Exam 2022 Sample: Question 1

Question

(a) **[APPLICATION QUESTION]** Describe two approaches to the location of the target in the images below as follows: (1) using edge detection and (2) using region detection. For each approach, your solution must consist of series of computer vision techniques and you must provide details of how the techniques will be applied including expected input and output for each technique.

[35 marks]

Edge detection Given the images are already in grayscale, there need not be any changes to the colour space. The first step is to apply a light Gaussian blur, this will prevent any interference from any noise, or colour aberrations. Once the blur has been applied we can use an edge detection technique. One technique which could be applied is Hough Transform for Circles. Given there are three concentric circles this could work quite effectively, we just want to find the largest circle in the image. Looking ahead to the next question though, this might not be the most effective way to solve this problem as Hough Transform is quite computationally expensive and relatively slow.

A more effective solution would be using Canny Edge Detection. This should produce an edge map with the three concentric circles, and potentially the numbers outlined. Next, we will find the contours from the edge map using contour segmentation techniques, such as Boundary Chain Codes. Finally, to identify the location of the target, we can place a bounding box around the contours that have been found (we can select the longest contour if we want to highlight specifically the target), then use this bounding box to find the center of the target.

Region Detection The first step is to once again apply a light Gaussian blur. This will limit the impact any noise or colour aberrations has on the analysis of the image. Next, we can do some connected component analysis on the blurred image to generate contours for potential regions. This will ideally output a region containing at the very least the outside ring of the target. A bounding box can be placed around this and the centre of the box treated as the centre location of the target.

Question

(b) **[APPLICATION QUESTION]** Given a video of a (shooting) target (sample frames shown above) taken from a static camera, describe how you would locate the target and score any shots which appear on the target. Your solution must consist of a series of computer vision techniques and you must provide details of how the techniques will be applied including expected input and output for each technique.

[15 marks]

To the naked eye it is clear the two frames shown are different, getting a computer to recognise this is a different story. Given the background of this scene (the target) remains the same

throughout the video, we can simply compare the current frame to the background frame (the picture on the left), to find any differences (holes). The difference image will show the changed pixels, which will ideally be our scored shots. We will threshold that difference image, anything that isn't black we want to keep. This will generate regions we can create contours from, count the contours and extrapolate out the running tally of scored shots.

Exam 2022 Sample: Question 2

Question

(a) **[COMPARE & CONTRAST QUESTION]** Compare and Contrast:

- tracking using SIFT

- tracking using mean shift

Also comment on which technique would be more appropriate for tracking the moving objects in the road scene below. You must provide a list of differences and similarities between techniques. Each of the differences and similarities must be clearly explained. **NOTE:** Marks will only be awarded for the detailed comparison of techniques. No marks will be awarded for separate descriptions of techniques.

[30 marks]

Question

(b) **[APPLICATION QUESTION]** One of the problems with tracking is locating the objects to track. Given a video of a road scene like that shown above describe how you would locate individual moving objects (which could then be tracked). Your solution must consist of a series of computer vision techniques and you must provide details of how the techniques will be applied including expected input and output for each technique.

[20 marks]

Exam 2022 Sample: Question 3

Question

(a) **[APPLICATION QUESTION]** Using back-projection (and other techniques) describe how you could locate the square blue signs in the image below and how you could transform them to a standard-size square image. Your solution must consist of a series of computer vision techniques and you must provide details of how the techniques will be applied including expected input and output for each technique.

[25 marks]

---

1. Reference image creation.

input Close up, high definition images of the signs

process isolate the signs, convert to HSV and calculate the colour histogram

output histogram representing colour distribution of the sign

2. Back-projection detection

input Images including signs we want to identify

process convert the images to HSV and reference histograms generated in the last step

output probability map indicating potential sign positions

3. Noise reduction and region isolation.

input probability map from back-projection

process apply thresholding to probability map and operations to clean up images and identify sign candidates

output a binary image highlighting likely sign regions

4. Region analysis

input binary image with likely regions

process perform CCA to label and analyse regions

output bounding boxes around regions that match criteria for exit signs

5. Further validation

input regions identified in previous step

process Check for presence of white boxes surrounding white character regions. Place bounding box around each box identified.

output final set of verified white box regions

---

6. Result presentation

input original images and verified locations of signs

process extract out regions and stretch/skew to be square if not already.

output original images seperated out to different square signs skewed or stretched to present the boxes as square

Back projection with different photos of the blue boxes (just the blue part) put together as the reference. This will make the back projection resilient to changing lighting conditions/degrees of wear. This will be used to highlight the blue signs in the image.

Question

(b) **[COMPARE & CONTRAST QUESTION]** Compare and Contrast:

- Chamfer Matching

- Statistical Pattern Recognition

- Robust object detection using a cascade of Haar classifiers

You must provide a list of differences and similarities between techniques. Each of the differences and similarities must be clearly explained. **NOTE:** Marks will only be awarded for the detailed comparison of techniques. No marks will be awarded for separate descriptions of techniques.

[25 marks]

Chamfer Matching Like Statistical Pattern Recognition and cascade of haar classifiers, Chamfer matching requires some preparation, except in place of a training phase a template is provided. When processing images, chamfer, just like the other two approaches, selects features to consider, although in the case of chamfer, these features tend to just edges which can be compared to the binary template provided to the chamfer matching function. Chamfer Matching can be resilient to basic transformation and scaling differences, but may struggle with rotation given it is effectively a generalized template matching function looking for more or less similar shapes (ie. placement of edges) between the template and a given sub-image. Chamfer matching is quite sensitive to noise or very cluttered scenes, additionally, as images get large chamfer matching gets computationally more expensive due to the growing of distance transform calculations.

Statistical Pattern Recognition Just like the other two, there is a "training" phase for SPR. During this training phase, the statistical model learns how to differentiate between different classes of objects using metrics like area, elongatedness, length to width ratio, rectangularity, concavities & holes, etc. SPRs resilience to transformations is dependant on what information was provided during the training phase, but given an adequate sample set, SPR should be fairly comprehensive. Finally, the processing of images for SPR is relatively efficient however the training phase may be more expensive as the statistical model classifies and iterates its model.

Robust Object detection using a cascade of Haar classifiers Haar is the most computationally efficient of the three recognition techniques here. Haar cascade object detection was specifically designed for real-time applications. It is trained on a large feature set which is used to generate classifiers for sub-images, because the classifiers are placed in sequence, if one classifier fails, the classifiers which follow are not tested at all, saving on signficant computational expense. Its training phase uses a machine-learning style approach, finding patterns in positive/negative examples rather than trying to ascribe qualities like SPR does for its training set. Because of this use of pattern recognition it can be used fairly robustly through scale, transformation, and rotation – this can be improved with a training set which includes many scaled and rotated objects.

Exam 2021 Paper: Question 3

Question

(a) **[APPLICATION QUESTION]** Using edge detection describe how you would find the road markings (i.e. the white and yellow lines) in a video from a moving camera giving images such as those in the scenes below. Discuss when your approach might give false positives and what you might do to deal with that. Your solution must consist of a series of computer vision techniques and you must provide details of how the techniques will be applied including expected input and output for each technique.

[25 marks]

Backprojection (using reference white + yellow lines) → thresholding → region classification (handling potential false positives of yellow info boxes)

Question

(b) **[COMPARE & CONTRAST QUESTION]** Compare and Contrast

- Static Background Model

- Median Background Model

- Gaussian Mixture Model

You must provide a list of differences and similarities between techniques. Each of the differences and similarities must be clearly explained. **NOTE:** Marks will only be awarded for the detailed comparison of techniques. No marks will be awarded for separate descriptions of techniques.

[25 marks]

Each of the models are used to compare a "background" image to the current frame to highlight differences. The differences between the models are primarily how (or if) they handle changing backgrounds. And how resilient they are to any changes.

Static Background Model This model does not handle changing background conditions at

all. It simply selects a background frame and compares each "current frame" to that background frame. This causes issues when analysing CCTV for example because lighting conditions may change, or there may be background objects which do not remain static (eg. trees swaying, water rippling).

Median Background Model This model uses the last k frames as a set, and then for each pixel finds the median histogram value. The current frame compares each pixel to the median histogram value to then classify changed pixels. This model is still not resilient enough for moving background items, but can handle changes in background colour (like changing lighting conditions) where the static model fails.

Gaussian Mixture Model This model is able to be resilient to moving background items. This builds a histogram distribution for the historical pixel data for each pixel. It is the most computationally expensive but due to modelling each pixel and the change in each pixel it is most effectively able to extract out the background despite changing background conditions and, curcially, moving background objects like trees swaying in the wind or water rippling.

Exam Assignment 2: Question 1

> **Question**
>
> (a) [APPLICATION QUESTION] Describe how you would reliably locate emergency exit signs such as those shown the top row of images below. Your solution must consist of a series of computer vision techniques and you must provide details of how the techniques will be applied including expected input and output for each technique. [25 marks]

> **Question**
>
> (b) [COMPARE & CONTRAST QUESTION] Compare and contrast:
>
> - Non-Maxima Suppression as used in first derivative edge detection.
>
> - Non-Maxima Suppression as used in the first stage of SIFT.
>
> - Non-Maxima Suppression as used in Moravec corner detection.
>
> - Non-Maxima Suppression as used in the Hough transform for circles where the radius is unknown.
>
> You must provide a list of the differences and similarities between the techniques. Each of the differences and similarities must be clearly explained. NOTE: Marks will only be awarded for the detailed comparison of techniques. No marks will be awarded for separate descriptions of the techniques [25 marks]

1st derivative edge detection reduces the effect lower gradients (first derivatives) might have on the result, reducing the number of double counts for an edge.

SIFT works similarly to moravec. Since we are looking for corners we are looking for the most cornery corner of an edge, non-maxima suppression works here by look for the

Moravec

Hough