2018 IEEE Intelligent Vehicles Symposium (IV)
Changshu, Suzhou, China, June 26-30, 2018

Pedestrian Dynamic and Kinematic Information Obtained from Vision Sensors

Santiago Gerling Konrad, Mao Shan, Favio R. Masson, Stewart Worrall and Eduardo Nebot

Abstract— **The estimation and prediction of pedestrian motion is of fundamental importance in ITS applications. Most existing solutions have utilized a particular type of sensor for perception such as cameras (stereo, monocular, infrared) or other modalities such as a laser range finder or radar. The advent of wearable devices with inertial sensors have led to the development of systems capable of the robust inference of pedestrian intention. Unfortunately, these devices do not have communications capabilities to broadcast this information to all vehicles in proximity, and also this strategy requires functioning devices on all pedestrians to work. This paper presents a robust perception method that is able to extract dynamic pedestrian information with accuracy comparable to that of typical gyroscopes and accelerometers installed in wearable devices. Experimental results are presented to demonstrate the potential for obtaining very comprehensive dynamic information from limbs representing the skeleton of a pedestrian. This work also demonstrates the accuracy of vision based systems by comparing these results to the rotation and acceleration measured directly on the pedestrian using a wearable device. The contributions of this paper demonstrate that it is possible to significantly improve both the detection and estimation of pedestrian intention by incorporating dynamic information obtained from vision sensors.**

I. Introduction

During the last few years, there has been significant progress in technology designed to prevent accidents between pedestrians and vehicles. Many vehicle manufacturers are introducing new technology to alert drivers if pedestrians are detected within the projected trajectory of the vehicle [1]–[3]. There are two fundamental problems that need to be addressed. The first problem is the robust detection and discrimination of people within a complex urban environment. The second problem is to estimate pedestrian behavior, which can be then used to make a prediction of the intended trajectory of the pedestrian. There are currently several outstanding contributions to address the first problem using different sensor modalities and sensor arrangements. The most successful applications in terms of commercial deployment are based on vision technology using monocular [4] and stereo [5] cameras. More recently, laser range and bearing sensors have demonstrated a high level of reliability to detect and discriminate pedestrians in urban scenarios [6], [7]. There have also been demonstrations of radar technology to detect pedestrians [8] with some level of success.

S. Gerling Konrad and F. R. Masson are with the Instituto de Investigaciones en Ingeniería Eléctrica, Departamento de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur (UNS)-CONICET, Bahía Blanca, Argentina.

M. Shan, S. Worrall and E. Nebot are with the Australian Centre for Field Robotics, The University of Sydney, Australia.

With respect to the second problem, there is significant research that is focused on detecting the current type of activity or behavior of a pedestrian. This information can be incorporated into detailed predictive models for estimating pedestrian intention if the information and models are of a sufficiently high quality. A particularly important type of pedestrian behavior to estimate is the intent to cross the street. This information is vital to reduce false alarms generated by collision avoidance systems.

Perception has been previously used to infer intention with limited success. This is due to the fact that a single image cannot capture dynamic information. New approaches to pedestrian safety are also exploring the use of inertial sensing to estimate behaviors. The inertial information usually consists of gyroscope rates and acceleration of the part of the body where the sensors are fixed. This information is becoming more available due to the widespread introduction of wearable devices such as mobile phones, smart watches and fitness devices. These devices can provide high quality, high frequency dynamic information that is required to reliably estimate pedestrian intention and behaviour. Unfortunately, the compatible technology to broadcast this information to vehicles and infrastructure is not yet widely available. In addition, this strategy requires all pedestrians to be fitted with equipment that is functional and compatible with the vehicle systems.

More recently, a number of vision based algorithms have been demonstrated to reliably detect pedestrians and to estimate a complete 3D pose in the form of a skeleton representation. This paper demonstrates that these vision based systems are capable of obtaining a dynamic representation of a pedestrian with information of similar quality to the inertial sensors used in wearable devices. This can enable the implementation of robust pedestrian intention algorithms based on vision information.

This work makes use of the *OpenPose* library developed by [9]–[11]. This tool is used to derive skeleton representations of pedestrians using a sequence of images. The library uses particular models to extract 18 points of the body including shoulder, elbow, wrist, hip, knee, ankle, etc., as shown in Figure 1a. This algorithm has been demonstrated to perform robustly even with a crowd of pedestrians in a single image, which is an essential requirement for pedestrian tracking in an urban environment.

The points of the skeleton are provided frame by frame in the video. By sampling images at the appropriate frequency, it is also possible to evaluate the dynamic information of the main relevant segments in the skeleton, i.e., the pedestrian

978-1-5386-4451-5/18/$31.00 ©2018 IEEE

limbs. We propose to use this information for the evaluation of rotation rates and accelerations of the main relevant limbs necessary for estimating pedestrian intention. We validate the dynamic information using data collected from a wearable device containing an inertial measurement unit (IMU) that is attached to particular limbs of the pedestrians. We then demonstrate that we can obtain dynamic information of all limbs using collected naturalistic pedestrian data, showing that this approach is very appealing for the development of accurate pedestrian prediction models.

The rest of the paper is organized as follows. A review of the literature is presented in Section II. Section III describes the experimental setup and vehicle platforms used to collect the data. The results are presented in Section IV. The first part of Section IV presents the validation of the proposed vision based approach using a comparison with inertial information obtained from the gyroscope and accelerometer of a wearable device attached to the pedestrian. The second part shows the dynamic information obtained from other limbs. Finally, Section V presents the conclusion and summarizes the main contributions of this paper.

II. Background

In this paper we are exploring the use of vehicle vision systems for the estimation of high quality pedestrian dynamic and kinematic information. In particular, we are interesting in pedestrian pose tracking that can be used in the robust estimation of pedestrian intent.

A. Pedestrian Pose Tracking

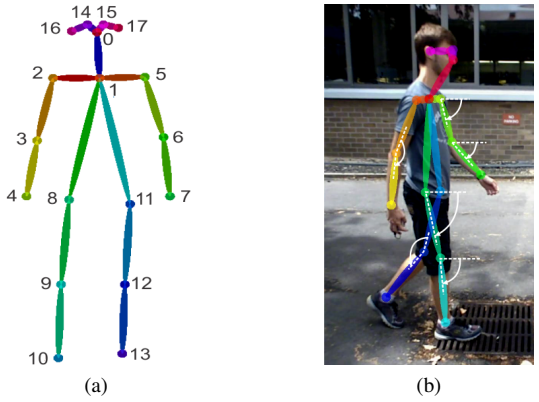
Obtaining an accurate pedestrian pose from an image is a challenging task. Ye *et al.* [12] propose a method to estimate the underlying pedestrian skeleton captured by a depth camera and match the pedestrian against a database. This database consists of high resolution point clouds of human models in various poses.

Quintero *et al.* [13] propose a method that exploits the general anthropometry of a standing human subject, photographed from a stereo camera. A top-down hierarchical search of pedestrian joints based on anthropometric measurements is done in order to estimate the pedestrian skeleton. A sample scoring algorithm allows the determination of skeleton segments and joints to be estimated. The score to be computed is adapted for each body part.

A novel approach for monocular 3D human pose estimation and tracking in realistic street conditions is presented by Andriluka *et al.* [14]. Their solution exploits tracking methods in a number of consecutive frames from a monocular camera, and matches it against possible pose solutions.

More recently, references [9]–[11] present very impressive results extracting skeleton representations of people from images. Their algorithms can detect and compose a number of skeletons in the same scene.

There has also been significant progress in the detection of pedestrians in an image through the use of semantic labeling. Current algorithms can operate in real time and classify each pixel of an image as belonging to a particular class including pedestrian. Zhou *et al.* [15] present experimental results adapting an existing model to a new environment by re-training a neural network with their own dataset describing new features of the local environment.



(a) (b)

Fig. 1: An example of the skeleton representation obtained using the *OpenPose* library. (a) shows points of the skeleton. (b) illustrates the skeleton and the corresponding angles.

B. Pedestrian Intention Estimation

Pedestrian limb tracking is probably the most relevant information used in the inference of pedestrian intention. Each pedestrian can perform one of a variety of actions in an urban environment, the most common of which are standing and walking. These two states, combined with other contextual and sequential information, could be used to obtain a prediction model of pedestrian intention.

Some authors have used cameras and lidars to detect pedestrian states [16] [17]. Others, for instance [18]–[20], use sensors such as accelerometers, gyroscopes, magnetometers and GPS that are attached to the pedestrians. These wearable sensors have been demonstrated to give comprehensive dynamic information of pedestrian movements. With appropriate fitting and location of the sensors, the acceleration and velocity of different limbs can be measured. Distributing a number of sensors over the pedestrian can provide a complete dynamic description of the movement of the pedestrian. Current wearable devices are used to reliably recognize activities such as standing, walking, and running. The data that they provide is accurate but they do not have the capabilities to broadcast this information to vehicles in proximity. In addition, this information is only available for pedestrians that are wearing the devices, and the devices must be functioning properly.

These limitations have prevented the use of this type of information in advanced driver assistance systems (ADAS) and autonomous vehicles applications. The fundamental question addressed in this work is if it is possible to obtain reliable and accurate dynamic and kinematic information of pedestrian limbs by using vision based sensors that are ubiquitous in intelligent vehicles.

Fig. 2: The autonomous electrical vehicle retrofitted with vision perception technology

III. Experiment Setup

The Australian Centre for Field Robotics (ACFR) is involved in a SMART city project that aims to bring new types of mobility to the university campus. In particular, the project addresses the seamless interaction between autonomous vehicles and people in a highly crowded environment. Inferring intention of pedestrians is of significant importance for a successful interaction between human and autonomous machines. We are currently using two electrical vehicles (EVs) with a set of sensors to facilitate autonomous driving, including six wide-angle cameras and an NVIDIA DRIVE PX2 automotive computer, which includes two GPUs for powerful parallel computation. The vehicle setup is shown in Figure 2. The six cameras mounted on the vehicle cover 360 degrees. Each camera has 100 degrees FOV and a framerate of 30 fps. The data used in this paper uses the images from the front camera.

The data was collected in public areas in streets surrounding the University of Sydney campus in Australia. The actions evaluated were: *standing*; *walking* (slow and fast); *starting to walk*; *walking then stopping*; and *running*. These are the activities of most interest that are performed by pedestrians in an urban environment. The dataset was collected in two parts. The first group of people were invited to perform standard activities as listed above. From this group, the images and dynamic information provided by an IMU was recorded while a single person was walking in front of the camera. The IMU was located on wrist and ankles of the pedestrian. The experiment was repeated with each person of the group. The IMU carried by pedestrians was based on a Raspberry Pi 2 single-board computer retrofitted with a Sense HAT module, which includes a 3-axis accelerometer, gyroscope and magnetometer. The IMU data were acquired at 50 Hz, which is considered sufficiently high to capture kinematic characteristics of human limbs. The Raspberry Pi was connected to the vehicle using the wireless network provided by the vehicle. The information of this device is used to validate the results of this paper. All data were captured, synchronized and stored within the ROS environment. The time-stamping and synchronization of the vision and inertial data are critical for this work.

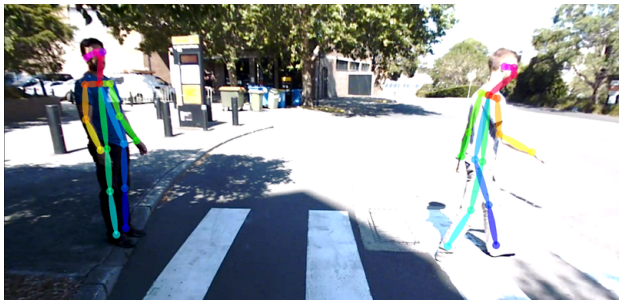


Fig. 3: Two pedestrians performing different actions in the same scene at the same time. One of them is standing on the border of the street, while the other is crossing the street.

A second dataset consisting of naturalistic data was obtained from a group of people walking in a public area performing the type of actions described above, as illustrated in Figure 3. This dataset is used to demonstrate the viability of the approach with all limbs and in real situations.

Although the images in the experiments were captured from a stationary vehicle platform, the proposed approach can be extended to work in moving platforms provided with good navigation capabilities. We are currently collecting datasets from a moving vehicle and operating under poor lighting conditions. We plan to make these datasets available to the research community.

IV. Results

This section shows the results obtained from the analysis of the datasets described in the previous section. The first part of the results shows the validation of the dynamic pedestrian tracking from vision data using a comparison with synchronized inertial data. The second part of the results, based on the naturalistic dataset, demonstrates the capabilities and robustness of the presented approach.

A. Calculation of Dynamic Information

The *OpenPose* library provides the skeleton representation in each image as a group of points, whose coordinates are used to calculate the angles of the limbs, the angular velocity and the corresponding linear accelerations. The angle of each limb is obtained by considering the endpoints of each line segment, which represents a limb in the skeleton, and computing its angle with respect to the horizontal line using trigonometric functions. Angular velocities are the first derivatives of the limb angles, computed by considering the change of angles in two consecutive images and the time interval. Lastly, linear accelerations are the second derivatives of the linear displacement of skeleton points from three consecutive images. The results are then smoothed using a moving average filter.

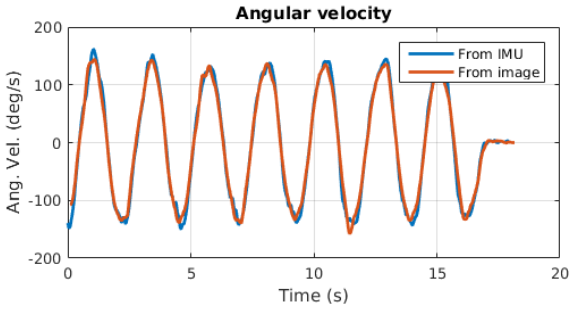


Fig. 4: Comparison of wrist angular velocity data. The blue line represents the data from the gyroscope's Z-axis while the orange line is the angular velocity obtained by processing the skeleton representation from the images.

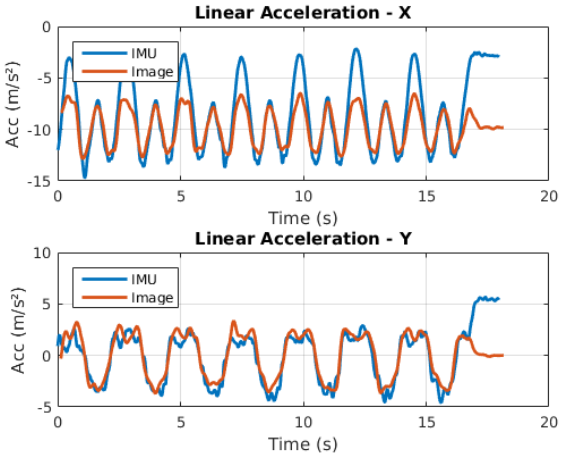


Fig. 5: Comparison of wrist linear acceleration data. In both plots, the blue line is the data provided by the IMU and the orange line is the data obtained by processing the images. The top plot shows the acceleration in X-axis while the bottom shows that of the Y-axis.

B. Approach Validation

This section presents a comparison of the measured inertial data with the limb dynamics estimated from the visual data. In this work, we compare the angular velocity and linear acceleration of wrists and ankles obtained from images with measurements from the accelerometer and gyroscope.

Figure 4 shows the angular velocity data from the gyroscope's Z-axis, which corresponds to the direction of the right arm, in blue line. The IMU was located on the right wrist of the person during walking. The orange line represents the angular velocity obtained by processing the same wrist of the skeleton representation derived from the images. It can be seen qualitatively that the values are very close. The similarity between the signals can be quantitively measured using cross-correlation, with a result correlation of 99.33%. We consider that this is an excellent result, the vision based sensor is seen to be comparable to the high accuracy of the wearable sensor.

The linear acceleration results were not as accurate as the angular velocity. Nevertheless, they still correlate very well with the acceleration measured by the wearable sensors. The frequency, shape and timing are very similar, the discrepancy appears only in the amplitude of the signals. This is more noticeable in the X-axis as shown at the top of Figure 5. This is due to the fact the precise mass of the limb of a random pedestrian walking on the street is not available when calculating the acceleration using the inverted pendulum model. The similarity in this axis is calculated as 79.76%. The linear acceleration results in Y-axis are closer to the IMU measurements, in terms of both gain and shape. In this case, the similarity in the Y-axis improves to 83.31%. This comparison is shown at the bottom of Figure 5.

Besides the skeleton pose, the vision data provides additional information including limbs' lengths, rotation ratios and linear acceleration of all limbs, frame by frame from the sequence of images. The skeleton kinematics were evaluated for different kinds of actions over a number of different pedestrians. Some of these are presented in this section. The results obtained and validated using the presented method are divided in accordance with different parts of the body as shown in the next section.

C. Full Skeleton Information

An accurate representation of the pedestrians skeleton is obtained by the *OpenPose* library. Despite this, there are certain conditions in which the results are incomplete or the algorithm outputs false detections. A common failure appears when one arm is hidden behind the body. In some cases the limb is not shown, but in other cases the limb appears in the incorrect place. This problem is addressed in this paper by selecting the most visible arm or leg of the subject of interest as the main limb to track.

Another problem that can manifest particularly in outdoor scenarios is that the algorithm produces false positives where skeletons are incorrectly detected. In the dataset collected for this paper, it was often a result of shadows or shapes similar to humans. This issue was resolved in this paper by using semantics information in images. The algorithm used in [15] was used to mask only the areas of the image labeled as "pedestrian", which helps address the false detection problem. A comparison sequence is shown in Figure 6 and in the video available in [21]. Figure 6a shows a standard image taken from the front camera of the vehicle. All images captured were rectified using the intrinsic parameters of the camera. The image in Figure 6b was processed to extract the skeleton of the person crossing the street. This image shows an erroneous result, where an extra skeleton was detected on the right side of the image. This object, marked with a yellow dotted oval in Figure 6b, is obviously not related to the pedestrian in the scene. Figure 6c shows the semantics classification outcome of the same image, which is used to filter out the phantom skeleton.

D. Limbs

The main points of interest in the skeleton representation are the arms and legs. Arms have 3 points that match the three main joints: shoulder, elbow and wrist (points 2-3-4

(a) (b) (c)

Fig. 6: Using semantics information to remove false detections. The original image captured by the vehicle's front camera is shown in (a). (b) illustrates false skeleton detection on the right side of the image, which is marked with a yellow dotted oval. In (c), semantics information is used to filter out the false detection. Pedestrian is colored in yellow, vehicles in red, drivable area in brown, vegetation in green, and sky in blue color.

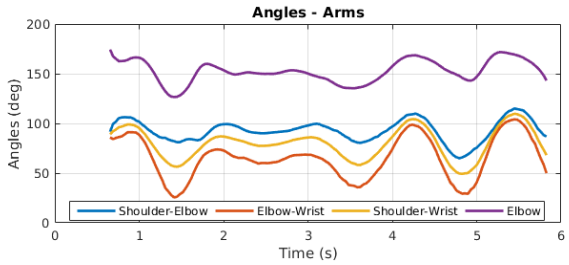


Fig. 7: Angles between different junctions and sections of arms with respect to the horizontal line, extracted from the skeleton of a person walking until $t = 2$ s, standing until $t = 3$ s and walking again after $t = 3$ s.

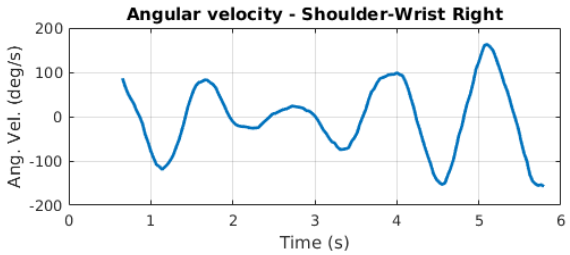


Fig. 8: Angular velocity of the right shoulder-wrist segment.

and 5-6-7 in Figure 1a). The angles and angular velocities between the horizontal line, as shown in Figure 1b, and the shoulder-elbow section and elbow-wrist section give information related to the action of the pedestrian. A single pedestrian performing different activities can be seen in Figure 7. This figure shows the angles of different arm joints extracted from the skeleton while a pedestrian walks in front of a car from $t = 0$ s to $t = 2$ s. Then the person stops until $t = 3$ s and starts to walk again afterwards.

The standing time (from $t = 2$ s to $t = 3$ s) can be recognized through the lack of angle change in each of the arm segments. The blue line shows the angle of the shoulder-elbow segment with respect to the horizontal line. The average of this angle is around 90 degrees and its variation is only a few degrees. The red line shows the angle of the elbow-wrist, from which one can clearly differentiate between walking and standing, and also find out the traveling direction of the person. In this case, according to our angle assumption shown in Figure 1b, the person can be determined to be walking from left to right within the image frame of reference, because the average is less than 90 degrees. When the person is moving from right to left, the average of this signal is over 90 degrees. The yellow line is the angle of a imaginary line that joins the shoulder and the wrist. The purple line is the angle of the elbow, composed by the shoulder-elbow and elbow-wrist segments.

Figure 8 presents the angular velocity of the shoulder-wrist segment. During the time the pedestrian is standing, the arms of the person were still moving a small amount because they could not decelerate to zero velocity instantaneously.

Legs are also divided into three junctions: hip, knee and ankle for each leg (points 8-9-10 and 11-12-13 on Figure 1a). The variation of the knee angle indicates activities of walking and stepping up or down the sidewalk. These are important activities to recognize, in particular whether or not the pedestrian is stationary. This state is observed if the angle between hip-knee right and left does not changed. This case is shown in Figure 9, where the angles between different segments of the legs are presented. The top two plots show the angles of the segment hip-knee and knee-ankle, respectively, with respect to the horizontal line for both legs. The standing time is clearly visible between $t = 2$ s and $t = 3$ s, during which the angles of two legs are 90 degrees, meaning both legs are not moving. The bottom left plot presents the angles of the hip-ankle imaginary segment, which show clear differences between the time spent walking and standing. The bottom right plot shows the angle of the knee, composed by the intersection of the hip-knee and knee-ankle segments.

Figure 10 shows the angular velocity of the knee-ankle segment. This provides very clear information to discriminate between a pedestrian walking or standing still. The signal pattern during walking is clearly distinguishable as the person moves a foot forward, while during standing the angular velocity is found zero.

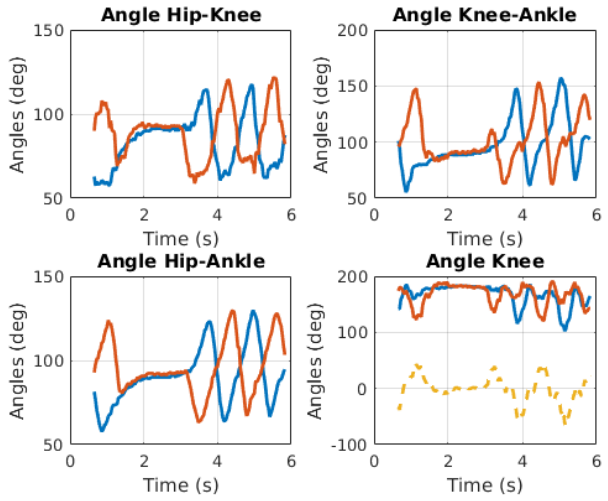


Fig. 9: Angles between different junctions and sections of legs with respect to the horizontal line. The last plot also shows the difference of the knee angle. The right leg is colored in blue, the left in orange and the difference in dotted yellow line.

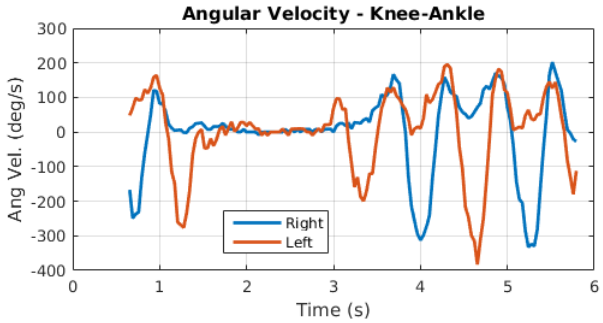


Fig. 10: Angular velocity of one segment of the leg. It is easy to recognize the different actions carried out by the pedestrian such as walking and standing.

The linear velocity of a pedestrian can be obtained by processing the linear acceleration of the limbs. Also, linear acceleration provides more information related to the state of the body, particularly, if it is starting to move from a resting state, as shown at time $t = 3$ s in Figure 11. Previous research such as [22], uses this kind of information to calculate the length of the leg using an inverted pendulum model.

An analysis of the utilization of the proposed approach has been presented in [23], where the authors use an IMU attached to the ankle to recognize the cycle of a stride. The foot forward creates a distinctive pattern for recognizing, according to the authors, six different activities including slow, normal and fast walking, running, climbing and descending stairs. The results obtained with our approach using only vision are comparable with those in [23] using inertial sensors.

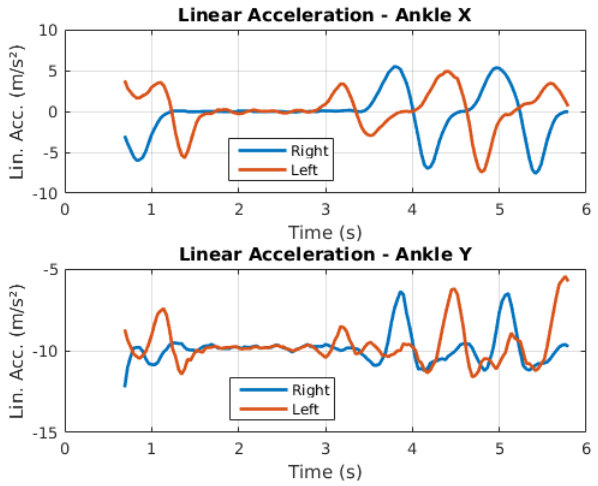


Fig. 11: Acceleration of the ankles in X and Y axes. The initial acceleration from standing to walking time is clearly shown at time $t = 3$ s.

V. Conclusion

The aim of this paper is to show that the dynamic and kinematic information from a pedestrian can be obtained from vision sensors. The approach presented in this paper is able to extract dynamic information from a single, or group of pedestrians with accuracy comparable with that of typical gyroscopes and accelerometers installed in wearable devices. The hardware used, based on an NVIDIA DRIVE PX2, is able to achieve real-time image acquisition and processing, and produce satisfactory results.

From the analysis of pedestrians, we obtain the skeleton representations and poses followed by different actions such as walking, standing, starting to walk, and stopping. From a sequence of skeleton poses, we can extract data relevant to lengths, angles, rotation rates and linear accelerations. The data are shown to be strongly correlated with measurements made by an inertial unit carried by pedestrians over different parts of the body. A dataset collected using pedestrians following a defined sequence of actions is used to validate the measurements obtained from the images. A second dataset composed of naturalistic data from real pedestrians is used to generalize the approach as described in Section IV.

We demonstrate in Section IV-B that data obtained from vision sensors is very similar to data provided by the IMU using a wearable device. This contribution not only improves the pedestrian detection but also leads to more relevant data such as angular velocities and linear accelerations. This information is of fundamental importance to estimate pedestrian activities and intent. We have also shown that the method proposed can be augmented with other techniques such as semantics to improve the robustness and reliability of the results. We consider the outcomes of this paper a significant contribution to address the fundamental problem of the seamless interaction between autonomous systems and humans.

REFERENCES

[1] E. Dagan, O. Mano, G. Stein, and A. Shashua, "Forward collision warning with a single camera," in *Intelligent Vehicles Symposium, 2004 IEEE*, June 2004, pp. 37–42.

[2] P. Hurney, P. Waldron, F. Morgan, E. Jones, and M. Glavin, "Review of pedestrian detection techniques in automotive far-infrared video," *Intelligent Transport Systems, IET*, vol. 9, no. 8, pp. 824–832, 2015.

[3] D. Gerónimo and A. M. López, *Vision-based pedestrian protection systems for intelligent vehicles*. Springer, 2013.

[4] Mobileye technologies limited. [Online]. Available: http://www.mobileye.com/

[5] C. Keller and D. Gavrila, "Will the pedestrian cross? a study on pedestrian path prediction," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 15, no. 2, pp. 494–506, April 2014.

[6] K. Fuerstenberg, K. Dietmayer, and V. Willhoeft, "Pedestrian recognition in urban traffic using a vehicle based multilayer laserscanner," in *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 1, June 2002, pp. 31–35 vol.1.

[7] (2014) Ibeo Automotive Systems GmbH. [Online]. Available: http://www.ibeo-as.com

[8] H. Ritter and H. Rohling, "Pedestrian detection based on automotive radar," in *Radar Systems, 2007 IET International Conference on*, Oct 2007, pp. 1–4.

[9] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *CVPR*, 2017.

[10] T. Simon, H. Joo, I. Matthews, and Y. Sheikh, "Hand keypoint detection in single images using multiview bootstrapping," in *CVPR*, 2017.

[11] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in *CVPR*, 2016.

[12] M. Ye, X. Wang, R. Yang, L. Ren, and M. Pollefeys, "Accurate 3d pose estimation from a single depth image," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 731–738.

[13] R. Quintero, J. Almeida, D. Llorca, and M. Sotelo, "Pedestrian path prediction using body language traits," in *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, June 2014, pp. 317–323.

[14] M. Andriluka, S. Roth, and B. Schiele, "Monocular 3d pose estimation and tracking by detection," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 623–630.

[15] W. Zhou, R. Arroyo, A. Zyner, J. Ward, S. Worrall, E. Nebot, and L. Bergasa, "Transferring visual knowledge for a robust road environment perception in intelligent vehicles," in *2017 IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2017.

[16] C. Benedek, B. Gálai, B. Nagy, and Z. Jankó, "Lidar-based gait analysis and activity recognition in a 4d surveillance system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 101–113, Jan 2018.

[17] H. Kataoka, Y. Aoki, Y. Satoh, S. Oikawa, and Y. Matsui, "Fine-grained walking activity recognition via driving recorder dataset," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Sept 2015, pp. 620–625.

[18] N. Abhayasinghe and I. Murray, "Human activity recognition using thigh angle derived from single thigh mounted imu data," in *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct 2014, pp. 111–115.

[19] J. B. Bancroft, D. Garrett, and G. Lachapelle, "Activity and environment classification using foot mounted navigation sensors," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Nov 2012, pp. 1–10.

[20] A. Bujari, B. Licar, and C. E. Palazzi, "Movement pattern recognition through smartphone's accelerometer," in *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, Jan 2012, pp. 502–506.

[21] Filtering pedestrian skeletons on images using semantics. [Online]. Available: https://youtu.be/BRcOneJ3vn0

[22] T. N. Do, R. Liu, C. Yuen, and U. X. Tan, "Design of an infrastructureless in-door localization device using an imu sensor," in *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Dec 2015, pp. 2115–2120.

[23] B. Beaufils, F. Chazal, M. Grelet, and B. Michel, "Stride detection for pedestrian trajectory reconstruction: A machine learning approach based on geometric patterns," in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sept 2017, pp. 1–6.