Machine Vision and Applications (2003) 13: 344–351
Digital Object Identifier (DOI) 10.1007/s00138-002-0100-3

Machine Vision and
Applications

Evaluation of statistical and multiple-hypothesis tracking for video traffic surveillance

Jeffrey E. Boyd[1], **Jean Meloche**[2]

[1] Department of Computer Science, University of Calgary, Calgary, AB T2N 1N4, Canada; e-mail: boyd@cpsc.ucalgary.ca
[2] Department of Statistics, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Received: 30 August 2001 / Accepted: 28 May 2002

Abstract. Conventional tracking methods encounter difficulties as the number of objects, clutter, and sensors increase, because of the requirement for data association. Statistical tracking, based on the concept of network tomography, is an alternative that avoids data association. It estimates the number of trips made from one region to another in a scene based on interregion boundary traffic counts accumulated over time. It is not necessary to track an object through a scene to determine when an object crosses a boundary. This paper describes statistical tracing and presents an evaluation based on the estimation of pedestrian and vehicular traffic intensities at an intersection over a period of 1 month. We compare the results with those from a multiple-hypothesis tracker and manually counted ground-truth estimates.

Key words: Tracking – Statistics – Network tomography – Statistical methods – Video – Motion

1 Introduction

Tracking normally refers to the estimation of an object's state trajectory, i.e., the value of an object's state as it changes over time. In this sense, tracking refers to a process that *follows* an object as it moves around. The need for tracking to interpret video would seem universal. There are numerous examples in video surveillance, including work by Mikic et al. (1998), Grimson et al. (1998), Cohen and Medioni (1998), and Boult et al. (1998).

The Kalman filter (Gelb 1974) in its simplest form estimates the state of a single object over time from a series of measurements corresponding to that object. Tracking is rarely that simple – there may be multiple objects with measurements for any subset of them, as well as clutter, i.e., measurements that do not correspond to any object. It is first necessary to associate each measurement with the correct object, i.e., solve the data-association problem. The use of multiple sensors further confounds data association. While methods exist to deal with these factors (e.g., Reid 1979; Barr-Shalom and Fortmann 1988; Cox and Hingorani 1996; Isard and Blake 1998), tracking remains a difficult problem in video applications, especially as the number of tracks, amount of clutter, and number of cameras increases.

Boyd et al. (1999) showed that, at least for some applications, estimating trajectories is not necessary for analyzing activity if only long-term patterns are required. Their method of *statistical tracking* is based on *network tomography* (Vardi 1996). Network tomography solves the inverse problem of estimating *source–destination* (SD) network traffic from link statistics alone. If we model a scene under surveillance as a network of connected regions, then the network nodes are physical spaces and the links are the passages between the spaces. Network tomography estimates the number of trips between any two spaces (SD traffic) using counts of traffic between adjacent spaces (link traffic). Statistical tracking simplifies the tracking process because it only requires detection of an object as it moves from one region to an adjacent region. Methods already exist for counting objects crossing a boundary in video sequences; e.g., Albiol et al. (2000) demonstrate a system for counting people entering and exiting a railroad car through a single door. This statistical approach simplifies processing of video sequences and avoids the difficult data-association problem in conventional tracking. In exchange, statistical tracking gives up the ability to estimate an object's trajectory at the time the object is moving – we can only know about the path after statistics have been collected over the entire network for a period of time. The method also gives up some resolution in the identification of individual trips.

While statistical tracking has been demonstrated (Boyd et al. 1999), it has yet to be evaluated in any empirical sense. The theoretical underpinnings of network tomography are sound and the numerical methods used to solve the tomography problem always provide a set of nonnegative SD traffic estimates. This paper presents an evaluation of statistical tracking when applied to the estimation of pedestrian and vehicle traffic intensities in an intersection on a university campus over a period of 1 month. As part of the evaluation we also look at the performance of a conventional tracker, the *multiple-hypothesis tracker* (MHT) described by Reid (1979) and Cox and Hingorani (1996), and ground-truth estimates based on manual tracking. We selected the MHT because it handles clutter, multiple targets, and a variable number of targets, and an im-

Correspondence to: J.E. Boyd

plementation that is easily adapted to the traffic surveillance problem is available (Cox and Hingorani 1996). In contrast, the joint probabilistic data-association filter (Bar-Shalom and Fortmann 1988) tracks a fixed number of targets, and condensation (Isard and Blake 1998) tracks only a single target (although it maintains a probability distribution that can represent multiple targets).

The evaluation does not show that statistical tracking is superior to conventional methods. Instead it shows the differences between the approaches and the applications for which each is best suited. The evaluation also indicates what performance can be expected given certain conditions, and the nature of the errors inherent in statistical tracking.

2 Statistical tracking

Network tomography provides the theoretical basis for statistical tracking. The following description is summarized from Vardi (1996).

Consider a network consisting of n nodes. Let r be the number of directed links in the network. There are $c = n(n-1)$ possible SD pairs in the network. Let $X_j^{(k)}$, $j = 1, \ldots, c$ be the number of transmitted packets for SD pair j during time period k. Assume that $X_j^{(k)} \sim \text{Poisson}(\lambda_j)$. $\boldsymbol{X}^{(k)}$ is the column vector $(X_1^{(k)}, \ldots, X_c^{(k)})^{\mathrm{T}}$. $\mathbf{A}$ is an $r \times c$ fixed routing matrix. Elements in $\mathbf{A}$ are defined by

$$a_{ij} = \begin{cases} 1 \text{ if link } i \text{ is in the path for SD pair } j \\ 0 \text{ otherwise.} \end{cases}$$

Thus, the columns of $\mathbf{A}$ indicate which links are in the path for each SD pair. We wish to know $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_c)^{\mathrm{T}}$ but we can only measure the link traffic, $\boldsymbol{Y}^{(k)} = (Y_1^{(k)}, \ldots, Y_r^{(k)})^{\mathrm{T}}$, where

$$\mathbf{A}\boldsymbol{X}^{(k)} = \boldsymbol{Y}^{(k)}. \tag{1}$$

Typically c is greater than r, so the solution of (1) is ill-posed.

The problem of network tomography can be solved using the method of moments. The first and second moments of the measurements are given by

$$E(Y_i) = \sum_{l=1}^{c} a_{i,l}\lambda_l, \quad i = 1, \ldots, r \tag{2}$$

$$\operatorname{cov}(Y_i, Y_{i'}) = \sum_{l=1}^{c} a_{i,l}a_{i',l}\lambda_l, \quad 1 \le i \le i' \le r, \tag{3}$$

which relies on the assumed Poisson distribution of $X_j^{(k)}$. Equations (2) and (3) give the $r(r+3)/2$ linear equations in the following system:

$$\begin{pmatrix} \bar{\boldsymbol{Y}} \\ \boldsymbol{S} \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \boldsymbol{\lambda}, \tag{4}$$

where $\boldsymbol{S}$ is a vector composed of the $r(r+1)/2$ sample covariances, and $\mathbf{B}$ is an $(r(r+1)/2) \times c$ matrix. Rows of $\mathbf{B}$ are indexed by $(i, i'), 1 \le i \le i' \le r$, and matches the indexing of $\boldsymbol{S}$. Row (i, i') of $\mathbf{B}$ is the element-wise product of rows i and i' of $\mathbf{A}$. Solution of (4) yields an estimate of SD traffic intensities.

To solve (4) we use the nonnegative least-squares method (NNLS) described by Lawson and Hanson (1974). NNLS finds $\hat{\boldsymbol{x}}$ to minimize $\|\mathbf{E}\boldsymbol{x} - \boldsymbol{f}\|$ subject to $\boldsymbol{x} \ge \mathbf{0}$. In the case of network tomography, we assign

$$\mathbf{E} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \text{ and } \boldsymbol{f} = \begin{bmatrix} \bar{\boldsymbol{Y}} \\ \boldsymbol{S} \end{bmatrix} \tag{5}$$

to solve for $\hat{\boldsymbol{\lambda}}$, the NNLS solution to (4). This finds a solution that gives equal weight to the first- and second-moment equations. The second-moment equations are less reliable because they rely on the independence of the SD processes and on their Poisson distribution. In order to put more emphasis on the first-moment equations, we use a weighted least-squares variation, where

$$\mathbf{E} = \begin{bmatrix} w\mathbf{A} \\ \mathbf{B} \end{bmatrix} \text{ and } \boldsymbol{f} = \begin{bmatrix} w\bar{\boldsymbol{Y}} \\ \boldsymbol{S} \end{bmatrix}. \tag{6}$$

If $w >> 1$, NNLS finds a nearly exact solution to the first-moment equations while relying as little as possible on the less reliable second-moment equations.

Alternative methods for numerical solution include using estimation-maximization as suggested by Vardi (1996), a variation of the downhill simplex method described by Press et al. (1990), or convex projections (Youla and Webb 1982) as described by Boyd et al. (1999).

The application of network tomography to video surveillance is described by Boyd et al. (1999). A single static camera connected to a VCR records images of a scene in which vehicles and pedestrians move about. Figure 1a shows a low-resolution image of a scene on the University of Calgary campus containing roads, sidewalks, parking lots, and vegetation. We divide the scene into regions in the image plane (Fig. 1b) that form a network (Fig. 1c). We then use the method described by Sudderth et al. (1998) to identify moving objects in the foreground and detect when an object moves from one region to another, which represents a transition across a link in the network. Counting the transitions over an extended period of time gives the link traffic data, $\boldsymbol{Y}^{(k)}$, required by network tomography to estimate the SD traffic intensities.

3 Evaluation

We evaluate the ability of competing algorithms to predict the traffic intensities in a scene given by a ground truth. In this case the competing algorithms are two variations of statistical tracking and an MHT. The scene is that shown in Fig. 1a. We base the evaluation on a trial that estimates the traffic for 22 weekdays spread over a period of 1 month in late January and early February, 1999. For each of these weekdays we have 30 min of digitized video for the time interval between 3:15 p.m. and 3:45 p.m. Weather conditions during the trial ranged from sunny and warm to overcast and snowing, i.e., a full range of winter weather conditions. Ground cover ranged from snow to exposed asphalt and grass on warmer days. The site is situated at a latitude of 51° N, so the shadows on clear, sunny days are long and moved quickly as the sun moved towards the horizon during the trial time interval.

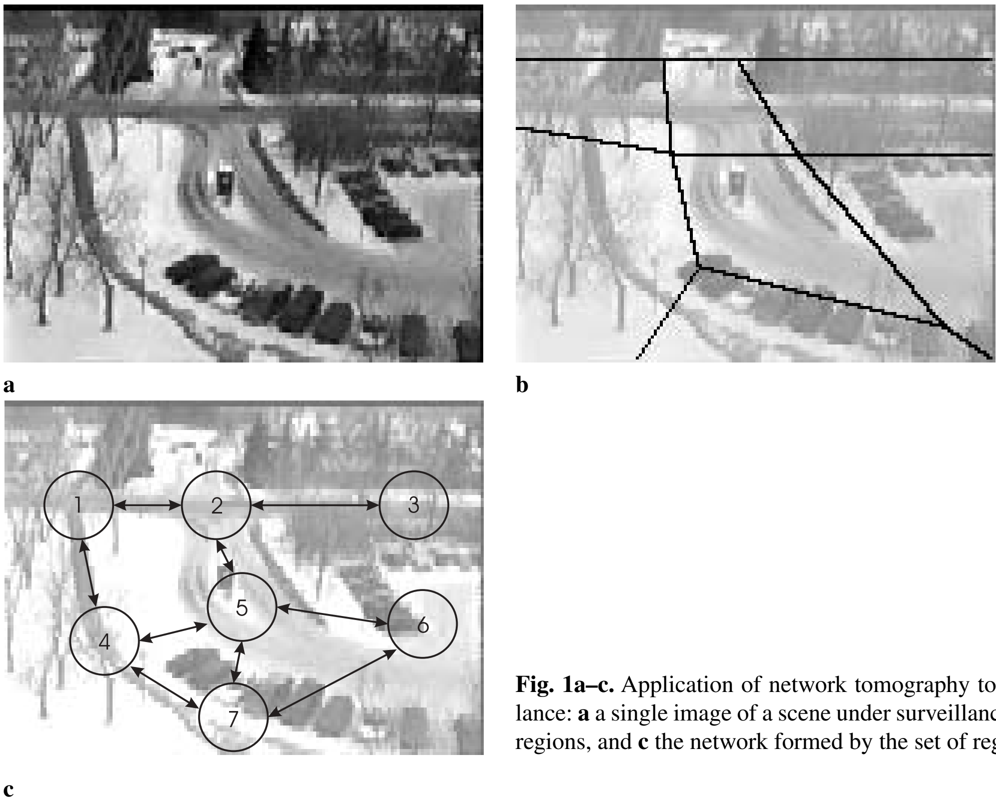


Fig. 1a–c. Application of network tomography to statistical tracking in video surveillance: **a** a single image of a scene under surveillance, **b** the partitioning of the scene into regions, and **c** the network formed by the set of regions

3.1 Data

Figure 2 describes the processing of the sample data to produce four estimates of the SD traffic in the scene corresponding to a ground truth and three machine estimates. The following describes each of these estimates. For the machine estimates, the analog video data is digitized from tape at 10 fps at a resolution of 160×120 pixels, full color. We then segment moving objects from the foreground to obtain a list of object centroids for each frame of video.

Ground truth

An observer manually tracks objects while watching the analog video during playback. From these tracks we derive the ground truth SD traffic estimates, $\boldsymbol{X}_{\mathrm{gt}}$, by looking at the starting and ending nodes for each track. Given $\boldsymbol{X}_{\mathrm{gt}}$, we use (1) to derive the corresponding set of link counts, $\boldsymbol{Y}_{\mathrm{gt}}$.

Simple transition counts

A simple algorithm for counting link transitions compares all the object centroids in one video frame with all the centroids in the previous frame. For each pair of centroids that differ in position by five pixels or less, the algorithm counts a transition for the corresponding link. Transitions within a region are ignored. The result is a vector of count data, $\boldsymbol{Y}_{\mathrm{sim}}$. This algorithm is easily confounded, but its simplicity makes it worthy of comparison to the more sophisticated algorithms.

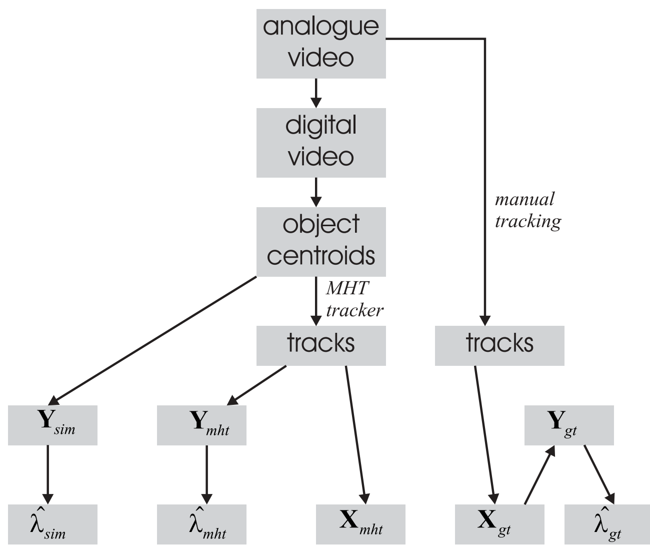


Fig. 2. Processing paths of data used to obtain the estimates of source–destination (SD) traffic used in the evaluation. *MHT*, multiple-hypothesis tracker; *gt*, ground truth; *sim*, simple algorithm

MHT-based transition counts

We use the MHT described by Cox and Hingorani (1996) to track the object centroids. The MHT has several parameters that require tuning for both the hypothesis-tree pruning and the embedded Kalman filter. We selected parameter values by trial and error to obtain reasonable tracks. From the MHT tracks we produce link-transition counts, $\boldsymbol{Y}_{\mathrm{mht}}$, by following the tracks across regions and counting the transitions. In essence, counts from the MHT tracks represent a more sophisticated counting algorithm that considers measurements from several frames before and after the transition.

MHT SD counts

Given the tracks produced by the MHT, we add up the SD traffic by looking at the starting and ending nodes for each track to obtain $\boldsymbol{X}_{\mathrm{mht}}$. This allows comparison of network tomography with estimates produced by conventional tracking.

Nonstationarities

The campus shown in Fig. 1 runs on separate timetables for Monday, Wednesday, and Friday (MWF), and Tuesday and Thursday (TR). This caused us to speculate that the traffic intensities might differ between the two timetables, i.e., the data may not be stationary. Since stationarity is a condition required by network tomography, we partitioned the data into three sets for analysis: the entire 22-day set, the 13-day MWF subset, and the 9-day TR subset.

3.2 Analytical tools

In order to compare numerically the accuracy of competing counting and tracking methods we use the Spearman rank correlation coefficient (Walpole and Myers 1978), r_{S}. This coefficient is similar to the well-know Pearson r value used in linear regression, except that it replaces numerical values in the data with their rank in the data set, making it better suited to data that may not be linear. For two n-element sets of data, $\boldsymbol{X}$ and $\boldsymbol{Y}$, r_{S} is given by

$$r_{\mathrm{S}} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \tag{7}$$

where d_i is the difference between the ranks assigned to x_i and y_i.

It is helpful to put some kind of bound on the estimates produced by statistical tracking. We do this numerically using the Monte Carlo method described in Vardi (1996), by synthesizing Poisson-independent SD counts based on the mean intensities measured in the ground truth. Repeated trials with synthesized data yield a covariance matrix, $\mathbf{C}$, that indicates the variation in the network tomography estimates. The various SD pairs are not independent, but for simplicity of presentation the following figures show error bars that are one standard deviation from the diagonal of $\mathbf{C}$.

3.3 Results

Ground-truth count data

Our first analysis step is to verify that network tomography estimates the SD traffic when we can count the link traffic perfectly. We apply network tomography to $\boldsymbol{Y}_{\mathrm{gt}}$ to obtain an estimate, $\hat{\boldsymbol{\lambda}}_{\mathrm{gt}}$. Figure 3a shows a scatter plot of $\hat{\boldsymbol{\lambda}}_{\mathrm{gt}}$ versus $\bar{\boldsymbol{X}}_{\mathrm{gt}}$, the mean of $\boldsymbol{X}_{\mathrm{gt}}$ over the trial period. The plot superimposes estimates for all days, MWF and TR. Rank correlation coefficients are given in Table 1.

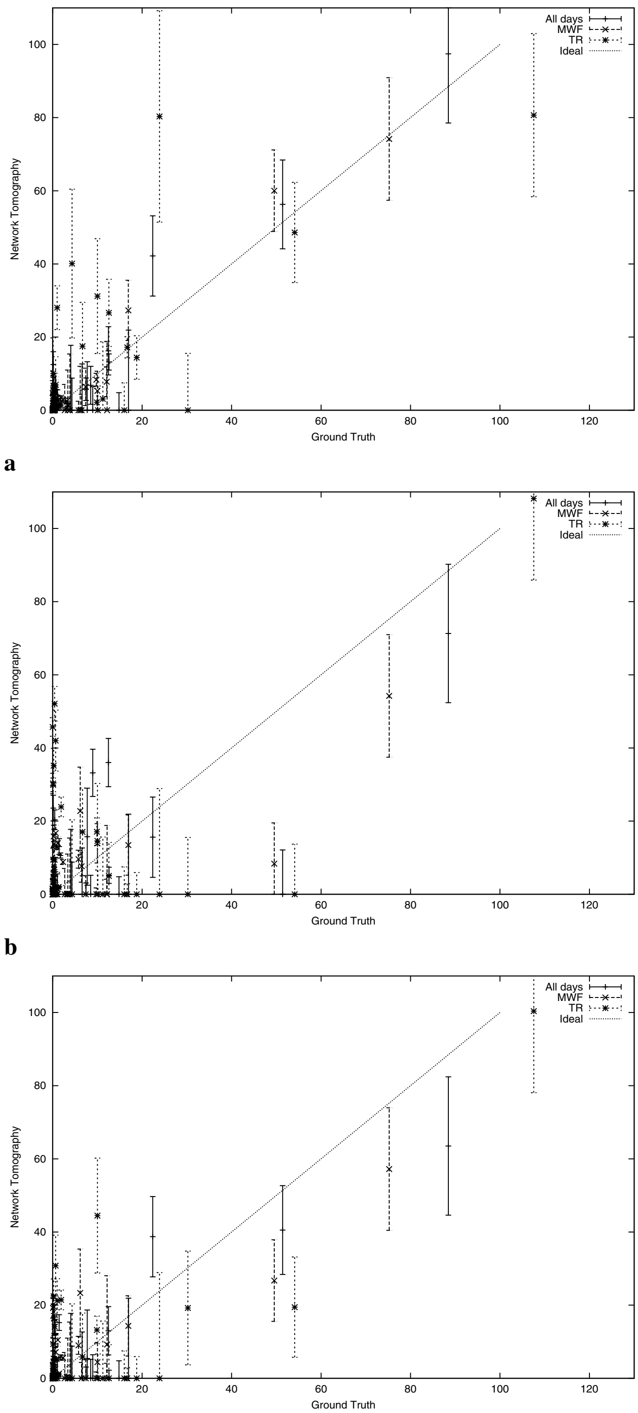


Fig. 3 a–c. Scatter plots showing network tomography estimates of SD traffic intensity versus $\bar{\boldsymbol{X}}_{\mathrm{gt}}$ for $\hat{\boldsymbol{\lambda}}_{\mathrm{gt}}$ **a**, $\hat{\boldsymbol{\lambda}}_{\mathrm{sim}}$ **b**, and $\hat{\boldsymbol{\lambda}}_{\mathrm{mht}}$ **c**. *Error bars* indicate one standard deviation. *MWF*, Monday, Wednesday, and Friday; *TR*, Tuesday and Thursday

Video-based count data

Estimates $\hat{\boldsymbol{\lambda}}_{\mathrm{sim}}$ and $\hat{\boldsymbol{\lambda}}_{\mathrm{mht}}$ are derived from $\boldsymbol{Y}_{\mathrm{sim}}$ and $\boldsymbol{Y}_{\mathrm{mht}}$ using network tomography. Figure 3b and c show scatter plots of $\hat{\boldsymbol{\lambda}}_{\mathrm{sim}}$ and $\hat{\boldsymbol{\lambda}}_{\mathrm{mht}}$ versus $\bar{\boldsymbol{X}}_{\mathrm{gt}}$. Rank correlation coefficients are given in Table 1.

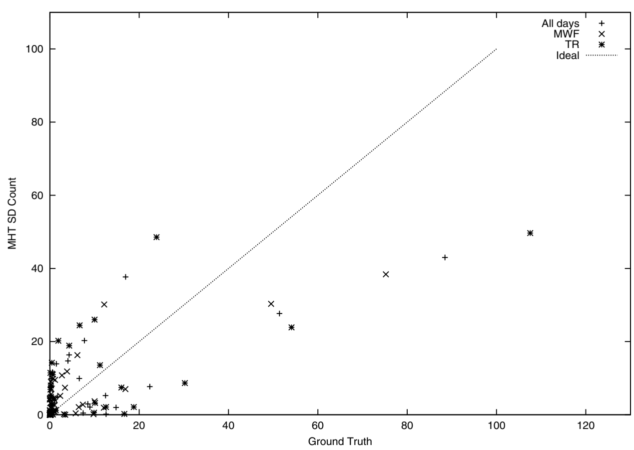


Fig. 4. Scatter plot showing conventional tracking (MHT) estimates of SD traffic intensity versus $\bar{\boldsymbol{X}}_{\mathrm{gt}}$

Table 1. Spearman rank correlation coefficients for source–destination (SD) traffic estimates based on different data sources versus ground truth. *MWF*, Monday, Wednesday, and Friday; *TR*, Tuesday and Thursday

	Data source			
Subset	$\boldsymbol{Y}_{\mathrm{gt}}$	$\boldsymbol{Y}_{\mathrm{sim}}$	$\boldsymbol{Y}_{\mathrm{mht}}$	MHT SD counts
All days	0.50	0.37	0.34	0.60
MWF	0.42	0.34	0.36	0.59
TR	0.60	0.12	0.37	0.57

Table 2. Spearman rank correlation coefficients for SD traffic estimates for split-time intervals and ground-truth counts

Subset	15 min	7.5 min
All days	0.52	0.47
MWF	0.59	0.75
TR	0.61	0.55

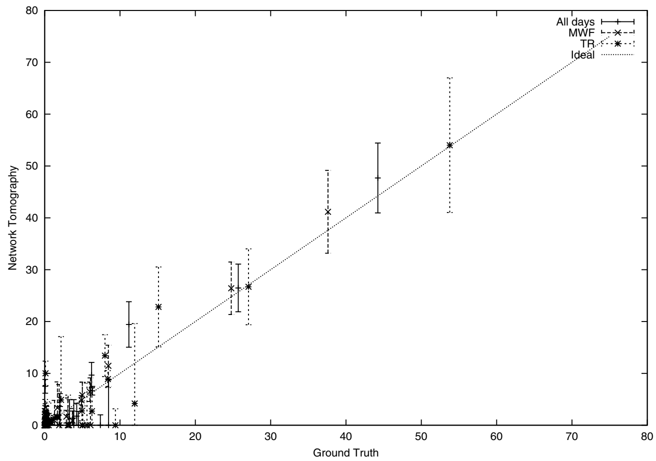
Multiple-hypothesis tracker. A scatter plot of $\bar{\boldsymbol{X}}_{\mathrm{mht}}$, the mean of $\boldsymbol{X}_{\mathrm{mht}}$ over the trial period, versus $\bar{\boldsymbol{X}}_{\mathrm{gt}}$ is shown in Fig. 4. Rank correlation coefficients are given in Table 1.

3.4 Split time intervals

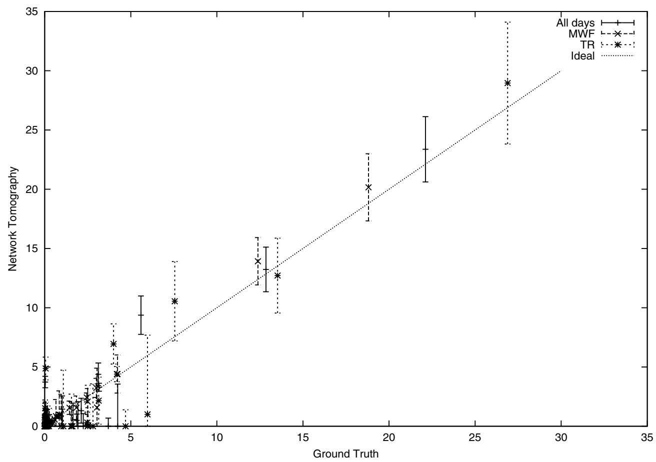
In some cases, it may be possible to improve the statistical estimates by splitting the time intervals to increase sample size. Issues surrounding this approach are discussed in Sect. 4.3. Figure 5 shows the estimates obtained from $\boldsymbol{Y}_{\mathrm{gt}}$ after splitting the 30-min time intervals into two 15-min intervals and four 7.5-min intervals. Rank correlation coefficients, given in Table 2, show an improvement for the MWF set, but not for the others.

4 Discussion

The plots in Figs. 3–5 and the coefficients in Tables 1 and 2 give different views of the data. To compare them, we consider differences inherent in methodology (i.e., statistical vs conventional) separately from differences due to measurement (counting) errors.



a



b

Fig. 5 a,b. Scatter plots showing network tomography estimates of SD traffic intensity versus $\bar{\boldsymbol{X}}_{\mathrm{gt}}$ for 15-min **a** and 7.5-min **b** intervals. Error bars are one standard deviation

4.1 Statistical versus conventional

Consider Figs. 3a and 4, which show a statistical estimate based on perfect count data versus a conventional tracking estimate. The rank correlation coefficients corresponding to these plots are comparable. The most predominant feature of the statistical estimates is the confidence in the estimate (shown graphically by the error bars). The error bars in the statistical estimates are large because the sample sizes are small, ranging between 9 and 22 days. Statistical tracking can improve confidence in estimates by increasing the sample size, i.e., counting over a longer period of time (see Sect. 4.3).

The NNLS solution to (4) contributes errors that are unique to the statistical method. Nonnegativity results in errors that have a positive bias for SD pairs with little or no traffic; we see this in the plots of Fig. 3. Data points are clustered along the vertical axes, indicating overestimation for the paths that have low traffic intensities. The least-squares nature of NNLS leads to another problem: errors in estimates tend to be spread throughout the entire set of SD pairs. However, an error in estimating a high-intensity SD pair, when spread to other pairs, results in a disproportionately large error for the low-intensity

SD pairs. This can be seen in the plots. Errors appear greatest where the true intensity is low.

In contrast to the statistical method, MHT errors are systematic. When the MHT loses the track of an object, and picks it up later, it appears as two tracks for two distinct objects. This breaking of tracks causes the count for the correct, longer path to be underestimated, while the counts for two shorter paths are overestimated. This is shown in Fig. 4 as two lines that diverge from the ground truth. The upper line is over-counted SD pairs for shorter paths, while the lower line is under-counted SD pairs for longer paths. Unlike statistical tracking, counting over longer periods of time will not improve the conventional tracking estimate. It is therefore necessary to prevent the systematic errors due to loss of tracks.

4.2 Measurement and counting

Measurement of link counts accounts for the differences between $\hat{\boldsymbol{\lambda}}_{\mathrm{gt}}$, $\hat{\boldsymbol{\lambda}}_{\mathrm{sim}}$, and $\hat{\boldsymbol{\lambda}}_{\mathrm{mht}}$, and is related to the divergence of the MHT SD counts. We are not able to count the link transitions perfectly and so find additional errors in traffic estimates, and lower r_{S} values.

A major source of link-count errors is partial occlusions. For example, in the scene under surveillance there are several areas where trees and shrubs partially block the view of traffic paths. The blockage is only partial because the branches are defoliated in winter. Two links affected by these partial occlusions are the 4–1 and 2–3 links (see Fig. 1). Further contributing to the problem is variations in lighting due to changes in weather. Count accuracy was best on overcast days with diffuse lighting, and worst on sunny days with high-contrast shadows. Finally, the low image resolution forced by data storage limitations added to the counting difficulties.

We cannot control the weather or remove vegetation for the sake of data collection. However, the problem of image resolution can be overcome by using more cameras. If, for example, we were to use one camera per link, positioning each camera with an optimal view for its link, then counting will be more reliable due to the elimination of problems caused by occlusions and lighting. This is demonstrated by Albiol et al. (2000), where one camera is used to monitor traffic through a single door. The strength of statistical tracking is evident in this approach. No data association is required so it is easy to use many cameras. Conventional tracking would require data association among the different cameras, something that is not easily implemented.

The same factors that confound tracking also affect the MHT. Occlusions, lighting, and limited resolution make it more difficult for the MHT to maintain a track correctly. In the trial described here, the MHT is further confounded by the segmentation algorithm. When an object becomes stationary for a prolonged period, the segmentation algorithm merges it with the background and the MHT no longer receives data for that object, and hence drops the track. This property is convenient when dealing with parking cars (the track for a car should be dropped after the car is parked), and does not affect transition counting. However, the MHT is confused when an object stops momentarily. It is possible to have the MHT maintain a track for a longer period of time, but this leads to the converse problem, that of wrongly merging tracks for two different object. In fact, this is the problem of data association in a nutshell.

4.3 Sample size

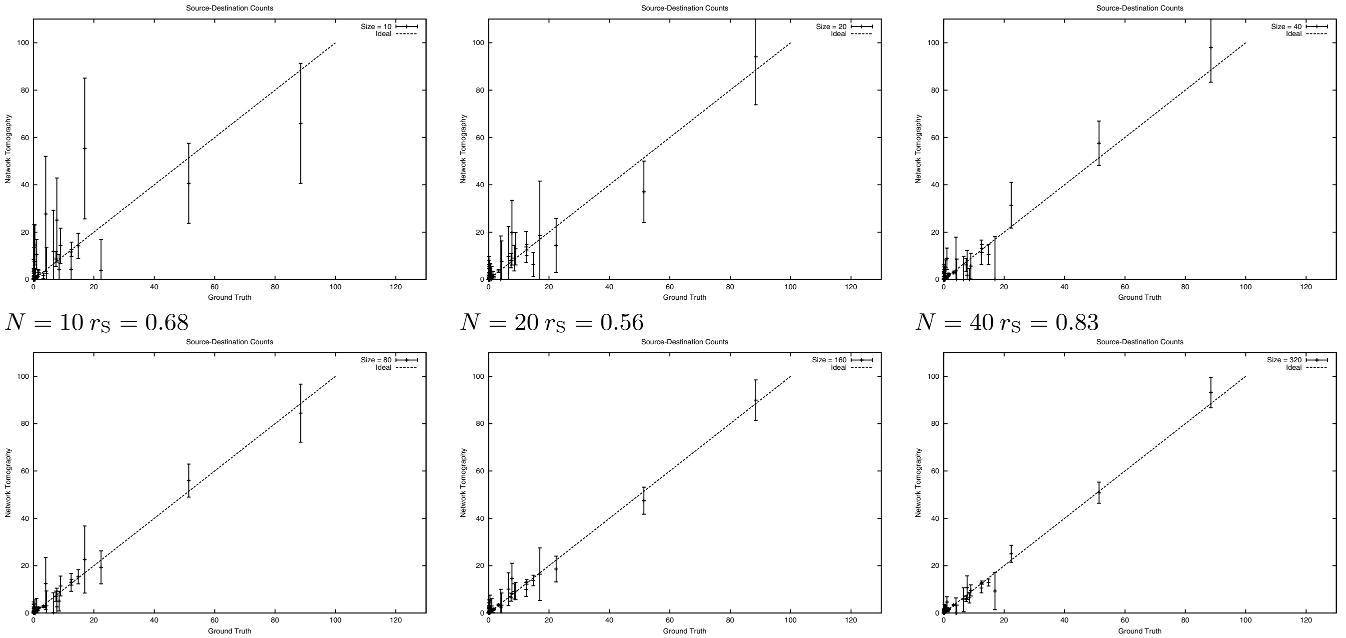
Practical limitations restricted the sample size of the trial, so to predict what is gained from larger samples we conducted trials on synthetic data. Figure 6 shows the results of these trials for sample size N ranging from 10 to 360. Data are for Poisson, independent SD traffic with mean intensities given by $\bar{\boldsymbol{X}}_{\mathrm{gt}}$. Error bars show one standard deviation, and are computed as described in Sect. 3.2.

For $N = 10$ and $N = 20$ the synthetic results are comparable to what is estimated for $\hat{\boldsymbol{\lambda}}_{\mathrm{gt}}$. As N increases so does r_{S}, with $r_{\mathrm{S}} > 0.9$ for $N \geq 320$.

There are two issues to consider when increasing the sample size to improve accuracy. The first is stationarity. Increasing the duration of a trial beyond a time interval for which SD traffic intensity is stationary cannot result in a better estimate. Thus adding more days to our trial may initially improve estimates, but will eventually worsen them as the traffic patterns change throughout the year. Taking more but smaller intervals has the same risk – it only makes sense if the smaller time intervals have similar traffic. The likely reason why splitting the time intervals only yielded an improvement for the MWF subset in Sect. 3.4 is that it was the only subset that was stationary across the 30-min intervals. The second issue is the duration of a typical SD trip. It is relevant when the time intervals in the sample are so short that a trip cannot be completed. For example, in the 22-day trial an SD trip typically takes between a few seconds and a minute. Thus splitting the 30-min interval into 60 30-s intervals is not wise, since many trips might not be completed within the shorter interval. While this may not seem to be a problem with the trial described here, consider the case of estimating traffic intensities for an entire city. Travel times for trips in a city can be 30 min or even an hour depending on the time of day and the location. In such a case, increasing sample size by using shorter time intervals is not an option.

5 Conclusions

With perfect count data, the statistical tracker produced results comparable to those of the MHT. Counting errors due to occlusions, weather, and limited image resolution confound counting and worsen the statistical estimates. Statistical tracking does not exhibit the systematic errors seen in the MHT due to incorrectly dropping the track of an object. Thus, traffic estimates from the statistical tracker can be improved by gathering data over a longer period, and by improving counting with multiple cameras. The ability to easily use multiple cameras with the statistical method is advantageous. It is ideally suited for tasks such as estimating vehicular traffic in a city freeway system or other large areas where data association for conventional tracking is not practical. In order to eliminate the need for data association, statistical tracking gives up the ability to provide current information about a scene. Finally, the ability to track over large areas with multiple cameras has the potential to adversely affect privacy. However, our results



$N = 10\; r_{\mathrm{S}} = 0.68$ $N = 20\; r_{\mathrm{S}} = 0.56$ $N = 40\; r_{\mathrm{S}} = 0.83$

$N = 80\; r_{\mathrm{S}} = 0.79$ $N = 160\; r_{\mathrm{S}} = 0.80$ $N = 320\; r_{\mathrm{S}} = 0.92$

Fig. 6. Statistical tracking of traffic estimates for varying sample size. The figures are based on synthetic Poisson data computed for the network in Sect. 3. $\boldsymbol{\lambda}$ is taken from the ground truth. N, sample size

show this potential is somewhat diminished by the fact that low-intensity traffic routes are difficult to estimate when sharing a network with much higher traffic.

References

Albiol A, Naranjo V, Mora I (2000) Real-time high density people counter using morphological tools. In: Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, Spain, 3–7 September, pp 652–655

Bar-Shalom Y, Fortmann TE (1988) Tracking and data association. Academic, Boston, Mass.

Boult T, Micheals R, Erkan A, Lewis P, Powers C, Qian C, Yin W (1998) Frame-rate multi-body tracking for surveillance. In: Proceedings of the 1998 Image Understanding Workshop, Monterey, CA, 20–23 November, pp 305–308

Boyd JE, Meloche J, Vardi Y (1999) Statistical tracking in video traffic surveillance. In: Proceedings of the Seventh International Conference on Computer Vision, Kerkyra, Greece, 20–27 September, pp 163–168

Cohen I, Medioni G (1998) Detecting and tracking moving objects in video from an airborne observer. In: Proceedings of the 1998 Image Understanding Workshop, Monterey, CA, 20–23 November, pp 217–222

Cox IJ, Hingorani SL (1996) An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. IEEE Trans Pattern Anal Mach Intell 18:138–150

Gelb A (ed) (1974) Applied optimal estimation. MIT Press, Cambridge, Mass.

Grimson WEL, Stauffer C, Romano R, Lee L (1998) Using adaptive tracking to classify and monitor activities in a site. In: Proceedings of Computer Vision and Pattern Recognition 98, Santa Barbara, CA, 23–25 June, pp 22–29

Isard M, Blake A (1998) Condensation – conditional density propagation for visual tracking. Int J Comput Vis 29:5–28

Lawson CL, Hanson RJ (1974) Solving least squares problems. Prentice-Hall, Englewood Cliffs, N.J.

Mikic I, Santini S, Jain R (1998) Video processing and integration from multiple cameras. In: Proceedings of the 1998 Image Understanding Workshop, Monterey, CA, 20–23 November, pp 183–187

Press WH, Vetterling WT, Flannery BP, Teukolsky SA (1990) Numerical recipes in C: the art of scientific computing. Cambridge University Press, Cambridge

Reid DB (1979) An algorithm for tracking multiple targets. IEEE Trans Autom Control 6:843–854

Sudderth E, Hunter E, Kreutz-Delgado K, Kelly PH, Jain R (1998) Adaptive video segmentation: theory and real-time implementation. In: Proceedings of the 1998 Image Understanding Workshop, Monterey, CA, 20–23 November, pp 177–181

Vardi Y (1996) Network tomography: estimating source–destination traffic intensities from link data. J Am Stat Assoc 91:365–377

Walpole RE, Myers RH (1978) Probability and statistics for engineers and scientists, 2nd edn. Macmillan, New York

Youla DC, Webb W (1982) Image restoration by the method of convex projections: part 1 – theory. IEEE Trans Med Imaging 1:81–94

Jeffrey E. Boyd is an assistant professor in the Department of Computer Science at the University of Calgary. He completed a PhD in computer science at the University of British Columbia in 1994, and a BSc and MSc in electrical engineering at the University of Calgary in 1983 and 1986, respectively. From 1994 to 1995 Jeffrey was a research assistant at TRIUMF in Vancouver, British Columbia. He was an NSERC postdoctoral fellow at the University of California, San Diego, from 1996 until 1998. His current research interest is computer vision focussing on the analysis of motion.

Jean Meloche obtained his PhD in 1989 at l'Université de Montréal. He is currently an associate member of the faculty on leave from the Department of Statistics at the University of British Columbia, and has been a researcher at Avaya Labs Research in Basking Ridge since 2001. His research interests are in smoothing methods, robustness, image processing, network tomography, environmental expert systems, data mining, and the network quality of service assessment methods.