

Anthony Poerio (adp59@pitt.edu)

CS1674: Homework 9 - Written

Due: 11/14/2016, 11:59pm

QUESTIONS

1) Briefly, what is the difference between window-based and part-based approaches?

In a **window-based approach**, we extract information (which we will later analyze) from some fixed-size window at each possible position in an image, and at a variety of different scales.

In a **parts-based approach**, we handle scale in a different way. Rather extracting information from fixed-sized windows (iteratively), we look at the image as a whole, and from there we examine the spatial relationship between a set of “parts” that we are interested in.

The essential difference is how we handle finding objects at a variety of scales. In windows-based, we take what is closer to a ‘brute force’ method, examining every possible window/scale combination and looking for our object. Whereas, in the parts-based approach, we examine the whole image at once, and try to fit its attributes against the ‘parts’ of our model, taking a variety of scales into account.

2) What expensive operation does detection involve?

No matter what method of detection we use, we will necessarily have to build and **train** an object model that we can match against.

Training, especially, can be a very computationally expensive operation, because we must analyze a large amount of data in order to find a good model.

3) **Boosting is an iterative process. Name one thing that changes between any two iterations.**

At each iteration of the boosting algorithm, we **change the weights** of features that were incorrectly classified.

This is necessary so that we can re-fit our weak-classifier in the next step, taking the incorrect values more heavily into account.

After a series of such series iterations, we can then combine our weak classifiers into a combination of weak classifiers that—when viewed as a single whole—become a more accurate model than any single weak-classifier alone.

4) **What two components does the score of an object hypothesis (which measures whether or not the object is predicted to exist at a certain location in the image) depend on?**

The object hypothesis' score depends on:

1. The Sum of **weighted appearance scores**
 - a. (i.e. – how closely each element of them object we are examining matches the object we are comparing against)
2. The Sum of **deformation costs**
 - a. (i.e. – we penalize for how far away each part *moved from its expected location* in the greater whole. The farther a part moved from its expected location, the more unlikely it is that we are looking at the expected object.)

Taken together, we first calculate the appearance scores (sum) and then subtract the summed deformation costs.

5) **On what types of bounding boxes do Deformable Part Models work best? (see failure analysis slides)**

Based on the slides available for this subject, it appears that the Deformable Parts Model works best on wide views, as opposed to small views.

But the problem with this generalization is that we are looking at **airplanes** specifically—and airplanes are wider than they are tall (i.e. – horizontal).

Really, what I think the point is, is this: we want to prioritize a bound a box whose orientation (aspect ratio) is matched to the expected aspect ratio of the object we are trying to identify.

As a contrasting example, a when searching for a **standing human**, we would most likely see fewer false positives when we use a taller (vertical) bounding box.

Essentially, we want to select a box that will collect as little irrelevant information as realistically possible.

Some of this may require a person's discretion (for selecting bounding box sizes), but I believe even this could be automated to some degree, based on the known characteristics of the objects we are training on and identifying.