CS131 Notes

Sean Wu

May 19, 2020

Contents

1 Introduction

1.1 What is Computer Vision and why is it hard

Computer Vision : extracting info from digital images OR developing algorithms to understand image content for other applications

- Computer Vision is a hard interdisciplinary problem that is still unsolved
- Hard to convert data storing RGB values in many pixels to semantic info (ex. this blob of black pixels is a chair)
- Vision (extracting meaningful info) is harder than 3D modelling

1.2 Definition of Vision and Comparisons to Human Vision

sensing device : captures details from a scene
interpreting device : processes image from sensing device to extract meaning

- Humans use eyes as sensing devices while computers use cameras
- For sensing devices, computer vision is actually better than human vision because cameras can see infrared, have longer range, and capture greater detail
- For interpreting devices, the human brain is way more advanced than computer systems

1.3 Human Vision Strengths and Weaknesses

- Human vision evolved to quickly recognize danger for survival
- It is very fast $\longrightarrow \sim 150$ ms to recognize an animal
- For speed, humans *focus* only on "relevant" *areas of interest*
- Thus, small signals/changes in the background can be difficult to detect and segment
- Humans also use *context* to infer clues
- Used to determine next area of focus, when to expect certain objects in certain positions, and colour compensation in shadows
- However, context can be used to trick human vision
- Context is very hard to include in computer vision

1.4 Extracting info from images

- 2 types of info extracted in computer vision: **measurements** and **semantic info**

1.4.1 Measurement in Vision

- Robots scan surroundings to make a map of its environment
- Stereo vision gives depth information (like 2 eyes) using triangulation
- Depth info represented as a depth map
- With multiple viewpoints of an object, a 3D surface can be created (or even a 3D model)

1.4.2 Obtaining Semantic Info from Vision

- Labelling objects (or scene)
- Recognizing people, actions, gestures, faces

1.5 Applications of Computer Vision

- Video special effects
- 3D object modelling
- Scene recognition
- Face detection
- Note: face recognition is harder than face detection
- Optical Character Recognition (OCR)
- Reverse image search
- Vision based interaction (ex. Microsoft Kinect)
- Augmented reality
- Virtual reality