CS-512 – Assignment 3

Fitting and recognition

Due by: March 20, 2023

Review questions

Answer the following questions. Make sure that your answers are concise. In questions requiring explanations, make sure your explanations are brief.

1. Line detection

   (a) Explain the problem of using the slope and y-intercept as line parameters when using the Hough transform.

   (b) When using the polar representation of lines, what does the vote of each point in the image looks like in the parameter plane?

   (c) Given the point $(1, 1)$ find the vote in parameter space it will cast for angle $\theta = 0$.

   (d) Explain how are lines detected by checking the parameter plane.

   (e) Explain the trade-off regarding bin size in the parameter plane.

   (f) Describe how voting in the parameter plane can be improved if the normal at each voting point is known.

   (g) When using Hough transform for circles, explain what should be the number of dimensions of the parameter space?

   (h) Assume an $m \times n$ image. Assume that the parameters of a line you detected using the Hough transform are given by (θ, d) and that $\theta \in [45^o, 135^o]$ (i.e. a "nearly horizontal" line). Start with the implicit line equation and show that the coordinates of pixels (x, y) on the line may be computed scanning $x \in [0, n]$ and computing $y = -\frac{\cos(\theta)}{\sin(\theta)} x + \frac{d}{\sin(\theta)}$. Explain why we scan x and compute y.

   (i) Assume an $m \times n$ image. Assume that the parameters of a line you detected using the Hough transform are given by (θ, d) and that $\theta \in [-45^o, 45^o]$ (i.e. a "nearly vertical" line). Start with the implicit line equation and show that the coordinates of pixels (x, y) on the line may be computed scanning $y \in [0, m]$ and computing $x = -\frac{\sin(\theta)}{\cos(\theta)} y + \frac{d}{\cos(\theta)}$. Explain why we scan y and compute x.

2. Model fitting

   (a) Explain the disadvantage of using the equation $y = a \cdot x + b$ for line fitting. What kind of lines cannot be fitted accurately using this equation?

   (b) Given a line with a slope of 45 degrees that passes at a distance of 10 from the origin, write the value of the coefficients a, b, c in the explicit line equation $a \cdot x + b \cdot y + c = 0$. To verify your answer, draw the line, find points on it, and see that they satisfy the explicit line equation.

(c) Given the points $(10, 10)$ and $(20, 20)$ write the implicit line equation they define. Write the normalized normal to this line.

(d) Given a line with a normal $(1, 2)$ and distance of 2 from the origin, write the value of the vector l representing this line in the implicit line equation $l^T x = 0$. Note that l is a 3×1 vector.

(e) Given line coefficients $(1, 2, 3)$ find the y-coordinate of the point on this line where $x = 2$.

(f) Explain how to fit a line using the implicit line equation. Write the equation that has to be solved for the unknown line parameters.

(g) Given the points $\{(0, 1), (1, 3), (2, 6)\}$ write the 3×3 matrix that has to be formed to find the parameters of the line that fits the points in homogeneous coordinates.

(h) Explain the difference between algebraic and geometric distance.

(i) Explain how the geometric distance of a point p from an implicit curve $f(p) = 0$ could be measured in an exact way and how it is approximated. Explain the reason for approximating the distance.

(j) Given an implicit curve f and a point p find the algebraic distance of the point p from the implicit curve f. Assume the value of f at p is 1, and that the gradient of f at the point closest to p is 2.

(k) Given an implicit curve f and a point p find an approximated geometric distance of the point p from the implicit curve f. Assume the value of f at p is 1, and that the gradient of f at p is 2.

(l) Write the objective function for active contours, and explain the components in it.

(m) Given three points on a discrete contour where $p_1 = (1, 2)$, $p_2 = (2, 3)$, $p_3 = (3, 4)$, find the continuity and curvature energy at p_2.

(n) Given a point with high curvature (e.g. corner) write the coefficient β that has to be set for that location to guarantee tight fitting of an active contour.

3. Robust estimation

(a) Explain what are outliers, and is the fundamental problem associated with them when fitting a model.

(b) Write the objective function that is used for robust estimation and explain its difference from the standard least squares objective function.

(c) Write the Geman-McClure function for robust estimation and explain its advantage. Explain how the bandwidth parameter σ can be adjusted in an iterative manner.

(d) Let $x = 1$, write the value of the Geman-McClure estimator with parameter $\sigma = 1$.

(e) Explain the principle of the RANSAC algorithm. Explain whether the number of points drawn at each attempt be large or small.

(f) Explain the parameters of the RANSAC algorithm. Write the formula for estimating the number of trials.

(g) Given the desired probability $p = 0.99$ of at least one experiment not having outliers, and a probability $w = 0.9$ that point is an inlier, compute the number of experiments needed to be performed in the RANSAC algorithm.

Programming questions

In this part you need to write a program (in Python using OpenCV) to perform several tasks. You may use a python file or a python notebook for your program. Whenever your implementation is slow (e.g. due to double loops) accelerate it using Cython or Numba.

1. Line fitting
2. Robust estimation
   (a) Generate points belonging to a line segment with given parameters of angle and distance from the origin. Test your line generation by generating points belonging to lines with angles between 0 and 90 degrees, and plotting the points.
   (b) Add Gaussian noise with specified mean and standard deviation to the points generated for a specific line segment and plot the points. Make sure that the noise is visible.
   (c) Given the set of points from the previous step, estimate the line parameters, and compute the error with respect to the known line parameters.
   (d) Plot a graph showing the error as a function of noise level.
   (e) Add x percent of outliers to the point set and estimate the line parameters. Plot a graph showing the error as a function of outliers percent.
   (f) Use the `cv2.fitLine` line fitting function to perform robust estimation using `CV_DIST_HUBER` distance and plot a graph showing the error as a function of outliers percent.
3. Image classification 1
   (a) Download the Kaggle's Cats and Dogs [1] dataset and select a subset of 2000 dogs and 2000 cat images to be used for training, validation, and testing. The use of subsets is intended to simulate the condition of limited data.
   (b) Build a convolutional neural network using several convolution, pooling, and normalization layers, followed by one or more dense layers. Flatten the data between the convolution and dense layers. Evaluate performance on the validation set.
   (c) Modify the data generator to perform data augmentation and evaluate performance on the validation set.
   (d) Replace your colvolution layers with a pre-trained convolution base of VGG16 and evaluate performance on the validation set.
4. Image classification 2
   (a) Download the CIFAR-10 [2] and load the pickled data into your program.
   (b) Build a basic convolutional neural networks with several convolution blocks, where each convolution block contains convolution, pooling, and normalization layers. Flatten the output of the convolution layers and pass it to a single dense layer that will produce the output using softmax activation. Evaluate performance on the validation set.
   (c) Replace the convolution blocks with inception blocks and test performance.
   (d) Replace the inception blocks with residual blocks and test performance.

[1] https://www.microsoft.com/en-us/download/details.aspx?id=54765
[2] https://www.cs.toronto.edu/~kriz/cifar.html

Submission Instructions

Follow the submission instructions of assignment 1.