Lecture 1
Introduction to Computer Vision

Lin ZHANG, PhD
School of Software Engineering,
Tongji University
Spring 2018
Course Info

Contact Information

Room 408L, Jishi Building
Email: csclinzhang@tongji.edu.cn

TA: Shiyu ZHAO, 1731558@tongji.edu.cn
QQ: 354538123

Course information can be found at

http://sse.tongji.edu.cn/linzhang
Materials

• Major materials
  • My slides

• References
  • Some papers
Examination

• Homework 30%: 3 times, and each time 10%.
• Project 30%: 2 or 3 people for one group
• Final paper exam 40%
• Being absent >=1/3 lectures, you will fail this course
• Bonus 5%: being active in class and answering my questions correctly
Today

- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview
What is vision?

“The plain man’s answer (and Aristotle’s too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is”  

David Marr, Vision 1982

David Marr (1945.1.19 – 1980.11.17), was a British neuroscientist and psychologist. The Marr Prize, one of the most prestigious awards in computer vision, is named in his honor.
What is computer vision?
To bridge the gap between pixels and “meaning”

What we see

What a computer sees

Lin ZHANG, SSE, 2018
What is computer vision?

• Computer vision is the science and technology of machines that can see
• Concerned with the theory for building artificial systems that obtain information from images
• The image data can take many forms, such as a video sequence, depth images, views from multiple cameras, or multi-dimensional data from a medical scanner
What is computer vision?

Image (or video) → Sensing device → Interpreting device → Interpretations

- garden, spring, bridge, water, trees, flower, green, etc.

Source: Feifei Li

Lin ZHANG, SSE, 2018
Our Eyes

Human vision
Human vision

Human Visual System (HVS)

- Optical Receptors
  - Image formation
- Visual Pathway
  - Encoding
  - Representation
- Primary Visual Cortex
  - Interpretation
Vision as a measurement device

Real-time stereo

Structure from motion

Reconstruction from Internet photo collections

NASA Mars Rover

Pollefeys et al.

Goesele et al.
Vision as a source of semantic information
Object categorization

- sky
- building
- flag
- banner
- face
- street lamp
- wall
- bus
- cars

slide credit: Fei-Fei, Fergus & Torralba
Scene and context categorization

- outdoor
- city
- traffic
- ...

slide credit: Fei-Fei, Fergus & Torralba
A little story about computer vision

• In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”

Marvin Minsky  
(1927-2016)  
Turing Award (1969)
Today

- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview
Why computer vision is difficult?

Challenges: viewpoint variation
Why computer vision is difficult?

Challenges: illumination
Why computer vision is difficult?

Challenges: scale
Why computer vision is difficult?

Challenges: deformation

Xu, Beihong 1943

Source: Feifei Li
Why computer vision is difficult?

Challenges: background clutter
Why computer vision is difficult?

Challenges: Motion
Why computer vision is difficult?

Challenges: object intra-class variation

Source: Feifei Li

Lin ZHANG, SSE, 2018
Today

- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview
Why study computer vision?
Lin ZHANG, SSE, 2018

Why study computer vision?

Vision is useful: Images and video are everywhere!

Personal photo albums

Movies, news, sports

Surveillance and security

Medical and scientific images

Google Image Search, Picasa, flickr, webshots, picsearch, YouTube
Why study computer vision?

Vision is useful: Images and video are everywhere!
Visual search—Google

Query image

Output
Mobile visual search—Google Goggles

Google Goggles in Action
Click the icons below to see the different ways Google Goggles can be used.

- Landmark
- Book
- Contact Info
- Artwork
- Places
- Wine
- Logo

Lin ZHANG, SSE, 2018
Photosynth

Project products of students from 2009 Media&Arts
Structure from motion

Bundler: Structure from Motion (SfM) for Unordered Image Collections ([https://www.cs.cornell.edu/~snavely/bundler/#S3](https://www.cs.cornell.edu/~snavely/bundler/#S3))
Automotive safety

- **Mobileye**: Vision systems in high-end BMW, GM, Volvo models
  - “In mid 2010 Mobileye will launch a world’s first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians.”
Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software

Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Source: S. Seitz
Videos based applications

- 视频拆分系统
  - 检测片头片尾
  - 检测新闻故事边界
  - 检测广告和看点

图 视频拆分标注系统

Lin ZHANG, SSE, 2018
Face detection

Many new digital cameras now detect faces

- Canon, Sony, Fuji, ...
Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

Source: S. Seitz
Vision-based biometrics

“How the Afghan Girl was Identified by Her Iris Patterns”
Login without a password...

Fingerprint scanners on many new laptops, other devices

Palmprint system

FKP Video Demo

Finger-Knuckle-Print system
Login without a password...

Palmprint and palmvein system developed in our group
Login without a password...

Towards contactless palmprint recognition: A novel device, a new benchmark, and a collaborative representation based identification approach

Lin Zhang, Lida Li, Anqi Yang, Ying Shen, Meng Yang

School of Software Engineering, Tongji University, China
Collaborative Innovation Center of Intelligent New Energy Vehicle, Tongji University, China
School of Computer Science and Software Engineering, Shenzhen University, China

ARTICLE INFO
Article history:
Received 5 October 2016
Revised 10 March 2017
Accepted 15 April 2017
Available online 19 April 2017

Keywords:
Contactless palmprint recognition
Collaborative representation
Block-wise features
Image quality assessment

ABSTRACT
Biometric authentication has been found to be an effective method for recognizing a person's identity with a high confidence. In this field, the use of palmprint represents a recent trend. To make the palmprint-based recognition systems more user-friendly and sanitary, researchers have been investigating how to design such systems in a contactless manner. Though substantial effort has been devoted to this area, it is still not quite clear about the discriminant power of the contactless palmprint, mainly owing to lack of a public, large-scale, and high-quality benchmark dataset collected using a well-designed device. As an attempt to fill this gap, we have at first developed a highly user-friendly device for capturing high-quality contactless palmprint images. Then, with the developed device, a large-scale palmprint image dataset is established, comprising 12,000 images collected from 600 different palms in two separate sessions. To the best of our knowledge, it is the largest contactless palmprint image benchmark dataset ever...
Face verification

National Stadium, Beijing Olympic Games, 2008
Robotics

NASA’s Mars Spirit Rover

http://www.robocup.org/
Medical imaging

3D imaging
MRI, CT
Demo: Parking-slot Detection
Demo: Fire Detection
You can find a good job!

- Many first-class companies now are developing CV related applications, to name a few
  - Google
  - Microsoft
  - HP
  - Facebook
  - Tencent
  - Baidu
  - iQIYI
  - DJI
  - Huawei
  - ...

Lin ZHANG, SSE, 2018
Today

- What is computer vision?
- Why is computer vision difficult?
- Why do we need to study CV?
- Course overview
Course content

• Introduction
• Local interest point detectors
• Local feature descriptors and matching
• Face detection and face recognition
• Biometrics: Theories and applications
• Basics for machine learning and its applications
• Applications of DCNNs
• *Introduction to numerical geometry
Some tips

• Prerequisites
  • Linear algebra
  • Calculus
  • Matlab Programming
  • C++ Programming

• Knowledge sources
  • IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
  • IEEE Transactions on Image Processing (TIP)
  • International Journal of Computer Vision (IJCV)
  • IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)
  • IEEE International Conference on Computer Vision (ICCV)
  • European Conference on Computer Vision (ECCV)
Thanks for your attention