☎ 02392 248250
🖷 02392 261014
www.smartcctvltd.com
info@smartcctvltd.com

White Paper

Comparing the results of manual enumeration and Video Analytics based enumeration of video surveys.

Introduction

It is estimated that over 30,000 surveys are under taken each year in the UK to count the number of people and / or vehicles travelling in a defined location. Historically these surveys have been conducted over a 12 hour period (normally 7.00am to 7.00pm) on a single day. This time base derives from the effective maximum time a human enumerator could be on site and the cost of labour.

More recently many surveys have been conducted by capturing a video record and processing this manually afterwards. This has the advantage that a permanent record is created allowing any unusual results to be reworked to see if there was an error in the enumeration. There has always been a concern as to both the accuracy of the results and more importantly the cost associated with the enumeration. The accuracy is not only dependent on the diligence of the human enumerators but significantly due to the short time frame of the survey. For example a road traffic incident a few miles away could significantly increase or decrease the traffic flows at the survey over four or five hours without the survey company being aware. The ability to increase the sample size by surveying over a number of days can reduce this type of effect and produce more detailed information as there maybe trends which are based over a weekly cycle.

If it is possible to automate the enumeration of the video automatically it should either be possible reduce the cost of enumeration on a 12 hour survey or increase the time of the survey without increasing the cost. Smart CCTV has been working on methods to undertake the enumeration of video surveys automatically and we have recently been running our algorithms over already enumerated video sequences to determine the level of consistency between the results of the manual enumeration and the automatic enumeration.

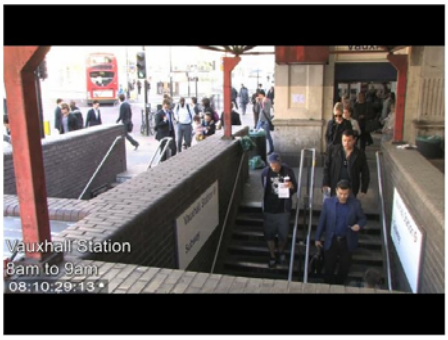
It should be noted that both methods will have an accuracy level, neither will produce an exact result.

Footfall Surveys

Counting people in a crowded environment is a highly complex problem for both human enumerators and for a computer algorithm. Occlusion (where one person moves in front of another hiding them from both the human enumerator and the computer algorithm) is common. People often take random paths through the

☎ 02392 248250
🖷 02392 261014
www.smartcctvltd.com
info@smartcctvltd.com

environment changing their direction through significant angles to avoid other people and fixed objects.



An example image taken from a survey of a railway station at rush hour is shown above.

In this case the survey requirement was to measure the number of people using the stairwells, during the rush hour this was almost 5000 people between 8.00am and 9.00 pm. When we run the computer algorithm (assuming we have good quality video and a sensibly place camera) against the manual enumeration we are getting between 90% and 95% agreement between the manual count and the computer generated account.

Traffic Surveys

In some ways traffic surveys are easier than pedestrian surveys in that vehicles tend to move in much more predictable paths. However it is often a requirement to classify the vehicles into different classes (which is not usually a requirement for pedestrians) as well as providing a count. These classes can be quite subtle; for example looking for HGVs with different amounts of wheels or axels. At this stage it isn't viable to classify except on a gross size basis giving a maximum of four classes: bikes (both pedal and motor), cars and light vans, light goods vehicles and large vehicles (HGVs and busses).

☎ 02392 248250
🖷 02392 261014
www.smartcctvltd.com
info@smartcctvltd.com

An example of a traffic survey image on a high speed road is shown above with identified vehicles being marked up.

We have again run a number of already enumerated surveys by outside consultancies through our image processing solution and we have had consistency between the results of between 95% and 98%.

Conclusion

It is noticeable that the computer enumerated result is always lower than the manual enumerated result. We believe that human enumerators are today better than the computer algorithms at dealing with occlusion. The computer will count two people / vehicles which are partially occluding as one object more often than a person. On the reverse we believe that it is a facet of human nature that when you count an object in frame 1 and notice it again a few frames of video later you sometimes are not sure if you have seen that individual / vehicle before and you are more prone to counting it again rather than ignoring it the second or subsequent times you notice it.

We therefore believe that in general human enumerators tend to overstate the count while the computer algorithm tends to understate the result. The real result is probably somewhere between the two counts.

We also believe that with a consistency of 90% or more against manual enumeration, the advantages of lower cost / more data throughput for the same cost, more than outweighs the discrepancies between the two methodologies.