Computer Vision Assignment 1

Student Guide / Answers / Feed-back

John Bastian

Australian Institute for Machine Learning
The University of Adelaide

May 2020

Patch Match Overview

- Assignment 1 involved experimenting with *Patch Match*:
  - a technique for quickly finding approximate, dense, image correspondences
- The aim of the assignment—
  - critically think about the algorithm and its limitations;
  - appreciate that finding correspondences is difficult;
  - understand the trade-offs between various design choices; and
  - write a concise report on your findings.

Overview 2

- The aim of this guide—
  - help frame the thought process for working through the assignment;
  - give an outline of what we hoped to see; and
  - some tips on writing a report for maximum marks :)
- Caveat—
  - this is a *non-exhaustive super-set* of what we hoped to see:
    - ie. it is entirely possible to get full marks without addressing everything in these slides.
  - the figures here are to illustrate ideas, but they're not necessarily the best/final statement in what should be in the report.
  - some of the ideas are hints at what you could have tried, but were *not* expected in the assignment.

PM overview

- Stereo vision requires identifying the same *scene point* in 2 (or more) images
- Matching points involves defining a *visual similarity metric*
  - e.g. based on edge distributions (SIFT, ORB, SURF, BRIEF, Daisy, ...)
  - template / colour matching (SSD, NCC, cosine angle, ...)
- Template matching is simple, but:
  - points are infinitely small→single colour
  - PM uses pulls in neighbouring image points to match *patches*

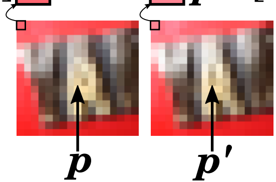
SSD Patch Matching

- Sum-of-squared-differences measures the difference in colour between corresponding pixels between each patch.
- Increasing patch size includes more information

> **Is more information (always?) better?**
> - why would you need more pixels?
> - under what conditions would an accurate match appear to be incorrect?

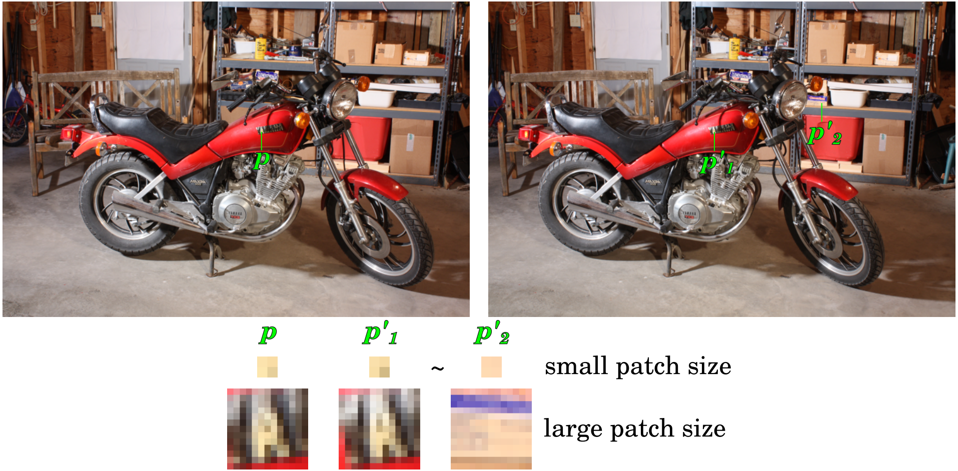
$$\sum\begin{pmatrix}255 & & 255\\ 100 & - & 132\\ 112 & & 147\end{pmatrix}^2 = 2249$$

$\boldsymbol{p}+[-w,-w]$ $\qquad$ $\boldsymbol{p}'+[-w,-w]$



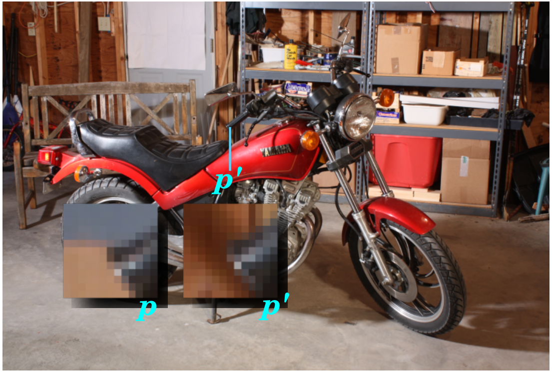
Advantages of Increased Patch Size

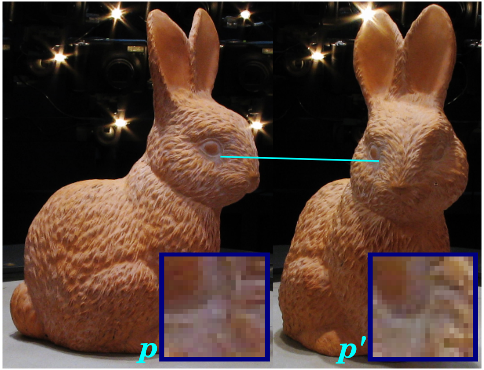
- More pixels → less ambiguity



Disadvantages of Increased Patch Size

- Increasing patch size only reliably works when the scene point and its neighbours move *consistently* between the two images.
- This is unlikely to hold for points
  - on occlusion boundaries; or
  - subject to large perspective changes





Classifying (or *Characterising*) Scene Points

- where template matching works well
  - unique: ie. contains lots of detail
- will struggle with
  - homogeneous regions and repeated textures
  - changes in projection from
    - occlusion (or 'missing' objects);
    - view-point;
    - rotation; and
    - scale
  - changes in lighting from
    - specular highlights;
    - shadows;
    - reflections; and
    - exposure

Spatial Coherence

- Matching points (/patches) independently is *difficult*
- Exhaustively searching over every pair is time-consuming
- Patch Match addresses both issues by exploiting spatial coherency: matches for neighbouring pixels are likely to be the same
    - share information across the image; and
    - randomly search to improve initial guesses

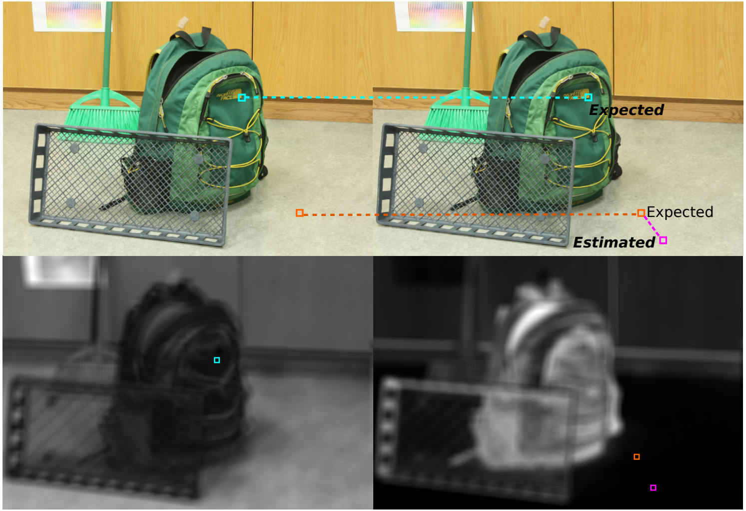
Task 1

- The goal was for you to think about where PM does and does not work
  - $\rightarrow$ find examples to tests those assumptions
- It was imperative that you clearly explain why a pair of images was chosen.
- Needed to include a variety and balance of different characteristics:
  - Not all combinations of issues needed to be tested!
  - Testing difficult pairs is fine, but if they were all unsuitable then it's possible you didn't understand PM. :-(
- Some people cropped larger images to generate ground-truth disparities. This was inspired. Kudos.
  - But this approach has limitations, too.
  - Advanced trick: generated synthetic images with ground-truth!

Task 2 pt1

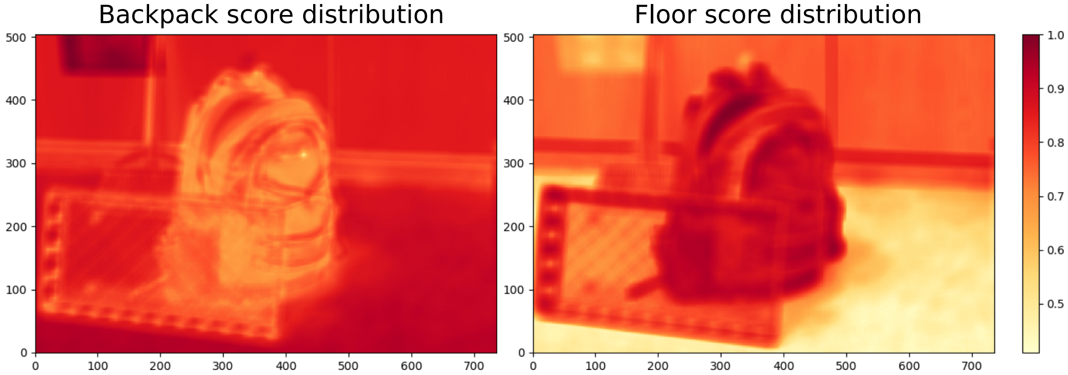
- Test a range of scene-points with different characteristics:
  - homogeneous regions; repeated features
  - unique features
  - advanced: occlusion boundaries, changes in perspective, etc.
- Justify your choice; what hypothesis are you testing?
- *Clearly* present your results
  - Scores close to 0 are more important than scores $> \tau$
  - Identify—
    - the scene-point in the first image and the patch size
    - the *best* match in the second image: not just in the score distribution, but the RGB image

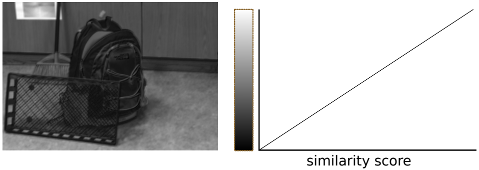
Task 2 Score Distribution 1

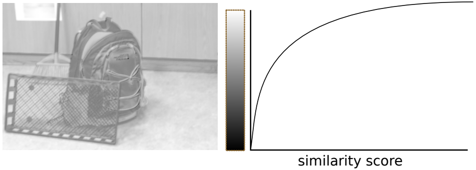


Backpack score distribution   Floor score distribution

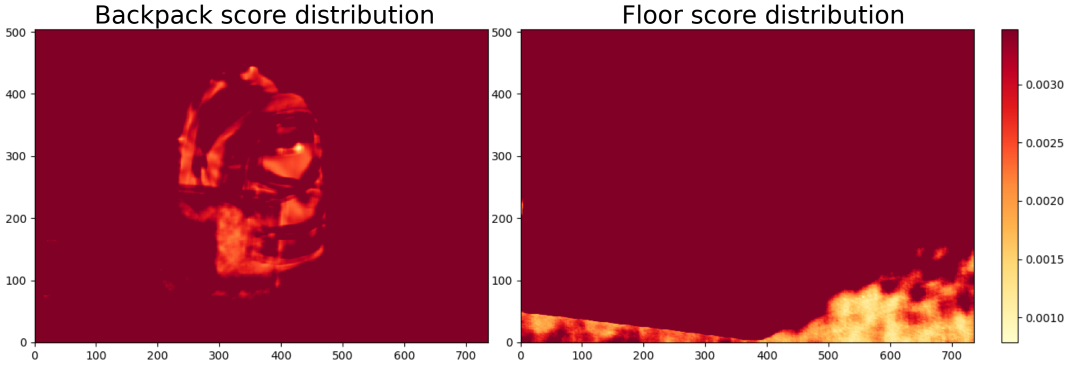
Task 2 Score Distribution 2



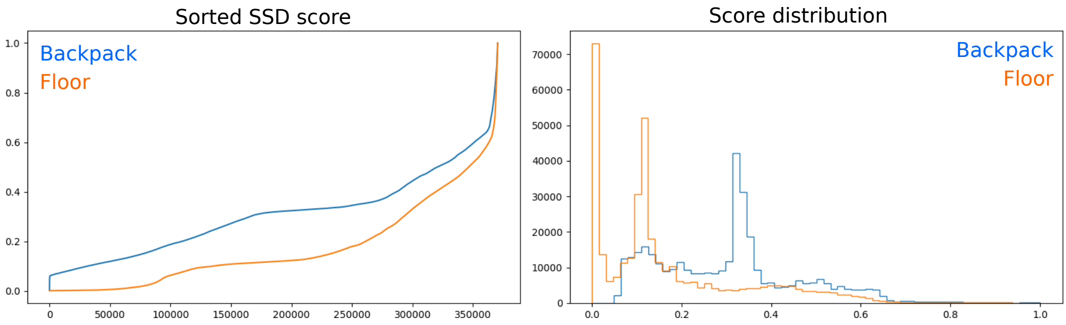




Task 2 Score Distribution 3



Task 2 Score Distribution 4



Task 2 pt 2

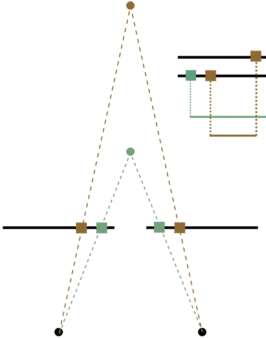
- Were the results what you expected? Why or why not?
  - You were assessed on the quality of your analysis, not whether you were proven right.
- Score-distribution is correlated with patch size and choice of scene-point
  - Implied trick: experiment with both at the same time.
  - Find examples to illustrate that there is not a single best patch size for all points.
- Normalisation should improved results for changes in *exposure*, not *reflectance*

Task 3

- PM uses spatial coherence as a prior
- ... but it doesn't search exhaustively
- Does propagation help?
  - only needed to find *one* case where a patch from T2 had a better estimate in T3!
  - super advanced: show improvement for all points
  - → this is where ground truth disparity + automation would have really helped
- Reconstructing images
  - Check that your PM is behaving sensibly!
  - Emphasises trade-off between patch size and match
    - Large changes in perspective / reflectance can affect results
    - Smaller patch-size gives better reconstructions at the expense of match quality
    - Larger patch-sizes tend not to localise quite as well
    - Errors can also be introduced from PM falling into local minima

Task 3: Disparity and Scene Structure

- Disparity corresponds to *depth*
  - depending on how well your PM works, this isn't always quite clear, though
- Disparity estimation struggles with occlusion
  - Needs to search further afield;
  - Can adversely affect propagation.



General comments

- the assignment was about
  - critically thinking about how the engine behaves
  - designing experiments to test your hypotheses
  - appreciating that *matching is difficult*
- We need to be convinced that you understand the problem
  - it is imperative you back conclusions with evidence
  - run tests, think about the results, and draw conclusions. Repeat as required. :D
- Presentation is extremely important
  - *Organise your thoughts*
  - Be concise
- Make your figures clear
  - give plenty of thought to what you want your figures to convey.
  - refer to them in your text.
  - label axes, including colour legend.
  - make it clear which points aid your discussion.

Structuring your report

- Group your thoughts into a logical, coherent narrative
  - Introduction; supporting statements; conclusion.
  - Separate tasks into sections
  - The report is not a diary!
- Write in paragraphs: one paragraph per idea. Each paragraph should have
  1. Topic sentence: a general statement about the idea contained within;
  2. Supporting statements which clarify and expand on the idea;
  3. Concluding sentence which summarises the idea and segues into the next paragraph.
- Review and revise
  - *Are you saying what you intended.*
  - You have to make it easy for us to give you marks :)

The key to the report is to test your hypothesis about the behaviour of the system and to support your conclusions with evidence. The aim of the practical is to understand the challenges with estimating correspondences, rather than simply implement the PatchMatch algorithm. When conducting your experiments, clearly articulate what hypothesis you are testing by explaining why a particular patch or stereo pair was chosen and try to give some insight into why a particular outcome was observed. Finally, please be assured that computing correspondences is a challenging problem, and *you should anticipate that your results to be incorrect for some cases*, so please do not be discouraged if your results are not perfect. Discovering where it does and does not work is the point of the practical!

Other remarks

- Be wary of not addressing the assignment
  - Don't use SIFT/etc. features or stereo block matching *instead* of PM.
- Not every experiment has value: summarise key findings.
  - e.g. don't report on different patches which have identical characteristics
  - But experiments which are unexpected *are* valuable if your hypothesis is sound!

Where this fits into the Bigger Picture

- Computing correspondences is a fundamental problem in computer vision
- It is challenging:
  - Scene-points can change appearance or be occluded.
- Extensions:
  - Estimate homographies between patches
  - Global optimisation with a smoothness prior and occlusion estimation
  - Better feature descriptors, geometric constraints, etc etc
- How do errors in correspondences affect 3D reconstruction → Assignment 2!
  - This is why you add noise to image points.
  - The second assignment is *not a about solving correspondences.*