

## Article

# Colorimetric Characterization of Color Imaging System Based on Kernel Partial Least Squares

Siyu Zhao <sup>1</sup>, Lu Liu <sup>1</sup>, Zibing Feng <sup>1</sup>, Ningfang Liao <sup>2</sup>, Qiang Liu <sup>1,3,\*</sup> and Xufen Xie <sup>1,\*</sup> 

<sup>1</sup> School of Information Science and Engineering, Dalian Polytechnic University, Dalian 116034, China; zhaosiyudlpu@126.com (S.Z.); zibingfengdlpu@126.com (Z.F.)

<sup>2</sup> National Key Lab of Colour Science and Engineering, Beijing Institute of Technology, Beijing 100081, China; liaonf@bit.edu.cn

<sup>3</sup> Research Center of Graphic Communication, Printing and Packaging, Wuhan University, Wuhan 430079, China

\* Correspondence: liuqiang@whu.edu.cn (Q.L.); xiexf@dlpu.edu.cn (X.X.)

**Abstract:** Colorimetric characterization is the basis of color information management in color imaging systems. In this paper, we propose a colorimetric characterization method based on kernel partial least squares (KPLS) for color imaging systems. This method takes the kernel function expansion of the three-channel response values (RGB) in the device-dependent space of the imaging system as input feature vectors, and CIE-1931 XYZ as output vectors. We first establish a KPLS color-characterization model for color imaging systems. Then we determine the hyperparameters based on nested cross validation and grid search; a color space transformation model is realized. The proposed model is validated with experiments. The CIELAB, CIELUV and CIEDE2000 color differences are used as evaluation metrics. The results of the nested cross validation test for the ColorChecker SG chart show that the proposed model is superior to the weighted nonlinear regression model and the neural network model. The method proposed in this paper has good prediction accuracy.

**Keywords:** color imaging system; colorimetric characterization; color space conversion; partial least squares



**Citation:** Zhao, S.; Liu, L.; Feng, Z.; Liao, N.; Liu, Q.; Xie, X. Colorimetric Characterization of Color Imaging System Based on Kernel Partial Least Squares. *Sensors* **2023**, *23*, 5706. <https://doi.org/10.3390/s23125706>

Academic Editor: Ludovic Macaire

Received: 2 May 2023

Revised: 5 June 2023

Accepted: 15 June 2023

Published: 19 June 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Color imaging systems have been widely used in industrial manufacturing, medicine, geology exploration, art and other fields. Color characterization [1–3] is an important means to evaluate the color characteristics of color imaging systems; it establishes the transformation relationship between the device-dependent RGB color space and the device-independent color space of the color imaging systems, and realizes the color consistency from color information acquisition to color information output [4]. Colorimetric characterization is very important for image high-fidelity display, colorimetric measurement, color reproduction, color analysis, color management of different devices [5] and color appearance prediction [6], and the requirements for the accuracy and stability of color characterization are increasing for different applications. Color analysis [7–10] plays a very important role in the evaluation of the performance of many technologies including paper-based devices that use colorimetric reactions and that researchers have used with respect to different color spaces such as weighted RGB, greyscale, etc.

According to international standards, color characterization can be generally divided into two categories: spectral response-based and target color-based characterization methods [11–13]. The spectral response-based characterization method is to find out the relationship between the spectral response and the CIE color-matching function under the condition of knowing the three-channel spectral response of the color imaging system, and establish the transformation relationship between the device-dependent RGB color space

and the device-independent CIEXYZ color space [14]. However, measurement for the three-channel spectral response of the imaging system not only requires special experimental equipment and measurement equipment, but also requires professional measurement methods [15]. Monochromators and radiometers are among the measurement equipment used, and experts are required to manage the operation in the laboratory [16]; the operation is difficult for general device users. The advantage is that if the spectral response is known, the tristimulus values of objects under any known spectral conditions of a light source can be predicted, and the estimation methods of camera spectral response characteristics are also constantly explored [17]. The target color-based color-characterization method establishes the mapping relationship between the input and output color spaces based on samples, including interpolation methods, regression models, etc. The method has low requirements for experiments and is widely used in practice. In the early stages of color characterization, interpolation methods were widely used [18]. One of the common color-characterization methods at that time was the three-dimensional lookup table method [19,20]. Based on different regression models, common target color-based color-characterization methods include least squares polynomial regression method [21,22], neural network [23–27], support vector machine [28], etc.; from the perspective of input features of the model, these methods can also incorporate different feature-selection and optimization techniques. RGB cross-polynomial features have been widely used in color measurement, information management and color correction for their high conversion accuracy and low computational cost [29–33]. Essentially, this is a technique that uses the nonlinearity of input features to perform dimensionality increase. In addition to polynomial expansion, other kernel-based increasing dimension techniques also need further exploration in the field of colorimetric characterization. There are many research methods based on the ordinary least squares method with RGB cross-polynomial dimensionality increase. In 2000, Hong et al. proposed a polynomial regression method [34], which uses different combinations of multiple inputs of R, G and B cross-terms. In 2007, Bianco proposed a pattern search optimization algorithm [35] to study the conversion from RGB to CIEXYZ space. It uses pattern search optimization of the least squares method for  $3 \times 3$  RGB to XYZ conversion. Meanwhile, for the problems of medium color spectral reflectance reconstruction and training sample validity in color measurement, various improved algorithms of the ordinary least squares method have also been applied in different studies. In terms of color spectral reconstruction, in 2010, Shen et al. proposed that the partial least squares regression (PLS) method can also be used to build a regression model based on the correlation between response values and spectral reflectance [36]. In 2013, Heikkinen et al. proposed a kernel ridge regression (KRR) method for spectral reflectance [37]. Essentially, KRR is a method that nonlinearly transforms low-dimensional camera responses to high-dimensional feature space; and it performs regularized least squares regression on the reflectance data in the feature space. In 2019, Xiao et al. proposed a new method for spectral reflectance reconstruction based on kernel partial least squares regression (KPLS) [38]. The problem of colorimetric characterization is different from that of color spectral reflectance reconstruction. Although the spectral reflectance of the object surface is a high-resolution expression of the object color, the spectral reflectance of the object surface needs to be combined with the light source to show the color information. And the spectral reflectance of the object surface cannot be directly used for colorimetric characterization and color space conversion; meanwhile, this method uses RMSE for parameter tuning without considering visual color difference, and cannot accurately optimize model parameters for visual color difference. In terms of training sample validity, in 2018, Amiri et al. proposed a weighted nonlinear regression method (WT-NONLIN), which enables commercial digital RGB cameras to be used for spectral and colorimetric color reproduction [39].

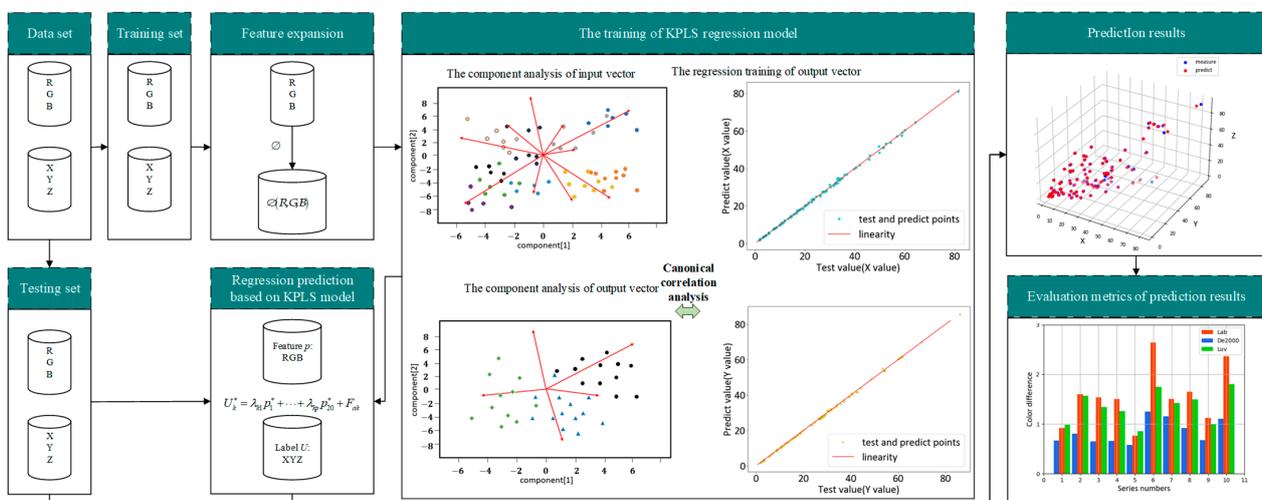
The accuracy of color space conversion is constrained by the regression model, the training samples and the feature extraction. The effectiveness of feature extraction is the key to improving the accuracy. We propose a new method for colorimetric characterization of color imaging systems based on KPLS. The method establishes the mapping relationship

between the three-channel response kernel function of the color imaging system and the CIE1931 tristimulus values. Through canonical correlation analysis, the direction vector with the strongest correlation between input and output is determined. Based on different direction vectors, a multivariate nonlinear regression model is established. The method considers the relevance between the input vector and the output vector; this method can promote predictive ability when a feature vector is selected. It can effectively solve the multicollinearity problem caused by kernel function expansion, and eliminate irrelevant input variables. It makes the model more concise and stable. It is proved that this method is an effective colorimetric characterization method by using two nested cross-validation.

## 2. Theory and Method

### 2.1. Colorimetric Characterization of Color Imaging System Based on KPLS

The basic idea of the colorimetric characterization method for a color imaging system is to establish the mapping relationship between the device-dependent space RGB and the device-independent space XYZ, so as to realize the transformation model from RGB to XYZ. In order to obtain the colorimetric characterization mapping from RGB to XYZ, it is usually necessary to use a standard color chart as a measurement sample, and measure its three-channel response values and tristimulus values at the same time, so as to obtain a series of calibration data pairs of XYZ and RGB. In this study, the KPLS method is used for modeling research. Firstly, the RGB dataset is expanded via the kernel function, then the dataset is divided. Next, the PLS regression model is trained. Finally, using color difference as evaluation indicators, the trained model is used to predict on the test set and evaluate the performance of the algorithm. The colorimetric characterization method based on KPLS proposed in this paper is shown in Figure 1.



**Figure 1.** Colorimetric characterization method based on KPLS.

### 2.2. Kernel Expansion of the RGB Color Value

A series of data pairs of  $(R, G, B)$  and  $(X, Y, Z)$  are obtained by simultaneously measuring the three-channel response values and the three-stimulus values. The input matrix  $\mathbf{N} = (n_{ij})_{m \times 3}$  and the output matrix  $\mathbf{U} = (u_{ij})_{m \times 3}$  are composed of  $m$  groups of data, where  $\mathbf{n}_i = (R_i, G_i, B_i), i = 1 \cdots m, \mathbf{u}_i = (X_i, Y_i, Z_i), i = 1 \cdots m$ . In order to deal with the nonlinear characteristics between variables, we use a nonlinear mapping function  $\varnothing$  to map the input matrix to the input matrix  $\mathbf{P}$  in the high-dimensional feature space. That is,  $(n_{ij}) \rightarrow \varnothing(n_{ij})$ . The matrix after dimensionality reduction is shown in Equation (1).

$$\mathbf{P} = [\varnothing(\mathbf{n}_1), \varnothing(\mathbf{n}_2), \cdots, \varnothing(\mathbf{n}_m)]^T \quad (1)$$

where  $\varnothing()$  denotes the basis function.

Since the dimension of  $\varnothing$  can be arbitrarily large or even infinite, we usually define the following kernel matrix  $\mathbf{K}$  to avoid explicitly using  $\varnothing$

$$\mathbf{K} = \mathbf{P}\mathbf{P}^T \quad (2)$$

where the element  $k_{i,j}$  of the  $i$ th row and  $j$ th column of  $\mathbf{K}$  is

$$k_{i,j} = \langle \varnothing(\mathbf{n}_i)^T, \varnothing(\mathbf{n}_j) \rangle = f_{\text{ker}}(\mathbf{n}_i, \mathbf{n}_j) \quad (3)$$

where  $\langle \rangle$  denotes the inner product operation, and  $f_{\text{ker}}$  is the kernel function. In this paper, we use three kernel functions, which are polynomial kernel, root polynomial kernel and Gaussian kernel, to perform dimensionality enhancement; these kernel functions are shown in Equations (4)–(6)

$$f_{\text{ker-P}}(\mathbf{n}_i^T, \mathbf{n}_j) = \sum_{s=0}^d \frac{d!}{s!(d-s)!} \left( (\mathbf{n}_i^T, \mathbf{n}_j)^s + s \sum_{t=1}^{s-1} \frac{(\mathbf{n}_i^T, \mathbf{n}_j)^t}{t!} + 1 \right) \quad (4)$$

$$f_{\text{ker-R}}(\mathbf{n}_i^T, \mathbf{n}_j) = \sum_{s=0}^d \frac{d!}{s!(d-s)!} \left( (\mathbf{n}_i^T, \mathbf{n}_j)^s + s \sum_{t=1}^{s-1} \frac{(\mathbf{n}_i^T, \mathbf{n}_j)^t}{t!} + 1 \right)^{1/s} \quad (5)$$

$$f_{\text{ker-G}}(\mathbf{n}_i^T, \mathbf{n}_j) = \frac{\exp(-\|\mathbf{n}_i - \mathbf{n}_j\|^2)}{2\sigma^2}, i, j \in m \quad (6)$$

### 2.3. Color Space Conversion Based on KPLS

To enhance the accuracy of the conversion model between the source color space (RGB) and the target color space (XYZ), we use a kernel function to expand the feature vector for regression. Taking into account the correlation among the expanded feature vectors, we apply the PLS method. PLS regression is related to principal component regression. However, instead of finding a hyperplane of maximum variance between the response and independent variables, it finds a linear regression model by projecting both the predicted and observed variables to a new space. First, the data are normalized.

$$\begin{cases} \mathbf{E}_0 = \left( (n_{ij} - \bar{n}_j) / S_{n_j} \right)_{m \times m} \\ \mathbf{F}_0 = \left( (u_{ij} - \bar{u}_j) / S_{u_j} \right)_{m \times 3} \end{cases} \quad (7)$$

$\mathbf{P}$  is normalized to  $\mathbf{E}_0$ , and  $\mathbf{U}$  is normalized to  $\mathbf{F}_0$ . To consider the correlation between the input and output vectors, we decomposed them and modeled the two direction vectors with the highest correlation [40]. The first component  $\mathbf{u}_1$  is extracted from the independent variable matrix  $\mathbf{E}_0$ , and the first component  $\mathbf{t}_1$  of the dependent variable matrix  $\mathbf{F}_0$  is extracted. Namely,  $\mathbf{t}_1 = \mathbf{E}_0 \mathbf{w}_1$ , where  $\mathbf{w}_1$  is the direction vector. Because  $\mathbf{E}_0$  is the normalized matrix,  $\mathbf{w}_1$  is the unit vector. It denotes the direction of the first axis in decomposition of  $\mathbf{E}_0$ . That is,  $\mathbf{w}_1^T \mathbf{w}_1 = 1$ . Similarly,  $\mathbf{u}_1 = \mathbf{F}_0 \mathbf{c}_1$ . The objective function can be described as in Equation (8)

$$\max_{\mathbf{w}_1, \mathbf{c}_1} \langle \mathbf{E}_0 \mathbf{w}_1, \mathbf{F}_0 \mathbf{c}_1 \rangle, s.t. \begin{cases} \mathbf{w}_1^T \mathbf{w}_1 = 1 \\ \mathbf{c}_1^T \mathbf{c}_1 = 1 \end{cases} \quad (8)$$

Using the Lagrange multiplier method to solve, we obtain  $\mathbf{w}_1$  and  $\mathbf{c}_1$ . Then we can obtain the components  $\mathbf{t}_1 = \mathbf{E}_0 \mathbf{w}_1$  and  $\mathbf{u}_1 = \mathbf{F}_0 \mathbf{c}_1$ . Then, we find the regression equations of  $\mathbf{E}_0$  and  $\mathbf{F}_0$  with respect to  $\mathbf{t}_1$ , respectively.

$$\mathbf{E}_0 = \mathbf{t}_1 \mathbf{h}_1^T + \mathbf{E}_1 \quad (9)$$

$$\mathbf{F}_0 = \mathbf{t}_1 \mathbf{r}_1^T + \mathbf{F}_1 \quad (10)$$

where  $\mathbf{E}_1$  and  $\mathbf{F}_1$  are the residual matrices of two regression equations, and the regression coefficient vectors are

$$\mathbf{h}_1 = \frac{\mathbf{E}_0^T \mathbf{t}_1}{\|\mathbf{t}_1\|^2} \quad (11)$$

$$\mathbf{r}_1 = \frac{\mathbf{F}_0^T \mathbf{t}_1}{\|\mathbf{t}_1\|^2} \quad (12)$$

To meet the accuracy requirement, we continued to search for the second pair of direction vectors with the highest correlation. The residual matrices  $\mathbf{E}_1$  and  $\mathbf{F}_1$  replace  $\mathbf{E}_0$  and  $\mathbf{F}_0$ . We find the second directions of the axis  $\mathbf{w}_2, \mathbf{c}_2$  from  $\mathbf{E}_1$  and  $\mathbf{F}_1$ , and then we obtained the second components  $\mathbf{t}_2, \mathbf{u}_2$ . Continuing this calculation, if the mathematical rank of  $\mathbf{P}$  is  $a$ , then Equations (13) and (14) can be eventually obtained

$$\mathbf{E}_0 = \mathbf{h}_a^T \mathbf{t}_1 + \mathbf{h}_a^T \mathbf{t}_2 + \cdots + \mathbf{h}_a^T \mathbf{t}_a, \quad (13)$$

$$\mathbf{F}_0 = \mathbf{r}_a^T \mathbf{t}_1 + \mathbf{r}_a^T \mathbf{t}_2 + \cdots + \mathbf{r}_a^T \mathbf{t}_a + \mathbf{F}, \quad (14)$$

where  $\mathbf{h}_1, \mathbf{h}_2 \cdots \mathbf{h}_a, \mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_a$  are the regression coefficient.

$\mathbf{t}_1, \mathbf{t}_2 \cdots \mathbf{t}_a$  can be expressed as a linear combination of the vectors of the standardized vector. Based on Equation (14), we can obtain the regression form as follows

$$\mathbf{F}_0 = \mathbf{r}_1 \mathbf{E}_0 \mathbf{w}_1^* + \cdots + \mathbf{r}_a \mathbf{E}_0 \mathbf{w}_a^* + \mathbf{F} \quad (15)$$

where  $\mathbf{w}_a^* = \prod_{j=1}^{a-1} (I - \mathbf{w}_j \mathbf{h}_j) \mathbf{w}_a$ ,  $I$  is the unit matrix. Then we can obtain

$$\mathbf{F}^* = \sum_j \lambda_j^* \mathbf{x}_j^* + \mathbf{F} \quad (16)$$

where  $\lambda_j^* = \sum_a \mathbf{r}_a \mathbf{w}_a^*$ . Finally, following the standardized inverse process, the regression equation of  $\mathbf{F}^*$  is reduced to the regression equation of  $\mathbf{P}$  to  $\mathbf{U}$ .

The color space conversion algorithm based on KPLS is shown in Figure 2.

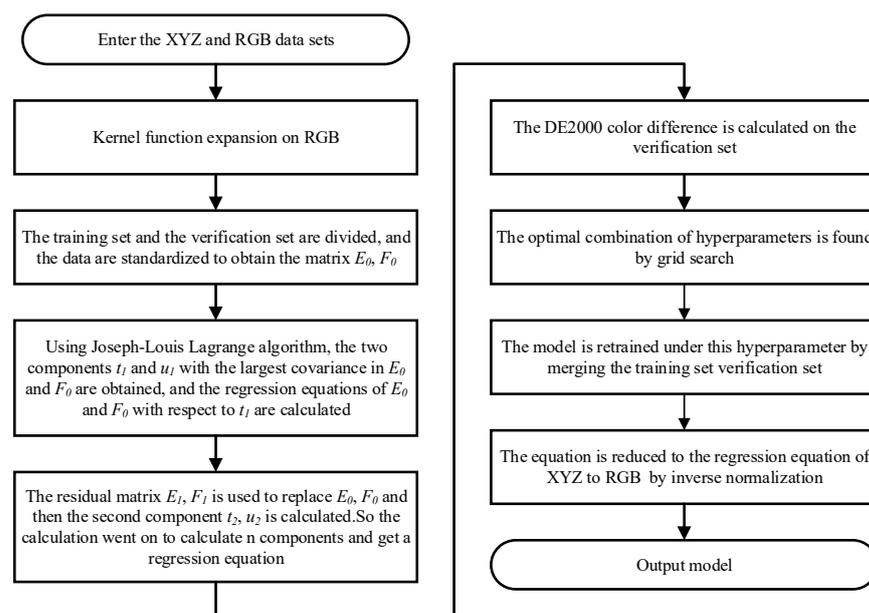


Figure 2. Color space conversion algorithm flow based on KPLS.

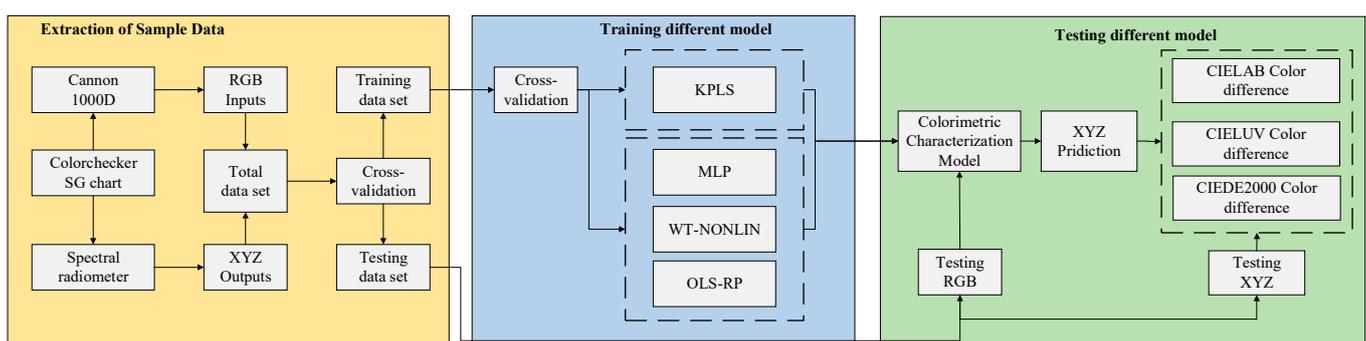
## 2.4. Evaluation Metrics

We chose three different color difference formulas according to different application domains to evaluate our model. CIEDE2000 is a widely used color difference formula that considers the needs of different fields [41]. CIELAB color space can better describe the psychological effects of object colors, and it is suitable for materials such as dyes, pigments and inks when the color difference is larger than the visual recognition threshold but smaller than the color difference between adjacent colors in the Munsell system [42,43]. CIELUV is also a relatively perceptually uniform color space, and it is suitable for applications of spectral colors and color vision. The CIELUV color difference formula performs better than other methods when predicting the color difference of illumination stimuli, especially when using a black background [44–46]. We used CIELAB, CIELUV and CIEDE2000 color differences to evaluate the generalization ability and prediction accuracy of the KPLS model [47–49].

## 3. Experiment

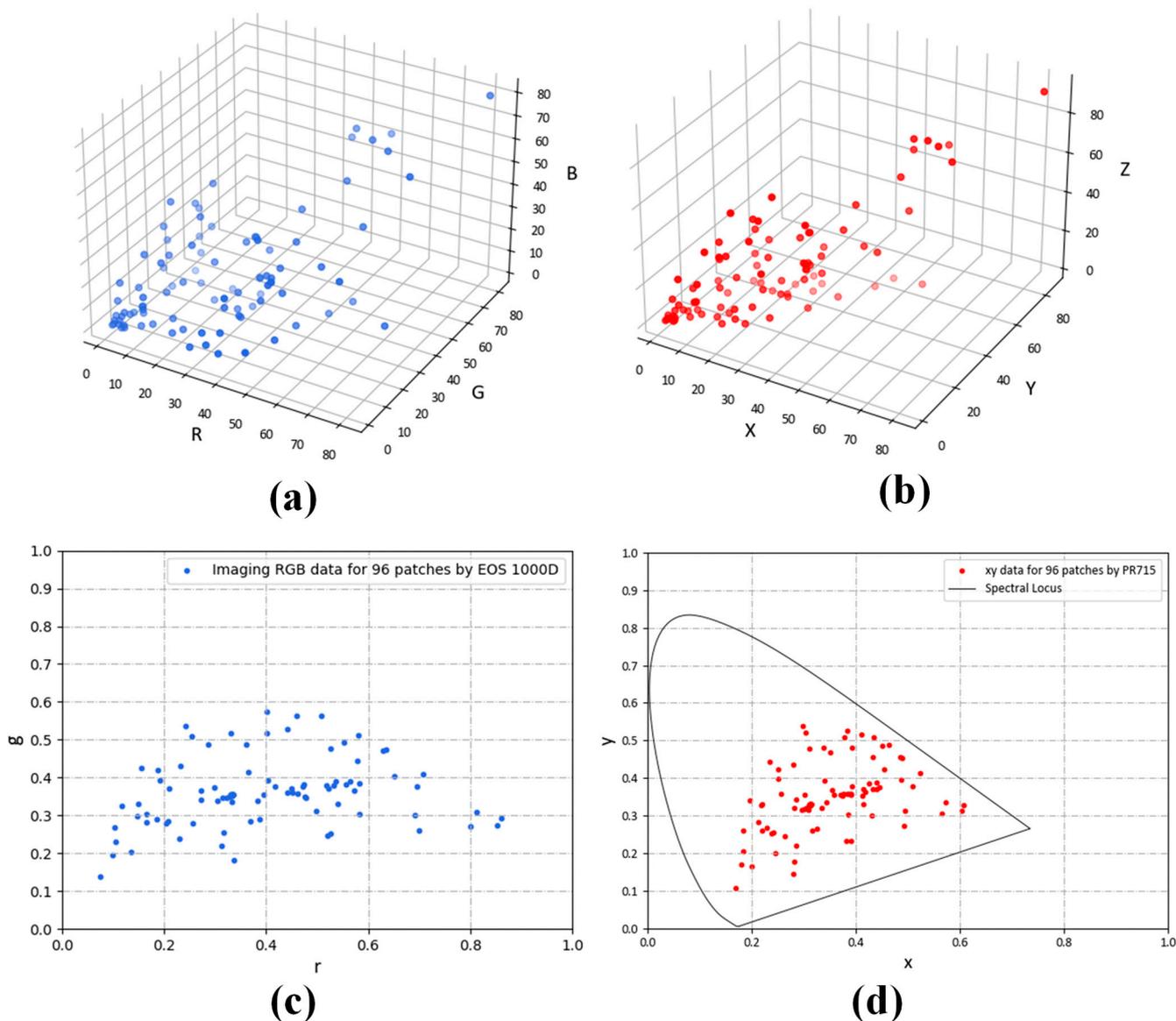
### 3.1. Experimental Scheme

To validate the proposed model, we first constructed a dataset using the RGB and XYZ values of the color samples. To verify the colorimetric characterization accuracy of the model proposed in this paper, we further divided the dataset into training and testing sets. We trained different colorimetric characterization models based on the training set, and evaluated these models based on the testing set. Finally, we compared the colorimetric characterization effects of different models. The experimental scheme is shown in Figure 3. The color samples used in the experiment are selected from the internationally common ColorChecker SG color chart (X-Rite, Grand Rapids, MN, USA); a D65 light booth (Datacolor, Lawrenceville, GA, USA) was used as the light source. A PR715 spectroradiometer (Photo Research, Syracuse, NY, USA) was used to measure the tristimulus values (CIE XYZ) of 96 non-neutral color patches in the SG color chart. A Canon EOS 1000D was used in manual operation mode. Its resolution was  $3888 \times 2592$  ( $22.2 \text{ mm} \times 14.8 \text{ mm}$ ), the ISO was 800 and the F number was 10. When obtaining the color block image, RAW format image was used to capture RGB values. A total of 96 pairs of data were obtained in the experiment, and we divided the 96-group dataset into training and testing sets in accordance with ten-fold cross-validation.



**Figure 3.** Experimental Scheme.

The measured sample dataset is shown in Figure 4. Figure 4a,b show the sample distribution, and Figure 4c,d show the chromaticity distribution.



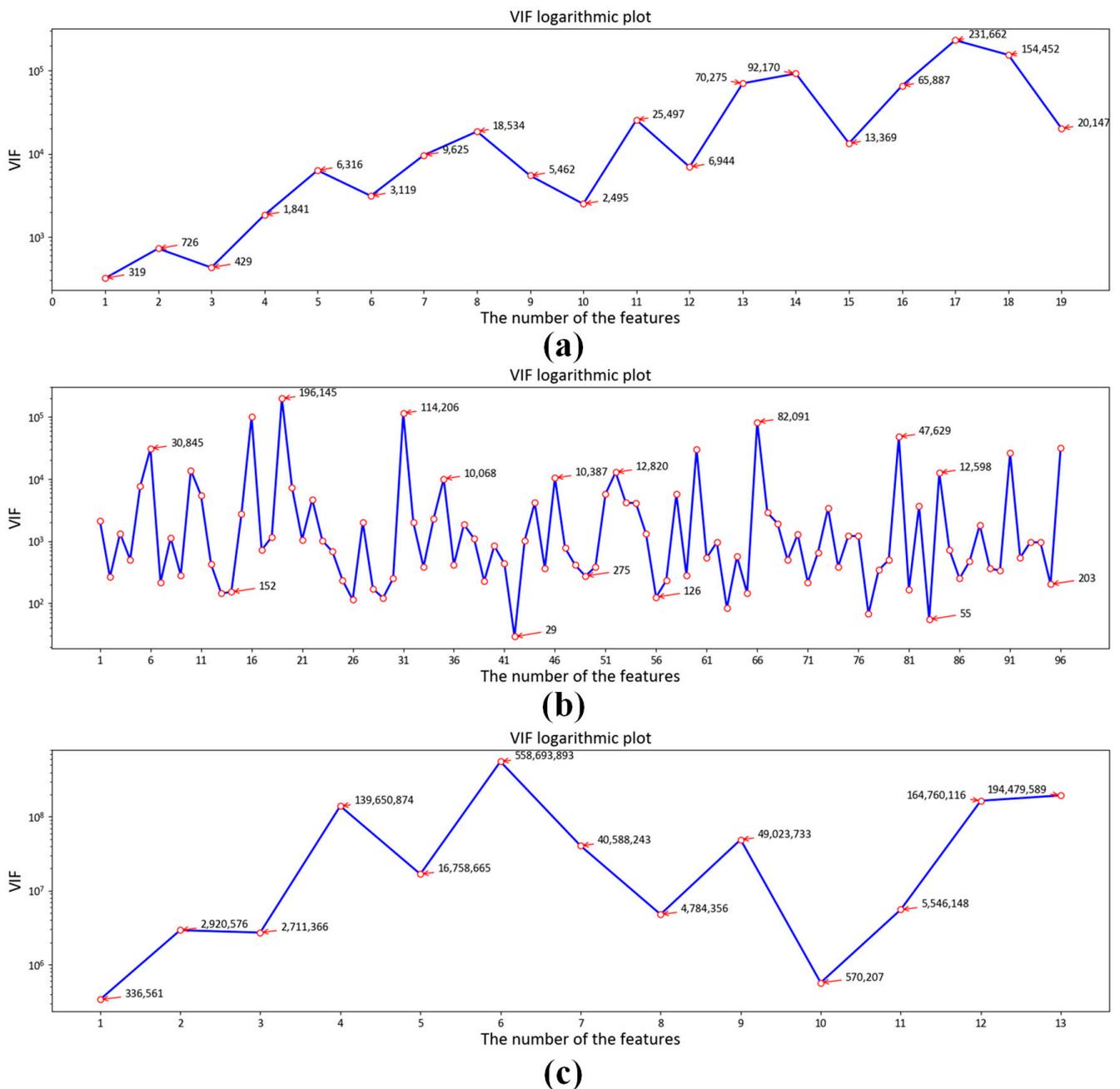
**Figure 4.** Color and colorimetric distributions of 96 patches collected in our experiment: (a) RGB distribution of 96 datasets; (b) CIE 1931XYZ tristimulus value distribution; (c) device-dependent RGB color space chromaticity distribution; (d) CIE1931XYZ chromaticity distribution.

### 3.2. Correlation Analysis of Input and Output Vectors

The input vector is the kernel function expansion of RGB. Therefore, there is multi-collinearity among input feature variables. The Variance Inflation Factor (VIF) represents the magnification of model parameter estimation variances between non-collinearity and multi-collinearity. It can be calculated with Equation (17)

$$V_j = (1 - R_j^2)^{-1} \tag{17}$$

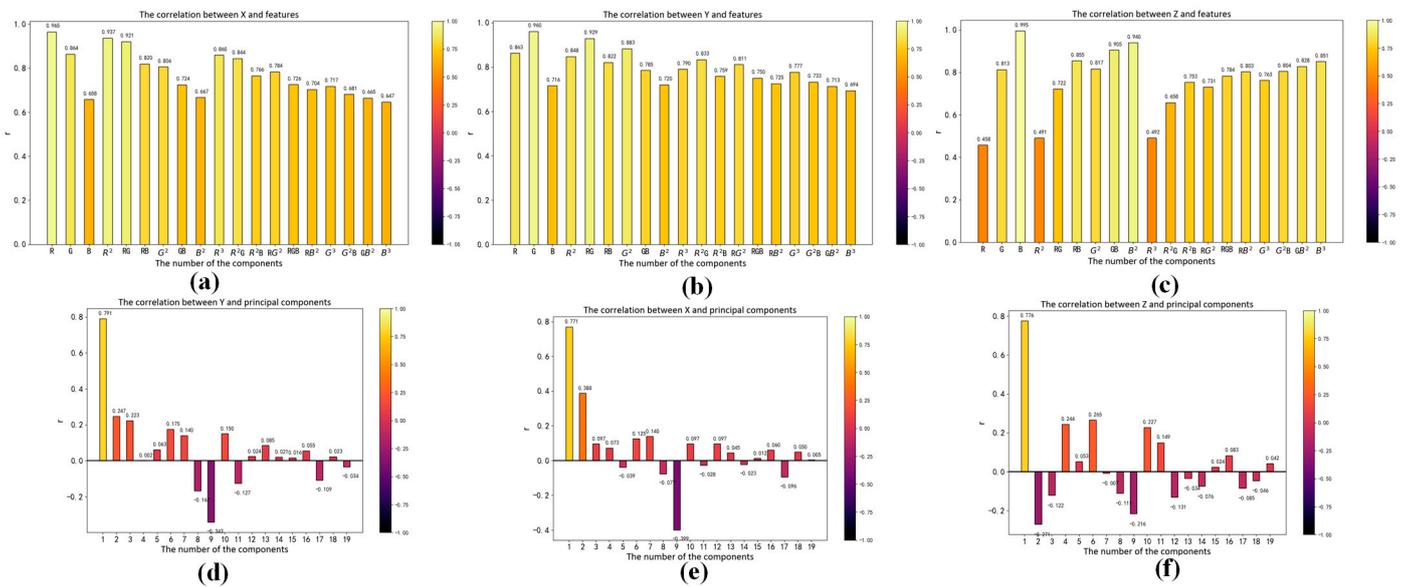
where  $V_j$  represents VIF corresponding to  $j$ th variable;  $R_j^2$  is the complex measuring coefficient of regression between dependent variables  $x_j$  and other independent variables  $x_i$ . The VIF of the input feature variables is shown in Figure 5. If it is more than 100 times larger than that of non-collinearity, the confidence interval is too wide, and the significance test is not credible. Those results reflect the severe collinearity of the R, G and B kernel function expansion.



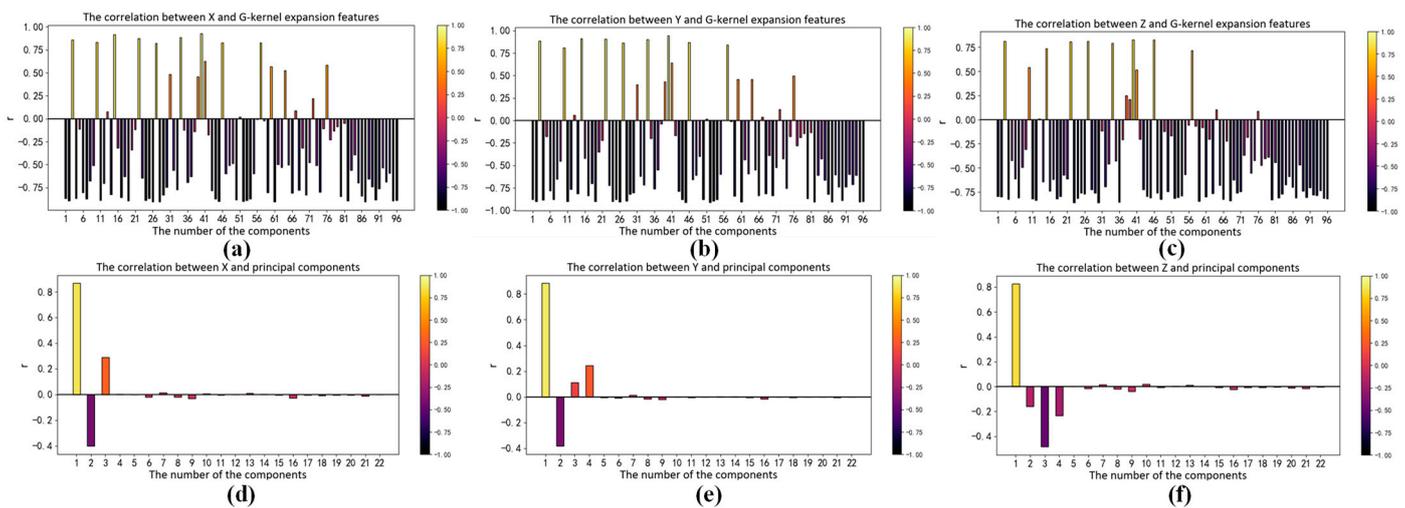
**Figure 5.** VIF of input feature: (a) Based on polynomial kernel function expansion; (b) Based on Gaussian kernel function expansion; (c) Based on root polynomial kernel function expansion.

The correlation between input and output vectors directly affects the accuracy of the regression model; therefore, we analyze the RGB kernel function expansion and its correlation with X, Y and Z. The absolute value of the correlation coefficient reflects the degree of correlation. When the absolute values of the correlation coefficient are greater than 0.6, that means high correlation between the two vectors. The Pearson correlation coefficients between each term and X, Y and Z are shown in Figures 6, 7 and 8a–c, respectively. We can see that the input vectors are highly correlated with X, Y and Z, respectively. Moreover, we applied principal component analysis (PCA) to reduce the multicollinearity among terms of input vectors. The VIF of each component after PCA is approximately 1. The Pearson correlation coefficients between each component after PCA and X, Y and Z are shown in Figures 6, 7 and 8d–f, respectively. There is only one component whose absolute value of the correlation coefficient is greater than 0.6. It suggests that most principal components have

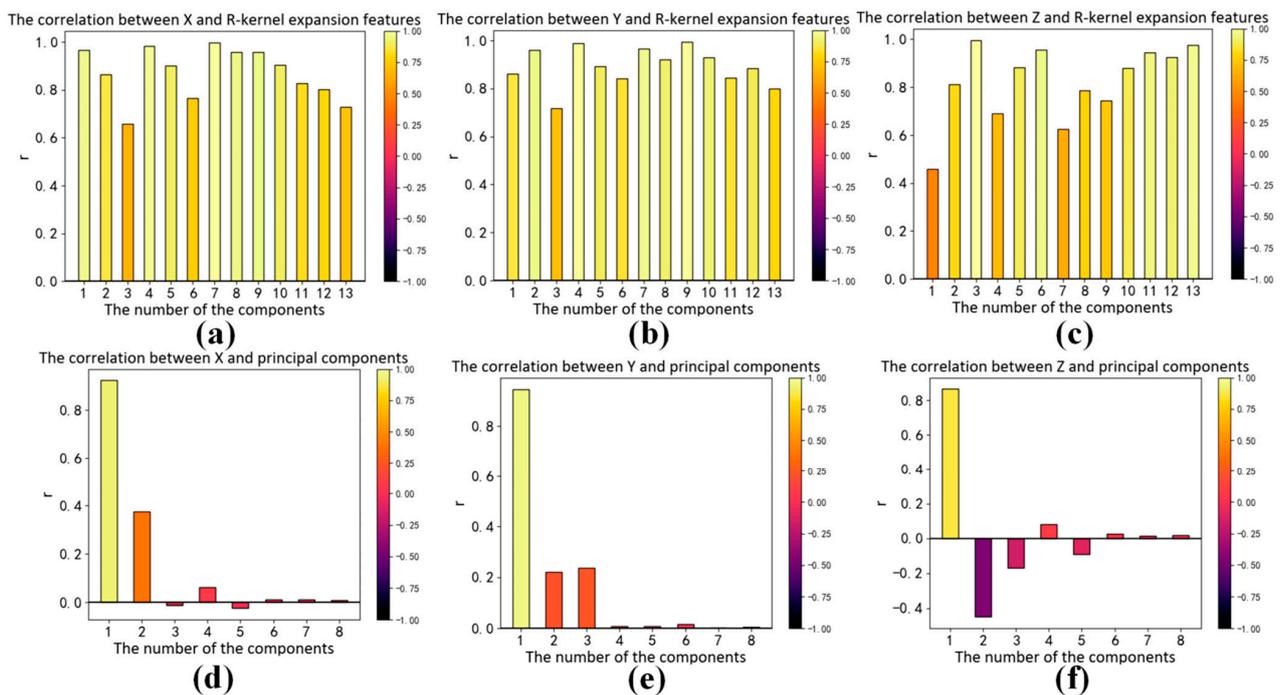
a low correlation with X, Y and Z. As shown in Figures 5–8, RGB kernel function expansion has a high degree of information overlap and low information validity. It represents high multicollinearity, which affects the accuracy of parameter estimation and the robustness of the model. Although PCA can eliminate collinearity, it also reduces the correlation with the dependent variable.



**Figure 6.** The correlation between XYZ and the polynomial kernel function expansion of RGB: (a) Correlation between X and features; (b) Correlation between Y and features; (c) Correlation between Z and features; (d) Correlation between X and features after principal component analysis; (e) Correlation between Y and features after principal component analysis; (f) Correlation between Z and features after principal component analysis.



**Figure 7.** The correlation between XYZ and the Gaussian kernel function expansion of RGB: (a) Correlation between X and features; (b) Correlation between Y and features; (c) Correlation between Z and features; (d) Correlation between X and features after principal component analysis; (e) Correlation between Y and features after principal component analysis; (f) Correlation between Z and features after principal component analysis.



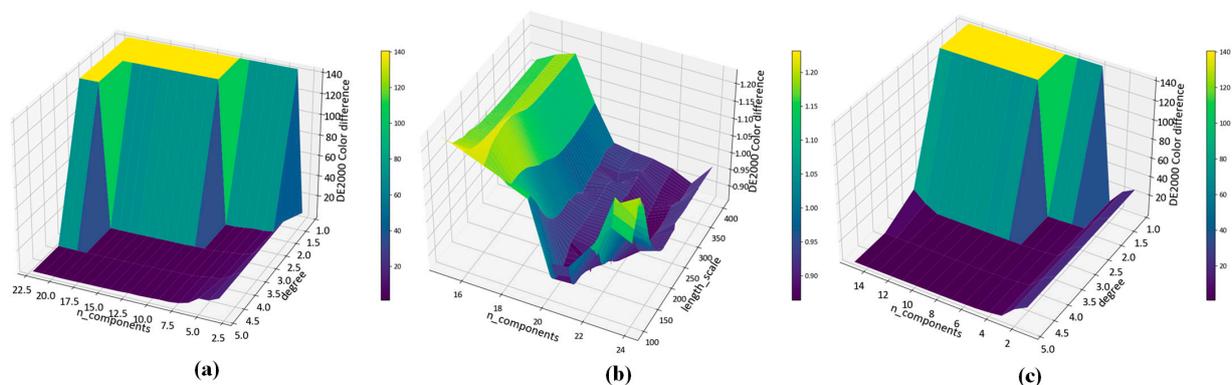
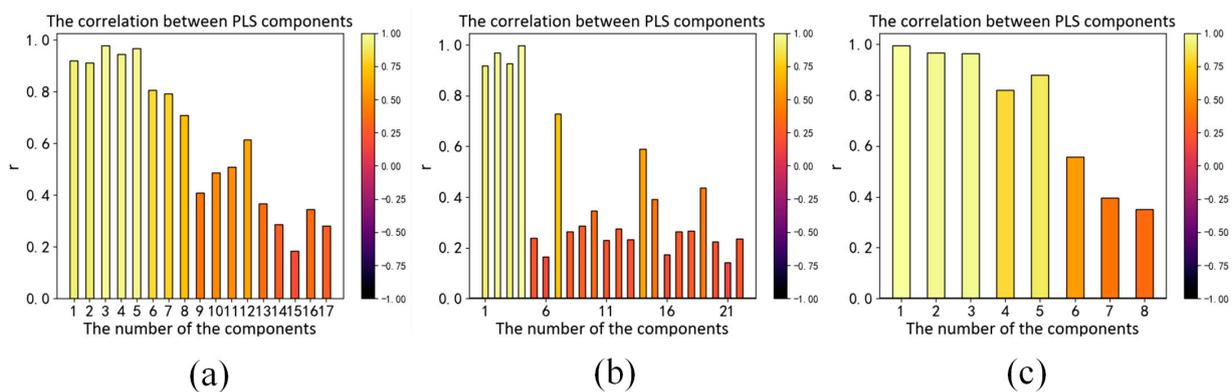
**Figure 8.** The correlation between XYZ and the root polynomial function expansion of RGB: (a) Correlation between X and features; (b) Correlation between Y and features; (c) Correlation between Z and features; (d) Correlation between X and features after principal component analysis; (e) Correlation between Y and features after principal component analysis; (f) Correlation between Z and features after principal component analysis.

### 3.3. Hyperparameter Selection

In this paper, we perform kernel expansion on the input features of the KPLS model using the polynomial kernel function, the Gaussian kernel function and the root polynomial kernel function. We optimize and evaluate their hyperparameters. The hyperparameters include the order of the polynomial,  $\sigma$  which represents the width parameter of the radial basis function and the numbers of components of the PLS model. Unlike the traditional k-fold cross-validation method, we use a nested cross-validation method to select the optimal hyperparameters and evaluate the performance of the model at the same time. We perform two-layer cross-validation. First, we split the training set and test set. Then we split each training set into training set and validation set. The validation set is used for model hyperparameter selection, and the test set is used for accuracy evaluation. The 10-fold cross-validation procedure for model hyperparameter optimization is nested inside the 10-fold cross-validation procedure for accuracy evaluation. The inner loop cross-validation applies grid search to find the optimal hyperparameters within the preset parameter range, so as to minimize the average CIEDE2000 color difference on the validation set; then we use this set of optimal hyperparameters to evaluate the model performance on the test set in the outer loop. Finally, we obtain the CIELAB, CIELUV and CIEDE2000 color differences on the test set. Table 1 shows the optimal hyperparameter combinations selected via the KPLS model based on polynomial kernel expansion, Gaussian kernel and root polynomial kernel expansion in ten training sets. Figure 9 shows the change in CIEDE2000 color difference on the test set under different hyperparameter combinations on the first training set. Figure 10 shows the correlation coefficients between the dependent variable and independent variable obtained via the KPLS model based on polynomial kernel expansion, Gaussian kernel expansion and root polynomial kernel function expansion, respectively. It can be seen that the KPLS model effectively eliminates the problem of multicollinearity between independent variables.

**Table 1.** The Hyperparameter optimization results of KPLS model in cross-validation.

The Series of Fold	P-Kernel			$\sigma$	G-Kernel		RP-Kernel		
	Order	Components	DE2000		Components	DE2000	Order	Components	DE2000
1	3	17	0.862	176	21	0.863	3	8	0.752
2	3	18	0.895	285	22	0.858	3	8	0.756
3	3	17	0.897	196	21	0.880	3	8	0.742
4	3	18	0.947	400	20	0.940	5	11	0.759
5	3	17	0.941	267	22	0.896	5	12	0.764
6	3	17	0.794	236	23	0.882	5	8	0.769
7	4	18	0.825	230	22	0.773	3	8	0.670
8	3	18	0.838	226	22	0.749	3	9	0.699
9	3	18	0.898	225	22	0.861	5	11	0.720
10	4	17	0.840	245	23	0.738	5	8	0.707

**Figure 9.** Color difference of CIEDE2000 based on KPLS model under different hyperparameter combinations: (a) Based on polynomial kernel function expansion; (b) Based on Gaussian kernel function expansion; (c) Based on root polynomial kernel function expansion.**Figure 10.** The correlation coefficient between the components of the KPLS model: (a) Based on polynomial kernel function expansion; (b) Based on Gaussian kernel function expansion; (c) Based on root polynomial kernel function expansion.

### 3.4. Experimental Results of This Paper

To assess the performance of the KPLS model in colorimetric characterization, we use three different color difference formulas—CIEDE2000, CIELAB and CIELUV—to calculate the average color difference on the test set in the outer loop. We collect the test set data from each fold of the cross-validation. We obtain a regression analysis of 96 pairs of actual and predicted data. Figures 11–13 show the regression results of the KPLS model based on polynomial kernel expansion, Gaussian kernel expansion and root polynomial kernel

expansion on the test set in the 10-fold cross-validation. Table 2 gives the average color difference of the model prediction on the test set of each fold.

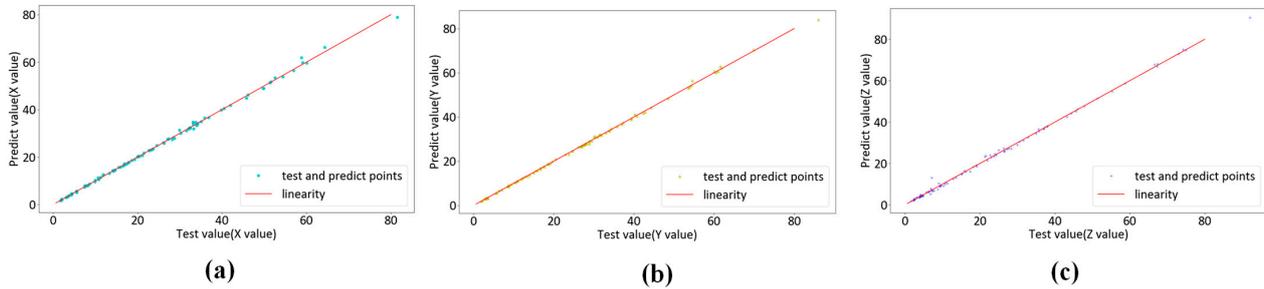


Figure 11. KPLS regression based on polynomial kernel expansion: (a) X test data and forecast data pairs; (b) Y test data and forecast data pairs; (c) Z test data and forecast data pairs.

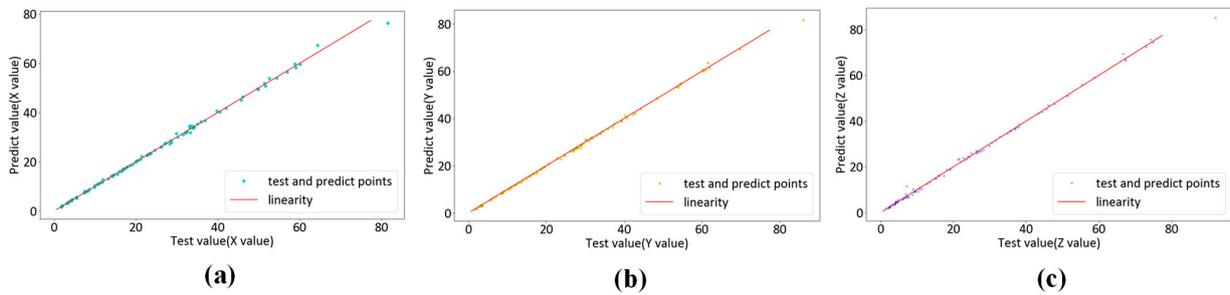


Figure 12. KPLS regression based on Gaussian kernel expansion: (a) X test data and forecast data pairs; (b) Y test data and forecast data pairs; (c) Z test data and forecast data pairs.

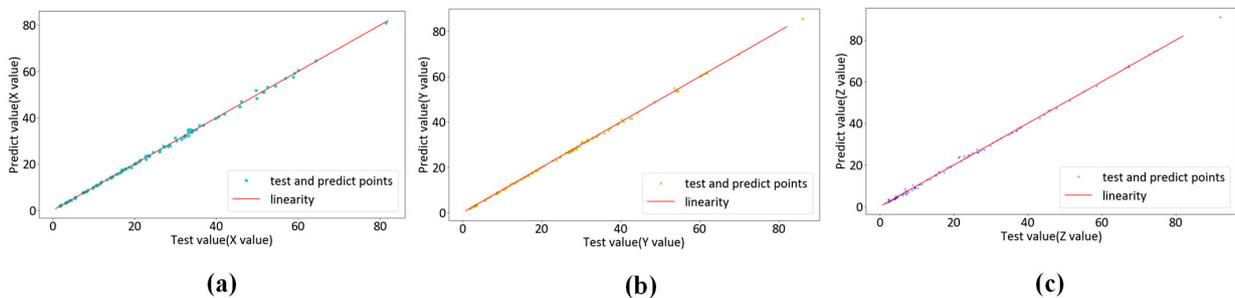


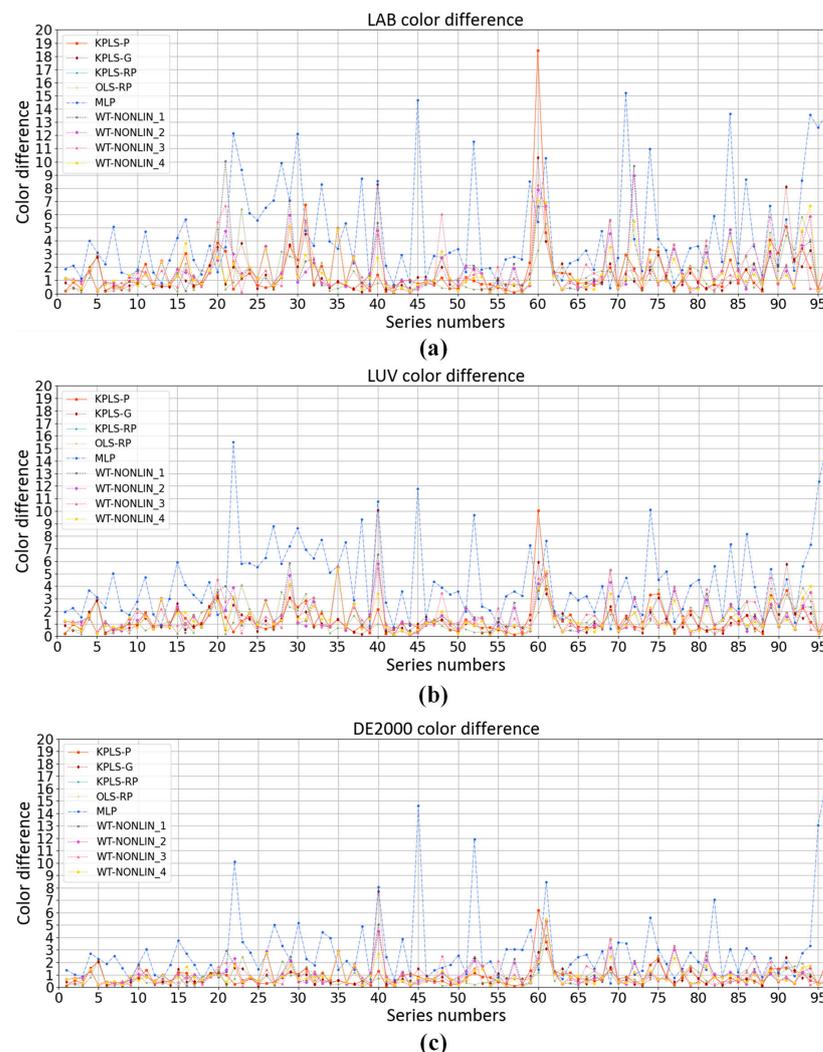
Figure 13. KPLS regression based on root polynomial kernel expansion: (a) X test data and forecast data pairs; (b) Y test data and forecast data pairs; (c) Z test data and forecast data pairs.

Table 2. Model average color difference per fold test set.

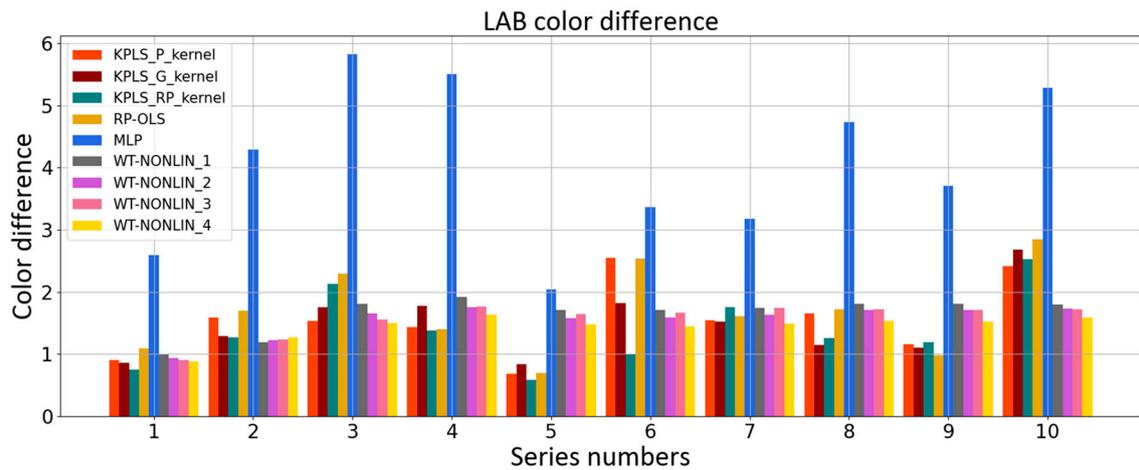
The Series of Fold	P-Kernel			G-Kernel			RP-Kernel		
	LAB	LUV	DE2000	LAB	LUV	DE2000	LAB	LUV	DE2000
1	0.907	0.975	0.666	0.855	0.933	0.613	0.746	0.848	0.527
2	1.583	1.558	0.801	1.294	1.401	0.676	1.273	1.344	0.597
3	1.535	1.351	0.640	1.755	1.444	0.844	2.130	1.718	0.918
4	1.429	1.207	0.635	1.772	1.784	1.150	1.376	1.221	0.665
5	0.681	0.798	0.541	0.834	0.835	0.631	0.581	0.714	0.442
6	2.550	1.666	1.206	1.823	1.372	1.053	1.007	0.753	0.603
7	1.542	1.452	1.188	1.517	1.429	1.144	1.749	1.557	1.311
8	1.655	1.471	0.898	1.146	1.164	0.711	1.258	1.391	0.799
9	1.154	1.010	0.665	1.099	1.056	0.696	1.194	1.232	0.733
10	2.411	1.847	1.116	2.684	1.954	1.113	2.522	1.865	1.003

### 3.5. Comparison with Other Methods

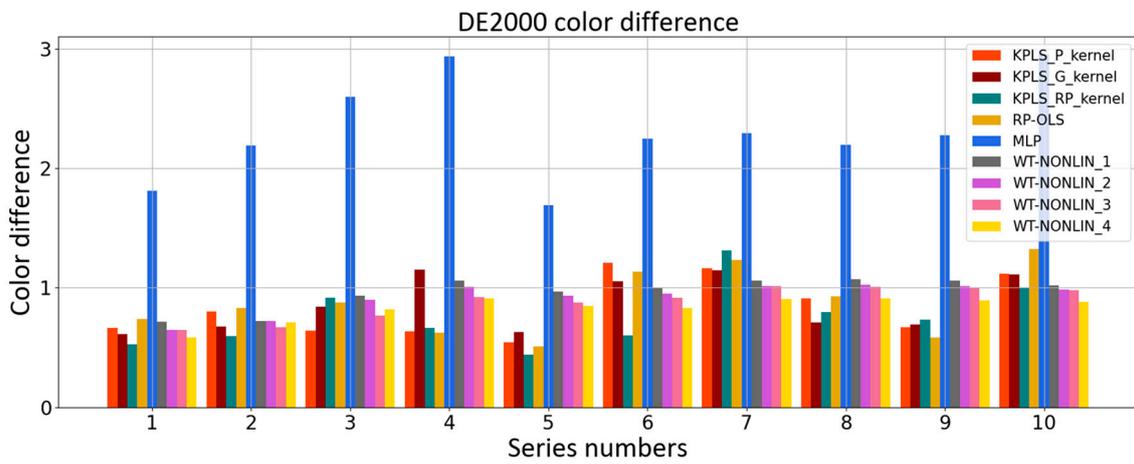
Polynomial regression model [21,22] and neural network model [23–27] are used in colorimetric characterization methods commonly. We compare the KPLS model with the weighted nonlinear polynomial regression model, neural network model and root polynomial regression model [32]. The comparison results are shown in Figures 14 and 15. MLP represents the neural network model. In the MLP model, the hidden layer is 10, the learning rate is 0.1 and the learning target is 0.00001. WT-NONLIN\_1, WT-NONLIN\_2, WT-NONLIN\_3 and WT-NONLIN\_4 represent the weighted regression model which uses four different formulae for calculating the distance here [39]. RP-OLS represents the root polynomial regression model. We choose 3 as the degree of the polynomial. The same dataset is used in this experiment. After dividing the dataset by nested cross-validation, color differences predicted by each model are compared. In Figure 14, the ten test sets are collected together. There are 96 samples. We compare the average color differences of different models under the same sample. In Figure 15, we compare the average color differences of different models of each fold test set. The CIELAB, CIELUV and CIEDE2000 color differences of the test set are calculated and compared. The experimental results are shown in Table 3. Table 3 shows the average color difference of the model prediction on the test set of each fold with different algorithms.



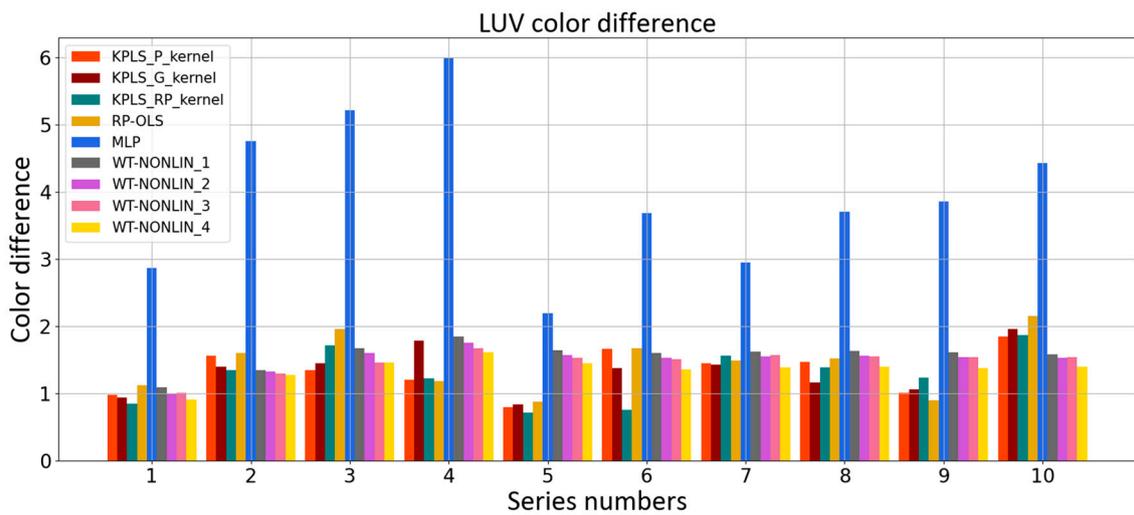
**Figure 14.** Comparison of colorimetric characteristics of different algorithms in the cross-validation method: (a) 96 non-neutral color prediction CIELAB color difference; (b) 96 non-neutral color prediction CIELUV color difference; (c) 96 non-neutral color prediction CIEDE2000 color difference.



(a)



(b)



(c)

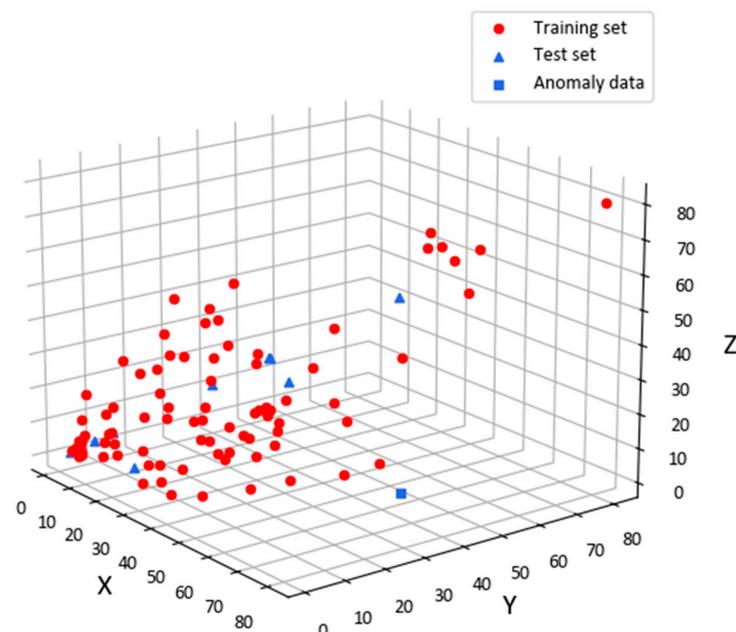
**Figure 15.** Comparison of average color difference of different algorithms: (a) Average CIELAB color difference of ten-fold test data; (b) Average CIELUV color difference of ten-fold test data; (c) Average CIEDE2000 color difference of ten-fold test data.

**Table 3.** The average color difference of different algorithms.

Model		CIELAB Color Difference	CIELUV Color Difference	CIEDE2000 Color Difference
KPLS	P-kernel	1.5447	1.3335	0.8356
KPLS	G-kernel	1.4779	1.3372	0.8631
KPLS	RP-kernel	1.3836	1.2643	0.7598
	RP-OLS [32]	1.4221	1.2933	0.7775
	MLP	4.052	4.4166	2.8895
	WT-NONLIN formula 1 [39]	1.7977	1.5839	1.0207
	WT-NONLIN formula 2 [39]	1.7272	1.5343	0.9858
	WT-NONLIN formula 3 [39]	1.7242	1.5429	0.9799
	WT-NONLIN formula 4 [39]	1.5878	1.4017	0.8847

#### 4. Discussion

We observed a large color difference point in the colorimetric characterization results of the KPLS model and the weighted nonlinear polynomial regression model. This point is the 60th pair of data, corresponding to the 6th fold of the ten-fold data. The sample distribution of the training and test sets for this group of data are shown in Figure 16. The red points are the training set; the blue points are the test set, and the blue square point is the 60th pair of data. We can see that these data are far away from the training set samples in the test set. They belong to the extrapolation prediction of the model. They exhibit a large difference in comparison with the data in the training set, so we cannot accurately predict them. Therefore, how to increase the coverage range of the color gamut of the training dataset or improve the accuracy of the model to enhance the prediction ability of the extrapolation points is the next research direction.

**Figure 16.** Test set distribution where the outliers are located.

#### 5. Conclusions

This paper studies a colorimetric characterization method for color imaging systems based on KPLS. The method uses device-dependent RGB space as input features and expands the feature vectors with kernel functions. According to feature analysis and canonical correlation analysis, a KPLS colorimetric characterization model for color imaging systems is established. The hyperparameters of the model are determined via nested

cross-validation and grid search, and the color space conversion model is constructed. The experimental verification is carried out by using a Canon 1000D commercial digital camera to shoot the SG color card, and CIELAB, CIELUV, CIEDE2000 color differences are used for evaluation. The results show that the color difference of the KPLS colorimetric characterization model based on root polynomial function expansion is better than that of the weighted nonlinear regression model, the neural network model and the root polynomial regression model. The colorimetric characterization method for color imaging systems based on KPLS is an effective method with good prediction accuracy and nonlinear fitting ability, which can support the cross-media color management of color imaging systems well.

**Author Contributions:** Conceptualization, N.L. and Q.L.; methodology, S.Z. and X.X.; software, S.Z., Z.F. and L.L.; validation, S.Z. and X.X.; writing—original draft, S.Z. and X.X.; writing—review and editing, S.Z. and X.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** National Natural Science Foundation of China (Nos. 61975012).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available on request from the corresponding author.

**Acknowledgments:** We are particularly grateful to the reviewers for their rigor and erudition. The proposal of the root kernel function in this paper was inspired by the reviewer.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mou, T.; Shen, H. Colorimetric characterization of imaging device by total color difference minimization. *J. Zhejiang Univ. Sci. A* **2006**, *7*, 1041–1045. [[CrossRef](#)]
2. Ma, Y.; Liu, H.; Liu, X. A research on the color characterization of digital camera. *J. Beijing Inst. Graph. Commun.* **2006**, *14*, 9–12.
3. Zhang, X. *Study of Color Reproduction Theory and Method for Digital Image*; Zhejiang University: Hangzhou, China, 2010.
4. Fu, S.; Cui, C.; Zhang, R. Colorimetric characterization modeling software for digital imaging device. *Opto. Electron. Eng.* **2010**, *37*, 88–92.
5. Green, P. *Color Management: Understanding and Using ICC Profiles*; Wiley: Hoboken, NJ, USA, 2010; pp. 20–34.
6. Danny, C.R. Publication CIE 159: A colour appearance model for colour management systems: CIECAM02. *Color Res. Appl.* **2006**, *31*, 156–159.
7. Charbaji, A.; Heidari-Bafroui, H.; Rahmani, N.; Anagnostopoulos, C.; Faghri, M. Colorimetric Determination of Nitrate after Reduction to Nitrite in a Paper-Based Dip Strip. *Chem. Proc.* **2021**, *5*, 9.
8. Berlina, A.N.; Ragozina, M.Y.; Komova, N.S.; Serebrennikova, K.V.; Zherdev, A.V.; Dzantiev, B.B. Development of Lateral Flow Test-System for the Immunoassay of Dibutyl Phthalate in Natural Waters. *Biosensors* **2022**, *12*, 1002. [[CrossRef](#)]
9. Kim, J.-H.; Lee, Y.-J.; Ahn, Y.-J.; Kim, M.; Lee, G.-J. In situ detection of hydrogen sulfide in 3D-cultured, live prostate cancer cells using a paper-integrated analytical device. *Chemosensors* **2022**, *10*, 27. [[CrossRef](#)]
10. Pomili, T.; Gatto, F.; Pompa, P.P. A Lateral Flow Device for Point-of-Care Detection of Doxorubicin. *Biosensors* **2022**, *12*, 896. [[CrossRef](#)] [[PubMed](#)]
11. *ISO 17321-1:2012*; Graphic Technology and Photography—Colour Characterisation of Digital Still Cameras (DSCs)—Part 1: Stimuli, Metrology, and Test Procedures. International Organization for Standardization: Geneva, Switzerland, 5 November 2012.
12. *ISO 17321-1:2012*; Graphic Technology and Photography—Colour Characterization of Digital Still Cameras (DSCs)—Part 2: Methods for Determining Transforms from Raw Dsc to Scene-Referred. International Organization for Standardization: Geneva, Switzerland, 12 October 2012.
13. *IEC 61966-9:200*; Colour Measurement and Management—Multimedia Systems and Equipment—Part 9: Digital Cameras. International Electrotechnical Commission: Geneva, Switzerland, November 2003.
14. Verdu, F.M.; Pujol, J.; Capilla, P. Calculation of the color matching functions of digital cameras from their complete spectral sensitivities. *J. Imaging Sci. Technol.* **2002**, *46*, 15–25. [[CrossRef](#)]
15. Chouikha, M.B.; Placais, B.; Pouleau, G. Benefits and drawbacks of two methods for characterizing digital cameras. In Proceedings of the IS&T CGIV 2006 3rd European Conference on Colour in Graphics, Imaging, and Vision, Leeds, UK, 19–22 June 2006; Society for Imaging Science and Technology: Springfield, VA, USA, 2006; pp. 185–188.

16. Lee, S.H.; Choi, J.S. Design and implementation of color correction system for images captured by digital camera. *IEEE Trans. Consum. Electron.* **2008**, *54*, 268–276. [[CrossRef](#)]
17. Rump, M.; Zinke, A.; Klein, R. Practical spectral characterization of trichromatic cameras. *ACM Trans. Graph.* **2011**, *30*, 170. [[CrossRef](#)]
18. Hung, P.C. Colorimetric calibration in electronic imaging devices using a look-up-table model and interpolations. *J. Electron. Imaging* **1993**, *2*, 53–61. [[CrossRef](#)]
19. Balasubramanian, R. Reducing the cost of look up table based color transformations. *J. Imaging Sci. Technol.* **2000**, *44*, 321–327. [[CrossRef](#)]
20. Johnson, T. Methods for characterizing colour scanners and digital cameras. *Displays* **1996**, *16*, 183–191. [[CrossRef](#)]
21. Rowlands, D.A. Color conversion matrices in digital cameras: A tutorial. *Opt. Eng.* **2020**, *59*, 110801. [[CrossRef](#)]
22. Jing, J.; Fang, S.; Shi, Z.; Xia, Q.; Li, Y. An efficient nonlinear polynomial color characterization method based on interrelations of color spaces. *Color Res. Appl.* **2020**, *45*, 1023–1039.
23. Liu, Y.; Yu, H.; Shi, J. Camera characterization using back-propagation artificial neural network based on Munsell system. *Proc. SPIE* **2008**, *6621*, 66210A.
24. Li, Y.; Liao, N.; Li, H.; Lv, N.; Wu, W. Colorimetric characterization of the wide-color-gamut camera using the multilayer artificial neural network. *J. Opt. Soc. Am.* **2023**, *40*, 629–636. [[CrossRef](#)] [[PubMed](#)]
25. Liu, L.; Xie, X.; Zhang, Y.; Cao, F.; Liang, J.; Liao, N. Colorimetric characterization of color imaging systems using a multi-input PSO-BP neural network. *Color Res. Appl.* **2022**, *47*, 855–865. [[CrossRef](#)]
26. Miao, H.; Zhang, L. The color characteristic model based on optimized BP neural network. *China Acad. Conf. Printing Packaging* **2016**, *369*, 55–63.
27. Wang, P.; Chou, J.; Tseng, C. Colorimetric characterization of color image sensors based on convolutional neural network modeling. *Sens. Mater.* **2019**, *31*, 1513–1522. [[CrossRef](#)]
28. Yang, B.; Chou, H.; Yang, T. Color reproduction method by support vector regression for color computer vision. *Optik* **2013**, *124*, 5649–5656. [[CrossRef](#)]
29. Gong, R.; Wang, Q.; Shao, X.; Liu, J.J. A color calibration method between different digital cameras. *Optik* **2016**, *127*, 3281–3285. [[CrossRef](#)]
30. Wu, X.; Fang, J.; Xu, H.; Wang, Z. High dynamic range image reconstruction in device-independent color space based on camera colorimetric characterization. *Optik* **2017**, *140*, 776–785. [[CrossRef](#)]
31. Molada-Tebar, A.; Lerma, J.L.; Marques-Mateu, Á. Camera characterization for improving color archaeological documentation. *Color Res. Appl.* **2018**, *43*, 47–57.
32. Finlayson, G.; Mackiewicz, M.; Hurlbert, A. Color correction using root-polynomial regression. *IEEE Trans. Image Process.* **2015**, *24*, 1460–1470. [[CrossRef](#)]
33. Yamakabe, R.; Monno, Y.; Tanaka, M. Tunable color correction for noisy images. *J. Electron. Imaging* **2020**, *29*, 033012. [[CrossRef](#)]
34. Hong, G.; Luo, M.; Rhodes, P.A. A study of digital camera colorimetric characterization based on polynomial modeling. *Color Res. Appl.* **2015**, *26*, 76–84. [[CrossRef](#)]
35. Bianco, S.; Gasparini, F.; Russo, A.; Schettini, R. A new method for RGB to XYZ transformation based on pattern search optimization. *IEEE Trans. Consum. Electron.* **2007**, *53*, 1020–1028. [[CrossRef](#)]
36. Shen, H.; Wan, H.; Zhang, Z. Estimating reflectance from multispectral camera responses based on partial least-squares regression. *J. Electron. Imaging* **2010**, *19*, 020501. [[CrossRef](#)]
37. Heikkinen, V.; Mirhashemi, A.; Alho, J. Link functions and matern kernel in the estimation of reflectance spectra from rgb responses. *J. Opt. Soc. Am. A* **2013**, *30*, 2444–2454. [[CrossRef](#)] [[PubMed](#)]
38. Xiao, G.; Wan, X.; Wang, L.; Liu, S. Reflectance spectra reconstruction from trichromatic camera based on kernel partial least square method. *Opt. Express* **2019**, *27*. [[CrossRef](#)] [[PubMed](#)]
39. Amiri, M.M.; Fairchild, M.D. A strategy toward spectral and colorimetric color reproduction using ordinary digital cameras. *Color Res. Appl.* **2018**, *43*, 675–684. [[CrossRef](#)]
40. Abdi, H.; Williams, L.J. Partial least squares methods: Partial least squares correlation and partial least square regression. In *Computational Toxicology: Volume II*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 549–579.
41. Georgoula, M. Assessing colour differences of lighting stimuli using a visual display. PhD Thesis, University of Leeds, Leeds, UK, 2015.
42. Commission Internationale de l’Eclairage (CIE). *Recommendations on Uniform Color Spaces—Color Difference Equations, Psychometric Color Terms*; CIE Publication: Vienna, Austria, 1978.
43. Liu, H.; Cui, G.; Huang, M. Color-difference threshold for printed images. *Appl. Mech. Mater.* **2013**, *469*, 236–239.
44. Luo, M.; Cui, G.; Georgoula, M. Colour difference evaluation for white light sources. *Light Res. Technol.* **2015**, *47*, 360–369. [[CrossRef](#)]
45. Wen, S. P-46: A color space derived from CIELUV for display color management. *SID Symp. Dig. Tech. Pap.* **2012**, *42*, 1269–1272. [[CrossRef](#)]
46. Schanda, J. CIE  $u'$ ,  $v'$  uniform chromaticity scale diagram and CIELUV color space. In *Encyclopedia of Color Science and Technology*; Luo, M.R., Ed.; Springer: Berlin/Heidelberg, Germany, 2015.
47. Hill, B.; Fw, V.; Roger, T. Comparative analysis of the quantization of color spaces on the basis of the CIELAB color difference formula. *ACM Trans. Graph.* **1997**, *16*, 109–154. [[CrossRef](#)]

48. Melgosa, M.; Trémeau, A.; Cui, G. *Colour Difference Evaluation*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 59–77.
49. Zhang, X.; Qiang, W.; Li, J. Estimating spectral reflectance from camera responses based on CIE XYZ tristimulus values under multi-illuminants. *Color Res. Appl.* **2017**, *42*, 68–75. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.