Link functions and Matérn kernel in the estimation of reflectance spectra from RGB responses

Ville Heikkinen,[1,*] Arash Mirhashemi,[1] and Juha Alho[2]

[1]*University of Eastern Finland, P.O. Box 111, FIN-80101 Joensuu, Finland*
[2]*University of Helsinki, P.O. Box 68, FIN-00014 Helsinki, Finland*
**Corresponding author: ville.heikkinen@uef.fi*

Received April 25, 2013; revised July 25, 2013; accepted October 4, 2013;
posted October 7, 2013 (Doc. ID 189432); published October 30, 2013

We evaluate three link functions (square root, logit, and copula) and Matérn kernel in the kernel-based estimation of reflectance spectra of the Munsell Matte collection in the 400–700 nm region. We estimate reflectance spectra from RGB camera responses in case of real and simulated responses and show that a combination of link function and a kernel regression model with a Matérn kernel decreases spectral errors when compared to a Gaussian mixture model or kernel regression with the Gaussian kernel. Matérn kernel produces performance similar to the thin plate spline model, but does not require a parametric polynomial part in the model. © 2013 Optical Society of America

OCIS codes: (100.3010) Image reconstruction techniques; (100.3190) Inverse problems; (150.0150) Machine vision; (330.1710) Color, measurement.
http://dx.doi.org/10.1364/JOSAA.30.002444

1. INTRODUCTION

Spectral reflectance image from object surface is a highly useful representation for object analysis and visualization. Here we concentrate on the problem of estimating pixel-wise reflectance spectra from RGB measurements using nonlinear estimation models and *a priori* knowledge. The motivation for using RGB cameras is that they are cheap and practical imaging devices and can rapidly collect many measurements from the same object in one shot. The results obtained with the RGB camera cannot compete with the quality of the traditional spectral imaging devices, but it is expected that the estimations are useful in many applications. For example, in the context of artwork imaging, it has been suggested that spectral reflectance estimations and RGB imaging systems (or other broadband systems with few bands [1]) are useful and practical [2,3].

Our earlier research in [4–6] suggests that one way to increase the accuracy of estimation is via the inclusion of *a priori* knowledge, such as the physically feasibility of all spectral reflectance values. Here we continue this work and produce physically feasible estimations via link functions and show that the precision of the estimations is increased especially in terms of spectral shape.

Our main focus is in the comparison of the performance of link functions when combined with so-called reproducing kernel Hilbert space (RKHS) based regression models [4–6]. We evaluate, apparently for the first time in reflectance estimation, the Matérn kernel and copula-based link functions. In particular, we demonstrate their performance relative to a model with the positive definite Gaussian kernel and conditionally positive definite thin plate spline (TPS) kernel. We compare the performance of copula-based link function to previously proposed transformations by using simulated and real data from the Munsell Matte collection [7,8]. Hence, we assume a context where the training and test sets come from same source, and a large, representative training set is used. Evaluation of models is done by using spectral and color differences in two extreme conditions: using ideal noise-free simulations and real noisy measurements with limited accuracy. In a simulated noise-free case, we use the Gaussian mixture model (GMM) as a reference method and show that performance of kernel models is significantly better in terms of standard spectral metrics. In the case of real data, we compare kernel-based models and show that a combination of link function and a model with a Matérn kernel decreases spectral errors when compared to previously evaluated kernel model with the Gaussian kernel.

2. ESTIMATION OF SPECTRAL REFLECTANCE

A. Observational Model

Let $\mathbf{q} \in \mathbb{R}^n$ be hyperspectral reflectance and suppose $x \in \mathbb{R}^k$ is a multispectral response of the form $\mathbf{x} = \mathbf{z} + \mathbf{e}$ where $\mathbf{e}$ is measurement noise that has expectation $E[\mathbf{e}] = \mathbf{0}$. We assume a linear observational model for multispectral response

$$\mathbf{z} = W\mathbf{q}, \tag{1}$$

where $W \in \mathbb{R}^{k\times n}$ is a known matrix that combines the effect of sensor and light source. The object to be estimated in following is $\mathbf{q} \in \mathbb{R}^n$ when $k < n$. We assume that a training set of the form

$$S = \{(\mathbf{x}_1, \mathbf{q}_1), \ldots, (\mathbf{x}_m, \mathbf{q}_m)\} \subset \mathbb{R}^k \times \mathbb{R}^n \tag{2}$$

is available, where $\mathbf{x}_i = W\mathbf{q}_i + \mathbf{e}_i,\ i = 1, \ldots, m$.

B. Estimation Using Wiener Model

For future reference, we note that if $\mathbf{u}_1$ and $\mathbf{u}_2$ are two jointly Gaussian random vectors with expectations $E[\mathbf{u}_i] = \mu_i$ and $i = 1, 2$ and covariances $\mathrm{Cov}(\mathbf{u}_i, \mathbf{u}_j) = \Sigma_{ij}$ and $i, j = 1, 2$, so that

$$\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right), \tag{3}$$

then the conditional distribution of $\mathbf{u}_2$ conditionally on $\mathbf{u}_1$ is ([9], p. 522)

$$\mathbf{u}_2|\mathbf{u}_1 \sim \mathcal{N}(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{u}_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}). \tag{4}$$

Assuming that we have Eq. (1) with $\mathbf{q} \sim \mathcal{N}(0, \Sigma_q)$ and $\mathbf{e} \sim \mathcal{N}(0, \sigma_e^2 I_k)$, i.e., we assume that the measurement errors are independent across response channels and have equal variances. Using Eq. (4), we have representation for the conditional mean $E[\mathbf{q}|\mathbf{x}]$ as

$$\hat{\mathbf{q}} = \Sigma_q W^T (W \Sigma_q W^T + \sigma_e^2 I_k)^{-1} \mathbf{x}. \tag{5}$$

Estimate in Eq. (5) is also called a Wiener estimate [10–14], and it is widely used in reflectance estimation. Because of the prior assumption, the values of estimations, however, are not constrained to be positive or in any bounded range. Also this model only utilizes the covariance matrix and mean of the spectral data in estimation.

C. Estimation Using a Gaussian Mixture Model

The Wiener model can be replaced by a much more general Gaussian mixture density for reflectance as suggested, e.g., by Murakami *et al.* [11]. Especially when an accurate (in terms of target reflectances) and sufficiently large set of spectral data is available, an estimation model based on the Gaussian mixture density can be much more accurate than the Wiener model. Therefore, as a reference method for reflectance estimation in the simulated case, we use a GMM.

In our case, it assumed that we have a training set in Eq. (2). Assuming that we have the observational model in Eq. (1) with $\mathbf{e} \sim \mathcal{N}(0, \sigma_e^2 I_k)$, we construct a Gaussian mixture density as a prior for future observations $\mathbf{q}$ of the form

$$p(\mathbf{q}) = \sum_{i=1}^{m} v_i p_i(\mathbf{q}), \tag{6}$$

where $v_i = 1/m$ is the prior probability that a new sample comes from a Gaussian distribution, $\mathbf{q} \sim \mathcal{N}(\mathbf{q}_i, \sigma_q^2 I_n)$, with density

$$p_i(\mathbf{q}) = \frac{1}{(2\pi)^{n/2}\sigma_q^n} \exp\left(-\frac{1}{2\sigma_q^2} \|\mathbf{q} - \mathbf{q}_i\|^2 \right), \qquad i = 1, \ldots, m. \tag{7}$$

If $\mathbf{q}$, corresponding to a new measurement $\mathbf{x}$, came from p_i, then

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{q} \end{bmatrix} = \mathcal{N}\left(\begin{bmatrix} W\mathbf{q}_i \\ \mathbf{q}_i \end{bmatrix}, \begin{bmatrix} \sigma_q^2 W W^T + \sigma_e^2 I_k & \sigma_q^2 W \\ \sigma_q^2 W^T & \sigma_q^2 I_n \end{bmatrix} \right). \tag{8}$$

Using notation $\gamma = \sigma_e^2/\sigma_q^2$ and Eq. (4) for conditional distribution, we have

$$\mathbf{q}|\mathbf{x} \sim \mathcal{N}(\tilde{\mathbf{q}}_i(\mathbf{x}), \Sigma) \tag{9}$$

with the mean and the covariance

$$\tilde{\mathbf{q}}_i(\mathbf{x}) = W^T (W W^T + \gamma I_k)^{-1} (\mathbf{x} - W\mathbf{q}_i) + \mathbf{q}_i, \tag{10}$$

$$\Sigma = \sigma_q^2 (I_n - W^T (W W^T + \gamma I_k)^{-1} W), \tag{11}$$

provided that $\mathbf{q}$ came from p_i.

Assuming Eq. (1) such that $\mathbf{q} \sim \mathcal{N}(\mathbf{q}_i, \sigma_q^2 I_n)$ with density Eq. (7), and independently $\mathbf{e} \sim \mathcal{N}(0, \sigma_e^2 I_k)$, we have

$$\mathbf{x}|\mathbf{q} \sim \mathcal{N}(W\mathbf{q}, \sigma_e^2 I_k), \tag{12}$$

for all $\mathbf{q}_i$. This common density is denoted as $p(\mathbf{x}|\mathbf{q})$. The joint density of $(\mathbf{x}, \mathbf{q})$ is denoted as $p_i(\mathbf{x}, \mathbf{q}) = p_i(\mathbf{q}|\mathbf{x}) p_i(\mathbf{x})$ where $p_i(\mathbf{q}|\mathbf{x})$ is the Gaussian density with mean Eq. (10) and covariance Eq. (11), and $p_i(x)$ is the marginal distribution. The joint density can be also written as

$$p_i(\mathbf{x}, \mathbf{q}) = p(\mathbf{x}|\mathbf{q}) p_i(\mathbf{q}). \tag{13}$$

We have the posterior $p(\mathbf{q}|\mathbf{x}) = p(\mathbf{x}, \mathbf{q})/p(\mathbf{x})$ where $p(\mathbf{x}, \mathbf{q}) = p(\mathbf{x}|\mathbf{q}) p(\mathbf{q})$. Using Eqs. (6) and (13), we can write

$$p(\mathbf{x}|\mathbf{q}) p(\mathbf{q}) = \sum_{i=1}^{m} v_i p_i(\mathbf{q}|\mathbf{x}) p_i(\mathbf{x}), \tag{14}$$

$$p(\mathbf{x}) = \sum_{i=1}^{m} v_i p_i(\mathbf{x}), \tag{15}$$

and the representation for the posterior distribution

$$p(\mathbf{q}|\mathbf{x}) = \frac{\sum_{i=1}^{m} p_i(\mathbf{q}|\mathbf{x}) p_i(\mathbf{x})}{\sum_{i=1}^{m} p_i(\mathbf{x})}. \tag{16}$$

The conditional probability that $\mathbf{q}$ came from p_i, given observed $\mathbf{x}$ is

$$\tilde{p}_i(\mathbf{x}) = p_i(\mathbf{x}) / \sum_{i=1}^{m} p_i(\mathbf{x}), \tag{17}$$

where

$$p_i(\mathbf{x}) \propto \exp\left(-\frac{1}{\sigma_q^2} (\mathbf{x} - W\mathbf{q}_i)^T (W W^T + \gamma I_k)^{-1} (\mathbf{x} - W\mathbf{q}_i) \right). \tag{18}$$

The estimate of $\mathbf{q}$ is then obtained from Eq. (16) as the conditional mean

$$\hat{\mathbf{q}} = E[\mathbf{q}|\mathbf{x}] = \sum_{i=1}^{m} \tilde{p}_i(\mathbf{x}) E[\mathbf{q}_i|\mathbf{x}] = \sum_{i=1}^{m} \tilde{p}_i(\mathbf{x}) \tilde{\mathbf{q}}_i(\mathbf{x}), \tag{19}$$

Defining $Q = [\mathbf{q}_1 \ldots \mathbf{q}_m]^T \in \mathbb{R}^{m \times n}$, the estimate in Eq. (19) can be formulated as a combination of a parametric and a nonparametric part

$$\hat{\mathbf{q}} = (Q^T - W^T(WW^T + \gamma I_k)^{-1}WQ^T)\tilde{\mathbf{p}}(\mathbf{x}) + W^T(WW^T + \gamma I_k)^{-1}\mathbf{x}, \quad (20)$$

where $\tilde{\mathbf{p}}(\mathbf{x}) = (\tilde{p}_1(\mathbf{x}), \ldots, \tilde{p}_m(\mathbf{x}))^T$ and $\tilde{p}_i(\mathbf{x})$ are calculated using Eq. (17). The first term is the weighted average of the residual from Wiener estimates, and the second term corresponds to a Wiener estimate in Eq. (5) with $\Sigma_q = \sigma_q^2 I_n$.

D. Estimation Using Kernel-Based Regression

We utilize a kernel-based regression model for reflectance estimation as suggested in [4–6]. The goal is to estimate a mapping $\mathbf{x} \to \mathbf{q}$ using a regression model, where we assume that $\mathbf{x} \in \mathcal{X}$ and $\mathcal{X}$ is the closed and bounded subset of $\mathbb{R}^k$.

In the simplest case, a vector valued regression model for mapping $\mathbf{x} \to \mathbf{q}$ is formulated with independent scalar regressions for components of vector $\mathbf{q}$. Let us write $Q = [q_{ji}]$, $j = 1, \ldots, m$, $i = 1, \ldots, n$. We denote a training set $S_i = \{(\mathbf{x}_1, q_{1i}), \ldots, (\mathbf{x}_m, q_{mi})\} \subset \mathbb{R}^k \times \mathbb{R}$ for $i = 1, \ldots, n$, and the scalar regression is formulated as the minimization problem

$$V[f_i] = \sum_{j=1}^{m} (q_{ji} - f_i(\mathbf{x}_j))^2 + \eta \|f_i\|_{\mathcal{H}}^2, \quad (21)$$

over $f_i \in \mathcal{H}$, where $\mathcal{H}$ is an RKHS of functions $f_i : \mathcal{X} \to \mathbb{R}$ induced by a symmetric, continuous, and positive definite function (Mercer kernel) $\kappa : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ [15,16]. The first term on the right-hand side measures the fit of the function to the training data under the squared error loss. The term $\|f\|_{\mathcal{H}}^2$ in Eq. (21) is the square of a norm of f, and η is a scalar parameter, which allows us to control the balance between lack of fit and the RKHS norm of the solution. The term $\eta\|f\|_{\mathcal{H}}^2$ is called as a regularization functional, and its purpose is to reduce wiggliness of f and improve generalization properties of models. By using the Mercer kernel κ of $\mathcal{H}$, the representation for the function $f_i \in \mathcal{H}$ that minimizes Eq. (21) can be written as

$$f_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{k}(\mathbf{x}), \quad (22)$$

where $\mathbf{a}_i = (a_{i1}, a_{i2}, \ldots, a_{im})^T \in \mathbb{R}^m$ is unique, and $\mathbf{k}(\mathbf{x}) = (\kappa(\mathbf{x}_1, \mathbf{x}), \kappa(\mathbf{x}_2, \mathbf{x}), \ldots, \kappa(\mathbf{x}_m, \mathbf{x}))^T \in \mathbb{R}^m$ [15,16]. The squared norm is written as $\|f_i\|_{\mathcal{H}}^2 = \mathbf{a}_i^T K \mathbf{a}_i$, where $K \in \mathbb{R}^{m \times m}$ is the kernel matrix of training data, with $K_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$.

We focus on this approach by using the squared error loss and the same model for all the $i = 1, \ldots, n$ components. The minimization problems for $i = 1, \ldots, n$ combined are written

$$\arg\min_{A \in \mathbb{R}^{m \times n}} V(A) = \sum_{j=1}^{m} \|\mathbf{q}_j - A^T \mathbf{k}(\mathbf{x}_j)\|^2 + \eta \,\mathrm{Tr}(A^T K A), \quad (23)$$

where $A = [\mathbf{a}_1 \ldots \mathbf{a}_n] \in \mathbb{R}^{m \times n}$, $\mathrm{Tr}(\cdot)$ denotes matrix trace, $K \in \mathbb{R}^{m \times m}$ is the kernel matrix of training data, and $\mathbf{k}(\mathbf{x}_j)^T = (\kappa(\mathbf{x}_1, \mathbf{x}_j), \ldots, \kappa(\mathbf{x}_m, \mathbf{x}_j)) \in \mathbb{R}^m$ is a row vector containing the kernel evaluations between the training set and point $\mathbf{x}_j$. The solution for Eq. (23) is written as

$$\hat{A} = (K + \eta I_m)^{-1} Q \quad (24)$$

and the estimate of $\mathbf{q}$ corresponding to $\mathbf{x}$ is evaluated as

$$\hat{\mathbf{q}} = \hat{A}^T \mathbf{k}(\mathbf{x}) = Q^T (K + \eta I_m)^{-1} \mathbf{k}(\mathbf{x}). \quad (25)$$

The estimate in Eq. (25) is a linear combination of training spectra, i.e., $\hat{\mathbf{q}} = Q^T \mathbf{p}(\mathbf{x})$, where $\mathbf{p}(\mathbf{x}) = (K + \eta I_m)^{-1} \mathbf{k}(\mathbf{x}) \in \mathbb{R}^m$ denotes coefficient vector corresponding to measurement $\mathbf{x}$.

In the following, we evaluate the nonparametric estimate in Eq. (25) by considering two strictly positive definite kernels. As a first choice, we have the Gaussian kernel

$$\kappa_1(\mathbf{x}, \mathbf{v}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{v}\|^2}{2\varsigma^2}\right), \quad (26)$$

where $\mathbf{x}, \mathbf{v} \in \mathcal{X}$, and $\varsigma > 0$ is the scale parameter. As a second choice, we consider a class of Matérn functions ([17], p. 85) and choose the kernel

$$\kappa_2(\mathbf{x}, \mathbf{v}) = \left(1 + \frac{\sqrt{3}\|\mathbf{x} - \mathbf{v}\|}{\varsigma}\right) \exp\left(-\frac{\sqrt{3}\|\mathbf{x} - \mathbf{v}\|}{\varsigma}\right), \quad (27)$$

where $\mathbf{x}, \mathbf{v} \in \mathcal{X}$ and $\varsigma > 0$ is the free parameter.

We compare the performance of these two kernels to the performance of strictly d-conditionally positive definite TPS (Duchon's spline) that was found to give best performance in [4,6]. Here we use the kernel

$$\kappa_3(\mathbf{x}, \mathbf{v}) = c_{d,k}\|\mathbf{x} - \mathbf{v}\|^{2d-k}, \qquad c_{d,k} = \frac{\Gamma(k/2 - d)}{2^{2d}\pi^{k/2}(d-1)!}, \quad (28)$$

where $2d - k > 0$ is odd, and Γ denotes the gamma function. For this kernel, the minimization problem in Eq. (23) needs to be modified to include a parametric polynomial part, so that we have the minimization

$$\arg\min_{B \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{N \times n}} V(B, C) = \sum_{j=1}^{m} \|\mathbf{q}_j - B^T \mathbf{k}(\mathbf{x}_j) - C^T \Psi(\mathbf{x}_j)\|^2 + \eta \,\mathrm{Tr}(B^T K B) \quad \text{s.t. } \Psi(X)^T B = \mathbf{0}, \quad (29)$$

where elements of $\Psi(\mathbf{x}_j) \in \mathbb{R}^N$ include all monomial terms of polynomial with k variables and total degree $d - 1$ and $\Psi(X) = [\Psi(\mathbf{x}_1) \cdots \Psi(\mathbf{x}_m)]^T \in \mathbb{R}^{m \times N}$ is the matrix of the polynomial terms as rows [6,16]. The semi-parametric estimate of $\mathbf{q}$ corresponding to $\mathbf{x}$ is

$$\hat{\mathbf{q}} = \hat{B}^T \mathbf{k}(\mathbf{x}) + \hat{C}^T \Psi(\mathbf{x}), \quad (30)$$

where $\hat{B}$ and $\hat{C}$ are the solution for Eq. (29) (coefficient matrices presented in [6]).

3. LINK FUNCTIONS

The values of reflectance spectrum $\mathbf{q}$ are positive or even restricted to region [0,1]. Especially for the Lambertian surfaces, all the measurements are in [0,1] due to physical reasons. However, it is usual also for practical measurements that reflectance data have values restricted to [0,1]. Motivated by these constraints, we utilize transformation for elements of reflectance vectors that provide physically feasible values in estimations (positive or in region [0,1] depending on transformation). We call the functions defining these transformations as *link functions*.

Here, the kernel-based learning models are combined with the transformation for $\mathbf{q}$ via link functions. In detail, the reflectance vectors $\mathbf{q}_j, j = 1, \ldots, m$ in Eqs. (23) and (29) are replaced with transformed vectors $\tilde{\mathbf{q}}_j = \mathcal{T}(\mathbf{q}_j)$ in the training phase of the regression model (and recovered via inverse link function when needed, i.e., $\mathbf{q} = \mathcal{T}^{-1}(\tilde{\mathbf{q}})$). We evaluated three link functions $\mathcal{T}$:

(1) *The square root* $\tilde{\mathbf{q}} = \sqrt{\mathbf{q}}$ with element-wise calculation and with inverse transformation $\mathbf{q} = \tilde{\mathbf{q}}^2$.

(2) *The logit function*

$$\tilde{\mathbf{q}} = \text{logit}(\mathbf{q}) = \log\left(\frac{\mathbf{q}}{1-\mathbf{q}}\right), \tag{31}$$

where $\log :]0, +\infty[\to \mathbb{R}$ is the natural logarithm evaluated element-wise for $\mathbf{q} \in]0,1[^n$. The inverse transformation is defined using component-wise *logistic function*

$$\mathbf{q} = \frac{\exp(\tilde{\mathbf{q}})}{1+\exp(\tilde{\mathbf{q}})}. \tag{32}$$

(3) *A Gaussian copula*

$$\tilde{\mathbf{q}} = \Phi^{-1}(F(\mathbf{q})) = \Phi^{-1}(F_1(q_1), F_2(q_2), \ldots, F_n(q_n))^T, \tag{33}$$

where F_i denotes the beta cumulative distribution function with parameters a_i, b_i, and Φ^{-1} as the inverse of the normal cumulative distribution function $\Phi(x|\mu = 0, \sigma = 1) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} \exp((-t^2)/2)\mathrm{d}t$ evaluated element-wise for $F_i(q_i)$, $i = 1, \ldots, n$. The transformed values are

$$\tilde{q}_i = \Phi^{-1}(F_i(q_i)) = \Phi^{-1}\left(\frac{\Gamma(a_i+b_i)}{\Gamma(a_i)\Gamma(b_i)} \int_0^{q_i} t^{a_i-1}(1-t)^{b_i-1}\mathrm{d}t\right), \quad i = 1, \ldots, n, \tag{34}$$

where Γ is the gamma function. The idea in this transformation is based on the assumption that the elements q_i of reflectance vector $\mathbf{q} = (q_1, q_2, \ldots, q_n)^T$ are in [0,1] and marginally distributed according to the beta distribution. By evaluating the beta cumulative distribution function $F_i(q_i)$ as above (parameters a_i, b_i estimated by using available data and maximum likelihood), the values become uniformly distributed. The uniformly distributed values are further transformed via the inverse of the normal cumulative distribution function Φ^{-1}, so that resulting values $\tilde{q}_i$ in Eq. (34) become normally distributed. The inverse transformation is defined element-wise using

$$q_i = F_i^{-1}(\Phi(\tilde{q}_i)), \qquad i = 1, \ldots, n, \tag{35}$$

where $\Phi(\tilde{q}_i) = \Phi(\tilde{q}_i|\mu = 0, \sigma = 1)$ is the normal cumulative distribution function, and F_i^{-1} is the inverse of the beta cumulative distribution function with same parameters a_i, b_i as used in Eq. (34).

Only the logit and copula link functions provide restriction to [0,1], whereas the square root can be used to provide positive estimates. In our approach, the link functions provided only the transformed reflectance data $\tilde{\mathbf{q}}$ in minimizations of Eqs. (23) and (29) without any transform to response vectors $\mathbf{x}$. After the evaluation of the model (and during parameter optimization), return to original reflectance scale was performed via inverse link function.

4. EXPERIMENTS

In the experiments, we used the reflectance data of the *Munsell Book of Color—Matte Finish Collection* (400–700 nm wavelength range with 5 nm sampling) [8] and corresponding RGB responses for evaluations. Using these data, we had reflectance $\mathbf{q} \in \mathbb{R}^{61}$ and corresponding observation $\mathbf{x} \in \mathbb{R}^3$.

A. Simulated and Real RGB Data

In the first experiment, we simulated RGB values for the Munsell Matte colors by using the combination of CIE $\bar{x}, \bar{y}, \bar{z}$ 1931 system and CIE D65 illuminant as the sensor system W in the observational model of Eq. (1). Before response simulation, the light source and the CIE $\bar{x}, \bar{y}, \bar{z}$ responsivities were scaled to unit vectors. In this experiment, we performed estimations by using simulations with a linear and noise-free observation model with the aim to give an indication about the performance of the models in ideal conditions.

In the second experiment, we used the measured RGB values (sRGB, JPEG) for the Munsell Matte colors using Fujifilm Finepix S1 Pro digital camera with a Nikon AFNikkor 25–50 mm zoom lens. The camera's f-number was set to F4.8, ISO was set to 400, white balance set to fluorescent ("Fluorescent 1"), and the exposure time to 1/45s. The pages of the Munsell book were attached to a vertical surface, and the camera was set up orthogonally toward the surface. Denoising of the RGB measurements was performed by calculating spatial averages of 5×5 pixel area from the middle of each color chip. We also removed 36 samples with saturated response values.

This same real RGB data were used in [4–6], with different processing for saturated responses. There are several limitations in these data that can be listed as (1) measured using RGB camera with a somewhat aged technique (announced January 2000) with a Fuji SuperCCD/RGBG filter with some unknown demosaicing method; (2) unknown spectral responsivity; (3) Fluorescent Philips DeLuxe90 TLD 18W/965 as a light source during measurements; (4) default spectral nonlinearity (without additional post-processing); (5) JPEG values with sRGB color space; (6) default spatial uniformity (without additional postprocessing). Due to these limitations, the quality of real RGB data correspond to somewhat extreme conditions and can be easily improved in terms of measurement approach, data format, and device properties.

Our evaluation setting therefore includes two extreme cases for RGB measurements: ideal noise-free simulated measurements with CIE system and real noisy responses (averaged, though) that have far from optimal quality. The details of the two experiments are summarized in Table 1.

B. Parameter Optimization and Evaluation

For training set S in Eq. (2), we randomized 90% of the Munsell samples for training and used the remaining 10% for testing. Before model constructions and evaluations, we combined the RGB responses from training and tests and scaled them to range [0,1]. All the free parameters in the models were optimized by using tenfold cross-validation [minimization of average of squared RMSE, see Eq. (36) below] in the training set [18]. For the GMM in Eq. (19), we fixed $\gamma = 0$ due to the

Table 1. Overview over Training and Test Sets[a]

Experiment	Training/Test Set	Sensor	Nonlinearities	Spectral Responsivity	Color Space	Format	Light	Noise	Denoising
1: Simulated	Munsell 90%/10%	CIE $\bar{x},\bar{y},\bar{z}$ 1931	—	Known	CIE XYZ	—	CIE D65	—	—
2: Real	Munsell* 90%/10%	Fujifilm Finepix Pro 1	Yes	Unknown	sRGB	JPEG	Fluorescent	Real	Averaging

[a]Munsell* denotes a set of 1232 Munsell samples (corresponding to nonsaturated RGB values). Reflectance data was represented in 400–700 nm range using 5 nm sampling.

noise-free model and optimized the σ_q in Eq. (18). For the kernel model in Eq. (25) with the Gaussian in Eq. (26) and the Matérn kernel in Eq. (27), we optimized the regularization η and kernel parameter ς. In case of the TPS model, we fixed $d = 3$ in Eq. (28) and optimized only the regularization parameter η in Eq. (29). In this case, the parametric part of TPS was a multivariate polynomial with 10 monomial terms and total degree 2.

Two consecutive cross-validation search grids (search lines for the GMM and TPS) for the free parameters were used: the first grid provided the coarse location of parameters, and the second was a refined grid in the neighborhood of coarse location. The first grid was defined so that the kernel parameters for the Matérn and Gaussian kernels were sampled in range $[2^{-4}, 2^{4}]$ using 17 sampling points ($\{2^{-4:0.5:4}\}$) and the regularization parameters (all kernel models) were sampled in range $[2^{-30}, 2^{-4}]$ using 27 sampling points ($\{2^{-30:1:-4}\}$). The second grid was defined similarly in the neighborhood of coarse location using 20×20 grid size. During the optimization of parameters, the link functions were used so that the errors were always evaluated in original reflectance scale (via inverse link function). The GMM was optimized so that the cross validation was using 60 sampling points in coarse range $[10^{-6},1]$. The finer sampling was defined using 91 points in the neighborhood of optimal coarse location.

By using the optimized parameters, we evaluated the accuracy for the independent test set and calculated several spectral and color error values as presented below. As a second case, we also produced 200 other data permutations randomly (via MATLAB and uniformly distributed pseudorandom numbers) for training (90% of data) and test set (10% of data) and evaluated the average performance of these permutations by using the same optimized parameters as found by the cross-validation method above. The purpose of this calculation is to reduce data-sampling effects from accuracy evaluation.

In the evaluation of results for each spectrum in the test set, we used RMSE,

$$\mathrm{RMSE}(\mathbf{q},\hat{\mathbf{q}}) = \sqrt{\|\mathbf{q}-\hat{\mathbf{q}}\|^2/n}, \tag{36}$$

and Pearson distance (PD)

$$\mathrm{PD}(\mathbf{q},\hat{\mathbf{q}}) = 1 - \frac{\mathbf{q}^T\hat{\mathbf{q}}}{\|\mathbf{q}\|\|\hat{\mathbf{q}}\|}. \tag{37}$$

The PD (1 – PD is called as GFC [19,20]) is independent of the magnitude and therefore gives information about the shape of estimations. Evaluation of colorimetric accuracy of estimations were performed by using DE2000 distance and CIE D65, CIE A, and F11 illuminants. We also calculated wavelength-wise MSE for each wavelength location:

$$\mathrm{MSE}(i) = \sum_{j=1}^{l}(q_{ji}-\hat{q}_{ji})^2/l, \qquad i = 1,\ldots,n, \tag{38}$$

where l denotes the number of test samples, and q_{ji}, $\hat{q}_{ji}$ denote the elements of reflectance matrices for test set samples

$$Q = [\mathbf{q}_1\mathbf{q}_2\cdots\mathbf{q}_l]^T = [q_{ji}], \qquad j = 1,\ldots,l, \qquad i = 1,\ldots,n \tag{39}$$

and estimated test set samples

$$\hat{Q} = [\hat{\mathbf{q}}_1\hat{\mathbf{q}}_2\cdots\hat{\mathbf{q}}_l]^T = [\hat{q}_{ji}], \qquad j = 1,\ldots,l, \qquad i = 1,\ldots,n. \tag{40}$$

5. RESULTS

The numerical results for Experiments 1 and 2 (Table 1) are presented in Tables 2 and 3. The results in Table 2 for spectral accuracy (RMSE and PD) are presented using average, maximum values and 95th percentile values. Color differences in Table 3 are presented using average and maximum values for D65, A, and F11 illuminants. In both tables, the values in parentheses represent the average values corresponding to 200 data permutations to training and test sets.

A. Experiment 1: Simulated RGB

The numerical results for simulated data in Tables 2 and 3 can be summarized as follows:

(1) The GMM and kernel models without link function (Table 2, upper half): in this case, all three kernel models provide lower average, maximum, and 95th percentile RMSE values than the GMM, but the maximum and 95th percentile of PD values indicate better performance for the GMM. Averaged results from 200 randomization for training and test set show similar performance for the RMSE, but in this, case the kernel models with the Matérn and TPS kernels provide the best results also for all the PD metrics. These two kernels also provide significant improvements when compared to performance of the Gaussian kernel. For example, when the Matérn kernel is considered, it can be seen that (when compared to the Gaussian) the average RMSE metric decreases 13.5% (averaged result decreases 12.6%), the maximum RMSE metric increases 7% (averaged result, however, is similar), and the 95th percentile of RMSE decreases 2.9% (averaged result decreases 2.8%). Similarly for the PD metric: the average PD metric decreases 18.2% (averaged result decreases 20.4%), the maximum PD metric decreases 30.8% (averaged result decreases 8.6%), and the 95th percentile of PD increases 9.3% (averaged result, however, decreases 19.3%). All the results for the TPS kernel are similar to the Matérn kernel results

Table 2. Spectral Error Values[a]

Method	RMSE Avg.	RMSE Max.	RMSE 95th	PD Avg.	PD Max.	PD 95th
Experiment 1 (Simulated RGB data)						
Link function: none						
GMM	0.0129 (0.0116)	0.0778 (0.0786)	0.0402 (0.0370)	0.00161 (0.00230)	0.0152 (0.0694)	0.00694 (0.00887)
Gaussian kernel	0.0111 (0.0103)	0.0532 (0.0599)	0.0342 (0.0289)	0.00165 (0.00186)	0.0399 (0.0419)	0.00689 (0.00768)
Matérn kernel	0.0096 (0.0090)	0.0572 (0.0595)	0.0332 (0.0281)	0.00135 (0.00148)	0.0276 (0.0383)	0.00760 (0.00620)
TPS kernel	0.0096 (0.0090)	0.0570 (0.0595)	0.0328 (0.0281)	0.00135 (0.00147)	0.0280 (0.0382)	0.00757 (0.00617)
Link function: square root						
Gaussian kernel	0.0107 (0.0099)	0.0512 (0.0610)	0.0323 (0.0291)	0.00120 (0.00154)	0.0242 (0.0417)	0.00508 (0.00608)
Matérn kernel	0.0093 (0.0087)	0.0554 (0.0597)	0.0319 (0.0278)	0.00101 (0.00131)	0.0137 (0.0401)	0.00538 (0.00513)
TPS kernel	0.0094 (0.0087)	0.0551 (0.0598)	0.0328 (0.0278)	0.00101 (0.00130)	0.0139 (0.0398)	0.00530 (0.00512)
Link function: logit						
Gaussian kernel	0.0103 (0.0097)	0.0508 (0.0601)	0.0333 (0.0288)	0.00104 (0.00150)	0.0154 (0.0443)	0.00462 (0.00569)
Matérn kernel	0.0092 (0.0086)	0.0558 (0.0598)	0.0326 (0.0276)	0.00088 (0.00125)	0.0083 (0.0417)	0.00420 (0.00467)
TPS kernel	0.0092 (0.0086)	0.0556 (0.0598)	0.0332 (0.0276)	0.00088 (0.00123)	0.0083 (0.0411)	0.00411 (0.00463)
Link function: copula						
Gaussian kernel	0.0105 (0.0098)	0.0517 (0.0601)	0.0327 (0.0289)	0.00113 (0.00151)	0.0201 (0.0424)	0.00480 (0.00592)
Matérn kernel	0.0094 (0.0087)	0.0563 (0.0597)	0.0330 (0.0278)	0.00096 (0.00127)	0.0110 (0.0402)	0.00482 (0.00487)
TPS kernel	0.0094 (0.0087)	0.0562 (0.0597)	0.0329 (0.0278)	0.00096 (0.00125)	0.0110 (0.0396)	0.00468 (0.00482)
Experiment 2 (Real RGB data)						
Link function: none						
Gaussian kernel	0.0148 (0.0150)	0.0533 (0.0621)	0.0336 (0.0373)	0.00138 (0.00142)	0.0288 (0.0327)	0.00724 (0.00537)
Matérn kernel	0.0132 (0.0139)	0.0493 (0.0630)	0.0317 (0.0351)	0.00117 (0.00128)	0.0207 (0.0325)	0.00595 (0.00507)
TPS kernel	0.0135 (0.0141)	0.0507 (0.0610)	0.0325 (0.0357)	0.00121 (0.00128)	0.0237 (0.0318)	0.00617 (0.00520)
Link function: square root						
Gaussian kernel	0.0148 (0.0149)	0.0528 (0.0619)	0.0328 (0.0372)	0.00142 (0.00133)	0.0432 (0.0325)	0.00615 (0.00496)
Matérn kernel	0.0134 (0.0139)	0.0494 (0.0615)	0.0336 (0.0351)	0.00124 (0.00121)	0.0285 (0.0306)	0.00581 (0.00481)
TPS kernel	0.0135 (0.0140)	0.0497 (0.0610)	0.0321 (0.0357)	0.00124 (0.00121)	0.0295 (0.0303)	0.00614 (0.00480)
Link function: logit						
Gaussian kernel	0.0146 (0.0148)	0.0511 (0.0619)	0.0318 (0.0374)	0.00141 (0.00129)	0.0432 (0.0314)	0.00661 (0.00482)
Matérn kernel	0.0133 (0.0139)	0.0493 (0.0618)	0.0324 (0.0352)	0.00124 (0.00119)	0.0301 (0.0303)	0.00545 (0.00472)
TPS kernel	0.0133 (0.0140)	0.0494 (0.0612)	0.0302 (0.0356)	0.00123 (0.00119)	0.0307 (0.0297)	0.00586 (0.00469)
Link function: copula						
Gaussian kernel	0.0147 (0.0148)	0.0518 (0.0621)	0.0329 (0.0373)	0.00138 (0.00130)	0.0384 (0.0312)	0.00640 (0.00486)
Matérn kernel	0.0133 (0.0139)	0.0495 (0.0617)	0.0329 (0.0351)	0.00123 (0.00120)	0.0288 (0.0306)	0.00559 (0.00476)
TPS kernel	0.0133 (0.0140)	0.0497 (0.0612)	0.0306 (0.0356)	0.00122 (0.00120)	0.0287 (0.0300)	0.00579 (0.00474)

[a]Number in parenthesis is average error from 200 randomizations to training and test sets.

(see also similarity in wavelength-wise errors for the Matérn and TPS in Fig. 1).

(2) Kernel models with and without link function (Table 2, upper half): the use of link functions provide slightly better results in terms of the RMSE and markedly better accuracy in terms of the PD metric. For example, for the Matérn kernel model with and without link function, in the best case (with the logit) the average PD metric decreases 34.8% (averaged result decreases 18.4%), the maximum PD metric decreases 69.9% (averaged result, however, increases 8.2%), and the 95th percentile of the PD decreases 44.7% (averaged result decreases 24.7%).

(3) Wavelength-wise error analysis in Fig. 1 shows that regions with the largest errors correspond to low sensitivity regions for the sensor. This figure also shows that the Matérn kernel model with copula link function provides better accuracy than the GMM model for all the wavelength locations. Wavelength-wise error analysis in Fig. 2 (figure on the left) shows that the Matérn kernel model with link function provides improvement in MSE accuracy especially in the range of 625–700 nm when compared to the Gaussian kernel model.

(4) Color errors (Table 3, upper half) are similar for all the models. The only exception is the case of the F11 illuminant, where the maximum color difference for the GMM shows approximately two units higher CIE2000 error when compared to the kernel models. The color error for the GMM in the case of the D65 illuminant is zero, since $\gamma = 0$, and the matrix W in Eq. (1) was defined using the CIE XYZ sensitivities and the D65 illuminant. Although the color errors are low and similar for all the models, they hide the fact that especially estimated values in the 600–700 nm region show significant differences spectrally (see Figs. 1 and 3).

In conclusion, it can be seen that new model, the Matérn kernel model with the logit link function provides the most accurate results in most cases (similar with the TPS model). For this model, we further analyzed the RMSE, PD, and DE2000 (under D65) errors for samples in the test set and compared those to the results for the TPS (with the logit) and Gaussian kernel model. The results are presented in Fig. 4, where results for the Matérn kernel model are ordered, and the corresponding result for the Gaussian kernel and TPS are also presented. The figures show that in terms of the

Table 3. Color Error Values, Corresponding to DeltaE2000 Errors under D65, A, and F11 Illuminant[a]

Method	ΔE Avg. (D65)	ΔE Max. (D65)	ΔE Avg. (A)	ΔE Max. (A)	ΔE Avg. (F11)	ΔE Max. (F11)
Experiment 1 (Simulated RGB data)						
Link function: none						
GMM	0 (0)	0 (0)	0.49 (0.47)	3.02 (2.53)	0.69 (0.69)	5.32 (4.73)
Gaussian kernel	0.05 (0.04)	0.27 (0.29)	0.42 (0.44)	2.25 (2.09)	0.68 (0.67)	3.13 (3.49)
Matérn kernel	0.00 (0.00)	0.01 (0.01)	0.38 (0.39)	2.47 (2.12)	0.60 (0.58)	3.34 (3.63)
TPS kernel	0.00 (0.00)	0.00 (0.00)	0.38 (0.39)	2.47 (2.12)	0.60 (0.58)	3.34 (3.63)
Link function: square root						
Gaussian kernel	0.14 (0.14)	0.53 (0.64)	0.43 (0.44)	2.31 (2.08)	0.67 (0.66)	3.13 (3.49)
Matérn kernel	0.09 (0.10)	0.58 (0.61)	0.38 (0.39)	2.51 (2.14)	0.60 (0.58)	3.39 (3.58)
TPS kernel	0.09 (0.10)	0.54 (0.57)	0.38 (0.39)	2.50 (2.13)	0.60 (0.58)	3.39 (3.59)
Link function: logit						
Gaussian kernel	0.20 (0.21)	0.95 (1.04)	0.45 (0.48)	2.30 (2.19)	0.68 (0.67)	3.08 (3.47)
Matérn kernel	0.13 (0.14)	0.96 (0.96)	0.39 (0.41)	2.52 (2.24)	0.61 (0.59)	3.36 (3.58)
TPS kernel	0.14 (0.14)	0.92 (0.91)	0.39 (0.41)	2.52 (2.22)	0.61 (0.58)	3.38 (3.59)
Link function: copula						
Gaussian kernel	0.15 (0.16)	0.74 (0.80)	0.43 (0.45)	2.28 (2.11)	0.67 (0.67)	3.10 (3.47)
Matérn kernel	0.10 (0.11)	0.77 (0.78)	0.38 (0.40)	2.50 (2.18)	0.61 (0.58)	3.30 (3.58)
TPS kernel	0.10 (0.10)	0.74 (0.74)	0.38 (0.40)	2.50 (2.17)	0.61 (0.58)	3.36 (3.59)
Experiment 2 (Real RGB data)						
Link function: none						
Gaussian kernel	1.28 (1.24)	8.81 (3.67)	1.3 (1.28)	8.99 (3.89)	1.44 (1.40)	9.18 (4.90)
Matérn kernel	1.23 (1.21)	8.49 (3.64)	1.25 (1.25)	8.88 (3.95)	1.38 (1.36)	8.85 (5.06)
TPS kernel	1.24 (1.21)	8.51 (3.61)	1.26 (1.24)	8.87 (3.91)	1.39 (1.35)	8.86 (4.92)
Link function: square root						
Gaussian kernel	1.28 (1.23)	8.74 (3.63)	1.29 (1.27)	8.9 (3.86)	1.44 (1.38)	9.11 (4.77)
Matérn kernel	1.23 (1.20)	8.5 (3.59)	1.25 (1.23)	8.86 (3.89)	1.38 (1.34)	8.85 (4.91)
TPS kernel	1.23 (1.20)	8.51 (3.59)	1.25 (1.24)	8.86 (3.90)	1.39 (1.34)	8.86 (4.88)
Link function: logit						
Gaussian kernel	1.29 (1.25)	8.75 (3.69)	1.29 (1.28)	8.93 (3.95)	1.43 (1.39)	9.1 (4.89)
Matérn kernel	1.23 (1.21)	8.49 (3.61)	1.24 (1.24)	8.86 (3.93)	1.37 (1.35)	8.85 (4.96)
TPS kernel	1.23 (1.21)	8.5 (3.61)	1.24 (1.24)	8.86 (3.93)	1.38 (1.35)	8.85 (4.94)
Link function: copula						
Gaussian kernel	1.28 (1.24)	8.77 (3.67)	1.29 (1.28)	8.95 (3.93)	1.43 (1.39)	9.13 (4.90)
Matérn kernel	1.23 (1.20)	8.5 (3.60)	1.24 (1.24)	8.87 (3.91)	1.38 (1.35)	8.85 (4.95)
TPS kernel	1.23 (1.21)	8.5 (3.61)	1.24 (1.24)	8.86 (3.92)	1.38 (1.35)	8.85 (4.94)

[a]Number in parenthesis is average error from 200 randomizations to training and test sets.

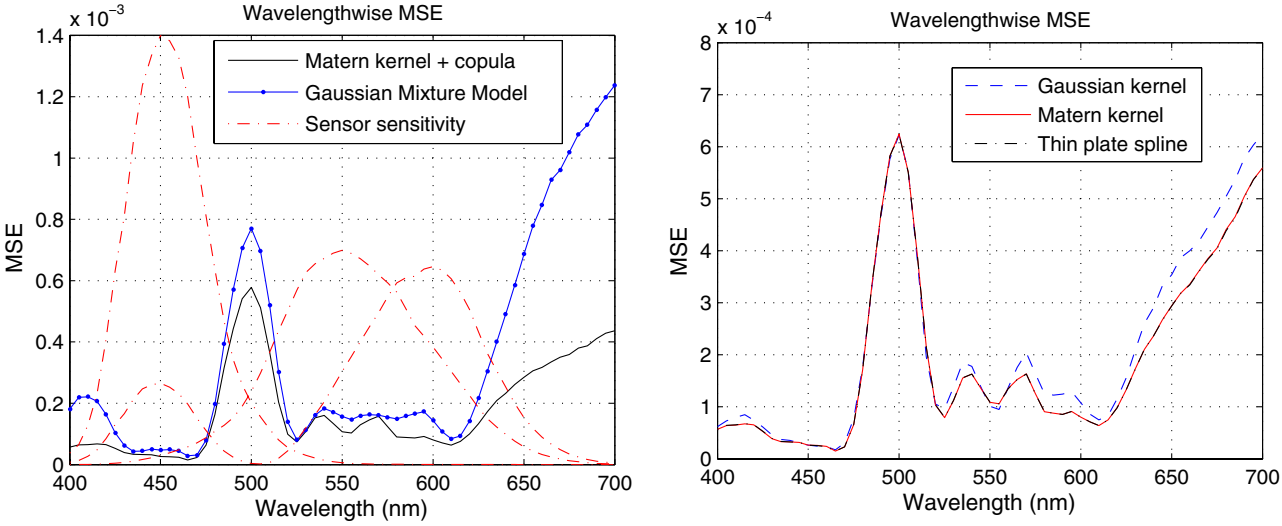

Fig. 1. Left: Location of wavelength-wise errors for the GMM and for a kernel model with a link function (a result corresponding to one data randomization in Experiment 1) with respect to the sensor sensitivity properties. Red curves represent relative sensitivity of CIE $\bar{x},\bar{y},\bar{z}$ 1931 sensor system. Right: Wavelength-wise errors for the Gaussian and Matérn kernel and TPS (Experiment 1).

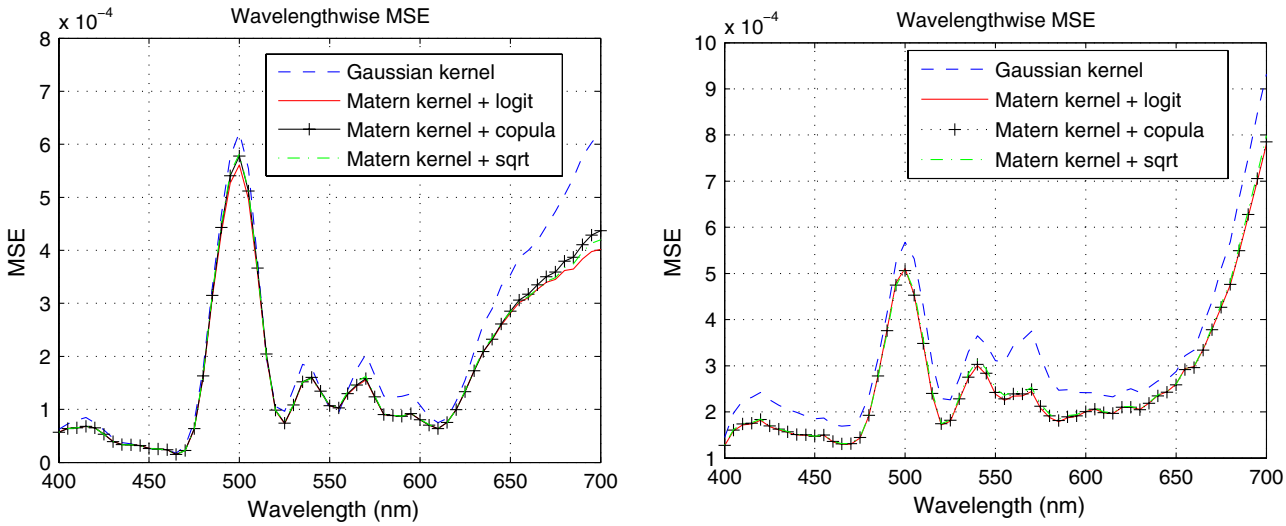

Fig. 2. Left: Wavelength-wise errors for the Gaussian kernel and for the Matérn kernel models with link functions (a result corresponding to one data randomization in Experiment 1). Right: Wavelength-wise errors for the Gaussian kernel and the Matérn kernel models with link functions (a result corresponding to one data randomization in Experiment 2).

RMSE and PD metrics, the Matérn kernel and the TPS model with the logit function provides significant improvements for several samples when compared to the Gaussian kernel, but also decreased performance for some samples. However, the figure shows that the decrease in accuracy (for some samples) is not as dramatic as is the significant increase in accuracy for some samples. The results for the Matérn kernel and for the TPS are similar with only small deviations for some samples.

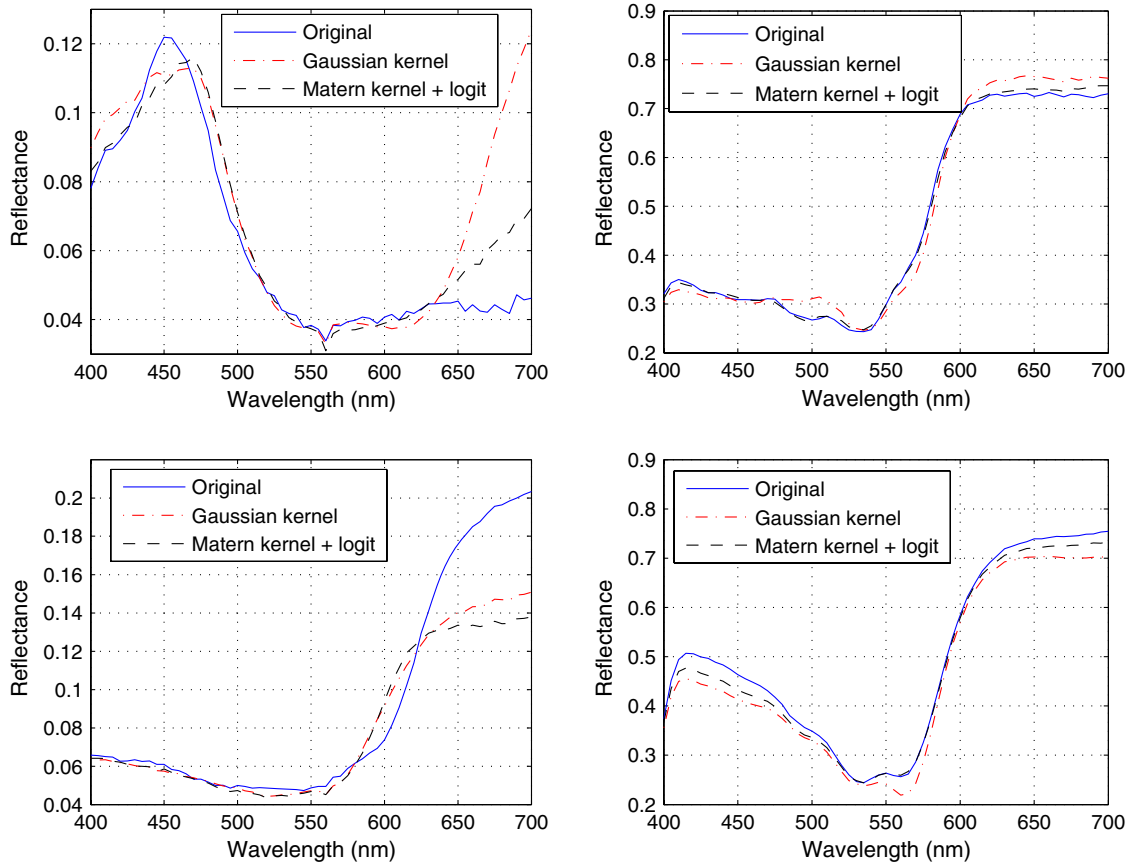

Fig. 3. First row (image on left): Sample corresponding to largest PD difference between the Gaussian kernel model (0.03987) and the Matérn kernel model with link function (0.00831) in Experiment 1 (simulated data). First row (image on right): Sample corresponding to the largest RMSE difference between the Gaussian kernel model (0.0261) and the Matérn kernel model with link function (0.0083) in Experiment 1 (simulated data). Second row (image on left): Sample corresponding to largest PD difference between the Gaussian kernel model (0.01043) and the Matérn kernel model with link function (0.01758) in Experiment 2 (real data). Second row (image on right): Sample corresponding to the largest RMSE difference between the Gaussian kernel model (0.0352) and the Matérn kernel model with link function (0.0179) in Experiment 2 (real data).

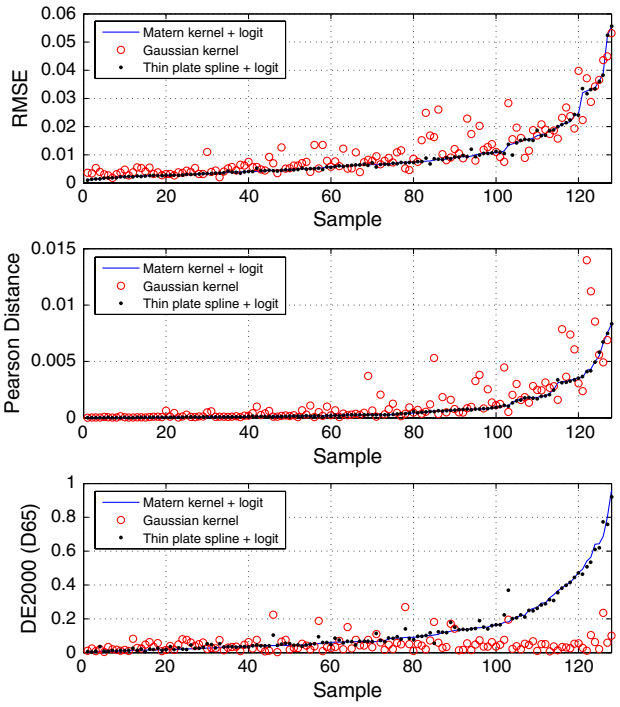

Fig. 4. Error values for individual test samples corresponding to Experiment 1 (simulated data) and tabulated results. Values are ordered according to (Matérn kernel+logit) results [maximum PD for Gaussian (0.0399) is excluded from the image].

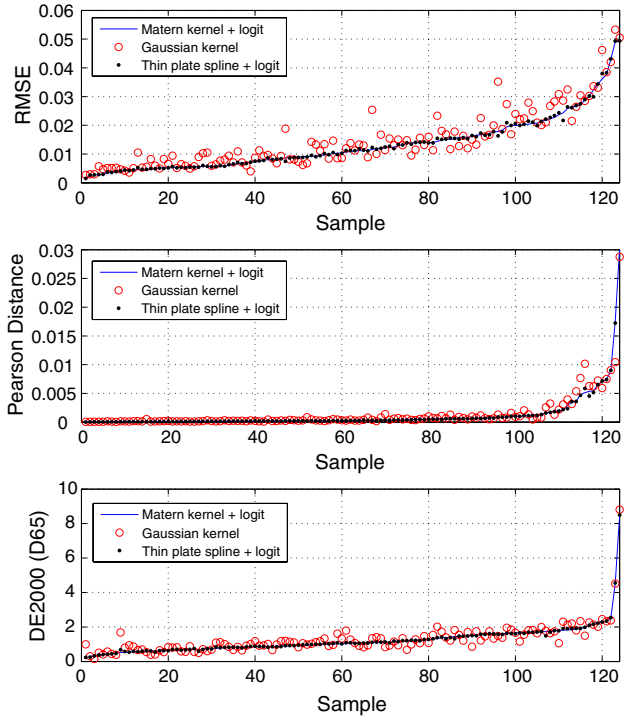

Fig. 5. Error values for individual test samples corresponding to Experiment 2 (real data) and tabulated results. Values are ordered according to (Matérn kernel+logit) results.

The last figure for the color errors show that the kernel models with link functions increase the color error (here D65 illuminant) when compared to the Gaussian kernel model, if the error is calculated under the illuminant that was used in the observation model (also D65) in Eq. (1). However, with this same observation model using the D65 illuminant, the distribution of color errors with the A and F11 illuminants behave similarly with or without link function.

B. Experiment 2: Real RGB

The numerical results for real data in Tables 2 and 3 can be summarized as follows:

(1) Kernel models without link function (Table 2, lower half): When the Matérn kernel is considered, it can be seen that (when compared to the Gaussian kernel): the average RMSE metric decreases 10.8% (averaged result decreases 7.3%), the maximum RMSE metric decreases 7.5% (averaged result increases 1.4%), and the 95th percentile of RMSE decreases 5.7% (averaged result decreases 5.9%). Similarly for the PD metric, the average PD metric decreases 15.2% (averaged result decreases 9.9%), the maximum PD metric decreases 28.1% (averaged results are similar), and the 95th percentile of the PD decreases 17.8% (averaged result decreases 5.6%). All the results for the TPS are similar with the Matérn kernel model (although more differences than in the simulated case).

(2) Kernel models with and without link function (Table 2, lower half): The RMSE results for all the kernels are similar with and without the link functions. There is some improvement when the link functions are used, but the positive effect of the link function is clearly dampened when compared to the simulated case. The performance of the link functions is somewhat mixed for the PD metric (all kernels): the average, maximum, and 95th percentile values are similar or increased in case of one training/test set randomization, but the averaged results over several training/test permutations (in parenthesis) are decreased when compared to case where link functions are not used.

(3) Relative shape of the wavelength-wise error curve in Fig. 2 (figure on the right) shows similar error distribution with the simulated case. This curve shows that the Matérn kernel model with some link function provides improvement in accuracy for all the wavelength locations when compared to the Gaussian kernel model without link function.

(4) Color errors (Table 3, lower half) are again similar for all the kernel models in spite of significant spectral differences (see Fig. 3, second row). However, the behavior of averaged (over 200 data randomizations) color errors deviate significantly from the simulated case. In this real case, the averaged maximum results are significantly lowered.

In same way as in the simulated case, we analyzed the RMSE, PD, and DE2000 (under D65) errors for each sample in the test set when using the TPS and Matérn kernel models with the link function. The results were again compared those for the Gaussian kernel model. The results are presented in Fig. 5, where results for the Matérn kernel model with the logit are ordered, and the corresponding result for the TPS (with the logit) and Gaussian kernel are also presented. The results for the Matérn kernel model and TPS are again similar as in the simulated case. For the RMSE and PD metrics, the Matérn kernel model and the TPS model with the logit function provide significant improvements for several samples when compared to the Gaussian kernel, but also decreased performance for some samples. The distribution of the PD and DE2000 values are somewhat different from the simulated case, since it can be seen that there are one to two samples with high errors.

6. DISCUSSION

The logit link function has been evaluated before in [5] (up to a scalar) and in [6]. Also a closely related link function for the square root has been used for reflectance estimation in a form $\tilde{\mathbf{q}} = \mathbf{a} - \sqrt{\mathbf{q}}$, with optimized $\mathbf{a} \in \mathbb{R}^n$ [3]. The copula model has not been used for reflectance estimation before, but the core element in our model, the beta distribution, has been proposed to be accurate distribution for reflectance in [21]. Nevertheless, it is possible to modify the copula model by using another distribution in place of the beta distribution and then follow the same steps when using cumulative distribution functions and corresponding inverse functions.

We did not evaluate the GMM with the link functions, since this model is based on the linear model in Eq. (1), and any nonlinear transformation for $\mathbf{q}$ would require some linearization procedure. In this case, the use of link functions would be significantly more complicated than in the case of empirical regression.

Here our main interest was in the evaluation of the global models in Eqs. (23) and (29) for the estimation with strictly (conditionally) positive definite kernels. These models allow to construct both regression estimators and interpolators that are independent of the spacing of data in the input domain. The model in Eq. (23) is also independent of dimension of data. Kernel models with the Gaussian [4–6,20,22] and the TPS (Duchon's spline) [4,6,22] kernels have been found to provide accurate reflectance estimation previously and therefore were natural choices for evaluation here. One aspect for all three kernels evaluated here is that the values $\kappa(\mathbf{x}, \mathbf{v})$ depend only on the Euclidean distance between $\mathbf{x}$ and $\mathbf{v}$, and there exists univariate functions f, such that $\kappa(\mathbf{x}, \mathbf{v}) = f(r)$, $f : [0, \infty) \to \mathbb{R}$, with $r = \|\mathbf{x} - \mathbf{v}\|$.

Kernel regression model based on the squared error loss has been found to provide a more accurate reflectance estimation when compared to several other models [4,5,20,22]. Especially, previous research in [22] has suggested that model based on the squared error loss is very competitive against relatively more complex kernel-based regression model with the ϵ-insensitive loss function, i.e., $|y - f(x)|_\epsilon = \max\{0, |y - f(x)| - \epsilon\}$.

We comment on two plug-in modifications for the regression models in Eqs. (23) and (29). First, the noisy observations $\mathbf{x}$ (experiment 2) in regression could be replaced with $\mathbf{z} = W\mathbf{q}$ if W is known. Second, if Eq. (1) need to be valid for the estimate, the kernel regression in Eq. (23) or Eq. (29) can be used only for modeling the Wiener residual. This approach produces a similar estimate as in the case of GMM in Eq. (20) and would be also similar to the approach in [23], where the Wiener residual was modeled with locally weighted linear regression. However, the best approach is the fitting of parametric and nonparametric parts at the same time as in the case of TPS.

7. CONCLUSION

Our results suggest that link functions (square root, logit, copula) and Matérn kernel improve the spectral accuracy for regression-based spectral reflectance estimation from RGB responses. Simulated and real data show similar relative performance for different kernels and link functions and indicate that especially the spectral shape (indicated by Pearson distance) is estimated more accurately via link functions. In our experiments with the Munsell Matte collection, the effect of link functions was more significant in the case of simulated data, partly attributable to limited accuracy of our real RGB responses. In most cases the kernel model with the Matérn kernel and the logit link function gave the best accuracy, although with highly similar accuracy to the model with the Matérn kernel and the copula link function. One useful aspect is that the logit model is independent of the training data and is therefore somewhat easier to calculate than the copula model. For simulated and real data, the combination of the logit link function and the Matérn kernel decreased several spectral errors significantly in most cases when compared to the kernel regression with the Gaussian kernel (with or without the link function). It was also found that the Matérn and TPS kernel produced similar results for simulated and real data. The Matérn kernel, however, is a strictly positive definite kernel and does not require parametric polynomial part in model and is therefore easier to use in computations. A simulated noise-free case indicated that all the evaluated kernel models with some link function provide more accurate estimation than the Gaussian Mixture Model in terms of the RMSE, Pearson distance, and wavelength-wise errors.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for advice and suggestions that lead to a clearer presentation of this work.

REFERENCES

1. H. Haneishi, T. Hasegawa, A. Hosoi, Y. Yokoyama, N. Tsumura, and Y. Miyake, "System design for accurately estimating the reflectance spectra of art paintings," Appl. Opt. **39**, 6621–6632 (2000).
2. R. S. Berns, "Color-accurate image archives using spectral imaging," in *Scientific Examination of Art: Modern Techniques in Conservation and Analysis* (National Academy, 2005), pp. 105–119.
3. F. Imai, R. S. Berns, and D. Tzeng, "A comparative analysis of spectral reflectance estimation in various spaces using a trichromatic camera system," J. Imaging Sci. Technol. **44**, 280–287 (2000).
4. V. Heikkinen, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, T. Jääskeläinen, and S. D. Lee, "Regularized learning framework in the estimation of reflectance spectra from camera responses," J. Opt. Soc. Am. A **24**, 2673–2683 (2007).
5. V. Heikkinen, R. Lenz, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, and T. Jääskeläinen, "Evaluation and unification of some methods for estimating reflectance spectra from RGB images," J. Opt. Soc. Am. A **25**, 2444–2458 (2008).
6. V. Heikkinen, Kernel methods for estimation and classification of data from spectral imaging, Ph.D. Thesis (School of Computing, University of Eastern Finland, 2011).
7. J. Parkkinen, J. Hallikainen, and T. Jääskeläinen, "Characteristic spectra of Munsell colors," J. Opt. Soc. Am. A **6**, 318–322 (1989).
8. University of Eastern Finland Color Group., "Spectral Database," http://uef.fi/spectral.
9. C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed. (Wiley, 1973).
10. W. K. Pratt and C. E. Mancill, "Spectral estimation techniques for the spectral calibration of color image scanner," Appl. Opt. **15**, 73–75 (1976).
11. Y. Murakami, T. Obi, M. Yamaguchi, and N. Ohyama, "Nonlinear estimation of spectral reflectance on Gaussian mixture distribution for color image reproduction," Appl. Opt. **41**, 4840–4847 (2002).
12. N. Shimano, "Recovery of spectral reflectances of objects being imaged without prior knowledge," IEEE Trans. Image Process. **15**, 1848–1856 (2006).

13. M. Yamaguchi, H. Haneishi, and N. Ohyama, "Beyond red-green-blue (RGB): spectrum-based color imaging technology," J. Imaging Sci. Technol. **52**, 10201 (2008).
14. P. Urban, M. R. Rosen, and R. S. Berns, "Spectral image reconstruction using an edge preserving spatio-spectral Wiener estimation," J. Opt. Soc. Am. A **26**, 1865–1875 (2009).
15. F. Cucker and S. Smale, "On the mathematical foundations of learning," Bulletin of the American Mathematical Society **39**, 1–49 (2002).
16. G. Wahba, *Spline Models for Observational Data*, SIAM CBMS-NSF Regional Conference Series in Applied Mathematics (SIAM, 1990), Vol. **59**.
17. C. E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning* (MIT, 2006).
18. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (Springer-Verlag, 2001).
19. J. Hernández-Andrés, J. L. Nieves, E. M. Valero, and J. Romero, "Spectral-daylight recovery by use of only a few sensors," J. Opt. Soc. Am. A **21**, 13–23 (2004).
20. E. M. Valero, Y. Hu, T. Eckhard, J. L. Nieves, J. Romero, M. Schnitzlein, and D. Nowack, "Comparative performance analysis of spectral estimation algorithms and computational optimization of a multispectral imaging system for print inspection," Color Res. Appl. Available online (2012), http://onlinelibrary.wiley.com/doi/10.1002/col.21763/abstract.
21. D. Attewell and B. J. Roland, "The distribution of reflectances within the visual environment," Vision Res. **47**, 548–554 (2007).
22. W.-F. Zhang and D.-Q. Dai, "Spectral reflectance estimation from camera responses by support vector regression and a composite model," J. Opt. Soc. Am. A **25**, 2286–2296 (2008).
23. J. M. Dicarlo and B. A. Wandell, "Spectral estimation theory: beyond linear but before Bayesian," J. Opt. Soc. Am. A **20**, 1261–1270 (2003).