Evaluation of the Accuracy of a Computer-Vision Based Crowd Monitoring System

Problem presented by

Joel Woon (CrowdVision)

ESGI138 was jointly hosted by

The University of Bath

The University of Bristol

with additional financial support from

Bath Institute for Mathematical Innovation (IMI)

The Engineering and Physical Sciences Research Council (EPSRC)

GW4 Alliance

The Infrastructure Industry Innovation Platform (I3P)

Industrially Focused Mathematical Modelling CDT (InFoMM))

Innovate UK's Knowledge Transfer Network (KTN)

Report Authors and Contributors

Facilitator: Jonathan A. Ward
Contributors: Giancarlo Antonucci, Edmund Barter, Frank Brooks-Tyreman,
Colm Connaughton, Michael Coughlan, Ronja Kuhne, Marcus Kaiser,
Victor Wang, Jonathan A. Ward

Contents

1 Executive Summary

Computer vision systems can be used to measure pedestrian flow rates, occupancy levels and queue times. It is difficult to assess the accuracy of such methods because the ground truth can be difficult to establish. Human counting is equally prone to error, even when using video recordings with no time constraints and the support of sophisticated software.

In this report, we consider how errors may arise directly from the images recorded by the cameras, due to both occlusion of people and image distortion due to a fisheye lens. We also develop a statistical model of human counting errors and attempt to estimate human accuracy from data. Finally, we attempt to relate human and computer accuracy on the basis of simplifying statistical approximations.

2 CrowdVision Problem Description

CrowdVision sells a software solution to monitor crowd levels in enclosed spaces in real-time, their main installations being at large international airports. Their monitoring method relies on using computer vision from an array of fixed overhead cameras to estimate occupancy levels, flow rates and queuing times within certain prescribed areas. One of the requirements of many of their clients is that they can demonstrate that their system is at least 95% accurate. However, it's not clear how to measure the accuracy of their system since the ground truth count is not known. CrowdVision have found that manual counts can vary from person to person, even when using video software that can be paused and played multiple times. Thus both machine-generated counts and human counts are imperfect. Moreover, the difference between two humans is often greater than the tolerance level required of the computer vision solution. It is therefore not possible to measure the system's accuracy directly.

The desired outcome of the study group was to help CrowdVision improve the method by which they validate their solutions. In particular, we were asked to develop a statistically robust protocol for verification of CrowdVision's system using the minimum number of human counts. Crowd vision also posed the following mathematical questions:

- Can we understand the true sources of variability of human approaches to counting occupancy levels? Are there specific events or types of crowd movements that lead to specific kinds of increased variability?
- What are the main sources of error in the CrowdVision system? Is the nature of the overlap region between cameras and how these are dealt with the main source of error? How does the error in the crowd size vary with the number of cameras used? Is a boot- strapping approach to error estimation possible?
- Can we look at specific errors in rate and queuing times? What is the main source of error when the camera is not directly overhead to the end of the queue?

Our work during the study group consisted of two distinct parts. In the first part, the geometric analysis described in Section 5, we considered how errors might arise due to the occlusion of people within the line of sight of a camera, and due to distortion effects resulting from the camera's fish eye lens. In the second part, the data analysis described in Sections 6 and 7, we analysed anonymised count data provided by CrowdVision and developed statistical models to relate the difference between both human and machine counts to estimate the repsective errors.

3 CrowdVision Software

There are two main techniques used to analyse video recordings of pedestrian movement: microscopic and macroscopic [1]. In the microscopic approach, a semi-automatic system is used, whereby a computer tracks individuals but a human manually selects people that the computer should track. This method results in trajectories of individual people, but it is typically only used over time-scales of minutes and when the density of pedestrians is low. It takes longer than real-time to perform, however the accuracy (i.e. the fraction of true trajectories counted) can be very high. In contrast, the macroscopic approach makes sole use of computer vision algorithms to count aggregate quantities such as pedestrian density, flow and velocity over timescales of hours or even days, and can be used in densely populated areas. The main benefit of the macroscopic approach is that it can be used in real-time, allowing operators to implement measures that reduce congestion or avoid incidents.

The video tracking system used by CrowdVision is based on the work presented in [1], which is essentially a two-step microscopic process: i) identification of an object, and ii) tracking of that object. The software is set up to work on video recordings, which consist of sequences of images, taken by a camera mounted overhead. Thus the software aims to identify people's heads, and it does this by analysing the pixels in each of the images in the video and searching for round objects using various filters, e.g. to separate the foreground from the background. This identifies candidate heads, but then an artificial neural network is used to improve this identification and distinguish between pedestrians, pedestrian-like objects, objects attached to pedestrians and people standing close to each other. This method is able to automatically identify new individuals when they appear on screen. In order to *track* a person, a simple probabilistic model that estimates the most likely position of a person in the next frame is used. Trajectories are mapped onto the floor using simple geometry under the assumption that everyone is the same height. The trajectories are then aggregated to produce the macroscopic measures of interest.

The development of the CrowdVision software was motivated by the identification of a number of failings in other systems [1], where there was little functionality to measure pedestrian motion; identification was small scale; there was little interactivity with the software to investigate new measurements; and a lack of automation reduced user-friendliness. Ultimately, the CrowdVision software has been implemented by airports and other public transport links where it is possible to quickly adapt and accommodate rapidly evolving situations caused by density increases in the pedestrian population. Whilst CrowdVision is built on macroscopic video-tracking software, typically an airport's independent means of verification of the software is via counts recorded by people in real-time physically present at the camera recording locations. These differing approaches are likely to give rise to

problems as highlighted in this report.

4 Literature Review

This section provides a short discussion about publications related to the accuracy of manual people counting. The literature suggests that hardly any extensive research on this question has been done, in particular there appears to be very little research specifically comparing manual counts to confirmed reference values taken in controlled environments. Previous research does indicate various sources of variablity in human approaches to counting.

- In the context of manual traffic counts [2]: Counts in this study were taken from video recordings and errors were analysed based on an aggregation level of five minutes. Human counters were asked to count short and long vehicles separately. It was observed that the total differences are small, and percentage errors were within $\pm 1\%$. However, there were large discrepancies regarding the classification of short and large vehicles.
- In the context of manual cell counts in hemocytometers [3]: Individuals were asked to count cells on a 9×9mm grid. The data demonstrated increased variation between different people's counts with both smaller areas and higher concentration. Errors were attributed to cells lying on the boundary of the observation area or cell clustering. It was remarked that results might be improved by extensive training of the human counters. Multiple counters counting the same chamber of the hemocytometers resulted in a coefficient of variation of up to 15.6%.
- The study "Pedestrian Counting Methods at Intersections: A Comparative Study" [4] compares the following three manual counting methods: manual counts using sheets, manual counts using clickers, and manual counts using video cameras. Their findings show that field counts with either sheets or clickers systematically underestimated pedestrian volumes with error rates from $8 - 25\%$. The error rate was observed to be greater at the beginning and end of the observation period, which was suggested to be related to the observer's lack of familiarity with the tasks or fatigue. Results further suggested that the pedestrian volume did not influence the error rate. However, later research [5] suggests that the high error rates in this study might be due to unscheduled breaks of the observer and the fact that the observer was tasked with a complex data collection including not only the number of pedestrians but also their gender and noticeable characteristics.
- On the other hand, the study "Effectiveness of a Commercially Available Automated

Pedestrian Counting Device in Urban Environments: Comparison with Manual Counts" [5] finds a high degree of inter-reliability between counts collected by field observers and through video recordings, with errors calculated for field counts being low, varying from -0.9% to 1.4%. It was suggested that the error rate is not directly related to pedestrian volumes but to the tendency of pedestrians to walk closely together.

4.1 Factors influencing the accuracy

Out of the above studies, the following conclusion about sources of variability of human approaches to counting might be drawn:

- Manual counts taken in real-time by people physically present at the camera locations using clickers might vary greatly (up to 25% according to [4]) from counts obtained using video footage of the same situation.
- A lack of clear rules on how to deal with the boundary of the observed area might lead to a higher error rate [3]. Hence, it is advisable to lay out a set of instructions for the manual counters on how to deal with borderline cases.
- Manual counters seem to deliver less accurate results at the beginning and end of their observation period [5] and further deliver more accurate results the more experienced they are [3].
- Clustering of the objects leads to a higher error rate of the manual counts [3].
- Regular breaks during the observation period and high motivation of the manual counters might increase the accuracy of manual counts [5].

Further factors influencing the accuracy of manual counts might be:

- Different methods of counting using video footage, for example placing dots on heads and re-watching the video several times.
- The quality and resolution of video recording, for example the distortion coming from fisheye lenses.

To further investigate the true sources of variability of human approaches to counting it might be advisable to conduct experiments in controlled environments where the actual number of people in the observed area is known. A further possibility might be to compare videos in which the manual counters marked their counted objects by dots. The latter might be especially interesting when addressing questions like errors coming from overlap regions of the cameras respectively distorted images due to the fisheye lenses.

Geometric Analysis

5 Image Geometry

On examining example footage of the camera feed and behaviour of the head-tracking algorithm from the company's website, it appeared to the group that the main factors that contribute to errors in human counts were due to the actual images the human counters have to work from. In particular, the group identified two main sources of this type of error: i) occlusion/overlapping of individuals within the camera's line of sight, and ii) distortion caused by the fisheye lens toward the edges of the image. We consider occlusion in Section 5.1 and distortion in Section 5.2. In Section 5.3 we consider the layout of the array of cameras.

5.1 Proximity Occlusion

In modelling the image errors we first assumed there was no focal distortion and focused primarily on the occlusion of individuals caused by camera height and distance between individuals. We consider a two-dimensional vertical plane and, in the first instance, we represent people by vertical lines, as illustrated in Figure 1. Using simple geometry, we will derive how much someone is 'shadowed' by someone else standing in front of them. This is a similar approach to the work presented in [1]. We assume that a camera is located at a point on the ceiling at a height h_c and each person has the same height $\bar{h}$. We suppose that in order to identify an individual, the image recorded by the camera must show at least $x \geq s_{\min}$ meters of a person in order to be included in the count. There are two main factors affecting the amount of x visible to the camera: i) the horizontal distance r from the camera to the person furthest away, and ii) the distance d between the two people. Let h denote the amount of the person in shadow that is visible from the camera's perspective. Using Figure 1, we derive the following equalities:

$$\tan(\alpha) = \frac{l}{x}$$

and

$$\tan(\alpha) = \frac{r}{h_c - \bar{h}},$$

so

$$r = \frac{l}{x} \cdot (h_c - \bar{h}).$$

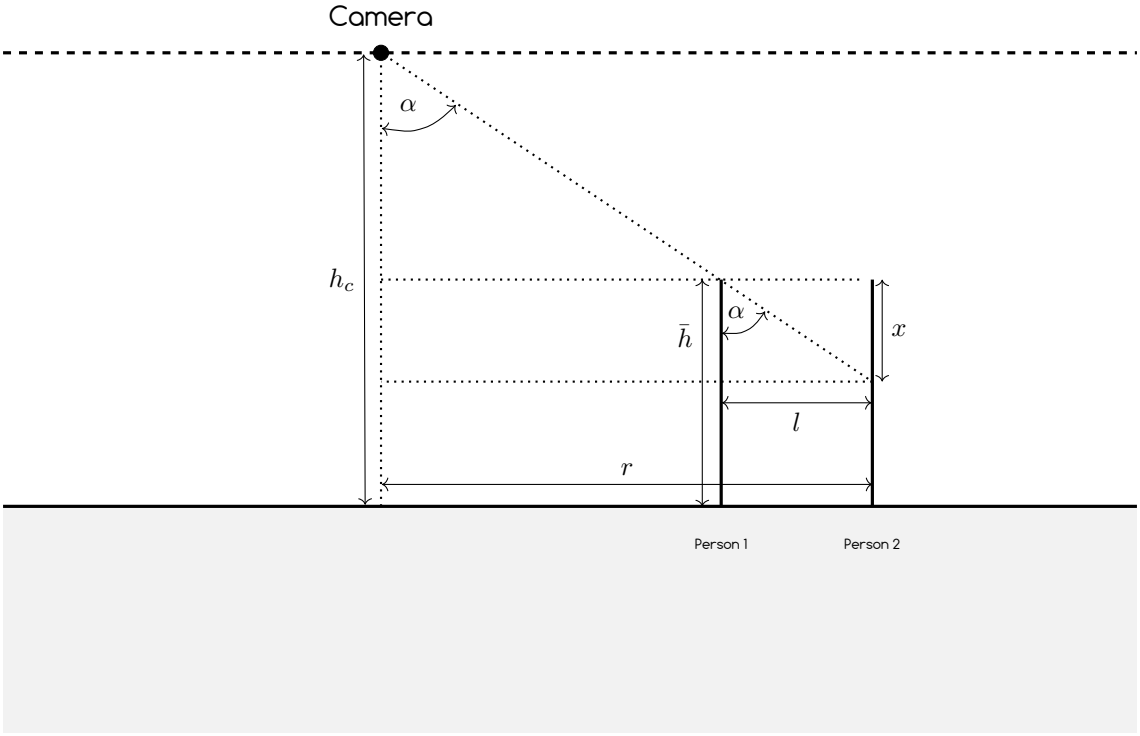

Figure 1: A diagram showing the simplified model used to calculate overlap errors.

Since x has to be greater than or equal to $s_{\min}$, we find

$$r \leq l \, \frac{h_c - \bar{h}}{s_{\min}}. \tag{1}$$

Since h_c, $\bar{h}$ and $s_{\min}$ are fixed parameters, (1) tells us that the radius within which people can be distinguished increases linearly with the distance between them. Using a suitable lower bound for the distance l between any two individuals and estimates for the average height of the observed humans, we can therefore determine the maximal radius of the circular area the camera can cover without introducing occlusion errors.

We now consider the more realistic case where people's heads have finite size and derive the fractional angular overlap. The scenario is illustrated in figure 2, where a camera is placed in a corridor a height h_c from the floor. We model two individuals' heads as discs of equal diameter d, a height h_d from the floor, such that the vertical distance from the discs' centres to the camera is $h = h_c - h_d$. We define the spacing between the discs as l. This spacing represents a density of people $\rho = 1/l$. With reference to the camera, this spacing subtends an angle $\Delta\theta$. We fix one of the discs at an angle θ from the optical axis of the camera. If the discs have angular sizes of δ_1 and δ_2 with respect to the camera, then the fractional overlap ϕ of the discs is given by

$$\phi = \frac{\delta_1 + \delta_2 - \Delta\theta}{\delta_2} \tag{2}$$

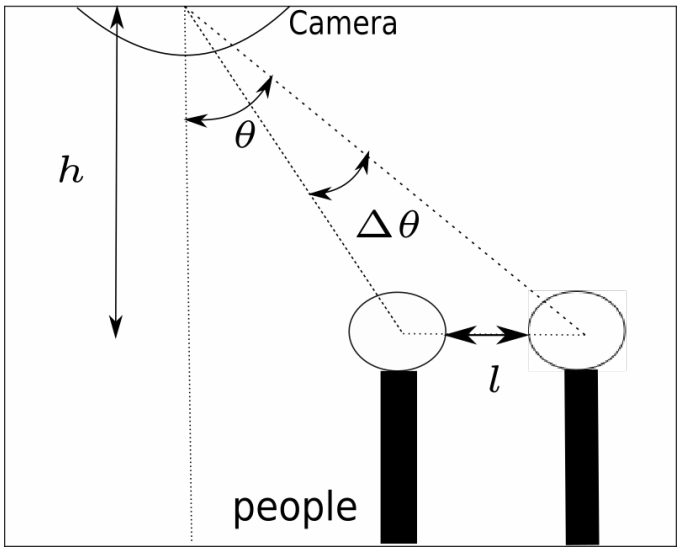

Figure 2: A schematic diagram of two individuals and their line of sight to the camera. Their heads are a vertical distance h from the camera and they are located at a horizontal distance l from one another. They are positioned at an angle of θ from the optical axis of the camera and are separated by an angle $\Delta\theta$ from one another with reference to the camera. As θ increases and $\Delta\theta$ decreases their lines of sight to the camera will begin to coincide.

as can be seen in figure 3.

We assume that $d \ll D$, where D is the distance of the discs from the camera, and make use of a small angle approximation to simplify expressions for angular size, i.e.

$$\delta = 2\,\sin^{-1}\left(\frac{d}{2D}\right) \sim \frac{d}{D}. \tag{3}$$

Making use of the approximation 3 and some simple trigonometry, we can express 2 as a function of the parameters d and h and the variables θ and ρ in the form

$$\phi(\theta,\rho) = 1 + \frac{\tan\theta\sin\theta}{(\tan\theta - \frac{1}{\rho h})\sin\{\tan^{-1}(\tan\theta - \frac{1}{\rho h})\}} - \frac{\tan\theta\sin\theta}{d\rho}. \tag{4}$$

We suppose that the probability of a counting error due to overlap is a fraction of $\phi(\theta,\rho)$ varying from 0 to 1. Lacking any further information for the cameras, the group chose a smoothstep function to demonstrate a calculation of the error. The smooth step function is based on Hermite polynomials and is often used as an interpolating function in computer graphics. With this choice,

$$P(\text{error}) = \begin{cases} 0 & \phi \le 0, \\ 6\phi^5 - 15\phi^4 + 10\phi^3 & 0 < \phi < 1, \\ 1 & 1 \le \phi. \end{cases} \tag{5}$$

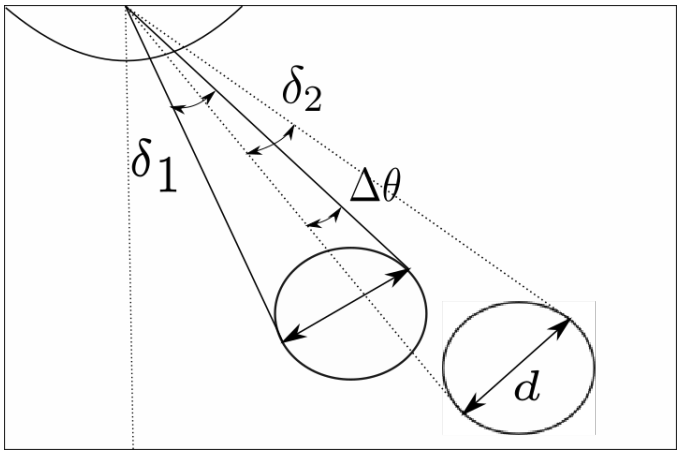

Figure 3: A more idealised schematic of two individuals as they begin to overlap as they appear in the image of the camera. Their heads (the discs) have angular sizes of δ_1 and δ_2 as they appear in the image and are separated by $\Delta\theta$. Their fractional overlap can be calculated as a function of θ, their angle from the optical axis of the camera, or how far away they are from one another, and $\Delta\theta$, which represents how far away they are from one another.

Figure 4 plots the error probability for overlap of two disc shaped heads with diameters of 10 cm each at a height of 2 metres from the ground.

5.2 Fisheye Distortion

In this section we consider errors introduced by distortion of the images due to the fisheye lens used in the cameras. In any camera image, the projection of our three dimensional world to two dimensions means that objects become more distorted the further away they are from the camera's optical centre. We have also seen that without a fisheye lens, the vertical amount of a person visible when partially shadowed by another decreases the further away they are. Both of these effects are exacerbated by fisheye lenses.

We assume that there is a scale length Δl for resolution in the image, below which the algorithm cannot distinguish an individual. The mapping between between distances in an image and the angular size of an object in reality in relation to a camera depend on the lens used.

For a typical fish-eye lens the mapping can be fitted as

$$r = f_1 \sin(f_2 \theta) \tag{6}$$

where r is the distance from the object to the image centre and θ is the angle from the object to the optical axis. The parameters f_1 and f_2 depend on the curvature of the lens.

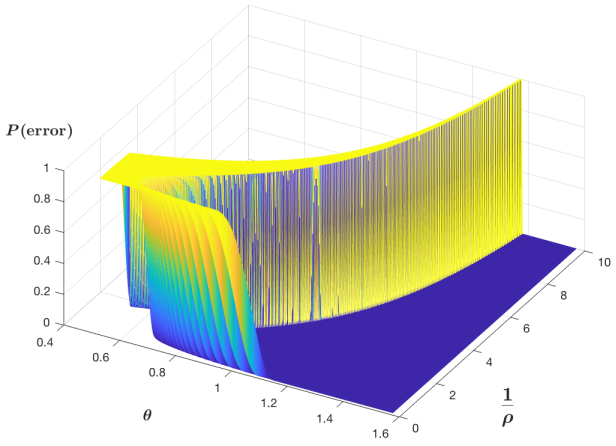

Figure 4: Numerical results for equation 5 using the expression 4 for the overlapping fraction of two heads, depending on the density or closeness of individuals (ρ or l respectively) and their angular distance from the camera's optical axis (θ).

Given that the exact mapping for the lenses used by the company was unknown to the group, we write

$$r = g(\theta). \tag{7}$$

In order to determine how the resolution length-scale maps to angular size, we employ a Taylor expansion about small deviations r_1 and θ_1 from r and θ respectively,

$$r - r_1 = (\theta - \theta_1)\frac{\mathrm{d}r}{\mathrm{d}\theta} + \mathcal{O}([\theta - \theta_1]^2) \tag{8}$$

and write

$$\Delta r = \Delta\theta\frac{\mathrm{d}r}{\mathrm{d}\theta} + \mathcal{O}(\Delta\theta^2), \tag{9}$$

thus

$$\Delta l = \Delta\theta g'(\theta)|_\theta. \tag{10}$$

For a camera a height h_c above the ground,

$$R = h_c \tan\theta. \tag{11}$$

The radius R_2 can then be found by solving 10 and 11 given a measure of the smallest angular diameter $\Delta\theta$ that the algorithm would need to resolve (i.e. the smallest human head the algorithm needs to identify).

5.3 Camera Array Geometry

We now consider how multiple cameras might be arranged to avoid introducing errors due to occlusion or distortion. We assume there is a radius R_1 from the camera inside which

occlusion is not an issue and assume there is a radius R_2 from the camera inside which the camera distortion is negligible. We assume $R_1 \ll R_2$. We examine figure 5, which sketches a

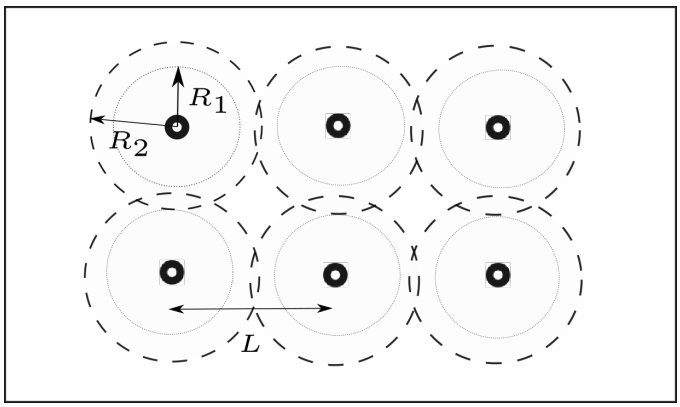

Figure 5: Plan view of an array of cameras (denoted by rings) in two dimensions. Possible radii, within which each source of error is negligible, are sketched. Spacing the cameras so that each point on the floor is within one of each radii for at least one camera presents a tiling problem.

two dimensional array of cameras that may represent the placement of cameras in a large open room, such as a departure lounge in an airport. In the following, we assume that the cameras are positioned in such a rectangular array. Consequently, in order to cover every point such that distortion effects are irrelevant, cameras must be spaced such that the spacing between them satisfies $L \leq \sqrt{2}R_2$. Occlusion or overlap errors can be reduced in a two dimensional array, as long as each individual lies either within R_1, such that occlusion is negligible, or as long as the individual obscured from view is visible to another camera. This second mechanism is possible at any angle in the interior of the array. To eliminate both sources of error, a lattice width of $L \leq 1/\sqrt{2}(R_1 + R_2)$ must be used for the array, i.e. so that the diagonal distance between cameras in the lattice is at most $R_1 + R_2$. Cameras should also be positioned within at least a distance $L_w \leq R_1/\sqrt{2}$ from a wall to eliminate both errors, as triangulation is not viable at the edges of the array, unless $R_2 > 3R_1$.

To summarize, for a lattice or array of cameras:

- cameras at the boundary of the array should be placed a distance $L_w \leq R_2/\sqrt{2}$ from walls and
- cameras in the interior of the array should be placed at distances of $L \leq 1/\sqrt{2}(R_1+R_2)$ from one another.

We examine the region in Figure 6, a line of cameras, which may represent the positioning of cameras in a corridor, which we assume is much longer than it is wide, and narrower

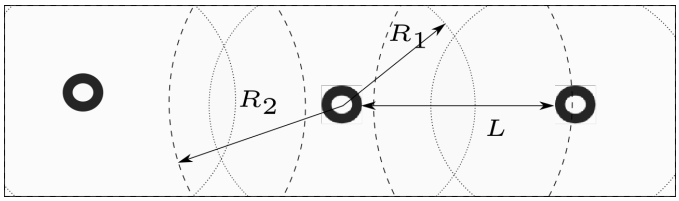

Figure 6: Plan view of an array of cameras (denoted by rings) in one dimension, in a corridor much longer than it is wide. In this case the spacing of cameras is dominated by the smaller error radius. Both error radii are sketched.

than R_1. We note that triangulation of individuals is not viable. In this case the spacing of cameras is limited to $L < R_1 + R_2$ in order to eliminate overlap errors, assuming that the algorithm can distinguish specific individuals in a stitch.

Statistical Analysis

6 Exploratory Data Analysis

We now give an overview of the data provided by CrowdVision. In Theory, one would like to use human counting data as a baseline for the counting results from the CrowdVision software. The problem with this is that the human counts are error-prone, which can be seen from Figure 7 for 5 different human counts of flow and occupancy.

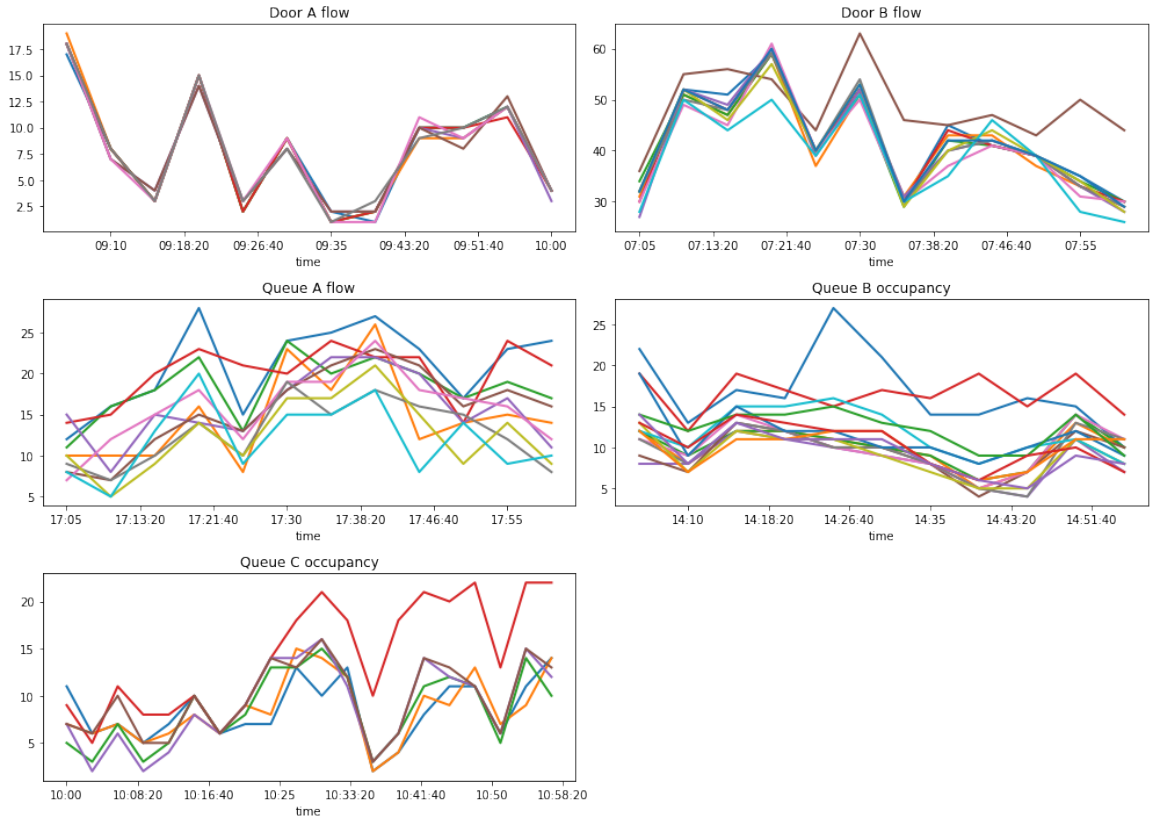

Figure 7: Plot of several data sets. Each of the plots represents one specific location and each lines corresponds to the counts of one human.

6.1 Data Sets

The data sets consist of Flow data (which describes the number of people crossing two given flow lines), Occupancy data (which consists of the number of people within a defined

area at a given time) and queuing times (the time a person spends in a given queue). The flow data and the occupancy data are recorded in pre-defined time intervals (usually every 5 minutes); for the flow data the number of people that cross the second flow line is recorded, whereas for occupancy data the number of people within the area of interest at that point in time is recorded. In contrast, the queuing times describe the time a single person is waiting in the queue, although these are only recorded for a sample of all people passing through the queue.

Further to the mentioned data (for flow, occupancy and queueing times) we also have access to raw data for flow counts (which consists of the precise times at which the software or the human counter observes a human crossing the second flow line). In total this amounts to four different types of data sets. Each of the given data sets contains information for a particular location (e.g. some open area in an airport or an area at the exit of a train). The recorded information stems either from human counters only, computer generated counts, or a comparison of both.

In the following we describe the structure of the given data sets and do a brief Explanatory Data Analysis (EDA).

6.1.1 Flow Data

This sub-section focuses on the flow count data; the other types of data will be considered in Section 6.1.2. We were given data that contains counts from two human counters as well as a count from the computer system. We will here focus on the data in the files Flow_Data named Airport3_FlowX_Manual.csv and Airport3_FlowX_System.csv (in what follows, we will drop the .csv extension from the file names). We first plot a histogram (Figure 8) of the accumulated counts and four individual trajectories (Figure 9).

Both plots show the natural fluctuations in the number of counts between the different people and the software. In particular, we can see in Airport3_Flow6 that the software differs substantially from the human counts. We can quantify this by looking at the correlation between the different counters, which is depicted in Table 1. In most cases the correlations are reasonably high, i.e. greater than 0.95. However, there is some evidence for systematic errors, e.g. in most cases the correlation between people is higher than between each person and the software.

To investigate this further, in Figure 10 we illustrate scatter plots comparing human-human and human-computer counts for the above four data sets. We observe different behaviour in the data sets in terms of bias (e.g. Person1 vs. Person2 in Airport3_Flow8 seems to have a slight bias, whereas the data points in Airport3_Flow6 seem to be perfectly aligned) and

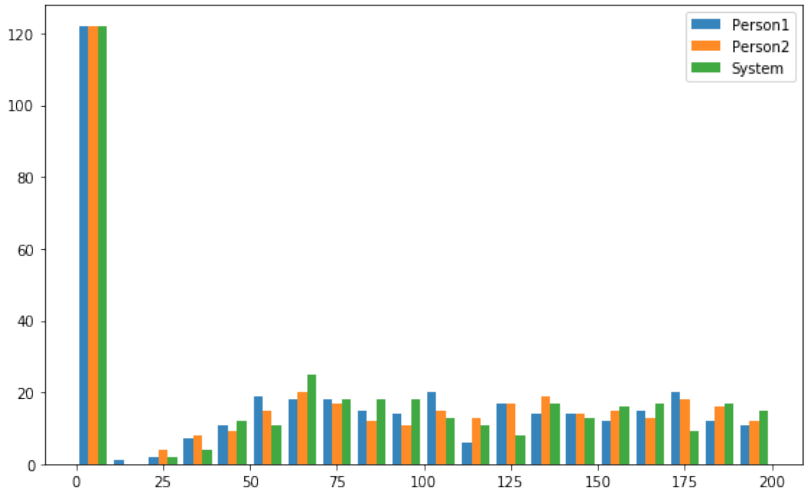

Figure 8: Histogram for the files Airport3_FlowX_Manual.csv and Airport3_FlowX_System.csv for $X = 1, 3, \ldots, 12$

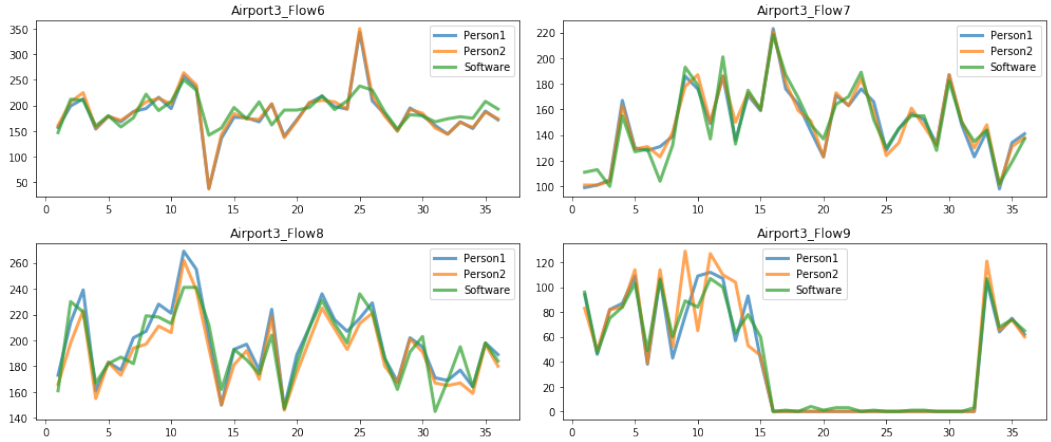

Figure 9: Plot of individual trajectories for for X=6,...,9.

different levels of heteroscedasticity. This suggests that one should analyse individual locations separately. An alternative way to compare human-human or human-computer counts is via a Bland-Altman plot, as illustrated in Figure 11. This is simply a plot of the difference between counts against the mean of the pair of counts. Figure 11 compares human-human counts from Airport4_Flow1_20180417_Manual. The solid black line indicates the mean of all the differenced data (0.12) and the dashed line corresponds to the 95 percentile confidence interval under a normal approximation (i.e. 1.95 standard deviations either side of the mean). We can see that there does seem to be a dependence of the count error on flow rates, particularly for flows over 80 people per five minutes.

Correlations	P1_vs_P2	P1_vs_S	P2_vs_S
Airport3_Flow4	0.935163	0.970543	0.987624
Airport3_Flow5	0.994892	0.932718	0.923975
Airport3_Flow6	0.994707	0.750548	0.764750
Airport3_Flow7	0.981363	0.947585	0.951722
Airport3_Flow8	0.987551	0.911267	0.892939
Airport3_Flow9	0.939150	0.985826	0.964066
Airport3_Flow10	0.939000	0.963832	0.912025
Airport3_Flow11	0.960971	0.963735	0.992650
Airport3_Flow12	0.983934	0.983538	0.966373
mean	0.968526	0.934399	0.928458

Table 1: Different correlations (Person1 versus Person2, Person1 versus Software and Person2 versus Software).

Additionally, we can consider differenced data. The idea behind this is to make the time series stationary, i.e. in stead of $(x_t)_{t\in\{1,\dots,N\}}$, we consider the differences $(x_t - x_{t-1})_{t\in\{2,\dots,N\}}$.

6.1.2 Other Types of Data

We now have a brief look at the other data types as well.

Occupancy Data

Since we chose to focus on pedestrian flows, we only had access to one occupancy data set (which can be found in Airport1_Occupancy.csv), see Figure 13. Again, we can consider the scatter plot for the individual variables (Figure 14). While no strong patterns are evident in this data, we expect to find similar features to those found in the flow data.

Queueing Times

We now turn to the data on queueing times. We obtained data for two scenarios. Each of the two data sets contains a count from a human counter and a computer generated count. The files are stored in Airport2_ExpQueueTimeX_CV.csv (generated by the CrowdVision system) and Airport2_ExpQueueTimeX_MV.csv (manual count) for X=1,2.

The plot below shows the recorded values for X=2. Note that the software and the human

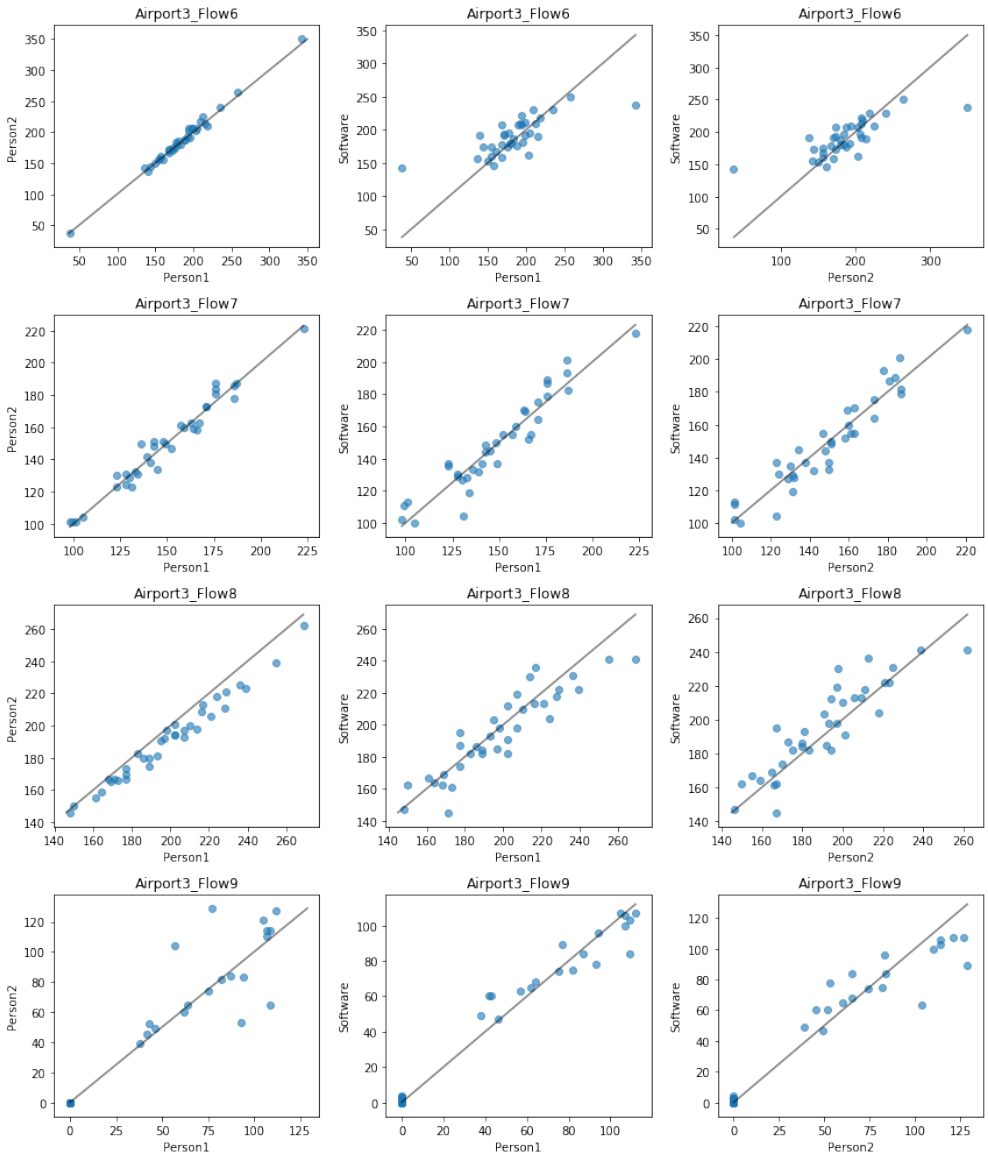

Figure 10: Scatter plot for the different counts. The closer the points are to the straight line on the diagonal, the more the two compared counts coincide.

counter use different time-intervals for the recorded counts. Therefore, the plots are not as easily comparable as in the other cases.

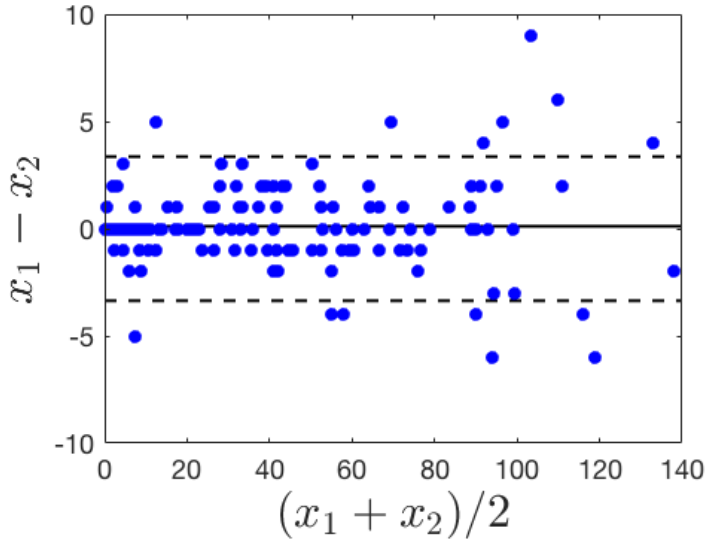

Figure 11: Bland-Altman plot comparing human counts from two different people, x_1 and x_2. The mean and difference between the two people's measurements are plotted on the horizontal and vertical axes respectively, each data point indicated a blue marker. The solid line is the mean of all the differenced counts which was 0.12. The dashed lines denote the 95 percentile under a normal assumption.

Raw Data

We finally consider the raw data for the queueing times. The given data sets provide recordings for two validators. Each data set contains the values 'Adjusted time', 'Epoch', 'Value' and 'Cumulative'. The latter two can be reconstructed from the time stamp contained in Adjusted time. We illustrate a histogram of the Adjusted in Figure 16.

7 Statistical Models

The empirical evidence presented in Section 6 suggests that manual counting is prone to large errors. Consequently, given a manual count what is the best estimate of the ground truth? And how likely is this to be correct? A broader consideration asks what the probability is of any particular value being the ground truth. If we can estimate human error rate, can we infer the computer vision error rate? In this section we address these questions by considering simple models of human counting with error.

As a first test of the data, we used the two sample Kolmogorov-Smirnov test to see if human-

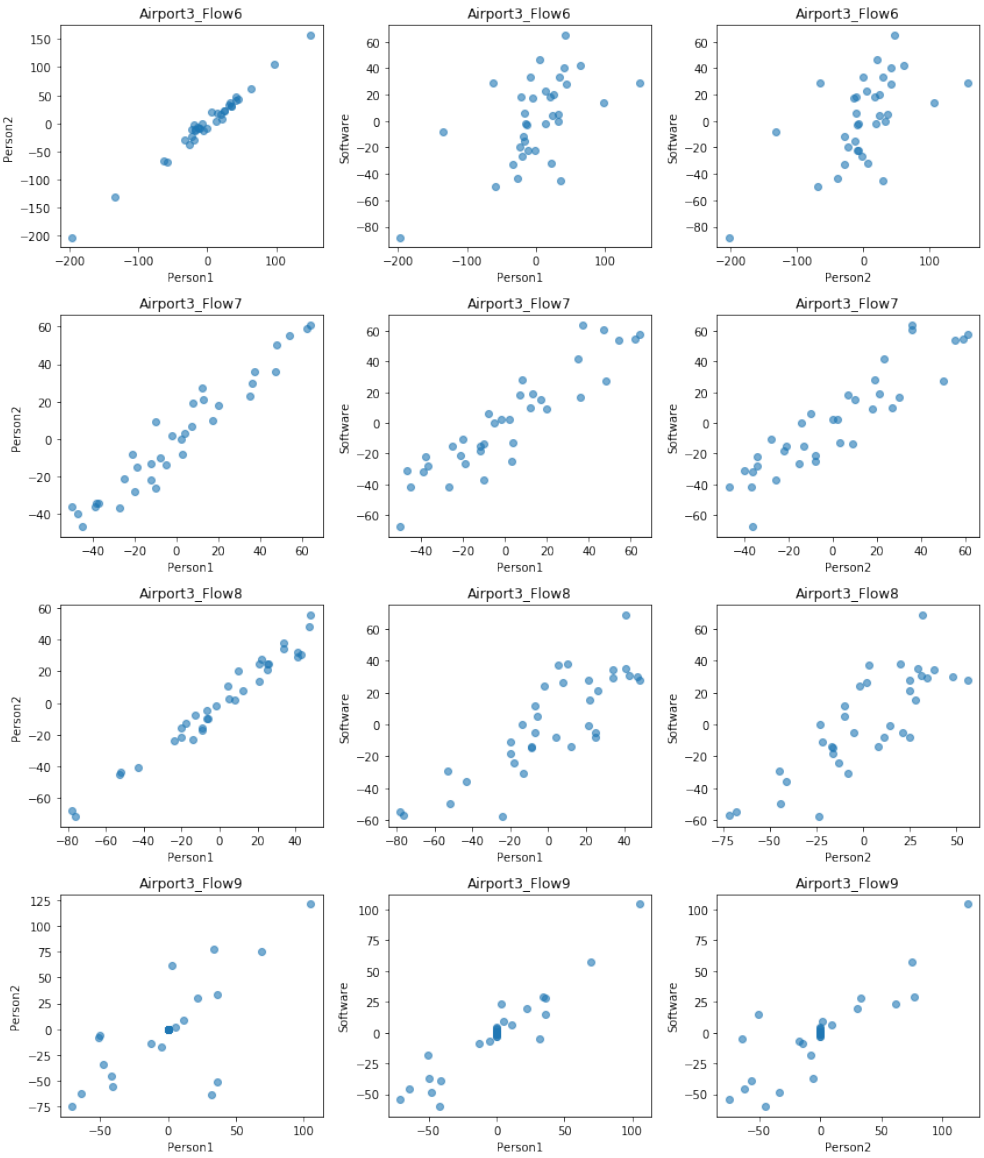

Figure 12: Scatter plot for the differenced data.

human and human-computer counts might come from different distributions. Even at the 1% significance level, the null hypothesis, that both sets of data are sampled from the same distribution, was not rejected. However, the variability of the ground truth could mask the differences between the distributions.

To model human counting error, we also need to have a model of the arrival process of people recorded on the camera. A basic approach is to assume that people arrive independently, and so the process is memoryless. This gives rise to an exponential waiting time distribution and Poisson flow counts. This was certainly *not* true of all of our datasets, for

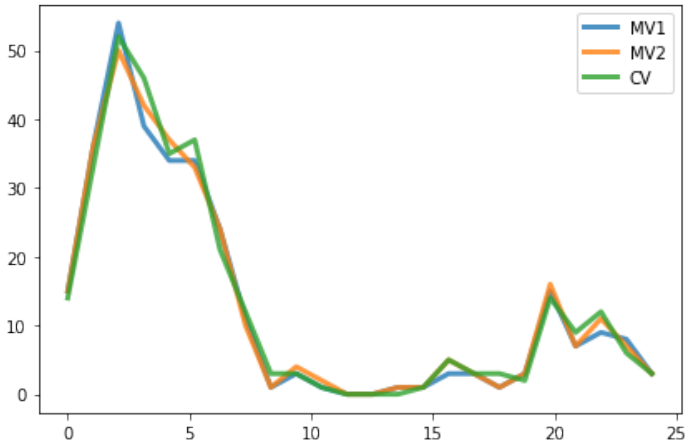

Figure 13: Plot of occupancy data from Airport1_Occupancy.csv

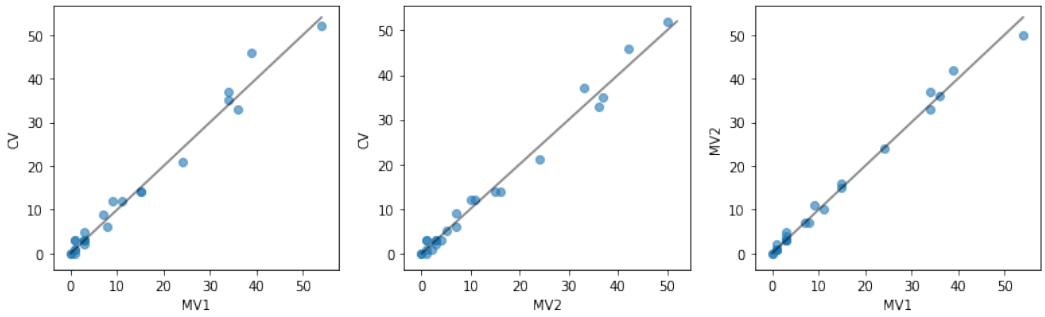

Figure 14: Scatter plot for occupancy data.

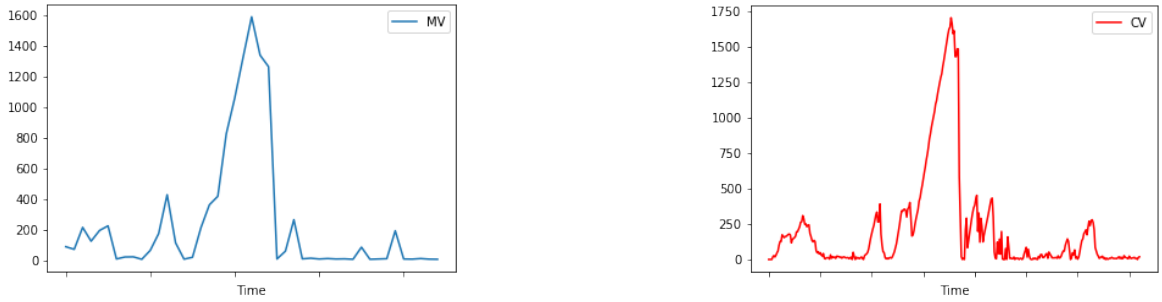

Figure 15: Plot of queueing times data from Airport2_ExpQueueTimeX_CV.csv and Airport2_ExpQueueTimeX_MV.csv

example we found that a χ^2 goodness of fit test rejected the hypothesis that the counts from Airport4_Flow1_20180417 are Poissonian. However, after removing one spuriously long period of zero counts, we found that the time intervals between arrivals computed from

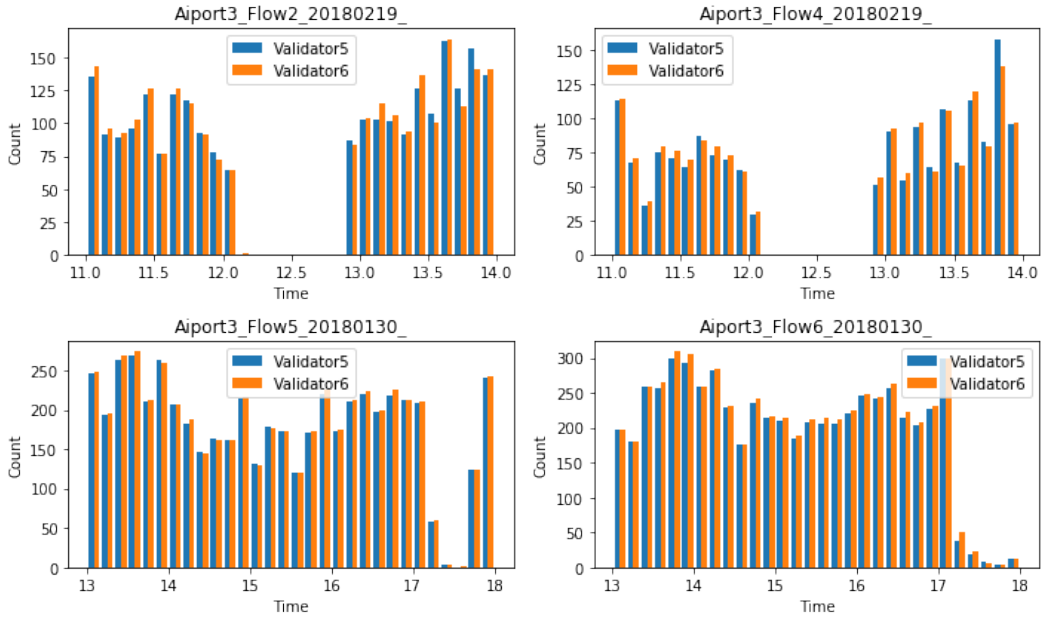

Figure 16: Histogram of raw data sets.

the raw human counting data for Airport3_Flow6_Manual were exponentially distributed, with mean arrival rate of around 1.6 people per second.

7.1 Human Error Models

Counting errors by a human validator take two forms: undercounts, where no addition is made to the count despite a person being present; and overcounts, where an addition is made to the count that does not correspond to a person. Both kinds of error can arise from a variety of processes, both in the visual act of counting, such as luggage being mistaken for a person, or in the data entry process, such as an accidental or unregistered click of the mouse.

For simplicity we will not consider the microscopic causes of errors, and only the two broad categories of overcounts and undercounts. As such, we can model a human validator counting the flow of people through a particular point as a sequence of many events, each one of three distinct types:

1. True counts, where a person passes through the area and is logged in the validation system;
2. Undercounts, where a person passes through the area but is not logged in the validation system;

3. Overcounts, where no person passes through the area but one is logged in the validation system.

The output from the manual validation of flow rates was presented in two ways:

1. a series of time stamps each representing a recorded event,
2. a series of flow rates, each giving the number of people who are thought to have passed through the box within some time interval, such as every five minutes.

These outputs are subject to the ground truth flow of people, but also the rate at which errors occur. By modeling the errors we seek to find the combination of ground truth and error rates that best explain the observation, i.e. the most likely ground truth value.

7.2 Single Person Counting with Over and Under Counting

We wish to build a statistical model of an observer imperfectly counting n objects in a set. Imperfect counting means that each object in the set has a probability of being missed (under-counting) or counted more than once (over-counting). For simplicity, let us assume that each object is counted at most twice. For each object in the set, we model the counting process as a random variable, x, with the Probability Mass Function (PMF):

$$\mathbb{P}(x = X) = \begin{cases} p & X = 0 \\ q & X = 1 \\ r & X = 2 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

We allow the probability of under-counting, p, to differ from the probability of over-counting, r. Clearly we must have $p + q + r = 1$. If we further assume that the observer counts each object in the set independently, the observer's estimate of the number of objects contained in the set is given by the trinomial random variable

$$y = \sum_{i=1}^{n} X_i \tag{13}$$

where each X_i is independent and identically distributed with PMF given by Eq. (12). Clearly y takes integer values m in the range $0 \leq m \leq 2n$. Our objective is to calculate the PMF of y as a function of p, q, r and n. To keep the notation compact, we will drop the explicit dependence on p, q and r throughout and use $\mathbb{P}(m, n)$ to denote the probability that y takes the value m given that there are n objects in the set.

We calculate $\mathbb{P}(m, n)$ using a generating function approach. The generating function of x is

$$G_X(z) = p + q\,z + r\,z^2.$$

Since the individual counts are independent, the generating function of y is

$$G_Y(z,n) = G_X(z)^n = (p + q\,z + r\,z^2)^n \tag{14}$$

and the corresponding probabilities are obtained by differentiation:

$$\mathbb{P}(m,n) = \begin{cases} \frac{1}{m!}\,\frac{d^m}{d\,z^m}\,G_Y(z,n)\big|_{z=0} & 0 \le m \le 2n \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

Expanding Eq. (14) gives

$$G_Y(z,n) = \sum_{\substack{j,k,l \ge 0 \\ j+k+l=n}} \binom{n}{j,k,l}\, p^j q^k r^l z^{k+2l} \tag{16}$$

where

$$\binom{n}{j,k,l} = \frac{n!}{j!\,k!\,l!}$$

are the trinomial coefficients. Only the terms containing z^m contribute to $\mathbb{P}(m,n)$. There are therefore 2 constraints on the indices, j, k and l appearing in the above sum:

$$\begin{aligned} j + k + l &= n \\ k + 2l &= m. \end{aligned}$$

If we take k to be the independent index, we can solve for j and l:

$$\begin{aligned} j &= \frac{2n - (m+k)}{2} \\ l &= \frac{m-k}{2}. \end{aligned}$$

To keep the indices j and l non-negative, the range of k is $0 \le k \le \min\{m, 2n-m\}$. To ensure that j and l are integers, both $m-k$ and $m+k$ are even. To ensure that $m-k$ is even we introduce a new independent integer index, i, and write $m - k = 2i$. This ensures that $m + k = 2(m-i)$ is also even. Writing j, k and l in terms of i we get

$$\begin{aligned} j &= n - m + i \\ k &= m - 2i \\ l &= i. \end{aligned}$$

The range $0 \le k \le \min\{m, 2n-m\}$ translates in terms of i to $\max\{0, m-n\} \le i \le \lfloor \frac{m}{2} \rfloor$ where $\lfloor x \rfloor$ is the smallest integer less than or equal to x. Putting this all together, the required coefficients in Eq. (16) give

$$\begin{aligned} \mathbb{P}(m,n) \quad = \quad & \sum_{i=\max\{0,m-n\}}^{\lfloor \frac{m}{2} \rfloor} \binom{n}{n-m+i, m-2i, i} \\ & \times\, p^{n-m+i} q^{m-2i} r^i. \end{aligned} \tag{17}$$

This formula is checked against the empirical distribution obtained by sampling from the distribution of y in Figure 17.

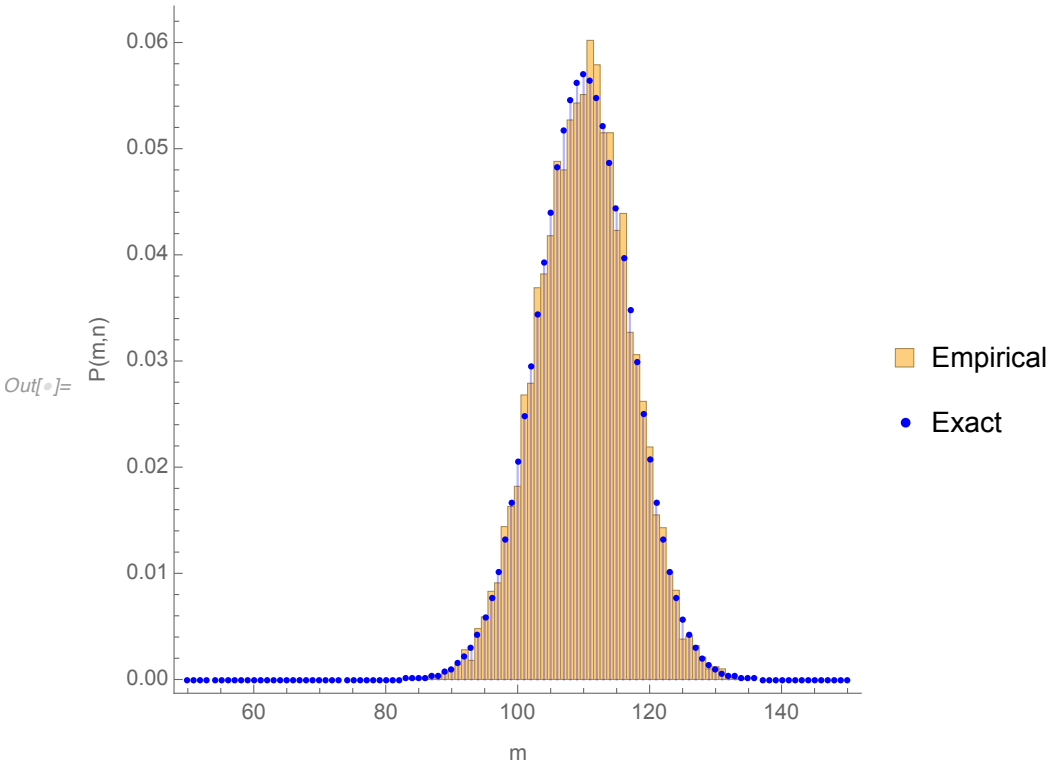

Figure 17: Empirical histogram of distribution of y for $n = 100$, $p = 0.2$ and $q = 0.3$ built from 10000 samples. The corresponding exact PMF, Eq. (17), is shown for comparison.

7.3 Single Person Counting with Only Under Counting

The analysis in the previous section assumed that the ground truth count n is known. To estimate the human error rate α and each true flow rate n, from a sequence of observed flow rates (counts) c, it is necessary to model both the actual flow and the counting error. As a first approximation we assume that, for a set of consecutive intervals, the true flow in each interval can be considered a random i.i.d variable drawn from a Poisson distribution with mean λ. A more realistic model is the Borel-Tanner distribution (see Appendix A), although we do not consider that in this report. As a further simplification, we consider a single validator who makes only under counting errors only, and does so at a constant rate α. Therefore, for each event the validator either: records it correctly, with probability $1 - \alpha$; or does not record anything, with probability α.

In a time interval the probability of the validator counting c given a flow of n is given by

$$P(c|n) = \binom{n}{c}(1-\alpha)^c\alpha^{n-c}, \tag{18}$$

and the probability of the validator counting c is therefore $P(c) = \sum_{N=c}^{\inf} P(c|n)$.

We can write the distribution of $P(c)$ using probability generating functions, i.e. power series representations of probability distributions, the properties of which allow for fast computation of combinatoric probabilities [6].

The distribution of true values n is generated by

$$G_n(x) = \exp^{\lambda(x-1)}$$

where x is a dummy variable. This is the generating function for a Poisson random variable. The probability distribution that a single person is counted, the result of a Bernoulli trial, can be generated by

$$G_B(x) = \alpha + (1-\alpha)x.$$

Then the probability distribution of people counted is generated by

$$G_c(x) = G_n(G_B(x)) \tag{19}$$

$$= \exp^{\lambda(1-\alpha)(x-1)}. \tag{20}$$

From equation 19 we can see that the distribution of observed counts is a Poisson distribution with mean $\lambda(1-\alpha)$. Therefore we can say that the distribution with the maximum likelihood of giving the observed counts is that with $(1-\alpha)\lambda = \overline{x}$ where, $\overline{x}$ is the mean of the observed count. However this means that all values of the mean of the true flow distribution, $\lambda > \overline{x}$, are possible, each with an associated value of human error, see Figure18. Thus

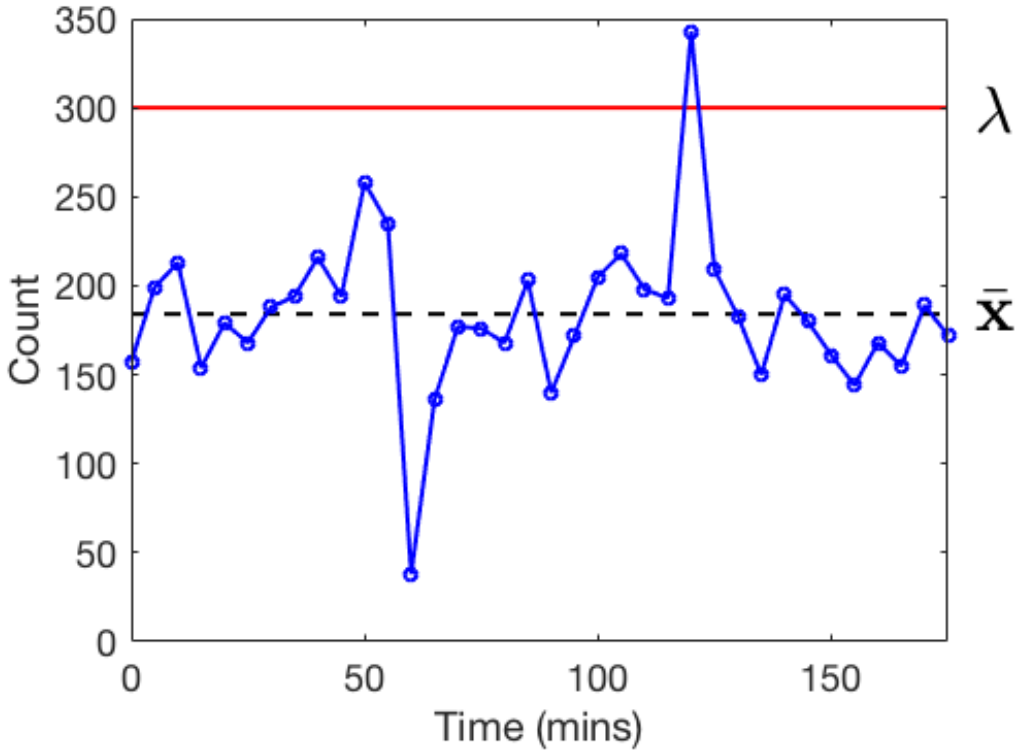

Figure 18: A single human validator does not restrict the possible true count. The mean value of the Poisson distribution describing the true count (red line), can take any value as long as the error is $\alpha = 1 - \overline{x}/\lambda$.

it is not possible to estimate the true flow rate with counts from a single person. Even if we extended our model to include over counting, this would still be true, since for any flow rate we can choose the ratio of over to under counting to fit the observed data.

7.4 Model of two validators counting errors

We have seen that extracting the true flow count n, from a single human count c is not possible without further information. This is due to not having enough information about the unknown number of individuals that are not counted. We now show that with two human validator estimation is possible. Having a second validator introduces the possibility that an individual is counted by one validator but no the other. By connecting this probability to the probability that both validators do not count an individual the true count can be estimated.

In this approach we consider raw data for each individual counting event recorded by two people, who we refer to as validator 1 and validator 2. For each validator we have a

series of times stamps, each corresponding to counting an individual crossing the flow line. To demonstrate the method we assume that only undercounts occur. We expect that the introduction of overcounting makes the analytical solution less tractable, however numeric computation should still yield results.

For each event in the true count, one of four possible counting events can occur. The person can be counted by both validators, just one of either validator 1 or 2, or not counted by both validators. Assuming errors are uncorrelated but occur at the same rate for both the validators, the probabilities of each of these occurrences are shown in Figure 19. For a

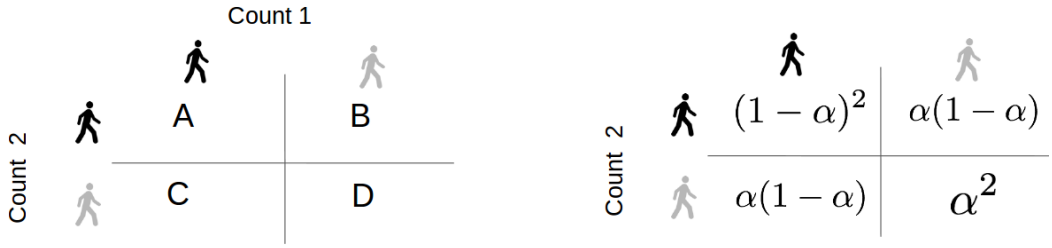

Figure 19: Two human validators with uncorrelated errors. With two human validators there are four possible outcomes. Black silhouettes indicate people that are recorded, grey silhouettes indicate people who are not recorded. Left) Notation used for the number of outcomes of each type, right) Probabilities of each outcome.

sequence of n such events

$$P(A,B,C,D|\alpha,n) = \frac{n!}{A!B!C!D!}(1-\alpha)^{2A}(\alpha(1-\alpha))^{B+C}\alpha^{2D} \tag{21}$$

where A is the number of people counted by both validators, B is the number counted by only validator 1, C is the number counted by only validator 2, and D is the number counted by neither.

Differentiating equation 21, and using $n = A+B+C+D$ we find that the value of α that gives the maximum likelihood given A,B,C and n is

$$\alpha^* = 1 - \frac{2A+B+C}{2n}. \tag{22}$$

Counting events in two validators series can be matched using the Hungarian algorithm [7], also known as the Munkres algorithm. Hence, we can extract from the data:

- A, the number of people counted by both validators;
- B, the number counted by only validator 1;

- C, the number counted by only validator 2.

Using this data, and $n = A+B+C+D$ we find the value of α that maximizes $P(A, B, C, D|\alpha, n)$, and subsequently the maximum value of $P(A, B, C, D|\alpha, n)$ for each value of n. From this the combination of α and n can be found numerically.

The result for a single time series is shown in figure 20 as an example. In the period validator 1 counted 1100 while validator 2 counted 1101, however we find that the most likely true count is 1109. The human error rate is 8%, which is not inconsistent with the literature reviewed in Section 4. The ability of this method to estimate both the error rate and the actual number

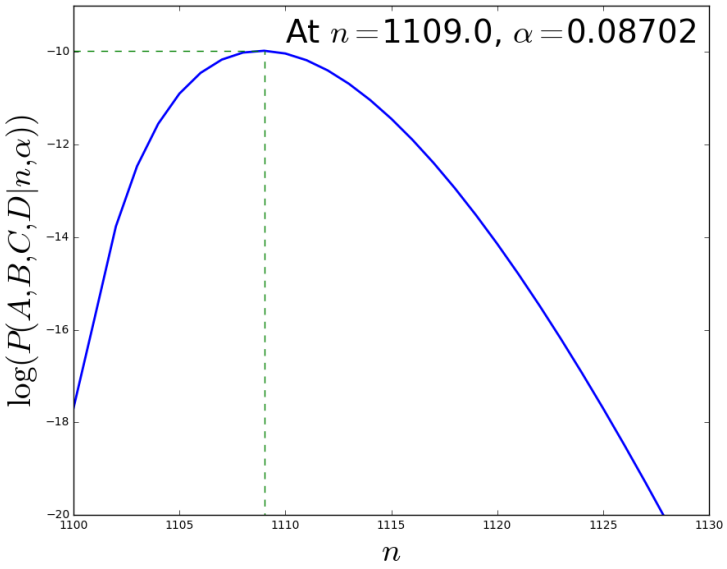

Figure 20: From two human validators we can estimate the true count and the error. The probability of observing A,B and C values in the data, is shown for the value of α giving the maximum likelihood for each value of n. The peak of the graph corresponds to the values of α and n that give the maximum overall likelihood of observing A, B and C.

of people is due to the large number of data points in the time series. As such we expect that a similar construction is possible even when increasing the complexity of the human counting model, for example by including over counting errors, different error rates for each validator or even correlating some errors between validators. A further advantage over the method for single validators is that there is no model needed for the ground truth.

7.5 Computer Error Models

CrowdVision are interested in finding out how the computer error relates to human counting error. For this reason, we estimate the optimal standard deviation of the computer counting given the datasets of human counting. Hence, we build a statistical model to describe the distributions of human and computer measurement errors. Then, we infer the parameters of the distributions from frequentist and Bayesian approaches. In particular, we compute the standard deviation of the computer measurement error.

In the following, we use the notation:

- x_t^c: the count recorded by the computer at time t.
- x_t^m: the manual count at time t.
- x_t: the ground truth value of the count at time t.
- ϵ_t^c: the counting error of the computer at time t.
- ϵ_t^m: the counting error of manual counting at time t.

The data that we focus on here includes one set of time series data of computer counts $\{x_t^c\}_{t=1}^T$ and two sets of time series data of manual counts $\{x_{t,i}^m\}_{t=1}^T$, $i = 1, 2$.

Using the notation introduced above, the following equation holds:

$$x_t = x_t^c + \epsilon_t^c = x_t^m + \epsilon_t^m \,. \tag{23}$$

We then proceed to our analysis starting from (23) and the following assumptions:

1. The computer counting error ϵ_t^c may be correlated to the manual counting error ϵ_t^m. We take this into account via
$$\text{corr}(\epsilon_t^c, \epsilon_t^m) = \rho_\epsilon \,. \tag{24}$$
2. We assume that both computer and human counts are not biased. Also, for simplicity we assume that both errors are normally distributed with means 0 and standard deviations
$$\epsilon_t^c \sim \mathcal{N}(0, \sigma_c^2) \,, \quad \epsilon_t^m \sim \mathcal{N}(0, \sigma_m^2) \,, \tag{25}$$
where σ_c and σ_m are the standard deviations of the computer and human errors, respectively. This is a rough assumption, as in reality the counts are discrete, but it is a sensible starting point.

7.6 Frequentist approach

First, we recast (23) into

$$\epsilon_t^c = x_t^m - x_t^c + \epsilon_t^m \,, \tag{26}$$

from which we derive that

$$\sigma^2(\epsilon_t^c) = \sigma^2(x_t^m - x_t^c) + \sigma^2(\epsilon_t^m) + 2\sigma(x_t^m - x_t^c, \epsilon_t^m) \,. \tag{27}$$

Since

$$\begin{aligned} \sigma(x_t^m - x_t^c, \epsilon_t^m) &= \sigma(\epsilon_t^c - \epsilon_t^m, \epsilon_t^m), \\ &= \sigma(\epsilon_t^c, \epsilon_t^m) - \sigma^2(\epsilon_t^m), \end{aligned}$$

using our assumption that $\rho_\epsilon = 0$, we can eliminate $\sigma(x_t^m - x_t^c, \epsilon_t^m)$ in (27), resulting in

$$\sigma_c = \sigma(\epsilon_t^c) = \sqrt{\sigma^2(x_t^m - x_t^c) - \sigma^2(\epsilon_t^m)} \,. \tag{28}$$

(Note that the quantity under the square root is always positive.) In order to make use of (28) we need to estimate all the parameters involved. We know that $\epsilon_{t,i}^m = x_{t,i}^m - x_t$, so, since we only deal with two manual counting datasets, we estimate ϵ_t^m as $\hat{\epsilon}_t^m = x_{t,1}^m - x_{t,2}^m$. Next, we estimate $\hat{\sigma}(\hat{\epsilon}_t^m)$ using the usual unbiased estimator, and then we compute $\hat{\sigma}^c$, which for the datasets described above yields $\hat{\sigma}^c \approx 1.78$. This simple procedure can be extended to datasets with more time series data.

7.7 Bayesian approach

In the Bayesian approach, we assume some distributional form for the parameters to be estimated. The estimation of the distribution starts from a predefined prior distribution, which will be continuously updated as more data are observed. As a result, by investigating how the estimates behave with more data, we can also gain insights into the number of data points that are required for a converged estimate.

First, from equation (23) we can derive

$$y_t := x_t^c - x_t^m = \epsilon_t^m - \epsilon_t^c \,, \tag{29}$$

By the assumptions that both errors are normally distributed, we have

$$y_t \sim \mathcal{N}(0, \sigma^2) \,, \tag{30}$$

where we denote $\sigma^2 = \sigma_c^2 + \sigma_m^2 - 2\rho_\epsilon \sigma_c \sigma_m$. The questions to be answered are twofold:

- What is the most likely estimate of σ, given the observed data?

- How many data points are needed for a converged estimate of σ?

If we denote the data set by D, and our prior knowledege of this problem by X, then by Bayes' rule we have

$$\mathbb{P}(\sigma|D,X) \propto \mathbb{P}(D|\sigma,X) \times \mathbb{P}(\sigma|X)\,. \tag{31}$$

We therefore want to maximise this probability given D and X. This requires a prior distribution of σ and a model for data generation given σ. To achieve this, we make two extra assumptions:

1. σ follows a log-normal distribution:

$$\ln\sigma \sim \mathcal{N}(\mu_\sigma,\eta)\,. \tag{32}$$

2. y_t is time-independent.

The first assumption enables us to write

$$\mathbb{P}(\sigma|X) = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}\eta}\exp\left(-\frac{(\ln\sigma-\mu_\sigma)^2}{2\eta^2}\right), \tag{33}$$

and the time-independence assumption leads to

$$\begin{aligned}\mathbb{P}(D|\sigma,X) &= \mathbb{P}(y_1,\ldots,y_T|\sigma,X)\\ &= \prod_{t=1}^{T}\mathbb{P}(y_t|\sigma,X)\\ &= \prod_{t=1}^{T}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{y_t^2}{2\sigma^2}\right).\end{aligned} \tag{34}$$

Then, we plug equations (33) and (34) into (31) and get the log-likelihood function of σ

$$\mathbb{P}(\sigma|D,X) \propto T\ln\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \sum_{t=1}^{T}\frac{y_t^2}{2\sigma^2} + \ln\left(\frac{1}{\sigma\sqrt{2\pi}\eta}\right) - \frac{(\ln\sigma-\mu_\sigma)^2}{2\eta^2}\,. \tag{35}$$

To maximise the log-likelihood, we make the first derivative with respect to σ equal to zero 0 and solve for σ, i.e.

$$\frac{\partial\ln\mathbb{P}(\sigma|D,X)}{\partial\sigma} = 0\,, \tag{36}$$

which leads to

$$(T+1)\sigma^3 - \sum_{t=1}^{T}y_t^2 + \frac{\sigma}{\eta^2}(\ln\sigma-\mu_\sigma) = 0\,. \tag{37}$$

In Figure 21 we show the solved $\hat{\sigma}$ from equation (37) by root-finding procedure, given different number of data points. We observe that the estimate converges to 1.53. This means that the standard deviation of the computer counting error is bounded by:

$$(1.53-\sigma_m)^+ \le \sigma \le 1.53+\sigma_m\,. \tag{38}$$

Moreover, it appears to require at least 20 data points to infer a converged estimate.

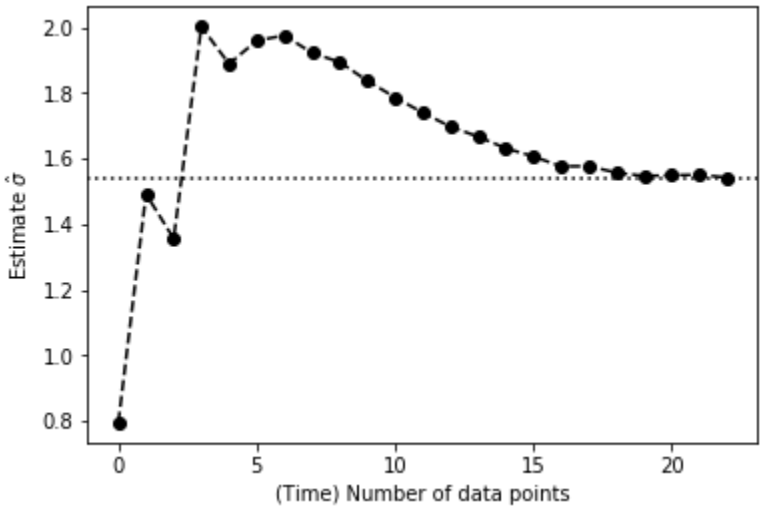

Figure 21: Estimates of σ with more data points being consumed

We choose some constants for μ_σ and η for estimating σ. For example, we set $\mu_\sigma = 0.1$, $\eta = 0.1$ for plotting Figure 21. To show how these choice will affect the estimates, we vary their values and then solve for σ. The results are shown in Figure 22. It turns out that the estimate of σ is insensitive to the choice of parameters in prior distribution.

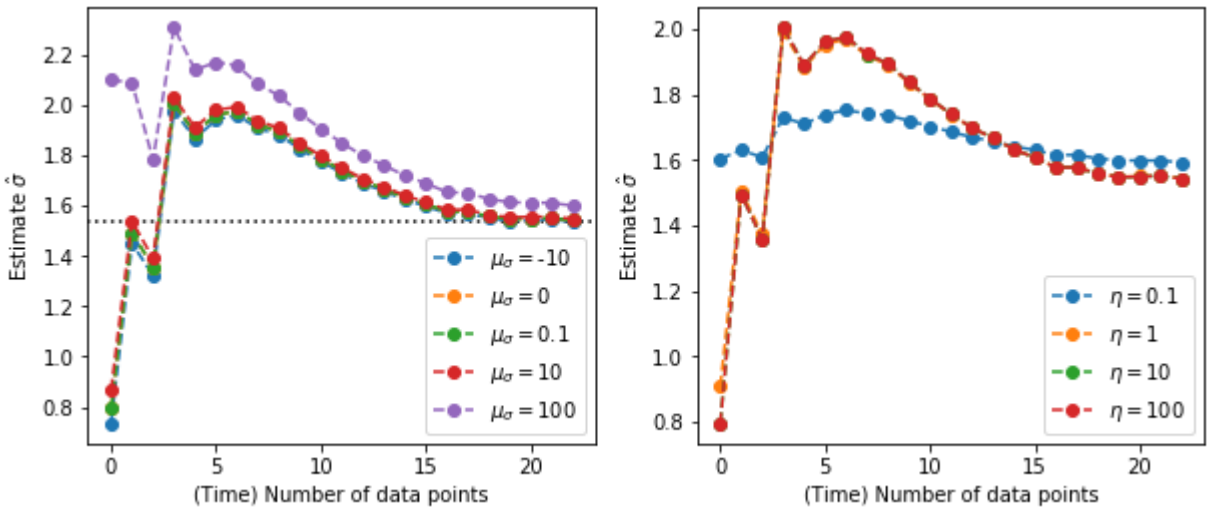

Figure 22: Sensitivity analysis of estimate $\hat{\sigma}$

8 Conclusion

In this report we have studied how errors arise when humans and computers attempt to track pedestrians recorded on ceiling mounted cameras. Given that no ground truth is known about the measures of interest, we have taken steps towards developing a statistically robust protocol to verify CrowdVision's software system. In summary:

- Our literature review found that there is relatively little research that investigates how errors arise when humans or computers are used to count objects of interest. Studies suggest that real-time, in-field counting errors can be large, there are indications that errors decrease with better training, errors are higher at the beginning and end of observations, and shorter observations give rise to relatively fewer errors.
- Using simple geometry, we related crowd density to errors due to occlusion and performed a preliminary analysis to identify the radius within which errors due to both occlusion and image distortion resulting from fisheye lens are negligible. We also consider how the size of a rectangular lattice of cameras that is needed to avoid such errors. These findings depend on the values of various parameters that need to measured in order to apply the results.
- We conducted an exploratory data analysis of the data provided by CrowdVision, focusing on their flow data. We constructed a simple statistical model of human errors in counting and showed how, with raw count data from two people, this could be used to approximate human counting error and estimate the most likely ground truth count. This could be extended to include over and under counting, although we would need to develop additional numerical procedures to determine the variables of interest. We could also extend this method to include multiple counters, and it would be interesting to see how this impacts on the accuracy of the estimations. The algorithm used to match counts could introduce errors, however, using the full trajectories recorded by CrowdVision's system could lead to a much better matching.
- We connected human counting error to computer error under simplifying assumptions of normality. This also allowed us to estimate how many measurements would be needed in order to get a good approximation for the computer error. The analytical tractability of our approach rests on the normality assumptions, although further testing is needed to see how good these approximations are. It seems possible that this analysis could still be performed if we relax the normality assumptions, but it is likely then to become a numerical problem. This could still be of use to CrowdVision.

From our findings, we recommend that the validation data should perhaps be limited to shorter durations in order to improve its accuracy. However, if this approach is adopted

then there should be an investigation into how the intervals are chosen so as best to avoid bias in the error estimation. We also recommend comparing trajectory data if possible since there the information lost through data aggregation can mask how the errors are manifested, since high over and under count errors overall can result in an approximation that is still close to the truth. Our statistical model ignores the particular causes of error, and this approach may be advantageous given these could differ considerably from person to person.

An important question that arises is how CrowdVision's clients independently validate CrowdVision's product? We have seen evidence that real-time in-situ human counting can be prone to very large errors, and so this approach should not be used. This may need to be made very clear to clients, perhaps through controlled demonstrations that they can participate in. However, validation is understandably very important for clients, and perhaps an alternative method can be developed, e.g. if clients were to have their own employees take part in the counting process.

We were not in a position to analyse CrowdVision's algorithm, e.g. at the level of comparing trajectories. Whilst time consuming, such an analysis may highlight particular features of their software that gives rise to errors. It also remains to analyse the occupancy and queue time data the have. However, since this is derived from the raw data, we expect that conclusions drawn at the level of the raw data will carry through to the data that is derived from them. The precise nature of how this happens could however be complex.

In summary we have shown that there is scope to develop a statistical framework within which the quantification of errors in CrowdVision's software can be made. We feel that this goal has wider importance, given that applications of computer vision and artificial intelligence algorithms are likely to increase in the future.

References

[1] A. Johansson, "Data-driven modeling of pedestrian crowds," 2008.

[2] P. Zheng and M. Mike, "An investigation on the manual traffic count accuracy," *Procedia-Social and Behavioral Sciences*, vol. 43, pp. 226–231, 2012.

[3] F. Hsiung, T. McCollum, E. Hefner, and T. Rubio, "Comparisons of count reproducibility, accuracy, and time to get results between a hemocytometer and the tc20 automated cell counter," *Bulletin*, vol. 6003, pp. 1–4, 2013.

[4] M. C. Diogenes, R. Greene-Roesel, L. S. Arnold, and D. R. Ragland, "Pedestrian counting methods at intersections: a comparative study," *Transportation Research Record*, vol. 2002, no. 1, pp. 26–30, 2007.

[5] R. Greene-Roesel, M. C. Diogenes, D. R. Ragland, and L. A. Lindau, "Effectiveness of a commercially available automated pedestrian counting device in urban environments: Comparison with manual counts," 2008.

[6] H. S. Wilf, *generatingfunctionology*. AK Peters/CRC Press, 2005.

[7] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

A Borel-Tanner

For sufficiently busy airports one assumes that for any queue there must be multiple servers to fit the demand. We can model this system as an $M/D/k$ queue: the arrival rate λ is Markovian and modelled according to a Poisson process; the service time μ is deterministic and there exists a finite k number of servers to the queue. Borel-Tanner is a distribution for the total number of customers served before the queue vanishes. Its probability density distribution is

$$p(x; L, k) = \frac{k}{(x-k)!} x^{x-k-1} L^{x-k} \exp(-Lx), \qquad x = k, k+1, \ldots, 0 < L < 1, \tag{39}$$

where $L = \lambda\mu$