

# Spatio-chromatic PCA of a mosaiced color image

*David Alleysson<sup>1</sup> and Sabine Süsstrunk<sup>2</sup>*

*<sup>1</sup>Laboratory of Psychology and Neurocognition, Université Pierre-Mendes, France*

*<sup>2</sup>Audiovisual Communications Laboratory (LCAV), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland*

## Abstract

In this paper, we analyze whether Principal Component Analysis (PCA) is an appropriate tool for estimating spatial information in spatio-chromatic mosaiced images. Ruderman et al. [1] have shown that the spatio-chromatic principal components of cone images contain first spatial information, followed by blue minus yellow and red minus green. However, their analysis is based on fully defined spatio-chromatic images. In case of a reduced spatio-chromatic set with a single chromatic value per pixel, such as present in the retina or in CFA images, we found that PCA is not an appropriate tool for estimating spatial information. By extension, we discuss that the relation between natural image statistics and the visual system does not remain valid if we take into account the spatio-chromatic sampling by cone photoreceptors.

## Introduction

The statistical analysis of natural scenes, as viewed by human observers, has given new insight in the processing and functionality of the human visual system. Pioneer work has shown the relation between redundancy reduction in natural scenes and the visual system's receptive fields [2, 3]. Using gray-scale natural scene imagery, Olshausen & Field [4] show that representing images with sparse (less redundant) code leads to spatial basis functions that are oriented, localized and band-pass, and resemble the receptive field structures of the primary cortex cells. Bell & Sejnowski [5] found that sparseness could be appropriately formalized using Independent Component Analysis (ICA), and show that independent components of natural scenes act as edge filters.

For the case of color, Buchsbaum & Gottschalk [6] use Principal Component Analysis (PCA) of L, M and S cone signals to derive post-receptoral mechanisms: luminance and opponent chromatic channels (blue minus yellow, and red minus green). Using a simple model of a natural scene (flat spectrum), they proved the emergence of post-receptoral mechanisms from cone signals and propose that this de-correlated coding reduces the information transmitted to the optical nerve. Later, Attick & Redlich [7] formalized the relation between natural

color scenes and retinal functions. They show that a retinal filter is consistent with a whitening process of the natural scene structure when noise is taken into account.

Finally, the use of hyperspectral images has allowed to precisely analyze the spatio-chromatic structure of natural scenes and confirmed previous studies [8]. Ruderman et al. [1] show that the principal components of natural color images, as sampled by cones, are consistent with post-receptoral receptive fields and provide reduced signals. Using ICA, Tailor et al. [9] and Lee et al. [10] show also that natural color image statistics could account for simple and complex color opponent receptive fields in the primary cortex.

From these studies, it seems that post-receptoral mechanisms of the human visual system correspond to a statistical analysis of natural scenes and provide a redundancy reduction. But all these studies do not take into account that cone receptor sampling already results in a reduced spatio-chromatic set. Doi et al. [11] have proposed a study where the cone mosaic is taken into account. They used a local arrangement of cones (127) from which they sampled LMS responses to construct vectors and perform ICA analysis. Although this method gives interesting results, it is still not realistic for simulating cone sampling since only a small part of the entire mosaic is used. Their study actually corresponds to analyzing the signal of a part of the retina moving along natural scenes. In this paper, we propose two novel methods for analyzing an entire mosaiced image.

In the visual system, the three types of cones L, M and S form a mosaic such that only a single chromatic sensitivity is sampled at each spatial location. Thus, the spatio-chromatic signals are already reduced by a factor of three compared to fully defined spatio-chromatic signals of a natural scene (or color image).

In this paper, we study whether statistical analysis of natural color images sampled with a spatio-chromatic mosaic still has a correspondence with the processing of the visual system. In this preliminary study, we investigate only a simple case. We restrict our analysis to Principal Component Analysis (PCA), a second order statistical analysis that performs a simple de-correlation of a signal. We use RGB color images instead of LMS images constructed from hyperspectral data, and we use a regular arrangement of RGB samples instead of a random

arrangement, such as given by the cone distribution in the retina. Actually, this experimental set-up coincides with many digital camera output, since most use a single CCD and a Color Filter Array (CFA) to provide color responses. Such systems sample a single chromatic sensitivity per pixel and need to interpolate the missing information to render color images. Thus, we can investigate if a spatio-chromatic analysis is able to help the reconstruction of a full spatio-chromatic image.

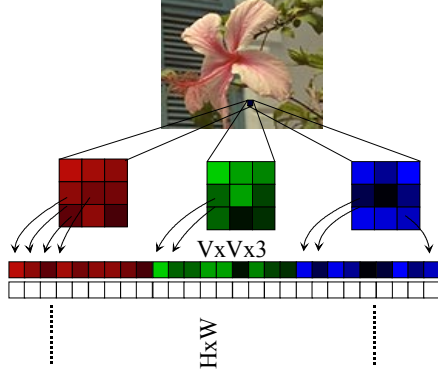


Figure 1: From an image, we construct a matrix that contains on each row the spatio-chromatic neighborhood of a pixel.

### PCA of color images

The main result of the work of Ruderman et al. [1] is that the spatio-chromatic PCA analysis of L, M and S responses to natural color images, computed from hyperspectral images, is consistent with post-receptoral mechanisms. We can reproduce this result using RGB images.

Given an image  $\mathbf{I}_{i,j,c}$ , defined by a three dimensional matrix of size  $H \times W \times 3$ , we can construct a matrix  $\mathbf{x}$  that contains for each row a column vector composed of spatial neighbors of size  $V$  of each pixel for each color layer (see Figure 1). Thus, the size of  $\mathbf{x}$  is  $(HW) \times (3V^2)$ . This matrix, on which we can apply an analysis, can be interpreted as containing on each row a representation of the spatio-chromatic random variables of a color image. We first compute the covariance of  $\mathbf{x}$  as follows:

$$\mathbf{C} = \text{Cov}(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}})^T (\mathbf{x} - \bar{\mathbf{x}}) / (HW - 1) \quad (1)$$

The resulting matrix  $\mathbf{C}$  of size  $3V^2 \times 3V^2$  is then decomposed into eigenvalues  $\mathbf{S}$  and corresponding eigenvectors  $\mathbf{U}$  of decreasing eigenvalue magnitude:

$$\mathbf{C} = \mathbf{U} \mathbf{S} \mathbf{U}^{-1} \quad (2)$$

The columns of  $\mathbf{U}$  are the eigenvectors and represent the basis functions of the transformation. We can represent these basis functions as spatio-chromatic samples and display their spatial and chromatic properties as color images (see Figure 2).

As shown by Ruderman et al. [1], the first principal components are mostly achromatic basis functions, followed by blue-yellow and finally red-green. Note that

the second component in Figure 2 is red. This result depends on our particular image; if we had chosen a set of RGB images instead of one single image, we probably would obtain a more accurate result.



Figure 2: (a) Image (b) Spatio-chromatic representations of eigenvectors of the covariance matrix (with  $V = 3$ ). Vectors are arranged in rows of decreasing eigenvalue magnitude.

The matrix  $\mathbf{U}$  represents a rotation matrix that transforms the original spatio-chromatic space into a space where components are de-correlated. If we call  $\mathbf{y}$  the de-correlated matrix corresponding to  $\mathbf{x}$ , we have:

$$\mathbf{y} = \mathbf{x} \mathbf{U} \Rightarrow (\mathbf{y} - \bar{\mathbf{y}})^T (\mathbf{y} - \bar{\mathbf{y}}) = \mathbf{U}^T \underbrace{(\mathbf{x} - \bar{\mathbf{x}})^T (\mathbf{x} - \bar{\mathbf{x}})}_{\mathbf{C}} \mathbf{U} = \mathbf{S} \quad (3)$$

Eq. 3 shows clearly that  $\mathbf{y}$  is a de-correlated data set. It is possible to partially reconstruct an image using only a few basis functions.  $\tilde{\mathbf{x}}$  is obtained from  $\mathbf{x}$  as follow:

$$\tilde{\mathbf{x}} = \mathbf{x} \mathbf{U} \mathbf{d} \mathbf{U}^{-1} \quad (4)$$

$\mathbf{d}$  is a diagonal matrix that contains zero or one depending if the corresponding vector is to be used or not. As shown in Figure 3, the first principal components of spatio-chromatic samples give a good approximation of the image. This is particularly true when the neighborhood remains small. Also, the achromatic basis functions only increase the resolution of the image (compare 3(b) and 3(c)) when the chromatic basis functions carry the color components of spatio-chromatic samples.

We can observe that basis functions seem to be decomposed into three categories. For the example of a  $3 \times 3$  neighborhood, there are 9 achromatic basis functions, 9 blue-yellow and 9 red-green. We may ask the question if the achromatic basis functions are able to reconstruct accurately the luminance information of the original image. To test this hypothesis, we can reconstruct the image using only achromatic basis functions. We then compare this image with the luminance image estimated as the mean of R, G and B at each pixel. Using basis functions 1, 3, 4, 6-11 (see Figure 2 (b)), we found a PSNR of 37.8 dB which is satisfying. Note, however, that if we use the first 11 principal

components, the PSNR equals 64dB, but using the first 9 components (i.e. leaving the last two achromatic basis functions out) results in a PSNR equal to 44 dB. This means that the chromatic 2<sup>nd</sup> and 5<sup>th</sup> components are important for the luminance reconstruction because they adjust the luminance level in the red and purple part of the image. This adjustment gives a similar estimate of luminance as adding the 10<sup>th</sup> and 11<sup>th</sup> achromatic components.



Figure 3: Partial reconstruction of the image. (a) Using only the first principal component (CPSNR=19.7) (b) Using 1<sup>st</sup> and 2<sup>nd</sup> (CPSNR = 21.6) (c) Using four (CPSNR = 24.4) (d) Using five principal components (CPSNR = 27.9).

### PCA of mosaiced color images

A mosaiced color image can be decomposed in luminance and opponent chromatic channels as illustrated in Figure 4. This decomposition keeps the full definition of spatial information in the luminance channel. Only opponent chromatic channels are sub-sampled according to the color mosaic of the image [12].

As we have shown in the previous section, it is possible to accurately estimate luminance from the spatio-chromatic samples of a color image using PCA. We now investigate if the method also works in case of a mosaiced color image.



Figure 4: (a) Mosaiced color image according to the Bayer CFA, decomposed as (b) luminance plus (c) sub-sampled opponent chromatic channels.

The construction of the vectors as shown in Figure 1 is no longer possible because the dimension of the mosaiced image with a single chromatic value per spatial location is  $H \times W$ . In other words, a mosaiced color image is already a scalar image. To avoid this problem, we thus propose to replace the missing colors by values 0

in the vector, as illustrated in Figure 5 (a). With this method, we keep a trace of the color components of the basis functions, since the first third of the vectors correspond to red, the second third to green and the last third to blue pixels. Figure 6 shows the resulting principal components of the mosaiced color image of Figure 4 (a).

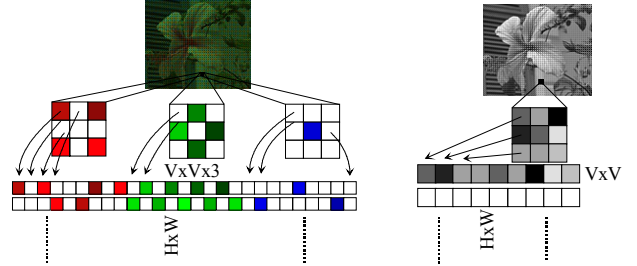


Figure 5: Method for constructing vectors of neighboring pixels in a mosaiced color image. (a) Inserting zeros for missing colors (b) Considering a color mosaic as a gray-scale image.

It can be seen that none of the basis functions have achromatic properties, meaning that the reconstruction of luminance is not possible. Actually, this is not exactly true; the 4<sup>th</sup> component does not contain the CFA structure. Using only this basis function, we can reconstruct luminance with a PSNR equal to 22.8. The resulting image is a low pass-filtered version of the luminance. Increasing the size of the neighborhood does not improve the result.

The reason for this failure is that the achromatic basis functions are weighted by the mosaic of color samples. For example, the 5<sup>th</sup> and 6<sup>th</sup> components of Figure 6 are edge functions, similar to the 3<sup>rd</sup> and 4<sup>th</sup> basis functions in Figure 2 (b), but weighted by the color mosaic. The 4<sup>th</sup> is the only basis function that lacks a color mosaic pattern.

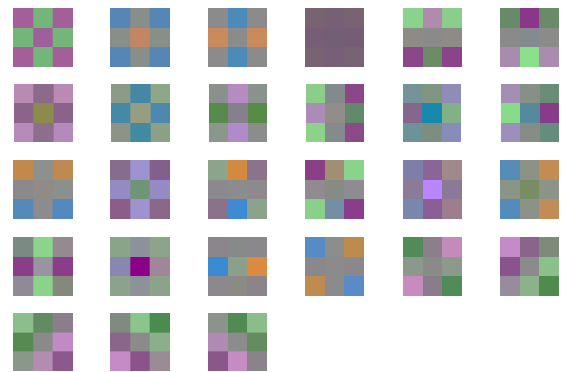


Figure 6: Result of the PCA of the mosaiced image of Fig. 4 (a) using the method of inserting zeros at missing color positions. The basis functions are arranged in rows of decreasing eigenvalue magnitude.

We have done the same analysis using only a Bayer CFA pattern. That is identical to sampling a constant white flat field by the CFA. We found that only three components have significant eigenvalues, and they

correspond exactly to the first three principal components of Figure 6. We tested if removing these components removes the grid effect due to the CFA, but that is not the case. The grid remains for each component except the 4<sup>th</sup>.

As shown in Figure 4, luminance has an important role for estimating spatial information in a CFA image. Considering that the method we have used above does not provide a good estimate of luminance, we propose a second method for analyzing mosaiced color images.

As illustrated in Figure 5 (b), we can also consider a CFA image as a gray-scale image and construct the vector of neighboring pixel as one would do it for a gray-scale image. Figure 7 (b) shows the result of the PCA in that case. In Figure 7 (a), we illustrate the result of a PCA analysis on luminance only, estimated as the mean of R, G and B at each pixel.

The PCA of the mosaic adds basis functions that are not present in the analysis of the “luminance only” image. By suppressing these functions (i.e. 2, 3, and 6), we obtain a PSNR equal to 29.2. Using a 5x5 neighborhood results in a PSNR equal to 29.6.



Figure 7: (a) PCA on Luminance (b) PCA on a mosaiced color image interpreted as a gray-scale image.

Thus, this method better estimates luminance than the previous one, and indicates that PCA could be used to estimate spatial information in mosaiced color images. We will further investigate if this method can follow the particular statistics of an image or image set.

## Conclusion

Principal Component Analysis allows efficient separation of the achromatic channel from the chromatic channels in color images because the achromatic component follows the second order statistics of a particular image. However, when using a mosaiced color image, it performs worse than a simple gaussian low-pass filter. The mosaic “pollutes” the basis functions and prevents good reconstruction. This is certainly due to the fact that the color mosaic and the color image are not de-correlated, and a de-correlation procedure cannot process them separately.

By extension, it seems that the de-correlation stage in the visual system, proposed by some authors as being a model of retinal processing, changes behavior when considering the sampling of a single color per cone location. Actually, this retinal sampling is already a redundancy reduction of the spatio-chromatic information of natural scenes, and might not necessary need a further de-correlation process.

In this study, we have restricted the statistical analysis to PCA. As Bell and Sejnowski [5] point out, PCA could model retinal processing, but for modeling

cortical processing, Independent Component Analysis is needed. It is possible that the color mosaic and the color image are independent rather than just de-correlated. In that case, a separation of information should be possible, and would confirm that the separation of spatial and chromatic information arises at a cortical level.

## References

- [1] Ruderman D.L., Cronin T.W., Chiao C.C., 1998. Statistics of cone responses to natural images: implication for visual coding. *J. Opt. Soc. Am.* 15. 2036-2045.
- [2] Field D.J., 1987. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4, 2379-2394.
- [3] Barlow H.B., 1989. Unsupervised learning. *Neural Computation* 1, 295-311.
- [4] Olshausen B.A., Field D.J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature* 381, 607-609.
- [5] Bell A.J., Sejnowski T.J., 1997. The independent components of natural scenes are edge filters. *Vis. Res.* 3327-3338.
- [6] Bauchsbaum G. & Gottschalk A., 1983. Trichromacy, opponent colours coding and optimum colour information in the retina. *Proc. R. Soc. Lond. B* 220, 89-113.
- [7] Atick J.J., Redlich A.N., 1992. What does the retina know about natural scenes? *Neural Computation* 4, 196-210.
- [8] Wachtler T., Lee T.W., Sejnowski T.J., 2001. Chromatic structure of natural scenes. *J. Opt. Soc. Am. A* 18. 65-77.
- [9] Tailor D.R., Finkel L.F., Buchsbaum G., 2000. Color-opponent receptive fields derived from independent component analysis of natural images. *Vis. Res.* 40. 2071-2076.
- [10] Lee T., Wachtler T. and Sejnowski T.J., 2002. Color opponency is an efficient representation of spectral properties in natural scenes. *Vis. Res.* 42, 2095-2103.
- [11] Doi E., Inui T., Lee T.W., Wachtler T., Sejnowski T.J., 2003. Spatio-chromatic receptive field properties derived from information-theoretic analyses of cone mosaic response to natural scenes. *Neural Comp.* 15. 397-417.
- [12] Alleysson, D., Süsstrunk S., Héroult J., 2002. Color demosaicing by estimating luminance and opponent chromatic signals in the Fourier domain. *Proc. IS&T/SID 10<sup>th</sup> Color Imaging Conference*, 331-336.