Lecture 12.2

- This lecture introduces linear models of neurons, describing how they are used to model the receptive fields of neurons in the retina, the LGN, and the *simple cells* in V1. We also describe complex cells in V1.
- Then we provide a different perspective of these cells as representing images and introduce overcomplete bases and sparse encoding.
- This lecture includes two exercises involving interactive demos: (12.2.1) Linear filters and convolution, and (12.2.2) Gabor filters.

Linear models of simplified cells

- This section introduces a model of a simplified cell.
- The cell receives inputs $\mathbf{I} = (I_1, I_2, ..., I_N)$ from *dendrites* that are weighted by *synaptic strengths* $\mathbf{w} = (w_1, w_2, ..., w_N)$.
- These are summed at the *soma* (cell body) to obtain:

$$\mathbf{w} \cdot \mathbf{I} = \sum_{i=1}^{N} w_i I_i$$

.

- The cells outputs a response $f(\mathbf{w} \cdot \mathbf{I})$ along its *axon*, indicated by the firing rate of the neuron. $f(.)$ is a monotonic function (see next slide) but in this lecture we use a linear approximation:

$$S = \mathbf{w} \cdot \mathbf{I} = \sum_{i=1}^{N} w_i I_i$$

The nonlinear function $f(.)$

- $f(.)$ is monotonic nonlinear function, which takes value 0 if the input is small, then increases linearly in the *linear regime* until it saturates at a maximum value.
- A typical choice of $f(.)$ is the sigmoid function $f(\mathbf{w} \cdot \mathbf{I}) = \sigma(\mathbf{w} \cdot \mathbf{I} - T)$, where T is a threshold and $\sigma(.)$ is a soft threshold.
- In this lecture, we ignore $f(.)$ and study the behavior of the model in the linear regime.
- Cells in the retina and the LGN are often modeled without the nonlinear function $f(.)$, adding instead a constant C to the output, to account for spontaneous firing of the cell, and yielding an output $\mathbf{w} \cdot \mathbf{I} + C$, see (Zhaoping, 2014).

Linear filter figure

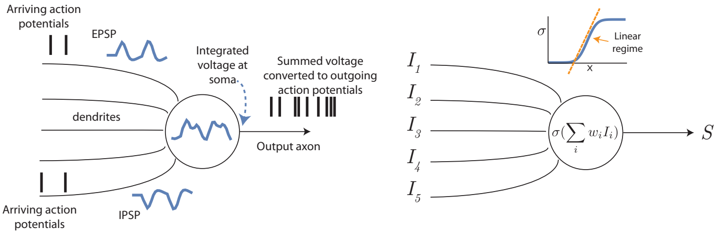

Figure 12 : Left: A neuron receives input – action potentials from other neurons – at its dendrites, which generate excitatory and inhibitory postsynaptic potentials (EPSPs and IPSPs respectively), whose voltages are integrated at the soma and converted to outgoing action potentials. Right: A simplified model of a neuron. There are inputs $(I_1, \ldots, I_5)$ at the dendrites, with synaptic strengths $w_1, \ldots, w_5$. These are summed at the soma, $\sum_i w_i I_i$, and the output S is given by a sigmoid function $\sigma(\sum_i w_i I_i)$. The sigmoid function $\sigma()$ (top right) has a linear regime (brown line) and low and high thresholds.

Linearity and superposition

- This model $S = \mathbf{w} \cdot \mathbf{I}$ is linear in two respects.
- First, it is linear in the input $\mathbf{I}$ so that if we double the input $\mathbf{I} \mapsto 2\mathbf{I}$, then the output doubles also $S \mapsto 2S$. Second, it is linear in the weights $\mathbf{w}$.
- Most importantly, it obeys the *principle of superposition*, so that if S^1, S^2 are the outputs to input $\mathbf{I}^1, \mathbf{I}^2$ respectively, then the output to input $\lambda_1 \mathbf{I}^1 + \lambda_2 \mathbf{I}^2$ is $\lambda_1 S_1 + \lambda_2 S_2$.
- This result is important for characterizing the response of simple neural cells, since it implies that we can determine the output of the cell to any stimulus by observing its response to a limited set of input stimuli $\mathbf{I}$.
- Note that this property still remains if we re-introduce the nonlinear function $f(.)$, provided the function is known.

Retinotopy (I)

- The retinotopic organization of the early visual system has two implications for these cells.
- *First*, the weights of the cell depend on its retinotopic position $\vec{x} = (x_1, x_2)$ and the positions $\vec{y} = (y_1, y_2)$ of its dendrites.
- We replace the input I_i by $I(\vec{y})$ and the weights w_i by $w(\vec{x} - \vec{y})$. The *receptive field* $w(\vec{x} - \vec{y})$ will typically be zero unless $|\vec{x} - \vec{y}|$ is small.
- The neuron is modeled by:

$$S(\vec{x}) = \sum_{\vec{y}} w(\vec{x} - \vec{y}) I(\vec{y}) = \mathbf{w} * I$$

Retinotopy (II)

- *Second*, retinotopy implies that there are cells with similar properties (e.g., the same weights $\vec{w}$) arranged roughly evenly in spatial position (apart from the log-polar transformations (Schwartz, 1980)).
- This can be thought of as having "copies" of the same cell at all positions in space. In terms of linear filter theory, these sets of cells are *convolving* the image $\vec{I}$ by a filter $\vec{w}$.

Receptive fields in retina and LGN.

- The receptive fields of the ganglion cells in the retina and the cells in the LGN can be determined by measuring the firing rate of the neurons in terms of their response to different input stimuli $\vec{I}$ and estimating a model for the response.
- The experimental findings are that many simple cells have a characteristic receptive field called *center-surround*. But these findings are the result of using synthetic stimuli, and cells' response may be more complex if they are studied using natural stimuli.
- Photoreceptors have different properties, see (Rieke et al., 1997).

On-center and off-center receptive fields

- There are two different types: on-center and off-center. The receptive field weights $w(\vec{x} - \vec{y})$ are radially symmetric and take the form of a "Mexican hat" or inverted Mexican hat, for on-center and off-center cells, respectively (Marr, 1982).
- These cell responses are usually thresholded, e.g., by the sigmoid function, so that they usually give only positive responses.
- The weights $w(\vec{x} - \vec{y})$ can be approximated by the *Laplacian of a Gaussian* (LOG) or by its negative:

$$w_{LOG}(\vec{x}) = -\{\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\}G(\vec{x} : \vec{0}, \sigma^2)$$

where $G(\vec{x} : \vec{0}, \sigma^2) = \frac{1}{2\pi\sigma} \exp\{-(x_1^2 + x_2^2)/(2\sigma^2)\}$.

Illustration of center-surround cells

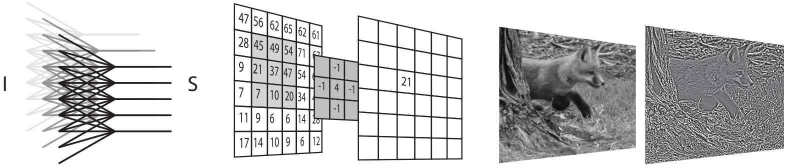

Figure 13 : This figure shows the input-output of a center surround cell (e.g., Laplacian of a Gaussian) in three different ways. First, in terms of the inputs and outputs of neurons (left). Second, in terms of the digitized input image, the filter, and the digitized output (center). The output at each pixel is given by the product of the filter to the appropriate intensity values in the input image, e.g., $4 \times 37 - 1 \times 49 - 1 \times 47 - 1 \times 10 - 1 \times 21 = 21$. Third, in terms of the input and output images (right).

Figure of Gaussians and derivatives of Gaussians

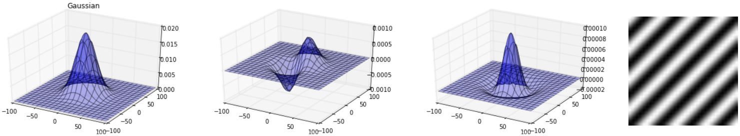

Figure 14 : A Gaussian filter (far left). The first derivative of a Gaussian (left). The Laplacian of a Gaussian, or Mexican hat (right). A sinusoid (far right).

Symmetry and properties of receptive fields

- These cells have two important properties:
  1. They are radially symmetric in the sense that $w_{LOG}(.)$ is invariant to rotation; e.g., suppose we express position $\vec{x}$ in terms of radial components: $x_1 = r\cos\theta, x_2 = r\sin\theta$, then $w_{LOG}(r\cos\theta, r\sin\theta)$ is independent of θ.
  2. The receptive field weights $w(.)$ sum up to zero. More precisely,

$$\sum_{\vec{x}} w_{LOG}(\vec{x}) = 0.$$

- Note that center-surround cells are often modelled as the *differences of two Gaussians*: $w_{DOG}(\vec{x}) = A_1 G(\vec{x} : \vec{0}, \sigma_1^2) - A_2 G(\vec{x} : \vec{0}, \sigma_2^2)$, where σ_1, σ_2 take different values (Zhaoping, 2014). This gives a similar model, if $|\sigma_1 - \sigma_2|$ and $|A_1 - A_2|$ are small.

Purpose of center surround cells: Dynamic range

- These center-surround cells are believed to help deal with the large dynamic range of images.
- Suppose we can express the image locally as $I(\vec{x}) = C(\vec{x}) + B$ where $C(\vec{x})$ is the *contrast*, which describes the local details of the image, and B is the *background*. Then filtering an image by a center-surround cell, whose receptive field sums to 0, removes the background term and preserves part of the contrast.
- More precisely:

$$S(\vec{x}) = \sum_{\vec{y}} w_{LOG}(\vec{x} - \vec{y}) I(\vec{y}) = \sum_{\vec{y}} w_{LOG}(\vec{x} - \vec{y})(C(\vec{y}) + B)$$

$$= \sum_{\vec{y}} w_{LOG}(\vec{x} - \vec{y}) C(\vec{y})$$

Encoding information for transmission

- Receptive fields of this type can also help efficiently encode the information at the retina in order to transmit it efficiently to the visual cortex.
- This can be studied using information theory and the statistics of natural images to predict properties of receptive fields and how they change in different environments (Atick & Redlich, 1992).
- This theory is beyond the scope of this chapter and we refer to the detailed exposition in (Zhaoping, 2014).

Is the retina more complex?

- These models of cells in both the retina and the LGN are well studied. Although many of their properties were estimated using synthetic input data, it has been shown that in some cases, the input image can be estimated from the response of cells in either the retina or the LGN using these types of models Warland et al., 1997; Dan et al., 1996; Carandini:2005).
- But others (Gollisch & Meister, 2010) argue that the retina is more complex, and that, in particular, the neurons may act more as *feature detectors* than as spatial-temporal filters.
- In particular, Gollisch & Meister (2010) describe many findings suggesting that the retina is more complex than the linear filtering model described above. It is known, for example, that if the light levels go down, then the receptive field size becomes larger (Zhaoping, 2014).

Temporal and color properties

- A more realistic model of the output is

$$S(\vec{x}, t) = \sum_{\vec{y},\tau} w(\vec{x} - \vec{y}, t - \tau) I(\vec{y}, \tau)$$

  where $w(\vec{x} - \vec{y}, t - \tau)$ is a space-time filter.
- There are two types of cells with different temporal properties:
  1. M-cells, whose receptive fields are spatially large but temporally small (faster), project to the dorsal stream.
  2. P-cells, whose receptive fields are spatially smaller but temporally larger (slower), project to the ventral stream.
- We can also model the dependence of the cells on the wavelength of the input light by

$$S(\vec{x}) = \int d\lambda w(\vec{x} - \vec{y}) w_c(\lambda) I(\vec{x}, \lambda),$$

  where λ denotes the wavelength and $w_c(\lambda)$ specifies the sensitivity of the cell to color, see (Zhaoping, 2014).

Tuning of receptive fields to sinusoids

- To determine the receptive field of a neuron, we study its response to a class of stimuli while varying the stimulus parameters (i.e., the perceptual dimensions). To find how well the neuron is *tuned* to particular stimulus parameters, see (Hubel, 1982).
- In this section, we analyze tuning when the stimuli are sinusoid gratings.
- We stimulate the receptive field of a neuron by a sinusoid grating

$$I(\vec{x}) = A\cos(\vec{\omega}\cdot\vec{x}+\rho)+I_0,$$

  where A is the *amplitude*, ρ is the *phase*, $\vec{\omega}$ is the *frequency*, and I_0 is the mean light level.
- The frequency specifies the orientation of the stimulus by the unit vector $\vec{\hat{\omega}} = \vec{\omega}/|\vec{\omega}|$, and the period of the oscillation by $|\vec{\omega}|$. The phase ρ shifts the center of the sinusoid. To see this, re-express $A\cos(\vec{\omega}\cdot\vec{x}+\rho) = A\cos(\vec{\omega}\cdot(\vec{x}-\vec{x}_0))$, where $\vec{x}_0 = -\rho\vec{\omega}/|\vec{\omega}|^2$ is the shift in position. If $\rho = 0$, the center occurs at $\vec{x} = 0$.

The response of a center-surround cell to sinusoids

- We assume that the neuron is a center-surround cell and its receptive field is a Laplacian of a Gaussian $w_{LOG}(\vec{x})$.
- The predicted response is:

$$\int d\vec{x} w_{LOG}(\vec{x}) A \cos(\vec{\omega} \cdot \vec{x} + \rho) = A(\cos\rho)(\vec{\omega} \cdot \vec{\omega}) \exp\{-(\sigma^2 \vec{\omega} \cdot \vec{\omega})/2\}.$$

- We deduce three properties:
  1. The response is biggest if the center of the sinusoid is aligned to the center of the cell, i.e., $\rho = 0$, falling to zero at $\rho = \pi/2$
  2. The cell responds best to frequencies with $|\vec{\omega} \cdot \vec{\omega}| = 2\sigma^{-2}$ (by maximizing the response with respect to $|\vec{\omega}|$)
  3. The cell is insensitive to the orientation of the stimuli.
- We can characterize a neuron by measuring its firing rate when it is stimulated with sinusoids. We can use these properties to determine if it is center-surround or not, and if it is, to estimate its parameter σ^2.

Simple cell receptive fields in V1

- The receptive field properties of *simple cells* in V1 were studied by Hubel and Wiesel (1962, 1968) who showed that many cells were *tuned* to the orientation of edges and to the size of bars of light.
- They also showed that these cells were spatially organized with hypercolumns and retinotopic organization. Further electrophysiological studies by Roner and Pollen (1981) and Jones and Palmer (1987) showed that the receptive field properties of these cells could be approximately modelled by *Gabor filters* (Daugman, 1985), which are the product of Gaussians and sinusoids. Derivative of Gaussian filters give an alternative model (Young et al., 2001).
- It was also reported that the receptive fields occur in quadrature pairs (Pollen & Roner, 1981), so that neighboring cells are 90 degrees out of phase (e.g., a cosine Gabor is paired with a sine Gabor).

Gabor filters

- Gabor functions are the product of a Gaussian
  $G(\vec{x};\vec{0},\Sigma) = \frac{1}{2\pi|\Sigma|}\exp\{-(1/2)\vec{x}^T\Sigma^{-1}\vec{x}\}$

  with covariance Σ times a sinusoid:
  $\exp\{i\vec{\omega}\cdot\vec{x}\} = \cos\vec{\omega}\cdot\vec{x} + i\sin\vec{\omega}\cdot\vec{x}$.
- This gives two basic types of Gabors:
  1. cosine-Gabors
  $$G_{\cos}(\vec{x}) = G(\vec{x};\vec{0},\Sigma)\cos\vec{\omega}\cdot\vec{x}$$
  2. sine-Gabors
  $$G_{sin}(\vec{x}) = G(\vec{x};\vec{0},\Sigma)\sin\vec{\omega}\cdot\vec{x}.$$
- These form a *quadrature pair*, because sin(.) and cos(.) are 90 degrees out of phase.

Properties of Gabor filters

- Gabor filters give a good trade-off between *localization* in position and in frequency.
- The Gaussian has good localization in position, in the sense that its response is very small if $|\vec{x}| > 2\sigma$. The sinusoid has perfect localization in frequency (due to the orthogonality of sinusoids) but is unable to localize in position (because a sinusoid does not tend to zero for large $\vec{x}$).
- Gabor derived the Gabor function by optimizing a criterion that balanced optimality in frequency with optimality in position (Daugman, 1985).

Illustration of Gabor filters

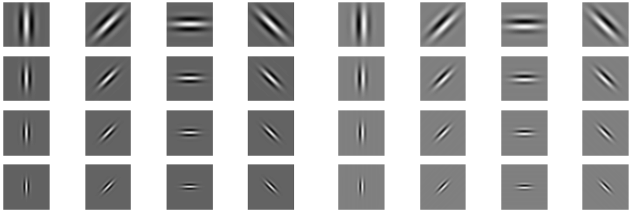

Figure 15 : A family of Gabor receptive fields. The panels show cosine-Gabors (left) and sine-Gabors (right) at different orientations (rows) and different scales (columns). Observe that the cosine-Gabors have biggest responses at their centers (because $\cos 0 = 1$), while the sine-Gabors have small responses there (because $\sin 0 = 0$).

The response of Gabor filters

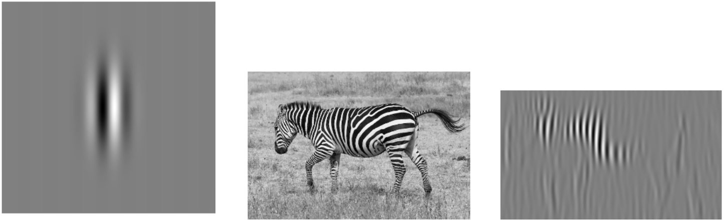

Figure 16 : A Gabor functions aligned to the vertical axis (left). The image of a zebra (center). The response of the vertical Gabor filter on the zebra image (right).

Modelling V1 neurons with Gabor filters

- It has been argued (Lee, 1996) that many simple cells in V1 could be modeled by a family of Gabor filters with specific relationships between the parameters of the Gaussian and the sinusoid, Σ and $\vec{\omega}$. The orientations of the Gaussian and the sinusoid are aligned, and the aspect ratio between the major and minor axes of the Gaussian is 4.
- In more detail, express the frequency of the sinusoid by $\vec{\omega} = \omega(\cos\theta, \sin\theta)$, where θ is its orientation and ω is the frequency. Then the covariance Σ of the Gaussian is proportional to $(1/4)(\cos\theta, \sin\theta)(\cos\theta, \sin\theta)^T + (-\sin\theta, \cos\theta)(-\sin\theta, \cos\theta)^T$ (T denotes vector transform).
- The sinusoid $\exp(i\vec{x} \cdot \vec{\omega})$ has its "propagating direction" along the shorter axis of the Gaussian, so the Gaussian smooths more in the direction perpendicular to the propagating direction, by a factor of $1/2 = \sqrt{1/4}$.

A family of Gabor filters

- This family is specified by:

$$\psi(\vec{x};\omega,\theta,K) = \frac{\omega^2}{4\pi K^2}$$
$$\times \exp\{-(\omega^2/8K^2)\{4(\vec{x}\cdot(\cos\theta,\sin\theta))^2 + (\vec{x}\cdot(-\sin\theta,\cos\theta))^2\}$$
$$\times \exp\{i\omega\vec{x}\cdot(\cos\theta,\sin\theta)\}\exp\{(K^2/2)\}\}.$$

- The variance is proportional to K^2. This is normalized so that $\int d\vec{x}\{\psi(\vec{x};\omega,\theta,K)\}^2 = 1$. $K \approx \pi$ for a frequency bandwidth of one octave, $K \approx 2.5$ for a frequency bandwidth of 1.5 octaves ("octaves" are the log ratio of the frequency – see Zhaoping, 2014).
- This family can also be scaled to give a form:

$$\psi_a(\vec{x};\omega,\theta,K) = \frac{1}{a}\psi_a(\vec{x}/a;\omega,\theta,K)$$

The tuning of Gabor filters (I)

- We study the tuning of Gabor cells by stimulating them with a family of stimuli of form $A\cos(\vec{\omega}\cdot\vec{x}+\rho)$ and varying $\vec{\omega}$ and ρ.
- We define $\omega_x = \vec{\omega}\cdot(\cos\theta, \sin\theta)$ and $\omega_y = \vec{\omega}\cdot(-\sin\theta, \cos\theta)$ to be the projections of the input sinusoid in the favored direction of the cell (i.e., $\vec{\omega}$) and in the orthogonal direction (i.e., $\omega_y = 0$ if the input sinusoid aligns perfectly with the orientation of the cell).

The tuning of Gabor filters (II)

- The responses of the cosine-Gabor G_{cos} and the sine-Gabor G_{sin} are given by:

$$\frac{A}{2}\cos\rho\exp\{-2K^2\omega_y^2/\omega^2\}$$
$$\times\{\exp\{-(K^2/2\omega^2)(\omega+\omega_x)^2\}+\exp\{-(K^2/2\omega^2)(\omega-\omega_x)^2\}\}\exp\{K^2/2\}$$

$$\frac{A}{2}\sin\rho\exp\{-2K^2\omega_y^2/\omega^2\}$$
$$\times\{\exp\{-(K^2/2\omega^2)(\omega+\omega_x)^2\}-\exp\{-(K^2/2\omega^2)(\omega-\omega_x)^2\}\}\exp\{K^2/2\}.$$

- The cosine-Gabor cell is tuned to $\rho = 0$, and the tuning falls off as $\cos\rho$. The cell also favors sinusoid stimuli, which are aligned to it (i.e., $\omega_y = 0$), and whose frequency $\omega_x = \pm\omega$.
- The sine-Gabor prefers stimuli with $\rho = \pi/2$ and has similar tuning to the frequency with $\omega_y = 0$ and $\omega_x = \pm\omega$.

Complex cells

- Complex cells are sensitive to orientation, but they are less sensitive than simple cells to the spatial position of the stimuli. This illustrates the standard theory of the ventral stream: visual processing proceeds up this stream using receptive fields, similar to simple and complex cells, which are increasingly tuned to more complex structures and are less sensitive to the precise positions of the stimuli.
- From this perspective, complex cells are the second stage after simple cells, forming a simple-complex cell module that gets repeated up the hierarchy.

Complex cells energy model

- We describe here the *energy model* where the complex cell receives input from two simple cells that are 90 degrees out of phase (i.e., cosine-Gabors and sine-Gabors). This is partly motivated by quadrature cells (Jones & Palmer, 1987) and partly by these cells being less sensitive than simple cells to the specific position of the stimuli.
- More precisely, the energy model of a complex cell gives response:

$$S(\vec{x}) = \{\psi_{\text{sin}} * I(\vec{x})\}^2 + \{\psi_{\text{cos}} * I(\vec{x})\}^2$$

where $*$ indicates convolution.

Tuning of complex cells

- We study the tuning of complex cells by measuring their response to sinusoid stimuli. The findings show that these cells are, like simple cells, tuned to orientation, frequency, and phase. But their tuning, particularly to phase, is less precise. Hence complex cells are less sensitive to the precise position of the stimuli. The response is given by:

$$\frac{A^2}{4}\exp\{K^2\}\exp\{-4K^2\omega_y^2/\omega^2\}$$
$$\{\exp\{-(K^2/\omega^2)(\omega+\omega_x)^2\}+\exp\{-(K^2/\omega^2)(\omega-\omega_x)^2\}$$
$$+2\cos 2\rho\exp\{-(K^2/\omega^2)(\omega+\omega_x)^2\}\exp\{-(K^2/\omega^2)(\omega-\omega_x)^2\}\}.$$

- Observe that the dependence on the phase ρ is much smaller (the dominant term in the second line is independent of ρ).

Illustration of complex cells

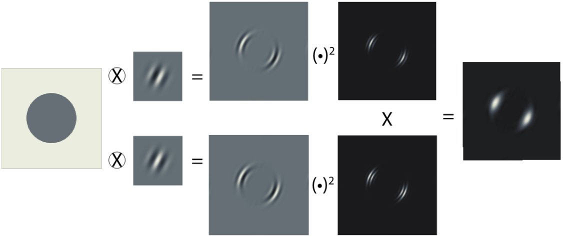

Figure 17 : A complex cell can be modeled as a quadrature pair of Gabor filters. The stimulus is a grey circle on a white background (far left). A quadrature pair of Gabor filters is applied to the stimulus, giving the largest responses when the orientation of the Gabors matches the orientation of the edge of the circle. The responses of the Gabors are squared and then summed to yield the final output (far right).

Complex cells: Complications

- In other models, complex cells are built from simple cells in alternative ways, but the complex cells retain their basic property of being tuned to orientation and frequency but being less sensitive to the position of the stimuli.
- But some researchers question whether complex cells receive input from single cells arguing that the computations could be done by nonlinear neurons that exploit the complexity of the dendritic tree (Mel et al., 1998).
- Other researchers argue (Mechler & Ringach, 2002) that there is no sharp dichotomy between simple and complex cells, but instead there is an continuum of cells with variable sensitivity to position.

Linear filtering and basis functions (I)

- We put these models into the context of the literature on linear filtering and Fourier analysis. This is an advanced section that gives greater understanding but is not required for a basic introduction.
- As discussed earlier, simple cell models apply *linear filters* to images and cells at different spatial locations, performing *convolution* $*$ by applying the same filter $\vec{w}$ across the image:

$$S(\vec{x}) = \vec{w} * I(\vec{x}) = \sum_{\vec{y}} w(\vec{x} - \vec{y}) I(\vec{y}).$$

- It is also convenient to approximate this (take the continuum limit) and express it as an integral:

$$S(\vec{x}) = \int_{\vec{y}} w(\vec{x} - \vec{y}) I(\vec{y}) d\vec{y}.$$

- This continuum limit is a good approximation, if the summation $\sum_{\vec{y}}$ is over a dense set of positions $\vec{y}$, and enables certain type of analysis (e.g., showing that a center-surround cell model sums, approximately, to zero).

Convolving by a Gaussian and derivatives of a Gaussian

Convolving an image by a linear filter produces an output image $S(\vec{x})$ whose form depends on the type of filter $\vec{w}$. For example, if $w(\vec{x})$ is a Gaussian function $G(\vec{x};\sigma) = \frac{1}{2\pi\sigma^2}\exp\{-(x_1^2+x_2^2)/(2\sigma^2)\}$, then convolution effectively just smooths the image by taking a linear weighted average. If $\vec{w}$ is a derivative of the Gaussian in the x_1 direction, $w(\vec{x}) = \frac{d}{dx_1}G(\vec{x};\sigma)$, then this filter gives a large response to *edges*, positions $\vec{y}$ where the intensity $I(\vec{y})$ changes abruptly, and has small responses in places where the image intensity changes slowly.

Linear filtering, basis functions: Fourier analysis (I)

We can better understand images, and linear filtering, by using *functional analysis*. This states that an image, or any signal, can be expressed uniquely as a weighted sum of *basis functions*:

$$I(\vec{x}) = \sum_i \alpha_i b_i(\vec{x}), \tag{1}$$

where the $b_i(\vec{x})$ are basis functions and the $\{\alpha_i\}$ are *coefficients*. These basis functions are usually chosen to be *orthonormal*, so that $\sum_{\vec{x}} b_i(\vec{x}) b_j(\vec{x}) = \delta_{ij}$ ($= 1$ if $i = j$ and $= 0$ if $i \neq j$). If the basis functions are orthogonal, then the coefficients α can be obtained by:

$$\alpha_i = \sum_{\vec{x}} I(\vec{x}) b_i(\vec{x}). \tag{2}$$

Superposition

- The principle of superposition states that we can determine the output S as a weighted combination of the outputs of the basis functions:

$$S(\vec{x}) = \sum_i \alpha_i S_i(\vec{x}), \quad \text{where } S_i(\vec{x}) = \sum_{\vec{y}} w(\vec{x} - \vec{y}) b_i(\vec{y}). \tag{3}$$

- This implies that if we know the response $S_i(.)$ to each basis function $b_i(.)$, then we can predict the response to any input. This is an attractive property that if it holds, enables us to measure the receptive field of a linear neuron, or a thresholded linear neuron, from a limited set of stimuli.

Linear filtering, basis functions: Fourier analysis (II)

Fourier analysis deals with a special class of basis functions. These are sinusoids, i.e., of form $\sin \omega x, \cos \omega x$. The α's are the *fourier transform* of the image. If we restrict ourselves to an image defined on a lattice (i.e., so that x_1, x_2 each take a finite number of values, as on a digital camera), then this is the *discrete fourier transform*. But if we allow x_1, x_2 to take continuous values, then we get the fourier transform:

$$I(\vec{x}) = \frac{1}{2\pi} \int \hat{I}(\vec{\omega}) \exp\{-i\vec{\omega} \cdot \vec{x}\} d\vec{\omega} \tag{4}$$

$$\hat{I}(\vec{\omega}) = \frac{1}{2\pi} \int I(\vec{x}) \exp\{i\vec{\omega} \cdot \vec{x}\} d\vec{x} \tag{5}$$

Here $\exp\{i\vec{\omega} \cdot \vec{x}\} = \cos(\vec{\omega} \cdot \vec{x}) + i \sin(\vec{\omega} \cdot \vec{x})$. Note that if $I(.)$ is symmetric, $I(\vec{x}) = I(-\vec{x})$, then $\hat{I}(\vec{\omega})$ is also symmetric, $\hat{I}(-\vec{\omega}) = \hat{I}(\vec{\omega})$. Observe that equations (4, 5) correspond to equations (1, 2) for special choices of the basis functions (and changing from discrete to continuous $\vec{x}$).

Linear filtering, basis functions: Fourier analysis (III)

Fourier analysis is particularly important because it gives us a way to represent nonlocal structure of images in terms of *frequencies* ω. The high frequencies (large $|\vec{\omega}|$) represent image patterns that change rapidly, while the lower frequencies (small $|\vec{\omega}|$) represent slowly changing patterns. In particular, if an image pattern is *periodic*, like the stripes on a zebra, then it can be expressed in form:

$$I(\vec{x}) = \sum_n A_n \cos(2\pi n \vec{\omega}_0 \cdot \vec{x}),$$

where $\vec{\omega}_0$ is the basic frequency and n denotes integers. Then the Fourier transform is only nonzero at integer multiples of the basic frequency $\vec{\omega} = \vec{\omega}_0$. Hence periodic image patterns, such as *textures*, have very simple descriptions in Fourier space.

Linear filtering, basis functions: Fourier analysis (IV)

- If we blur the image, by convolving with a Gaussian $G(\vec{x};\sigma)$, to obtain $G * I(\vec{x})$, then the high frequencies of the image $\vec{I}$ will be smoothed out. By the *convolution theorem*, the Fourier transform of $G * I(\vec{x})$ is the product of the Fourier transforms of G and $\vec{I}$. The F.T. of a Gaussian is also a Gaussian $\exp\{-|\vec{\omega}|^2(\sigma^2/2)\}$. Hence we can express the convolved image as a weighted combination of sinusoids, where the high-frequency weights are decreased by $\exp\{-|\vec{\omega}|^2(\sigma^2/2)\}$:

$$\vec{I}(\vec{x}) = \frac{1}{2\pi}\int \hat{I}(\vec{\omega})\exp\{-i\vec{\omega}\cdot\vec{x}\}\exp\{-|\vec{\omega}|^2(\sigma^2/2)\}d\vec{\omega}.$$

- If we increase the blurring, by increasing the variance σ^2, we will make the high-frequency coefficients small. Blurring the image can be obtained by defocusing your eyes so that the image is seen out of focus. The receptive fields of cells occurs at a range of different scales, corresponding to convolving with Gaussians of different variances.

Linear filtering, basis functions: Fourier analysis (V)

The superposition principle, combined with the use of basis functions, shows that we can determine the receptive fields of linear neurons by stimulating them with sinusoids. Sinusoids can be used as basis functions, and superposition can be used to predict the response to stimuli that have not been seen yet (i.e., as superpositions of those stimuli to which the response is known). This, however, is rarely done.