Representing Spectral Functions by a Composite Model of Smooth and Spiky Components for Efficient Full-Spectrum Photorealism

Yinlong Sun, Mark S. Drew, and F. David Fracchia
School of Computing Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6
{sun, mark, fracchia}@cs.sfu.ca

Abstract

We propose a new model called the "composite model" to represent spectral functions. This model is built on the idea of decomposing all spectral functions into smooth and spiky components, each with its own distinct representation. A smooth spectrum can be expressed with coefficients of a set of given basis functions, and the discrete spikes in a spiky spectrum with their locations and heights. For the smooth part, we propose re-sampling functions that are reconstructed from the coefficients in a linear combination to improve efficiency. Spectral multiplication is thus greatly reduced in complexity. This new model shows remarkable advantages in representing spectral functions with aspect to accuracy, compactness, computational efficiency, portability, and flexibility, and it has a great application potential in color science, realistic image synthesis, and color image analysis. Here we apply it to rendering images involving real spiky illuminants as well as objects with light dispersion. The composite model is shown to surpass other models in these applications.

1. Introduction

Spectral functions, such as spectral power distributions (SPDs) of lights and reflectances of materials, are essential in realistic image synthesis, color image analysis, and color science. Because natural spectral functions can be of arbitrary shape, they may need an infinite number of parameters to describe exactly. In practice, however, describing every spectral function with even a finite but large number of parameters is not feasible. For example, a photograph usually involves hundreds or even thousands of spectral functions. Therefore, it is highly desirable that we develop a method that can represent all spectral functions through a small number of parameters while guaranteeing sufficient accuracy.

Specifically, spectral representations should meet the following criteria:

- *Accuracy*: A representation must be capable of achieving very high accuracy.
- *Compactness*: Spectral functions should be represented with as few parameters as possible. This can greatly reduce the memory required in computations.
- *Efficiency*: In many cases, the overall performance efficiency is largely determined by the efficiency of *multiplying* two spectral functions (a computation associated with each reflection and transmission).
- *Portability*: The data representing spectral functions should be portable. This facilitates exchanges and aggregations of spectral data from various sources.
- *Flexibility*: A representation should allow convenient adjustment according to application requirements (e.g., accuracy is a priority in realistic image synthesis while efficiency is more important in real-time simulations).

In this paper, we propose a new model called the *composite model* [20] to represent spectral functions. The basic idea is that we classify all spectral functions into smooth and spiky spectra, and represent the two categories differently. A smooth spectrum can be effectively expressed as a linear combination of a finite set of given basis functions and therefore be represented through the coefficients in the expansion. The discrete spikes of a spiky spectrum can be represented through their locations and heights, and the smooth background (if any) can be represented in the same way as if it were a smooth spectrum. This new model demonstrates remarkable advantages in representing spectral functions and has great application potential in color science, realistic image synthesis and color image analysis.

2. Previous Spectral Representations

2.1. Color models

Color models, such as the CIE XYZ or RGB models, are spectral representations based on projections from the infinite-dimensional spectrum space to three-dimensional color spaces. The CIE XYZ model is described by

$$X_k = \kappa \int_{\lambda_{\min}}^{\lambda_{\max}} S(\lambda) x_k(\lambda) d\lambda, \quad k = 1,2,3 \qquad (2.1)$$

where $S(\lambda)$ is a spectral function, $x_k(\lambda)$ are the CIE XYZ color matching functions, $[\lambda_{\min}, \lambda_{\max}]$ is the visible wavelength range, and κ is a constant. Color models are advantageous in compactness, efficiency and portability. Their disadvantage is insufficient accuracy [11,12], as much information can be lost during the transformation

from a spectral function into a color. As most natural spectral functions are very smooth, color models seem to work fine in many applications [1]. But in general they are not accurate and provide no potential for achieving reliable accuracy. Finally, they provide little flexibility.

2.2. Point sampling method

This method represents a spectrum through sample values at a set of discrete wavelengths, with sampling intervals typically from 1 nm to 10 nm. An advantage of this method is its capability of achieving high accuracy if the sampled points are dense enough. This method is also efficient in performance in that products of spectral functions are computed by simply multiplying corresponding functional values at sample points,

$$S_3(\lambda_i) = S_1(\lambda_i)S_2(\lambda_i), \quad i = 1,..., N, \tag{2.2}$$

with cost O(N). Moreover, the point sampling method is portable and is flexible (one can use different numbers of sampled points). Its disadvantage is poor compactness. In particular, this method is not suitable for representing spectra with spikes such as in fluorescent SPDs.

2.3. Analytical method

Typically, we can represent spectral functions with fitting polynomials [6,18]. The accuracy and compactness are thus determined by the polynomial degree M. Unfortunately, when M becomes large ($M \geq 7$), a computational instability arises [5]. Therefore, the fitting polynomials have to be of low degree, which limits the potential in achieving accuracy. Also, this method is not very efficient because multiplying two spectral functions is O(M^2). Finally, although portable, it is not flexible. If we increase M for better accuracy, all the polynomial coefficients have to be recomputed.

2.4. Linear model

The key point of linear model is to express a spectrum as a linear combination of a set of basis functions $B_m(\lambda)$

$$S(\lambda) = \sum_{m=1}^{M} a_m B_m(\lambda), \tag{2.3}$$

so that spectra are represented using coefficients $a_1,\ldots, a_M$. In principle, any set of basis functions can be used as long as they are linearly independent. In practice, however, the basis functions are usually derived numerically in such a way that they can best represent all spectral functions for a specified domain [4,13,16,19], e.g., choosing the basis functions to minimize

$$\sum_{\{S(\lambda)\}} \int [S(\lambda) - \sum_{m=1}^{M} c_m B_m(\lambda)]^2 \, d\lambda, \tag{2.4}$$

where the summation is over all relevant spectral functions within the domain, including SPDs, reflectances and transmittances, and perhaps their combinations.

The linear model demonstrates advantages in both accuracy and compactness. It is also flexible because, once the basis functions are determined, one can use various numbers of the basis functions. However, this model is not efficient. First, computing the product of two spectral functions is O(M^2). Second, there is a computational cost in determining the basis functions. If we use a fixed set of basis functions for different scenes, the basis functions may possibly not work well for some spectra *outside* the spectral domain from which the basis functions are derived [4]. On the other hand, if we determine the basis functions every time for different scenes, the repetitive computations can be a significant bottleneck. Also, the model is not very portable because the basis functions are domain-dependent.

2.5. Summary

As shown in Table 1, the previous representation methods have different advantages but none of them meets all the criteria. In comparison, the composite model proposed in this paper can meet all the criteria.

Table 1: Comparing different spectral models.

	Color models	Point sampling	Analytical	Linear model	Composite model
Accuracy	No	Yes	No	Yes	Yes
Compactness	Yes	No	Yes	Yes	Yes
Efficiency	Yes	Yes	No	No	Yes
Portability	Yes	Yes	No	No	Yes
Flexibility	No	Yes	Yes	Yes	Yes

The color models, in spite that they are ubiquitous in software, are not sufficient to be the foundation for various applications because of their limited accuracy. For the same reason, the analytical method is not appropriate either. As for the point sampling method, its weakness in compactness cannot be eliminated because neighboring sampled points are inherently redundant. The linear model offers freedom in choosing the basis functions according to specific spectral domains, and this freedom could be applied successfully in representing either smooth or spiky spectral functions. However, it is this freedom that leads to the model's weaknesses in efficiency and portability. In applications, to use the linear model efficiently we must use a uniform format (the same basis functions) to represent both smooth and spiky spectral functions, but smooth and spiky spectra are inherently opposite to each other. Smooth ones are dominated by low-frequency components and are more effectively described in the *frequency* (inverse of wavelength) space, while spikes contain significant high-frequency components and are

more effectively described in the *wavelength* space. For this reason, the linear model has to rely on statistical and subjective calculations in choosing the basis functions numerically, but cannot benefit from using simple and neutral analytical functions (such as the Fouriers) as the basis functions.

3. New Method: the Composite Model

3.1. Basic concept

Fig. 1 shows examples of smooth and spiky spectral functions. The curves in (a) and (b) are reflectances of seven MacBeth Checkers [8,14] and they are smooth spectral functions. The spectrum in (c) is the SPD of the CIE "daylight" fluorescent lamp [22] and it is a spiky spectral function. Note that the spikes are actually 10 times higher, but for better visibility they are drawn with width of 10 nm while maintaining the same areas (or weights) of the spikes.

The composite model [20] is built on the idea of classifying all spectral functions into smooth and spiky categories, and representing them differently. A smooth spectrum can be expressed as a linear combination of a set of basis functions and represented through the corresponding coefficients. For a spiky spectral function, its smooth background (if any) can be represented likewise, and its discrete spikes can be represented through the locations and heights of the spikes.

If $S(\lambda)$ is an smooth spectrum, we can express it as a Fourier series over the visible range $[\lambda_{\min}, \lambda_{\max}]$:

$$S'(\lambda) = \sum_{n=0}^{\infty} \{a_n \cos[2\pi n(\lambda - \lambda_{\min})/L] + b_n \sin[2\pi n(\lambda - \lambda_{\min})/L]\}, \tag{3.1}$$

where $L = \lambda_{\max} - \lambda_{\min}$ and the coefficients are

$$\begin{cases} a_0 = \dfrac{1}{L}\displaystyle\int_{\lambda_{\min}}^{\lambda_{\max}} S(\lambda)d\lambda, \\ b_0 = 0, \\ a_n = \dfrac{2}{L}\displaystyle\int_{\lambda_{\min}}^{\lambda_{\max}} S(\lambda)\cos[2\pi n(\lambda - \lambda_{\min})/L]d\lambda, \quad n = 1,2...,\infty \\ b_n = \dfrac{2}{L}\displaystyle\int_{\lambda_{\min}}^{\lambda_{\max}} S(\lambda)\sin[2\pi n(\lambda - \lambda_{\min})/L]d\lambda, \quad n = 1,2...,\infty \end{cases} \tag{3.2}$$

For smooth $S(\lambda)$, we can ignore its high-frequency components, say those above N, so that $S(\lambda)$ is approximated by

$$S'(\lambda) = \sum_{n=0}^{N} \{a_n \cos[2\pi n(\lambda - \lambda_{\min})/L] + b_n \sin[2\pi n(\lambda - \lambda_{\min})/L]\}, \tag{3.3}$$

with $2N+1$ coefficients a_0, a_1, …, a_N and b_1, b_2, …, b_N. For a spiky spectrum, we decompose it as

$$S(\lambda) = S_{\text{smooth}}(\lambda) + S_{\text{spiky}}(\lambda) = S_{\text{smooth}}(\lambda) + \sum_{m=1}^{M} w_m \delta(\lambda - l_m), \tag{3.4}$$

where each spike is described by a delta function, which is a good approximation because all spikes are very narrow (typically under 1 nm). Thus, all spectral functions can be represented through scheme A:

(1) $a_0, a_1, ..., a_N$ and $b_1, ..., b_N$, lowest Fourier coefficients
(2) $(l_1, w_1), (l_2, w_2), ..., (l_M, w_M)$, spike locations and weights

(Scheme A)

3.2. Classifying smooth and spiky spectra

Table 2 summarizes the occurrences and conditions of smooth and spiky spectral functions [15, 21]. The smooth category includes the SPDs of thermal light sources

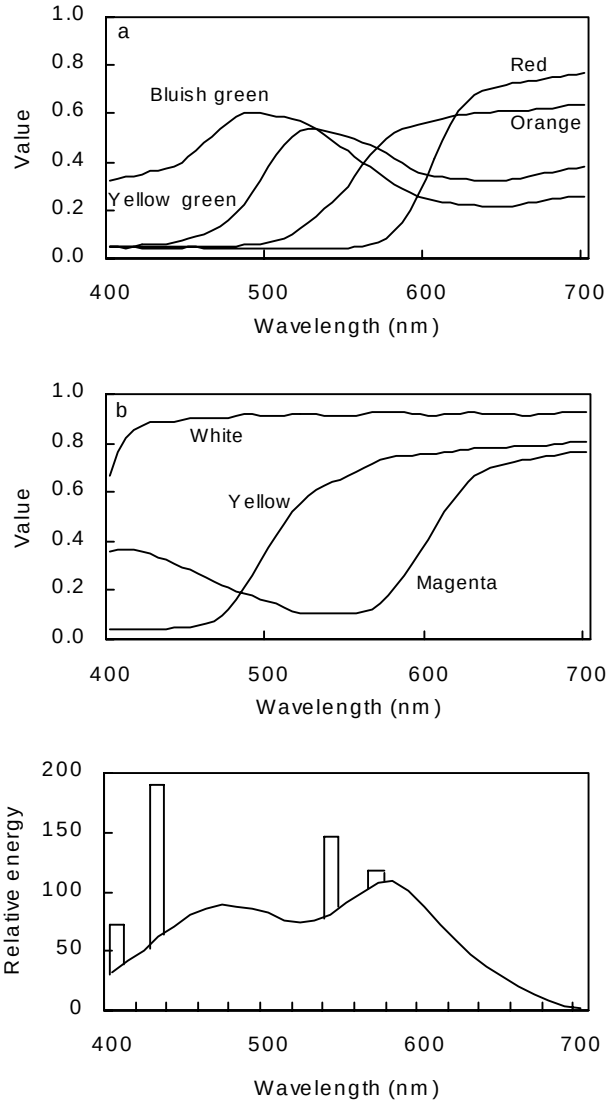


Fig. 1: Smooth and spiky spectral functions.

Table 2: The sources, generations, and conditions of smooth and spiky spectral functions.

Generations and conditions	Smooth spectral functions	Spiky spectral functions
Light sources	Thermal sources (sunlight, incandescent lamps, and ordinary flames)	Fluorescent light sources, mercury lamps and sodium lamps, and lasers
Reflectances and transmittances of materials	Non-fluorescent materials (the case of most natural materials)	Fluorescent materials, very few
After reflections and transmissions	The outgoing SPDs are smooth if the incident SPDs are smooth and the materials are non-fluorescent	The outgoing SPDs are spiky if the incident SPDs are spike, or the materials are fluorescent
Rayleigh scattering, interference, diffraction	The outgoing SPDs are smooth if the incident SPDs are smooth	The outgoing SPDs are spiky if the incident SPDs are spiky
After light dispersion		The result SPDs are monochromatic; may vanish if the incident SPDs are spiky but without smooth background

(sunlight, incandescent lamps, and ordinary flames), reflectances and transmittances of non-fluorescent materials, and their combinations. Smooth spectral functions usually remain within the category when subjected to physical optical events such as Rayleigh scattering, interference and diffraction. Typical examples of spiky spectral functions are the SPDs of lights from fluorescent sources, vapor discharge sources (mercury and sodium lamps), and lasers. Spiky SPDs that undergo reflections, transmissions, or the physical optical events remain spiky. Finally, when lights interact with *fluorescent* materials, even if the incident SPDs are smooth the outgoing SPDs can be spiky. However, if the incident SPDs are spiky but do not have a smooth background, the resultant SPDs after dispersion may vanish in certain outgoing directions.

Natural spectral functions seldom contain peaks as narrow as between 1 nm and 10 nm. This can be understood in terms of light coherence in physics. Suppose a spectrum is dominated by a single peak, then the coherence length of the corresponding light is give by [2]

$$\Delta l \cong \frac{\lambda_0^2}{\Delta\lambda}, \tag{3.5}$$

where λ_0 is the peak center and $\Delta\lambda$ is the peak width. According to this equation, the narrower the peak, the larger the coherence length. If λ_0 = 550 nm and $\Delta\lambda$ = 10 nm, then Δl ~ 30000 nm, over which the light should be regarded as coherent light. However, it is well known that common light sources (i.e., except lasers) only generate incoherent light, and incoherent light remains incoherent while interacting with regular materials. An interesting case is fluorescence, an effect in which light is absorbed and then re-radiated at longer wavelengths. This effect involves electronic transitions between discrete energy levels and thus the outgoing spectra are characterized with very narrow spikes, which can be treated separately in our composite model. Overall, it is practically correct to ignore the occurrences of peaks of widths between 1 nm and 10 nm, and to regard the smooth (without peaks of widths below 10 nm) and spiky (with peaks of widths below 1 nm) categories as being *complementary* to each other.

3.3. Proof: Sufficiency to represent all spectral functions through a few parameters

It is important to show that all spectral functions can be represented through a small number of parameters while still achieving sufficient accuracy. Since a spectrum can be decomposed into a smooth component and a collection of spikes, where the number of spikes is small (typically below 5) [22], we only need to show that the *smooth* component can be sufficiently represented through a small number of parameters.

We start by performing Fourier transformations for the three CIE color matching functions

$$x_k(\lambda) = \sum_{n=0}^{\infty} \{A_{n,k} \cos[2\pi n(\lambda - \lambda_{\min})/L] + B_{n,k} \sin[2\pi n(\lambda - \lambda_{\min})/L]\}, \tag{3.6}$$

where the coefficients are

$$\begin{cases} A_{0,k} = \dfrac{1}{L}\displaystyle\int_{\lambda_{\min}}^{\lambda_{\max}} x_k(\lambda)d\lambda, \\ B_{0,k} = 0, \\ A_{n,k} = \dfrac{2}{L}\displaystyle\int_{\lambda_{\min}}^{\lambda_{\max}} x_k(\lambda)\cos[2\pi n(\lambda - \lambda_{\min})/L]d\lambda, \quad n = 1,2...,\infty \\ B_{n,k} = \dfrac{2}{L}\displaystyle\int_{\lambda_{\min}}^{\lambda_{\max}} x_k(\lambda)\sin[2\pi n(\lambda - \lambda_{\min})/L]d\lambda, \quad n = 1,2...,\infty \end{cases} \tag{3.7}$$

Substituting Eqs. (3.1) and (3.6) into (2.1), the tristimulus values can be expressed as

$$X_k = L\sum_{n=1}^{\infty}\left[a_n A_{n,k} + b_n B_{n,k}\right], \quad k = 1,2,3 \tag{3.8}$$

This implies that if the high-frequency coefficients of either the color matching functions or the smooth component are negligible, $S(\lambda)$ can be represented with good accuracy through its low-frequency Fourier coefficients. Fig. 2 shows the first 21 Fourier coefficients of the three CIE color matching functions plotted against the coefficient ordinals (in the sequence a_0, a_1, b_1, …, a_N, b_N). Coefficients of color matching functions $x_1(\lambda)$, $x_2(\lambda)$, and $x_3(\lambda)$ are represented by solid circles, empty circles and empty squares, respectively. Their amplitudes become negligibly small when the ordinals are above 11.

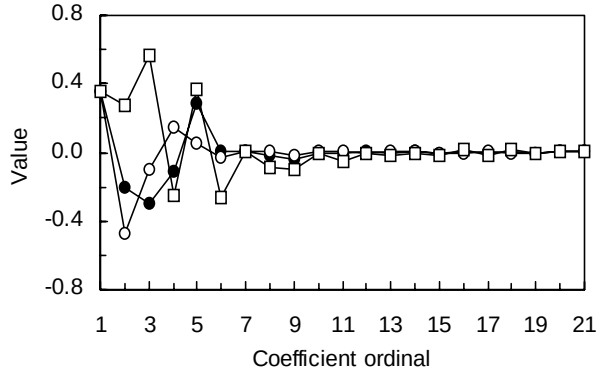


Fig. 2: Fourier coefficients of the CIE color matching functions against the coefficient ordinals.

Now let us characterize analytically precisely what number of frequencies must be retained in smooth spectra, when multiplying the color matching functions. First, we have found that the CIE color matching functions can be accurately approximated in terms of Gaussian functions

$$g(\lambda) = h\exp[-4(\ln 2)(\lambda - \lambda_c)^2 / w^2], \tag{3.9}$$

if the values of parameters h, λ_c, and w are chosen appropriately. Here, parameters h, λ_c, and w correspond to the height, center, and width (at half height) of a Gaussian function, respectively. Note that the CIE color matching function $x_1(\lambda)$ needs two Gaussian functions to fit it because it has two peaks. Fitting parameter values that we found are shown in Table 3. Using these values, the fitted curves are very close to the original CIE color matching functions, as shown in Fig. 3.

Table 3: Fitting values of the color matching functions.

small peak of $x_1(\lambda)$	large peak of $x_1(\lambda)$	$x_2(\lambda)$	$x_3(\lambda)$
$\lambda_{c,1a} = 445$	$\lambda_{c,1b} = 595$	$\lambda_{c,2} = 560$	$\lambda_{c,3} = 450$
$w_{1a} = 45$	$w_{1b} = 80$	$w_2 = 100$	$w_3 = 55$
$h_{1a} = 0.38$	$h_{1b} = 1.06$	$h_2 = 1.0$	$h_3 = 1.8$

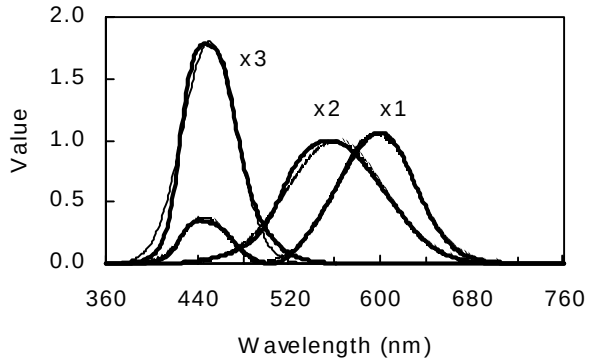


Fig. 3: CIE color matching functions (thick lines) and fitted Gaussian functions (thin lines).

Therefore, we represent the three color matching functions analytically via

$$\begin{cases} x_1(\lambda) = h_{1a}\exp[-4(\ln 2)(\lambda - \lambda_{c,1a})^2 / {w_{1a}}^2] \\ \qquad + h_{1b}\exp[-4(\ln 2)(\lambda - \lambda_{c,1b})^2 / {w_{1b}}^2], \\ x_2(\lambda) = h_2\exp[-4(\ln 2)(\lambda - \lambda_{c,2})^2 / {w_2}^2], \\ x_3(\lambda) = h_3\exp[-4(\ln 2)(\lambda - \lambda_{c,3})^2 / {w_3}^2] \end{cases} \tag{3.10}$$

Substituting into (3.7) and calculating the integrals (see the Appendix), we obtain

$$A_{n,k}, B_{n,k} \propto \exp[-\pi^2 n^2 {w_k}^2 / 4(\ln 2)L^2] \tag{3.11}$$

For smooth spectra, it is safe to assume

$$a_n \propto \frac{1}{n}, \quad b_n \propto \frac{1}{n}, \tag{3.12}$$

because the Fourier coefficients should vanish in speed no slower than $1/n$. Step functions and linear functions with non-zero slopes (triangular functions) are such worst cases. From (3.11) and (3.12), we obtain the ratios of high-frequency terms to the first term in Eq. (3.8)

$$\frac{|a_n A_{n,k}|}{a_0 A_{0,k}} \le \frac{4}{n} e^{-\pi^2 n^2 {w_k}^2 / 4(\ln 2)L^2} \tag{3.13}$$

With L = 300 nm and the parameter values given in Table 3, to make the above ratio less than 0.05, a sufficient condition is that n be not less than 5. In other words, if we keep the first 2×5+1 = 11 Fourier coefficients, the first ignored term only causes a relative error less than 0.05, and the following terms vanish rapidly (and they tend to cancel each other because they are related to the sine and cosine components of high frequencies). For most natural smooth spectral functions, the Fourier coefficients vanish much faster than $1/n$ and therefore the required number of parameters will be less than 11. Thus we have proved that we can represent all spectral functions through a small number of parameters while achieving sufficient accuracy. Note that a similar conclusion was obtained based on statistical studies [13], but ours is the first analytical proof.

3.4. Improving performance by re-sampling

However, scheme A is not very efficient. First, the cost for multiplying two smooth spectral functions is $O(N^2)$. Second, the multiplication between spiky and smooth spectra invokes evaluations of sine and cosine functions. Let a spiky spectrum be expressed as

$$\tilde{S}_1(\lambda) = S_1(\lambda) + \sum_{m=1}^{M} w_m \delta(\lambda - l_m), \tag{3.14}$$

where $S_1(\lambda)$ is smooth. If $S_2(\lambda)$ is a smooth spectrum, then

$$\begin{aligned} \tilde{S}_1(\lambda)S_2(\lambda) &= S_1(\lambda)S_2(\lambda) + S_2(\lambda)\sum_{m=1}^{M} w_m \delta(\lambda - l_m) \\ &= S_1(\lambda)S_2(\lambda) + \sum_{m=1}^{M} S_2(l_m) w_m \delta(\lambda - l_m) \end{aligned} \tag{3.15}$$

To compute $S_2(l_m)$, the value of $S_2(\lambda)$ at the location of spike m, we need to evaluate several sine and cosine functions using Eq. (3.3).

To improve the efficiency, we propose *re-sampling* the functions that are reconstructed with the Fourier coefficients in scheme A. Suppose $S'(\lambda)$ is reconstructed with $2N+1$ coefficients, then the highest frequency involved is N/L. According to the Shannon sampling theorem, the re-sampling interval in the visible range should be less than $L/(2N+1)$ to avoid information loss. To meet this requirement, we take $2N+2$ sampled points in the visible range. Therefore, a smooth spectral function can be represented through $s_0, s_1, \ldots, s_{2N+1}$ that are functional values of the reconstructed $S'(\lambda)$ at the $2N+2$ sampled wavelengths $\lambda_0, \lambda_1, \ldots, \lambda_{2N+1}$

$$\begin{cases} s_i = S'(\lambda_i) = S'(i\Delta\lambda + \lambda_{\min}), & i = 0,1,...,2N+1 \\ \Delta\lambda = L/(2N+1) \end{cases} \quad (3.16)$$

Thus spectral functions can be represented in "scheme B":

(1) $s_0, s_1, ..., s_{2N+1}$, values of reconstructed functions
(2) $(l_1, w_1), (l_2, w_2), ..., (l_M, w_M)$, spike locations and weights

(Scheme B)

In scheme B, products between smooth spectra can be obtained by multiplying the corresponding functional values, which involves $2N+2$ multiplications. When multiplying a spiky and a smooth function, the value of $S'(l_m)$ in Eq. (3.15) can be given by a linear interpolation

$$\begin{cases} S'(l_m) = \dfrac{(l_m - \lambda_i)}{\Delta\lambda} S'(\lambda_i) + \dfrac{(\lambda_{i+1} - l_m)}{\Delta\lambda} S'(\lambda_{i+1}) \\ \lambda_i \le l_m \le \lambda_{i+1} \end{cases} \quad (3.17)$$

where $m = 1,2,\ldots M$.

One may question the necessity of scheme A, as we could use scheme B based on direct sampling on the raw spectral data derived from measurements. However, we believe that scheme A is necessary. First, direct sampling on raw spectral data may result in significant errors due to noises or local peaks or valleys. In contrast, the re-sampling based on scheme A can mitigate such errors because the noises, peaks or valleys can be filtered out. Second, scheme A is useful for making tradeoffs between accuracy and compactness, by changing the number of representing parameters N for smooth spectra.

3.5. Representing concrete spectral functions

Table 4 displays the values of Fourier coefficients in schemes A and B for the SPDs of the CIE light sources A and C (smooth spectra), the fluorescent daylight (spiky with a smooth background), and a 90w low-pressure sodium lamp (spiky without smooth background).

In summary, the composite model satisfies all the representation criteria. In particular, scheme A can mitigate errors and is useful for making tradeoffs between accuracy and compactness. Scheme B is recommended for computations that involve many reflections and transmissions, because it can significantly improve performance.

Table 4 (I): The first 7 Fourier coefficients in scheme A.

	a0	a1	b1	a2	b2	a3	b3
Source A	191.9	2.5	-63.5	-1.7	-29.6	-2.5	-19.5
Source C	99.3	-5.4	15.7	-7.8	5.2	-7.0	-3.4
Daylight	65.4	-32.7	13.6	-6.5	20.4	-0.7	-3.8
Sodium lamp	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 4 (II): The spikes in the spectra.

	Locations and weights of all spikes
Source A	No spikes
Source C	No spikes
Daylight	(404.7, 50), (435.8, 132), (546.1, 66), (577.8, 14)
Sodium lamp	(589.0, 31.95)

Table 4 (III): Parameters in scheme B.

	s1	s2	s3	s4	s5	s6	s7	s8
Source A	91.6	9.6	50.7	52.6	87.8	98.4	136.3	143.6
Source C	77.5	111.5	127.5	109.8	99.6	103.7	89.7	88.4
Daylight	24.1	67.8	86.2	82.6	75.6	106.9	87.4	44.1
Sodium lamp	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

4. Applications in Realistic Image Synthesis

In this paper, we focus on applying the composite model to image rendering by ray tracing.[1] First, we show that the composite model does indeed capture important spectral information correctly, by comparing with an RGB model rendering of Gray Object Metamers 5 and 11, with reflectances shown in Fig. 4, metameric with respect to the 1931 CIE Standard Observer for source C [22]. Our rendering results show that a spectrum-based approach is crucial.

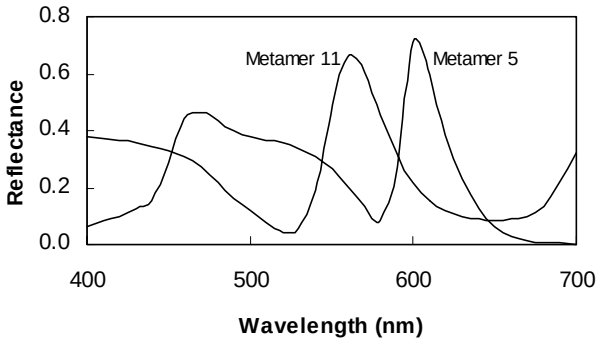


Fig. 4: Reflectances of Gray Object Colors Metamers 5 and 11 [22].

[1] The rendered color images with the composite model can be viewed at URL: http://www.cs.sfu.ca/~graphics/pubs/pmcvg99/.

Second, we compare the composite model with the point sampling method with respect to fluorescent illumination. We render images of seven spheres whose reflectances are of the "red", "orange", "yellow", "yellow-green", "bluish-green", "magenta", and "white" of the MacBeth Checkers (refer to Fig. 1), respectively. For the composite model, although there are noticeable errors when using 6 sampling points, using 11 re-sampling points is almost undistinguishable from the highly accurate result, and increasing the re-sampling points does not make much difference. In contract, in the point sampling method, the results are substantially different from the accurate one, even with 61 sampling points. Also, when increasing the number of the sampling points, the results are not stable but vibrating. The vibration is caused by whether or not some sampling points happen to be inside narrow spikes. This situation can be stabilized only when we use a very large number of sampling points so that it is guaranteed that there are a number of sampling points inside every spike. Because all spikes in the fluorescent SPD are made uniform with width of 10 nm, as shown in Fig. 1(c), we need at least 61 sampling points to stabilize the results. Note that we should not make the spike widths wider, as that will introduce errors due to the different distributions with respect to wavelength. This shows the advantage of the composite model over the point sampling method.

One special capability that the composite model offers is rendering images for monochromatic illuminations. Another special capability is rendering light dispersion. For this purpose, a regular ray tracer should be generalized to describe the refractive dependency on wavelength for dispersive materials. In other words, a single refractive ray in a regular ray tracer is now split into a series of dispersive rays with consecutive wavelength and refractive directions. Note that a split ray contains a spectral component in a narrow wavelength range, which can be represented naturally with the composite model. While effects of monochromatic illuminations and light dispersion are difficult (even impossible) to render with other spectral models, they are very easy to render with the composite model.

5. Future Directions and Conclusion

Much research effort has focused on generating realistic computer images based on full spectral information [10], such as using the point sampling method [3,9,11,12], the polynomial method [6,18], and the linear model [16,17]. However, due to their limitations, these previous methods are effective only under restrictive conditions but are not robust enough to be a universal basis for various problems. Since the composite model eliminates the limitations of previous methods, we expect that this model will push forward both research and applications in realistic image synthesis.

In above, we have shown the advantages of the composite model in rendering effects of spiky illuminations and light dispersion. Besides these effects, other modeling situations that require at least a partial full-spectrum mode may also benefit from the composite model idea. These include such physical optics phenomena as scattering, interference, diffraction, volume absorption, and light reflection from rough surfaces.

The composite model can be further improved in several ways. In our discussion above, all spikes are regarded as having zero widths and infinite heights (but their products are finite). This treatment can be generalized so that all spikes have finite widths and heights. With this generalization, the composite model can more accurately and flexibly describe spectral peaks with various widths and heights. Also, the range of wavelengths could be extended beyond the visible to include the ultra-violet region when a scene involves fluorescence. The composite model provides a basis to effectively represent spectra for fluorescence effects, which involve spectral spikes; then the spectral multiplication in Eq. (2.2) is not valid and a representation in form of a triangular matrix is required [7,22].

In conclusion, we have proposed the composite model to represent spectral functions based on decomposing smooth and spiky components. The smooth components can be expressed as a linear combination of a set of given basis functions and the spikes can be represented through their locations and heights. Specifically, we have proposed two schemes for representing the smooth components: scheme A using low-frequency Fourier coefficients, and scheme B using the re-sampled values of functions reconstructed with the Fourier coefficients. Scheme A provides appropriate high accuracy and flexibility, while scheme B is recommended for achieving performance. As the composite model meets all the representation criteria of accuracy, compactness, computational efficiency, portability, and flexibility, it possesses great application potential in various application areas.

Appendix A

With the color matching functions in Eq. (3.7) in terms of the Gaussians in (3.9), we can replace the integrals over the interval $[\lambda_{\min}, \lambda_{\max}]$ with $(-\infty, +\infty)$. This is a good approximation because the Gaussian functions are negligibly small outside range $[\lambda_{\min}, \lambda_{\min}]$. Thus, we obtain

$$\begin{cases} A_n = \dfrac{2h}{L}\displaystyle\int_{-\infty}^{+\infty} \exp^{-4(\ln 2)(\lambda-\lambda_c)^2/w} \cos\left[2\pi n(\lambda-\lambda_{\min})/L\right] d\lambda \\ B_n = \dfrac{2h}{L}\displaystyle\int_{-\infty}^{+\infty} \exp^{-4(\ln 2)(\lambda-\lambda_c)^2/w} \sin\left[2\pi n(\lambda-\lambda_{\min})/L\right] d\lambda \end{cases} \tag{A.1}$$

Note that we have dropped the index k. Now let us make variable change $\lambda = \lambda + \lambda_0$ and let $\alpha = 2(\ln 2)^{1/2}/w$, $\beta = 2\pi n/L$, $\gamma = 2h/L$, and $\lambda' = \lambda_c - \lambda_{\min}$. Thus we have

$$\begin{aligned} A_n &= \gamma \int_{-\infty}^{+\infty} e^{-\alpha^2\lambda^2} \cos[\beta(\lambda+\lambda')]d\lambda \\ &= \gamma \cos(\beta\lambda') \int_{-\infty}^{+\infty} e^{-\alpha^2\lambda^2} \cos(\beta\lambda)d\lambda \\ &\quad - \gamma \sin(\beta\lambda') \int_{-\infty}^{+\infty} e^{-\alpha^2\lambda^2} \sin(\beta\lambda)d\lambda \\ &= 2\gamma \cos(\beta\lambda') \int_{0}^{+\infty} e^{-\alpha^2\lambda^2} \cos(\beta\lambda)d\lambda \end{aligned} \tag{A.2}$$

Here we have used the fact that $\cos(\beta\lambda)$ is even while $\sin(\beta\lambda)$ is odd over $(-\infty, +\infty)$. Similarly,

$$\begin{aligned} B_n &= \gamma \int_{-\infty}^{+\infty} e^{-\alpha^2\lambda^2} \sin[\beta(\lambda+\lambda')]d\lambda \\ &= \gamma \sin(\beta\lambda') \int_{-\infty}^{+\infty} e^{-\alpha^2\lambda^2} \cos(\beta\lambda)d\lambda \\ &\quad + \gamma co(\beta\lambda') \int_{-\infty}^{+\infty} e^{-\alpha^2\lambda^2} \sin(\beta\lambda)d\lambda \\ &= 2\gamma \sin(\beta\lambda') \int_{0}^{+\infty} e^{-\alpha^2\lambda^2} \cos(\beta\lambda)d\lambda \end{aligned} \tag{A.3}$$

Now using the fact that

$$\int_0^{+\infty} e^{-\alpha^2\lambda^2} \cos(\beta\lambda)d\lambda = \frac{\sqrt{\pi}}{2\alpha} e^{-\beta^2/4\alpha^2}, \tag{A.4}$$

we obtain

$$\begin{cases} A_n = \dfrac{\sqrt{\pi}\gamma \cos(\beta\lambda')}{\alpha} e^{-\beta^2/4\alpha^2} \\ B_n = \dfrac{\sqrt{\pi}\gamma \sin(\beta\lambda')}{\alpha} e^{-\beta^2/4\alpha^2} \end{cases} \tag{A.5}$$

Substituting with the original parameters, we have

$$\begin{cases} A_n = \dfrac{\sqrt{\pi}hw}{\sqrt{\ln 2}L} \cos[2\pi n(\lambda_c - \lambda_{\min})/L] e^{-(\pi n w)^2/4(\ln 2)L^2} \\ B_n = \dfrac{\sqrt{\pi}hw}{\sqrt{\ln 2}L} \sin[2\pi n(\lambda_c - \lambda_{\min})/L] e^{-(\pi n w)^2/4(\ln 2)L^2} \end{cases} \tag{A.6}$$

Specifically for the first coefficient,

$$A_0 = \frac{\sqrt{\pi}hw}{2\sqrt{\ln 2}L} e^{-(\pi n w)^2/4(\ln 2)L^2} \tag{A.7}$$

Reference

[1] Carlos F. Borges, "Trichromatic approximation method for surface illumination," *J. Opt. Soc. Am.*, Vol. 8, No. 8, Aug. 1991, pp. 1319-1323.

[2] Max Born and Emil Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Pergamon Press, Oxford, 1975.

[3] Robert L. Cook and Kenneth E. Torrance, "A Reflection Model for Computer Graphics," *ACM TOG*, Vol. 1, No. 1, Jan. 1982, pp. 7-24.

[4] Mark Drew and Brian Funt"Natural Metamers," *CVGIP: Image Understanding*, Vol. 56, No. 2, Sep. 1992, pp. 139-151.

[5] G. E. Forsythe, "Generation and use of orthogonal polynomials for data fitting with a digital computer," *J. Soc. Ind. Appl. Math.*, Vol. 5, 1957, page 57.

[6] Robert Geist, Oliver Heim, and Stephen Junkins, "Color representation in virtual environments," *Color Res. Appl.* Vol. 21, No. 2, Apr. 1996, pp. 121-128.

[7] Andrew S. Glassner, "A model of fluorescence and phosphorescence," in *Proc. of the Fifth Eurographics Workshop on Rendering*, Stefan Haas, Stefan Muller, George Sakas, Peter Shirley, editors, Springer-Verlag, 1994, pp. 57-68.

[8] Andrew Glassner, *Principles of Digital Image Synthesis*, Morgan Kaufmann, San Francisco, 1995.

[9] Jay S. Gondek, Gray W. Meyer, Jonathan G. Newman, "Wavelength dependent reflection functions," *Computer Graphics, Proc. of ACM SIGGRAPH 94*, 1994, pp. 213-219.

[10] For a recent and comprehensive review, see Donald P. Greenberg, Kenneth E. Torrance, Peter Shirley, James Arvo, James A. Ferwerda, Sumanta Pattanaik, Eric P. F. Lafortune, Bruce Walter, Sing-Choong Foo, Ben Trunbore, "A framework for realistic image synthesis," *Computer Graphics, Proc. of ACM SIGGRAPH 97*, 1997, pp. 477-494.

[11] Roy A. Hall and Donald P. Greenberg, "A Testbed for Realistic Image Synthesis," *IEEE Computer Graphics & Appl.*, Vol. 3, No. 6, Nov. 1983, pp. 10-20.

[12] Roy A. Hall, *Illumination and Color in Computer Generated Imagery*, Springer-Verlag, New York, 1989.

[13] Laurence T. Maloney, "Evaluation of Linear Models of Surface Spectral Reflectance with Small Numbers of Parameters," *J. Opt. Soc. Am.*, Vol. 3, No. 10, Oct. 1986, pp. 1673-1683.

[14] C. S. McCamy, H. Marcus, and J. G. Davidson, "A color-rendition chart," *J. Appl. Photogr. Eng.*, Vol. 2, No. 3, 1976, pp. 95-99.

[15] Kurt Nassau, *The Physics and Chemistry of Color: The Fifteen Causes of Color*, John Wiley & Sons, New York, 1983.

[16] Mark S. Peercy, "Linear Color Representations in Full Spectral Rendering," *Computer Graphics, Proc. of SIGGRAPH 93*, ACM Press, New York, 1993, pp. 191-198.

[17] Mark S. Peercy, Daniel R. Baum, and Benjamin M. Zhu, "Linear color representations for efficient image synthesis," *Color Res. Appl.*, Vol. 21, No. 2 Apr. 1996, pp. 129-137.

[18] Maria Raso and Alain Fournier, "A piecewise polynomial approach to shading using spectral distributions," in *Graphics Interface 91*, 1991, pp. 40-46.

[19] Javier Romero, Antonio Garcia-Beltran, and Javier Hernandez-Andre, "Linear bases for representation of natural and artificial illuminants," *J. Opt. Soc. Am.*, Vol. 14, No. 5, May 1997, pp. 1007-1014.

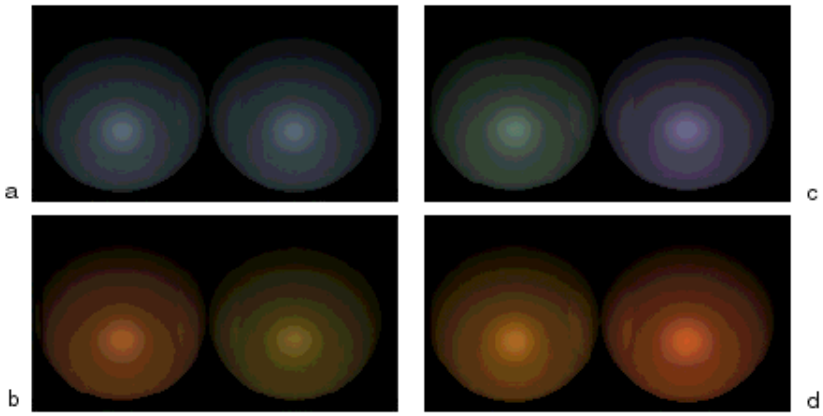
[20] For detailed discussion of the composite model, see Yinlong Sun, F. David Fracchia, and Mark S. Drew, "A composite model for representing spectral functions," Simon Fraser University, Technical Report SFU CMPT TR 1998-18, available at ftp://fas.sfu.ca/pub/cs/TR/1998/.

[21] Samuel J. Williamson and Herman Z. Cummins, *Light and Color in Nature and Art*, John Wiley and Sons, New York, 1983.

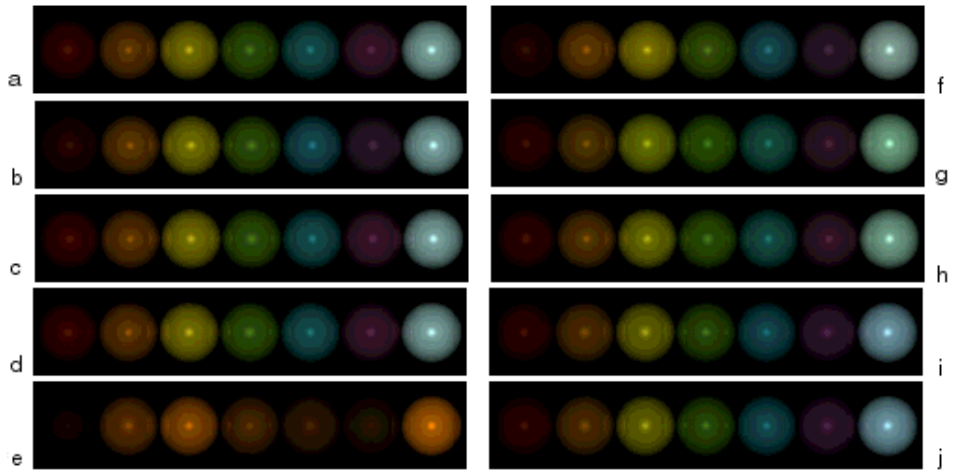
[22] G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, Second Edition, Wiley, New York, 1982.

"Representing Spectral Functions by a Composite Model of Smooth and Spiky Components for Efficient Full-Spectrum "

(Recommend: View this page with high-resolution monitor)

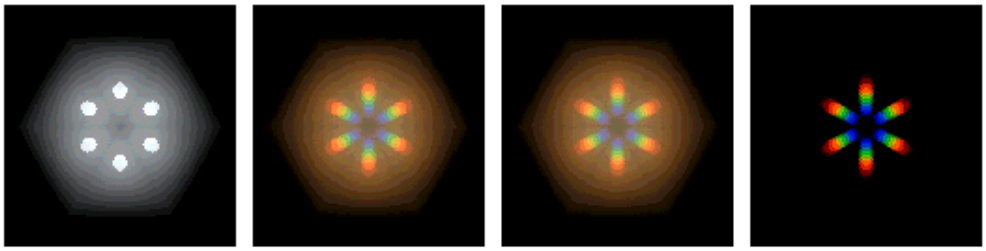


Comparison between the composite model and with an RGB model.
Images (a) and (b) are rendered with the composite model, and (c) and (d) with an RGB model. In every image, the left sphere has the reflectance of Metamer 5 and the right one of Metamer 11. The illuminant is source C for images (a) and (c) and source A for images (b) and (d). The metamerism is preserved in image (a), which is generated using the composite model. However, as expected, in the RGB-generated image (c) the two spheres are substantially different in color. These results show that a spectrum-based approach is crucial.



Comparison between the composite model and the point sampling method.
Images of seven spheres whose reflectances from left to right are the "red", "orange", "yellow", "yellow-green", "bluish-green", "magenta", and "white" of the MacBeth Checkers (refer to Fig. 1), respectively. Except for (e), illumination for all images is fluorescent daylight (refer to Fig. 1). Image (a), which is highly the accurate, is rendered with the point sampling method using 1201 sampled points with all spike widths of 1 nm. Images (b), (c)

and (d) are rendered with the composite model using 6, 11 and 21 sampling point for the smooth background, respectively. Images (f)-(j) are rendered with the point sampling model using 6, 11, 21, 31 and 61 sampling point, respectively. For the composite model, although there are noticeable errors when using 6 sampling points, i.e., the 1st and 6th spheres in image (b) are darker, image (c) using 11 re-sampling points is almost undistinguishable from the highly accurate image (a), and increasing to 21 re-sampling points does not make much difference. In contract, using the point sampling method, the results are substantially different from the accurate one, even with 61 sampling points. Also, when increasing the number of the sampling points, the results are not stable but vibrating (note the colors of the white sphere). In fact, the white sphere in image (j) appearing "bluer" is caused by the larger spike width used (i.e., 10 nm). These results clearly demonstrate the advantage of the composite model over the point sampling method.



Light dispersion rendered with the composite model.
The object is a transparent hexagon pyramid (i.e., the bottom face is a regular hexagon). The camera points to the pyramid from its top, while a white surface light source is located below the pyramid. The pyramid is non-dispersive in image (a) and dispersive for the rest. The pyramid surface is not perfectly smooth for images (a), (b) and (c), which are the reason for the non-highlight shades, but perfectly smooth for image (d), which has no non-highlight shades. Image (b) has 6 dispersive spawned rays while image (c) has 10, which results in the smoothness of the dispersive color bands in image (c). Note that the yellowish non-highlight colors in image (b) and (c) are due to the overall effect that long-wavelength components are more likely to penetrate.