Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014, Kraków, Poland 19-23 July 2014
Edited by T. Ahram, W. Karwowski and T. Marek

Comparative Analysis of Driving Maneuvers Using 3D Body Posture Data

Alexandra Kondyli [a], Virginia P. Sisiopiku [b], Liangke Zhao [c] and Angelos Barmpoutis [c]

[a] *Department of Civil and Coastal Engineering*
University of Florida
Gainesville, FL 32611-6580, USA

[b] *Department of Civil, Construction, and Environmental Engineering*
University of Alabama at Birmingham
Birmingham, AL 35294-4440, USA

[c] *Digital Worlds Institute*
University of Florida
Gainesville, FL 32611-5800, USA

ABSTRACT

Research has shown that driver inattention is the most prevalent cause of traffic collisions accounting for an estimated 25 to 56% of crashes in the US. Driver inattention may result from drivers engagement in secondary activities (such as texting or cellphone use), or lack of awareness of the surrounding environment. The main objective of this research is to investigate the relationship between potentially unsafe driving events and the actual driver body posture and movements when performing a driving maneuver under different traffic configurations. The paper presents results from a pilot study that captured the 3-D posture and activity of three drivers while performing both mandatory (merging) and discretionary (lane changing) maneuvers on freeway and arterial segments in Gainesville, Florida. The body posture of the drivers was captured through the use of a low-cost infrared depth sensor. A 7-point human skeletal model was fit to the captured depth frame sequences using our proposed framework. The comparative analysis of the participants' body movements while performing the maneuvers revealed differences between the participants' body activity when performing the same maneuvers. The findings of this research provide significant insights regarding which body movements may hide unsafe situations while performing a driving maneuver that requires the attention of the surrounding environment.

Keywords: Infrared depth sensors, driver behavior, merging process, lane changing, body posture

INTRODUCTION

Driver error constitutes a major cause of traffic crashes internationally (Peden et al., 2004). According to the World Health Organization (WHO), annually there are over 1.2 million fatalities and over 20 million serious injuries worldwide. In the US, the 100-car naturalistic study (Dingus et al., 2006) concluded that driver distraction is responsible for about 80 percent of crashes and 65 percent of near crashes. In an effort to assist drivers in the driving process and reduce the degree of uncertainty due to driver error, a number of advanced (or intelligent) driver assistance systems (ADAS) have been introduced (Shaout et al., 2011). Such systems are typically designed to provide assistance or warnings to drivers by taking into account the position of the vehicle or other vehicle-related components. Examples of driver-assistance systems include lane departure warning, pedestrian protection, blind spot detection, collision avoidance, lane change assistance, and automatic parking.

Recognizing the importance of driver assistance with respect to traffic safety, the Federal Highway Administration (FHWA) and the US Department of Transportation (USDOT) introduced the connected-vehicle research program. This program offers a mixture of cutting edge technologies such as advanced wireless communications, on-board computer processing, advanced vehicle-sensors, GPS navigation, and smart infrastructure, to identify and warn the drivers on unsafe roadway conditions. The connected-vehicle program supports vehicle-to-vehicle (V2V) and vehicle-to-infrastructure communication research activities. The vehicle-to-vehicle communication is related to the exchange of data (e.g., speed, acceleration, heading angle, etc.) over wireless network that provide information on surrounding vehicles status and allows for performing calculations and issue driver warnings to avoid crashes. The communication option is based on Dedicated Short Range Communications (DSRC). Although the development of the communication component of this program is not complete to date, a number of crash avoidance systems (e.g., blind spot and lane changing warning, forward collision warning, etc.) have been established so far. In August 2012, DOT launched the Safety Pilot "model deployment" in collaboration with University of Michigan's Transportation Research Institute (UMTRI), where nearly 3,000 vehicles were deployed in the largest-ever road test of V2V technology. DOT testing is indicating interoperability of V2V technology among products from different vehicle manufacturers and suppliers and has demonstrated that they work in real-world environments (NHTSA, 2014).

It should be noted that currently available ADAS systems are designed to look at vehicle's external environment whereas systems that focus on the drivers' behaviors and actions inside the vehicle are still limited. However, the next generation of advanced driver assistance systems should also consider human factors issues and take into account interactions among drivers and vehicles, either in relation to their own vehicle or the surrounding vehicles. A multifaceted approach that looks at both the ergonomics aspect of human factors research as well as the psychological aspects of the engaged participants is expected to advance the state-of-the art of ADAS systems in the future.

Toward this direction, this paper investigated the relationship between potentially unsafe driving events and the actual driver body posture and movements when performing a driving maneuver (e.g., lane changing, merging) under different traffic configurations in a naturalistic setting. The ultimate objective of this research is to develop a framework for constructing an in-vehicle driver-assistance system that accounts for the driver's body posture and movements, rather than considering solely the vehicle position relative to other vehicles on the road.

LITERATURE REVIEW

A significant amount of research has focused on understanding driver intentions and actions (e.g., inattention or distraction states), using vision-based systems. Research studies primarily look into tracking of the head and the face of the driver. For example, Huang and Trivedi (2004) and Murphy-Chutorian and Trivedi (2010) developed a system that monitors and predicts drivers' head pose using video detection. Their head detection system was designed for identifying drivers' inattention and distraction, however, the authors did not specifically address how inattention was detected. Similarly, Braathen et al. (2001), developed an approach for identifying spontaneous facial expressions, such as blinking, to monitor alertness and anxiety. Huang et al. (2003) and Wu and Trivedi (2008) proposed a model that combines head pose detection with actual vehicle movement direction.

In addition, past research (e.g., Tijerina et al., 2005; Trivedi et al., 2007; McCall et al., 2005; Doshi et al., 2011) analyzed combination of head pose and gaze data for identifying and predicting driver's intent to change lanes and perform a maneuver. Research has also studied the position of hands and the grasp in conjunction with monitoring the head pose for lane change intent analysis and prediction (Cheng and Trivedi, 2010) or for driver distraction monitoring (Tran and Trivedi, 2009). A system that was developed to track the 3D body movement combined with head pose was also introduced in Tran and Trivedi (2010), where preliminary results of body posture and lane changing activity were collected in a simulation environment. Tran et al. (2012) used video-based analysis in conjunction with pedal sensor measurements and looked at drivers' foot behavior. They developed prediction models for braking and acceleration and concluded that the foot behavior depends greatly on the driver type.

A significant amount of research has dealt with predicting automobile driving posture for purposes of vehicle interior design. Reed et al. (2002) developed a model that predicts driver posture through a series of cascading regression functions. The final model is constructed to produce the best fit for the eye and the hip position using data from drivers with varying anthropometric characteristics. Fatollahzadeh (2006) proposed mathematical models that predicted truck drivers' comfortable sitting posture and position. Mathematical models using multiple regression analyses on selected body landmarks as well as anthropometrical measures were developed which proposed a linear correlation between parameters. Kovacevic et al. (2010) explored the impact of anthropometric measurements on ergonomic driver posture and concluded that anthropometric measurements of drivers and surrounding measurements of car controls affect traffic safety and driver's fatigue.

Overall, the literature review reveals that a significant amount of research has been involved with the development of advanced driver-assistance systems; however, most of these systems rely on the automobile position on the roadway and do not necessarily consider the drivers posture and actions. Apart from that, the lane trajectory and position of the vehicle could potentially differ from the driver's intent to change lanes. In addition safety research has focused on eye tracking as a means of capturing driver's attention, fatigue, or drowsiness; however, the entire body posture of drivers when performing a maneuver as well as different postures between various groups of drivers may also reveal behaviors that contribute to unsafe driving conditions.

METHODOLOGICAL FRAMEWORK

This research proposes a novel approach for studying the actual movements of drivers inside the vehicle, when performing specific maneuver types or while engaging to secondary tasks that require a certain body movement. With the use of a low-cost infrared depth sensor, the 3D shape of selected participants is being constructed, as they are performing various driving maneuvers and/or as engaged in secondary tasks while driving. A brief description of the methodology undertaken to collect and analyze the 3D data is presented here; however, more detailed information can be found in Kondyli et al. (2013). Each data frame captured by a digital depth sensor is a two dimensional array of depth values (i.e., distance between the sensor and objects). Similarly, a collection of frames is a three dimensional array that can be represented as $D \in R^{W \times H \times N}$, where N denotes the total number of recorded frames, and W and H denote the number of pixels across the width and height of the depth frame respectively. The depth value in a particular pixel with coordinates (i, j) on frame i is denoted by $D_{i,j,t} \in R^{+}$. In practice, each depth camera has a specific range of operation, which restricts accordingly the range of the recorded values (see depicted field of view in Figure 1). The depth frames can be equivalently expressed as quadratic meshes given by $X_{i,j,t} = (i - i_c) D_{i,j,t} f^{-1}$, $Y_{i,j,t} = (j - j_c) D_{i,j,t} f^{-1}$, $Z_{i,j,t} = D_{i,j,t}$, where (i_c, j_c) denote the coordinates of the central pixel in the depth frame, and *f* is the focal length of the depth camera. One of the advantages of the quadratic mesh representation of the depth frames is that they can be easily visualized using virtual lighting, shading, perspective and point of view using standard computer graphics techniques (Faugeras, 1993). An example of the quadratic mesh of a captured depth frame is shown in Figure 1 (left).

The 3D shape of the body of the driver and part of the vehicles' cabin are clearly captured in the depth frame. Optionally, the color information from a video frame can be applied as a texture to the quadratic mesh of the depth frame (Figure 2) and can also be used to enhance the detection of the body features.

The primary goal of our data processing method was to trace body features using the captured depth frame sequences. The body features of our interest included the *X, Y, Z* coordinates of the wrists, elbows, and shoulders as well as the orientation of the torso. The values of these quantities can be estimated by fitting a human skeletal model to each of the depth frames in our datasets. The main challenge in the skeletal fitting process is that the human body in our particular field of view is very close to other objects such as the driver's seat, the steering wheel and the driver's door. Any generic skeletal fitting algorithm performs better when the human body is clearly visible and at a distance from nearby objects, and therefore will fail in our in-cabin setting. For instance, the skeleton tracking algorithm included in the Microsoft Kinect Software Development Kit (SDK) fails in detecting the driver's body as it was previously reported in Kondyli et al. (2013).

Figure 1. Left: visualization of a depth frame. Right: The corresponding mask with enhanced boundaries between objects, computed using our framework

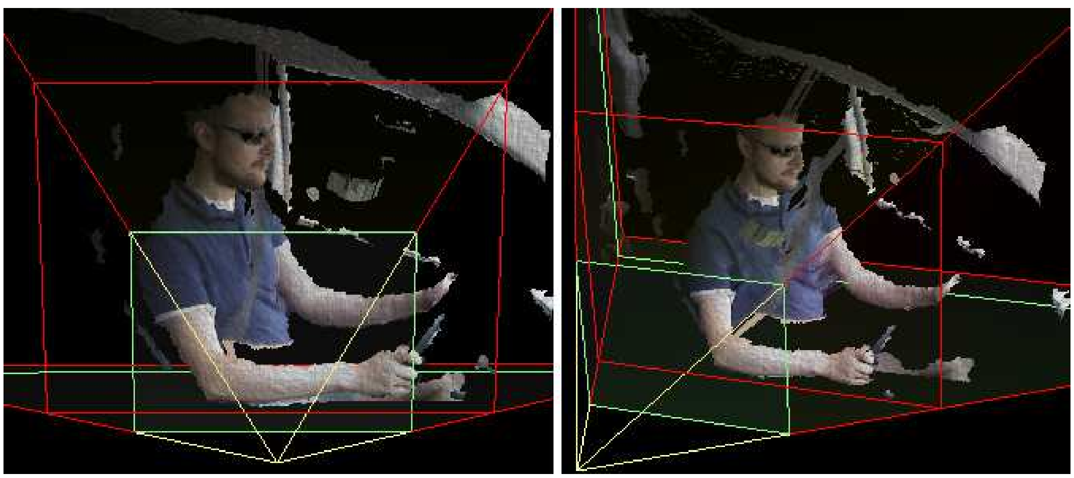

Figure 2. Two 3D views of the same frame from the recorded dataset. The video and depth frames are presented as a sequence of textured 3D frames. The field of view of the depth camera is also shown as a trapezoid.

In order to overcome the aforementioned skeleton fitting challenges we developed a novel algorithm that was designed to fit a 7-point skeletal model to the body of the driver using a sequence of depth frames. Our skeletal model included the line segments between the following joints: right wrist, right elbow, right shoulder, neck, left shoulder, left elbow, and left wrist. The skeletal model is visualized in Figures 5 and 6. In our visualization we also show the triangle formed by the left shoulder, the right shoulder and the neck, whose normal vector was used as an indicator of the torso orientation.

The proposed skeleton fitting algorithm scans the depth frames in a diagonal fashion from upper right to lower left, pixel stripe by pixel stripe until the entire image is covered. In each diagonal pixel stripe the medial points of the masked regions (see mask in Figure 1) are detected. The medial points that belong to the same region of the mask are connected so that they form medial line curves in all the masked regions. It should be noted that the medial points and curves are calculated in 3D and not in the 2D coordinates of the frames. After that the detected medial curves are filtered so that potential noise caused by the depth sensor is removed. Finally the curves that correspond to the arms are detected by using spatial constraints as well as geometrical constraints regarding the size, orientation and curvature of the arms. This process fits our 7-point skeletal model to the best matching medial curves. The proposed algorithm has linear complexity, which allow us to perform the fitting of the skeleton in real time in less than 15 milliseconds per depth frame in the computer configuration described in the next section.

DRIVER BEHAVIOR DATA COLLECTION

The field data obtained for this research were collected along a 2.6 mi stretch of Interstate 75 (I-75) in the southbound (SB) and northbound (NB) directions, and a 0.7 mile long arterial segment (Newberry Road eastbound and westbound approaches) in Gainesville, FL. The freeway segment along I-75 has three lanes per direction and the arterial segment has three through lanes per direction, several median openings, and includes a total of six signalized intersections. A schematic of the data collection sites is presented in Figure 3.

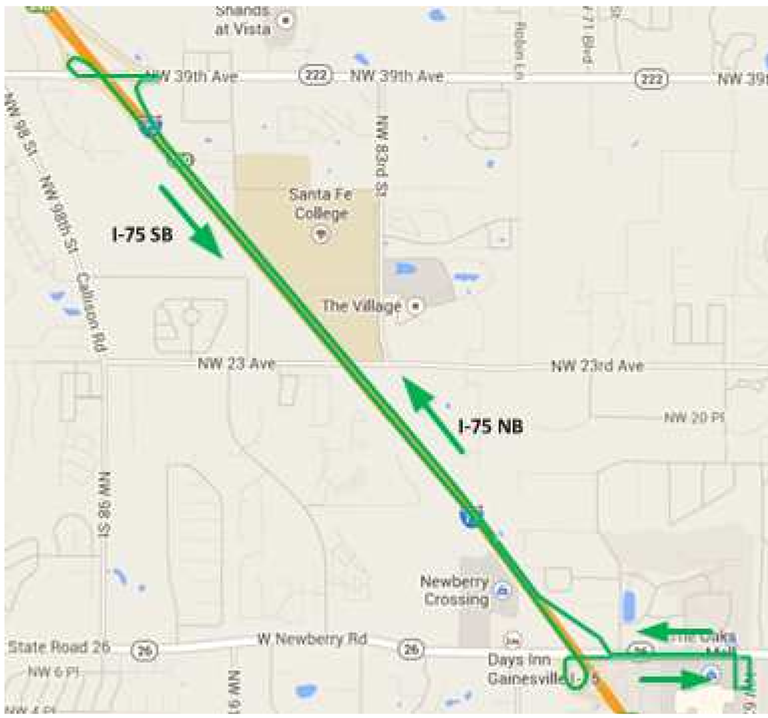

Figure 3. Map of the data collection route along I-75 and Newberry Road in Gainesville, FL.

The data collection effort took place on Sunday, September 1st 2013, between 10 am and noon. Traffic conditions were generally uncongested and free-flowing, especially on the freeway segment. Traffic on the arterial segment was light, although towards the end of the data collection effort the flows were considerably increased. For the purposes of this pilot study, three participants affiliated with the research team were asked to complete one route along the freeway and arterial segments. The participants performed two mandatory lane changes (i.e., merging onto the freeway) and several discretionary lane changes on the freeway and the arterial street. The entire duration of the experiment for each participant was approximately 20 minutes.

The real-time driver behavior data were acquired using the PrimeSenseTM depth sensor contained in the Microsoft Kinect^{TM} sensor. The device was connected (via a USB 2.0 port) to a 64-bit computer with Intel Core i5 (quad core) CPU at 2.53GHz and 4GB RAM. The computer and the sensor were both powered using a 75 Watt car power inverter. The resolution of the depth camera was 320 × 240 pixels with horizontal field-of-view angle (FoV) angle of 57^{o}. The resolution of the video camera was 640 × 480 pixels with horizontal FoV of 62^{o}. The range of the camera was calibrated so that it records depth values in the range from 0.5m to 3.0m, which is suitable for the limited space of the cabin of a typical passenger vehicle. The sensor was fixed on the front passenger's door, so that the driver is within the field of view of the depth and video cameras. **Error! Reference source not found.** shows the field of view of the depth camera. The green rectangle depicts the closest plane of sensing, which is located 0.5m in front of the sensor (shown as the tip of the yellow pyramid in Figure 2).

DATA ANALYSIS

The video and depth sequences captured during this study were segmented into several fragments that correspond to the merging and lane changing maneuvers that were performed as participants were driving along the freeway and the arterial segment. Each of the fragments was analyzed independently using the framework presented earlier, and a comparative analysis was performed across the corresponding datasets from different participating drivers. The proposed framework was implemented in Java using the Java-for-Kinect library introduced by Barmpoutis (2013).

One characteristic of the Microsoft Kinect is that it has embedded an accelerometer that can be used to obtain the signature of the roadway segment that was used in this study. The output of the accelerometer is a unit vector, which typically indicates the direction of gravity in the case of a steady sensor. If the sensor moves due to the motion of the vehicle, the output of the accelerometer may be affected by the slope of the roadway as well as changes in the speed of the vehicle. Figure 4 shows the the X and Z coordinates of the accelerometer reading for the three drivers participated in this study. It should be noted that the Y dimension is perpendicular to the roadway and therefore is mainly affected by the vector of gravity and therefore it does not capture significant information about the signature of the road. Although the time stamp varies by driver due to differences in their driving speeds, the actual profiles of the accelerometer signature are very consistent among all three drivers. These signatures essentially describe changes in the elevation and slope of the roadway segment. For example, the first part of the graph clearly depicts the vertical alignment change due to vehicle driving from the arterial street on the on-ramp (upgrade) and onto the freeway (downgrade). Similarly, the last part of the graph shows the change in elevation due to the vehicle exiting the freeway through the off-ramp (downgrade). The information from the accelerometer is very useful for identifying the exact locations for obtaining the video and depth sequences of interest (e.g., merging onto the freeway).

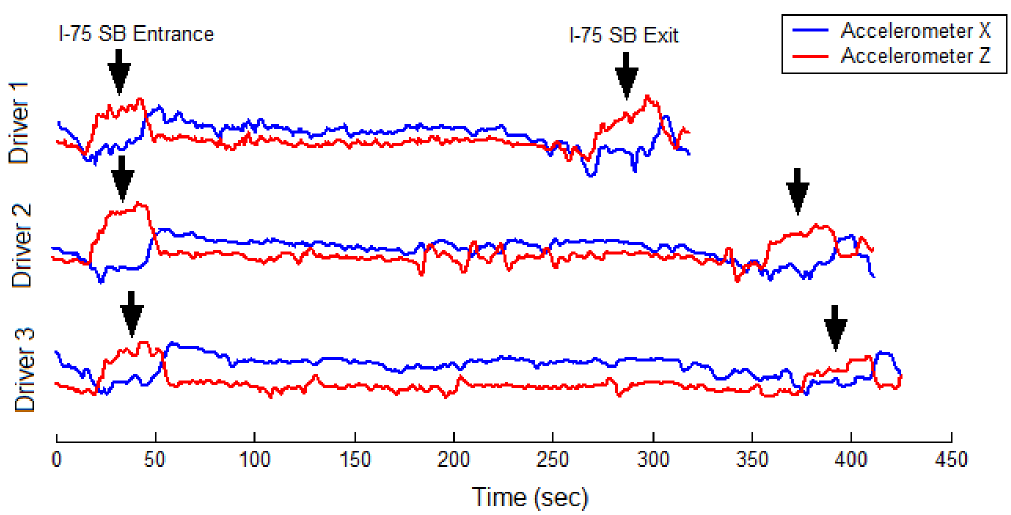

Figure 4. Microsoft Kinect accelerometer data obtained for three participants.

Using the framework presented earlier, this study examined differences in the body posture during a lane change maneuver for two of the three participants. Figure 5 shows the seven point skeleton model before and after a lane change maneuver for Driver 1 and Driver 3. The differences in the body posture between the two drivers are apparent from this figure. The torso of Driver 1 remains practically unaltered during the maneuver, whereas Driver 3 clearly shifts her body to the left in order to have a better visual of the traffic at the next lane. On the other hand, Driver 1 shifts only the head to identify potential conflict at the next lane through the rear mirror.

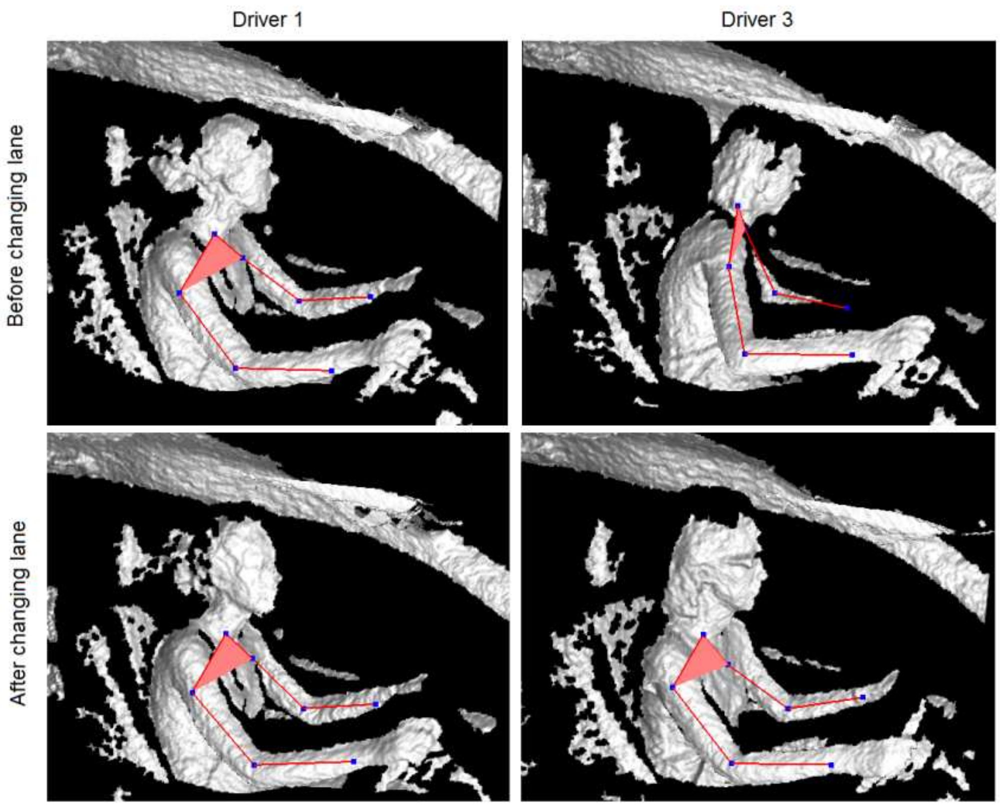

Figure 5. Change in body posture due to a lane change maneuver shown by the fitted skeletons in the depth frames.

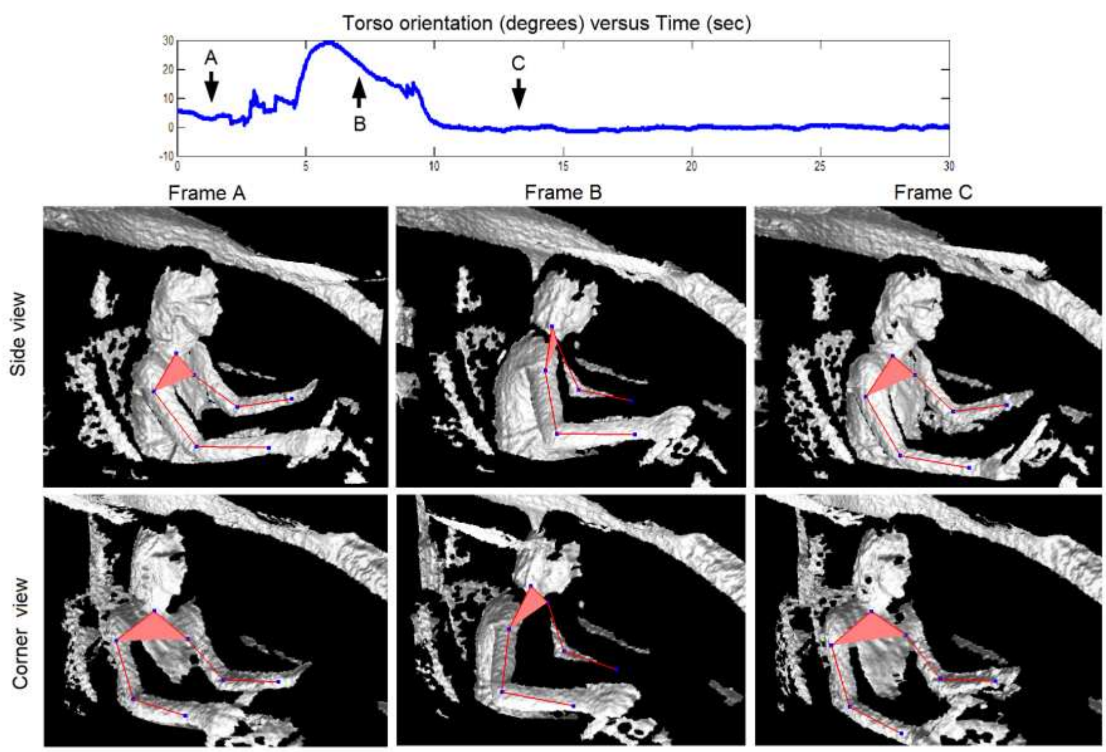

Figure 6. Change in body posture due to a merging maneuver for Driver 3. The plot shows the torso orientation during the merging maneuver. The depth frames and fitted skeletons of 3 frames are shown from two 3D perspectives.

In addition to the lane change maneuver, a comparative analysis of the body posture during a merging maneuver was also performed. Figure 6 presents the frame sequence during a merging maneuver for Driver 3, along with the corresponding time-series of the torso orientation. In this graph the torso orientation represents the rotation in degrees from the torso position perpendicular to the steering wheel. The orientation is positive for left-turn rotation and negative for right-turn torso rotation. Frames A, B and C are taken as before, during, and after the execution of the merging maneuver. From these graphs it is clear that the torso rotation of Driver 3 is considerably increased during the merging task. Driver 3 torso orientation during this merging maneuver is consistent with the lane changing example shown in Figure 5.

The same analysis was performed for the remaining two participants while merging on the freeway. The time-series plots of the torso orientation are presented in Figure *7*.

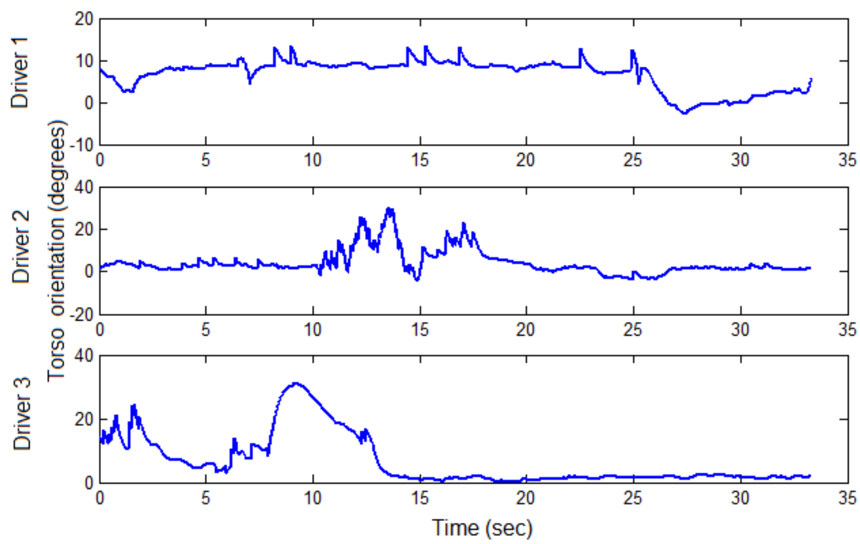

Figure 7. Comparison of the torso orientation during a merging maneuver at the same ramp merge junction between the three participants.

A comparative analysis of the time-series of the torsi orientation shows considerably different profiles of the torsi during the merging maneuver task. Comparable to Driver 3, Driver 2 also displayed increased torso rotation during the merging task; however, the profile of the torso rotation is different, as this is evident by the unsmooth time-series plot. Driver 1 torso orientation differs significantly from the other two. As shown in Figure *7*, Driver 1 does not seem to have an apparent torso rotation, but there is some rotation throughout the merging task (there is a small detected rotation at around 10 degrees for the first 25 seconds of the task).

Similarly to the torso orientation, a comparative analysis of other parts of the participants' body motion can be performed. Figure 8 shows the time-series of the X, Y, Z coordinates of the wrists, elbows, shoulders for Driver 3, during the entire duration of the driving task. Using the data shown in Figure 8 it is easy to obtain instances where there is significant body activity by identifying spikes in the respective graphs, and further analyze the underlying conditions for these instances.

By observing Figure 8 it is evident that there is a more frequent arm motion detected during the arterial segments compared to the freeway segments as we anticipated. For example in this dataset the driver started merging onto the freeway at 100 sec. and exited at 300 sec. which correspond to intense arm activity as indicated by a significant change to the coordinates of the wrists and the elbows. During the freeway segment (between 100 sec. and 300 sec.) there was no significant change of posture detected and the coordinates of the traced joints change only occasionally as it was also anticipated. This smooth driving pattern is significantly different compared to the one observed during the arterial segments which corresponds to 0 sec. - 100 sec., 300 sec. - 480 sec., and 650 sec. - 700 sec. During these segments the driver stopped at red traffic lights and followed a path that included many 90-degree turns as it is shown in Figure 3. All of these instances were naturally associated with arm activity, which corresponds to changes in the coordinates of the wrists slightly as it is shown in Figure 8. Finally, the segment from 480 sec. - 650 sec. corresponds to the northbound freeway segment, which was associated with occasional body motion according to the plots in Figure 8.

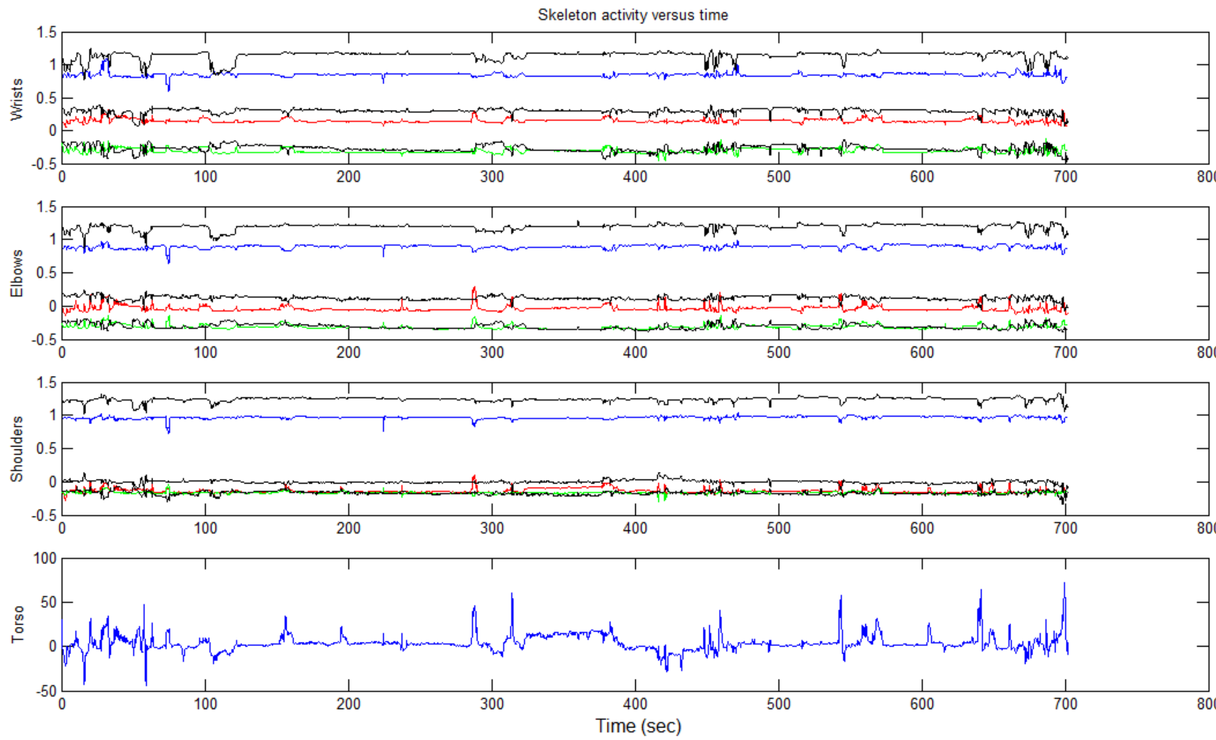

Figure 8. Skeleton activity versus time for Driver 3.

CONCLUSIONS AND RECOMMENDATIONS

In this paper a pilot study was conducted to capture the 3-D posture and activity of three drivers while performing various driving maneuvers. A low-cost infrared depth sensor was used to capture the 3-D posture. The participants were observed to perform merging and lane changing maneuvers on freeway and arterial segments in Gainesville, Florida. A comparative analysis of the participants' body movements while performing the maneuvers was conducted which revealed differences between the participants' body activity when performing the same maneuvers. More specifically, it was observed that there are considerable differences in terms of the torso activity for both lane changing and in merging maneuvers between different drivers. It is recommended to further examine the torso activity in conjunction with the activity of the shoulder, elbows and wrists, in order to obtain more complete representation of the entire body posture. It is further suggested to evaluate the effect of actual traffic conditions, such as proximity of the subject vehicle to the adjacent lead/lag vehicles, to the body posture while performing these maneuvers. The findings of this research can provide significant insights regarding which body movements may hide unsafe situations while performing a driving maneuver that requires the attention of the surrounding environment.

ACKNOWLEDGMENT

Funding for this study is provided by US DOT/RITA through STRIDE, the Alabama Department of Transportation (ALDOT) and the Florida Department of Transportation (FDOT). The support of the study sponsors is greatly appreciated.

REFERENCES

Barmpoutis, A. (2013), "Tensor Body: Real-time reconstruction of the human body and avatar synthesis from RGB-D," IEEE Transactions on Cybernetics, Volume 43, No. 5, pp. 1347–1356.

Braathen, B., Bartlett, M., Littewort-Ford, G., Movellan, J. (2001), "3-d head pose estimation from video by nonlinear stochastic particle filtering," in Proceedings of the 8th Joint Symposium on Neural Computation.

Cheng, S., Trivedi, M. (2010), "Vision-based infotainment user determination by hand recognition for driver assistance," IEEE Transactions on Intelligent Transportation Systems, Volume 11, No. 3, pp. 759–764.

Dingus, T. A., Klauer, S., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., Knipling, R. (2006), "The 100-car naturalistic driving study, phase ii results of the 100-car field experiment, report no. DOT HS 810 593," National Highway Traffic Safety Administration (NHTSA), Tech. Rep.

Doshi, A., Morris, B., Trivedi, M. (2011), "On-road prediction of driver's intent with multimodal sensory cues," Pervasive Computing, IEEE, Volume 10, No. 3, pp. 22–34.

Fatollahzadeh, K. (2006). "A laboratory vehicle mock-up research work on truck driver's seat position and posture", Doctoral thesis, Stockholm, Sweden.

Huang, K., Trivedi, M., Gandhi, T. (2003), "Driver's view and vehicle surround estimation using omnidirectional video stream," Proceedings in Intelligent Vehicles Symposium, IEEE, pp. 444–449.

Huang, K., Trivedi, M. (2004), "Robust real-time detection, tracking, and pose estimation of faces in video streams", Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004,Volume 3, pp. 965–968.

Kondyli, A., Sisiopiku, V., Barmpoutis, A. (2013), "A 3D experimental framework for exploring drivers' body activity using infrared depth sensors", IEEE International Conference on Connected Vehicles and Expo, Las Vegas, NV.

Kovacenic, S., Vucinic, J., Kirin, S., and Pejnovc, N. (2010) Impact of anthropometric measurements on ergonomic driver posture and safety, Period Biol, Volume 112 No 1, pp. 51-54.

McCall, J., Trivedi, M., Wipf, D., Rao, B. (2005), "Lane change intent analysis using robust operators and sparse bayesian learning," Computer Vision and Pattern Recognition - Workshops, CVPR Workshops, IEEE Computer Society, pp. 59–59.

Murphy-Chutorian, E., Trivedi, M. (2010), "Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness," IEEE Transactions on Intelligent Transportation Systems, Volume 11, No. 2, pp. 300–311.

NHTSA (2014), NHTSA Announcement available at http://www.safetypilot.us/images/documents/NHTSA_Announcement_2-3-14.pdf

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A., Jarawan, E., Mathers, M. (2004), "World report on road traffic injury prevention," World Health Organization, Tech. Rep.

Reed, M. P., Manary, M. A., Flannagan, C. A. C., Schneider, L. W. (2002), "A Statistical Method for Predicting Automobile Driving Posture", Human Factors Volume 44 No. 4.

Shaout, A., Colella, D., Awad, S. (2011), "Advanced driver assistance systems - past, present and future," in Computer Engineering Conference (ICENCO), 2011 Seventh International, pp. 72–82.

Tijerina, L., Stoltzfus, D., Parmer, E. (2005), "Eye glance behavior of van and passenger car drivers during lane change decision phase," Transportation Research Record: Journal of the Transportation Research Board, Volume 1937, p. 3743.

Tran, C., Trivedi, M. (2009), "Driver assistance for "'keeping hands on the wheel and eyes on the road"'," IEEE International Conference on Vehicular Electronics and Safety (ICVES), pp. 97–101.

Tran, C., Trivedi, M. (2010), "Towards a vision-based system exploring 3d driver posture dynamics for driver assistance: Issues and possibilities," in Intelligent Vehicles Symposium (IV), 2010 IEEE, pp. 179–184.

Tran, C., Doshi, A., Trivedi, M. (2012), "Modeling and prediction of driver behavior by foot gesture analysis," Computer Vision and Image Understanding, Volume 116, pp. 435–445.

Trivedi, M., Gandhi, T., McCall, J. (2007), "Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety," Transactions on Intelligent Transportation Systems, IEEE, Volume 8, No. 1, pp. 108– 120.

Wu, J., Trivedi, M. (2008), "A two-stage head pose estimation framework and evaluation," Pattern Recognition, Volume 41, pp. 1138–1158.