BLIND PRIMARY COLORANT SPECTRAL SEPARATION COMBINING ICA AND POCS NON-NEGATIVE MATRIX FACTORIZATION

Chunghui Kuo

Nexpress Solutions LLC.
1447 St. Paul Street, Rochester, NY 14653, USA

ABSTRACT

Traditional printing technologies use metamerism to match color between original objects and reproduced images under certain lighting constraints. These techniques become unsatisfactory for applications demanding high color accuracy where lighting conditions are subject to change, for example, artwork reproduction. Therefore, it is desirable to reproduce the image using colorants whose spectrum is closest to the original print. To estimate the original object primary colorant spectrum, we propose a blind spectral separation algorithm by combining Independent Component Analysis (*ICA*) and Non-Negative Matrix Factorization (*NNMF*). Our experimental results show a satisfactory match between the original and the estimated primaries' spectrums.

1. INTRODUCTION

Conventional printing technologies only achieve metameric color match between original objects and reproduced images under certain lighting constraints. However, the results become unsatisfactory for applications demanding high color accuracy where lighting conditions are subject to change, for example, artwork reproduction, product catalog and museum image archiving. The perceived image is determined by the reflected light of objects reaching retina, and it is governed by the following integral equation:

$$S_r(\lambda) = \int R(\lambda) S_i(\lambda) d\lambda \tag{1}$$

where $S_i(\lambda)$ and $S_r(\lambda)$ represent the incident and reflected light spectrum, and $R(\lambda)$ is the object reflectance curve. The human retina composes of three types of light receptors behaving as bandpass filters which map the perceived light from an infinite dimensional space to a three dimensional subspace. Hence, it is possible for two objects with different light reflectance curves to be perceived as possessing the same color under certain $\hat{S}_i(\lambda)$, noted as *metameric pair*, because of the above many-to-one mapping. In this case, it is desirable to reproduce the image using colorants whose spectrum is closest to the original object. Therefore, first estimating the primary colorants' reflectance spectrum of the original object becomes essential. In this paper, we address this spectral separation problem by assuming the original objects to be color prints such as art paintings, while the same spectral separation problems also exist in other situations such as remote sensing and hyperspectral data analysis. [1,2]

The spectral separation problem can be divided into two sections: transforming the signal to a linear spectral-mixing space and designing a blind signal separation algorithm. A linear spectral-mixing space allows the following separation problem to be more tractable because it can be formulated as an under-determined algebraic linear system. The spectral-mixing characteristics on prints can be modelled using multiple light absorption and reflection by the primary colorants and various interfaces, which is highly nonlinear. Various theories have been proposed to describe this spectral mixing characteristics, such as Yule-Nielsen model and Kubelka-Munk theory [1,3]. We will compare these two models in the following section to verify its linearity after applying the corresponding transformation.

We propose a spectral separation algorithm that combines Independent Component Analysis *ICA* and NonNegative Matrix Factorization *NNMF*. It has been argued that natural constraints of spectral components and nonnegativity of the primary colorants amount make *NNMF* a viable approach [4]. Assuming that the reflectance spectrum of the primary colorants are independent of one another, we first calculate a set of independent component signals as the initial condition for the ensuing *NNMF* algorithm. The non-negative constraint can be shown to be a intersection of two convex sets. As a result, we propose a *NNMF* algorithm based on the principle of projection onto convex sets, *POCS* [5]. Finally, we verify our algorithms by comparing the recovered primary reflectance spectrum with the measured primary reflectance spectrum. Our experiment results demonstrate satisfactory match.

2. SUBTRACTIVE SPECTRAL MIXING MODEL

A print is perceived when light reflected from that print reaches our eyes. Thus, the portion of spectrum absorbed by the colorant will be missing from the perceived light, i.e. subtractive spectral mixture. We set linear spectral mixing spaces using two models, the Yule-Nielsen model and Kubelka-Munk theory, respectively. Considering the colorant, substrate and colorant/substrate interaction, the Yule-Nielsen model describes the perceived spectrum, $R(\lambda)$, as following [6]:

$$R(\lambda)^{\frac{1}{n}} = (1-A)R_s(\lambda)^{\frac{1}{n}} + AR_c(\lambda)^{\frac{1}{n}} \tag{2}$$

where $R_c(\lambda)$ and $R_s(\lambda)$ represent the reflectance spectrum of the colorant and the substrate respectively, and A is the relative substrate area percentage covered by colorants. The *N-factor*, n, in Equation (2) takes into account of multiple reflection at air-colorant-substrate interfaces. Equivalently, Yule-Nielsen model suggests that the nth root of the reflectance spectrums $\bigcup_{i=1}^{k} R_c^i(\lambda)$ and $R_s(\lambda)$ form a linear spectral mixing space.

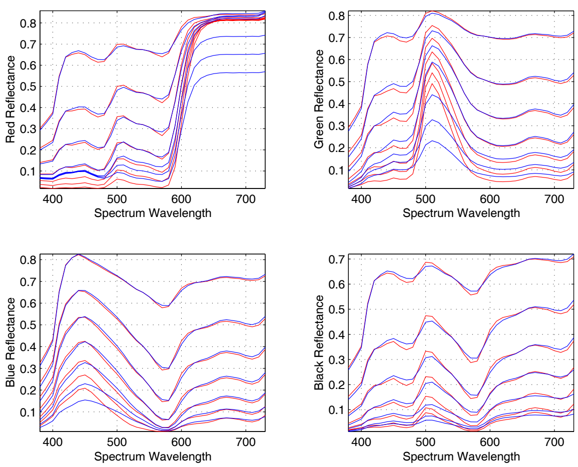

Fig. 1. Bi-Chrome Yule-Nielsen Model Fit

Let $K(\lambda)$ and $S(\lambda)$ be the absorption and scattering coefficients of a primary colorant, and the Kubelka-Munk theory describes the light reflection via a differential equation system [3,6]:

$$\frac{dI_i(w)}{dw} = [K(\lambda)+S(\lambda)]I_i(w) - S(\lambda)I_r(w) \tag{3}$$

$$\frac{dI_r(w)}{dw} = S(\lambda)I_i(w) - [K(\lambda)+S(\lambda)]I_r(w) \tag{4}$$

where $I_i(w)$ and $I_r(w)$ are the incident and reflected light intensity at position w within the colorant. $w = 0$ is the substrate-colorant interface and $w = W$ represents the colorant/air interface. Thus, the reflectance spectrum $R(\lambda) = I_r(w)/I_i(w)$ and $R_s(\lambda) = I_r(0)/I_i(0)$. Further assume that $K \gg S$, and the above differential equation system can be solved as following:

$$K(\lambda)W = \psi(\lambda) = -0.5\ln(R(\lambda)/R_s(\lambda)). \tag{5}$$

Because $\psi(\lambda)$ is proportional to the light absorbed by the colorant and previous literatures have demonstrated its mixture linearity in paint and plastics [1], we will also test the above logarithmic transformation to linearize the colorant mixing problem.

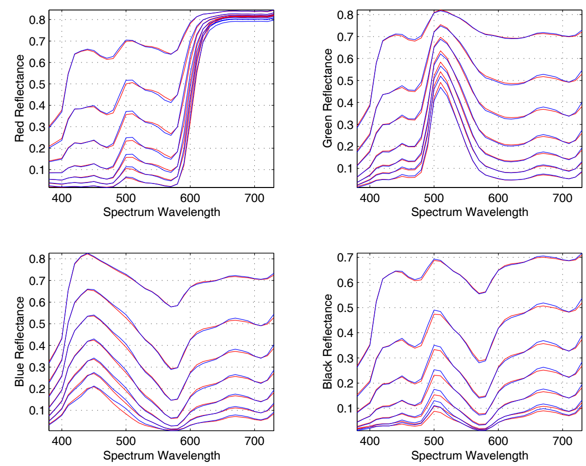

Fig. 2. Bi-Chrome Kubelka-Munk Theory Fit

A set of known primary colorants are adopted to create patches with two different colorants. The fitness between the measured and the fitted reflectance spectrums are used to verify the linearity of two transformations as noted in Equations (2) and (5), where the red curve is the measured reflectance and blue curve is the fitted reflectance. Figure 1 and 2 show that the Kubelka-Munk theory provides a better linear spectral mixing transformation. Note that $n = 3$ is selected in the Yule-Nielsen model.

3. PROPOSED ALGORITHM

Based on previous spectral mixture linearity analysis, we will first map the measured reflectance spectrum to $\psi(\lambda)$ via Equation (5). Let $\bigcup_{i=1}^{u} \psi_i(\lambda)$ be the set of unknown primary signals in the transformed spectral space, and each reflectance measurement can be decomposed as following:

$$\phi(\lambda) = \sum_{i=1}^{u} \gamma_i \psi_i(\lambda) \tag{6}$$

where γ_i and $\psi_i(\lambda)$ are unknown independent vectors with nonnegative elements. Assume the reflectance spectrum is measured at p locations and q frequency bands, and the above decomposition can be written as a matrix form:

$$\Phi = \Psi\Gamma, \tag{7}$$

where Φ, Γ and Ψ are $q \times p$, $p \times u$ and $q \times u$ matrices respectively, and $p, q \gg u$. This is similar matrix factorization as the *PCA* and *ICA* except for the nonnegative constraint on the elements of Γ and Ψ.

3.1. Independent Component Analysis

Let $\psi_i(\lambda_k) = \mu_i + \bar{\psi}_i(\lambda_k)$ such that $\sum_{k=1}^{q} \psi_i(\lambda_k) = q\mu_i$ and $\sum_{k=1}^{q} \bar{\psi}_i(\lambda_k) = 0$. Equation (6) can be rewritten as following:

$$\phi = \sum_{i=1}^{u}(\gamma_i\mu_i + \gamma_i\bar{\psi}_i) = \sum_{i=1}^{u}\gamma_i\mu_i + \bar{\phi}, \qquad (8)$$

where $\sum_{k=1}^{q} \bar{\phi}(\lambda_k) = 0$. Hence, considering the vector components with zero mean, Equation (7) is updated as

$$\bar{\Phi} = \Gamma\bar{\Psi}. \qquad (9)$$

Because each primary reflectance can be assumed to be independent of each other, a set of independent components, denoted as $\bigcup_{i=1}^{u} \bar{\varphi}_i(\lambda)$, with zero mean can be estimated via *ICA* decomposition [7]. Note that $\bar{\varphi}_i(\lambda)$ is not necessarily equivalent to the basis components, $\bar{\psi}_i(\lambda)$, thus the *DC* bias, μ_i, still needs to be estimated. As a result, we can compute $\varphi_i = \alpha_i + \bar{\varphi}_i$ satisfying the nonnegative constraint with minimal α_i in order to serve as the initial condition for the following *NNMF* algorithm. It can be readily shown that

$$\varphi_i = |\min_{\lambda} \bar{\varphi}_i(\lambda)| + \bar{\varphi}_i. \qquad (10)$$

3.2. POCS Non-Negative Matrix Factorization

The *NNMF* algorithm has been adopted to solve problems where the nonnegativity constraint, C, on the matrices elements arise. Various algorithms based on minimizing a designed cost function have been proposed [8,9,10]. In this paper, we propose to solve this problem by first representing C as an intersection of two convex sets. Consequently, a *POCS NNMF* algorithm can be formulated, and it can be shown to weakly converge to a point in C.

Assuming Ψ is known in Equation (7), the nonnegativity constraint $C_1 \equiv \{\gamma_{\bullet i} \geq 0, i = 1 \cdots p\}$ is a convex set, and the projection operator P_1:

$$P_1 \equiv \{\gamma_{\bullet i} | \min_{\gamma_{\bullet i} \in C_1} \|\phi_{\bullet i} - \Psi\gamma_{\bullet i}\|, i = 1 \cdots p\} \qquad (11)$$

can be easily solved independently via mathematical programming [11]. On the other hand, if Γ is known, the nonnegative constraint $C_2 \equiv \{\psi_{i\bullet} \geq 0, i = 1 \cdots q\}$ is also a convex set with the projection operator P_2:

$$P_2 \equiv \{\psi_{i\bullet} | \min_{\psi_{i\bullet} \in C_2} \|\phi_{i\bullet}^T - \Gamma^T\psi_{i\bullet}^T\|, i = 1 \cdots q\}. \qquad (12)$$

P_2 can be implemented by the same mathematical programming routine as P_1. Because $C = (C_1 \cap C_2) \neq \emptyset$, where C_1 and C_2 are closed convex sets, based on the following theorem of *POCS* proved by Gubin, Polyak and Raik, the operator $\bar{P} = P_2P_1$ will converge weakly [5].

Fundamental Theorem of POCS : Let C be an intersection of M closed convex sets in a Hilbert space H, and $T_i = I + \xi_i(P_i - I)$, where P_i is the projection operator onto C_i and $\xi_i \in (0, 2)$. Denote $T = T_M \cdots T_1$, and sequence $\{T^n x, n = 1 \cdots \infty\}$, $\forall x \in H$, converges weakly to a point in C.

$\xi_i = 1$ in our algorithm, but it is possible to be modified to improve the convergence rate.

4. EXPERIMENT RESULTS

The *ISO*-Standard test target as shown in Figure 3 is used to verify the proposed algorithm. This test target is reproduced by using four primary colorants: cyan, magenta, yellow and black. The reflectance spectrums of four primary colorants are measured a priori and serve as the *ground truth*, which is only used to compare with the recovered basis components.

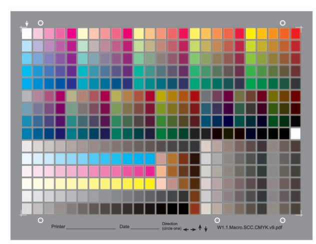

Fig. 3. W1.1 CMYK Test Target

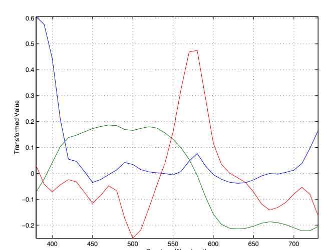

Fig. 4. Estimated 3 Independent Components

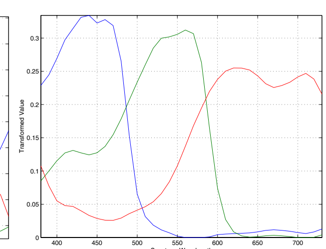

Fig. 5. *POCS NNMF* basis Components

Figure 4 shows that the first three independent components in the transformed spectral space, and Figure 5 is the corresponding recovered basis components using the proposed *POCS NNMF* algorithm. It is obvious that the components estimated by Equation (10) is less smooth than the final recovered basis components which exhibit as a set of bandpass filters with distinct passbands. Because the black colorant can be considered as an all-stop filter in the visible light frequency range, it can be replaced approximately by a linear combination of other primary colorants. As a result, a progressive recovery algorithm is adopted where one single basis component is estimated from points with less satisfying spectrum matches. Figure 6 demonstrates a good agreement between the measured and recovered reflectance spectrums of four primary colorants, represented by red and blue curve respectively. Finally, it shows in Figure 7 that the proposed *POCS NNMF* algorithm converges only after three iterations.

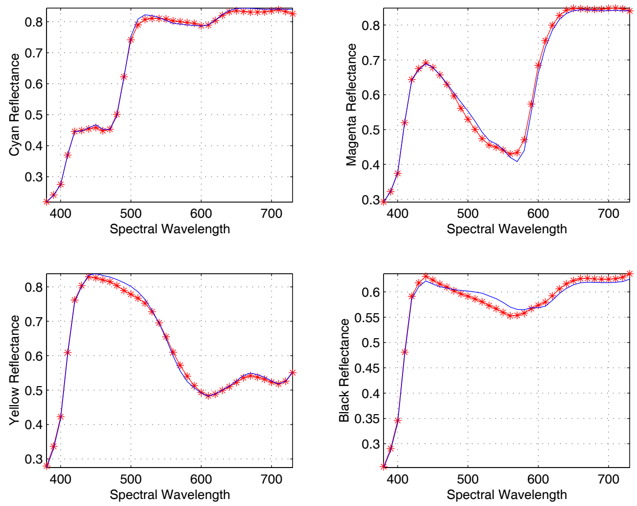

Fig. 6. Primary Colorant Reflectance Separation

5. CONCLUSION AND FUTURE WORK

A blind primary colorant spectral separation algorithm is proposed by combining *ICA* and *POCS NNMF*. By first mapping to a linear spectral-mixing space, a set of independent components are estimated via *ICA* and used as the starting point for the following *POCS NNMF* algorithm. We successfully demonstrate a good match between the measured and recovered reflectance spectrums of primary colorants. We will test our algorithm using actual printed images. This spectral separation algorithm can be further extended in the future to other applications such as remote sensing and hyperspectral data analysis.

6. REFERENCES

1 D-Y Tzeng and R. Berns, "Spectral-Based Ink Selection for Multiple-Ink Printing I: Colorant Estimation of Original Objects", *Proc. 6th IS&T Color Imaging Conf.*, pp.106-111, 1998.

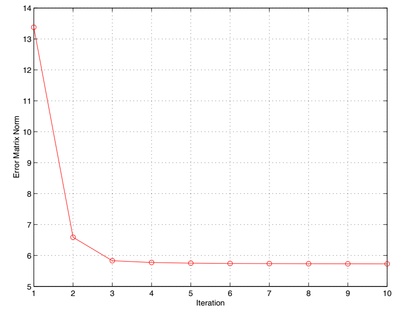

Fig. 7. *POCS NNMF* Algorithm Error Convergence

2 J. Bayliss, J. Gualtieri and R. Cromp, "Analyzing Hyperspectral Data with Independent Component Analysis", *Proc. SPIE AIPR Workshop*, Vol 9, 1997.

3 P. Emmel and R. Hersch, "An Unified Model for Color Prediction of Halftoned Prints", *J. Imaging Science and Technology*, Vol 44, pp.351-385, 2000.

4 D. Donoho and V. Stodden, "When does Non-Negative Matrix Factorization Give a Correct Decomposition into Parts", *Technical Report, Stanford Univ.*, 2003.

5 H. Stark and Y. Yang, "Vector Space Projections", *Wiley*, 1998.

6 H. Kang, "Color Technology for Electronic Imaging Devices", *SPIE Press*, 1997.

7 A. Cichocki and S. Amari, "Adaptive Blind Signal and Image Processing", *Wiley*, 2002.

8 P. Paatero and U. Tapper, "Positive Matrix Factorization: A non-negative factor model with optimal utilization of error estimate of data values", *Environmetrics*, Vol 5, pp.111-126, 1994.

9 D. Lee and S. Seung, "Learning the parts of objects by nonnegative matrix factorization", *Nature*, 401:788-791, 1999.

10 D. Lee and S. Seung, "Algorithms for Non-Negative Matrix Factorization", *Advances in Neural Information Processing*, 2000.

11 D. Bertsekas, "Convex Analysis and Optimization", *Athena Scientific*, 2003.