Motion periodicity-based pedestrian detection and particle filter-based pedestrian tracking using stereo vision camera

Khalid Al-Mutib*, Muhammad Emaduddin, Mansour AlSulaiman and Hedjar Ramdane

Department of Computer Engineering,
College of Computer Science and Information,
King Saud University,
P.O. Box 51178, Kingdom of Saudi Arabia
E-mail: muteb@ksu.edu.sa
E-mail: memaduddin@ksu.edu.sa
E-mail: msuliman@ksu.edu.sa
E-mail: hedjar@ksu.edu.sa
*Corresponding author

Ebrahim Mattar

Electrical and Electronics Engineering Department,
College of Engineering,
University of Bahrain,
P.O. Box 13184, Kingdom of Bahrain
E-mail: ebmattar@uob.edu.bh

Abstract: A methodology that detects harmonic motions of limbs and body during a typical human walk is presented. It temporally propagates the position, stride, direction and phase using a particle filter. This is based on a human limb-motion model, and is able to track the walking pedestrians in a heavily occluded environment. Potential 3D point clusters belonging to arms and feet are extracted employing an adapted version of RANSAC based surface detection algorithm. The periodicity feature is established via a Fourier-transform based periodogram that confirms the walk periodicity for each point-cluster representing limbs. RGB or intensity data from the stereo-vision input is completely ignored and the proposed method completely relies upon 3D data produced by the stereo-vision sensor. This independence from light-based information, produces reliable illumination invariant pedestrian detection and tracking results in outdoor environment using Daimler stereo pedestrian detection dataset.

Keywords: robotics; pedestrian detection and tracking; NARF feature-based pedestrian tracking; stereo-vision; particle filter; gait periodicity analysis.

Reference to this paper should be made as follows: Al-Mutib, K., Emaduddin, M., AlSulaiman, M., Ramdane, H. and Mattar, E. (2014) 'Motion periodicity-based pedestrian detection and particle filter-based pedestrian tracking using stereo vision camera', *Int. J. Computer Applications in Technology*, Vol. 50, Nos. 1/2, pp.113–121.

Biographical notes: Khalid Al-Mutib is an Assistant Professor in Intelligent Systems and Robotics at the Department of Software Engineering College of Computer and Information Sciences, King Saud University, Kingdom of Saudi Arabia. He received his PhD from University of Reading, UK in 1997. Also, he received his MSc in Electrical and Computer Engineering in 1985 from University of Kansas, USA. Currently, his primary research interest lies in mobile robotics systems.

Muhammad Emaduddin received his Masters degree in Intelligent Robotics from the University of Southern California in 2008. Currently, he is working as a Research Engineer at King Saud University. His primary research interest lies in computer vision-based SLAM. He has also worked extensively on intelligent pedestrian detection and tracking techniques.

Mansour AlSulaiman is an Associate Professor at the Department of Computer Engineering, College of Computer and Information Sciences at King Saud University, Kingdom of Saudi Arabia. His research interest includes self organising feature maps for Arabic phonemes, architectures for neuro-fuzzy controllers. In addition, he is interested in Arabic speech recognition using neural networks.

Hedjar Ramdane received his BSc and PhD from USTHB University and MSc from BLIDA University in Algiers in Electronics and Electrical Engineering in 1988, 1992 and 2002, respectively. He is an Associate Professor at King Saud University. His research interests include predictive and intelligent control of non-linear systems with specialisation in machines and robotics.

Ebrahim Mattar is an Associate Professor of Intelligent Control and Robotics at University of Bahrain, Bahrain. He received his BSc in Electrical Engineering (from University of Bahrain in 1986), he studied MSc in Electronics in 1989 (from University of Southampton, UK), and in 1994 he received from University of Reading his PhD in Cybernetics and Robotics Control. His research interest includes in computational intelligence, robotics control, modelling and control.

This paper is a revised and expanded version of a paper entitled 'Motion periodicity based pedestrian detection and particle filter based pedestrian tracking using stereo vision camera' presented at the 2012 19th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), Auckland, New Zealand, 28–30 November 2012.

1 Introduction

1.1 Introducing pedestrian detection theme

Pedestrian detection through visual observations, is an active research area primarily due to its key role in deployment of autonomous ground vehicles in urban environments. By proposing this novel pedestrian detection method researchers strive to add to the ability of urban robots to operate safely in the presence of moving objects such as pedestrians and other vehicles. With the recent advent of inexpensive depth sensors which operate at real-time frame rates such as Microsoft Kinect, rich point cloud data is now available to the researchers for a wide array of applications including pedestrian detection. A typical human walk or running motion is periodic. It is proposed use the periodic motion of human limbs as the primary cue for pedestrian detection problem. Since human walking motion is not strictly regular, therefore we introduce limited randomness within human walk model used by our proposed tracker. In this paper, we attempt to devise a pedestrian detection and tracking method that is able to process input point clouds from RGB+D devices such as Microsoft Kinect as well as stereovision cameras. A Point Grey Bumblebee stereo vision camera is used as primary input sensor for our proposed method due to its relatively high accuracy and performance profile. The stereo camera gathers point cloud using (640 × 480) resolution images @15FPS in wide baseline configuration. The point cloud from the stereo vision sensor is processed using an already devised recursive surface detection algorithm (Emaduddin et al., 2012). This algorithm chooses multiple regions of interest (ROI) within the point cloud that are later on processed for validation as pedestrians. The validation step utilises mean-shift clustering approach to detect clusters for limbs and then a temporal correlation analysis step to validate detection of a pedestrian. Another alternative approach was to extract NARF features from the point cloud and perform periodicity validation upon the points representing these features. This approach proved to be vital for reducing the overall computational load but at the cost of detection accuracy. Furthermore pedestrians are tracked via their walk periodicity, stride and speed features using an adapted particle filter proposed by Emaduddin and Shell (2009). The experiments indicate that even pedestrians with concealed limbs translate into 3D points that exhibit periodicity. Although the clusters of such points are not very large, yet we were able to reasonably detect such pedestrians.

1.2 Surveying the literature

Among the most closely related work is that of Vinicius and Borges (2011), who claim to detect blobs representing human limbs and body parts that exhibit

1 a cyclic behaviour in the blob trajectory (signifying the periodicity in human hand and feet motion during a walk)

2 a correlation between blob size and vertical position of the blob (signifying the mild change in height that occurs while taking a step).

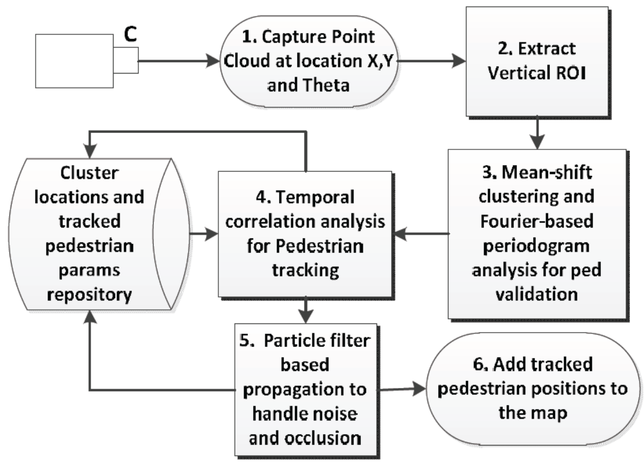
We have used the motion periodicity feature outlined by case 1 and 2 for pedestrian validation. The adopted approach differs with the approach used in Vinicius and Borges (2011) though as this work detects positional periodic changes of blob within 2D image space and periodic colour intensity changes for each pixel associated with a blob. Our approach instead uses 3D point data to detect '3D blobs' and detect periodicity in 3D space. Furthermore, the presented approach is additionally robust to false positives encountered in Vinicius and Borges (2011). This is due to the fact that, we do not rely upon changes in intensity value of pixels. NARF features are also employed by suggested algorithm in order to track points belonging to pedestrian limbs. Since NARF features are typically used in object recognition (Steder et al., 2010), these are fairly robust to camera pose change and illumination changes. Use of NARF features also eliminates the need of additional sensor calibration for cameras having separate dedicated sensors for RGB and distance values (such as Microsoft Kinect). Other classical approaches used to detect periodicity in the pedestrian walk include phase-locked loop and autoregressive moving average models (ARMA). These approaches are discussed by Ran et al., (2007). In this work frequencies of sinusoidal signals generated by pedestrian motion are detected using the afore-mentioned classical approaches.

Howard et al. (2007) suggested an intelligent filter to select ROI, i.e., potential pedestrians via detection of vertical surfaces within 3D point-clouds. This ROI extraction approach is utilised within our proposed method but Howard et al. approach differs from our approach in terms of pedestrian validation since it uses derived feature-based Bayes classifier in order to detect pedestrians. Nedevschi et al. (2009) calculates 3D optical flow vectors for 3D points provided by a stereo system. The periodicity in the change of direction for these optical vectors is calculated after some pre-processing. This periodicity is then used to establish whether the subject is walking pedestrian or not. Although Nedevschi et al. provides a relatively practical approach in terms of detection of pedestrians using 3D point-cloud, it lacks cluster detection step and a generic walk model that can be used to predict pedestrian periodic motion and location. Many other methods such as Soga et al. (2005) and Li et al. (2010) use artificial neural network (ANN) and support vector machine-based (SVM) classifiers for detection of pedestrian body parts using shape-based classification. These approaches work reliably under most circumstances but fail due to heavy occlusions and lack of comprehensive training sets. 3D shape classifiers also fail due to a huge number of poses involved in training sets along with the variety involved in human posture, clothing and body structure.

1.3 Manuscript contribution

The proposed method utilises a much richer set of point-cloud data than any of aforementioned techniques while keeping the method real-time and independent of any requirement such as NN or SVM-based classifier training. We used mobile robot programming toolkit (MRPT), point cloud library (PCL) and robot operating system (ROS) to implement various modules of the method. The system follows a client-server architecture where Pioneer-3 Powerbot mobile research platform acts as a client and an Intel Corei7-based, GPU mounted high-end Machine serves as a server. The client application onboard the robot consists of a stereo camera driver layer, a point-cloud compression module, a robot localisation module and a network communication module. The server application consists of all modules except module 1 as depicted in Figure 1. We have employed parallel processing using both CPU and GPU threads over the server machine. Module 2 is assigned a separate GPU thread while module 3 and 4 are assigned to another CPU thread. Module 5 and 6 execute on a separate CPU thread as well.

Figure 1 Stereo vision-based pedestrian detection and tracker method – flowchart



1.4 Manuscript organisation

After such a brief introduction about the previous work conducted over pedestrian detection using stereo camera in Section 1, Section 2 will illustrate the system by describing ROI extraction, the pedestrian validation and tracking steps. In Section 3, we shall present the results of conducted experiments and the rate of detection and false positives over few walk sequences from daimler stereo pedestrian detection dataset. Future work and concluding remarks will be presented in Section 5.

2 System architecture description and ROI extraction

2.1 Point-cloud capture

The client application executing onboard the robot, constantly captures the 3D point-cloud from the stereo camera, compresses the data and transmits the point-cloud to the server. Point-cloud is compressed using Octree-based point-cloud compression technique (Schnabel and Klein, 2006) in order to minimise the observation transmission time and keep the method viable for real-time processing. It is assumed that the robot speed is considerably slow as it takes on an average over 500 ms for the data to be captured and transmitted over the network. It must be mentioned here that the compression module also clips off the unwanted points from the point cloud, i.e., any point ranging above seven feet is clipped since this is a sound upper limit for maximum human height. The Bumblebee camera pose, $Pose_{BB}$ is associated with each point cloud observation. This enables our method to map each detected obstacle including pedestrians to the any vertex $U_{x,y}$ of the global grid-map G as the mobile robot navigates through the environment.

2.2 Extract vertical ROI

The point cloud received from the client application is fed into a well-tested accurate plane fitting technique detailed in Emaduddin et al. (2012). An adaptation needs to be made into the standard plane detection algorithm listed in Emaduddin et al. (2012). The algorithm requires a normal to floor surface □□□□ in order to detect the floor surface as illustrated by the following expression.

$$(\hat{n}, P') = \text{RANSAC\_planefitting}(P', \alpha, e, \hat{n}) \tag{1}$$

Here P' on LHS represents the output point-cloud and on RHS represents the input point cloud. e represents the acceptable error in the Euler angles of the plane whose normal is $\hat{n}$. α represents the distance threshold for a 3D point to be part of the plane. We can assign new values to normal $\hat{n}$ that are orthogonal to floor normal and larger angular values for e in order to detect walls or any other vertical planar surfaces placed on floor. Thus, after several executions of equation (1), when no more walls or vertical planar surfaces can be detected within the environment, the resultant point-cloud P' will contain only non-planar sub-point-clouds that are good candidates for pedestrian detection. The plane fitting technique also removes noisy points from the input point-cloud, hence we do not need to explicitly perform this step.

2.3 Mean-shift clustering and Fourier-based periodogram analysis

This step performs cluster extraction and motion-based analysis of clusters belonging to a pedestrian. The points from point-cloud P' are clustered using mean shift clustering algorithm. The module requires the size of cluster parameter at this point which is empirically determined to be equivalent to $S = 0.008$ cubic meter. Selecting clusters of this size produces the most optimal results in tracking step. Cluster centres are stored into a time-stamped set with k detected clusters $C_t = \{c_1, \ldots, c_n\}$. Here $c_i = \{x_i, y_i, z_i\}$, where (x-y) plane represents the (2D) stretch of the point-cloud map and z-coordinate represents the height of obstacle points within the point-cloud map. The (x, y and z) coordinates of each cluster are then analysed for periodicity independently using the assumptions followed by Ran et al. (2007).

The following set of equations, are the evaluation expressions used in the process.

$$X_{c_i}(\omega) = \frac{2}{N}\left\|\sum_{t=0}^{N-1} x_{c_i} e^{\omega t}\right\| \tag{2}$$

$$Y_{c_i}(\omega) = \frac{2}{N}\left\|\sum_{t=0}^{N-1} y_{c_i} e^{\omega t}\right\| \tag{3}$$

$$Z_{c_i}(\omega) = \frac{2}{N}\left\|\sum_{t=0}^{N-1} z_{c_i} e^{\omega t}\right\| \tag{4}$$

In the previous set of equations, the term (ω) represents the given frequency. Furthermore, $(x_{c_i}, y_{c_i}, z_{c_i})$ refer to the (x, y and z) coordinates of the detected cluster c_i for which periodicity analysis is required. In addition, t represents the observation index from ($t = 0, \ldots, k - 1, k, k + 1 \ldots, n - 1$). In reference to Theodoridis and Koutroumbas (2009), and as per the guidelines suggested to us by Theodoridis and Koutroumbas (2009), centre of mass (COM) for the periodograms was calculated as suggested by expressions in equations (2) to (4). The following formula outlines the procedure to calculate COM for each one the Fourier-transform-based periodograms.

$$COM\_X_k = \frac{\sum_{m=0}^{M} m\left|x_{c_i}(m)\right|}{\sum_{m=0}^{M}\left|x_{c_i}(m)\right|} \tag{5}$$

$$COM\_Y_k = \frac{\sum_{p=0}^{P} p\left|y_{c_i}(p)\right|}{\sum_{p=0}^{P}\left|y_{c_i}(p)\right|} \tag{6}$$

$$COM\_Z_k = \frac{\sum_{q=0}^{Q} q\left|z_{c_i}(q)\right|}{\sum_{q=0}^{Q}\left|z_{c_i}(q)\right|} \tag{7}$$

In the above set of equations, the indices (m, p and q) denote the frequencies detected by expression equations (2) to (4). k represents the total number of observations taken into the account in order to calculate the periodogram. As per the relevant literature (Emaduddin and Shell, 2009; Vinicius and Borges, 2011; Ran et al., 2006), the typical values of COM range between (1 Hz) for slow walking to 3 Hz for running pedestrians. Thus in order to establish for a cluster c_i to be associated to a walking pedestrian at least any two of the COM values yielded by equations (5) to (7) should render values between (1 Hz) and (3 Hz). As an additional experiment the mean-shift Clustering step is repeated for only NARF feature points instead of all the points belonging to an ROI. Detected clusters C_t now have their detected centres c_i representing the mean location of all NARF feature points belonging to a same cluster.

2.4 Temporal correlation analysis

As a result of last step we retrieve a set of clusters $C_t' = \{c_1, \ldots, c_l\}$ associated with pedestrians. In order to reduce the computational burden on the server, all 3D points associated with the set of clusters C_t', are projected on detected floor plane. This transformation leaves us with 2D points associated to 2D clusters whose centres ($c_i.xyloc$) are re-calculated after the 3D to 2D projection. The next step is to iterate through the set C_t' in order to associate a sub-set of set C_t' to a particular pedestrian. The routine of iteration continues to be repeated until no un-associated clusters are left within the set C_t'. The criteria based upon which this association is decided, states that each pedestrian can be associated with a maximum of nine clusters. This number was suggested by the superior detection and tracking performance during our experiments. Thus for the *first run* a maximum of nine closest 2D clusters, the centres of which lie within a range of 0.2 m are associated to a single pedestrian. We thus populate a set of pedestrian locations at time $t = 0$, referred to as $Ped_t = \{p_1, \ldots, p_n\}$. Here $p_i = \{c_0, \ldots, c_j\}$, where $j \le 9$. The mean location of a pedestrian is calculated at $t = 0$ by taking the mean of the centres of associated clusters.

For the subsequent observations, i.e., ($t \ge 1$), the pedestrian identity is maintained using the pedestrian location from the tracker module. Thus in this case we associate the identity of a tracked pedestrian p_i with a cluster c_j, wherever the following constraint is true:

$$
\begin{aligned}
&if\left(dist\left(p_i.xyloc \in Ped_{t-1}, c_j.xyloc \in C_t'\right) \le 0.4m\right)\\
&AND\left(cardinality\left(Ped_t\right) < 9\right)\\
&\left\{p_i = p_i \cup \left\{c_j\right\};\ \text{where } p_i \in Ped_t, c_j \in C_j'\right\};
\end{aligned}
$$

2.5 Particle filter-based pedestrian tracking

The tracker module estimates the parameters of motion and location attached with a pedestrian based on updates from temporal correlation analysis step. This module uses particle filter technique to estimate the position, stride, direction and phase of a pedestrian walk similar to a particle filter proposed in Shao et al. (2007). It was important to make insignificant modifications in the proposed walk model documented by Shao et al. (2007), in order to tune it to match motion of arms as well as feet. This led us to initialise 200 particles per pedestrian for smooth and efficient tracking. The tracker functions via three traditional tracking steps:

1 *Update step*: The tracker associates weights to each of the pedestrian's particles. These weights are proportional to the distance between each particle's position and the centre of the cluster associated to the particular pedestrian.

2 *Sampling step*: The tracker now samples the weighted particles where the likelihood of any particle to be chosen is proportional to its weight. Thus a certain predefined number of particles Q are chosen.

3 *Propagation step*: In this step, the sampled Q particles are propagated through a multidimensional space representing the motion of the tracked pedestrian similar to the walk model described in detail in Shao et al. (2007).

The propagation step thus updates the pedestrian position ($p_i.xyloc$), stride ($p_i.s$), direction ($p_i.\theta$) and the limb motion phase ($p_i.\O$) of a pedestrian. The propagation step is carried out regardless of whether a pedestrian has any associated clusters present in last observation update or not. This continuous propagation of pedestrians based upon limb motion model helps the tracker to track partially or completely occluded pedestrians. We declare the tracker to be lost for a pedestrian if the tracker does not find any update for over 30 seconds for that particular pedestrian.

2.6 Updating tracked pedestrian locations on a map

The pedestrian locations are rotated and translated as per the camera pose received form the client application. After necessary transformations the points belonging to clusters associated with each pedestrian are mapped onto grid map. The grid map can be used to perform grid-based path planning employing methods such as A* or D*. Additional safety features can be added in terms of limiting robot velocity and extra space padding between the robot and pedestrian location.

3 Experimentation and results

At this stage we shall verify the adopted algorithm. In this sense, we gathered stereo point cloud sequences through the stereo-vision camera mounted on top of a mobile robotic system research platform, refer to Figure 2. Such sequences were gathered while the mobile robot was in motion and robot/camera pose was bundled with each observation. Observations were processed online by the proposed method and in real-time.

Figure 2 Sensor mounted POWERBOT used for experimentation at KSU (see online version for colours)

For method verification and evaluation purposes, it was decided to rely upon Daimler stereo pedestrian detection dataset. We selected multiple walk sequences from the dataset that focused on distinct problems faced by modern pedestrian detection and tracking systems. The detection rate and false positives results are highlighted in Table 1. Figure 3 illustrates the limb and head point clouds detected during sequence capturing activity.

Table 1 Detection rate and false positive results for seven scenarios from Daimler stereo pedestrian detection dataset

Scenario	*Results*		
	Total peds	*Detected peds*	*False positives*
Single pedestrian, walking across, near cam, no noise, no occlusion.	01	01	00
Single pedestrian, walking towards cam, no noise, no occlusion.	01	01	00
Multiple peds, walking across, walking towards cam far away, with occlusion, no noise	02	02	00
Multiple peds, walking together, far away, with occlusion, bicycle riders	03	04	01
Multiple peds, walking in all directions, far away, with and without occlusion, tree trunks moving	10	07	00
Multiple peds, walking in all directions, both near and far away, with and without occlusion, bicycle riders, camera shaking	08	06	01
Single pedestrian, walking across, near cam, totally occluded except head, no noise	01	01	00

Note: The results are categorised into various grades of scenario complexity.

Figure 3 3D point clouds of clusters belonging to a single pedestrian, exhibiting harmonic-motion during a walk sequence (see online version for colours)

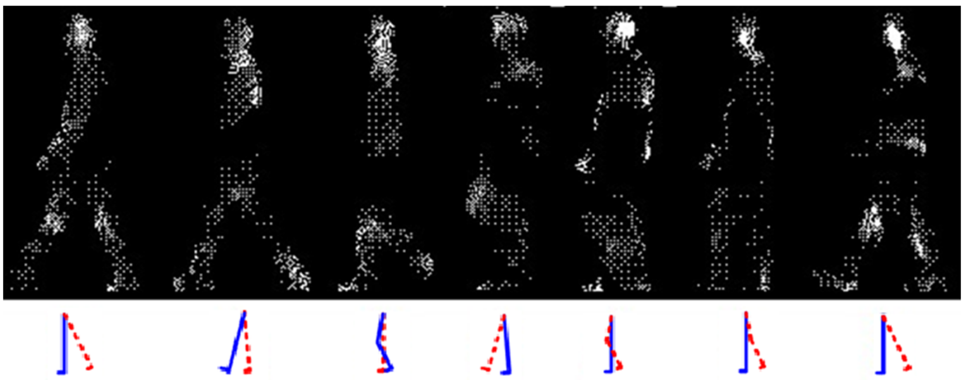


Figure 4 Images from Daimler stereo detection dataset and corresponding point clusters with harmonic motion (see online version for colours)

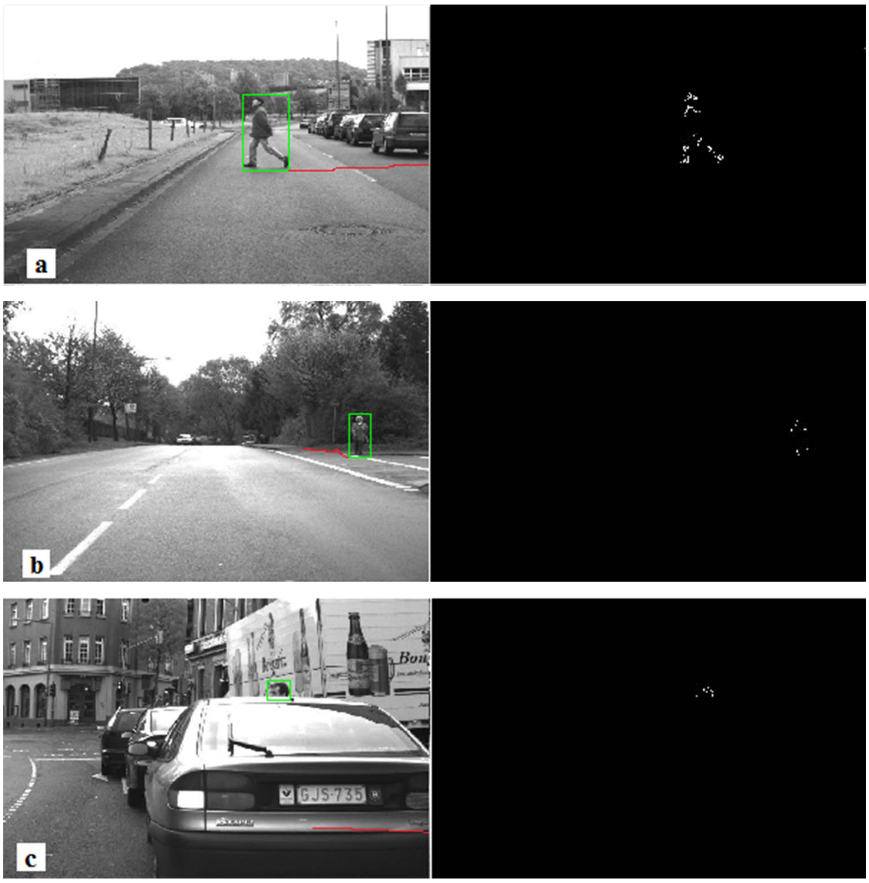
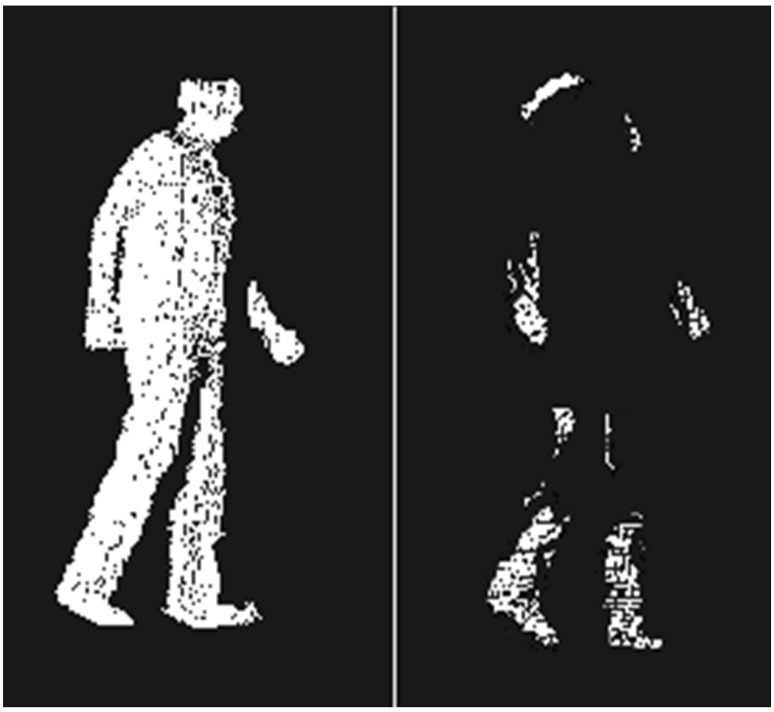


Figure 5 Pedestrian point cloud segmentation output for clusters with harmonic motion (last frame of a four seconds walk sequence)



Apart from dataset results, laboratory experiments also confirm the soundness of Fourier-based periodogram analysis for detection of periodic motion following clusters. In this sense, Figure 5 shows the pedestrian point cloud and final output clusters generated by the proposed method. No noise was present during the laboratory experiments. Additionally, the experiments was involving a single pedestrian. Another sequence of output clusters generated by the algorithm is shown in Figure 4. Here a single walking pedestrian is observed during various stages of a typical walk. Clusters exhibiting periodicity for different frames during a walk are shown in this figure. It is worth mentioning here that these clusters were generated by the system after observing the pedestrian for at least eight seconds. Less observation time produces fewer clusters for a typical pedestrian walk.

4 Discussions

One of the major observations noted during our testing on Daimler stereo pedestrian detection dataset was that various noise elements caused false positives for example swinging tree trunks, bicycle riders pedalling their way through the scene and entire point cloud harmonic motion due to jerks encountered while driving. Many of the noise factors are obviously easy to rectify but pedalling bicycle rider noise demands more accurate walking model to be incorporated into our method since the frequency of cycling sometimes falls within the range of 1 Hz to 3 Hz.

Considering results of image ‘c’ within Figure 3, the harmonic motion of pedestrian head was exploited to its maximum. The fact that no arms or legs or body is visible throughout a (three second) sequence and the method was able to detect and track the pedestrian for at least (two seconds), highlights the utility of motion-based detection itself. Thus it is believed that although motion-based detection methods may not be applied to standing pedestrians but output of these methods can be fused with other shape-based detection methods in order to increase detection confidence in troublesome scenarios.

Table 2 highlights the real-time execution times for two implementations of proposed algorithm. The process was implemented while using an INTEL CORE i7 machine running Windows 7 Professional (32 bit) for testing purposes. The modules were compiled using Microsoft Visual Studio 2010 release version.

The per frame execution times for both implementations show a marked improvement when it comes to considering only the NARF feature points for periodicity validation instead of using all points belonging to a limb cluster. This improvement comes at a cost of roughly (13%) decrease in pedestrian detection accuracy. Since NARF features may not be detected for points belonging to certain limbs, there exists a greater chance of missing out a complete pedestrian limb. On a positive note whenever NARF features are discovered for a certain limb, these exhibit well pronounced periodicity.

5 Conclusions

In this manuscript, a motion periodicity-based pedestrian detection and tracking method was suggested. The technique uses a robotic mounted stereo vision camera as its primary sensor. The foremost emphasis of the work in this paper is the motion periodicity detection with a 3D point cloud. Also a walk model described in Shao et al. (2007) was shown to successfully propagate limb motion via a particle filter. By successfully testing this method over sequences from Daimler stereo pedestrian detection dataset and achieving an average of 85% detection rate for walking pedestrians, the method contends to be a viable solution for pedestrian detection using stereo vision sensor. The technique can additionally improve its tracking performance by introducing an additional layer of SVM-based classifier in order to learn and classify periodic walk motion more accurately. Image intensity-based blob-tracking results can also be fused with the currently deployed particle filter-based tracking results to cater for situations with highly occluded pedestrians.

Table 2 Per frame average execution time for multiple scenarios

Scenario	*Average execution time per frame*	
	Considering all points belonging to a cluster	*Considering only the NARF feature points*
Single pedestrian, walking across, near cam, no noise, no occlusion	371 ms	88 ms
Single pedestrian, walking towards cam, no noise, no occlusion	437 ms	65 ms
Multiple peds, walking across, walking towards cam far away, with occlusion, no noise	851 ms	110 ms
Multiple peds, walking together, far away, with occlusion, bicycle riders	922 ms	139 ms
Multiple peds, walking in all directions, far away, with and without occlusion, tree trunks moving	1,073 ms	229 ms
Multiple peds, walking in all directions, both near and far away, with and without occlusion, bicycle riders, camera shaking	1,082 ms	238 ms
Single pedestrian, walking across, near cam, totally occluded except head, no noise	137 ms	29 ms

Acknowledgements

This work is supported by NPST programme by King Saud University (Project No. 08-ELE300-02).

References

Emaduddin, M. and Shell, D. (2009) 'Estimation of pedestrian Distribution in Indoor environments using multiple pedestrian tracking', Paper presented at the *Workshop on People Detection and Tracking, IEEE International Conference on Robotics and Automation (ICRA)*, Kobe, Japan.

Emaduddin, M., AlMutib, K., AlSulaiman, M., Hedjar, R. and Mattar, E. (2012) 'Accurate floor detection and segmentation for indoor navigation using RGB+D and stereo cameras', Paper presented at the *International Conference on Image Processing, Computer Vision, & Pattern Recognition*, Las Vegas, USA.

Howard, A., Matthies, L., Huertas, A., Bajracharya, M. and Rankin, A. (2007) 'Detecting pedestrians with stereo vision: safe operation of autonomous ground vehicles in dynamic environments', Paper presented at the *13th International Symposium of Robotics Research*.

Li, H., Toulminet, G. and Nashashibi, F. (2010) *An Improved Flies Method for Stereo Vision: Application to Pedestrian Detection*, IMARA – INRIA Open Archive.

Nedevschi, S., Bota, S. and Tomiuc, C. (2009) 'Stereo-based pedestrian detection for collision-avoidance applications', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 10, No. 3, pp.380–391.

Ran, Y., Weiss, I., Zheng, Q. and Davis, L. (2006) 'Pedestrian detection via periodic motion analysis', *International Journal of Computer Vision*, Vol. 71, No. 2, pp.143–160.

Ran, Y., Weiss, I., Zheng, Q. and Davis, L. (2007) 'Pedestrian detection via periodic motion analysis', *International Journal of Computer Vision*, Vol. 71, No. 2, pp.143–160.

Schnabel, R. and Klein, R. (2006) 'Octree-based point-cloud compression', Paper presented at the *Eurographics Symposium on Point-Based Graphics*, pp.111–120, The Eurographics Association.

Shao, X., Zhao, H., Nakamura, K., Katabira, K. and Shibasaki, R. (2007) 'Detection and tracking of multiple pedestrians by using laser range scanners' Paper presented at the *IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 29–November 2, Vol. 1, No. 1, pp.2174–2179, Tokyo.

Soga, M., Kato, T., Ohta, M. and Ninomiya, Y. (2005) 'Pedestrian detection with stereo vision', Paper presented at the *Data Engineering Workshops, 21st International Conference*, 05–08 April, Vol. 1, No. 1, p.1200, Toyota Central R&D Labs.

Steder, B., Rusu, B., Konolige, K. and Burgard, W. (2010) 'NARF: 3D range image features for object recognition', Paper presented at the *Workshop on Defining and Solving Realistic Perception Problems in Personal Robotics at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, October 18–22, Vol. 44, Taiwan.

Theodoridis, S. and Koutroumbas, K. (2009) *Pattern Recognition*, China Machine Press, Elsevier, USA, ISBN-13: 978-1597492720.

Vinicius, P. and Borges, K. (2011) 'Blob motion statistics for pedestrian detection', Paper presented at the *International Conference on Digital Image Computing: Techniques and Applications*, pp.442–447, December, Noosa, Australia.