Summary Computational Vision

# Lecture 1

**Topics**: Anatomical basics

**Keywords**

* [Anatomy of the eye]
  + **Fovea**: About 1.5 mm wide, highest density of cones, virtually no rods. Along with the absence of blood vessels and non-photosensitive neurons, it accounts for the high visual acuity capability at the fovea. In primate fovea, the ratio of ganglion cells to photoreceptors is about 2.5 => visual acuity is limited only by the density of cones. Not exactly in the middle of the retina
  + **Blind spot** (aka **optic disc**): The location where ganglion cell axons exit the eye. There are no rods or cones. The optic disc represents the beginning of the optic nerve. The optic disc is also the entry point for the major blood vessels that supply the retina. Carries from 1 to 1.2 million neurons from the eye towards the human brain. The brain deduces the missing information at the blind spot from neighboring information.
  + The photoreceptors are nearest to the brain, while the ganglion cells are farthest from the brain.
* [Basic retinal circuitry]
  + 4 stages of image processing  
    Photoreception, transmission to bipolar cells, transmission to ganglion cells and transmission along the optic nerve. At *each* synaptic stage (i.e. transmission stage) there are also laterally connecting horizontal and amacrine cells.
* [Visual pathways]
  + Image is inverted in retina, i.e. left half of retina contains right half of the visual field. The left primary visual cortex gets input from both eyes, but only from the right visual field.
  + Image processing after retina  
    Retina, (-> optic tract aka optic nerve->), optic chiasm, LGN, (-> optic radiation ->), V1, rest of visual cortex, extrastriate cortex
* The **thalamus** is a symmetrical structure within the human brain. Its function includes relaying sensory and motor signals to the cerebral cortex, along with the regulation of consciousness, sleep, and alertness.
* **[Lateral geniculate nucleus] (LGN)**: Primary relay and filter center for visual information received from the retina. Since it is located in the thalamus, there is one LGN per brain hemisphere.  
  The LGN has six layers, piled onto each other. (???) The LGN in the left hemisphere receives input from both eyes, but only from the left visual field (vice versa with right hemisphere). Moreover, the **ipsilateral** (same side) eye, projects to layers 2, 3 and 5, whereas the **contralateral** (other side) eye projects to layers 1, 4 and 6. Therefore, every LGN neuron is monocular.

|  |  |  |  |
| --- | --- | --- | --- |
| Type | Size | Source / Type of Info | Location |
| M: **Magnocellular** cells | Large | Rods; perception of movement, depth, and small differences in brightness | Layers 1, 2 |
| P: **Parvocellular** cells | Medium | M- and L-cones; perception of color and form | Layers 3, 4, 5, 6 |
| K: **Koniocellular** cells (or “interlaminar”) | Small | S-cones | Between each of M and P layers |

|  |  |
| --- | --- |
| Input | Output |
| Retina (via optic tract)  Layer 6 from V1 (recurrent connection) | Magno layers: V1 layers 4 upper, 6  Parvo layers: V1 layers 4 lower, 6  Konio layers: V1 layers 1, 2 , 3, 5 |

* [Visual cortex] V1: Primary visual cortex: receives all visual input. Begins processing of color, motion and shape. Cells in this area have the smallest receptive fields. V2, V3 and VP: Continue processing: cells of each area have progressively larger RF. MT/V5: Detects motion. V8: Processes color vision.
* Extrastriate cortex [Human extrastriate cortex, Macaque extrastriate cortex]
* **Visual field**: The area which one sees. This area can be mapped onto the retina.
* **Receptive field**: The RF of a visual cell is the region of visual space (or any other space for general cells) in which light influences the cell’s firing activity. The RF of a cell is formed by the RF which synapse (converge) on it. E.g. the RF of a ganglion cell is determined by the RF of “its” photoreceptors.
* [**Lateral inhibition**] (exam question)
  + Receptive field of a bipolar cell: A bipolar cell has a round RF with a center that is excitatory with inhibitory surround (or vice versa). A horizontal (lateral) cell bundles some photoreceptors in such a way that this RF emerges.  
    This kind of wiring is called lateral inhibition and is widely used in the central nervous system. In this case, the horizontal cell is the inhibitory interneuron.
  + Bipolar cell responses: Center: On, Surround: Off stimulus is excitatory (center and surround of cell react excitatory). Conversely, Center: Off, Surround: On stimulus is inhibitory. Center and surround: On is weakly excitatory.
  + Ganglion cells adapt to high voltage of bipolar cells after some time and would completely adapt. Thanks to REM, the signal never remains constant over the time needed.
  + Retinal neurons normally don’t spike (besides GC) but vary membrane potentials.

# Lecture 2

**Topics**: Chromatic vision

**Keywords**

* **Rod cells** are *photoreceptor cells* in the retina. Rods are concentrated at the outer edges of the retina and are used in *peripheral vision*. More sensitive than cone cells, rod cells work best in *dim light conditions*. On average, there are approximately *125 million* rod cells in the human retina.
  + Rod cells respond maximally around wavelength 498nm [Wavelength sensitivity]
* **Cone cells** are *photoreceptor cells* in the retina. Cone cells are densely packed in the fovea centralis, but quickly reduce in number towards the periphery of the retina. Cones work better in *bright light conditions*. There are about *six to seven million* cones in a human eye.
  + In humans, there are 3 types of cones: S, M and L-cones. Their intensity maxima are around wavelengths 437nm (S), 533nm (M), 564nm (L) [Wavelength sensitivity]
* [**Distribution of rods and cones**] In the fovea, cones are smaller in diameter than in the periphery. Within the cones, S-cones are quite rare and completely absent in the fovea. [Cone mosaic].  
  The ratio of S to L and M cones is constant but that of L to M cones can vary significantly between individuals.
* [**Principle of univariance**]: At wavelength-specific intensities, the response curve of a cone cell is the same for all wavelengths. => With only a single cone, it is impossible to distinguish different wavelengths.
* Response of S, M and L cones

,  
where I = Illumination, R = Reflection, S = Wavelength sensitivity. Response is a three-dimensional vector.

* The physical color can consist of infinitely many wavelengths, whereas the perceived color has only three components.
* **Metamery**: The physical color space is infinite-dimensional, since a physical color can consist of arbitrary many wavelengths. Cones map the physical color space to a three-dimensional color space. This mapping is of course not unique. Different spectral wavelength distributions of different light sources can lead to the same color perception. This ambiguity is called metamery.
* **Color constancy**: Color perception doesn’t change drastically even if illumination changes drastically.
* Cones adapt to background illumination => at lower illumination, cones are more sensitive. [Adaptation to dark background]

# From Lecture 7

* [Absolute wavelength sensitivities]: The graph shows the share of blue, green and red light in a testlight between 400 and 700 nm. E.g. for 600nm, there is very much red light and only some green light. Negative values mean that light of a color had to be shone on the testlight area to have a color match. The distribution perfectly matches S, M and L-cone sensitivities.
* [**Receptive fields of LGN neurons**] There are two types of receptive fields of LGN Parvocellular cells: Red-green cells and yellow-blue cells. => Red, green, yellow and blue is preferred in the LGN
* In cortical cells, there are no color preferences over all cells, but the individual cells are much more selective [Cells’ response to color variations]
* Three stages of color processing:
  + Cones: Mapping the infinite-dimensional physical color space to a three-dimensional color space with components: Blue, green, red
  + LGN cells: Preference of opponent colors: Red-green and yellow-blue
  + Cortex: Higher order mechanisms

# Lecture 3

**Topics**: Receptive fields (RF), Ganglion cells (GC)

**Keywords**

* [Responses of an On-Center-Cell] If the center is stimulated, the center reacts with a higher activity. Conversely, if the center is not-stimulated, the center reacts with a decline in activity. The stimulus and the spiking frequency of the GC are in close correlation, although a small adaptation is visible. Open question!
* [Responses of an Off-Center-Cell]

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Frames | Center light stimulus | Surround light stimulus | (Off-)Center response | (On-)Surround response | Overall (center + surround) response |
| 6 | 0 | 0 | 1 | -1 | 0 |
| 2 | 0 | 1 | 1 | 1 | 2 |
| 4 | 1 | 0 | -1 | -1 | -2 |
| 1,3,5,7 | 1 | 1 | -1 | 1 | 0 |

* Assumptions: An Off-Center responds to a off-center light stimulus with an increased spiking frequency relative to the normal frequency, and to an on-center light stimulus with decreased spiking frequency relative to the normal frequency. Vice versa does the On-surround.  
  If center and surround are converged (summed up in this example) by a horizontal cell, an increase or decrease of the normal spiking rate is only caused by a completely “matching” light stimulus (increase) or “anti-matching” light stimulus (decrease).
* Applied to any picture, a RF like the one of a GC serves for edge detection [Response of a ganglion cell to edges], much like a Gaussian filter. The GC or Gaussian filter corresponds to the following discrete stencil:

|  |  |  |
| --- | --- | --- |
| -0.1 | -0.1 | -0.1 |
| -0.1 | 1 | -0.1 |
| -0.1 | -0.1 | -0.1 |

Analytically, the Gaussian filter is a convolution (multiply and sum)

, where R is the filtered image, F is the filter and I is the original image. [Gaussian filter]

* **Linear model of RF**: Modeling the GC as a Gaussian filter only works if receptive fields are linear, i.e.
  + Homogeneity:
  + Superposition:
* Linearity can be checked by using sinusoidal stimuli:
  + Reponses to sinusoids are sinusoids
  + Response can be predicted from the shape of the RF and from the stimulus frequency
* **Nonlinearities** (Lecture 5):
* **Thresholding:** Response is zero until threshold, after that it’s linear.
* **Saturation:** After a certain response intensity (spike /s), the response doesn’t increase anymore. Remarkably, saturation is not a biophysical limitation, but depends for example on the contrast.
* [**Masking**]: If a mask is applied test stimulus, this mask can act inhibitory on the actual stimulus. The response to the mask stimulus is always zero, while it is increasing for the actual stimulus. However, a strong mask stimulus influences the actual stimulus in an inhibitory way. The stronger the mask stimulus, the higher is the “penalty” for the actual stimulus. Mathematically, this problem arises because negative responses are not possible.
* [**Nonlinear model for V1 simple cells**]
* **[Chevreul illusion], Mach bands**: Because the curve of ganglion cells’ response rapidly increases to a maximum if the edge is exactly in the middle of the receptive field, the Chevreul illusion exists. At edges, the darker side seems to be even darker and vice versa, effectively sharpening the edge.
* [**Spatial frequency sensitivity curve**]: A stimulus is ideally as wide as the On-center of a RF captures. Counting the different sizes of On-center widths the distribution as in the figure emerges. In [Spatial frequency sensitivity example], a real example is given, in which one can verify the shape of the curve.

# Lecture 4

**Topics**: LGN cells, simple cells, complex cells

**Keywords**

* [**RF of LGN neurons** (aka circular simple cells)]: Similar to RF of GC: Circular center surround shape. Reacts best on a circular spot of light at a precise position.
* [**RF of V1 simple cells**]: Line and edge detectors. Reacts best at a given orientation of the line. The position of the line must be precise.
* [**RF of V1 complex cells**]: In complex cells, On- and Off-regions are not separated. Complex cells only care about the *change* in light intensity. The preferred orientation is not clear from the receptive field, but the preferred orientation is determined by the simple cells that project on the complex cell (according to the wrong Hubel and Wiesel model)
* ToDo: Hypercomplex cells. An extended line produces a reduction or complete abolition of the response (Hubel & Wiesel)
* [Hubel & Wiesel’s **feed-forward model of simple cells**] Multiple LGN cells can be wired together in such a way that the converged receptive field is similar to the receptive field of a V1 simple cell neuron.
* [Hubel & Wiesel’s **feed-forward model of complex cells**] Multiple simple cells are “overlapped” such that the characteristic complex cell’s receptive field arises. The model doesn’t work.
* The directional selectivity of V1 simple cells is extremely sharp. Sharp means that the cell reacts only on a small frequency range. Frequency means spatial freq. of a sine wave.
* **Retinotopy** [Retinotopic map] is the mapping of the retina to neurons of V1. Neighboring neurons cover neighboring areas of the retina.
* **Cortical magnification**: The fovea has a much larger resolution than the periphery. Cone density in the fovea is much larger than in periphery. Plus, the cortex devotes more space to a degree of visual field in the fovea than in the periphery. That way, the cortex is remarkably uniform and still guarantees to have more machinery ready for the fovea than for the periphery.
  + **Aggregate visual field**: By making a vertical section through the layers of cortex, the corresponding visual fields are slightly scattered. The union of these visual fields is called the aggregate visual field.
  + A shift of 1-2 mm is enough to receive a completely disjunctive receptive field. This holds for the whole cortex, regardless whether the fovea or the periphery is affected. => The smaller the visual fields are, the bigger their number is.

# Lecture 5

**Topics**: Orientation and direction selectivity, Orientation and ocular dominance columns, Nonlinear model, Adaptation

**Keywords**

* **Ocular dominance**: “Cortical cells receive qualitatively similar inputs from the two eyes but for any given cell the densities of the two inputs are not necessarily the same”. E.g. a cell may receive 80% of its input from the left eye and only 20% from the right eye. Extreme case: Monocular cell
* [**Orientation columns**]: The cortex can be structured after the preferred orientation of each cell. This structure is 2D since cells in a column mostly have the same preferences. The structure has no obvious ordering.
* [**Orientation and ocular dominance columns**]: Ocular dominance columns and orientation columns are perpendicular.
* **Ice-cube model by Hubel and Wiesel**: The ice-cube model states that the cortex is structured along the three axes in the following way: x-axis: different orientations, y-axis: different receptive fields in visual field, z-axis: other features like direction selectivity
* [**Adaptation in V1 neuron**]: There are two adaptation effects in a V1 neuron
* Adaptation during response: During the adaptation stimulus, the response of the neuron declines.
* Intracellular adaptation: The response to the test stimulus of the other eye declines as well, compared to the response of the test stimulus before the adaptation stimulus on the other eye.
* Hebbian Learning (Exam question): Cells that fire together, wire together.
* McCullough effect: Vertical lines are associated with green, horizontal with red.

# Lecture 6

**Topics**: Visual motion

**Keywords**

* [**Reichardt detector**]: Explanation of circuit elements:
  + Triangle: Delay of signal
  + Cross: Multiplication; Two signals must arrive simultaneously; otherwise the signal is not forwarded.
  + Minus: Simultaneously arriving signals are subtracted. Needed to ensure direction selectivity of symmetrical Reichardt detector.
* Example: Flow diagram of a symmetrical Reichardt detector for a left to right movement. Delay time: 1ms, all other transfer times: 0ms

|  |  |  |  |
| --- | --- | --- | --- |
| Time [ms] | 0 | 1 | 2 |
| L (Input) | 1 |  |  |
| R (Input) |  | 1 |  |
| M\_L |  | (1 x 1 = ) 1 |  |
| M\_R | (1 x 0 = ) 0 |  | (0 x 1 = ) 0 |
| S | 0 | 1 | 0 |
| Output | 0 | 1 | 0 |

* [Space-time receptive fields and direction selectivity]
* Reverse correlation: A series of randomized black or white light spots on grey frames is displayed in a fast sequence. The response of a cell is recorded. Whenever a spike occurs, the frames from that spike until some point in the past are summed and averaged over time. Black spots count -1, white spots +1 and grey background 0.  
  In principle you just need to set random points over a space-time domain and then you’ll find out the shape of the RF, in a similar way as Monte-Carlo integration.
* **LGN X cell**: Mess, On-Center, Off-Surround. Suddenly, the polarity reverses, then again returning to a mess.
* [**V1 simple cell separable**]: RF(x,t) = G(x)H(t)
* [**V1 simple cell inseparable**]: RF(x,t) ≠ G(x)H(t)
* V1 complex cell
* [**Aperture problem**]. In a small field, it may not be clear, in which direction a line moves. S. en.wikipedia.org/wiki/Motion\_perception. Such a small field is given by the receptive fields. Detecting the direction of motion requires therefore the integration over all receptive fields.  
  Solution: [Intersection of constraints]
* Area MT (aka V5)
* **[Component and pattern direction selectivity]** (polar graph, exam question)
  + Plaids (grids) consist of two gratings (lines), which are oriented perpendicular, to each other. Gratings move always in the perpendicular direction of their orientation. E.g. a vertical bar moves horizontally.
  + Since plaids are a superposition of gratings, their responses can be superimposed as well. However, the prediction is not unique, because the relation between orientation and direction doesn’t hold for a plaid.
  + Therefore, two predictions are possible: The straight line shape shows component selectivity, because plaids with movement directions 45 or 135 degrees correspond to plaids with a vertical bar component. The dotted line shape shows pattern selectivity, because the preferred movement direction matches, but not the orientation of any component.
* Population analysis: In V1, there are virtually only component selective cells, while in MT, there are cells of both types.
* [Perceptual and neural sensitivity] Based on 5 to 10 MT cells, monkey’s performance (guessing right direction) corresponds to cell’s performance. Influence on MT cell shows direct influence on performance [Microstimulation in MT cells]
* [Functional map of direction selectivity in MT]

# Lecture 8

**Topics**: Depth perception

**Keywords**:

* Monocular depth cues: Depth cues that work with only one eye
  + Size
  + Lighting and shadows
  + Interposition
  + Clarity and elevation
  + Perspective
* Binocular depth cues
* **Fixation point**: Point in space to which both foveas project.
* **Binocular disparity**: In binocular vision, images aside the fixation point is projected to two different locations in the left and right retina. Their positions are different even in relation to the fovea.
* **Retinal correspondence**: In the two retinas, there are cell pairs, which are wired together to enable binocular vision. The images from both cells are, if possible, fused together.
* **[Horopter]**: The horopter is an ellipsoid line, which is defined by all intersections arising from the projections of retinal corresponding points. The exact position of the horopter varies along the fixation point
* [**Panum’s fusion area]** / Diplopia: Zone where binocular vision is possible, i.e. the two separate images from both eyes are fused together. In the diplopian zone, objects can’t be fused, and are seen double.
* **Binocular rivalry**: If two completely different images are shown to the two retinas, the brain switches back and forth between the two images.
* Strictly spoken, there are always two RF for a binocular cell: An RF for the right eye and one for the left eye.

Reading Webvision

Reading Ferrier Lecture

There is a constant displacement distance for distinct receptive fields of about 1-2 mm. That means that the smaller receptive fields, the more of them there are.

Because of this distinction, the visual cortex processes stimuli individually for each portion of the visual field. The whole picture is therefore analyzed at other, subsequent stages of the cortical machinery.

Ocular dominance columns are arranged in a very orderly manner: Neighboring cells most likely have the same ocular dominance.

The input cells of a cell X in layer II or III most likely lie in layer IV and are enshrined to a region in the order of millimeters.

Orientation columns: Remarkably orderly; clock- or counterclockwise. There might be sudden eye dominance shifts.

A column may be looked as a functional unit. Columns are in fact rather slabs piled up vertically.

Reading Primers