



计算机视觉表征与识别

Chapter 11: Course Summary

王利民修改

媒体计算课题组

<http://mcg.nju.edu.cn/>



Course summary



- Chapter 1. Introduction (1.5 学时)
- Chapter 2. Images and Filter (2.5 学时)
- Chapter 3. Frequency Domain and Sampling (1 学时)
- Chapter 4. Template, Pyramid, and Filter Banks (4 学时)
- Chapter 5. Edges (2 学时)
- Chapter 6. Segmentation and Grouping (2 学时)
- Chapter 7. Interest Points: detector (1 学时)
- Chapter 8. Interest Points: descriptor (1 学时)
- Chapter 9. Alignment & Transformation (1 学时)
- Chapter 10. Recognition (1 学时)

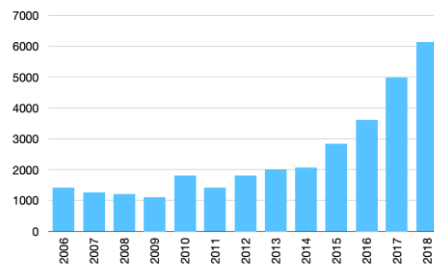


Introduction to computer vision

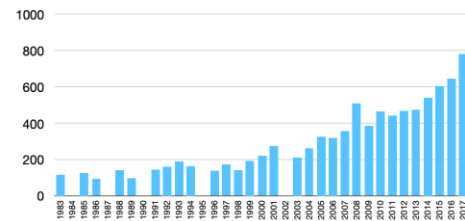


- What is computer vision
 - Extract “meaning” from pixels
 - Geometric information (measurement)
 - Semantic information (perception and interpretation)
- Computer vision is useful
 - Many application: face recognition, surveillance, driver safety, medical images
- Computer vision is difficult
 - Gap between low level signal and high level meanings
- Computer vision is fast developing

CVPR Attendance



CVPR Papers





Images and Filtering



- Basic image formation: light and color
- Image as matrix
- Filtering operation
 - Enhance an image (denoise, resize, etc)
 - Extract information (texture, edges, etc)
 - Detect patterns (template matching)
- Image filters in spatial domain
 - Filter is a mathematical operation on values of each patch
 - Smoothing, sharpening, measuring texture
- Image filters in the frequency domain
 - Filtering is a way to modify the frequencies of images
 - Denoising, sampling, image compression
- Pyramid representation: Gaussian pyramid, Laplacian pyramid, Steerable pyramid
- Texture: Texture classification: filter bank, texon representation, Texture synthesis



Edges



- Gradient and edges
 - Noise, Gaussian Smooth, Derivative of Gaussian, LoG
- Canny edge detector
 - Non-maximum suppression, Linking and thresholding
- Object contour
 - Human segmentation vs. gradient magnitude
 - Berkeley Segmentation Data Set
- Pb edge detector.
 - Brightness, Color, Texture
 - Learning based classifier
- Recent advances in edge detection.
 - Global Pb, Random forest, Deep networks
- Straight line detection
 - Second moment matrix



Segmentation and grouping



- Grouping problem and segmentation
- Inspiration from human perception
 - Gestalt properties
- Segmentation via clustering
 - K-means, GMM, Mean-shift
- Segmentation via graph based method
 - Normalized cut
- Superpixel algorithms
 - Watershed, Felzenszwalb and Huttenlocher graph-based
- Multiple segmentations
 - Hierarchical segmentation, region proposals, vary segmentation parameters



Correspondence and alignment



- Interest point detectors
 - Harris, Hessian, LoG, DoG etc.
 - Region detection: scale, rotation, affine etc.
- Interest point descriptors
 - SIFT, SURF, GIST etc.
 - Deep learning based descriptor
- Fitting and alignment algorithm
 - Least Squares, Hough Voting, RANSAC etc.
- 2D Transformation
 - Translation, similarity, affine, projective etc.
- Instance recognition by alignment
 - Matching, voting, verification
- Image stitching by alignment



Image Recognition



- Introduction to categorization
 - Object, Places, Action, Fine-grained
- Basic task
 - Image classification, object detection, pixel labeling
- Classification pipeline and general issue
 - Spatial pyramid + BoW
 - Feature, Classifier
- Deep learning: CNN & Transformer
 - Large scale dataset, representation learning
 - Transfer learning
- Object Detection:
 - HoG + SVM
 - R-CNN line of detectors
 - YOLO
 - Query based detectors



Not covered



- Image recognition
 - Pixel labeling
 - Human pose estimation
- Geometry & reconstruction
 - Camera model and calibration
 - Epipolar Geometry and stereo
 - Structure from motion
- Video analysis
 - Tracking
 - Optical flow
 - Action recognition
- More advanced topics
 - Vision and language
 - Deep learning and geometry (3D)
 - Weakly supervised learning, self-supervised learning
 - Transformer for vision



Vision as part of an intelligent system



3D Scene



Feature
Extraction

Texture

Color

Optical
Flow

Stereo
Disparity

Grouping

Surfaces

Bits of
objects

Sense
of
depth

Motion
patterns

Interpretation

Objects

Agents
and
goals

Shapes
and
properties

Open
paths

Words

Action

Walk, touch, contemplate, smile, evade, read on, pick up, ...



Important note:

In general, computer vision does not work
(except in certain situations/conditions)

**Hope you learn
something useful!**