

1 Introduction

Computer vision is a field concerned with using computers to process and analyse images. It is an attempt to augment the human ability to perceive visual information. In doing so, it allows users to increase their own capabilities to accurately and quickly extract information from an image.

1.1 Computer Vision

Recognising an object is one of the most important things that human vision is able to do. This means that the brain is able to decipher the basic information received from the eyes, deducing any objects the information describes. For the average human, seeing the local environment and immediately recognising objects within is expected. The act of object recognition is a subconscious one, and the complexity associated with it is not known. The level of information contained in a single moment of vision is extremely high. A normal room contains hundreds or thousands of distinct shapes, all of which must be pieced together to form a cohesive object. Light levels typically vary from one position in a room to another, so an object may appear different on two sides. The human brain is still able to analyse this data extremely rapidly, and make decisions about objects in the image. This task is very intricate task and not simple to replicate.

Consider describing any single object. The adage “a picture is worth a thousand words” has a particular meaning here; fully describing the object would be an extensive task to say the least. Expressing this in a logically defined way increases the level of complexity; exactly how shiny is an apple, what precise shade of brown is the shoe, what is the exact shape of the knife? These are all questions, the answer to which, is difficult to begin to define.

Unsurprisingly, implementing this on a computer is not simple. The normal description of a computer is a machine that excels at calculating mathematical operations at speed. Modern computers are binary devices, designed to store and act upon precise values. Human vision, as we understand it, does not directly map onto the operations a computer excels at.

Despite these difficulties, computer vision is an impressive field. Algorithms such as Scale-Invariant Feature Transform (SIFT) have been developed which are able to find known objects under various transformations. It is used in numerous areas, from robotics control, to automation defect recognition in manufacturing and on to tracking user motions with devices like the Kinect. Computer vision has the potential to be useful in countless fields, and in myriad aspects of everyday life.

1.2 High Performance Computer Vision

One of the major blockades for widespread use of computer vision is the limiting relationships between speed, accuracy and generality. The human eye can detect most objects both quickly and accurately. Computer vision is not so advanced. Achieving real time speeds often comes at the cost of accuracy or generality. On the other hand, creating a fast or generalised system, if possible, comes at the cost of a greatly increased completion time.

For each way to describe an object, several different techniques emerge to locate the object. Each technique has varying requirements and reliability. Some need specific descriptions (e.g. a sample image) and produce very reliable results. Others do not require such precise descriptions, but do not create similarly dependable

results. By combining these techniques together, the reliability and robustness of computational object recognition can be improved. Using all available techniques will considerably increase the time the vision program takes to complete. This can be mitigated by computing each techniques in parallel.

An intuitive method of parallelism for this type of problem is the task farm. This means that general tasks (in this case, each method of identifying an object) will be run concurrently. Task farms are useful in this case, because it allows the use of serial libraries and algorithms that would otherwise complicate or prevent parallelism. Computing results simultaneously, a task farm allows computer vision programs to produce results quickly, reliably and generally.

1.3 Where's Wally as a Test Case

A good way of testing the task farm method is with Where's Wally puzzles. These are simple cartoons, normally a large image filled with various characters, who wear simply coloured clothing, see Figure ?? . One of these characters is the eponymous Wally, who is dressed distinctly from most others characters, see Figure ?? . Similarly dressed characters exist, Figure ?? , adding some complexity to finding Wally correctly. The cartoon nature of the characters means that shapes are boldly coloured and often bordered by a black line. As Where's Wally is a puzzle, sometimes Wally will be hard to find; he is often obscured, camouflaged or simply small. Creating a program to solve Where's Wally is non-trivial, requiring a combination of computer vision techniques. Despite this, the puzzle provides a simple base to implement parallel object recognition.

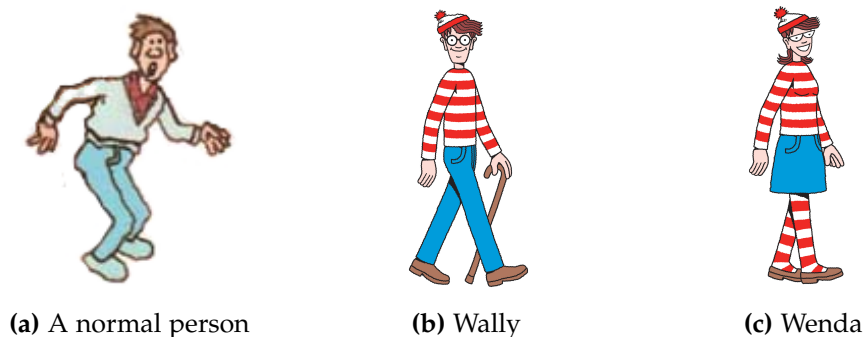


Figure 1: Characters from Where's Wally

1.4 Potential Use Cases

Extending beyond Where's Wally puzzles, parallel object recognition is a useful technique for many areas. The UK Missing Person Bureau released data on missing persons in 2010-2011 [?]. During this period, two thirds of missing people in the UK were under the age of 18. Such a group is particularly vulnerable to abduction and abuse, if left unsupervised. Although the majority of missing people were found within the 5 miles of their homes, up 21% of people were further out. A 5 mile radius is a large area to look for one person, especially in an urban area. Further out, and it becomes untenable to search efficiently. Almost 4/5 of missing people are found within the first 16 hours. According to information released from a survey of young runaways [?], 34% of said they had been harmed or in a risky experience more than once, and 11% expressly said they had been hurt or harmed. At risk individuals, such as children, need to be found rapidly. It is critical to find these children, before they are harmed. Computer vision is already used to facilitate searches, but the vol-

ume of data available can exceed the capabilities of existing programs. Furthermore, surveillance data from CCTV equipment produces data at real time speeds, so faster than real time processing speeds are desired. Implementing a parallelised version of image recognition tools should greatly reduce the time an image needs to search. Governmental agencies, such as the police, could have access to very large computer systems, allowing for a high degree of parallelism.

Wildlife conservation is a delicate task, which could benefit from rapid computer vision techniques. To accurately know which species are endangered, it is important to have an accurate count of the members of the species for a given region. Actively surveying the population, by means of physically interacting with members can have adverse effects on the population. Nielsen [?] discusses the negative effects of electrofishing on rare fish populations. She goes on to indicate the lack of non-invasive methods of surveying the population, without which population counts cannot be maintained. Directly surveying endangered species can be inefficient, slow or dangerous to either the researcher or the animal in question. Passive techniques, such as photography, allow the researcher to estimate populations without interacting with the environment. Ideally a researcher would be constantly vigilant and able to immediately identify each species correctly. This is rarely the case; a single human is fallible and a team may be beyond the funding of the endeavour. Instead, with access to any modern laptop and a digital camera, parallel computer vision may be able to assist in many ways. A video feed would allow observation for as long as the battery lasts, and a database of the features of regional species would help with identification. Parallel species recognition would allow multiple species to be surveyed at a time. It could even allow non-experts to survey the populations, freeing up researchers for more critical tasks.

The final suggestion for the use of parallel object recognition is with nationwide traffic monitoring. Using existing roadside cameras, such as CCTV or speed cameras, a network could be built that monitors traffic on a large scale. This would help to improve commute times and general congestion issues. Parallelism would be of use here, as the sheer quantity of data for a large scale system like this would prevent real time analysis.

1.5 Goals

In this report, Where's Wally puzzles will be used as a testbed to determine if High Performance Computer Vision is advanced enough for everyday use. This will include the production of a suite of functions that, though tuned to locating Wally, could be used to find other characters. The use of directive based parallelism will be added, to determine if a user-friendly system is viable. The generality of the suite of functions will be tested on non-Wally puzzles, to determine how generalised the system is.

1.6 Overview of Report

This report will begin by discussing the underlying information required to easily comprehend parallel object recognition. Included within, will be literature reviews as appropriate. The next section will contain a discussion of the patterns used to recognise Wally. Each will include an analysis of the algorithm selection and a discussion of the level of parallelism that can be exposed. This will then be combined to produce an implementation of a Where's Wally solver, with an examination of the parallelism that can be used.

The report will comment on the results produced by the patterns and the solver. Following this, will be the concluding statements and ideas, along with recommendations for future use. The report will close with an evaluation of the project as a whole, including the differences between the preparation phase and the actual report period.