COMPUTER VISION 2018 - 2019
> EXAM TRAINING

UTRECHT UNIVERSITY

RONALD POPPE

OUTLINE

Exam criteria

Example questions

Assignment

GRADING

Assignments:

- Follow the steps and you’ll get an 8
- For missing/incorrect steps, points will be deducted
- For extra steps (or insights/testing), points will be added
- Different weight per assignment (bigger ones count more)

Practical assignments: 50%, Written exam: 50%

Retake only if weighted overall grade is >= 4

Final score must be at least 5.5 to pass (minimum for assignments/exam: 4)

No assignment retake!

EXAM CRITERIA

EXAM CRITERIA

You will be graded on:

- Theoretical knowledge
- Conceptual knowledge/insight

Both multiple choice and open questions

- MC mostly for theory. Multiple correct answers per question possible
- Open questions to test understanding, often cross-topic

EXAM CRITERIA^2

Theoretical knowledge. Be able to explain:

- How a method works (SIFT, voxel reconstruction, k-means)
  - Different steps
  - Input/output of each step
  - Relevance of each step
  - (Dis)advantages/limitations of the method
- Differences between methods (histogram vs. GMM)
  - Relative (dis)advantages
- Explain the different terms in specific equations used
  - What are they, what do they represent?
  - How can they be obtained?
  - Derivation of equations if these have been discussed

EXAM CRITERIA[3]

Conceptual knowledge/insight:

- Why are things the way they are (why do we look at the hue channel?, why do we remove keypoints in even areas?)
    - Explain (dis)advantages/limitations
    - Combinations/parallels between topics
- How would you address a certain problem (tracking a red ball with a moving camera)
    - Step-by-step process
    - Explain (pseudo-code or brief sentences) how it works
    - You might be asked to write the pseudo-code for a problem

EXAM CRITERIA[4]

How to answer:

- Brief: longer is not needed
- Specific: I need to be sure (not guess) that you understood

Example: given a set of voxels with (x,y,z) positions, corresponding to several people, how can we find the positions of the people on the ground?

- NOT: “apply a model to optimize the results”
- NOT: “look at how the voxels are spread, and then find the positions such that the people are separated and blabla”
- YES: “apply K-means with Euclidian distance on the (x,z) position of each voxel. The cluster centers correspond to the (x,z) center of each person. The number of clusters should correspond to the number of people.”

EXAM CRITERIA[5]

Example: what are two main limitations of K-means for the clustering of voxels

- NO: sometimes it doesn’t work for people
- YES: the number of clusters should be known in advance and K-means can get stuck in local minima”

Example: write the step-by-step algorithm of silhouette-based 3D reconstruction using look-up table. The algorithm must include 3 functions: background subtraction, initialization of the look-up table, and 3D reconstruction using the look-up table

- You will be provided with a format for your answer

EXAM CRITERIA[6]

You are allowed to make drawings (examples) but these cannot replace your textual explanations

- I should be able to understand your answer just from the text

If you add irrelevant or incorrect information, I might deduct points

- Avoid “hitting all buttons”

Be careful when using “vague” terms such as “much more” and “sometimes”

SOURCES

Slides on the website

Additional reading (for your own understanding):

- Links to websites
- Links to lectures

Insights that were gained during the lectures, in the practical sessions and while working on the assignments

DEPTH

In general, you should be able to:

- Understand each statement in a slide
- Be able to explain it
- Be able to give an example of how something should be applied
- Be able to give an example of a case in which something does/doesn’t work

If you cannot do this, use the additional reading material!

Can you can also post your questions on Slack

TOPICS

1. Introduction computer vision
2. Image formation
3. Silhouette-based volume reconstruction
4. Clustering
5. Image features
6. Optical flow
7. Training, classification, detection
8. Performance measures
9. Neural networks
10. Convolutional neural networks
11. Recent advances in deep learning

1. INTRODUCTION

General background:

- Definition of CV, difference between image processing, graphics
- Applications of CV
- Challenges in CV (in which applications are these important)

2. IMAGE FORMATION

Camera geometry:

- Intrinsics/extrinsics/camera matrix: how to calculate (equations), what is each element
- Calibration: how does it work (algorithm), which are the important parameters, which are the assumptions

Camera radiometry:

- Sensors: how do they work, how do we measure color?
- Distortions: what are they and how/when do they occur?

3. VOLUME RECONSTRUCTION

Depth from images:

- Which ways are there to get depth/3D from images?
- 3D reconstruction: Voxel vs. mesh models: (dis)advantages
- Silhouette-based reconstruction: how does it work (algorithm), look-up table, what can we model (limitations), how to improve speed/memory requirements, how to obtain a mesh model (algorithm)

Background subtraction

- How does it work, equation, assumptions, challenges

4. VOXEL-BASED CLUSTERING

Clustering:

- Algorithm (K-means), steps, role of distance function, limitations

Appearance models:

- Color spaces: which are there, (dis)advantages
- Gaussian mixture models, color histograms: what are they, how to construct them, how to use them?

Voxel-based tracking:

- Algorithm, how to project from 2D to 3D, challenges

5. IMAGE FEATURES

General:

- Applications, properties, invariancies
- Edge descriptors (equations), Canny edge detection (algorithm)

Histograms of oriented gradients:

- What are they, (dis)advantages, elements, how to calculate (algorithm)

Scale-invariant feature transforms:

- What are they, (dis)advantages, algorithm, what is Gaussian filtering, what are pyramids, keypoint detection, scale selection, how is the orientation determined, SIFT matching

6. OPTICAL FLOW

Optical flow:

- Concept, assumptions, challenges, limitations, aperture problem

Lucas-Kanade:

- Concept, rationale, limitations, use of pyramids, KLT tracking

Horn-Schunck:

- Concept, rationale, terms in the equations, limitations

Deepflow:

- Concept, relation to LK and HS, coarse steps

7. TRAINING, CLASSIFICATION, DETECTION

Detection vs. classification

- Common vision tasks
- Input/output, sliding window, image pyramid, Selective Search
- Role of image descriptors
- Bag of words: concept, steps, assumptions

Image classification:

- Role of training/testing, classifier, intra/inter-class variation, challenges, supervised learning process (steps), generalization, overfitting, cross-validation, parameter tuning, negative examples

8. PERFORMANCE MEASURES

Performance measures:

- Precision/recall, F1,PR-curve, AUC, average precision, ROC-curve, AUC
- Single vs. multiclass: confusion matrix
- Detection: IOU, non-maximum suppression

Data augmentation

- Motivation, options

Hard negative mining

- Motivation, how to use, risks

9. NEURAL NETWORKS

Deep learning

- Neurons: binary, ReLu, perceptrons, concepts
- Networks: feed-forward, hidden units, recurrent, limitations, challenges, low vs. high-level features

Function minimization

- Role in training, loss functions, convexity, role of backpropagation
- Overfitting vs. underfitting

10. CNNS

General idea

- Convolutions,inputs/outputs, activation maps, padding, calculating number of parameters, dimensions
- Types of layers: concept, calculation
- Low vs. high layers, semantics

Training:

- Mini-batch, convergence, learning rate
- Speed-up tricks using learning rate

11. RECENT ADVANCES

Regularisation: Overall goal, dropout, batch normalization, drop layers, weight decay, early stopping, model averaging, concepts, when to use, assumption, limitations

Residual connections: Concept, implementation, Inception module, concept

Transfer learning: Rationale, limitations, advantages

Combining models: Rationale, mixture vs. product, dropout, averaging

Normalization: inputs, batch normalization

EXAMPLE QUESTIONS

Four test “exams” online:

- 2015: NOT Q4, Q6, Q7b
- 2016: NOT Q4, Q7b-c, Q8a
- 2018 test: only MC, answers at the end
- 2018: NOT Q14

Actual exam includes MC and open questions:

- Will be similar in length, topics and difficulty to 2018 exam
- Exam questions are for you to understand the type of questions
- Don’t rely on these exams to “guess” which questions will be asked

REQUESTS

“How and when are the DOG results per octave (scale size) combined?”

Does the “coarse scales are interpolated” refer to that?

SIFT[5]

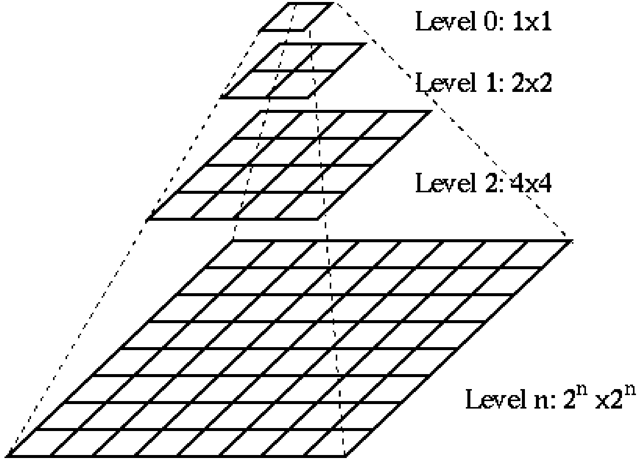
SIFT takes an image and analyses it at different scales

- Each octave/level is half the previous one
- All images together form a pyramid

At each octave, Gaussian filters are applied

- Different levels of variance

Cope with objects of different sizes



SIFT[6]

Example with 3 octaves and 6 levels (increasing σ) of Gaussian filtering

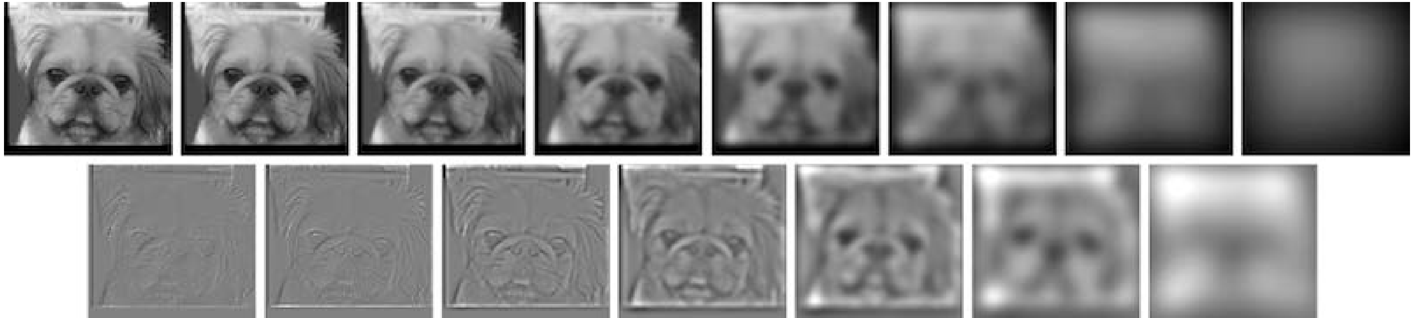
SIFT[7]

Images with subsequent Gaussian filtering are compared pairwise:

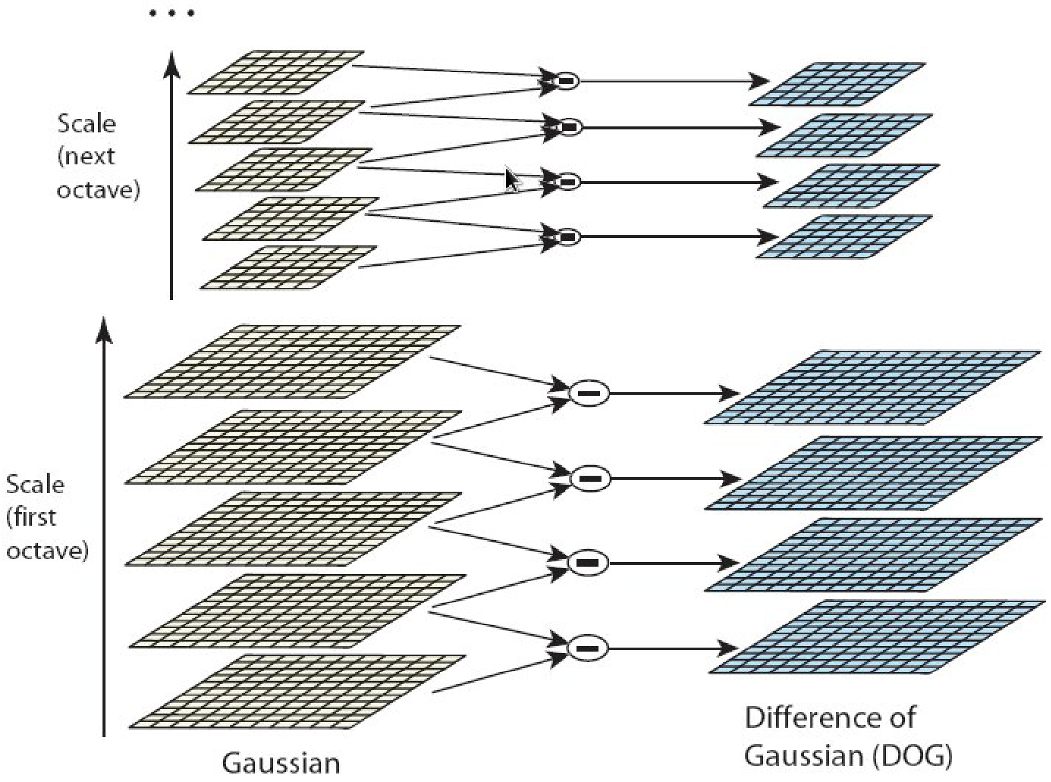
- This is termed a Difference of Gaussian (DOG)

Larger differences correspond to pixels that differ from their surroundings

- These locations are interesting
- Typically edges and corners



SIFT[8]

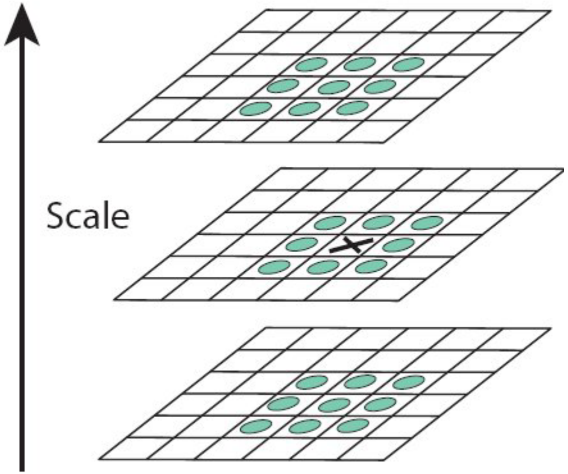


. . .
Scale
(next
octave)
Scale
(first
octave)
Gaussian
Difference of
Gaussian (DOG)

SIFT[9]

Once we have the DOG, we need to select the local minima/maxima:

- Interpolation when going from one octave to the next (remember: always half the size)



Compare each pixel to:

- Its 8 neighbors on the same level
- Its 9 neighbors from scale above
- Its 9 neighbors from scale below

Pixel is selected if it is the maximum

REQUESTS^2

x: $\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$, ***y***: $\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$, ***t* (1st)**: $\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$, **(2nd)** $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ **(Roberts)**

- There is also an alternative (Roberts' cross): *x*: $\begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix}$, *y*: $\begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}$

Where is the hotspot?

- Typically, left upper corner.
- We can also use: *x:* $\begin{bmatrix} 0 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$, *y*: $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & +1 \\ 0 & -1 & 0 \end{bmatrix}$
- Slight shift in correspondence between edge map and image

REQUESTS[3]

x*:** $\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$**,** ***y*:** $\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$**,** ***t **(1st):** $\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$**, (2nd)** $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ **(Roberts)**

How does this work? (*t* first, *t* second?)

- Value to be written on the hotspot is:
  $(I_{x,y,t+1}+I_{x+1,y,t+1}+I_{x,y+1,t+1}+I_{x+1,y+1+t+1}) - (I_{x,y,t}+I_{x+1,y,t}+I_{x,y+1,t}+I_{x+1,y+1+t})$

QUESTIONS?

FINALLY…

COURSE EVALUATION

I hope you have enjoyed the course!

Please give us feedback by filling in the course evaluation form. We always like to improve the course:

- If you have suggestions
- If you thought something was bad
- If you enjoyed something

ASSIGNMENT

Assignment 5:

- Deadline: Sunday April 14, 23:00
- Don’t underestimate the time required
- Reporting is important!

Need help?

- This Thursday last assignment help session
- Slack

EXAM

Tuesday April 9, 13:30-16:30

- EDUC-MEGARON

No books, no slides, no calculator, no phones

But do bring a pen and your student ID card

Any questions, just ask

FINALLY…

Good luck with the exam and assignments!