Neurons

Visual processing

pictures taken from: Natural Image Statistics (Hyvarinen, Hurri, Hoyer; 2009)
Visual processing

- **Photo receptors.** Mainly:
  - Rods ($> 10^8$)
  - Cones ($\approx 6 \cdot 10^6$, most in fovea)
- $\rightarrow$ **Retinal ganglion cells**
  - Multiple different types ($\approx 1.5 \cdot 10^6$ in total)
- $\rightarrow$ **Optic nerve**
- $\rightarrow$ **Lateral geniculate nucleus (LGN)**
- $\rightarrow$ **Visual Cortex**
  - Ventral stream (V1 $\rightarrow$ V2 $\rightarrow$ V4)
  - Dorsal stream (V1 $\rightarrow$ V2 $\rightarrow$ MT)

Spikes and firing rate

- Neurons communicate information via series of *action potentials* travelling along the axon.
- The *firing rate* depends on the total input collected from other neurons via its dendrites.
- Action potentials and firing rates are complex chemical processes.
- They can be modeled with differential equations whose solutions yield highly idealized models of the dependency of firing rate on total input.

Spikes and firing rate

- This makes it possible to abstract away the spikes and think of real numbers being sent along wires.
- Because of phenomena like *refractory periods* firing rates saturate at a maximum value.
- Some neurons have non-zero minimum firing rate called *spontaneous firing rate*.
- The input collected from other neurons depends on synaptic efficacies, which can be modeled as a weighted sum.

Idealized input-output behaviour

- Together this yields the common idealized input-output model
  $\text{output} = \frac{1}{1 + \exp(-w^T x)}$
- Other sigmoidal functions can be used.
- *(Or completely other non-linearities...)*
Idealized input-output behaviour

- The use of an element-wise nonlinearity which is applied to a linear combination of inputs is sometimes referred to as linear/nonlinear model.
- It is the building block of most practical neural network models.
- And it is key to building multi-layer systems.

Hebbian learning

- The fact that synaptic efficacies can change is the basis for learning.
- The most famous learning rule is Hebbian learning, introduced by Donalb Hebb in 1949: “When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.”
- “Fire together, wire together”.

Hebbian learning

- Hebbian learning is local so it doesn’t need information to travel far to perform weight updates.
- Interestingly, Hebbian-type local learning rules (with some simple twists) can implement most of the modules needed in visual recognition, such as:
  - Linear regression
  - Logistic regression
  - PCA
  - Sparse coding, feature learning

PCA via Hebbian learning

- Hebbian learning with linear neurons:
  \[ \mathbf{w} \leftarrow \mathbf{w} + \eta (\mathbf{w}^T \mathbf{x}) \mathbf{x} \]
- Average weight change:
  \[ \langle \Delta \mathbf{w} \rangle = \langle \eta (\mathbf{w}^T \mathbf{x}) \mathbf{x} \rangle = \langle \eta \mathbf{x} \mathbf{x}^T \mathbf{w} \rangle = \eta \mathbf{C} \mathbf{w} \]
  where \( \mathbf{C} \) is the covariance matrix of \( \mathbf{x} \).
- Weight dynamics as differential equation:
  \[ \frac{d\mathbf{w}}{dt} = \eta \mathbf{C} \mathbf{w} \]
- Solution:
  \[ \mathbf{w}(t) = e^{\eta \mathbf{C} t} \mathbf{w}(0) \]
PCA via Hebbian learning

- Consider discrete time for simplicity (but this works analogously for continuous time):
- Write $w(t)$ as a linear combination of eigenvectors of $C$:
  \[
  w(t) = \sum_{k=1}^{D} a_k(t)v_k
  \]
- Repeated multiplication of $w(0)$ by $C$ yields
  \[
  w(t) = C \cdots C w(0) = C^t w(0) = C^t \sum_k a_k(0)v_k = \sum_k a_k(0)C^t v_k = \sum_k a_k(0)\lambda_k^t v_k
  \]
- “But $\|w\|$ will grow.”
  Solution: Renormalize after every step (eg. Oja’s rule).
- “What about the other eigenvectors $v_2, \ldots, v_D$?”
  Solution: Orthogonalize after each step (eg. Sanger’s rule).

Receptive fields

- The firing rate of a neuron can be highly selective wrt. the input.
- The set of stimuli parameters that make a neuron fire is called receptive field.
Receptive fields

▶ a: Typical receptive fields of retinal ganglion cells and of cells in LGN.
▶ b: Typical receptive fields of cells in V1. These are also known as simple cells.

Retinotopy

▶ The receptive fields of neighboring neurons in LGN and V1 tend to cover neighboring areas of the visual field.
▶ This is known as retinotopic organization.
▶ It means that their joint responses yield the “visual analog of an STFT”.
▶ A convolution layer could be used to simulate this (But this is not the only organizing principle that can be found in visual cortex.)
Gabor filter response

- The response map is the visual analog of an STFT.
- (a) input image, (b,c) filters, (d,e) response maps, (e) amplitude map (more on this later).

Introspective evidence for frequency channels

- To compute local Fourier amplitudes (the spectrogram, (f)) it is necessary to sum over the squares of two filter responses.
- In practice, this requires convolving the image with a non-linear function.

Complex cells

- There do seem to be cells in V1 whose responses are phase independent. They are called complex cells (also proposed by Hubel and Wiesel)
- A simple mathematical model for these (called “energy mechanism”) is to sum over squared simple-cell responses.
- A Fourier feature pair with 90 deg phase difference is known as quadrature pair. (Better than saying “sin/cos”, because phasors in the Fourier expansion may have arbitrary different phase.)
Another organizing principle, besides retinotopy, is spatial organization according to frequency and orientation of filters. This is often called topographic organization. The phase of the filters, in contrast, is random. All attempts to model topography seem to require complex cell like mechanisms. Topographic organization is ubiquitous across the cortex.

Topography (synthetic)

(a) x-position, (b) y-position, (c) orientation, (d) frequency, (e) phase
Saccades

(Wikipedia)