AUTOMATED COLLECTION OF PEDESTRIAN DATA USING COMPUTER VISION TECHNIQUES

Karim Ismail, M.A.Sc.

Research Assistant, Department of Civil Engineering
University of British Columbia
karim@civil.ubc.ca

Tarek Sayed, PhD, P.Eng.

Distinguished University Scholar
Professor of Civil Engineering
University of British Columbia
tsayed@civil.ubc.ca

Nicolas Saunier, PhD

Research Associate, Department of Civil Engineering
University of British Columbia
saunier@civil.ubc.ca

Word Count = 7020 words

[5020 words + 2 Tables + 6 figures, *1020 references*]

AUTOMATED COLLECTION OF PEDESTRIAN DATA USING COMPUTER VISION TECHNIQUES

ABSTRACT

Pedestrian data collection is critical for the planning and design of pedestrian facilities. Most pedestrian data collection efforts involve field observations or observer-based video analysis. These manual observations are time consuming, limited in coverage, resource intensive and error prone. Automated video analysis which involves the use of computer vision techniques can overcome many of these shortcomings. Despite advances in the field of computer vision applications for pedestrian detection and tracking, the technical literature shows little use of these techniques in pedestrian data collection practices. The likely reasons are the technical complexities that surround the processing of pedestrian videos. To extract pedestrian trajectories automatically from video, all road users must be detected, tracked at each frame and classified by type, at least as pedestrians and non-pedestrians. This is a challenging task in busy open outdoor urban environment. Common problems include global illumination variations, multiple object tracking and shadow handling. Specific problems arise when dealing with pedestrians because of their complex movement dynamics, varied appearance and non-rigid nature. The main objective of this study is to present a system for automated collection of pedestrian walking speed using computer vision techniques. The system is based on a previously developed feature-based tracking system for vehicles which was significantly modified to adapt to the particularities of pedestrian movement and to discriminate pedestrian and motorized traffic. The system was tested on real video data collected at Downtown area of Vancouver, British Columbia. This study is unique in so far as it tests the system under a variety of daylight conditions, crowd densities, movement context, and the video analysis approach. Promising results were obtained and several conclusions were drawn using statistical analysis of the automatically extracted pedestrian trajectories.

INTRODUCTION

Walking is the most basic means of traveling and is a main driver for a sustainable, healthy, clean, resource-efficient and livable urban environment. Therefore, new urban planning concepts have been redefining the function and mode-assignment of streets by emphasizing walkability as well as changing industry standards and professional practice in order to accommodate the pedestrian as a key road user (1). The emergence of the pedestrian as a key road user in an urban environment is an element in a larger theme that concerns the creation of a more sustainable transportation system. The revival of the theme is likely a public response to global changes in energy resources as well as a desire for improving the quality of life in urban areas.

Despite findings in the literature that corroborate the importance of non-motorized traffic and in particular pedestrians, these modes of transportation are in general overlooked, and understudied relative to vehicular traffic. For example, current trip counts capture 16-33% of actual non-motorized trips (2), while collecting reliable non-motorized traffic information remains challenging (3). Planning for pedestrian facilities and modeling of pedestrian demand are areas of research that are yet to be developed to a level that matches vehicular traffic (4).

Real data is critical for the development and calibration of design and planning models for pedestrian facilities. Many design applications involve individual (microscopic) observations of pedestrian movement. For example, microscopic observational data is required to investigate the ability of individual pedestrians to vary their walking speed based on a signal indication, potential conflict with motorized traffic (5) or in response to external stimuli (6). In addition, microscopic pedestrian observations can provide valuable insight for pedestrian modeling, e.g. inter-person spacing and pedestrian maneuvering (7) and obstacle navigation (8). Although at a relatively advanced stage in theory and analysis, pedestrian simulation models are generally based on limited understanding of microscopic pedestrian behavior (8) and limited validity that stems from real data (7) (9).

Collecting observational data for pedestrians is particularly challenging due to the less organized nature of pedestrian traffic compared to vehicular traffic (10). The main methods are: manual field observations, manual observations from videos, semi-automated video analysis, and automated video analysis. Manual field observation, which is the common method of pedestrian data collection, is in general more expensive, error-prone, and time consuming compared to video analysis (11). Generally, the use of video sensors has several advantages. First, it captures naturalistic pedestrian movement with limited risk of stirring the attention of observed subjects, who may behave unnaturally if felt being watched (12). Other advantages include the relative ease of installation, the richness of the data that can be extracted (i.e. complete trajectories), the large area that can be covered and their low cost. However, manual video observations are time consuming, resource intensive, and error-prone. Semi-automated analysis, or time-lapse analysis, of pedestrian movement involve the use of image processing tools to manually mark or track pedestrians in a sequence of video images, e.g. (13) (14). Manual operations in semi-automated video analysis are laborious and limited in terms of data volume that can be analyzed compared to automated methods. Automated video analysis which involves the use of computer vision techniques can overcome many of the shortcomings associated with manual field observations and manual video analysis.

The transportation literature contains few studies that involved applying computer vision techniques to collect pedestrian data in real settings, especially in busy "open" outdoor urban environment, such as areas around an intersection and transit hubs. Open environment refers to the mixed traffic, including motorized vehicles and pedestrians, the variable environment, the multiple flows of moving objects that may enter and leave the scene, and stop for varying amounts of time in the field of view. Automated pedestrian data collection in such environments remains a largely unsolved problem in the field of computer vision. Most published work is limited to idealized conditions using small datasets.

The primary objective of this study is to document the development and testing of a prototype system that is capable of extracting real-world pedestrian tracks from a video taken at traffic intersections. The study is unique in regard to the developed video analysis technique as well as in testing the developed system under different conditions of lighting, crowdedness, and traffic mix in an open and uncontrolled environment. The paper discusses the technical issues that arose during the system development are described along with techniques for resolving these difficulties. The walking speed automatically calculated by the system was validated in comparison to walking speeds extracted by human observers. The system accuracy in automatically measuring pedestrian speed was satisfactory and provided support and reliability for analysis results. A case study is introduced using video data collected for pedestrian movement in a main commercial corridor in the Downtown area of Vancouver, British Columbia. The case study was validated and demonstrated satisfactory accuracy of the system. The paper includes a statistical analysis of the case study results and reports the findings.

The next section reports a review of previous work on the subjects of pedestrian walking speed and pedestrian detection and tracking, followed by a description of the developed system for automatically collecting pedestrian walking speed data from video sequences. Following sections report the data collection effort, system testing, and validation results. The paper concludes with statistical analysis of the walking speed data obtained from the testing datasets and a summary of conclusions drawn from the entire study.

PREVIOUS WORK

Pedestrians Walking Speed

Walking speed is a fundamental characteristic of pedestrian flow that supports a wide range of theories and applications. The application contexts in transportation engineering that require an assumption regarding walking speed include planning and management of crowd movement, developing pedestrian simulation models, and designing pedestrian and traffic signals. The ability to predict pedestrian movement under different external circumstances and individual attributes of pedestrians is an important underpinning for the process of planning and design of pedestrian facilities (15). There are several contextual and individual variables that influence walking speed. Examples of studies that involved substantial walking speed observations are presented in Table 1. The Table also lists the variables that were considered to impact walking speed. For a good review of the evolution of walking speed refer to research cited in (16) (17).

As shown in Table 1, none of the key studies in the literature made use of automated pedestrian speed collection. Current methods used in practice to collect pedestrian data are also unable to capture microscopic changes in speed and position (18). This highlights the shortcomings of the current techniques used for pedestrian data collection and signifies the practical need for this research work.

Automated Pedestrian Data Collection

Automated pedestrian data collection relies mostly on video sensors, including visible spectrum, infrared (11) and thermal imaging cameras, as well as sometimes on Light Detection and Ranging (LIDAR) sensors (19). The work presented in this paper uses video cameras (in the visible spectrum) as alternative sensors are still more expensive, less widely available, and their resolution in space and time is typically more limited (11).

To extract pedestrian data automatically from video, all road users must be detected, tracked from one frame to the next and classified by type, at least as pedestrians and non-pedestrians. This is a challenging task in busy open outdoor urban environment as described earlier. Common problems for all environments are global illumination variations, multiple object tracking and shadow handling. Specific problems arise when dealing with pedestrians because of their complex movement dynamics, varied appearance and non-rigid nature. For a good survey of the challenges, the readers are referred to (20), although it is geared towards the study of human motion at a finer scale than this study requires. In (20), the different techniques for the detection and tracking of pedestrians are classified into:

- Tracking by detection: detection of objects is done using background modeling and subtraction with the current image (9) (11) (21) (22), or deformable templates, i.e. a model of image appearance using color distribution, edge characteristics, and texture. Image classifiers can be trained on labeled data to detect pedestrians (23). In many cases, especially if the objects are well separated, this approach works well.
- Tracking using flow: selecting good interest points, features, and matching them between successive images provides feature tracks that can be clustered into object trajectories. This approach is also called feature-based tracking and has been applied to traffic monitoring in (24) (25), and pedestrian counting in (26).
- Tracking with probability: it is convenient to see tracking as a probabilistic inference problem in a Bayesian tracking framework. In simple cases, independent Kalman filters can be run successfully for each target (Extended Kalman Filters are used for individuals and groups of pedestrians in (27), but will fail in scenes where the objects interact and occlude each other. This is called the data association problem and can be addressed using particle filters and Markov chain Monte Carlo methods for sampling.

Although great progress has been made in recent years, tracking performance are difficult to report and compare, especially when the systems are not publicly available, and when benchmarks are rare and not systematically used. Tracking pedestrian and mixed traffic in crowded scenes is still an open problem. Most vision-based pedestrian data collection took place in idealized conditions, e.g. heads and feet present all the time (10), low pedestrian volume (21) (22), or heavily controlled indoor experiments including markers on pedestrians (10) (11). The collected datasets are typically small and in some cases, require significant manual input to correct the automated results and to supplement with additional data (22).

PROTOTYPE SYSTEM DEVELOPMENT

The main stages of development are:

1. Define the system structure, function of each component, and the inter-component data exchange.
2. Implement and document each system component.
3. Find a set of detection and tracking parameters.

Figure 1 shows the structure of the prototype system. The following is a brief description and algorithm documentation of system components:

Camera Calibration

The main objective of camera calibration is to find a set of parameters that constitute a mapping from world coordinates to image plane coordinates, so that world coordinates can in turn be recovered from detection in images. The extrinsic parameters specify the translation and rotation of the camera coordinates relative to the world coordinates. The intrinsic parameters describe the perspective projection of the road scene onto the image plane. Both sets of parameters can be obtained by minimizing the difference between the projection of geometric entities, e.g. points and lines, onto world or image plane spaces and the actual measurements of these entities in projection space. A more efficient approach is to make use of the regularities (e.g. parallel lane markings, signal poles) abundant in traffic scenes to perform more informed calibration (28). The presence of these geometric primitives provides additional constraints as to the comparison between actual and projected entities. The mapping from homogeneous world coordinates **P** to homogeneous image plane coordinates **p** is as follows:

$$\mathbf{p} = \mathbf{A} \cdot [\mathbf{R} \,|\, \mathbf{t}] \cdot \mathbf{P} \tag{1}$$

where **A**, **R** and **t** are the intrinsic projection, rotation and translation matrices respectively. The intrinsic parameters considered in this study are focal lengths and skew angle. The mapping in Equation (1) imposes a reduction in dimensionality due to projecting on a plane. The inverse projection is defined only if one of the world coordinates, or a relationship thereof, is known. In the current application, image plane coordinates are re-projected onto the road surface, i.e. the plane Z=0. The world coordinates were obtained from an orthographic satellite image of the traffic scene obtained from Google Maps (29). The following generic objective function makes use of low-level features and geometric primitives:

$$\mathrm{F}(\mathbf{c}) = w_1 \sum_i \left\| \mathrm{p}_i - \mathrm{p}_i^c \right\| + w_2 \sum_i \left\| \mathbf{P}_i - \mathbf{P}_i^c \right\| + w_3 \sum_j \left| \mathbf{D}_j - \mathbf{D}_j^c \right| + w_4 \sum_k \left| \mathbf{N}_k - \mathbf{N}_k^c \right| \tag{2}$$

where

- **c** is the vector of all camera parameters,
- w_{1-4} are weight factors,
- p_i and p_i^c are the projected and measured image plane coordinates of calibration points respectively,
- $\mathbf{P}_i$ and $\mathbf{P}_i^c$ are the re-projected and actual world coordinates of calibration points respectively,

- $\mathbf{D}_j$ and $\mathbf{D}_j^c$ are projected and measured distances respectively,
- $\mathbf{N}_k$ and $\mathbf{N}_k^c$ are the calculated and actual angles respectively between pairs of calibration lines.

The weight factors are used to form an aggregate objective function. Based on trial and error, values used in this study are $w_{1-4} = [0.8 \quad 0.2s \quad s \quad 1]$, where s is the approximate number of pixels/meter in the image plane. Note that Equation 2 describes a generic non-coplanar calibration since world coordinates are in 3D. Also, angles between pairs of lines can capture conditions in which lines are parallel and perpendicular. The selection of the calibration entities should be well distributed over the camera field of view and of balanced densities. Local concentration of the calibration entities can possibly lead to the convergence to suboptimal set of parameters and degraded projection quality in areas in the field of view that are not covered with calibration entities.

The objective of camera calibration is to find the set of camera parameters **c** that minimizes the objective function $\mathrm{F}(\mathbf{c})$ described in Equation 2. Due to its good convergence rate compared to other algorithm available in the Matlab Optimization Toolbox, the optimization algorithm used was based on the Nelder-Mead simplex method. Standard methods in the literature for finding initial estimates depend on the extension of parallel lines in the image scene, e.g. lane marking, to find their vanishing point (30) (31) (28). In this study, the monitored traffic scenes were too limited in their field of view to observe a reasonable convergence of parallel lines toward a vanishing point. Initial estimates for the camera parameter, as was evidenced by numerous trials, are critical for an optimal solution to be found. An initial estimate for the camera position was obtained using an approximate position for the camera set-up location and the rotation angles using an orthographic satellite image.

The calibration accuracy obtained by applying the previous procedure to a Vancouver intersection (as will be described later in the case study) was satisfactory. The average percentage error in linear measurements was 4%. Figure 2 shows the projection of a sample of pedestrian tracks on an orthographic satellite image of the scene. Similar studies in the literature used artificial construction of an orthographic image using video image rectification e.g. (32). The approach followed in this study by projecting the video data on an independent site map proved helpful in visually verifying the accuracy of projection - especially with the difficulties faced in obtaining calibration data. In addition, it was possible to collate pedestrian tracks obtained from different camera settings into a single site map, whereas video image rectification produces a setting-dependent site map.

Video Formatting

Depending on the video source, it may be necessary to encode the video in a suitable format for later processing, as well as correct recording artifacts such as interlacing.

Feature Tracking and Feature Grouping

A feature-based tracking system was initially developed for vehicle detection and tracking as part of a larger system for automated road safety analysis (25)(33). Feature-based tracking is preferred because it can handle partial occlusion. Tracking features is done through the well known Kanade-Lucas-Tomasi feature tracker. Stationary features and features with unrealistic motion are filtered out, and new features are generated to track objects entering the field of view. Since a moving object can have multiple features, the next step is to group the features, i.e. deciding what set of features belongs to the same object, using cues like spatial proximity and common motion. The grouping method described in (34) was extended to handle intersections (25). A graph connecting features is constructed over time. Two parameters are crucial for the success of the method: the connection distance $D_{connection}$, i.e. the maximum distance between two features for their connection, and the segmentation distance $D_{segmentation}$, i.e. the maximum difference between the minimum and maximum distance between two features. The tracking accuracy for motor vehicles has been measured between 84.7% and 94.4% on three different sets of sequences (25). This means that most trajectories are detected by the system, although over-grouping and over-segmentation can still occur.

High-level Object Processing

Difficulties occur in scenes where the traffic is mixed and the road users have very different sizes, e.g. passenger cars and pedestrians, and the connection and segmentation distances can only be adjusted for one type of road user. To address this issue, the original system has been extended by obtaining the type of the road users. The parameters are set for pedestrians, and consequently the cars are over-segmented. Once the groups of features belonging to cars are identified, the feature are processed a second time by the grouping algorithm using larger connection and segmentation distances.

In the current system, a simple test on the maximum speed reached of road users is sufficient to discriminate between pedestrians and motorized road users in most cases. This will be improved in the future by using object classifiers based on background subtraction and image appearance (23).

System Operator and User

The point of an automated system is to minimize user input, especially to eliminate the need for continuous supervising. Global optimization methods to adjust parameters are still lacking, as performance is difficult to evaluate completely automatically. The role of the system operator is therefore to find good parameter values by trial and error, and visual inspection of the results. Since the world coordinates are recovered, the parameters can be used unchanged in various scenes. The system was developed in an open manner in order to provide data for analysis and visualization purposes. The results are currently stored in plain text files, but could be as well stored in a database, and can be mined for the needs of the end user.

CASE STUDY

This section describes the analysis of video sequences collected from an open busy environment, in the Vancouver Downtown area. The objective of this analysis is to test the ability of the system to correctly measure the walking speed of pedestrians in a variety of settings. The validation study adopted the following steps:

- Select an intersection on a main commercial corridor in Vancouver, British Columbia with a nearby camera setting location. The intersection should contain a variety of pedestrian facilities. Also, the location should be on the main course of crowd movement outbound of a concurrent event in order to test the system.
- Record high-definition video data for the intersection in day- and night-time conditions.
- Select a random sample that represents 10% of the detected and tracked pedestrians (individuals or groups).

- Calculate the average walking speed by measuring the time the elapses during observing the crossing between two check lines, e.g. road marking.
- Compare the system-based and observer-based walking speeds.

Videos were collected for pedestrian movement at a traffic intersection on Robson St. which is a major commercial and business corridor in Vancouver Downtown area with active walking environment. A total of seven footages were recorded from 8:00 PM till 12:00 PM in order to capture normal night-time pedestrian movement as well as crowd movement to and from a fireworks event that took place in the same time. The timing of the video survey was intended to be concurrent with the fireworks event in order to capture higher pedestrian volumes and to provide walking speed information for local transportation authorities in order to assist in predicting outbound crowd movement in future events.

The camera was set on the 29th floor of a high-rise building that overlooks that intersection. Figure 2 shows a video image and an orthographic satellite image of the intersection along with real-world tracks of pedestrian movement as obtained using the video analysis system.

The recorded video sequences covered a wide variety of observation conditions that often exist in pedestrian facilities. Various pedestrian density conditions were monitored, ranging from crosswalks with low pedestrian volumes to concentrated crowd movement. Videos were collected in day- and night-time conditions. Pedestrian movement was monitored at sidewalks, crosswalks, and along a thoroughfare that was closed for motorized traffic.

Data Analysis

The implementation of this camera calibration procedure faced two obstacles: first, the road surface in the monitored intersection was recently repainted, thus leaving a handful of common features on both the orthographic satellite image and the video images. Second, it was not possible to conduct a lab-based camera calibration in order to find all the intrinsic camera parameters apart from the focal length. The first obstacle was addressed by collecting linear field observations of the true length of a total of 15 entities that appeared in the video images. The calibration process was mainly guided by the linear component of the objective function. The second obstacle required that all camera parameters be estimated based on information collected from the traffic scene. This increased the processing time required for the convergence criterion to be met. Accurate camera parameters were required in this study since the magnitude of error in speed estimate that results from position estimate can be significant at low speeds. This obstacle was addressed by following the previous camera calibration procedure.

Tracks shown in Figure 2(b) depict the movement of individual pedestrians as well as groups of pedestrians. Tracked objects, i.e. individuals and groups that reached a speed higher than a specific threshold, 3.5 m/s, were classified as motorized traffic and filtered out. Pedestrian tracks are clustered using the K-means algorithm. Each track is represented by a four-dimensional vector, each element being the average movement orientation over a section of the track. The first and last sections cover 20% of the entire duration during which the pedestrian object existed, starting from both ends. The two intermediate sections were selected at one third of each pedestrian track with a length of 10% of the track duration. This selection of several clustering variables is necessary to capture turning pedestrian movement through the intersection. The number of clusters was selected based on visual observation of the prevalent streams of pedestrian movement in each video record. The four trajectory clusters that appear in

Figure 2(b) are: pedestrians moving East-West (1), pedestrians moving West-East (2), pedestrian crossing movement (3) and Vehicles (4).

Night-time footage was the most challenging to analyze due to the poor visibility of pedestrians in dim corners of the intersection. A specific set of feature tracker parameters has to be used to recover more feature data. As shown in Figure 3, the results obtained are generally satisfactory. Data however could not be recovered from low-light areas. In addition, dark-clothed pedestrians were difficult to detect without rendering the integration of large volume of uninformative and low-quality features.

Walking speed data was collected at user-defined registration areas for each tracked object that falls in a specific movement cluster. The definition of a registration area is necessary for gathering walking speed data in desirable specific spatial context. Since walking speed varies during the time a tracked object was present within the registration area, the average walking speed within this duration was recorded. Figure 4 shows the registration area defined for the indicated crosswalk. Registration areas were defined for other pedestrian facilities (two sidewalks, two unmarked crosswalks, and another marked crosswalk) in order to gather walking speed data. Summary of walking speed statistics are presented in Table 2. Figures 5(a) and 5(b) show sample distributions of pedestrian walking speed for crossing and sidewalk movements respectively.

Validation

Validation of object detection and tracking is generally poorly studied in the literature. There is an absence of a standard evaluation method that follows a systematic approach and is based on a public testing database. Hence, the validation process in this study is limited to walking speed measurements. Average walking speed for a 10% random sample drawn from tracked pedestrian objects was compared to manual video observation of the walking speed. Walking speed was manually calculated based on the time required by moving objects to traverse the shortest distance between two check lines. The check lines were selected to be the road markings of the crosswalk across Robson St. Figures 6 (a) and (b) show a comparison between measured and automatically calculated walking speeds. There is an excellent agreement between manual and automated walking speed values (RMSE = 0.0725 m/s and 0.0548 m/s). The residual errors can be attributed to inaccuracy of manual speed calculation in which the pedestrians are unrealistically assumed to follow the shortest path between two check lines, inaccuracy in camera calibration, and irregularities in pedestrian tracks due to noise in feature detection.

Discussion of Results

The case study was intended to monitor pedestrian movement under several conditions. The monitored pedestrian facilities are crosswalk, sidewalks, and unmarked crosswalks. Data was also collected for crowd movement during a road closure and is presented in Table 2. Pedestrians moving from West to East had to walk up a 5% longitudinal grade. The average walking speed for all pedestrian objects is 1.217 m/s and the average and 15th percentile crossing speed is 1.315 and 0.93 m/s respectively. This value is consistent with studies in the literature as shown in Table 1. There is a statistically significant ($p < 0.05$) difference between walking speed at crosswalks and at sidewalks, walking uphill (from West to East) and opposite direction. There is no statistically significant ($p = 0.0616$) difference between walking speed along marked and unmarked crosswalks. However this result is deemed as inconclusive since it was measurably close to statistical significance. There is a statistically significant difference between West-East walking speed at night during a road closure and at day time along the sidewalks. This is likely due to the larger space afforded for pedestrians during a road closure as well as the leisurely nature of walking back from a night event.

As discussed before, one of the major advantages of video-based data collection is to capture walking speed variability. It was observed that pedestrians walked faster along unmarked crosswalks in case of approaching vehicles. The variability in crossing speed, quantified by the standard deviation of speed measurements over the time interval within a registration area, was recorded for movements along marked and unmarked crosswalks. There is a statistically significant ($p < 0.0001$) higher variability of walking speed at unmarked crosswalks compared to marked crosswalks.

CONCLUSIONS

Pedestrian walking speed has been the subject of continuous research. There has been a recent revival in pedestrian studies that is motivated in part by demographic changes. It is believed that future data collection is necessary to develop a better understanding of pedestrian movement and the factors that influence walking speed.

The majority of commercial techniques developed for automatically collecting traffic data focus on vehicular traffic. The technological aspects of automated pedestrian data collection are generally more involved than vehicular traffic. The majority of walking speed studies in the literature does not make use of automated video analysis for collecting pedestrian data. In this study, an automated system for collecting pedestrian walking speed using video analysis was developed and tested. A system previously developed for vehicle detection and tracking was significantly modified to adapt for particularities of pedestrian movement and to discriminate pedestrian and motorized traffic. The system was tested on real video data collected at Downtown area of Vancouver, British Columbia, during day- and night-time conditions. It was found that pedestrians walk faster at marked crosswalks than sidewalks. Walking speed was more variable at unmarked crosswalks compared to marked crosswalks. Gradient and lighting conditions were identified as statistically significant variables that affect walking speed.

Several conclusions can be drawn from this research work. First, the accuracy of walking speed calculations was sensitive to camera calibration parameters. Several challenges were faced during the recovery of the camera parameters due to site-specific conditions. A robust camera calibration technique was developed and reported in this study. Second, night-time conditions proved to be the most difficult as expected because of the obscurity of pedestrian outlines and video recording noise. A special set of detection parameters was used for night videos and results obtained are satisfactory. Third, there is a lack of a systematic procedure for evaluating video analysis techniques. There is also no public benchmark to compare various techniques. Finally, the literature of pedestrian observational studies is yet to benefit from automated video analysis techniques. It is expected that the system presented in this study will be further improved by adding other appearance-based techniques.

ACKNOWLEDGEMENTS
The data collection was made possible by the support of the following parties: The City of Vancouver and The Empire Landmark Hotel and Conference Centre. The authors are indebted to Clark Lim for the continuous help during the data collection process.

REFERENCES

1. *Toward a New Urbanist Transportation Agenda.* **Greenberg, Ellen.** 2005.

2. **Litman, Todd.** *Non-Motorized Transportation Demand Management, Sustainable Transport: Planning for Walking and Cycling in Urban Environments.* [ed.] Rodney Tolley. s.l. : Woodhead Publishing Ltd., 2003.

3. *How Much Do Americans Walk? An Analysis Of The 2001 NHTS.* **Weinstein, Asha and Schimek, Paul.** Washington, D.C. : s.n., 2005. Transportation Research Board 84th Annual Meeting.

4. *Assessment of Models to Measure Pedestrian Activity at Signalized Intersections.* **Pulugurtha, Srinivas Subrahmanyam and Repaka, Sudha R.** s.l. : Transportation Research Board, 2008.

5. **Stollof, Edward R., McGee, Hugh and Eccles, Kimberly A.** *Pedestrian Signal Safety for Older Persons.* AAA Foundation for Traffic Safety. 2007.

6. *Walking in Waikiki: Measuring Impact of Street Performers on Pedestrian Level of Service in Urban Resort District.* **Kim, Karl, et al.** 2007, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1982, pp. 104-112.

7. *Collecting, Processing and Calculating Pedestrian Flow Data in Real-time.* **Kerridge, J. and Chamberlain, T.** 2005.

8. *Human movement behaviour in urban spaces: implications for the design and modelling of effective pedestrian environments.* **Willis, Alexandra, et al.** 2004, Environment and Planning B: Planning and Design, Vol. 31, pp. 805–828.

9. *Behavioral Priors for Detection and Tracking of Pedestrians in Video Sequences.* **Antonini, Gianluca, et al.** 2006, International Journal of Computer Vision, Vol. 69, pp. 159-180.

10. *Extracting microscopic pedestrian characteristics from video data: results from experimental research into pedestrian walking behavior.* **Hoogendorn, Serge P., Daamen, W. and Bovy, P.H.L.** 2003. Transportation Research Board Annual Meeting.

11. *Using Low-Cost Infrared Detectors to Monitor Movement of Pedestrians: Initial Findings.* **Kerridge, J., et al.** 2004, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1878, pp. 11-18.

12. **Bechtel, R.** *Human movement in architecture, Environmental Psychology.* s.l. : Rinehart \& Winston; New York, 1970.

13. *Pedestrian Flow Characteristics in Hong Kong.* **Lam, W.H.K., Morrall, J. F. and Ho, H.** 1995, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1487, pp. 56-62.

14. *A speed-concentration relation for bi-directional crowd movements with strong interaction.* **AlGhadi, S., Mahmassani, H.S. and Herman, R.** s.l. : Springer, Berlin, Germany, 2002. pp. 3-20.

15. *Modeling Pedestrian Walking Speeds on Sidewalks.* **Al-Azzawi, Marwan and Raeside, Robert.** 2007, Journal of Urban Planning and Development, Vol. 133, pp. 211-219.

16. *The Continuing Evolution of Pedestrian Walking Speed Assumptions.* **LaPlante, John N. and Kaeser, Thomas P.** 2004, Institute of Transportation Engineers. ITE Journal, Vol. 74, pp. 32-40.

17. **Fitzpatrick, Kay, et al.** *TCRP Report 112/NCHRP Report 562. Improving Pedestrian Safety at Unsignalized Crossings.* Transportation Research Board. 2006.

18. *Research on Pedestrian Behavior and Traffic Characteristics at Unsignalized Midblock Crosswalk.* **Shi, Jiangang, et al.** 2007, Transportation Research Record: Journal of the Transportation Research Board, Vol. 2038, pp. 23-33.

19. *Fusion of Detection and Matching Based Approaches for Laser Based Multiple People Tracking.* **Cui, Jinshi, Zhao, Huijing and Shibasaki, R.** 2006. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol. 1, pp. 642-649.

20. *Computational Studies of Human Motion: Part 1, Tracking and Motion Synthesis.* **Forsyth, D.A., et al.** 2005, Foundations and Trends in Computer Graphics and Vision, Vol. 1, pp. 77-254.

21. *Video-Based Monitoring of Pedestrian Movements at Signalized Intersections.* **Malinovskiy, Yegor, Wu, Yao-Jan and Wang, Yinhai.** 2008.

22. *Empirical Study of Pedestrian-Vehicle Interactions in the Vicinity of Single-Lane Roundabouts.* **Chae, K. and Rouphail, N. M.** 2008. Transportation Research Board Annual Meeting Compendium of Papers. 08-2898.

23. *Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors.* **Wu, B. and Nevatia, R.** s.l. : Kluwer Academic Publishers, 2007, International Journal of Computer Vision, Vol. 75, pp. 247-266.

24. *Vehicle Segmentation and Tracking in the Presence of Occlusions.* **Kanhere, N. K., Birchfield, S. T. and Sarasua, W. A.** 2006. Transportation Research Board Annual Meeting Compendium of Papers.

25. *A feature-based tracking algorithm for vehicles in intersections.* **Saunier, N. and Sayed, T.** s.l. : IEEE, 2006.

26. *Counting Crowded Moving Objects.* **Rabaud, V. and Belongie, S.** 2006. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. Vol. 1, pp. 705-711.

27. *Estimating pedestrian counts in groups.* **Kilambi, P., et al.** s.l. : Elsevier Science Inc., 2008, Computer Vision and Image Understanding, Vol. 110, pp. 43-59.

28. *Using Geometric Primitives to Calibrate Traffic Scenes.* **Masoud, Osama and Papanikolopoulos, Nikolaos P.** 2007. Transportation Research Part C: Emerging Technologies.

29. *Google Maps.* s.l. : Google.

30. *Dynamic Camera Calibration of Roadside Traffic Management Cameras for Vehicle Speed Estimation.* **Schoepflin, Todd N. and Dailey, Daniel J.** 2003. IEEE Transactions on Intelligent Transportation Systems. Vol. 4.

31. *Efficient method for camera calibration in traffic scenes.* **Zhaoxue, C. and Pengfei, S.** 2004, Electronic Letters, Vol. 40.

32. *Automated video analysis as a tool for analysing road user behaviour.* **Laureshyn, Aliaksei and Ardö, Håkan.** 2006. ITS World Congress.

33. *Automated Road Safety Analysis Using Video Data.* **Saunier, N. and Sayed, T.** s.l. : Transportation Research Record: Journal of the Transportation Research Board, 2007, Vol. 2019, pp. 57-64.

34. *A Real-time Computer Vision System for Measuring Traffic Parameters.* **Beymer, D., et al.** s.l. : IEEE Computer Society, 1997. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. pp. 495-501.

35. **Dahlstedt, S.** *Walking speed and walking habits of elderly people.* National Swedish Road and Traffic Research Institute. 1978.

36. *Another Look at Pedestrian Walking Speed.* **Fitzpatrick, Kay, Brewer, Marcus A. and Turner, Shawn M.** 2006, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1982, pp. 21-29.

37. *Give Elderly Pedestrians More Time To Cross Intersections.* **Guerrier, Jose H. and C., Jr. Sylvan.** 1998.

38. *Recommended walking speeds for timing of pedestrian clearance intervals based on characteristics of the pedestrian population.* **Gates, Tim J., et al.** 2006, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1982, pp. 38-47.

39. *Pedestrian Walking Speed, Step Size, and Step Frequency from the Perspective of Gender and Age: Case Study in Beijing, China.* **Hui, Xiong, et al.** 2007. Source Data: Transportation Research Board Annual Meeting 2007 Paper No. 07-1486.

40. *Field studies of pedestrian walking speed and start-up time.* **Knoblauch, R. L., Pietrucha, M T. and Nitzburg, M.** 1996, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1538, pp. 27-38.

41. *Pedestrian Speed Flow Relationships for Walking Facilities in Hong Kong.* **Lam, W.H.K. and Cheung, C.** 2000, Journal of Transportation Engineering, Vol. 126, pp. 343-349.

42. *Variation of Walking Speeds on a Unidirectional Walkway and on a Bidirectional Stairway.* **Lee, Jodie and Lam, William.** 2006, Transportation Research Record: Journal of the Transportation Research Board, Vol. 1982, pp. 122-131.

43. *Pedestrians' Normal Walking Speed and Speed When Crossing a Street.* **Montufar, Jeannette, Michelle, Jorge Arango and Nakagawa, Satoru.** 2007, Transportation Research Record: Journal of the Transportation Research Board, Vol. 2002, pp. 90-97.

44. *Walking Behavior and Pedestrian Flow Characteristics for Different Types of Walking Facilities.* **Ye, Jianhong, et al.** 2008. Source Data: Transportation Research Board Annual Meeting 2007 Paper No. 08-1991.

45. *Pedestrian Cross Flow Characteristics and Performance.* **Khisty, C. J.** 1985, Environment and Behaviour, Vol. 17, pp. 679-695.

46. *The effects of density, sex and group size upon pedestrian walking velocity.* **Boles, W.** 1981, Man – Environment Systems, Vol. 11, pp. 37-40.

LIST OF TABLES

LIST OF FIGURES

TABLE 1 **Sample Walking Speed Studies with Reported Observation Methods and Factors Affecting Walking Speed**

Study	Reported 15th Percentile Walking Speed	Reported 50th Percentile Walking Speed	% difference from standards [1]	Number of subjects	Method	Significant Factors [2]	Insignificant Factors
Dahlstedt(35)	0.67	-	-26%	N/A	1[3]	1[4]	-
Fitzpatric et al.(36)	0.9	-	0%	2552	2	1	5,8,6,2
Guerrier et al.(37)	0.66	-	-27%	263	2	1	-
Gates et al. (38)	0.92	-	2%	1947	1,2	1,5,6	2
Hui et al.(39)	-	1.22	-6%	1882	2	1,2	-
Knoblauch et al.(40)	0.97	-	8%	7123	1	1,3	2,4-8
Lam and Cheung(41)	Model	-	N/A	16453	3	4,6,9,10,11	-
Lam and Morrall(13)	Model	-	N/A	N/A	2	4,6,9,11	-
Lee and Lam(42)	Model	-	N/A	14886	3	4,11	-
Montufar et al.(43)	0.88	-	-2%	1792	1	1,3,4	-
Stolloff et al.(5)	1.03-1.16	-	-64%	2603	1,2	1	-
Ye et al.(44)	Model	-	N/A	2089	2	11	-

1 We refer to the most recent recommended updates for MUTCD as standards (1.3 m/s average and 0.9 m/s 15th percentile)

2 Significance is statistical and/or practical. The assessment of the practical significance of walking speed factors was either directly reported in the studies or performed by the authors of this study. Insignificant factors were treated ins similar manner.

3 Number indications: *1)* Field observations*, 2)* Manual video analysis*, 3)* Semi-automated video analysis*,*

4 Number indications: *1)* Age and/or walking problems*, 2)* Gender*, 3)* Season /weather (precipitation, snow, temperature)*, 4)* Pedestrian facility type (Crosswalk, sidewalk, stairway, midblock crossing, experiment setting)*, 5)* Group size*, 6)* Traffic control (Pedestrian signal type, unsignalized, speed limit)*, 7)* site specifications (Marking, geometry, road classification, median, lane usage)*, 8)* Vehicular traffic*, 9)* Indoor/outdoor*, 10)* Activity area (Shopping, commercial, recreational, etc.)*, 11)* Pedestrian traffic characteristics (flow, density, directional split).

TABLE 2 Summary of Walking Speed Statistics

Movement	No. Pedestrian objects	Average (m/s)	Stan. Dev. (m/s)	P-value (difference in means between column and row movement types)		
				East-West UCW	East-West SW	West-East UCW & SW
East-West UCW[1]	907	1.406	0.262	-	-	<0.0001
East-West SW[2]	1148	1.0436	0.2797	-	-	
West-East UCW	289	1.2627	0.3031	<0.0001	-	-
West-East SW	44	0.9657	0.2365	-	0.0333	
MCW[3]	162	1.315	0.3722	0.0002	-	0.0069
Night-time	656	1.1316	0.2061	-	-	<0.0001

1 UCW: unmarked crosswalk 2 SW: sidewalk 3 MCW: marked crosswalk

FIGURE 1 Layout of system components

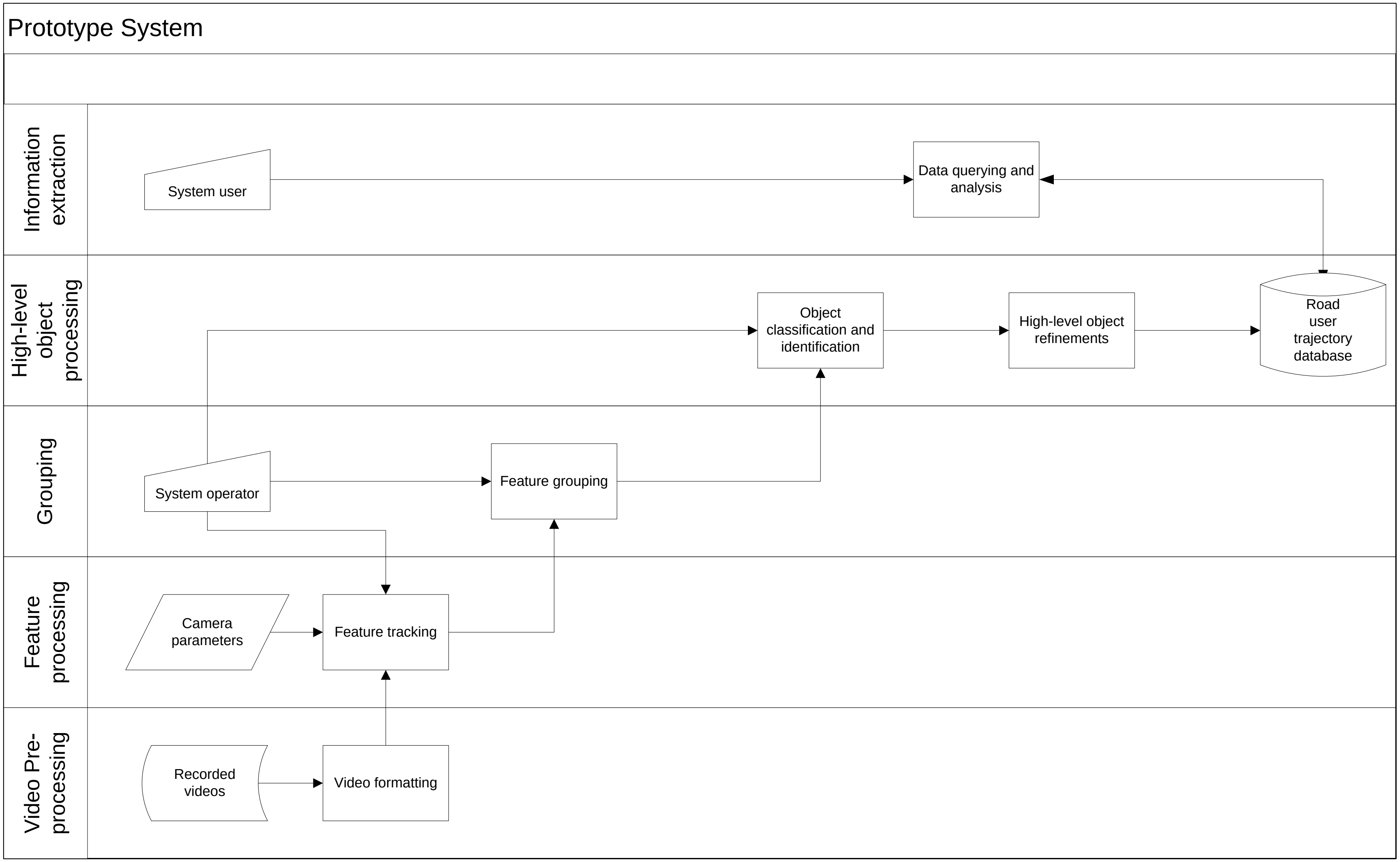

FIGURE 2 Pedestrian tracks at Site 1. Left figure shows tracks in the image plane. Right figure shows the same tracks projected on an orthographic image. The trajectories are classified by object type (vehicles or pedestrians) and direction. Clusters 1 to 3 are for pedestrians moving East-West, West-East and Crossing respectively, while cluster 4 is for vehicles.

Figure 2(a)

Figure 2(b)

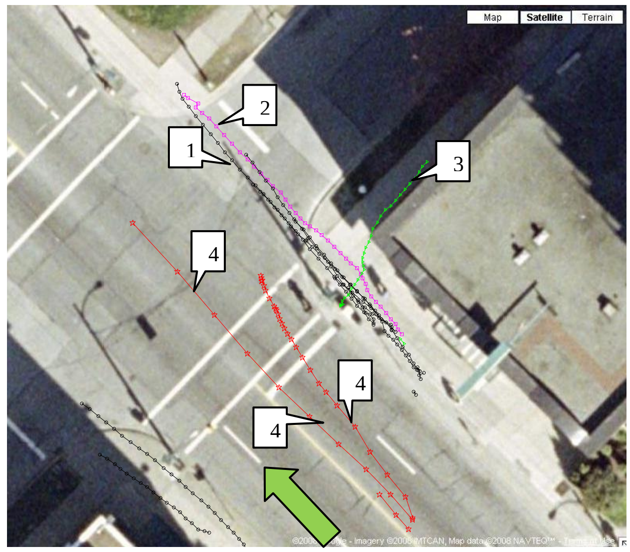

FIGURE 3 A sample frame from night-time video analysis. Displayed are red bounding boxes around pedestrian objects and walking speed.

FIGURE 4 the figure shows pedestrian trajectories that crossed through the marked data collection area. Trajectories are collated and projected to the world image from different videos with different fields of view and hence may be truncated in different regions.

FIGURE 5 (a) walking speed distribution for pedestrians moving through the data collection area shown in Figure 4 across Robson St. (b) walking speed distribution for pedestrians moving from East to West through corresponding data collection areas on both sidewalks of Robson St.

Figure 5(a)

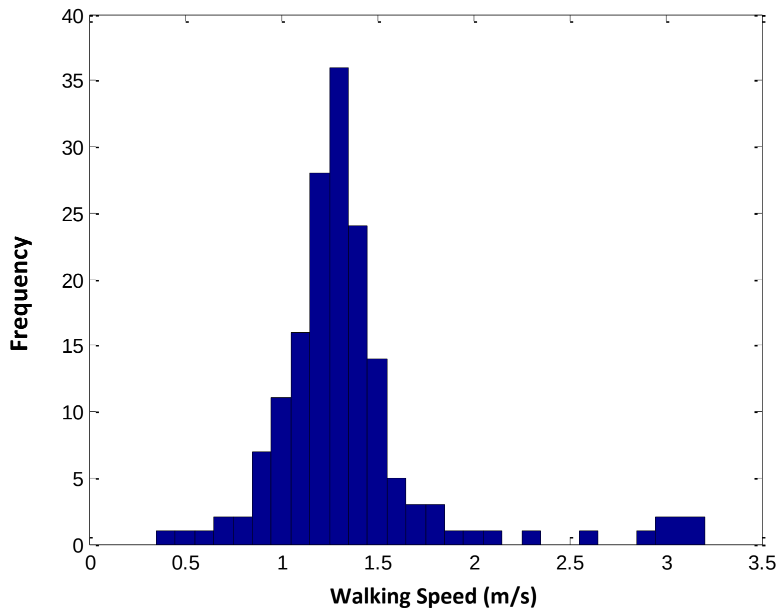

Figure 5(b)

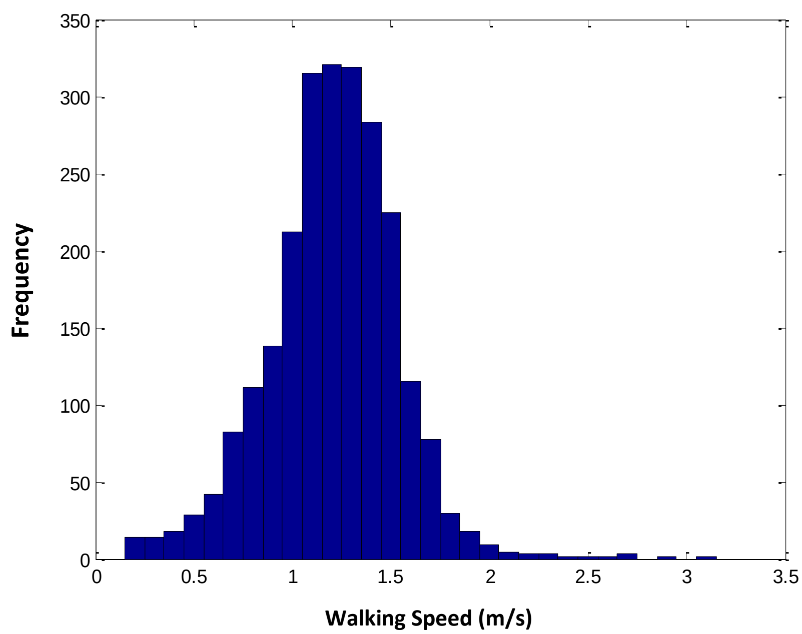

FIGURE 6 (a) Validation of walking speed measurements. Horizontal axis depicts walking speed based on the time interval required to walk between two check lines. Vertical axis depicts the average walking speed within the same time interval based on automated pedestrian tracking. (b) Validation of walking speed measurements of night-time conditions.

Figure 6(a)

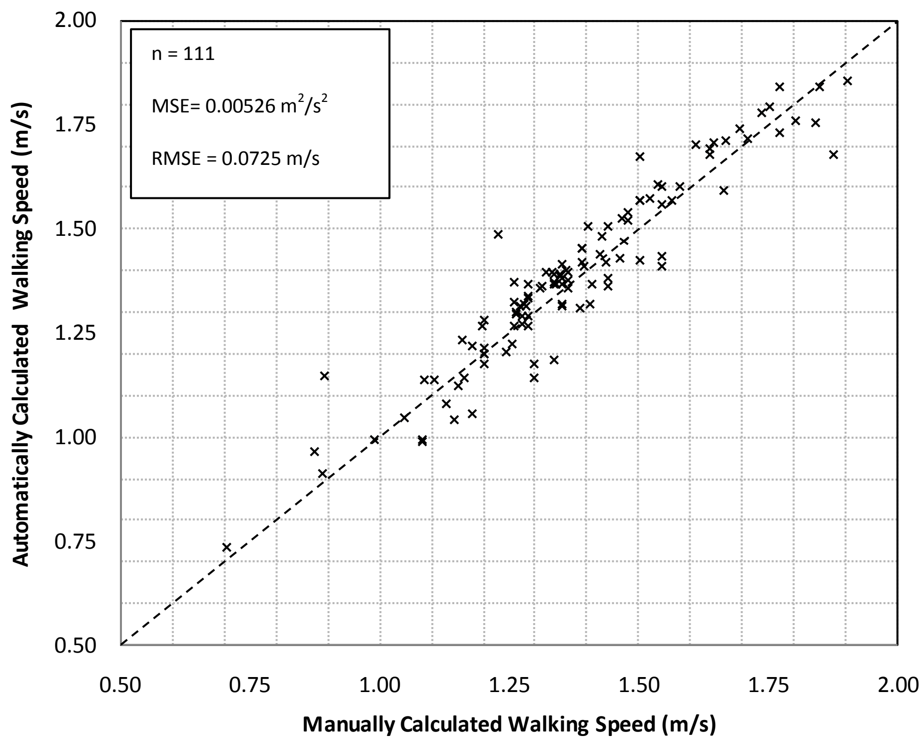

Figure 6(b)