Computation & Robot Vision: Lecture 1

Sam Barrett

February 3, 2021

1 Camera and image formation

1.1 The human eye

The ultimate goal of Computer vision is to allow computers to *see* in a similar (or better) way to humans.

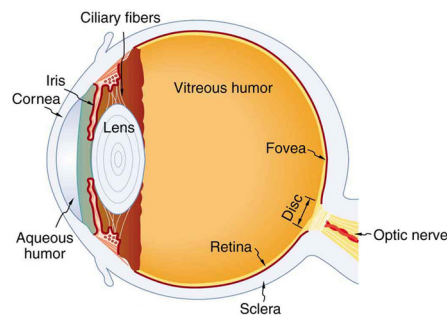We must first start with what we are trying to emulate: the human eye.

Figure 1: Diagram of the human eye

The human eye works by focusing the light entering via the pupil using the lens. This forms an image on the retina.

The retina is made up of many different types of cells, ultimately terminating in the ganglion cells which feed information to the brain via optic nerve fibres. It's composition can be seen in Figure 2.

The ganglion cells collect information regarding the visual world from bipolar and amacrine cells. The information takes the form of chemicals messages emitted by receptors on the ganglion cell's membrane.

1.2 Evolution of cameras

The evolution of artificial cameras started in the renaissance with the likes of Da Vinci experimenting with camerae obscurae (pinhole image projections) and perspective. Later, photographic film was invented, allowing for the capture of
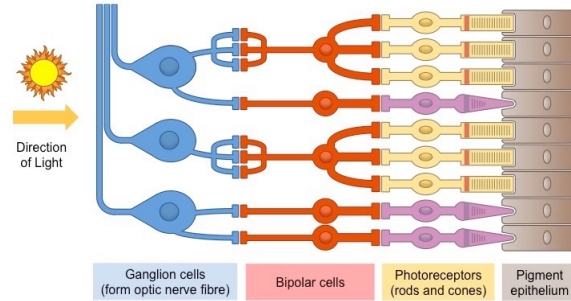
Figure 2: Diagram of the retina

images by exposing the film to different levels of light, focused by a lens the same as we have seen in Figure 1.

We have since gone on to develop entirely digital cameras but the principals remain the same. Here, the photographic film containing photosensitive particles is replaced with a camera *sensor* which are arrays of light-sensitive diodes which convert photons to electrons.
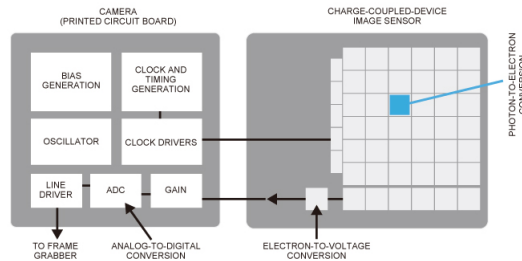

Figure 3: Basic photosensor construction

Digital images are a 2D grid (or matrix) of integers which show the continuous fluctuations in levels of light.

1.2.1 Encoding colour

How do we encode colour into this matrix of integers? We first need to decide on a representation of colour that a computer can understand. We do this by decomposing colours into their primary components. I.e. the different levels of red, green and blue. We can represent **any** colour using weighted combinations of these colours.

Unfortunately, each photodiode is *colour blind* and cannot tell the colour of the light hitting it, only it's intensity. So in order to detect the colour of an image, sensors deploy filters such as the Bayer filter. These are layouts of photo-diodes such that each receptor site is known to only detect light from

a particular part of the spectrum. The Bayer filter can be seen in Figure 4. Using such a filter we can estimate the RGB at each cell using the values of it's neighbours.
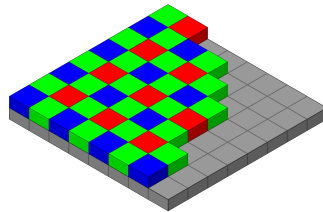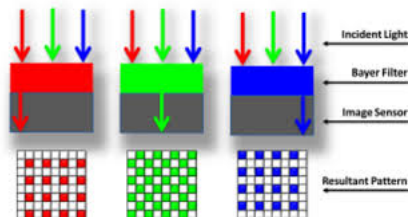

Figure 4: Bayer filter overtop of a sensor

Figure 5: Diagram showing effect of the Bayer filter

Why are there twice as many green as red and blue pixels?

This is done to mimic the higher sensitivity the human eye has towards green light.