB Manual*. Natick, MA.

[16] C. Noel Proudfoot, editor. *Handbook of Photographic Science and Engineering*. The Society for Imaging Science and Technology, Springfield, Virginia, 1997.

[17] William F. Schreiber. Today's challenge in device-independent color. In *Device-Independent Color Imaging*, pages 2–15, San Jose, CA, 1994.

[18] Gaurav Sharma and H. Joel Trussell. Digital color imaging. *IEEE Transactions on Image Processing*, 6(7):901–932, 1997.

[19] Gaurav Sharma, Michael J. Vrhel, and H. Joel Trussell. Color imaging for multimedia. *Proceedings of the IEEE*, 86(6):1088–1108, 1998.

[20] Gaurav Sharma, editor. *Digital Color Imaging Handbook*. CRC Press, New York, New York, 2003.

[21] Ludwik Silberstein and A. P. H. Trivelli. Relations between the sensitometric and the size-frequency characteristics of photographic emulsions. *Journal of the Optical Society of America*, 28:441–459, 1938.

[22] Frank Wang and Alan Edelman. A geometry-based alternative nonnegative matrix factorization. In preparation.

[23] Gunter Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley & Sons, New York City, New York, second edition, 1974.