Computer Vision 2 Exercises

These exercises are grouped into three parts. Part I contains routine exercises to help you understand the ideas directly presented in lectures. The exercises in Part II are designed to grow and deepen your understanding of principles that underpin computer vision. These exercises extend the lecture material and invite you to think about 'why' questions and about optimisations to what was lectured as well as alternatives to what was lectured. Part III are open-ended, stretching questions that you could tackle in your mini research project.

I hope you enjoy working through these questions, puzzles and research questions!

John Fawcett, July 2023

Part I

1. Why might the correspondence problem be difficult to solve?
2. What is the asymptotic cost of the naive approach to solving the correspondence problem for a set of N edges? What is the asymptotic cost of SIFT on the same input data?
3. The Spatio-Temporal Fourier Transform is actually quite easy to implement – have a go!
4. Explain how intensity gradients can be used to calculate the speed of motion for objects with sharp (in-focus) edges. Why do you not get zero, infinity, or a divide-by-zero error?
5. Would you use the first or second derivative operator for the dynamic zero crossings approach to estimating the speed and direction of motion?
6. Why is it important that codons are invariant to scale, rotation and position?
7. Do we get a *unique* description of a (closed) curve using codons?
8. How would you produce the codon description for an arbitrary curve? How would you match it against the outline descriptions for known objects?
9. Why does the "eigenfaces" approach have that name?
10. Why can eigenfaces provide excellent compression? Would it work for any kinds of image or does it only work on faces?
11. Would Gabor Wavelets give better compression than eigenfaces, on arbitrary images?
12. What advantage can we exploit when a basis set is *orthogonal*?

Part II

1. What extra difficulty arises with the extension of codon descriptions into 3 dimensions?
2. How would you match an eigenface to a database of known faces?
3. Suggest a real-world situation where it might be necessary to run a pre-processing step based on 3D facial modelling.
4. How is it possible for a set of weak classifiers to yield a better overall classifier than any of the individual classifiers?

Part III

1. What is *functional streaming* in the human brain visual system?