https://doi.org/10.2352/issn.2169-2629.2021.29.300

An analysis of spectral similarity measures

Mirko Agarla[1]***, Simone Bianco***[1]***, Luigi Celona***[1]***, Raimondo Schettini***[1]***, Mikhail Tchobanou***[2]
[1] ***University of Milano-Bicocca, Milano, Italy***
[2] ***Moscow Research Center, Huawei Technologies Co. Ltd, Russia***

Abstract

In this paper we analyze the most used measures for the assessment of spectral similarity of reflectance and radiance signals. First of all we divide them in five groups on the basis of the type of errors they measure. We proceed analyzing their mathematical definition to identify unintended behaviors and types of errors they are blind to. Then exploiting the Munsell atlas we analyze the correlation between metrics in terms of both Pearson's Linear Correlation Coefficient (PLCC) and Spearman's Rank Order Correlation Coefficient (SROCC). Finally we analyze the behaviour of the selected metrics with respect to two different color properties: the Chroma and the Lightness computed in the CIE $L^*a^*b^*$ color space.

The source code of the spectral measures considered is available at the following link: `https://celuigi.github.io/spectral-similarity-metrics-comparison/`.

Introduction

Hyperspectral imaging is an important visual modality that has gained increasing interest for a wide range of applications in astronomy, agriculture, molecular biology, biomedical imaging, etc. Unlike RGB or multispectral acquisition devices, the goal of hyperspectral imaging is to acquire the complete spectral signature reflected from each observable point. Although this information is rich and very useful, the diffusion of Hyperspectral Imaging Systems (HIS) is hindered by the fact that existing devices have a limited resolution (spatial, spectral and/or temporal), and moreover they are bulky and expensive.

Although the acquisition of full spectral signatures has evolved considerably in the past few decades [8, 6], more and more studies attempt to recover hyperspectral information using only RGB cameras [11, 12, 1], or more in general, to recover hyperspectral information from tristimulus values [22, 5]. This growth of interest is due both to the organization of several international challenges (e.g. [18, 3, 4]) and to the collection and availability of large databases for spectral reconstruction from RGB [2, 4, 14]. Additionally, measures for the evaluation and comparison of the proposed algorithms have been used or introduced. Since these measures have different peculiarities and assess reconstruction errors in profoundly different ways, it is important to analyze: to which extent these measures are reliable; whether and which measures are correlated.

In this article we provide a thorough analysis of the measures that are commonly used for the benchmark of spectral reconstruction methods, and more in general for the assessment of similarity between spectra. In particular we include in the analysis 14 different measures, that we group according to the type of error they measure, and finally we assess: (i) measures peculiarities, (ii) how measures relate to each other, (iii) the behaviour of measures with respect to different color properties.

Spectral similarity measures

In this section we identify the commonly used measures for evaluating spectral similarity and we then divide them into five groups according to the type of error that they measure.

Mean squared error measures. The first group consists in what could be called mean error measures, which are substantially based on an euclidean measure of the error between the recovered and ground-truth spectra, or their RGB backprojections. measures belonging to this group are: the Mean Square Error (MSE), the Root Mean Square Error (RMSE), the Mean Relative Absolute Error (MRAE) which operates in the spectral domain and its correspondent in the RGB domain which is the Back-Projection MRAE (BPMRAE) [21], and the Peak Signal-to-Noise Ratio (PSNR). The definitions of the previous measures are the following:

$$MSE = \frac{\sum_x \left\| R(x,\lambda) - \hat{R}(x,\lambda) \right\|_2^2}{N}, \tag{1}$$

$$RMSE = \sqrt{MSE}, \tag{2}$$

$$MRAE = \frac{1}{N} \sum_{x,\lambda} \frac{|R(x,\lambda) - \hat{R}(x,\lambda)|}{R(x,\lambda)}, \tag{3}$$

$$BPMRAE = \frac{1}{N} \sum_{x,\lambda} \frac{|\mathrm{CRF} \times R(x,\lambda) - \mathrm{CRF} \times \hat{R}(x,\lambda)|}{\mathrm{CRF} \times R(x,\lambda)}, \tag{4}$$

$$PSNR = 20 \times \log_{10}, \left(\frac{p\_max}{MSE} \right), \tag{5}$$

where $R(x,\lambda)$ and $\hat{R}(x,\lambda)$ are the actual and reconstructed spectral reflectances, N is the size of the actual image (pixel count $\times$ number of spectral channels), CRF is the camera response function, and $p\_max = 2^{16} - 1$, i.e. 65,535, corresponds to the maximum possible value of each pixel (it depends on image encoding).

Similarity measures. The second group of measures regards measures that instead of computing the euclidean distance, estimate the similarity between recovered spectra and ground-truth. Each of them uses a different definition of similarity. To the second group belong: the Goodness-of-Fit Coefficient (GFC) [16] that independently compares image pixels, the Mean Structural SImilarity Measure (MSSIM) [20] that instead computes error between actual and reconstructed images on small patches with a sliding window approach. The definition of GFC is the following:

$$GFC = \frac{1}{N} \sum_x \frac{|\sum_\lambda R(x,\lambda) \hat{R}(x,\lambda)|}{\sqrt{\sum_\lambda [R(x,\lambda)]^2} \sqrt{\sum_\lambda [\hat{R}(x,\lambda)]^2}}, \tag{6}$$

where $R(x,\lambda)$ and $\hat{R}(x,\lambda)$ are the actual and reconstructed spectral reflectances, while N is the size of the actual image (pixel count $\times$ number of spectral channels). The equation of MSSIM is the following:

$$\begin{aligned} SSIM_{i,n} &= \frac{(2\mu\mu^* + C_1)(2\hat{\sigma} + C_2)}{(\mu^{*2} + \mu^2 + C_1)(\sigma^{*2} + \sigma^2 + C_2)}, \\ MSSIM &= \frac{1}{M \times W \times H} \sum_{i=1}^{M} \sum_{n=1}^{W \times H} SSIM_{i,n}, \end{aligned} \tag{7}$$

where μ^* and σ^{*2} are the mean and variance for the n^{th} $N \times N$ window in the i^{th} wavelength-indexed band on the reconstructed spectral image. Similarly, μ and σ^2 account for the mean and variance of the window in the actual image. Also, $C_1 = k_1 L$, and $C_2 = k_2 L$ are introduced to avoid division by zero when the mean or covariance values are close to zero. M is the number of wavelengths, W and H are width and height of the image.

Angular measures. The third group of measures consists of measures that measure the angle between the recovered and ground-truth spectra. In this way, they focus more on the shape of the spectra than on their absolute values: two spectra that differ in just a global scaling are considered to have zero error. To this group belong the Angular error measures: Spectral Angle Mapper (SAM) [10], Mean Angular Error (MAngE), and Average Per-Pixel Spectral Angle (APPSA) [18]. SAM calculates the average spectral angle between the spectra of the actual and the reconstructed hyperspectral images and is based on this equation:

$$SAM = \frac{1}{m} cos^{-1} \left(\sum_{j=1}^{m} \frac{(p_h^{(j)})^T (p_c^{(j)})}{\left\| p_h^{(j)} \right\|_2 \left\| p_c^{(j)} \right\|_2} \right). \tag{8}$$

$p_h^{(j)}$, $p_c^{(j)} \in \mathbb{R}^C$ represent the spectra of the j-th hyperspectral pixel in real and estimated hyperspectral images (C is the number of bands), and m is the total number of pixels within an image. MAngE is expressed by the formula:

$$MAngE = \frac{1}{N} \sum_i angle(r_i, \hat{r}_i), \tag{9}$$

where r_i and $\hat{r}_i$ denote the i-th actual and reconstructed spectra; N is the size of the actual image (pixel count $\times$ number of spectral channels). APPSA is computed using this equation:

$$\begin{aligned} \rho &= \arccos \left(\frac{\sum_{i=1}^{M} (\mathbf{I}_i^* \odot \mathbf{I}_i)}{\sqrt{\sum_{i=1}^{M} (\mathbf{I}_i^* \odot \mathbf{I}_i^*)} \sqrt{\sum_{i=1}^{M} (\mathbf{I}_i \odot \mathbf{I}_i)}} \right) \\ APPSA &= \frac{1}{W \times H} \mathbf{1}^T \rho \mathbf{1}, \end{aligned} \tag{10}$$

where $\mathbf{I}_i^*$ is the matrix corresponding to the i-th wavelength-indexed channel in the super-resolved image, I_i, is the array for the channel indexed i in the reference, and M is the total number of wavelengths.

Colorimetric error measures. In the fourth group the colorimetric errors are included. These are used in order to have a measure that is well correlated with human perception. Differently from the measures in the previous groups, the measures in this group need an illuminant spectral power distribution when the spectra to be compared are reflectances.

Several versions of the ΔE measure have been proposed in the literature. We consider here the color difference ΔE_{76} [15], ΔE_{94}, and ΔE_{00} [17, 13], respectively.

To this category also belongs the euclidean distance in proLab space [9]. Transformation from CIE XYZ to proLab space is achieved via the following steps:

1. Remove the illuminant;
2. Convert into homogeneous coordinates;
3. Multiply by the optimized projection matrix

$$Q = \begin{bmatrix} 75.5362 & 486.661 & 167.387 & 0 \\ 617.7141 & -595.4477 & -22.2664 & 0 \\ 48.3433 & 194.9377 & -243.281 & 0 \\ 0.7554 & 3.8666 & 1.6739 & 1 \end{bmatrix};$$

4. Divide the first coordinates by the fourth coordinate (i.e. return to non-homogeneous coordinates).

Other measures. Finally, the fifth groups actually includes just one measure: the Mean Spectral Information Divergence (MSID) [7], which can be used to measure the spectral similarity between two pixel vectors x and y. The equation defining the previous measure is the following:

$$\begin{aligned} SID &= D(\mathbf{x} \,\|\, \mathbf{y}) + D(\mathbf{y} \,\|\, \mathbf{x}) \\ D(\mathbf{x} \,\|\, \mathbf{y}) &= \sum_{l=1}^{L} p_l \log(p_l / q_l) \\ D(y \,\|\, x) &= \sum_{l=1}^{L} q_l \log(q_l / p_l), \\ MSID &= \frac{1}{M} \sum_{i=1}^{M} SID_i. \end{aligned} \tag{11}$$

For a given multispectral ground-truth pixel vector $\mathbf{x} = (x_1, ..., x_L)$ and the predicted pixel vector $\mathbf{y} = (y_1, ..., y_L)$, each component x_l and y_l is a pixel of band image B_l. We normalize $\mathbf{x}$ and $\mathbf{y}$, $p_j = x_j / \sum_{l=1}^{L} x_l$ and $q_j = y_j / \sum_{l=1}^{L} y_l$ to get the respective probability values p_j and q_j.

The different measures considered are summarized in Table 1 together with complementary information detailing the applicability of each measure. Specifically, the column "Loss" indicates whether the metric has been used as a loss function for optimizing prediction models (e.g. neural networks) in the literature investigated by the authors. The column "Domain" reports in which domain the metric operates (i.e. RGB or spectral). Finally, the columns "Requirements" specify what extra information is needed to compute the metric (e.g. illuminant, illuminant white point, camera response functions).

Analysis

We conduct an analysis of the evaluation measures presented in the previous section to understand if they are correlated to each other, and how they behave on different colors. In particular we want to understand if their behaviour is homogeneous in the color space, if they tend to respond more on dark or light colors, if they tend to respond more on saturated or un-saturated colors.

Dataset

For this analysis we consider the 1269 Munsell chips of the Munsell atlas. We decomposed the ground-truth reflectances using Principal Component Analysis (PCA) and then reconstructed them considering only the first 6 PCA components following [19]

Table 1: Main characteristics of the evaluation measures considered. Column "Loss" indicates that the measure is used as loss function, CRF stands for Camera response function.

Name	Loss	Domain		Requirements		
		RGB	Spectral	Illum.	White point	CRF
MSE	✓		✓			
RMSE	✓		✓			
MRAE	✓	✓	✓			
BPMRAE [21]	✓	✓	✓			✓
PSNR	✓		✓			
GFC [16]	✓		✓			
MSSIM [20]			✓			
SAM [10]			✓			
MAngE			✓			
APPSA [18]			✓			
ΔE [15, 17, 13]				✓	✓	
proLab [9]				✓	✓	
MSID [7]	✓		✓			

that found that less than 7 basis are sufficient to obtain a visually acceptable approximation. We use the CIE Standard D65 illuminant to compute the XYZ and CIE L*a*b* values of the reconstructed and ground-truth chips.

Experimental results

Measures behaviour. First of all, we analyzed all the measures to understand if they have any problem in their definition and to what type of difference they are blind.

We noticed that MRAE and BPMRAE have a highly asymmetric behavior and should be avoided nonetheless MRAE is one of the official measures used in the NTIRE2020 spectral reconstruction challenge [4]. For this analysis we consider gray reflectances, i.e. $r(\lambda) = \alpha$ for $\lambda \in [400, \dots, 700]$, and $\alpha = \{0.001, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 1.00\}$. We computed the measures value for each couple of spectra obtained by scanning the α values from the smallest to the largest for what concerns the ground-truth, and from the largest to the smallest for the recovered spectra.

Figure 1 reports a graphical representation of such combinations. For each combination we compute the MRAE, RMSE, PSNR and MSE values, that are plotted in their respective plots in the same figure. We can observe how in the left part of the plot of MRAE (and consequently also BPMRAE) the error is several orders of magnitudes higher than the right part of the plot, although the absolute difference between the ground-truth and the prediction is the same. This does not happen with other measures such as RMSE, MSE and PSNR suggesting these should be preferred. This suggests that MRAE gives more importance to errors on dark colors.

Another problem may arise using angular measures: in fact these measures estimate the error between the ground-truth and the prediction by looking solely at their shape. This means that if they differ from just a multiplicative factor, they are judged to be equal. This may not represent a problem if in an image all the spectra are scaled by the same scaling factor, but it may constitute a problem if this scaling factor varies spatially. Unfortunately these measures cannot distinguish the two behaviors and that is why they should not be used in isolation but should be combined with other measures.

Another type of problem affects those measures that projects the spectra on a set of filters, such as BPMRAE and the ΔE variants. In fact there are spectra that although different, results in the same projection thus leading to zero errors. These type of spectra are called metamers.

Correlation between measures. To understand if the measures are related to each other, we compare two measure at time using the Pearson's Linear Correlation Coefficient (PLCC) and the Spearman's Rank Order Correlation Coefficient (SROCC) over all the Munsell chips. PLCC is defined as follows:

$$PLCC = \frac{\sum_i^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^N (x_i - \bar{x})^2}\sqrt{\sum_i^N (y_i - \bar{y})^2}}, \tag{12}$$

where N is the number of samples, x_i and y_i are the sample points indexed with i, and finally $\bar{x}$ and $\bar{y}$ are the means of each sample distribution. On the other side, the SROCC estimates the monotonic relationship between the actual and the predicted scores, and it is calculated as follows:

$$SROCC = 1 - \frac{6\sum_i^N d_i^2}{N(N^2-1)}, \tag{13}$$

where N is the number of samples, and $d_i = (\text{rank}(x_i) - \text{rank}(y_i))$ is the difference between the two ranks of each sample.

Tables 2a and 2b respectively report the correlation between each pair of measures in terms of both PLCC and SROCC.

Since we are not interested in the direction of the correlation, the tables are color coded on the basis of the absolute value of the correlation coefficients. Low to high absolute correlations correspond to a transition from light to dark blue. First of all we can notice how SROCC tends to be higher than PLCC, thus highlighting the monotonic relationship existing among the measures. We can also observe how, unsurprisingly, the strongest correlation occurs with measures belonging to the same group, with the highest correlation occurring between angular measures. On the other hand we can also notice that MRAE is the one showing the highest cross-group correlation.

Chroma and Lightness. In order to better understand how the different measures behave with respect to color chroma and dark/light colors we computed the PLCC and SROCC correlation coefficients of each measure with respect to Chroma C^* and Lightness L^* in the CIE L*a*b* color space. The results are reported in Tables 3a and 3b. Since now we are interested also in the sign of the correlation, the coefficients are color coded taking in consideration their signed value: values in the range $[-1, 1]$ are now mapped to transitions from green to red.

The results reported show that some measures are neutral with respect to Chroma, e.g. SAM, MangE, GFC, and ΔE_{00}. Some are positively correlated and thus give more importance to errors on colors with high Chroma, e.g. RMSE, MSE, MRAE. Some are negatively correlated, thus giving more importance to errors on colors with low Chroma, e.g. PSNR.

Concerning the analysis of the correlations with respect to Lightness we can observe how many measures behave differently from how they behaved with respect to Chroma. In fact only two are neutral with respect to chroma, e.g. BPMRAE and MSID. Some are positively correlated, thus giving more importance to errors on colors with high Lightness, e.g. RMSE, MSE, and GFC. The majority of them shows instead a negative correlation and as such give more importance to error on color with low lightness, e.g. MRAE, PSNR, SAM, MangE, APPSA, and ΔE_{00}. In particular we can observe how the highest absolute correlation coefficients is obtained by the angular measures, thus further confirming our analysis at the beginning of this section and the problems highlighted in Figure 1.

Table 2: Correlation analysis of the different measures considered, in terms of (a) PLCC and (b) SROCC.

Group		Mean error measures				Angular measures			Similarity measure	Colorimetric error measures				Other measures
	measure name	MSE	MRAE	BPMRAE	PSNR	SAM	MangE	APPSA	GFC	ΔE00	ΔE94	ΔE00	proLab	MSID
Mean error measures	RMSE	0.9500	0.4991	0.1904	-0.9379	0.2498	0.7646	0.2498	-0.2298	0.3356	0.3944	0.5283	0.5987	0.3933
	MSE		0.5495	0.2003	-0.7951	0.2968	0.7590	0.2968	-0.2884	0.3419	0.3998	0.5177	0.5848	0.5078
	MRAE			0.3257	-0.4129	0.6916	0.7423	0.6916	-0.5994	0.6116	0.6221	0.7745	0.6589	0.5940
	BPMRAE				-0.1607	0.1380	0.1935	0.1380	-0.1037	0.2294	0.2229	0.3782	0.2307	0.1478
	PSNR					-0.1981	-0.7004	-0.1981	0.1676	-0.3208	-0.3722	-0.4936	-0.5616	-0.2795
Angular measures	SAM						0.8034	1.0000	-0.9025	0.6655	0.6955	0.5539	0.6526	0.3423
	MangE							0.8034	-0.7193	0.6209	0.6846	0.6752	0.8071	0.4485
	APPSA								-0.9025	0.6655	0.6955	0.5539	0.6526	0.3423
Similarity measure	GFC									-0.5481	-0.5764	-0.4362	-0.5508	-0.4115
Colorimetric error measures	ΔE00										0.9820	0.8387	0.8020	0.2981
	ΔE94											0.8576	0.8746	0.2965
	ΔE76												0.8766	0.3669
	proLab													0.3426

(a)

Group		Mean error measures				Angular measures			Similarity measure	Colorimetric error measures				Other measures
	measure name	MSE	MRAE	BPMRAE	PSNR	SAM	MangE	APPSA	GFC	ΔE00	ΔE94	ΔE00	proLab	MSID
Mean error measures	RMSE	1.0000	0.3336	0.3372	-1.0000	0.2014	0.7064	0.2014	-0.2014	0.3121	0.3467	0.4920	0.5256	0.7665
	MSE		0.3336	0.3372	-1.0000	0.2014	0.7064	0.2014	-0.2014	0.3121	0.3467	0.4920	0.5256	0.7665
	MRAE			0.7524	-0.3336	0.8889	0.7786	0.8889	-0.8889	0.4929	0.5227	0.6696	0.6112	0.8086
	BPMRAE				-0.3372	0.6998	0.6901	0.6998	-0.6998	0.7476	0.7664	0.8714	0.8138	0.7235
	PSNR					-0.2014	-0.7064	-0.2014	0.2014	-0.3121	-0.3467	-0.4920	-0.5256	-0.7665
Angular measures	SAM						0.8020	1.0000	-1.0000	0.6131	0.6428	0.6693	0.6577	0.7075
	MangE							0.8020	-0.8020	0.6535	0.6928	0.7771	0.7843	0.9486
	APPSA								-1.0000	0.6131	0.6428	0.6693	0.6577	0.7075
Similarity measure	GFC									-0.6131	-0.6428	-0.6694	-0.6577	-0.7075
Colorimetric error measures	ΔE00										0.9884	0.9072	0.9068	0.6072
	ΔE94											0.9295	0.9390	0.6405
	ΔE76												0.9725	0.7899
	proLab													0.7624

(b)

Table 3: PLCC (a) and SROCC (b) correlation of the different measures with respect to chroma and lightness properties of colors.

Group	Mean error measures					Angular measures			Similarity and other measures		
measure name	RMSE	MSE	MRAE	BPMRAE	PSNR	SAM	MangE	APPSA	GFC	MSID	ΔE00
Chroma	0.505	0.416	0.484	0.251	-0.442	0.052	0.052	-0.208	0.011	0.406	0.029
Lightness	0.377	0.259	-0.363	-0.015	-0.288	-0.630	-0.630	-0.831	0.449	-0.084	-0.285

(a)

Group	Mean error measures					Angular measures			Similarity and other measures		
measure name	RMSE	MSE	MRAE	BPMRAE	PSNR	SAM	MangE	APPSA	GFC	MSID	ΔE00
Chroma	0.507	0.507	0.502	0.294	-0.507	0.182	0.182	-0.064	-0.182	0.499	-0.092
Lightness	0.489	0.489	-0.511	-0.317	-0.489	-0.663	-0.663	-0.933	0.663	-0.055	-0.233

(b)

Conclusions

In this paper we have presented a novel analysis of measures for spectral similarity assessment. The measures were first divided into five groups according to the type of error they measure. The implementation of the considered measures was made publicly available to prevent different results due to different interpretations of the same metric.

Using as a dataset the Munsell atlas, we focused on the behaviour and correlation between the most commonly used error measures and their behavior with respect to Chroma and Lightness of the colors.

The analysis on the measures highlighted how certain measures have asymmetric behavior (e.g. MRAE and BPMRAE) and others are blind to specific spectral differences (e.g. angular measures) and thus should be carefully used and possibly in combination with other measures.

The correlation indexes show how measures belonging to the same group have a very high correlation among them. Results also show how metrics belonging to different groups can also have a high correlation. These observations suggest that when it is important to assess the spectral similarity considering multiple aspects, selecting just one measure for each group could be sufficient.

Finally the analysis on of the behavior of the measures with respect to color properties such as Chroma and Lightness show that the majority of measures is positively correlated with Chroma and negatively correlated with Lightness.

Acknowledgment

This research was supported by Huawei Technologies Co. Ltd. Russia.

References

[1] Jonas Aeschbacher, Jiqing Wu, and Radu Timofte. “In defense of shallow learned spectral reconstruction from rgb images”. In: *ICCVW*. 2017, pp. 471–479.

[2] Boaz Arad and Ohad Ben-Shahar. “Sparse recovery of hyperspectral signal from natural rgb images”. In: *ECCV*. Springer. 2016, pp. 19–34.

[3] Boaz Arad, Ohad Ben-Shahar, and Radu Timofte. “Ntire 2018 challenge on spectral reconstruction from rgb images”. In: *CVPRW*. IEEE/CVF. 2018, pp. 929–938.

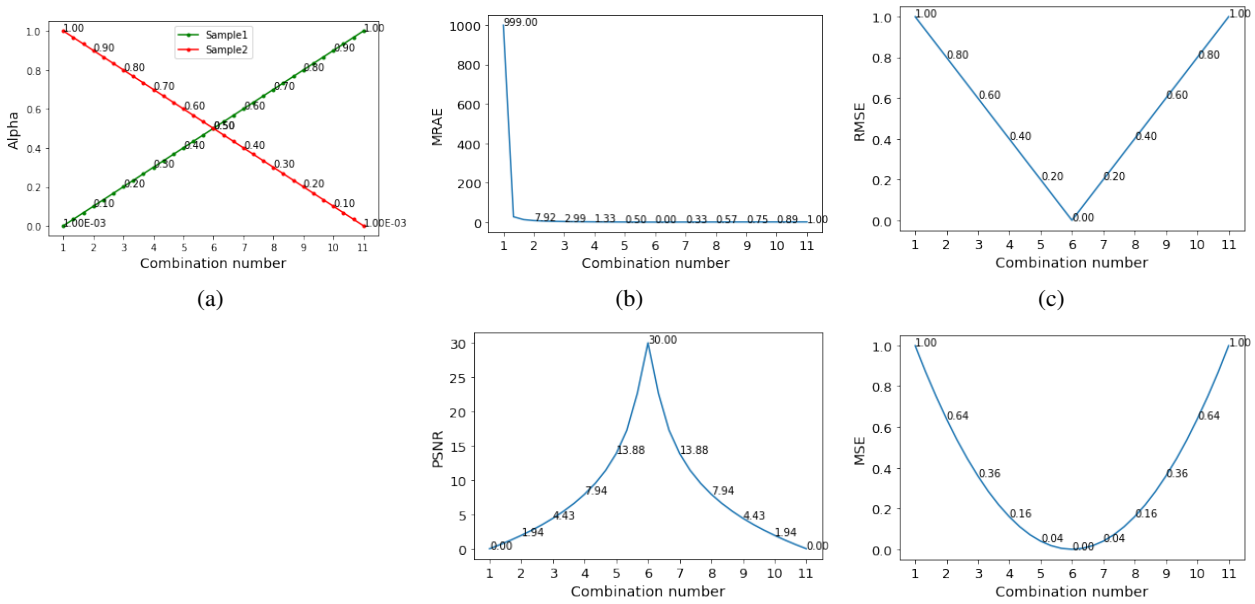

Figure 1: Behaviour of measures for gray reflectances, i.e. $r(\lambda) = \alpha$, with different α values. (a) Combinations considered for metric computation varying α for the two samples compared. (b)–(e) report the estimated error with different measures between the two samples on the ordinate and the combination number on the abscissa. Specifically, (b) MRAE (BPMRAE has the same behavior and thus is not reported). (c) RMSE. (d) PSNR. (e) MSE.

[4] Boaz Arad et al. "Ntire 2020 challenge on spectral reconstruction from an rgb image". In: *CVPRW*. IEEE/CVF. 2020, pp. 446–447.

[5] Simone Bianco. "Reflectance spectra recovery from tristimulus values by adaptive estimation with metameric shape correction". In: *Journal of the Optical Society of America A* 27.8 (2010), pp. 1868–1877. ISSN: 1084-7529. DOI: 10.1364/JOSAA.27.001868.

[6] Xun Cao et al. "High resolution multispectral video capture with a hybrid camera system". In: *CVPR*. IEEE. 2011, pp. 297–304.

[7] Chein-I Chang. "Spectral information divergence for hyperspectral image analysis". In: *International Geoscience and Remote Sensing Symposium*. Vol. 1. IEEE. 1999, pp. 509–511.

[8] David William Fletcher-Holmes and Andrew Robert Harvey. "Real-time imaging with a hyperspectral fovea". In: *Journal of Optics A: Pure and Applied Optics* 7.6 (2005), S298.

[9] Ivan A. Konovalenko et al. "ProLab: a perceptually uniform projective color coordinate system". In: *IEEE Access* (2021), pp. 1–1. DOI: 10.1109/ACCESS.2021.3115425.

[10] Fred A Kruse et al. "The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data". In: *Elsevier Remote sensing of environment* 44.2-3 (1993), pp. 145–163.

[11] Jiaojiao Li et al. "Adaptive weighted attention network with camera spectral sensitivity prior for spectral reconstruction from RGB images". In: *CVPRW*. IEEE/CVF. 2020, pp. 462–463.

[12] Yi-Tun Lin and Graham D Finlayson. "Physically Plausible Spectral Reconstruction". In: *MDPI Sensors* 20.21 (2020), p. 6399.

[13] M Ronnier Luo, Guihua Cui, and Bryan Rigg. "The development of the CIE 2000 colour-difference formula: CIEDE2000". In: *Color Research & Application* 26.5 (2001), pp. 340–350.

[14] Rang MH Nguyen, Dilip K Prasad, and Michael S Brown. "Training-based spectral reconstruction from a single RGB image". In: *ECCV*. Springer. 2014, pp. 186–201.

[15] AR Robertson. "The CIE 1976 Color-Difference formulae, Color Research Aplied, 2, 7". In: *Genotype Prinos mesa (%) Klasa polutki VJ* 56 (1977).

[16] Javier Romero, Antonio Garcıa-Beltrán, and Javier Hernández-Andrés. "Linear bases for representation of natural and artificial illuminants". In: *JOSA A* 14.5 (1997), pp. 1007–1014.

[17] Gaurav Sharma, Wencheng Wu, and Edul N Dalal. "The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations". In: *Color Research & Application* 30.1 (2005), pp. 21–30.

[18] Mehrdad Shoeiby et al. "Pirm2018 challenge on spectral image super-resolution: Dataset and study". In: *ECCV-W*. 2018.

[19] Michael J Vrhel, Ron Gershon, and Lawrence S Iwan. "Measurement and analysis of object reflectance spectra". In: *Color Research & Application* 19.1 (1994), pp. 4–9.

[20] Zhou Wang et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612.

[21] Yuzhi Zhao et al. "Hierarchical regression network for spectral reconstruction from RGB images". In: *CVPRW*. IEEE/CVF. 2020, pp. 422–423.

[22] Silvia Zuffi, Simone Santini, and Raimondo Schettini. “From Color Sensor Space to Feasible Reflectance Spectra”. In: *IEEE Transactions on Signal Processing* 56.2 (2008), pp. 518–531.