The Spectral Independent Components Of Natural Scenes

Te-Won Lee[1,2], Thomas Wachtler[1], ∗ and Terrence J. Sejnowski[1,2,3]

tewon@salk.edu, thomas@salk.edu, terry@salk.edu

[1] Howard Hughes Medical Institute, Computational Neurobiology Laboratory, The Salk Institute, La Jolla, California 92037, USA

[2] Institute for Neural Computation, University of California, San Diego, La Jolla, California 92093, USA

[3] Department of Biology, University of California, San Diego, La Jolla, California 92093, USA

∗Present address: Universität Freiburg, Biologie III / Neurobiologie und Biophysik, 79104 Freiburg, Germany

Abstract

We apply independent component analysis (ICA) for learning an efficient color image representation of natural scenes. In the spectra of single pixels, the algorithm was able to find basis functions that had a broadband spectrum similar to natural daylight, as well as basis functions that coincided with the human cone sensitivity response functions. When applied to small image patches, the algorithm found homogeneous basis functions, achromatic basis functions, and basis functions with overall chromatic variation along lines in color space. Our findings suggest that ICA may be used to reveal the structure of color information in natural images.

1   Learning codes for color images

The efficient encoding of visual sensory information is an important task for image processing systems as well as for the understanding of coding principles in the visual cortex. Barlow [1] proposed that the goal of sensory information processing is to transform the input signals such that it reduces the redundancy between the inputs. Recently, several methods have been proposed to learn gray-scale image codes that utilize a set of linear basis functions. Olshausen and Field [10] used a sparseness criterion and found codes that were similar to localized and oriented receptive fields. Similar results were obtained in [3, 8] using the infomax ICA algorithm and a Bayesian approach respectively. In this paper we are interested in finding efficient color image codes. Analysis of color images have mostly focused on coding efficiency with respect to the postreceptoral signals [4, 12]. Buchsbaum et al. [4] found opponent coding to be the most efficient way to encode human photoreceptor signals. In an analysis of spectra of natural scenes using PCA, Ruderman et al. [12] found principal components close to those of Buchsbaum. While cone opponency may give an optimal code for transmitting chromatic information through the bottleneck of the optic nerve, it may not necessarily reflect the chromatic statistics of natural scenes. For example, how the photoreceptor signals should be combined depends on their spectral properties. These however may not be determined solely by the spectral statistics in the environment, but by other functional (effects of infrared or UV sensitivity) or evolutionary (resolution) requirements. Therefore, opponent coding may not be the ultimate goal the visual system wants to achieve. And in fact, it is known that, while neurons in the Lateral Geniculate Nucleus (LGN) of trichromatic primates show responses along the coordinate axes ('cardinal directions') of cone-opponent color space [6], cortical cells do not adhere to these directions [7, 13]. This suggest that a different coding scheme may be more appropriate to encode the chromatic structure of natural images. Here, we use ICA to analyze the spectral and spatial properties of natural images.

2 Independent Component Analysis

ICA is a technique for finding a linear non-orthogonal coordinate system in multivariate data. The directions of the axes of this coordinate system are determined by the data's second- and higher-order statistics. The goal of the ICA is to linearly transform the data such that the transformed variables are as statistically independent from each other as possible [2, 5]. We assume that a data vector $\mathbf{x}$ can be modeled as a linear superposition of statistically independent source components $\mathbf{s}$ $(p(\mathbf{s}) = \prod_{i=1}^{M} p_i(s_i))$ such that

$$\mathbf{x} = \mathbf{As}, \tag{1}$$

where $\mathbf{A}$ is a $N \times M$ scalar matrix. The columns of $\mathbf{A}$ are called the basis functions. The learning algorithm can be derived using the information maximization principle [2] or the maximum likelihood estimation formulation. The data likelihood can be expressed as:

$$p(\mathbf{x}|\mathbf{A}) = \frac{p(\mathbf{s})}{|\det(\mathbf{A})|} \tag{2}$$

Maximizing the log-likelihood with respect to $\mathbf{A}$ and using the natural gradient gives

$$\Delta \mathbf{A} \propto -\mathbf{A} \left[\mathbf{I} - \varphi(\mathbf{s})\mathbf{x}^T\right] \tag{3}$$

where $\varphi(\mathbf{s}) = -\frac{\partial p(\mathbf{s})/\partial \mathbf{s}}{p(\mathbf{s})}$. Our primary interest is to learn efficient codes, and we choose a Laplacian prior $(p(s) \propto \exp(-|s|))$ because it captures the sparse structure of coefficients ($\mathbf{s}$) for natural images. This leads to the simple learning rule which we used for our analysis

$$\Delta \mathbf{A} \propto -\mathbf{A} \left[\mathbf{I} - \text{sign}(\mathbf{s})\mathbf{s}^T\right]. \tag{4}$$

3 Independent Components of Single Hyperspectral Pixels

We analyzed a set of four hyperspectral images of natural scenes[1]. A detailed description of the images is given in Parraga et al (1998). Briefly, the data set consists of 29 images. Each image has a size of 256x256 pixels, for which radiance values are given for 31 wavebands, sampled in 10

---

[1]Dataset provided by Parraga et al. (http://www.crs4.it/∼gjb/ftpJOSA.html).

nm steps between 400 and 700 nm. Pixel size is 0.056x0.056 deg of visual angle. The images were recorded around Bristol, either outdoors, or inside the glass houses of Bristol Botanical Gardens. We chose four of these images which had been obtained outdoors under apparently different illumination conditions (figure 1). Training was done in 1000 sweeps, each using a set of spectra of 40000 pixels,

Figure 1: Four hyperspectral color images of natural scenes.

which were chosen randomly from the four images. We used the logarithm of radiance values, as in the study by [12]. The data were not preprocessed otherwise. The resulting basis functions for the pixel spectra are shown in figure 2. The basis functions are plotted in order of decreasing L_2 norm. Figure 2 shows the corresponding relative contributions of the basis functions to the pixel spectra. The first basis function has a broadband spectrum, with a higher contribution in the short wavelength range. Its overall shape resembles typical daylight spectra [14]. Basis functions two to five show properties related to photoreceptor sensitivities: A comparison between the first five basis functions and human cone sensitivities is shown in figure3. Basis functions two and four have peaks that coincide with the peak of the M cone sensitivity. Note that basis functions are rescaled and sign-corrected to have positive peaks in this figure. Basis function three aligns with the short wavelength flank of the L cone sensitivity, and with the long wavelength flank of the M cone sensitivity. Finally, basis function five has a peak beyond the wavelength of the L cone sensitivity, where the difference between L and M cones is largest. These basis functions may represent object reflectances. Osorio et al. [11] showed that the human cone spectra are related to naturally occurring reflectance spectra. The remaining basis functions are mostly very narrow band, and their contributions are small.
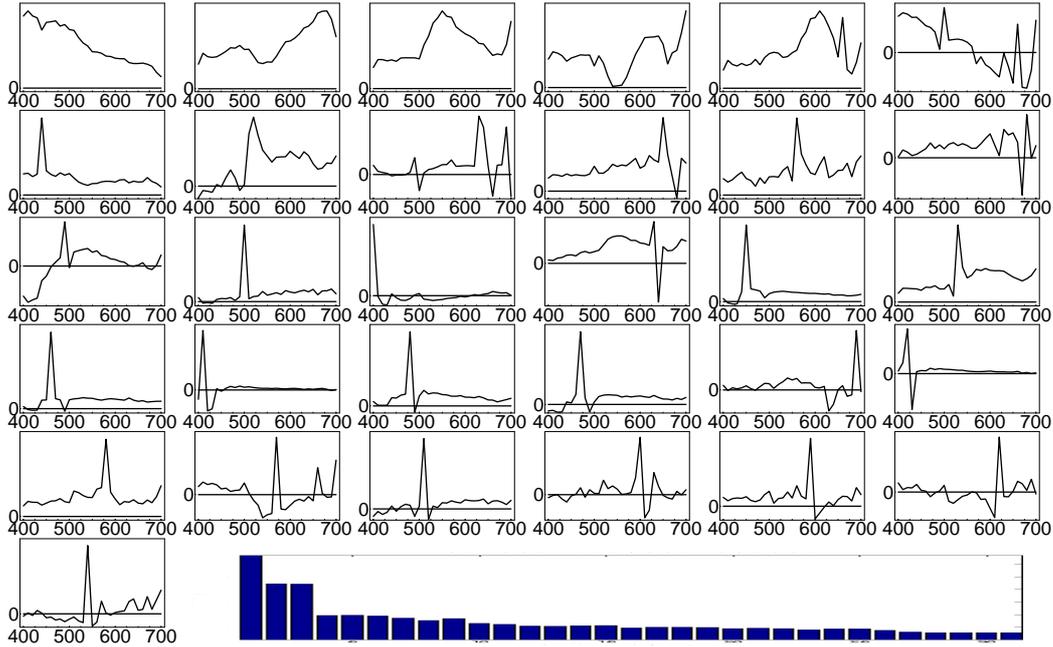
3

Figure 2: The spectra of the learned basis functions ordered in decreasing L_2-norm their relative contributions (inset).

4 Independent Components of Hyperspectral Natural Images

To analyze the spatial properties of the data set, we converted the spectra of each pixel to a vector of 3 cone excitation values (long-, medium-, and short-wavelength sensitive). This was done by multiplying the radiance value for each wavelength with the corresponding values for the human cone sensitivities as provided by Stockman et al (1993) (http://www-cvrl.ucsd.edu), and summing over the resulting values. From these data, 7x7 image patches were chosen randomly, yielding 7x7x3 = 147 dimensional vectors. Training was done in 500 sweeps, each using a set of spectra of 40000 image patches, which were chosen randomly from the four images. To visualize the resulting components, we plot the 7x7 pixels, with the color of each pixel indicating the combination of L, M, and S cone responses as follows. The values for each patch were normalized to values between 0 and 255, with 0 cone excitation corresponding to a value of 128. This method was used in [12]. Note
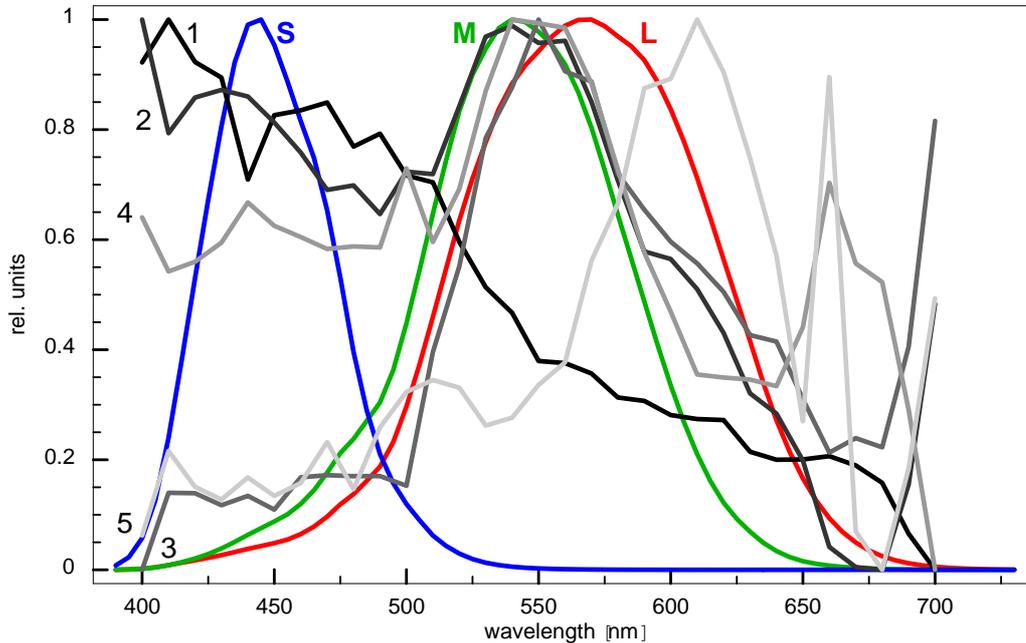
4

Figure 3: Comparison between the first five basis functions and the human cone sensitivity response function.

that the resulting colors are not the colors that would be seen through the corresponding filters. Rather, the red, green, and blue components of each pixel represents the relative excitations of L, M, and S cones, respectively. In figure 4 A, we show the first 100 of the 147 components, ordered by decreasing L_2 norm. The first three, homogeneous, basis functions contribute on average 25% to the intensity of the images. Most of the remaining basis functions are achromatic, localized and oriented filters similar to those found in the analysis of grayscale natural images [3]. There are also many basis functions with color modulated between light blue and dark yellow. For both types of components, low spatial frequency components tend to have higher norm than components with higher spatial frequency. To illustrate the chromatic properties of the filters, we convert the L, M, S values for each pixel to its projection onto the isoluminant plane of cone-opponent color space. This space has been introduced by [9] and generalized to include achromatic colors by [6]. In our plots, the x axis corresponds to the response of a L cone versus M cone opponent mechanism, the y axis corresponds to S cone modulation. Note that these axes do not coincide with colors we perceive
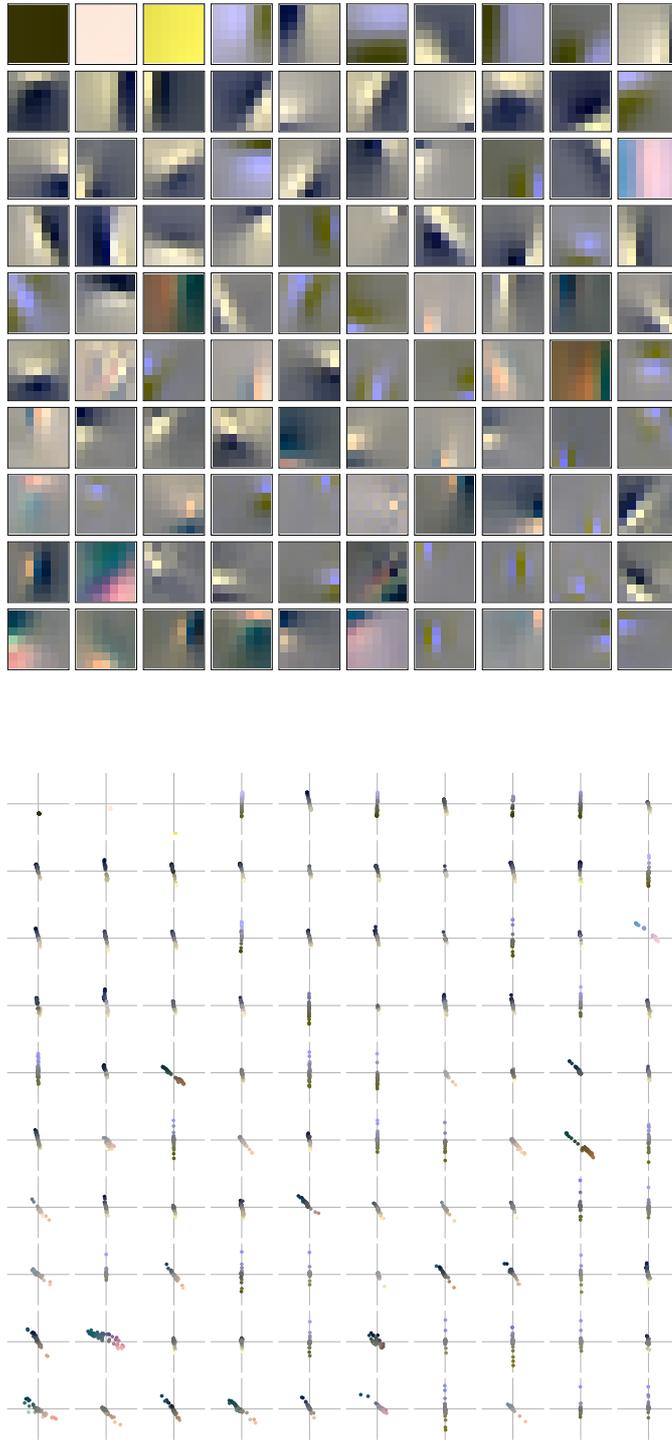
Figure 4: (Top) 100 of 147 total learned basis functions (7 by 7 pixels and 3 colors) ordered in decreasing L_2-norm. (Bottom) Corresponding color-space diagrams for the 100 basis functions.

6

as pure red, green, blue and yellow. For each pixel of the basis functions, we plot a point at its corresponding location in that color space. The color of the points are the same as used for the pixels in figure 4 (top). Thus, although only the projection onto the isoluminant plane is shown, the third dimension can be inferred by the brightness of the points. Interestingly, almost all components show chromatic variation along a line in color space. Only a few, weak, basis functions show color coordinates which do not form a line. The blue-yellow basis functions lie almost perfectly along the vertical S cone axis. The achromatic basis functions lie along lines that are slightly tilted away from this axis. This reflects the direction of variation of natural daylight spectra, whose coordinates in this color space lie along a line which is tilted counterclockwise with respect to the vertical axis. Notably, the yellow end of this line correlates with brighter colors (objects lit by sunlight), the blue end to darker colors (objects in shadow, lit by bluish skylight). The chromatic basis functions except the S cone modulated ones tend to lie along lines with orientations corresponding to (greenish) blue versus yellow/orange. There are no basis functions in the direction of L versus M cone responses (horizontal axis).

5 Discussion

We used ICA to analyze the spectral and spatial properties of natural images. In single hyperspectral pixels, we found that ICA was able to find basis functions with broadband spectra and basis functions related to human cone sensitivities. When applied to small image patches, ICA found homogeneous basis functions, achromatic and chromatic basis functions. Most basis functions showed pronounced opponency, i.e. their components form lines through the origin of color space. However, the directions of these lines do not always coincide with the coordinate axes. While it is known that chromatic properties of neurons in the LGN corresponds to variation restricted to these axes [6], cortical neurons show sensitivities for intermediate directions [7]. This suggests that the opponent coding along the 'cardinal directions' is used by the visual system to transmit visual information to the cortex, where the information is recoded, maybe to better reflect the statistical structure of the visual environment. Interestingly, ICA found only few basis functions with strong red-green opponency.

7

The reason for this may lie in the fact that our images did not contain flowers or other strongly colored objects. Also, chromatic signals that are ecologically important [11] may not be typical or frequent in natural scenes. Using PCA, Ruderman et al. (1998) found components which reflect the opponent mechanisms to decorrelate chromatic signals, given the human cone sensitivities. The basis functions found by PCA are a result of the correlations introduced by the overlapping sensitivities of human cones. In contrast to PCA, ICA tries to discover the underlying statistical structure of the images. Our results are consistent with previously reported results on gray-scale images and we suggest that ICA may be used to reveal the structure of color information in natural images.

Acknowledgments

References

[1] H. Barlow. *Sensory Communication*, chapter Possible principles underlying the transformation of sensory messages, pages 217–234. MIT press, 1961.

[2] A. J. Bell and T. J. Sejnowski. An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, 7:1129–1159, 1995.

[3] A. J. Bell and T. J. Sejnowski. The 'independent components' of natural scenes are edge filters. *Vision Research*, 37(23):3327–3338, 1997.

[4] G. Buchsbaum and A. Gottschalk. Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal Society London B*, 220:89–113, 1983.

[5] J-F. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Trans. on S.P.*, 45(2):434–444, 1996.