

Probabilistic Theories of the Visual Cortex

A.L. Yuille and D. Kersten

Abstract

THE VERY EARLY VISUAL SYSTEM

This lecture first briefly reviews the structural organization of V1, the properties of simple cells, and divisive normalization. The lecture also illustrated principles such as sparsity, independence, and inverting generative models.

A. Review: From Retina and LGN to V1

Light is captured in the retina, transmitted to the LGN, and then to area V1 of the visual cortex. Receptive field properties of neurons in retina and LGN are generally believed to be modelled by symmetric center-surround cells – i.e. the Laplacian of a Gaussian filter, which looks like a Mexican Hat. This may be an over-simplification (e.g., see Meister for an alternative viewpoint) but Yang Dan reports that it is possible to reconstruct the input image from the responses of neurons in retina or LGN (which would seem to be impossible if the standard models were badly wrong).

There is an expansion (by a factor between 80 and 400) as we move from the LGN to V1. This is not surprising because V1 starts the hard problem of interpreting the image (while the retina and LGN perform the simpler tasks of capturing the image and transmitting it to the cortex – at least this is the standard picture). Things get more complicated as we get to V1. There are a large variety of cells – e.g., simple and complex – and nobody has succeeded in reconstructing the input image by estimating it from the recordings of neurons in V1.

B. Organization of V1

The organizational structure of V1 was mapped by Hubel and Wiesel (who got a Nobel prize for doing this). They showed that the receptive fields of cells in V1 (unlike those in the retina and LGN) were tuned to stimuli at different orientations (e.g., bars, sinusoid gratings). They also showed that there was a clear organizational structure, see figure (1) (see also Blasdel). Durbin and Mitchison (A Dimension reduction framework for understanding cortical maps. *Nature*. 1990) suggested that this structure could be derived as an optimal way to map a four dimensional feature space (2D position, orientation, scale) onto the two dimensional cortical structure while encouraging that cells tuned to similar properties (in feature space) are located at neighboring positions in the 2D cortical surface.

Caveats. Mapping the visual cortex involves systematically estimating the receptive fields of neurons on the cortical surface. But there is "selection bias" because not every neurons can be tested (there are too many) and, in practice, scientists select neurons which respond a lot to the input stimuli being tested. So neurons which are sensitive to stimuli outside this stimuli set may be missed. See Carandini, M.; Demb, J. B.; Mante, V.; Tolhurst, D. J.; Dan, Y.; Olshausen, B. A.; Gallant, J. L.; Rust, N. C. (2005). "Do We Know What the Early Visual System Does?". *Journal of Neuroscience* 25 (46).

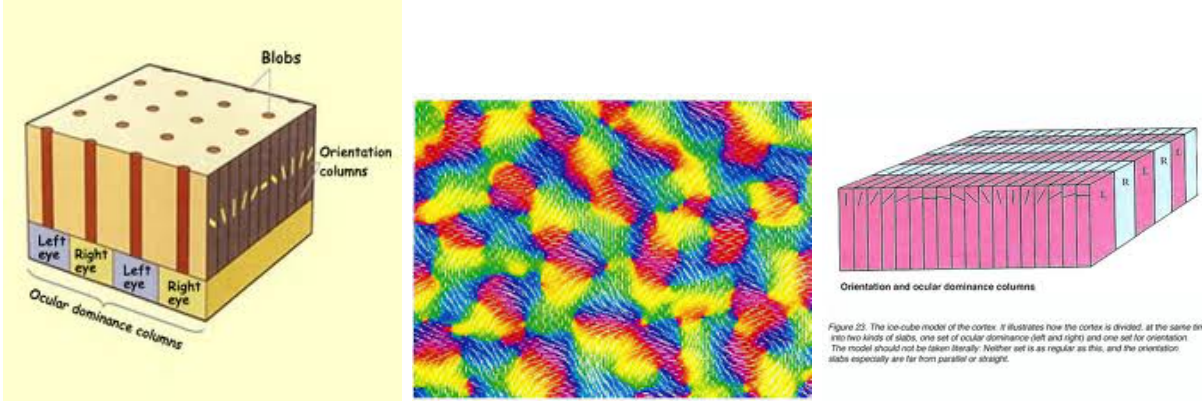


Fig. 1. Left: Blobs and orientation columns. Center: orientation tuning map. Right: orientation and ocular dominance columns.

C. Basic Models

1) *Oriented Stimuli and Edge Detection.*: Early studies of neurons in V1 found that cells were often sensitive to oriented stimuli. Cells in the visual cortex were classified by their receptive fields as simple, complex, and hyper-complex cells [2][3]. Simple and complex cells are sensitive to stimuli at specific orientations. Simple cells can be modeled by linear models (see previous lectures). The difference is that the filters $T(\cdot)$ are oriented. They can be modeled as the real and imaginary parts of Gabor functions $T(\vec{x}) = \exp\{i\vec{\omega} \cdot \vec{x}\}G(\vec{x} : \vec{0}, \Sigma)$ [1], where the covariance Σ of the Gaussian is aligned to the direction of $\vec{\omega}$. This gives sinusoid filters of form $Gabor_{sine}(\vec{x}) \sin(\vec{\omega} \cdot \vec{x})G(\vec{x} : \vec{0}, \Sigma)$ and $Gabor_{cosine}(\vec{x}) = \cos(\vec{\omega} \cdot \vec{x})G(\vec{x} : \vec{0}, \Sigma)$ which form quadrature pairs (where the covariance Σ of the Gaussian is chosen so that its principal direction is aligned to the direction $\vec{\omega}$ of the sinusoid). Alternatively these cells can be modeled as directional derivatives of derivatives of Gaussians [18]). Jones and Palmer [5] gave experimental support for Daugman's conjecture about Gabor functions. See figure (2).

Complex cells are sensitive to orientation but they are less sensitive to the spatial position of the stimuli. A popular model is to take the sum of the squares of the responses of two simple cells whose receptive fields are out of phase:

$$O(\vec{x}) = \{Gabor_{sine} * I(\vec{x})\}^2 + \{Gabor_{cosine} * I(\vec{x})\}^2. \quad (1)$$

There is some evidence - Roner and Pollen 1981 – that simple cells in the cat are organized in quadrature pairs at 90 degree different phase shift.

These cells can be used to design edge detectors. The simplest method is to label a pixel as an edge if the magnitude of the response of a derivative filter is above threshold. Better performance can be obtained if we consider the statistics of natural images and combine derivative filters at different scales (explain "scales" in more detail). See [6] which quantifies the effectiveness of filterbanks for performing edge detection. There have been recent attempts (Mel et al.) to relate this to local neural circuits. Simple and complex cells can perform several computational functions. In particular they can be used for edge detection [10].

Caveats: but are simple cells in V1 well modeled by Gabors? And are simple cells and complex cells really different? It has been suggested that simple and complex cells are not distinct classes but instead are part of a continuum of receptive field types with greater or less sensitivity to spatial position [9]. It is also reported that many complex cells do not have simple cells as inputs, which throws doubt about the simple quadratic model above.

2) *Frequency and Spatial structure of images*: Images can be represented by Fourier series expansions:

$$I(\vec{x}) = \sum_n a_n \exp\{i\vec{\omega}_n \cdot \vec{x}\}, \quad (2)$$

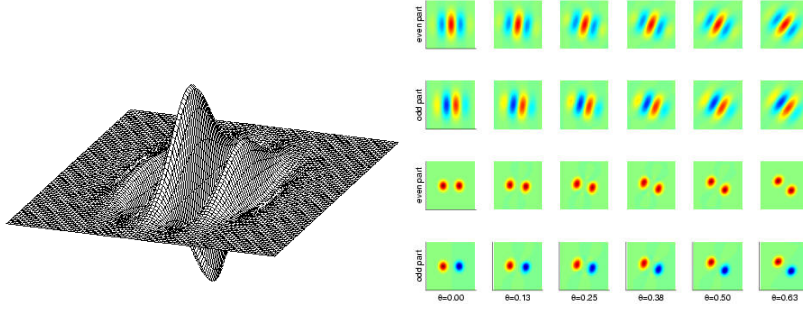


Fig. 2. Left: a Gabor filter. Right: a family of Gabor filters.

where ω denotes the frequencies and the a are the coefficients. These coefficients can be calculated by $a_n = \frac{1}{2\pi} \int d\vec{x} I(\vec{x}) \exp\{-i\vec{\omega}_n \cdot \vec{x}\}$.

The low frequencies (small $|\vec{\omega}|$) represent the large scale global structure of the image. The high frequencies (large $|\vec{\omega}|$) gives the small scale structure. If we blur the image, by convolving with a Gaussian G (with variance $\sigma^2 = t/2$), we rapidly reduce the coefficients of the high frequency components:

$$I(\vec{x}, t) = G_t * I(\vec{x}) = \sum_n a_n \exp\{-|\vec{\omega}|^2 t\} \exp\{i\vec{\omega}_n \cdot \vec{x}\}. \quad (3)$$

The classic "block illusion" (or Lincoln illusion), shown in figure (3), illustrates what happens when we remove the high frequencies from images yielding a block type image which obscures the identity of the person in the image. At first glance, the image seems impossible to interpret. But if we blur the image, by squinting our eyes, then the image becomes clearer and we can interpret it easily (removing our sensitivity to large frequencies).

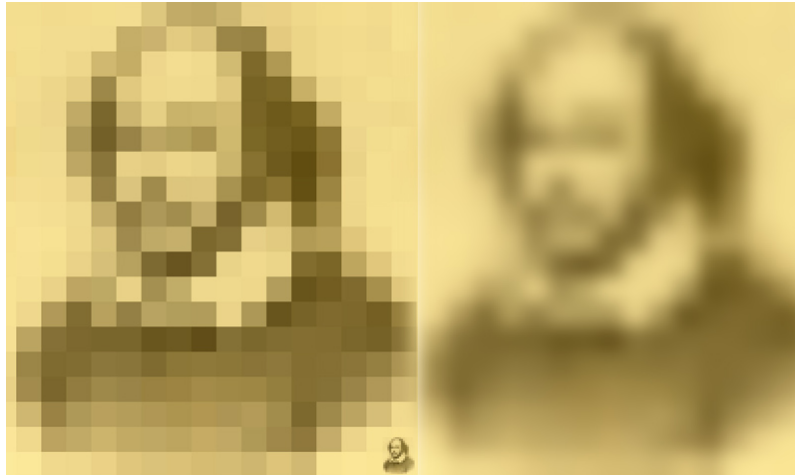


Fig. 3. Left: a "block" image where the high frequencies have been removed. But if you squint then you can recognize the image. Right: the same image blurred.

3) *The Gabor function:* The fourier transform enables us to represent images in terms of spatial frequencies. This representation is attractive since it allows us to interpret the image at different scales. But there is one big disadvantage: the fourier transform depends on global properties of an image. We would like a way to represent an image which depends on the local properties of images.

The Gabor function is a compromise between localization in space and in frequency (J. Daugmann. "Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters". J. Opt. Soc. America. vol.2. No. 7. July. 1985). We define optimality in space by

a criterion $\int x^2 |G(x)|^2 dx$ (we require $\int dx |G(x)|^2 = 0$). The function that minimizes this is the delta function $G(x) = \delta(x)$ which is precisely localized in space. But this function contains all frequencies – $\int dx \delta(x) \exp\{-i\omega \cdot x\} = 1$ – and so is completely unlocalized in frequency space. Conversely, a sinusoid $\exp\{-i\Omega x\}$ is completely unlocalized in space, but is perfectly localized in frequency space (i.e. its fourier transform is $\delta(\omega - \Omega)$).

Now try to optimize the joint locality in space and frequency – $\{\int x^2 |G(x)|^2 dx\} \{\int d\omega |FG(\omega)|^2 d\omega\}$, where $FG(\cdot)$ is the fourier transform of G . (Also require G, FG to be normalized). Then it is straightforward to show that the Gabor optimizes this criterion.

Note: this trade-off result in frequency and space was known to Gabor. But Daugmann extended it to two-dimensions and suggested Gabors as models of receptive fields of simple cells in V1.

4) *Filterbanks and texture description*: Many authors have suggested that filterbanks of Gabor filters (and similar filters) provide local description of images which can be used to enable other visual processing. An important example is to model image texture (R. Jain, Perona and Malik). T.S. Lee (reference) has proposed that cells in V1 represent these filterbanks.

These filterbanks can also be used for matching tasks, such as stereo and motion correspondence (discussed in later sections).

D. Learning Receptive Fields

There have been many attempts to derive these type of receptive fields. These attempt falls into two claases. The first are motivated by biologically plausible learning rules, such as variants of Hebbs rule, but uses only limited assumptions about the statistics of images (namely shift invariance). The second are motivated by the statistics of natural images. Both end up with receptive field models that look like Gabor filters.

1) *Hebbian-Type Models*: Research in the 1980's showed that qualitatively similar receptive fields could be obtained by learning algorithms whose input was random dot. More specifically, Linsker [8][7] showed that these receptive field properties could result from unsupervised learning algorithms provided the input images were shift-invariant and had specific correlation statistics.

We now describe a simple model of this type. The output $O(t)$ is a weighted function of the inputs $I_i(t)$, and the weights $\omega_i(t)$ are updated by a variant of Hebbs rule due to Oja:

$$O(t) = \sum_j \omega_j(t) I_j(t),$$

$$\frac{d\omega_i(t)}{dt} = O(t) \{I_i(t) - O(t) \omega_i(t)\}. \quad (4)$$

This can be expressed as a single update equation (substituting for $O(t)$):

$$\frac{d\omega_i(t)}{dt} = \sum_j \omega_j I_i(t) I_j(t) - \sum_{jk} \omega_i \omega_j \omega_k I_j(t) I_k(t). \quad (5)$$

Next we make the key assumption that the weights ω_i change at a slower rate than the input images. This enables us to replace the terms $I_i(t) I_j(t)$ in equation (5) by their expectation $F_{ij} = \langle I_i(t) I_j(t) \rangle$ where the expectation is taken with respect to the distribution $P(I)$ on the input images. This gives update rule:

$$\frac{d\omega_i(t)}{dt} = \sum_j \omega_j F_{ij} - \sum_{jk} \omega_i \omega_j \omega_k F_{jk}. \quad (6)$$

It can be shown that the weights $\omega_i(t)$ converge to the eigenvectors of the image correlation matrix F_{ij} (note: there are some technical conditions which must be satisfied – the result is more direct by modifying the learning rule).

Now suppose that the image correlation is shift invariant – i.e. the correlation between image intensity at sites i and j depends only on the relative positions of i and j . This is a very reasonable assumption. In this case, the eigenvectors can be shown to be sinusoids (simple properties of fourier theory).

Hence this class of models suggests that receptive fields should have sinusoid properties. But receptive fields on neurons are limited by the geometrical positions of the dendrites. This puts an additional constraint on receptive fields which cause them to fall off with distance. If this fall-off is a Gaussian then we recover Gabor functions (as can be shown by simulations).

Hebbs rule uses the first term of equation (??) – the synaptic weight is modified based on the correlation of the output with the input. (Kandel got a Nobel prize for showing that some animals – Aplysia – learnt by a variant of Hebbs rule.

Miller (reference) showed that the ocular dominance columns could also be obtained by similar unsupervised learning rules. Other work showed that quadrature pairs arose naturally by adding inhibition between neighboring neurons [19].

Why are the receptive fields oriented even though the basic set-up is completely symmetric? This can be explained (Kammen and Yuille) by the phenomenon known in Physics as symmetry breaking. The system dynamics minimizes a function which is symmetric but where the minima consist of a continuous set of points (e.g., the brim of a mexican hat) and not just a single point (e.g., a regular cap). Hence the dynamics converge to one of a set of minima, and the minima selected depends on small randomness in the initial conditions and the dynamics. (Note; symmetry breaking was studied in physics by Nambu and Goldstone, and is the basis of the Higgs mechanisms of giving mass to the particles that mediate nuclear forces, and which resulted in several Nobel prizes).

2) *Image Properties and Receptive Fields*: More recent works attempts to learn the receptive fields using statistics of natural images. This has been done by Independent Component Analysis (ICA) [16] or by sparse coding [12][14][13]. See also [4][11]. The next section describes an approach based on sparse coding.

Sparsity and Sparse Coding

Sparsity has been proposed as a general principle for the visual cortex. Barlow has argued for sparsity on the grounds that only a small proportion of neurons in V1 (and elsewhere??) fire in response to a given image and hence the response is "sparse". This is desirable for neurons because it means that only a few of them need to be active and expend energy (the brain consumes more energy than the rest of the body).

Filterbanks give an over-complete representation of images. Hence,

The principle behind sparse coding is that the vision system has an over-complete set of basis functions. It tries to represent each image in terms of a small set of these functions. This over-completeness means that the the basis functions can be tuned to interesting features of the image. As mentioned above, the sparsity principle can also be used to learn receptive fields from natural images (refs!!).

Following functional analysis, an image can be represented in terms of a linear combination of basis functions $\{A_i(x, y)\}$ by:

$$I(x, y) = \sum_{i=1}^n A_i(x, y) s_i, \quad (7)$$

where (x, y) are the position in the image, the $A_i(x, y)$ are the basis functions, and the s_i are the coefficients.

An important special case occurs when the number of basis vectors is equal to the number of pixels in the image. In this case the model can be inverted by linear operations to calculate the coefficients s uniquely. In addition, if the basis functions are orthonormal – i.e. $\sum_{x,y} A_i(x, y) A_j(x, y) = \delta_{ij}$, $\forall i, j$ – then we can compute the coefficients by $s_i = \sum_{x,y} A_i(x, y) I(x, y)$. This includes Fourier transforms, as discussed above.

But if we have a over-complete representation, like a filterbank, then there is no unique solution of the coefficients s in equation (7). We need to introduce restrictions on the coefficients $\{s_i\}$ in order to get a

unique solution.

One possibility is to put a penalty on the squares of the coefficients. This was proposed (in the context of brain modeling) by Daugman and Kammen (ref!!) and gives a solution for the coefficients s in terms of the Moore-Penrose psuedo-inverse. This approach starts by formulating the problem in terms of an optimization problem which contains two terms. The first penalizes an error between the the image and its representation in terms of basis functions. The second penalizes the strength of the coefficients $\{s_i\}$. Note that without penalizing the coefficients there are too many ways to represent the image in terms of basis functions and hence the representation is ambiguous. The criterion is specified as:

$$E_2[s] = \sum_{x,y} \{I(x,y) - \sum_i s_i A_i(x,y)\}^2 + \mu \sum_i s_i^2. \quad (8)$$

The $\{s_i\}$ are obtained by minimizing $E_2[s]$. By differentiation, this reduces to solving the following linear equations:

$$s_i = \{\mu \delta_{ij} + A_{ij}\}^{-1} \sum_{x,y} A_i(x,y) I(x,y),$$

$$\text{where } A_{ij} = \sum_{x,y} A_i(x,y) A_j(x,y), \text{ and } \{..\}^{-1} \text{denotes matrix inverse.} \quad (9)$$

In this model the response is partially specified by a linear filter – the $\sum_{x,y} A_i(x,y) I(x,y)$ term – but is also modulated by the other basis vectors, as expressed by the A_{ij} term. These could be implemented by lateral interactions between neurons.

But an alternative, and arguably much better, approach is to use an alternative L_1 penalty $\sum_i |s_i|$ instead of the quadratic penalty, or L_2 norm, shown above. As discussed below, this approach helps impose sparsity which encourages some of the coefficients s_i to be zero (while, by contrast, the L_2 norm encourages terms to be small but non-zero).

This *sparsity criteria* was developed by Olshausen and Field as a way to learn receptive fields of neurons [12]. It gives a reconstruction criteria:

$$E_1(s) = \sum_{x,y} \{I(x,y) - \sum_i s_i A_i(c,y)\}^2 + \mu \sum_i |s_i|. \quad (10)$$

This criteria can be extended to multiple images and used to learn receptive fields. It requires a dataset of images $\{I^\mu(x,y) : \mu \in \mathcal{D}\}$ and determines the basis functions by minimizing the criteria below with respect to $A_i(x,y)$ and the s_i^μ

$$E[A, s] = \sum_{\mu \in \mathcal{D}} \sum_{x,y} \{I^\mu(x,y) - \sum_i s_i^\mu A_i(c,y)\}^2 + \mu \sum_{\mu \in \mathcal{D}} \sum_i |s_i^\mu|. \quad (11)$$

This results in receptive field models which are similar to those measured [12]. Note that similar receptive fields can be obtained by assuming a similar model for the image, see equation (7), but imposing different assumptions on the form of the s_i . In particular, independent component analysis (ICA) gives similar receptive field models [16]. Hyvarinen [4] explains this by showing that both types of models – L1 sparsity and ICA – both encourage that the s_i are strongly peaked at 0, but can occasionally have large non-zero values (this contrasts with the Gaussian model – e.g. pseudo-inverse – where the s_i are strongly discouraged from taking large values).

Finally, we mention that all these models can be given a statistical interpretation (to lie within the spirit of this chapter). For both models, we specify a generative model $P(I|A, s) = \frac{1}{Z} \exp\{-\sum_{x,y} (I(x,y) - \sum_i s_i A_i(x,y))^2\}$ and a prior $p(s) = \frac{1}{Z} \exp\{-\mu \sum_i |s_i|\}$, which is a Gaussian for the pseudo-inverse.

3) *L₂ and L₁ norms and sparsity*: To illustrate the difference between L₂ and L₁ norms we consider a simple one dimensional example. Suppose we want to estimate a quantity x and have a noisy measurement a . We define a quadratic penalty $(x - a)^2$.

First suppose, we use an L₂ penalty kx^2 and solve for:

$$x^*(a) = \arg \min_x \{(x - a)^2 + kx^2\}$$

This gives a solution $x^*(a) = a/(1 + k)$, which is non-zero even if a is very small.

Secondly, use an L₁ penalty $2k|x|$. Now solve for x by minimizing $f(x) = (x - a)^2 + 2k|x|$. It can be shown that minimum of a is now given by:

$$x^*(a) = 0, \text{ if } |a| < k, \quad x^*(a) = a - k \text{ if } a > k, \quad x^*(a) = a + k \text{ if } a < -k$$

Hence the L₁ will make $x^*(a) = 0$ provided a is small enough, and hence enforces sparsity. This result is exploited in the theory of compressed sensing.

To understand this result. Re-express $f(x)$ in terms of two functions $f_+(x), f_-(x)$ valid in the domains $x > 0$ and $x < 0$ respectively:

$$f_+(x) = (x - a)^2 + kx, \quad f_-(x) = (x - a)^2 - kx.$$

If $a > k$, then it can easily be checked that we can calculate $x^*(a)$ by setting the derivative of $f_+(x)$ to be zero and solving for x . Similarly if $a < -k$, we can obtain $x^*(a)$ by differentiating $f_-(x)$. But if $-k < a < k$, then the minima of $f_+(x)$ and $f_-(x)$ both occur at $x = 0$, which yields $x^*(a) = 0$ in this case.

E. Divisive Normalization: inverting generative models to obtain independent factors

Standard wisdom states that the simple cells can be described by the oriented linear filter models described earlier in this lecture. But there are also local interactions between the neurons. Experiments showed that the response of neighbouring neurons was often normalized by the response of their neighbors (Heeger – reference!!).

Now we describe a why to explain this divisive normalization using the graphical model formulation. An important example is the use of probabilistic models [17] to account for divisive gain control (e.g. contrast normalization). The main idea is that filters with similar preferences for orientation representing nearby spatial locations in a scene have striking statistical dependencies, which can be removed by divisive normalization. Specifically, if we plot the statistics of two linear filters f_c, f_s (center and surround) then the magnitudes of f_c, f_s are coordinated in a straightforward way, which has a characteristic shape of a Bow-Tie.

This can be modeled by assuming there are hidden variables ν which affect both responses and hence induces correlation between the responses. For example, ν could represent the local average image intensity which could affect the response of both filters but, after the filter response could be made independent by conditioning on the average intensity. Suppose ν has a prior distribution $P(\nu) = \nu \exp\{-\nu^2/2\}$ for $\nu \geq 0$. We have a pair of filters $\{l_i : i = 1, 2\}$ which are related to gaussian models $\{g_i : i = 1, 2\}$ (the analysis can be generalized to arbitrary number of filters [15]). The claim is that we can model the activation of the set of filter responses:

$$P(l_1, l_2) = \int d\nu P(\nu) \prod_{i=1}^2 P(l_i | \nu, g_i) P(g_i), \quad (12)$$

where $P(l_i | \nu, g_i) = \delta(l_i - \nu g_i)$. In this model the filter responses are generated by independent processes, g_1, g_2 , but then are multiplied by the common factor ν .

In particular, for each filter we can compute $P(g_i|l_1, l_2)$. After some algebra, this is computed to be:

$$P(g_1|l_1, l_2) = \frac{g_1^{-1} \exp\{-\frac{g_1^2 l^2}{2\sigma^2 l_1^2} - \frac{l_1^2}{2g_1^2}\}}{B(0, l/\sigma)}, \quad (13)$$

where $l = \sqrt{l_1^2 + l_2^2}$, and $B(., .)$ is a Bessel's function. To get intuition, note that $g_1 = l_1/\nu$ and $g_2 = l_2/\nu$. So if ν is small then $|l_1|$ and $|l_2|$ are likely to be small together, while if ν is large, then $|l_1|$ and $|l_2|$ are both likely to be large.

Assume that the goal of a model unit is to estimate the g_i from the observed filter responses $\{l_i : i = 1, 2\}$, which gives the non-linear response of the cell. It follows, from analysis above, that

$$E(g_1|l_1, l_2) \propto \text{sign}\{l_1\} \sqrt{|l_1|} \sqrt{\frac{|l_1|}{\sqrt{l_1^2 + l_2^2 + k}}}. \quad (14)$$

The $\sqrt{l_c^2 + l_s^2 + k}$ term sets the gain and performs the divisive normalization.

The model has also been applied to explain the classic tilt illusion in perception [15]. In the “simultaneous” tilt illusion, a set of vertically oriented lines appears to tilt right when surrounded by an annulus of lines tilted left—an effect called “repulsion”. However, for large differences between the center orientation and surround (tilted left), the center vertical lines can appear to tilt left—an effect called “attraction”. In their model, the population of neurons responding to the surround tilted lines contribute to divisive normalizing of the neurons responding to the center stimulus. This results in a change of their neural tuning curves which, together with the degree of coupling between center and surrounds, accounts for repulsion and attraction.

F. Summary

This lecture discussed the basic properties of simple cells in V1. The standard wisdom is that simple cells can be modeled as oriented linear filter which are then subject to divisive normalization. The receptive field properties can be learnt by a variety of models where the typical form – sinusoid modulated by a gaussian fall-off – relates to the shift-invariance properties of images.

REFERENCES

- [1] J. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, 2:1160–1169, 1985.
- [2] D. Hubel and T. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.
- [3] D. Hubel and T. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1):215–243, 1968.
- [4] A. Hyvärinen. Statistical Models of Natural Images and Cortical Visual Representation. *Topics in Cognitive Science*, 2(2):251–264, Apr. 2010.
- [5] J. Jones and L. Palmer. The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6):1187–1211, 1987.
- [6] S. Konishi, A. Yuille, J. Coughlan, and S. Zhu. Statistical edge detection: Learning and evaluating edge cues. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(1):57–74, 2003.
- [7] R. Linsker. From basic network principles to neural architecture: Emergence of orientation-selective cells. *Proceedings of the National Academy of Sciences of the United States of America*, 83(21):8390, 1986.
- [8] R. Linsker. From basic network principles to neural architecture: Emergence of spatial-opponent cells. *Proceedings of the National Academy of Sciences of the United States of America*, 83(19):7508, 1986.
- [9] F. Mechler and D. L. Ringach. On the classification of simple and complex cells. *Vision Research*, 42(8):1017–1033, 2002.
- [10] M. C. Morrone and D. Burr. Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, pages 221–245, 1988.
- [11] E. Oja and A. Hyvärinen. Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4-5):411–430, 2000.
- [12] B. A. Olshausen. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, 1996.
- [13] B. A. Olshausen. Learning sparse, overcomplete representations of time-varying natural images. *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, 1:1–41–4 vol. 1, 2003.
- [14] B. A. Olshausen and D. Field. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*, 37(23):3311–3325, 1997.
- [15] O. Schwartz, T. J. Sejnowski, and P. Dayan. Perceptual organization in the tilt illusion. *Journal of Vision*, 9(4):19.1–20, 2009.
- [16] J. H. van Hateren and D. Ruderman. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1412):2315–2320, 1998.
- [17] M. J. Wainwright and E. Simoncelli. Scale mixtures of Gaussians and the statistics of natural images. *Advances in neural information processing systems*, 12(1):855–861, 2000.
- [18] R. Young, R. Lesperance, and W. Meyer. The Gaussian derivative model for spatial-temporal vision: I. Cortical model. *Spatial Vision*, 3(4):261–319, 2001.
- [19] A. Yuille, D. Kammen, and D. Cohen. Quadrature and the development of orientation selective cortical cells by Hebb rules. *Biological Cybernetics*, 61(3):183–194, 1989.