AUTOMATING COLLECTION OF PEDESTRIAN DATA USING COMPUTER VISION TECHNIQUES

Simon Li
Tarek Sayed
Mohamed H. Zaki
Department of Civil Engineering
The University of British Columbia
Vancouver, BC, Canada

Greg Mori
Ferdinand Stefanus
Bahman Khanlooa
School of Computing Science
Simon Fraser University
Vancouver, BC, Canada

Nicolas Saunier
Dept. of Civil, Geological and Mining Engineering
École Polytechnique de Montréal
Montréal, Québec, Canada

Word count: 4177 Text + 3 Tables + 9 Figures = 7177 words

ABSTRACT

New urban planning concepts are being redefined to emphasize walkability and to accommodate the pedestrian as a key road user. However, the availability of reliable pedestrian traffic information remains a major challenge inhibiting a better understanding of many pedestrian issues. Therefore, the importance of developing new techniques for pedestrian data collection cannot be overstated. This paper demonstrates the use of computer vision techniques for the automated collection of pedestrian data. Several applications are described, which include pedestrian counting, tracking, and walking speed measurements. An efficient pedestrian tracking algorithm, the MMTrack, is used. The algorithm employs a large margin learning criterion to effectively combine different sources of information. The applications are demonstrated using a real-world data set from Vancouver, British Columbia. The data set includes 1135 pedestrian tracks. Manual counts and tracking were performed in order to validate the results of the automatic data collection. The results show a 5% average error in counting which are considered reliable. The results of the walking speed validation show an excellent agreement between manual and automated walking speed values (RMSE = 0.0416 m/s, R^2 = 0.9269). Further analysis was conducted on the mean walking speed of pedestrians with regard to several factors. Gender, age, and the group size were found to significantly influence the pedestrian mean walking speed. The results demonstrate that computer vision techniques have the potential to collect microscopic road user data at a degree of automation and accuracy that cannot be feasibly achieved by manual or semi-automated techniques.

INTRODUCTION

New urban planning concepts place considerable emphasis on sustainable modes of transportation, particularly walking. Consequently, a solid understanding of pedestrian behaviour is of considerable interest. However, one of the main challenges in conducting detailed analysis of pedestrian behaviour is the lack of reliable data. Collecting reliable pedestrian data is often labour-intensive and time consuming as it is usually collected by manual counts or measurements. This lack of reliable data can have a significant impact on several transportation engineering and planning aspects. Two of the important areas of pedestrian data collection are volume counts and average walking speed measurement. Inconsistent and scarce pedestrian counts make it difficult to justify capital for implementing pedestrian facilities. Lack of data for pedestrian walking speed leads to average values being applied across substantially large regions, ignoring any variations that arise due to factors such as geographical and behavioural differences. Therefore, accurate walking speed data is important since it is used in various applications such as traffic signal timing, and commute time calculations for planning purposes.

In recent years, some degree of automation has been implemented for pedestrian data collection. For instance, turnstile and overhead laser counters (1) are used to count transit users. However, these approaches have geometric restrictions such as the need for physical choke points or the feasibility of implementing overhead structures at pedestrian walkways. As a result, most pedestrian counts in open space are still performed manually.

Given the aforementioned issues, there is a significant need to explore methods of increasing the availability of pedestrian data. One of these methods is to develop reliable automated techniques for data collection such as the use of computer vision techniques. Computer vision techniques are not new to the transportation field. For example, in recent years computer vision has been used to track vehicle and pedestrian trajectories to study traffic conflicts and their implications on traffic safety (2,3). This paper will demonstrate the use of a set of computer vision techniques for the automated collection of pedestrian data. The pedestrian tracks obtained are then employed to perform screen line based pedestrian counting as well as pedestrian speed measurements. The tracking algorithm used in this paper is the MMTrack (4), which employs a large margin learning criterion to combine different sources of information effectively. The applications are demonstrated using a real-world data set from Vancouver, British Columbia. The data set includes 1135 pedestrian tracks. Further analysis was conducted on the mean walking speed of pedestrians with regards to several factors.

PREVIOUS WORK

The two main areas of research that are relevant to the work described in this paper are related to the analysis of pedestrian walking speed, and automatic pedestrian tracking techniques.

Considerable literature exists on studying pedestrian walking speed. One of the main sources of published pedestrian walking speed for calculation purposes is within the Manual of Uniform Traffic Control Devices (MUTCD). The MUTCD recommended average walking speed has changed over the years and there are recent proposals for using varying values for different applications or across different groups. Generally, studies report walking speed ranging between 3.0 to 4.95 feet-per-second (5). Some studies have also investigated the effect of variables such as noise (6), temperature (7), area type (8) and mobility impairments (9) on average pedestrian walking speed. A significant difference between the walking speed of young and older

pedestrians was reported in (10,11). A large study of over 3500 pedestrians in Jordan found that age, gender, group size, and street width had an impact on the mean walking speed (12). These studies were conducted using manual measurements of time to cross a predetermined distance, and the classification of pedestrians was done by direct observations or questionnaires. One drawback of these studies is that they can only study average speed over a certain time span or a distance, and cannot calculate instantaneous velocity. Thus, further investigation on notions such as stride length or pedestrian acceleration cannot be performed.

Most of the efforts to automate the collection of pedestrian data have made use of sophisticated methods through the use of thermal and infrared cameras (13). There has also been past use of pedestrian tracking with video cameras. Ismail et al. have proposed an algorithm for such a purpose (3). One study combined the use of vehicle and pedestrian tracking techniques to study pedestrian-vehicle conflicts (3). Yang et al. used appearance model techniques to track multiple pedestrians within a video stream (14). Some difficulties they have cited include over-grouping in cases where multiple pedestrians have attire of similar color and intensity. They proposed that further research is needed to implement face recognition features to separate pedestrians. Hoogendoorn et al. (15) used data from microscopic simulation models to detect pedestrians. The results were fairly accurate, but the study was conducted in an idealized setting with visually marked test subjects (15). The quality of the tracking in the study of Hoogendoorn et al. (15) may not be representative, as one of the concerns with computer vision is tracking performance in open and busy spaces, defined by mixed use of space and a concentration of users. Nevertheless, using test subjects they were able to investigate tracking ability in various situations such as bottle-necking. Most of these algorithms are not fully automated, and require considerable human intervention and additional data (16). A study by Kelly et al. (17) attempted to track pedestrians within crowded areas. They found that their algorithm was fairly robust but they had difficulty optimising the code to reduce the computation burden (17).

Generally, most of the work undertaken in the area of automated pedestrian tracking cites problems with respect to difficulties in separating objects, negative correlation between accuracy and object density, lack of study on real settings, and the lack of computational power for the use of complex algorithms (15, 16). The algorithm used in the current paper was shown to address many of the abovementioned shortcomings and was able to obtain relatively accurate results based on video footage from real life settings (4).
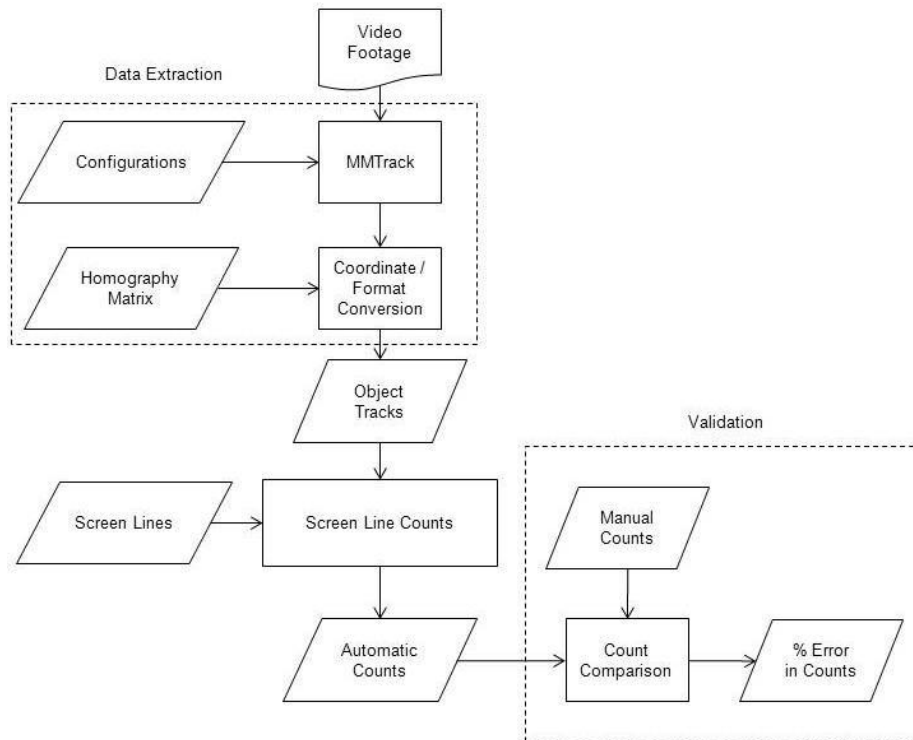
METHODOLOGY

Figure 1 illustrates the pedestrian counting methodology. The first component is data extraction which is divided into two steps. The first step is to use a computer vision algorithm (MMTrack) to detect pedestrians as objects and to track them. The algorithm employs a cluster-based appearance modeling and online tracking approach. Further details of the algorithm can be found in (4). The second step is to create a mapping from world coordinates to image plane coordinates using a homography matrix (camera calibration). This mapping enables the recovery of real-world coordinates of points that appear in the video. The matrix parameters specify the translation and orientation of the camera coordinates relative to the world coordinates. The parameters are obtained by minimizing the difference between the projection of geometric entities, e.g. points and lines, onto world or image plane spaces and the real-world measurements of these entities. The process is described in detail in (18).

A screen line counts procedure will identify the pedestrian tracks crossing a predefined screen lines in a selected location. The sum of the counts represents the final output of the

process, which are the automatic count of pedestrians crossing the screen line. The results of the automatic counting process are validated with manual counts made for the same screen line and same footage. A comparison of the counts is used as a measure of the accuracy the counts.

The methodology for performing automatic mean speed measurements is illustrated in Figure 2. The object tracks from the tracking algorithm are used directly as an input for mean speed measurements. Two parallel screen lines with a known distance apart are specified on the screen, and the time elapsed for each track to pass through the two lines is recorded. Dividing the known distance by the elapsed time generates the mean speed of that particular track.

**Figure 1. Screen Line Counts Methodology**
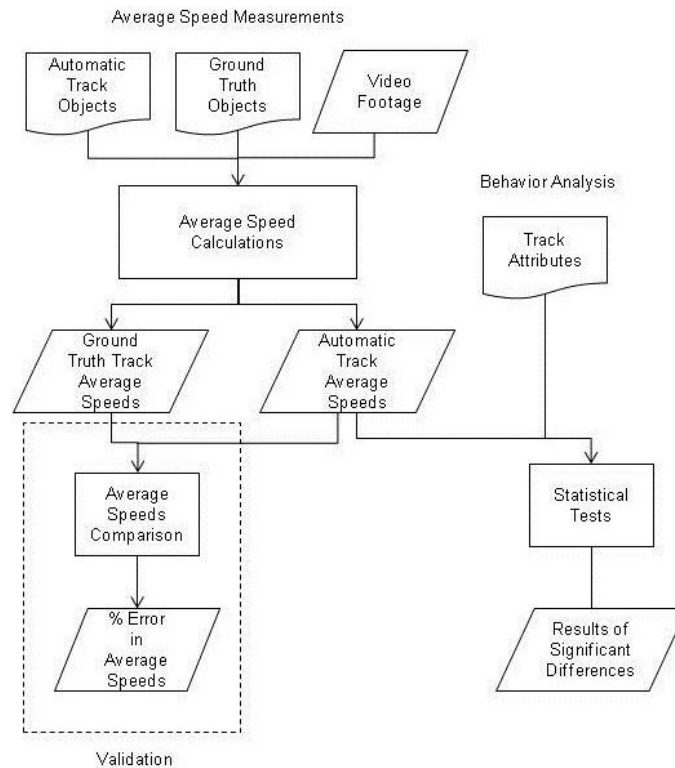
Figure 2. Walking Speed Measurements Methodology

To validate the results of the speed measurement, a specific length on the image with known physical distance is used. This section of road space is selected such that there is a high degree of overlap with the section defined by the screen lines in the automatic speed calculations. This ensures that the results of the speed measurements are comparable. A certain number of pedestrians with proper tracks are selected, and the time it takes for them to transverse the specific distance is obtained. By dividing the length of the known distance by the time elapsed, the average walking speed is obtained for the selected pedestrians. The manually calculated speed is compared to the automatically measured speed and the Root Mean Square Error (RMSE) is calculated. After the validation, the mean speed calculated from the automatic tracking can be used for further analysis. Pedestrians are classified based on their attributes, and statistical tests are used to test for significant differences between the groups.

CASE STUDY

This section describes the analysis of video sequences collected from an open busy environment in the Vancouver Downtown area, as shown in Figure 3. Videos were selected from a data set collected for pedestrian movement at the traffic intersection of Robson and Broughton Streets. Robson Street is a major commercial and business corridor in Vancouver Downtown area with an active walking environment. The same video footage was used by Saunier et al. in a study of pedestrian stride length and frequency using pedestrian tracking from a different algorithm (19). The intersection is a four-leg and is signal-controlled. The two-way streets are line separated. Each leg has one traffic lane and one parking/reserved lane. A 2-phase signal is used in the intersection. A total of 31 minutes video footage was selected and the timing of the video survey

was intended to be concurrent with a nearby special event in order to capture higher pedestrian volumes. East is roughly on the left side of the image. Figure 4 and Figure 5 show the intersection with the calibration demonstrated by corresponding grids in the video image and world (an orthographic satellite image) coordinates respectively. Figure 6 shows a video image along with real-world tracks of pedestrian movement obtained using the video analysis system.
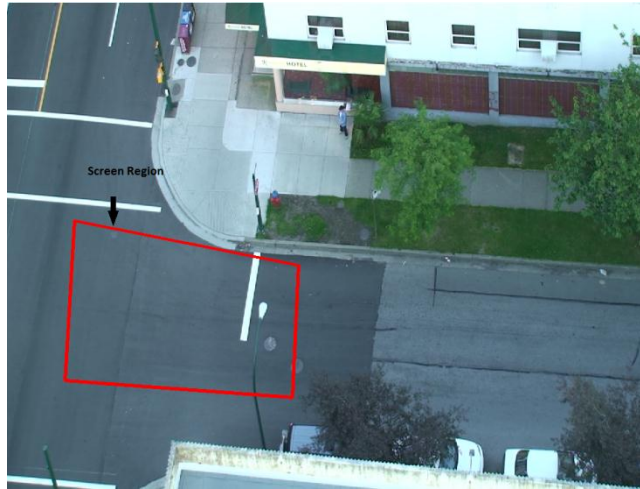
Figure 3. Screen Region

Figure 4. Sample Grid on Image Space

Figure 5. Projected Sample Grid on World Space

Figure 6. Sample Pedestrian Tracks

Validation

The quality of the tracking is assessed based on two measures:
1. Ability to accurately detect pedestrians, measured by the validation of counts made with screen lines
2. Ability to track pedestrians accurately in space and time, measured by the validation of average speed measurements taken from the tracks

Pedestrian Counting

The counting validation was performed by analyzing the tracks within MatLab. A screen line was placed spanning midway across the southern crosswalk and a MatLab script was devised to count the number of times a track crossed this line. The results of the script indicated that the MMTrack algorithm detected 1369 tracks that cross this screen line. Manual counting for the same screen line and footage found 1445 pedestrians. The resulting error rate of the counts is 5.3%, which is

fairly acceptable. Sample pedestrian tracks along with a selected screen are shown in Figure 7. Several anomalies were found in the tracking. These anomalies can be defined as:

Oversegmentation: Oversegmentation occurs when more than one track is attributed to a single pedestrian. In terms of counting, the effect of this is an inflation of counts in the case that both the tracks will pass the screen line. A common cause for this is if the person is accompanied by a large object such as a stroller.

Overgrouping: Overgrouping occurs when several pedestrians are grouped together as a single object. This can occur when several people are walking at a fairly close distance, dress similarly, or have similar walking speed. The effect of this is a deflation of actual counts. One emphasis is that overgrouping is defined as when several pedestrians are grouped together; this means that the centroid of the object should be near the centroid of the group.

Misdetection: Sometimes the tracking algorithm will pick up noise within the video, and detect a non-pedestrian as an object. A few non-pedestrians that were tracked include the manhole within the sidewalk, and one of the stop signs. It is hard to determine what the effects of misdetection on pedestrian counts are because some mis-detected objects have stationary tracks, while some may jump across the screen, triggering a count if it happens to pass one of the screen lines.

Missed Detection: Occasionally, a pedestrian is simply not tracked. This should be separated from overgrouping. In overgrouping, the pedestrian is clearly tracked but simply grouped with others. For missed detection, there is no evidence that the pedestrian follows any nearby trajectories. Ultimately, missed detection results in a deflation of pedestrian counts.

Overall, oversegmentation, overgrouping, and misdetection account for 37%, 42%, and 21% of the errors, respectively. For every 100 pedestrians that enter the screen, only about 1-2 are missed.
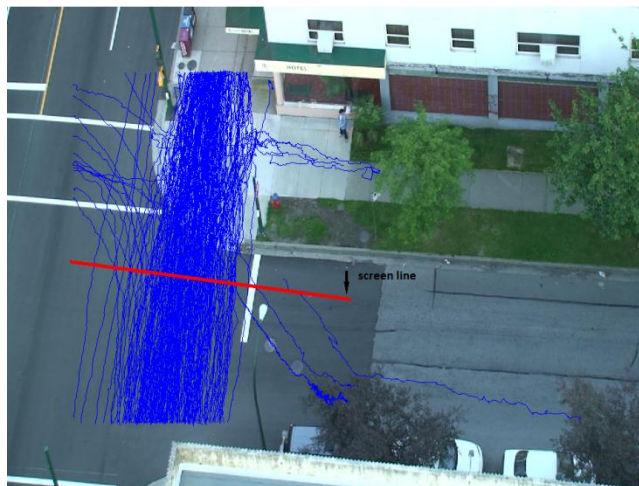
Figure 7. Screen-based Counting

Average Speed Measurements

The average speed validation was performed in a similar manner as count validation. Two parallel screen lines were placed on the southern crosswalk, with a known physical distance between them. The time it took a track to cross both lines is recorded and divided by the distance in order to measure average speed. Seventy five pedestrians with good tracking (continuous throughout) were selected. The length of the traffic stop line near the bottom of the screen was determined to be 4.02m. The time it took for the 75 pedestrians to travel across this distance was timed manually through the video footage. The results of the validation show an excellent agreement between manual and automated walking speed values (RMSE = 0.0416 m/s, R^2 = 0.9269). Sources of residual errors include the assumption that the pedestrians follow the shortest path between two check lines, camera calibration inaccuracy, and pedestrian tracks noise.
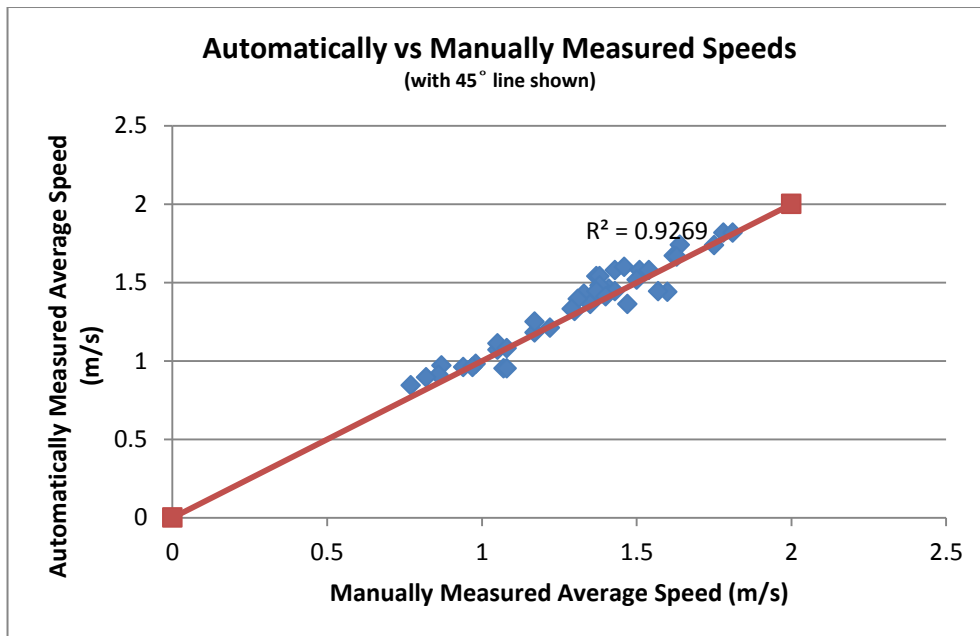

Figure 8. Results of Validation of Walking Speed Measurements

PEDESTRIAN BEHAVIOUR ANALYSIS

Further analysis was conducted on the data to investigate the effect of pedestrian characteristics on walking speed. Several variables were considered including: age, gender and group size. The effect of the direction of travel was also investigated because of the presence of a grade in one direction. This study is non-intrusive, so the values of the attributes have to be assigned to pedestrians based on discretion and good judgement. As a result, certain variables, such as age, which is fundamentally a continuous variable, have to be divided up into categorical groups. The final rubric for pedestrian classification is summarized in Table 1.

Variables	Categories			
Age Group	< 16	16-35	35-55	55+
Gender	Male	Female	-	-
Group Size	1	2	3	4+
Direction	Eastbound	Southbound	Westbound	Northbound

**Table 1. Classification Groups**

The above variables are the ones that were deemed discernable through a non-intrusive approach. A maximum of 4 categories for each variable was specified, either because of finite number of options, or because this was the maximum number that was considered possible for a non-intrusive classification. A summary of the classification is shown in Table 2.

	Categories	< 16	16-35	35-55	55+
Age Group	**Count**	58	634	395	48
	Proportion	5.10%	55.90%	34.80%	4.20%
	Categories	**Male**	**Female**		
Gender	**Count**	597	538		
	Proportion	52.60%	47.40%		
	Categories	**1**	**2**	**3**	**4+**
Group Size	**Count**	226	572	229	108
	Proportion	19.90%	50.40%	20.20%	9.50%
	Categories	**Eastbound**	**Southbound**	**Westbound**	**Northbound**
Direction	**Count**	233	19	879	4
	Proportion	20.50%	1.70%	77.40%	0.40%

Table 2. Classification Results

Effect of the Direction of Travel

The intersection has a slight upgrade in the Eastbound direction. An initial hypothesis is that the direction of travel, related to grade, will have a significant impact on the mean walking speed, such that it will overshadow the influence of other variables. To verify this hypothesis, the distribution of the data is plotted and a two-sample T-test is performed between the Eastbound and Westbound groups.
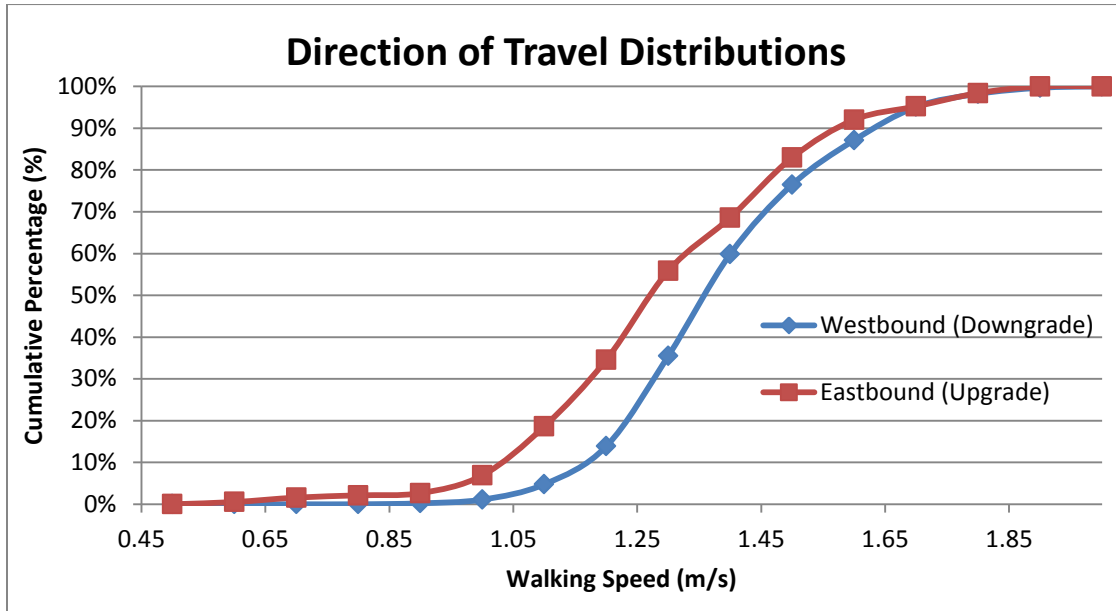
Figure 9. Cumulative Distribution of Walking Speed by Direction of Travel

As shown in Figure 9 and Table 3, there appears to be a significant difference in mean walking speed between the Eastbound and Westbound group. This difference is statistically significant ($p<0.0001$). Without isolating its effect, the influence of grade may bias the results of investigating the effect of other variables. As a result, the Eastbound direction was excluded from further analysis.

Attribute	Variable	Statistics		Significance (p-value)
		Count	**Mean Speed (SD)**	
Direction	Westbound	879	1.38 (0.18)	<0.0001
	Eastbound	233	1.29 (0.23)	
Gender (Excluding Eastbound)	Female	444	1.37 (0.17)	0.049
	Male	458	1.39 (0.19)	
Age (Excluding Eastbound)	<16	30	1.34 (0.14)	<0.0001
	16 - 35	628	1.40 (0.18)	
	36 - 55	228	1.34 (0.16)	
	>55	16	1.32 (0.14)	
Group Size (Excluding Eastbound)	1	162	1.46 (0.20)	<0.0001
	2	497	1.37 (0.17)	
	3	186	1.35 (0.15)	
	4+	57	1.32 (0.21)	

Table 3: Results of Pedestrian Behaviour Analysis

Table 3 indicates that gender, age, and group size significantly influence pedestrian walking speed. Male subjects had a slightly, but significantly higher mean speed compared to female subjects. Knoblauch et al. (10) and Tarawneh (12) also found significantly higher mean walking speeds for males compared to females. Mean walking speed was highest for the age group 16-35 years old and decreased with increasing age. The mean walking speed for the age group 16-35 was also greater than that of the younger age group. Older pedestrians (age 55 or higher) had the lowest mean walking speed. The results are generally consistent with findings by Knoblauch et al. (10), Fitzpatrick et al. (11), and Tarawneh (12). Table 3 also shows that the mean walking speed decreases with increasing group size. Knoblauch et al. (10) found the mean walking speed to be significantly higher for pedestrians walking alone compared to those walking in groups. Tarawneh (12) found the mean walking speed of pedestrians walking in groups of 3 or more to be significantly higher than those walking alone or in pairs.

DISCUSSION AND CONCLUSIONS

The result of the automatic counting and tracking were generally acceptable in terms of their accuracy (average percentage of error). This demonstrates that accurate pedestrian counts and tracking can be performed using computer vision techniques. This can expand the possibilities for pedestrian data collection significantly. The automation of pedestrian data collection can widen the range of feasible data both geographically (different locations) and temporally (for longer periods of time).

Several conclusions can be drawn from the analysis of pedestrian walking speed behaviour. Gender, age, and group size were found to significantly influence the pedestrian mean walking speed. Males had a slightly, but significantly greater mean walking speed compared to females. The mean walking speed was highest for the age group 16-35 years old and decreased with increasing age. The mean walking speed for the age group 16-35 was also greater than that of the younger age group. Older pedestrians (age 55 or higher) had the lowest mean walking speed. The mean walking speed decreases with increasing pedestrian group size.

The main limitations of this study lie within the pedestrian classification. The first limitation is related to the accuracy of the assignment to groups which is based on discretion and good judgement. Accuracy is dependent on quality and resolution of video, and as mentioned before, the range that each category covers. Because this study is non-intrusive, no validation can be performed for the classification. Second, because the study is non-intrusive, continuous variables such as age have to be divided into categorical ranges. Analysis tools such as multivariate linear regression are much more applicable on continuous data as opposed to categorical data.

One area of potential research is to assess a relationship between the accuracy of tracking and object density. Visual inspection reveals that some pedestrians were over-segmented (multiple tracks for a single pedestrian) or over-grouped and occluded, particularly when pedestrian density was high. Of the various possible errors in tracking, overgrouping appeared to be the most prominent. Another potential area of further research is to study the influence of grade and walking surface on pedestrian. Lack of multiple locations and lack of adequate sidewalk distance within the video footage are reasons that these variables could not be studied within this paper.

REFERENCES

1. Cekander E.J. Design of a Computer Controlled Turnstile. San diego: WESCON/94. 'Idea/Microelectronics'. Conference Record 1994.
2. Saunier, N., Sayed, T., and Ismail, K. (2010) Large Scale Automated Analysis of Vehicle Interactions and Collisions, Transportation Research Record: Journal of the Transportation Research Board, Vol. 2147, pp. 42-50, DOI: 10.3141/2147-06.
3. Ismail, K., Sayed, T., and Saunier, N. (2010) Automated Analysis of Pedestrian-vehicle Conflicts: A Context for Before-and-after Studies, Transportation Research Record: Journal of the Transportation Research Board, Vol. 2198, pp.52-64.
4. Khanloo, B., Stefanus, F., Ranjbar, M., Li, Z., Saunier, N., Sayed, T. and Mori, G. (2010) "Max-Margin Offline Pedestrian Tracking with Multiple Cues", Seventh Canadian Conference on Computer and Robot Vision (CRV 2010), Ottawa, Ontario, 31 May - 2 June 2010.
5. LaPlante J, Kaeser T.P. A History of Pedestrian Signal Walking Speed Assumptions. 3rd Urban Street Symposium. Seattle, Washington. 2007.
6. Boles W.E, Hayward S.C. Effects of Noise and Sidewalk Density upon Pedestrian Cooperation and Tempo. Journal of Social Psychology 1978; 104:29-35.
7. Walmsey DJ, Lewis GJ. The Pace of Pedestrian Flows in Cities. Environmental and Behavior 1989; 21(2):123-150.
8. Al-Masaeid H, Al-Suleiman T, Nelson N. Pedestrian Speed-Flow Relationship for Central Business District Areas in Developing Countries. Transportation Research Record: Journal of Transportation Research Board 1993; 1396:69-74.
9. Arango, J. and Montufar, J. Walking Speed of Older Pedestrians Who Use Canes or Walkers for Mobility. Transportation Research Record: Journal of Transportation Research Board 2073:79-85.
10. Knoblauch R.L, Peitrucha M.T, Nitzburg M. Field Studies of Pedestrian Walking Speed and Start-Up Time. Transportation Research Record: Journal of Transportation Research Board 1996; 1538.
11. Fitzpatrick, K., Brewer M.A, and Turner S.M. Another Look at Pedestrian Walking Speed. In Transportation Research Record: Journal of the Transportation Research Board, No. 1982, 2006, pp. 21–29.
12. Tarawneh M.S. Evaluation of Pedestrian Speed in Jordan with Investigation of Some Contributing Factors. Journal of Safety Research, 2001; 32(2):229-236.
13. Kerridge J, Chamberlain T. Collecting, Processing and Calculating Pedestrian Flow Data in Real-Time. Transportation Research Board Annual Meeting, Washington D.C., 2005.
14. Yang H.D, Sin B.K, Lee S.W. Automatic Pedestrian Detection and Tracking for Real-Time Video Surveillance. Lecture Notes in Computer Science 2003; 2688:1055.
15. Hoogendorn S, Daamen W, Bovy P. Extracting Microscopic Pedestrian Characteristics from Video Data: Results from Experimental Research into Pedestrian Walking Behavior. Transportation Research Board Annual Meeting, Washington D.C., 2003.
16. Chae K, Rouphail N. Empirical Study of Pedestrian-Vehicle Interactions in the Vicinity of Single-Lane Roundabouts. Transportation Research Board Annual Meeting, Washington D.C., 2008.
17. Kelly P, O'Conner N.E, Smeaton A.F. Robust Pedestrian Detection and Tracking in Crowded Scenes. In Image and Vision Computing, 2009; 27:1445-1458.

18. K. Ismail, T. Sayed and N. Saunier. Camera Calibration for Urban Traffic Scenes: Practical Issues and a Robust Approach. In Transportation Research Board Annual Meeting Compendium of Papers, Washington D.C., January 2010. 10-2715.
19. Saunier, N., El Husseini, A., Ismail, K., Morency, C., Auberlet, J., Sayed, T. (2011) Pedestrian Stride Frequency and Length Estimation in Outdoor Urban Environments using Video Sensors, in print, Transportation Research Record: Journal of the Transportation Research Board.