Queen Mary University of London

ECS709 - Introduction to Computer Vision

Coursework

1) Transformations.

Rotation, translation and skew are useful operations for matching, tracking, and data augmentation.

a) Write a function that takes as input an image *I*, rotates it by an angle θ_1 and horizontally skews it by an angle, θ_2. Write the matrix formulation for image rotation R(.) and skewing S(.). Define all the variables. Note that the origin of the coordinate system of the programming environment you use might be different from the one shown in the lectures.

b) Create an image that contains your name written in Arial, point 72, capital letters. Rotate clockwise the image you created by 30, 60, 120 and -50 degrees. Skew the same image by 10, 40 and 60 degrees. *Complete the process so that all the pixels have a value*. Discuss in the report the advantages and disadvantages of different approaches.

c) Analyse the results when you change the order of the two operators: R(S(*I*)) and S(R(*I*)).
   i) Rotate the image by $\theta_1 = 20$ clockwise and then skew the result by $\theta_2 = 50$.
   ii) Skew the image by $\theta_2 = 50$ and then rotate the result by $\theta_1 = 20$ clockwise.
   Are the results of (i) and (ii) the same? Why?

2) Convolution. (Use Dataset A)

Convolution provides a way of multiplying two arrays to produce a third array. Depending on the designed filter and the intended effect, the kernel can be a matrix of dimensions, for example, 3x3, 5x5 or 7x7.

a) Code a function that takes an input image, performs convolution with a given kernel, and returns the resulting image.

b) Design a convolution kernel that computes, for each pixel, the average intensity value in a 3x3 region. Use this kernel and the filtering function above, and save the resulting image.

c) Use the kernels provided below, apply the filtering function and save the resulting images. Comment on the effect of each kernel.

kernel A

1	2	1
2	4	2
1	2	1

kernel B

0	1	0
1	-4	1
0	1	0

d) Use the filtering function for the following filtering operations: (i) A followed by A; (ii) A followed by B; (iii) B followed by A. Comment the results.

3) Video Segmentation. (Use Dataset B)

A colour histogram $h(.)$ can be generated by counting how many times each colour occurs in an image. Histogram intersection can be used to match a pair of histograms. Given a pair of histograms, e.g., of an input image I and a model M, each containing n bins, the intersection of the histograms is defined as $\sum_{j=1}^{n} \min[h(I_j), h(Mj)]$.

a) Write a histogram function that returns the colour histogram of an input image. Visualize the histogram and save the corresponding figure. For a given video sequence, use the above function to construct the histogram of each frame.

b) Write a function that returns the value of the intersection of a pair of histograms. For a given video sequence, use the histogram intersection function to calculate the intersection between consecutive frames (e.g. between I_t and I_{t+1}, between I_{t+1} and I_{t+2} and so on). Find how to normalize the intersection. Does that change the results? Plot the intersection values over time and the normalised intersection values, and save the corresponding figures. Show and comment the figures in the report.

c) Discuss in the report the following: What does the intersection value represent for a given input video? Can you use it to make a decision about scene changes? How robust to changes in the video is the histogram intersection? When does it fail?

4) Texture Classification. (Use Datasets A and B)

The Local Binary Pattern (LBP) operator describes the surroundings of a pixel by generating a bit-code from the binary derivatives of a pixel.

a) Write a function that divides a greyscale image into equally sized non-overlapping windows and returns the feature descriptor for each window as distribution of LBP codes. For each pixel in the window, compare the pixel to each of its 8 neighbours. Convert the resulting bit-codes (base 2) to decimals (base 10 numbers) and compute their histogram over the window. Normalize the histogram (which is now a feature descriptor representing the window). Show in the report the resulting images.

b) Come up with a descriptor that represents the whole image as consisting of multiple windows. For example, you could combine several local descriptions into a global description by concatenation. Discuss in the report alternative approaches. Using the global descriptor you created, implement a classification process that separates the images in the dataset into two categories: face images and non-face images (for example, you could use histogram similarities). Comment the results in the report. Is the global descriptor able to represent whole images of different types (e.g. faces vs. cars)? Identify problems (if any), discuss them in the report and suggest possible solutions.

c) Decrease the window size and perform classification again. Comment the results in the report.

d) Increase the window size and perform classification again. Comment the results in the report.

e) Discuss how LBP can be used or modified for the analysis of dynamic textures in a video.

5) Object Counting. (Use Dataset C)

Moving objects captured by fixed cameras are the focus of several computer vision applications.

a) Write a function that performs pixel-by-pixel frame differencing using, as reference frame, the *first frame* of an image sequence. Apply a classification threshold and save the results.

b) Repeat the exercise using the *previous frame* as reference frame (use frame I_{t-1} as reference frame for frame I_t, for each t). Comment the results in the report.

c) Write a function that generates a reference frame (background) for the sequence using for example frame differencing and a weighted temporal averaging algorithm.

d) Write a function that counts the number of moving objects in each frame of a sequence. Generate a bar plot that visualizes the number of objects for each frame of the whole sequence. Discuss in the report the implemented solution, including advantages and disadvantages.