Colour

The visible spectrum of light corresponds wavelengths roughly from 400 to 700 nm.
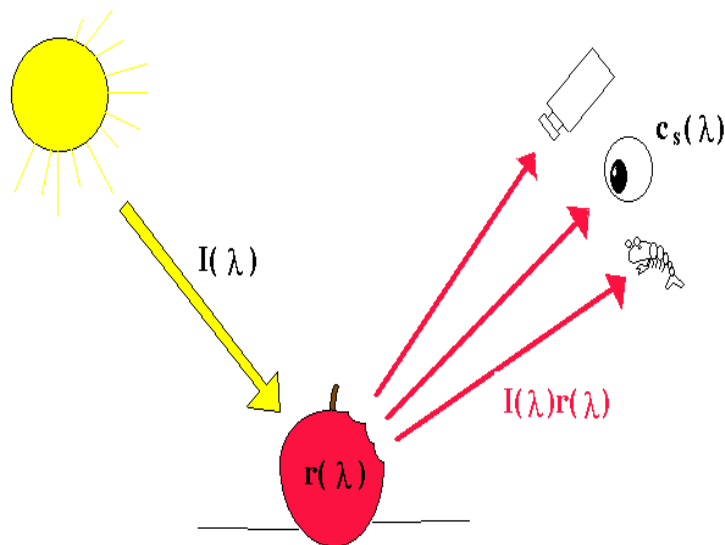
Visible Spectrum

(Apprx) Wavelength (nm)
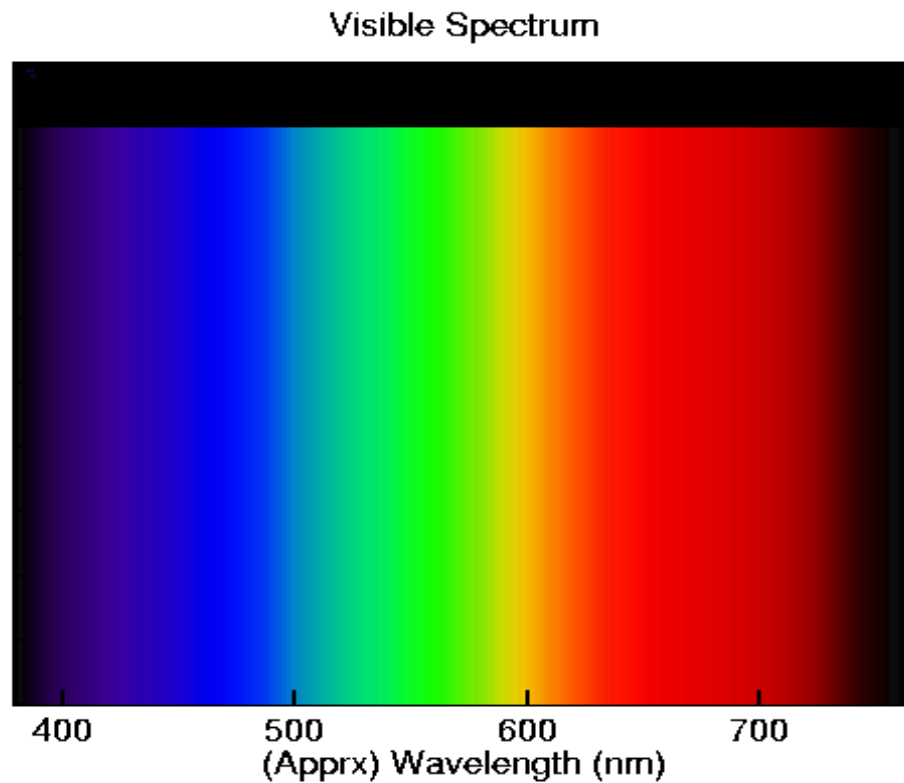
The image above is not colour calibrated. But it offers a rough idea of the correspondence between colours and wavelengths.

Reference: Matlab colourTutorial.m in utvisToolbox.

Colour Image Formation

The (R, G, B)-response of a pixel in an eye or camera is the combined result of four components:

- The illuminant spectral density, $I(\lambda)$,

- the reflectance function of surfaces, $r(\lambda)$,

- the geometry of the scene (eg. the surface orientation),

- the spectral sensitivities, $S_\mu(\lambda)$, of the photo-receptors in the observer's eye or camera.

We briefly discuss each of these terms below.

Spectral Power Distribution of Light

We describe light in terms of the power present at each wavelength in the visible spectrum, say $I(\lambda)$.

Here $I(\lambda)$ is called the spectral power distribution (SPD). It is measured in units of Watts (i.e. power) per unit wavelength (i.e. λ), per unit cross-sectional area.

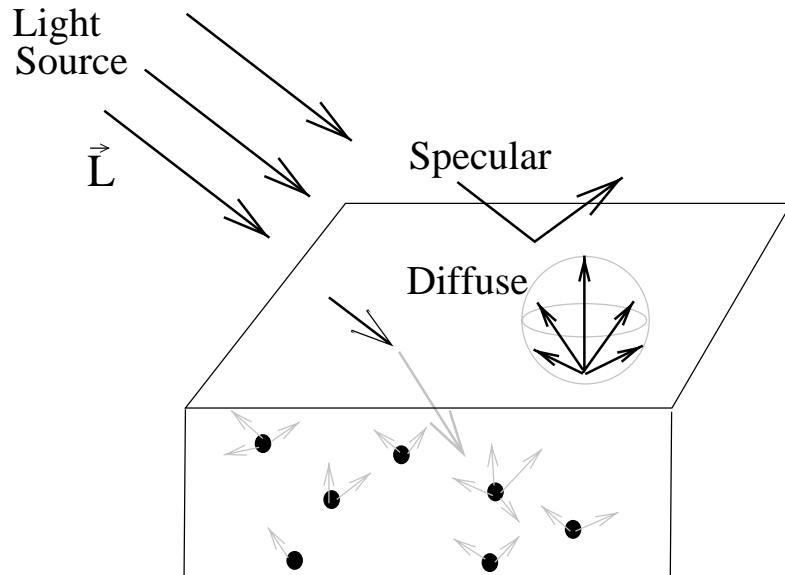The figure below shows typical SPDs for daylight.
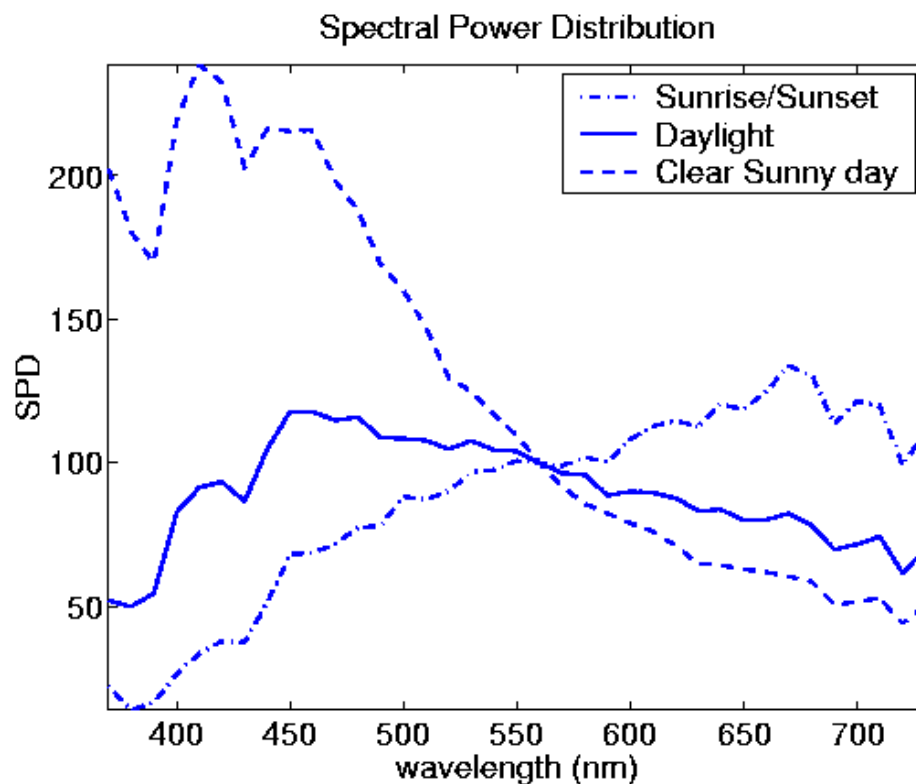
These have been normalized to 100 at the wavelength 550nm.

Notice the variation in the proportion of blue (short wavelength) and red (long wavelength) light at different times of day.

Surface Reflectance

Two types of reflectance:

- **Specular Reflectance**: Reflectance from the surface, in the "mirror reflection" direction. For non-metals the spectral distribution of this reflected light is roughly proportional to the SPD of the incident light.

- **Diffuse Reflectance**: Light is absorbed and re-emitted from the body, scattering in all directions. The spectral distribution of the reflected light depends on the pigmentation of the object.

Surface Reflectance (Cont.)

For incoming light travelling in the direction $\vec{L}$ and hitting a surface patch, consider the light diffusely reflected to a viewer in the direction $\vec{V}$.
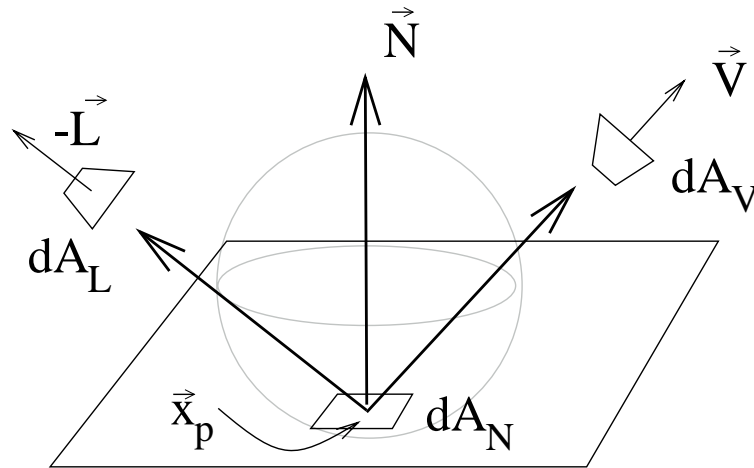

The SPD for the reflected light $I_r(\lambda)$ can be modelled as

$$I_r(\lambda) = r(\lambda) \max(-\vec{N} \cdot \vec{L}, 0) I(\lambda)$$

- $I(\lambda)$ is the SPD for the incoming light, arriving in direction $\vec{L}$,

- $I_r(\lambda)$ is the SPD for the reflected light (for simplicity we are ignoring an additional dependence on the solid angle of the scattered light),

- $\vec{N}$ is the surface normal,

- $r(\lambda)$ is the reflectance distribution for the surface,

- $-\vec{N} \cdot \vec{L}$ causes shading as the surface tilts away from the incoming light (if $-\vec{N} \cdot \vec{L}$ is negative then the surface faces away from the light, i.e. it is in shadow).

Munsell Chips

The Munsell set is a large collection of calibrated colour chips (as in paint 'chips'), the reflectances $r(\lambda)$ are systematically chosen to span a wide range of possibilities.

Various Munsell Reflectances
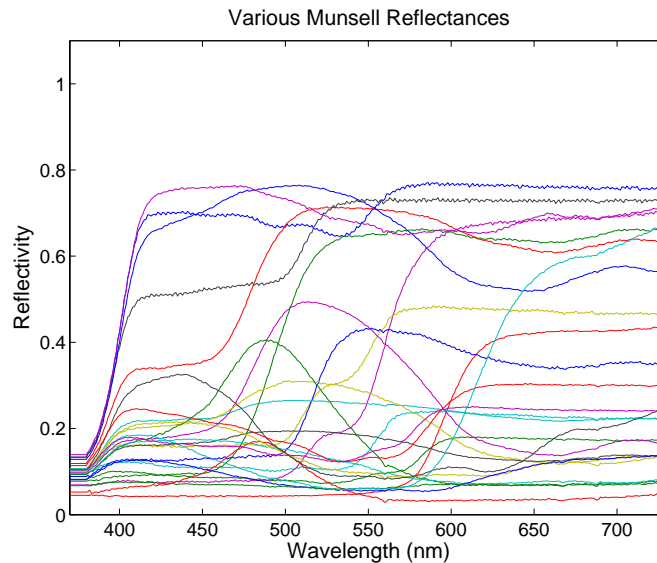
Munsell chips are named for their perceived colour, the colour is specified by three parameters:

- **Hue:** Specifies the colour name (i.e. R, Y, G, B, P, . . .).

- **Lightness:** Indicates the brightness of the chip.

- **Saturation:** How rich, or different from white.

Spectral Sensitivities

The pixel response is a function of the energy absorbed by a pixel. The absorbed energy is

$$e_\mu = C_T \int_0^\infty S_\mu(\lambda) I_r(\lambda) d\lambda \ \text{ for } \mu = R, G, B.$$

Here $I_r(\lambda, \vec{x}_I, \vec{n}_I)$ is the SPD for the incident light arriving from the scene. Also $S_\mu(\lambda)$ is the spectral sensitivity of the μth colour sensor, and C_T is a constant (eg. 1/(shutter speed)).
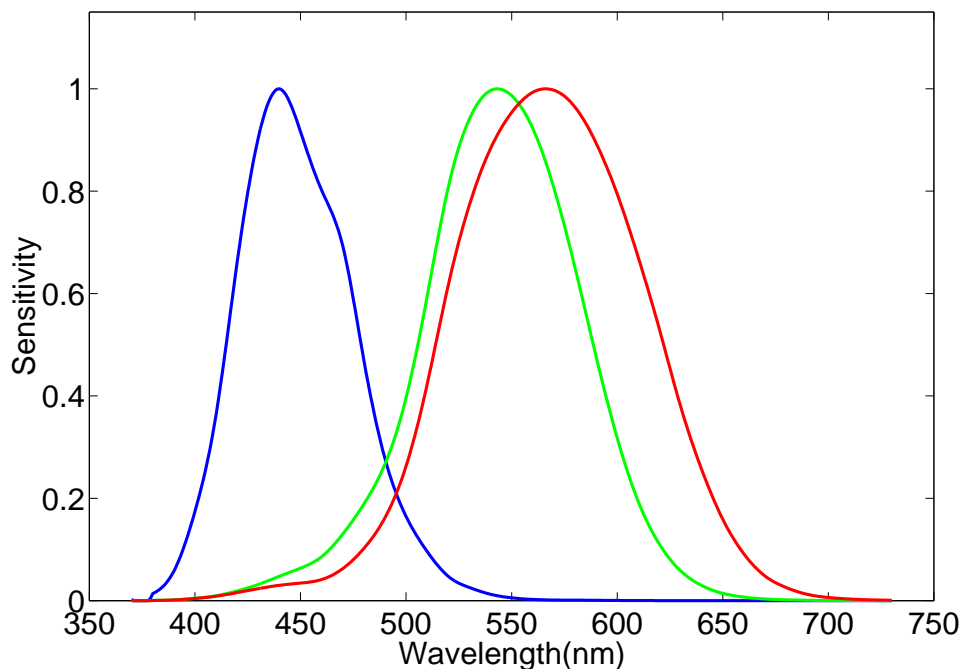
Colour images are formed (typically) using three spectral sensitivities, say $\mu = R, G, B$ for the 'red', 'green' and 'blue' channel. The normalized spectral sensitivities in the human retina are plotted above.

Absorbed Energy to Pixel Response
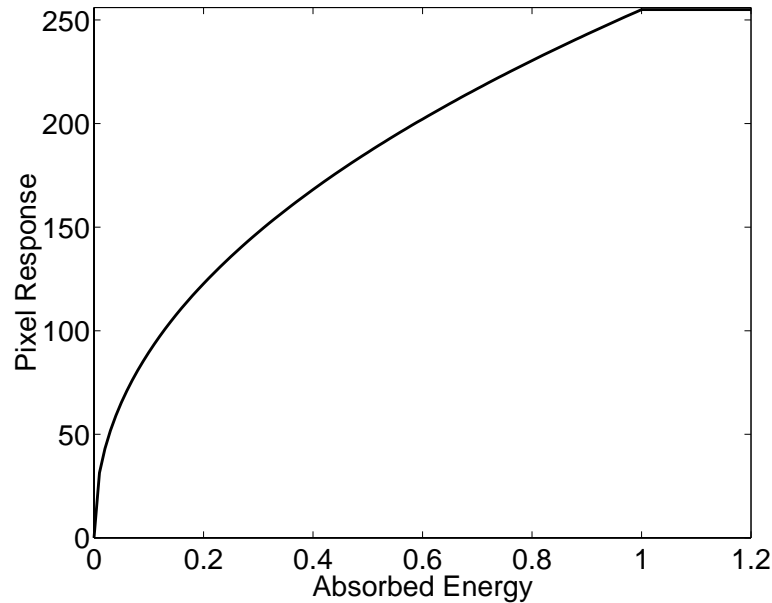
Gamma Correction. Finally, the absorbed energy e_μ is converted to a quantized pixel response, say r_μ, through a nonlinear function called a gamma correction,

$$r_\mu = \min(\beta \, [e_\mu]^{\frac{1}{\gamma}} , 255), \quad \text{for } \mu = R, G, B.$$

The value of γ can vary, values between 2 and 3 are common. This response r_μ is (typically) cropped to [0,255] and quantized to 8 bits.

Conversely, when an image is displayed, due to nonlinearities in the display process the brightness b_μ of a pixel in a particular colour channel is related to the pixel response r_μ through (roughly) the inverse relationship

$$b_\mu = [r_\mu / \beta]^\gamma , \quad \text{for } \mu = R, G, B.$$

The End Result is a Colour Image

colourTutorial.m answers the following questions (and more!):

- What are the typical values for $I(\lambda)$ (daylight,sunny day,sunrise/sunset) and $r(\lambda)$ (Munsell/paint chips)? Together these generate typical SPDs of reflected light $I_r(\lambda) \sim r(\lambda) \times I(\lambda)$.

- What are the spectral sensitivities for an average human? Three types of sensors: long(R), medium(G), and short(B) wavelengths.

- What are metamers?

- What are CIE XYZ and xy-colour coordinates?

- Why are only 3 colour channels used?

Metamers

Colour Matching Principle. If two light SPDs $I_1(\lambda)$ and $I_2(\lambda)$ cause the same energy absorptions

$$e_\mu = C_T \int_0^\infty S_\mu(\lambda) I_j(\lambda) d\lambda, \quad \text{for } \mu = R, G, B, \tag{1}$$

then they are perceptually indistinguishable. Note e_μ **does not depend** on j.

Colour Metamers. Two such light sources are said to be metamers.

This is closely related to our previous discussion of sampling continous signals:

- The SPD $I_r(\lambda)$ of the received light is the continuous signal,

- The broad sensor spectral sensitivities $S_\mu(\lambda)$ provide the preblur before sampling,

- The different sensors provide (only!) three discrete samples,

- Metamers are different continuous signals $I_1(\lambda)$ and $I_2(\lambda)$ aliased to the same absorbed energies $\vec{e} = (e_R, e_G, e_B)^T$ (see p.17 for an example).

CIE X,Y,Z Colour Matching Functions

The International Commission on Illumination (CIE) has carefully calibrated 2 and 3D colour spaces to specify perceptually identical chromatic stimuli.

The CIE colour matching functions $X(\lambda)$, $Y(\lambda)$, and $Z(\lambda)$ are three linear combinations of average human sensor sensitivities $S_\mu(\lambda)$, $\mu = R, G, B$.
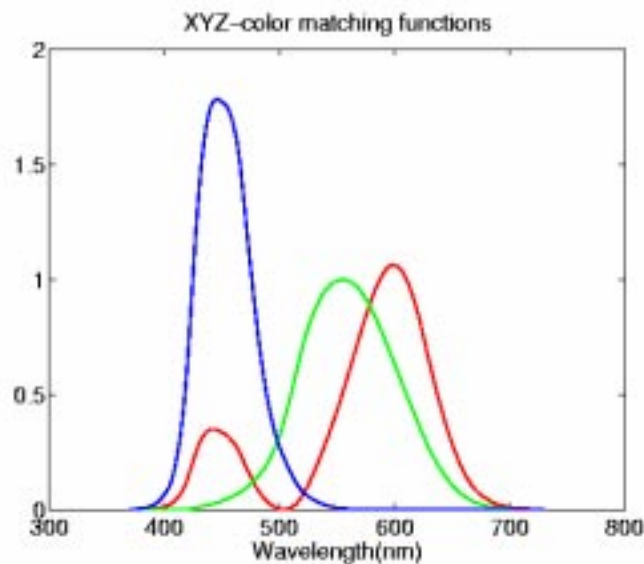
The (X, Y, Z) coordinates for a given SPD $I(\lambda)$ are just the integrals of these colour matching functions with $I(\lambda)$, i.e. $X = \int_0^\infty X(\lambda)I(\lambda)d\lambda$ and similarly for Y and Z.
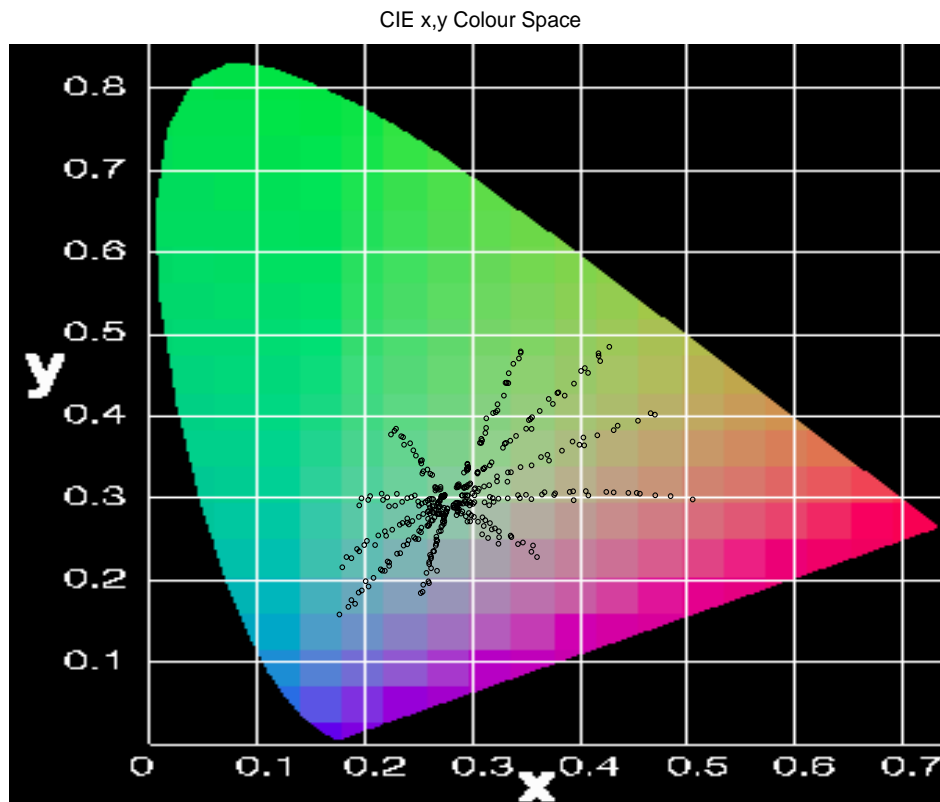
Two lights $I_1(\lambda)$ and $I_2(\lambda)$ with the same (X, Y, Z) coordinates will produce the same energy absorptions in the three colour sensors, and will therefore be metamers.

CIE xy-Colour Coordinates

Dividing the X and Y coords by the sum of the X,Y, and Z coords, normalizes the colour coordinates by the total brightness.

$$x = X/(X + Y + Z),$$
$$y = Y/(X + Y + Z).$$

CIE x,y Colour Space

Points along the curved boundary correspond to monochromatic stimuli (as on page 1 of these notes). The plotted points are various Munsell chips illuminated by a standard daylight. A white chip would appear at the center of the star pattern (near $(x, y) = (0.3, 0.3)$). Only hue and saturation are represented in this space.

Approximate CIE Coordinates

We can approximate the integrals using discrete sums

$$X \equiv \int_0^\infty X(\lambda) I_r(\lambda) d\lambda \approx \vec{X}^T \vec{I}_r \Delta_\lambda,$$

$$Y \equiv \int_0^\infty Y(\lambda) I_r(\lambda) d\lambda \approx \vec{Y}^T \vec{I}_r \Delta_\lambda,$$

$$Z \equiv \int_0^\infty Z(\lambda) I_r(\lambda) d\lambda \approx \vec{Z}^T \vec{I}_r \Delta_\lambda,$$

where $\vec{I}_r, \vec{X}, \vec{Y}, \vec{Z}$ are the corresponding functions $I_r(\lambda), X(\lambda), Y(\lambda),$ $Z(\lambda)$ evaluated at the discrete values $\lambda_1, \lambda_2, \ldots, \lambda_N,$ and $\Delta_\lambda = \lambda_{k+1} - \lambda_k$. For example, $\vec{I}_r \equiv (I_r(\lambda_1), I_r(\lambda_2), \ldots, I_r(\lambda_N))^T$.

Then the CIE X,Y,Z coordinates of the reflected light $I_r(\lambda)$ are (approximately) given by

$$\vec{c} \equiv \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = A\vec{I}_r, \quad \text{where } A = \Delta_\lambda \begin{pmatrix} \vec{X}^T \\ \vec{Y}^T \\ \vec{Z}^T \end{pmatrix} \in \mathcal{R}^{3 \times N}. \quad (2)$$

With this approximation, two lights $\vec{I}_1$ and $\vec{I}_2$ are metamers iff

$$\vec{c} = A\vec{I}_1 = A\vec{I}_2,$$

which we can easily check.

Colour Gamut

Consider a CRT or LCD display with three colour channels. For simplicity we will assume that the light generated by each pixel depends linearly on the coefficients $\vec{p} = (p_R, p_G, p_B)^T$ for each of the colour channels. Then the displayed SPD is

$$I_d(\lambda) = \sum_{\mu \in \{R,G,B\}} D_\mu(\lambda) p_\mu. \tag{3}$$

Here $D_\mu(\lambda)$ is the SPD for the μth colour channel of the display, and $p_\mu(\vec{x}) \in [0, 1]$ is the μth channel of the colour image $(I_R(\vec{x}), I_G(\vec{x}), I_B(\vec{x}))$ normalized to be between 0 and 1 (i.e. $p_\mu(\vec{x}) = I_\mu(\vec{x})/255$).

Equation (3) maps the unit cube in $\vec{p}$-space to SPD functions. This set of SPDs is called the **colour gamut** for the display. It represents the range of SPDs that the display can generate.
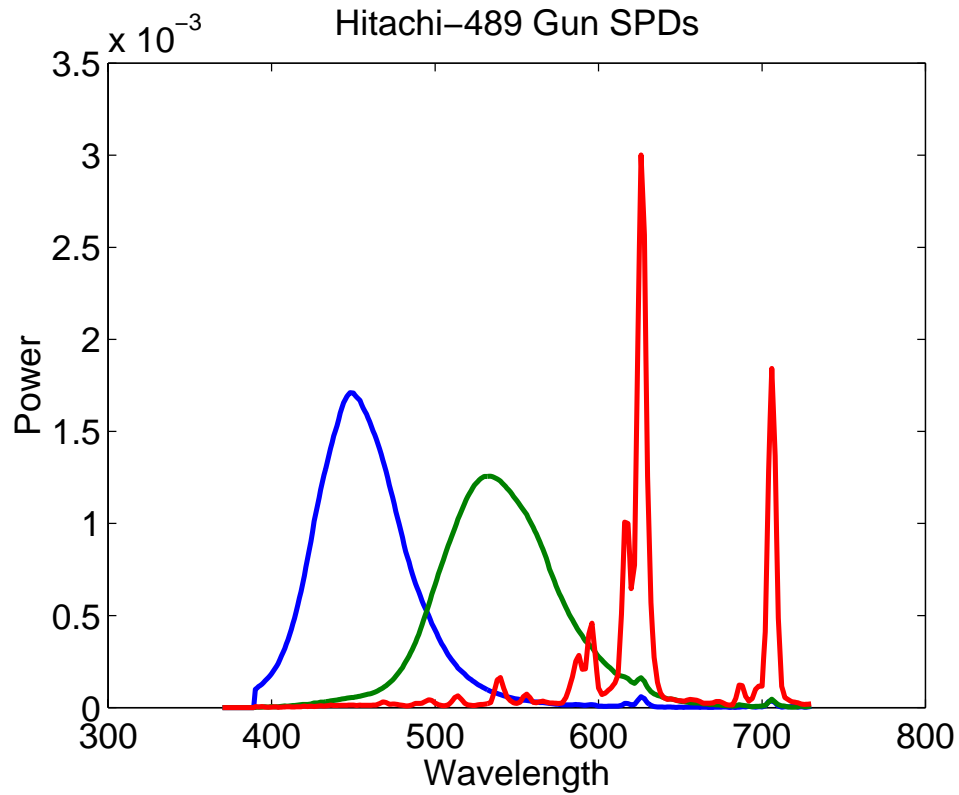
In terms of matrix-vector notation, we can write equation (3) as

$$\vec{I}_d = D\vec{p}, \ \ \text{for } D = \left(\vec{D}_R, \vec{D}_G, \vec{D}_B, \right) \in \mathcal{R}^{N \times 3}. \tag{4}$$

where, as above, each of these vectors corresponds to the functions $I_d(\lambda)$, $D_R(\lambda)$, $D_G(\lambda)$, $D_B(\lambda)$, sampled at $\lambda_1, \lambda_2, \ldots, \lambda_N$.

Example Display SPDs

For example, the colour gun SPDs for one display (a Hitachi model 849) are shown below:

Hitachi−489 Gun SPDs

Power vs. Wavelength

Colour Gamut in x, y-coordinates

When the colour gamut is plotted in CIE x, y-coordinates, the result is a triangle.

CIE x,y Colour Space

This triangle specifies the range of perceptually distinct lights that the Hitachi display can generate. Lights having (x, y) coordinates outside this triangle cannot be displayed (eg. any monochromatic light).

Colour Matching

Suppose we wish to display a particular "colour". As far as human perception is concerned, we can specify the desired display colour by its (X, Y, Z)-coordinates $\vec{c}$. Given $\vec{c}$, we need to choose coefficients $\vec{p}$ for the three colour guns such that the display SPD $\vec{I}_d = D\vec{p}$ satisfies

$$\vec{c} = A\vec{I}_d = AD\vec{p}, \quad \text{that is } \vec{p} = (AD)^{-1}\vec{c}. \tag{5}$$

Here A and D are $3 \times N$ and $N \times 3$ matrices, respectively, so AD is a 3×3 matrix. The resulting coefficients of $\vec{p}$ may need to be cropped to the interval $[0, 1]$.

Munsell Stimuli (g) and Display Metamer (r)
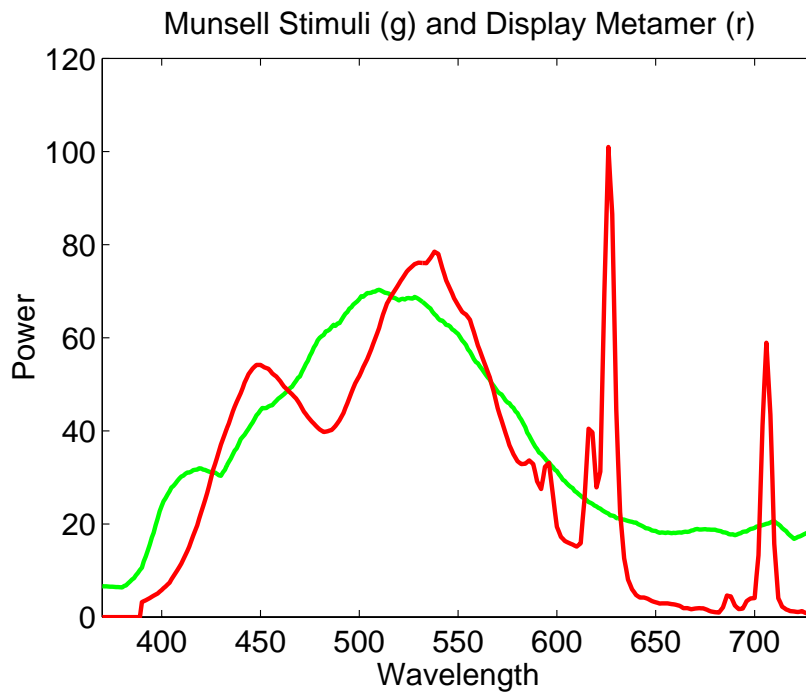
The display SPD $\vec{I}_d = D\vec{p}$ (red curve, above) is a linear combination of the previous Hitachi guns, and is a metamer for the SPD produced by a Munsell chip illuminated with a standard daylight (green curve).

Why did Nature Choose Only 3 Colour Channels?

Consider the "natural" stimuli

$$I_m(\lambda) = L(\lambda)R_m(\lambda) \tag{6}$$

where $L(\lambda)$ is a natural daylight SPD and $R_m(\lambda)$ is the reflectance function for the mth Munsell chip. This provides a large set of sample stimuli, namely $\{I_m(\lambda)\}_{m=1}^{M}$ for $M = 1269$.

What is the **effective dimension** of this set of natural stimuli?

For example, suppose we can write each $\vec{I}_m$ (i.e. $I_m(\lambda)$ sampled at $\lambda = \lambda_1, \lambda_2, \ldots, \lambda_N$) as a mean vector $\vec{I}_0$ plus a linear combination of K orthonormal basis vectors $\vec{U}_k$, plus a small error $\vec{e}_m$. That is,

$$\vec{I}_m = \vec{I}_0 + \sum_{k=1}^{K} \vec{U}_k a_{k,m} + \vec{e}_m.$$

If we were willing to ignore the small errors $\vec{e}_m$, then we could say that the data set $\{I_m(\lambda)\}_{m=1}^{M}$ effectively had dimension (at most) K.

The **singular value decomposition (SVD)** of a matrix can be used to compute the effective dimension K along with an appropriate set of basis directions $\{\vec{U}_k\}_{k=1}^{K}$.

Singular Value Decomposition

Let E be the matrix formed from the difference of each stimuli from the mean $\vec{I}_0 = \frac{1}{M} \sum_{m=1}^{M} \vec{I}_m$, that is

$$E \equiv (\vec{I}_1 - \vec{I}_0, \vec{I}_2 - \vec{I}_0, \ldots, \vec{I}_M - \vec{I}_0) \in \mathcal{R}^{N \times M}. \qquad (7)$$

In colourTutorial.m we sample the visible spectrum of wavelengths 370nm to 730nm with a spacing of 1nm, so the number of samples is $N = 361$. Also, the number of Munsell chips is $M = 1269$.

The SVD is a matrix factorization of the form

$$E = U \Sigma V^T \qquad (8)$$

where U is a $N \times N$ orthogonal matrix, Σ is a $N \times M$ matrix which is zero except along the main diagonal, with $\Sigma_{n,n} = \sigma_n$ for $n = 1, \ldots, \min(N, M)$, and V is a $M \times M$ orthogonal matrix. Here the **singular values** σ_n are non-negative and are sorted in decreasing order (i.e. $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N \geq 0$).

Matlab provides a built-in SVD function. The SVD is a well-behaved (numerically stable) matrix factorization with many applications.

Low-Dimensional Approximations

Consider the K-dimensional approximation of the signals formed using only the first K basis directions in U. That is,

$$\vec{I}_m = \vec{I}_0 + \sum_{k=1}^{K} \vec{U}_k a_{k,m} + \vec{e}_m, \tag{9}$$

where $\vec{U}_k$ is the kth column of the matrix U in the SVD (8) of E.

Since the columns of U are orthonormal, it follows that the appropriate choice for the coefficient $a_{k,m}$ is simply

$$a_{k,m} = \vec{U}_k^T (\vec{I}_m - \vec{I}_0). \tag{10}$$

This choice minimizes the Euclidean norm of the error, $||\vec{e}_m||$. It can be shown that sum of the squared errors is then

$$SSE \equiv \sum_{m=1}^{M} ||\vec{e}_m||^2 = \sum_{k=K+1}^{N} \sigma_k^2. \tag{11}$$

Moreover, an important property of the SVD is that this is **minimum possible sum of squared errors for any choice of K basis vectors $\vec{U}_k$.** Thus the SVD gives us the best possible K-dimensional representation for the data set, for each choice of K.

This is called a **principal coordinate analysis (PCA)** of the data set. The vectors $\vec{U}_k$ are the PCA basis functions.

Fraction of Variance Plot

The total variance of the data set is

$$V \equiv \sum_{m=1}^{M} ||\vec{I}_m - \vec{I}_0||^2 = \sum_{n=1}^{N} \sigma_n^2. \tag{12}$$

The right hand side here is a special case of (11) using a zero dimensional approximation (i.e. $K = 0$), with the error $\vec{e}_m = \vec{I}_m - \vec{I}_0$. We can normalize the sum of squared errors by the total variance,

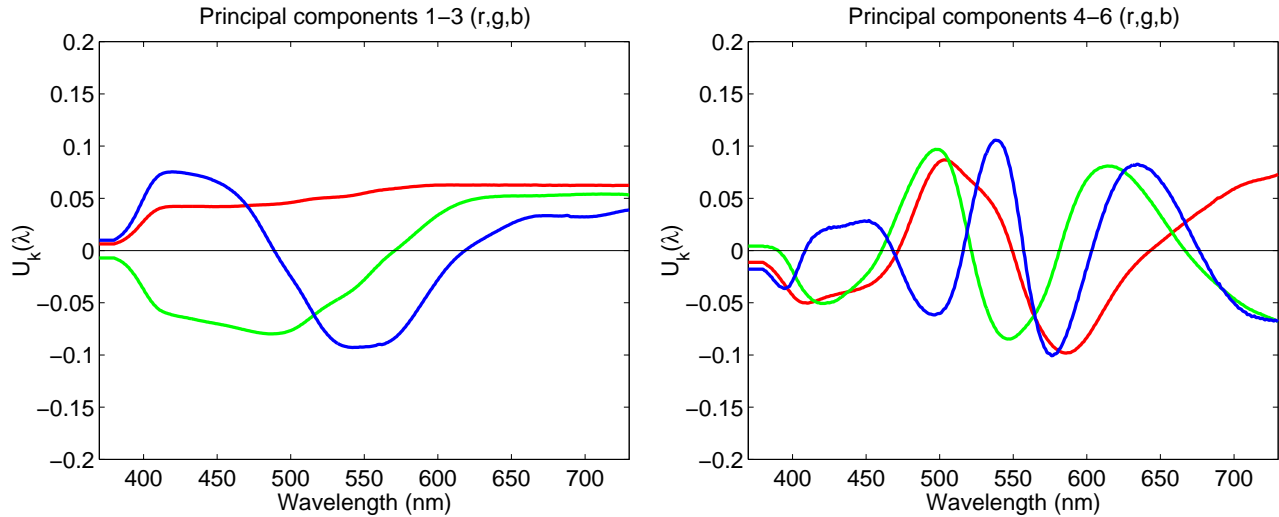$$Q_K \equiv \frac{1}{V} \sum_{k=K+1}^{N} \sigma_k^2, \tag{13}$$

and plot the proportion of variance accounted for, namely $1 - Q_K$.

Proportion of variance accounted for using dim K.
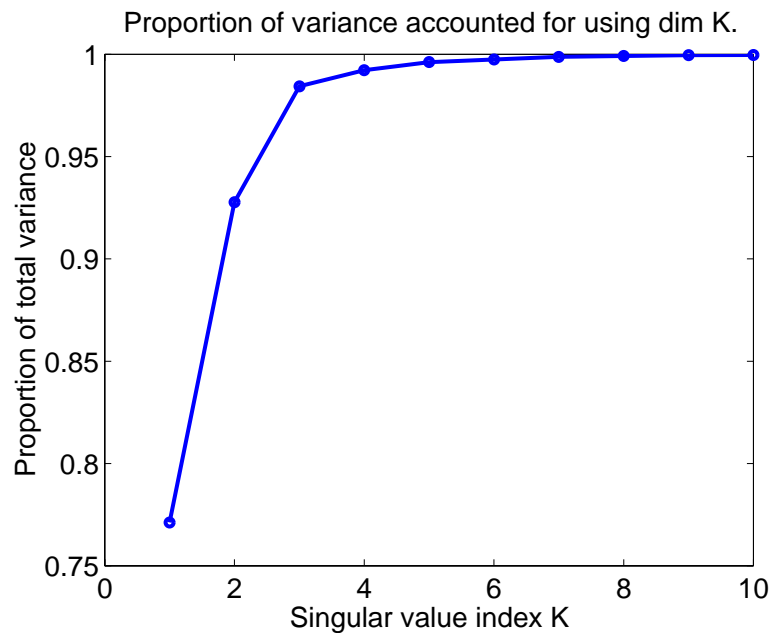
Conclusion. Over 98% of the variance in the Munsell reflectance data set is explained by a 3-dimensional basis!

PCA Basis Vectors

The first six basis vectors $\vec{U}_k$ for the matrix E are plotted below:

- The first principal component $U_1(\lambda)$ represents overall brightness.

- The second and third principal components $U_2(\lambda)$, $U_3(\lambda)$ represent (roughly) yellow-blue and green-purple variations, respectively.

- Higher order principal components represent more rapid variations of the SPD in λ.

- This is very roughly similar to the sines and cosines of Fourier basis functions.

- We saw the first 3 basis vectors account for more than 98% of the signal. These can be adequately sampled using the human spectral sensitivities (see p.7).

- The remaining 2% of the variance will be aliased.

The Munsell Chips in PCA Coordinates

Recall the SPD for the mth Munsell chip illuminated by a standard daylight is expanded in terms of the PCA basis as follows,

$$\vec{I}_m = \vec{I}_0 + \sum_{k=1}^{K} \vec{U}_k a_{k,m} + \vec{e}_m.$$

Scatter plots of the coefficients $a_{k,m}$ are shown below:
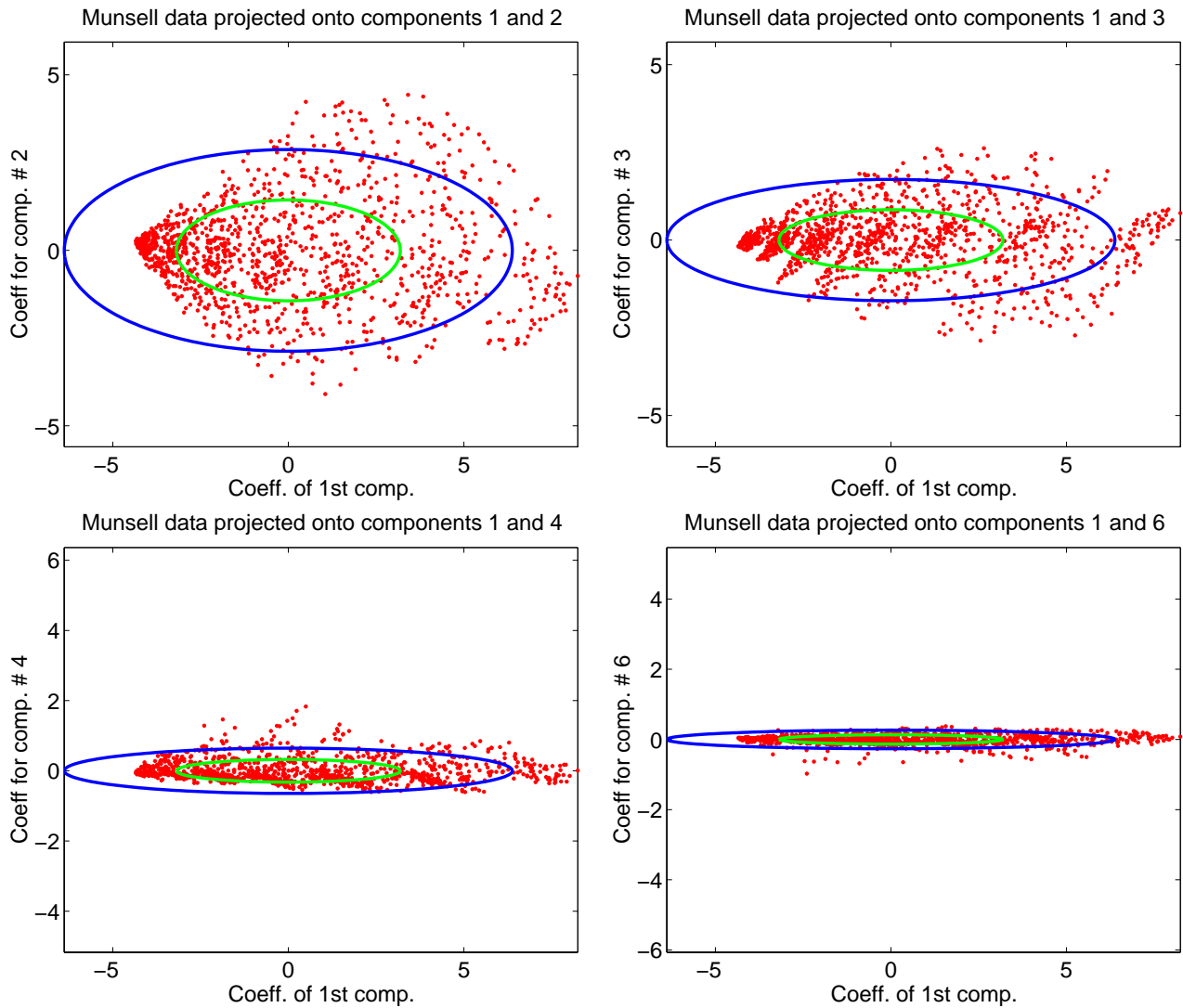
PCA Coordinates (Cont.)

- The green and blue curves in the previous plots indicate one and two standard deviations of the sample distribution.

- Note the standard deviation in direction $\vec{U}_k(\lambda)$ rapidly decreases as k increases. It is relatively small even for $k = 4$, illustrating that the data set is effectively three dimensional.

- Most species of animals have three or fewer colour channels. Honey bees are an exception, with 4 channels. They have an additional low-wavelength sensor, in the near ultra-violet range.

- Similar PCA results are obtained with other choices for the set of SPD signals (eg. different natural daylights, different natural materials, and considering the log of the SPDs).

- However, artificial light sources can have much more spikey spectral power distributions (see p.15), and produce significant responses in higher order PCA coefficients.

- As we saw before in the Fourier sampling theorem, a key ingredient in the representation and reconstruction of signals from a discrete set of samples is the (effective) dimension of the space of signals we are considering. The effective dimension must be less than or equal to the number of samples.

Alternative Formulation of PCA

An equivalent formulation of principal component analysis (PCA) is to consider the sample covariance matrix:

$$C = \frac{1}{M} \sum_{m=1}^{M} (\vec{I}_m - \vec{I}_0)(\vec{I}_m - \vec{I}_0)^T, \tag{14}$$

where $\vec{I}_0 = \frac{1}{M} \sum_{m=1}^{M} \vec{I}_m$ is the sample mean.

From the $N \times N$ covariance matrix C we can compute the variance in any direction $\vec{u}$ as follows (here $||\vec{u}|| = 1$):

$$\vec{u}^T C \vec{u} = \frac{1}{M} \sum_{m=1}^{M} |\vec{u}^T(\vec{I}_m - \vec{I}_0)|^2. \tag{15}$$

The principal component directions of the data set are defined to be the eigenvectors of the covariance matrix C. Since C is a symmetric matrix, there exists an orthogonal $N \times N$ matrix U such that

$$C = U \Lambda U^T, \tag{16}$$

with Λ a diagonal $N \times N$ matrix. The columns of U are the eigenvectors for C. They are the principal axes of the ellipsoids defined by

$$(\vec{I} - \vec{I}_0)^T C^{-1} (\vec{I} - \vec{I}_0) = \tau, \tag{17}$$

that is, surfaces of points $\vec{I}$ which are exactly τ (a constant) standard deviations from the mean $\vec{I}_0$ (see p.23).

Alternative PCA Formulation (Cont.)

To relate this approach to the SVD of E, note that

$$C = \frac{1}{M}\sum_{m=1}^{M}(\vec{I}_m - \vec{I}_0)(\vec{I}_m - \vec{I}_0)^T = \frac{1}{M}EE^T,$$

where E is given by equation (7). By the SVD decomposition of E (i.e. equation (8)) we find

$$C = \frac{1}{M}U\Sigma V^T V\Sigma^T U^T = \frac{1}{M}U\Sigma\Sigma^T U^T.$$

Here we have used the fact that V is an orthogonal matrix, so $V^T V = Id$. Finally, since the only nonzero elements in Σ appear along the main diagonal, $\Sigma\Sigma^T$ is a square diagonal matrix. Comparing this with equation (16) we can identify

$$\Lambda = \frac{1}{M}\Sigma\Sigma^T. \tag{18}$$

Thus the matrix U provided by the SVD analysis is a matrix of eigenvectors for the sample covariance C. Moreover, the associated eigenvalues of the sample covariance are just $\frac{1}{M}\sigma_n^2$, for $n = 1, \ldots, N$.

Thus principal component analysis can be formulated in terms of eigenvectors of the sample covariance matrix C or, equivalently, in terms of the SVD of the data matrix E.