

# Computational Vision

## Revision Notes

James Brown

March 31, 2017

Contents

1 Introduction 1

2 Human Vision 1

2.1 Image Formation . . . . . 1

2.2 Retina Processing . . . . . 1

2.3 The Visual Pathway . . . . . 1

3 Edge Detection 1

4 Noise Filtering 2

5 Advanced Edge Detection 2

6 Hough Transform 2

7 SIFT 2

8 Face Recognition 2

9 Motion 2

10 ROC Analysis 2

11 Object Recognition 2

12 Model Based Object Recognition 2

# 1 Introduction

These are notes I have written in preparation of the 2017 Computation Vision exam. This year the module was run by Hamid Deghani (H.Deghani@cs.bham.ac.uk).

Computational vision is the acquisition of knowledge about objects and events in the environment through information processing of light emitted or reflected from objects. In short - we want to make a computer know what is where, by looking through information. We can also use computational vision to do automatic inference of properties of the world from images.

## 2 Human Vision

As humans we have evolved eyes which perceive the visible section of the electromagnetic spectrum, which falls between the wavelengths of 380nm - 760nm. Red light lies at the longer end (760nm) of visible light, and purple at the shorter end (380nm). Visible light is strongly absorbed in the human eye because it can cause an electron to jump to a higher energy levels - yet it does not have enough energy to ionize cells. The evolutionary process of evolving eyes began more than 3 billion years ago with the formation of photopigments. These are molecules where light incident upon them will trigger a physical or chemical change. Photopigments capture photons which lead to the release of energy in the photopigment. This is may be used for photosynthesis or a behavioural reaction (a nerve reaction). A single photoreceptor contains multiple layers to catch light, not just one. This increases the chance of catching any one individual photon - if it's not caught by the first layer it's much more likely to be caught by the second and so on.

### 2.1 Image Formation

Photoreceptors contain a light sensitive patch of photopigments. Using a single cell we can capture light in 1 dimension, but we can't really 'see'. All we can do is tell if the light is on, or off. With multiple cells we can have better direction resolution and with multiple cells we have a very wide aperture - we can't tell exactly where the image is. The image formed will be very bright but extremely fuzzy. This became curved over time and this which really helped with direction resolution as light incident on the left side of the curve must have come from the right. Images formed this way are still very blurry and will result in multiple projections of the same image. Over time eyes evolved to become pinhole cameras which form sharp yet dim images by allowing light to come from a single source (the pinhole) - effectively throwing away loads of potential information about the image. The solution to form sharp and bright images was to use a lens at the front of the eye. The lens focuses all incoming light to a single point and from there we can use our simple pinhole camera for forming images. It should be noted that the image formed is upside down to what exists in reality.

### 2.2 Retina Processing

The human retina contains two kinds of photoreceptors which respond to incident light - **rods** (around 120 million) and **cones** (around 6 million). Rods are extremely sensitive photosensors and respond to a single photon of light. Multiple rod cells converge to the same ganglion cell and therefore neuron within the which results in poor spatial resolution. Rods are responsible simply for detecting the presence of light, and as a result make up the entirety of our night-vision. On the other hand, cones are active at much higher light levels and responsible for the detection of different colours of light. Cones have a much higher resolution as they are processed by several different neurons. Within the eye we have a receptive field, which is the area on which light must fall for a neuron to be stimulated. It should be noted that receptive field in the center of the eye is much smaller than it is for the periphery of the eye.

## **2.3 The Visual Pathway**

Vision is generated by photoreceptors in the retina. All the information captured leaves the eye by way of the optic nerve. There is a partial crossing of axons at the optic chiasm. This is only partial as information from both eyes is sent to both sides of the brain - this allows us to process depth. After this chiasm, the axons are called the optic tract. The optic tract wraps around the midbrain to get to the lateral geniculate nucleus (LGN). The LGN axons fan out through the deep white matter of the brain and ultimately to the visual cortex.

## **3 Edge Detection**

An intensity image is a data matrix whose values represent intensities within some range. Each element of the matrix corresponds to one image pixel.

## **4 Noise Filtering**

## **5 Advanced Edge Detection**

## **6 Hough Transform**

## **7 SIFT**

## **8 Face Recognition**

## **9 Motion**

## **10 ROC Analysis**

## **11 Object Recognition**

## **12 Model Based Object Recognition**

## Index

chiasm, 2

cones, 1

ganglion cell, 1

lateral geniculate nucleus, 2

lens, 1

optic nerve, 2

photocell, 1

photopigment, 1

photoreceptors, 1

pinhole camera, 1

receptive field, 1

retina, 1

rods, 1

visible light, 1

visual cortex, 2