Hands on the wheel: a Dataset for Driver Hand Detection and Tracking

Guido Borghi, Elia Frigieri, Roberto Vezzani and Rita Cucchiara
DIEF - Department of Engineering "Enzo Ferrari"
University of Modena and Reggio Emilia
Modena, Italy
{name.surname}@unimore.it

Abstract— The ability to detect, localize and track the hands is crucial in many applications requiring the understanding of the person behavior, attitude and interactions. In particular, this is true for the automotive context, in which hand analysis allows to predict preparatory movements for maneuvers or to investigate the driver's attention level. Moreover, due to the recent diffusion of cameras inside new car cockpits, it is feasible to use hand gestures to develop new Human-Car Interaction systems, more user-friendly and safe. In this paper, we propose a new dataset, called *Turms*, that consists of infrared images of driver's hands, collected from the back of the steering wheel, an innovative point of view. The *Leap Motion* device has been selected for the recordings, thanks to its stereo capabilities and the wide view-angle. Besides, we introduce a method to detect the presence and the location of driver's hands on the steering wheel, during driving activity tasks.

I. INTRODUCTION

In last decades, the study of hand detection and tracking has been widely investigated in the computer vision community. In particular, hand motion analysis is an interesting topic in many fields, like Human Computer Interaction [1], Human Behavior Understanding [2] and Action Recognition [3], due to the fact that hands are a rich source of information to study the motion, the interaction between humans and the environment and people behavior. In this paper, we focus in particular on the *automotive* context, a research field dominated by some challenging elements, like frequent light changes and hands occlusion. Moreover, we are witnessing to the increasingly presence of car cameras, to look both inside and outside the vehicle. These cameras are able to acquire data in a completely non-invasive manner. Besides, the recent spread of commercial light-invariance and high-quality cameras, like infrared, depth and thermal ones, gives the opportunity to collect a great amount of data and new ways to acquire images during driving activity.

In the automotive context, hand detection and tracking systems are key elements for many tasks.
First of all, through the hand detection and tracking is it possible to analyze the driver's attention level. According to the *National Safety Administration* (NHTSA), the driving

Fig. 1. Typical points of view for hand detection and tracking datasets in automotive. In particular, (a, b) are taken from [8], (c) [9], (d) [10], (e) [11] and (f) shows our innovative point of view. The fish-eye infrared camera is placed back to the steering wheel, very close to the hands.

distraction is generally defined as "an activity that could divert a person's attention away from the primary task of driving" [4]. As trivially expected, the inattentive driving plays a crucial role in road crashes [4]. Crayek *et al.* in [5] and [6] report three categories of driving distractions:

- **Manual distraction**: driver's hands are not on the wheel, the driver is not ready to quickly correct the trajectory or to react to eventual obstacles on the road;
- **Visual distraction**: driver's eyes are not looking at the road but they are engaged in other tasks, such as looking at the phone or reading a newspaper;
- **Cognitive distraction**: driver attention is not focused on driving activity, due to bad physical conditions or *fatigue*, that "concerns the inability or disinclination to continue an activity, generally because the activity has been going on for too long" [7];

Smartphones are one of the most important cause of fatal

driving crashes: as reported in [8], they cause about 18% of fatal driver accidents in North America and involve all three distraction categories mentioned above. Besides, drivers today are increasingly engaged in secondary tasks behind the wheel [12]. Above mentioned elements suggest that hand detection can play a main role in monitoring driver's behavior.

Several techniques have been proposed for hand detection in car, based on different input types [13]: *physiological signals*, like EEG, ECG and EMG, *vehicle signals*, acquired from the car bus, and *physical signals*, collected from inside cockpit cameras. In our work, we focus on *non-intrusive physical signals*.

Secondly, hand tracking allows the analysis of preparatory movements for maneuvers, and this is could be useful for providing alerts [14] and support to the driver [8], [15], [16] during driving activity.

Finally, hand tracking could be the first step for new entertainment and infotainment systems based on *Natural User Interfaces* [17], in which driver can interact with the car using hands and gestures only. Generally, these new systems are characterized by the concept of *user-friendly* usability [18], [19]. They also allow an incrementation of the level of velocity and so the safety of the Human-Car Interaction.

For these reasons, we aim to propose a new dataset, namely *Turms*, that contains infrared images of hands acquired placing an infrared camera – a *Leap Motion* device (Fig. 2) – that is able to provide high-quality and short-range images with a wide scene, thanks to its fish-eye lens. The presence of a stereo camera allows the possibility of recovering depth and 3D information. We collect images from an innovative point of view – back the steering wheel, very close to the hands – as shown in Figure 1 (f). Moreover, we propose a real time method to detect if the driver hands are or not on the wheel and in which position. The presented method can be considered as a first step to develop innovative *Human-Car Interaction* systems, where the driver can interact with the car *without* moving away the hands from the steering wheel.

The work is organized as follows: in the next section, related datasets and methods on hand detection and tracking for the automotive context are analyzed; in Section III, the *Turms* dataset is presented. In Section IV, the proposed method is detailed, in particular the image pre-processing (IV-A), homography transformation (IV-B) and image unrolling (IV-C) parts are shown. Finally, Section V draws the conclusions.

II. RELATED WORKS

Hand detection and tracking are the goal of several datasets and works in computer vision field, in particular in the automotive context [20]. Following, we report only automotive-based works and datasets.

Existing method can be divided depending on the type of images they rely on: RGB, RGB-D or both.

In [21] a new and challenging dataset containing a variety of in-car automotive data has been introduced, in particular for hand detection task under naturalistic driving settings. Some challenges in this dataset are: varying illumination condition, skin-colored non-hand objects, occlusions and different viewpoints. The acquisition viewpoint is from the top of the driver seat.

Ohn-Bar *et al.* [8] proposed a vision-based framework to analyze hand activity in real time. Hands are detected and tracked from a monocular camera, then motion trajectories are studied in settings of activity recognition, activity prediction and abnormal event detection. To do that, a SVM classifier [22] and a set of features elaborated on trajectories, displacement and transition histograms of displacement are used. Martin *et al.* [23] used two different cameras: one camera looking at the driver hands and the other looking at the driver head. In this way, visual information are extracted, data about hands and head are integrated and a driver behavior analysis can be done. [24] used a side-mounted camera looking at a driver's profile and utilizes the silhouette appearance obtained from skin-color segmentation for detecting activities.

Ohn-Bar *et al.* [25] provided a solution to localize the driver hands by partitioning RGB and depth images into disjoint sub-regions (wheel, lap, hand rest, gear, infotainment region), acquired by a RGB-D sensor placed in front of the wheel. Then different feature extraction methods are proposed and analyzed, based on modified HOG features and a linear kernel SVM, to perform hand or not hand detection in each region of interest.

In [11] a vision-based system exploiting both RGB and depth descriptor has been developed in order to classify hand gestures. Hand detection and gesture recognition are performed in a specific region located in the infotainment area, next to the wheel: for this reason, driver has to employ one hand to perform gestures.

Also in [27] a gesture recognition framework is proposed, based both on RGB and depth data; gestures have to be performed in front-row seat occupants and are designed for

Fig. 2. Different views of the *Leap Motion* device. The position of stereo cameras and infrared LEDs are shown. The original image is taken from the official website [26].
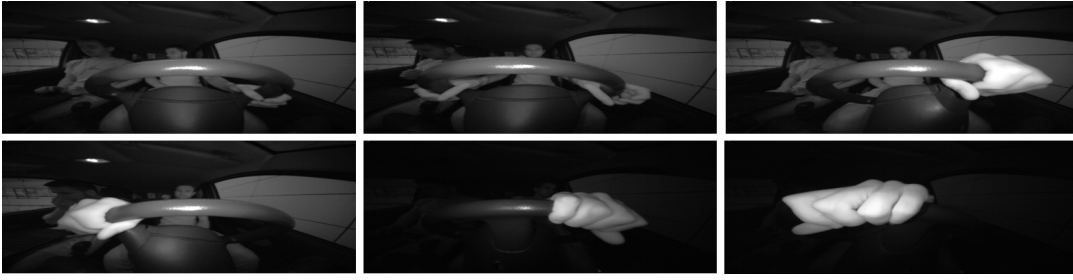
Fig. 3. Sample frames taken from the *Turms* dataset. As shown, various scenes are present, such as both hands or only one hand grasping the steering wheel.

accessing the infotainment system. HOG features [28] and a SVM classifier are used to perform classification.

3D movement of driver head and hands from multiview video are analyzed in [9], limited to the upper body parts. Thanks the aid of an upper body and head trackers, driver basic activities are determined: two, one or no hands on the wheel, head looking left, straight or right. Unlike methods listed above, Cheng *et al.* [10] used a RGB and infrared camera to acquire images of the front-row seat area in the vehicle, to detect hands performing user classification (driver or passenger). This approach exploits HOG features, a linear SVM and median filtering over time to classify each acquired frame. Besides, [15] used a thermal and color cameras to acquire images and the CAN bus of the vehicle for steering wheel angle data. Hand grasp analysis, head tracking and driver activity recognition are performed.

III. THE TURMS DATASET

To acquire a dataset in an automotive context, the choice of the most appropriate acquisition device is fundamental: an infrared camera has to been exploited, to overcome the traditional limit of RGB sensors that produce bad quality images with limited or even missing light sources. Moreover, due to the particular position that we adopted – *i.e.* at the back of the steering wheel –, the camera *must* be able to acquire short range images with a wide view-angle. Thanks to this particular position, we suppose to have a small number of hand occlusions produced by driver's body during driving activity. Besides, a small-sized acquisition device is required, to be integrated inside the dashboard. Considering the above elements, we decided to use the *Leap Motion* [26] device, shown in Figure 2, to acquire the *Turms* dataset.

Leap Motion is a device originally designed to interact with digital contents in virtual and augmented reality, based on natural body language: this device is able to track both hand and finger movements, with a low latency, converting them to 3D input. It is usually placed at the same plane of a laptop keyboard, on the desktop, in front of the user and with its cameras pointing upwards.

We collect the *Turms* dataset in a real car context, placing the acquisition device at the back of the steering wheel. Leap Motion device fits automotive requirements, for these reasons:

- **Infrared camera**: infrared light allows to develop vision-based systems that are able to work in different light conditions, for example during the night or during bad weather events;
- **Stereo camera**: depth or 3D information could be retrieved thanks to the presence of stereo images;
- **Short range**: thanks to its fish-eye lens, Leap Motion is able to capture a 150-degree scene; in this way, a single frame can include the steering wheel, right and left hands and the upper part of the driver body;
- **Small size**: a fundamental requirement because of the limited dimensions of the car instrument panel, for an eventual future integration in the cockpit. Leap Motion device is $7 \times 1.2 \times 3$ cm [26].
- **Real time performance**: high frame rate (more than 200 fps) allows to capture fast hand motions;

Besides, *Leap Motion* has two infrared LEDs, dynamically regulated to light up the scene. The novel and particular position of the acquisition device allows to obtain images with limited hand occlusions. Moreover, it is possible to perform gestures without moving away the hands from the steering wheel. Finally, the device is able to capture images in a completely non-invasive manner.

Turms dataset has the following features:

- **Data**: the dataset consists of 14k frames, acquired in a stereo mode, divided into sequences. The spatial resolution is 640×240. For each sequence, different actions have been performed by the subjects;
- **Variety**: a great variety is guaranteed due to the presence of 7 subjects (5 males, 2 females), each recorded twice;
- **Annotations**: since the original SDK of Leap Motion does not work in our particular setup due to the presence of occlusions of the wheel, we manually annotated the position $(x, y$ coordinates) of both left and right hands in each frame. We also provide the initial position of the steering wheel (5 points of the ellipse fitted in the external contour of the wheel).

The dataset is publicly available. [1]

---

[1] http://imagelab.ing.unimore.it/turms

IV. Proposed method

In this section, the proposed method is detailed. We develop a frame-by-frame algorithm to detect the presence of hands on the steering wheel, in three steps.

In the first step, the input image is pre-processed, in order to enhance the visual quality and to detect the ellipse fitting the external border of the steering wheel. In the following step, the ellipse is turned into a circle, through a homography transformation. Finally, the steering wheel image is unrolled, creating a new image space, in which driver's hands are detected.

A. Image acquisition and pre-processing

Both Leap Motion stereo cameras are geometrically calibrated to rectifying frames and to eliminate the fish-eye effect from images. Moreover, contrast stretching is applied in order to enhance the contrast.

Then, steering wheel detection is performed through 5 points manually selected from the rectified image in the external ellipse of the wheel: even if this step is not strictly required by the next stages, it is useful to visually verify the steering wheel detection. Through the least squares method an ellipse is fitted [29], [30].

Indeed, an ellipse can be defined as:

$$f(a, (x, y)) = D \cdot a = 0 \quad (1)$$

where

$$D = (x^2, xy, y^2, x, y, 1)$$
$$a = (a_{x,x}, a_{x,y}, a_{y,y}, a_x, a_y, a_1). \quad (2)$$

The ellipse can be fitted to N data points $\mathbf{x}_i, i = 1, ..., N$ by minimizing the distance:

$$\Delta(a, \mathbf{x}) = \sum_{i=1}^{N} (f(a, x_i))^2 \quad (3)$$

The standard method to solve the above minimization problem is to introduce a Lagrange multiplier λ:

$$L(a) = \Delta(a, \mathbf{x}) - \lambda(a^T C a + \phi). \quad (4)$$

where C is a matrix (6×6) filled with zeros, with the exception for $C_{1,3} = C_{3,1} = 2$ and $C_{2,2} = -1$ and $\phi = a^T C a$, resulting from the derivative of $L(a)$. Minimizing $L(a)$ as a function of a gives the expected result. This goal function can be rewritten (for all and detailed passages see [30], [31]) as a generalized eigenvalue problem:

$$\frac{1}{\lambda} a = S^{-1} C a \quad (5)$$

where $S = \sum D_i^T D_i$ is a scatter matrix. The solution a is the eigenvector corresponding to the largest eigenvalue $\frac{1}{\lambda}$. This routine is applied only on the first frame at the beginning of each video sequence, since we can suppose the steering wheel is fixed in the same position inside the car cockpit.

(a)

(b)   (c)

Fig. 4. Image (a) is the original frame acquired by the *Leap Motion* device; (b) is the same frame after image rectification and contrast stretching; (c) is the final frame after homography transformation. In (b) and (c) wheel center and borders are highlighted in red color.

B. Steering wheel homography

In this step, through an homography geometric transformation, the ellipse of the wheel is turned into a circle.

This kind of geometric operation implies a set of challenges. Our starting assumption is that the position and the orientation of the acquisition device are unknown. From this point of view, a 2D homography has 8 grades of freedom, so a 3×3 matrix H has to been discovered, to perform a correct geometric transformation.

At first glance, it may appear that this mapping is quite hard to be determined, with so many unknowns, but we can consider some elements to reduce the complexity: first, we know that some of the observed points in the scene lie on some unknown plane, identified by the steering wheel, and so it is possible to establish a homography between the camera and our region of interest – *i.e.* the plane that contains the wheel and the hands –. Besides, we can assume that there is only one rotation along x-axis, because the acquisition device

---

Algorithm 1 Iterative Homography Matrix Research

1: **procedure** Find homography matrix
2:     p_e: ellipse points manually acquired
3:     i_{min}, i_{max}: min and max rotation on x-axis
4:     x_{min}, x_{max}: min and max translation on x-axis
5:     y_{min}, y_{max}: min and max translation on y-axis
6:     **for** x from x_{min} to x_{max} **do**
7:         **for** i from i_{min} to i_{max} **do**
8:             **for** y from y_{min} to y_{max} **do**
9:                 H = get homography matrix
10:                 p_c = project points p_e through H
11:                 $crls$ = fit circles with p_e
12:                 d_{dist} = compute center distance of $crls$
13:                 r_{diff} = compute radius difference of $crls$
14:                 **if** $d_{dist} \leq th_{dist}$ and $r_{diff} \leq th_{diff}$ **then**
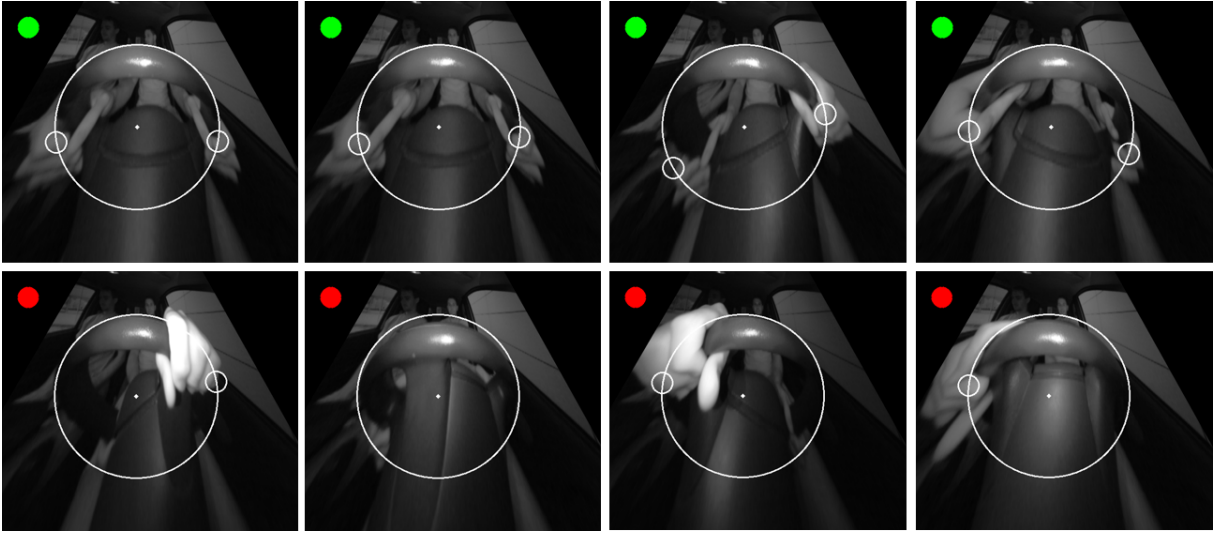15:   homography matrix found

Fig. 5. Sample output frames of the proposed hand detection systems. In particular, the framework checks if both driver's hands are present (green circle) or not (red circle) on the steering wheel, in the traditional "10:00/2:00" hand position. White circles denote the position of hands, if present.

is parallel with the wheel – it is placed into the instrument panel, so the rotation along y-axis is not present and the rotation along the z-axis is not significant for our scope – and so there is only one relevant translation on the same axis. Then, the distance between the acquisition device and the wheel is fixed (150 mm in our experiments). Therefore, we can assume the presence of only small translations along x and y axis.

We find the homography matrix with an iterative procedure. The key idea is to define a temporary homography matrix and then use this matrix to project input points belonging to the ellipse (see Section IV-A) in the new space; after that, 3 circles are fitted through 3 of the 5 points projected. If the homography transformation is correct, each point lies on a circumference, then all three obtained circles have about the same center and the same radius length. If different radius lengths are found, the procedure is repeated. If differences between the position of the center – all circles must be *concentric* – and the length of radius are under a certain threshold, we can assume that the correct homography matrix is found. The above procedure is detailed in Algorithm 1. This process is CPU bound, but it is conducted only on the first frame of the sequence. In this way, at the end of that procedure we know the center and the radius of the wheel.

C. Steering wheel unrolling

In order to perform hand detection, a new image space is created passing from a circular to a linear space, exploiting a well-known approach in the iris analysis field [32], [33]. Each pixel with (x, y) coordinates is mapped onto a (ρ, θ) space, where ρ is the pixel position with respect to the wheel center (in degrees), and θ is its distance. Thanks to this step, the hand detection algorithm detailed in the next section is simplified.

As the previous step, also this process is CPU bound, then to meet computational requirements a static table is calculated

before the unrolling procedure. The obtained table maps (ρ, θ) and (x, y) coordinates once, because the relations are the same for all images.

Unrolled images, as depicted in Figure 6, are cropped, keeping only the thickness of the steering wheel, empirically set: this makes more easier hand detection, removing all external noises around them. Pixel values are taken backward, namely starting from unrolled space and doing a bi-linear interpolation to avoid missing pixel.

D. Hand Detection

The hand detection is carried out as follows. The sum of pixel values is computed for each column of the unrolled images, the final result is a intensity histogram as shown in Figure 7. Since the unrolled images contain only the steering wheel and, usually, the steering wheel has a dark color (as in *Turms* dataset), we could suppose that the presence of strong discontinuities (*peaks* over a certain threshold) in histogram values could be related to the presence of driver's hands. Besides, we observe that we always obtain a intensity peak in the central position of the histogram, due to the presence of the infrared light emanated by Leap Motion.
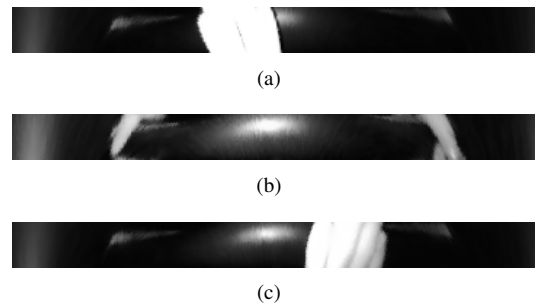

(a)

(b)

(c)

Fig. 6. Three examples of unrolled frames. In the first the steering wheel is turned to the left, in (c) to the right and in (b) both hands grasp the wheel.
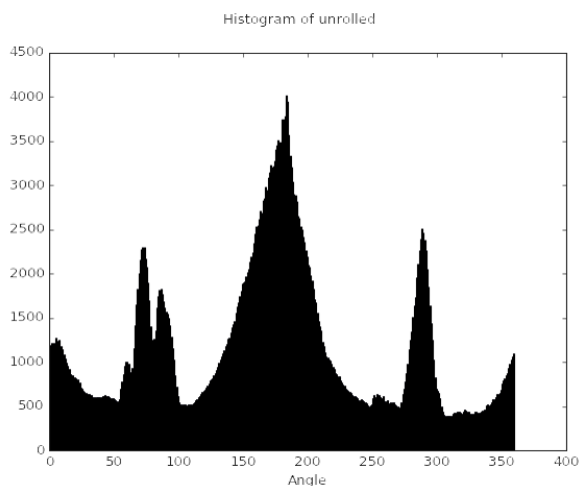
Fig. 7. Histogram of an unrolled image. The central peak is related to the infrared light emanated by the acquisition device. Left and right peaks are related to the presence of corresponding hands on the steering wheel.

The hand detection algorithm searches frame by frame for two local maxima on the left and right parts of the central peak and a threshold is empirically set. These two peaks, if present, are related with the presence of the right and left driver's hands. The corresponding value on the x-axis value represents the angle of the hand on the steering wheel.

We use the whole dataset to test the proposed pipeline. We use only the first frame of each sequence to manually acquire points for the initial steering wheel ellipse detection. We achieve an accuracy of more that the 95% in the prediction of the presence of both hands on the wheel. We note that the proposed algorithm is not able to detect the hand position in the central part of the histogram where the main peak is present.

A comparison with other literature hand detection approaches is not feasible. Due to the original point of view of our acquisition device and the presence of occlusions caused by the steering wheel, state-of-art hand detection techniques [34] generally fail. The proposed system runs at more than 30 fps, achieving real time performance, with low memory requirement. All tests have been carried on a *Intel i7-4790* (3.60 GHz).

Starting from this hand detection algorithm, a variety of applications can be developed. For instance, as represented in Figure 5, we developed a tool to detect the presence, depicted with a green or red circle in the top left corner of frames, of both hands on the steering wheel in the traditional "10:00/2:00" hand position. Moreover, we also show hand locations, depicted as small white circles.

V. CONCLUSION

In this paper, a new dataset, namely *Turms*, and a related driver's hand detection method have been presented.

The dataset consists of more than 14k infrared images containing driver's hands, acquired during real driving activities through the *Leap Motion* device, from a rear point of view,

oriented to avoid driver's body occlusions. The presence of stereo images allows the possibility to retrieve depth information, if needed.

The proposed method is able to detect and locate driver's hands in real time and shows good promise in real world applications, thanks to its reliability, accuracy, real time performances and low requirements in hardware setup.

A rich possibility for future works is possible: for instance, we are planning to include a tracking algorithm, to perform driver's behavior analysis and assuring temporal coherence during hand detection. Besides, we are investigating the consequences to include gloves and different skin colors in the proposed dataset.

Finally, this method can be considered as a first step for innovative systems for Human-Car Interaction, where users can perform gestures and can interact with the car, without the need to move away the hands from the wheel.

REFERENCES

[1] S. Mitra and T. Acharya, "Gesture recognition: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 3, pp. 311–324, 2007.

[2] M. Pantic, A. Pentland, A. Nijholt, and T. S. Huang, "Human computing and machine understanding of human behavior: a survey," in *Artifical Intelligence for Human Computing*. Springer, 2007, pp. 47–71.

[3] F. Baradel, C. Wolf, and J. Mille, "Human action recognition: Pose-based attention draws focus to hands," in *ICCV Workshop on Hands in Action*, 2017.

[4] "distraction.gov, official us government website for distracted driving," http://www.distraction.gov/index.html, accessed: 2016-09-14.

[5] C. Craye and F. Karray, "Driver distraction detection and recognition using RGB-D sensor," *CoRR*, vol. abs/1502.00250, 2015. [Online]. Available: http://arxiv.org/abs/1502.00250

[6] G. Borghi, R. Gasparini, R. Vezzani, and R. Cucchiara, "Embedded recurrent network for head pose estimation in car," in *Intelligent Vehicles Symposium (IV), 2017 IEEE*. IEEE, 2017, pp. 1503–1508.

[7] H. Croo, M. Bandmann, G. Mackay, K. Rumar, and P. Vollenhoven, "The role of driver fatigue in commercial road transport crashes," *Eur. Transp. Safety Council, Brussels, Belgium, Tech. Rep*, 2001.

[8] E. Ohn-Bar and M. M. Trivedi, "Beyond just keeping hands on the wheel: Towards visual interpretation of driver hand motion patterns," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2014, pp. 1245–1250.

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

[10] S. Y. Cheng and M. M. Trivedi, "Vision-based infotainment user determination by hand recognition for driver assistance," *IEEE transactions on intelligent transportation systems*, vol. 11, no. 3, pp. 759–764, 2010.

[11] E. Ohn-Bar and M. M. Trivedi, "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2368–2377, 2014.

[12] S. G. Klauer, F. Guo, J. Sudweeks, and T. A. Dingus, "An analysis of driver inattention using a case-crossover approach on 100-car data: Final report," Tech. Rep., 2010.

[13] M. Venturelli, G. Borghi, R. Vezzani, and R. Cucchiara, "From depth data to head pose estimation: a siamese approach," in *n Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2016.

[14] A. Doshi and M. M. Trivedi, "Tactical driver behavior prediction and intent inference: A review," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2011, pp. 1892–1897.

[15] S. Y. Cheng, S. Park, and M. M. Trivedi, "Multi-spectral and multi-perspective video arrays for driver body tracking and activity analysis," *Computer Vision and Image Understanding*, vol. 106, no. 2, pp. 245–257, 2007.

[16] E. Ohn-Bar, A. Tawari, S. Martin, and M. M. Trivedi, "Predicting driver maneuvers by learning holistic features," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*. IEEE, 2014, pp. 719–724.

[17] M. N. K. Boulos, B. J. Blanchard, C. Walker, J. Montero, A. Tripathy, and R. Gutierrez-Osuna, "Web gis in practice x: a microsoft kinect natural user interface for google earth navigation," *International journal of health geographics*, vol. 10, no. 1, p. 45, 2011.

[18] T. Nawaz, M. S. Mian, and H. A. Habib, "Infotainment devices control by eye gaze and gesture recognition fusion," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 2, 2008.

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

[20] D. Tran and M. M. Trivedi, "Vision for driver assistance: Looking at people in a vehicle," in *Visual Analysis of Humans*. Springer, 2011, pp. 597–614.

[21] N. Das, E. Ohn-Bar, and M. M. Trivedi, "On performance evaluation of driver hand detection algorithms: Challenges, dataset, and metrics," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 2953–2958.

[22] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[23] S. Martin, E. Ohn-Bar, A. Tawari, and M. M. Trivedi, "Understanding head and hand activities and coordination in naturalistic driving videos," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*. IEEE, 2014, pp. 884–889.

[24] H. Veeraraghavan, S. Atev, N. Bird, P. Schrater, and N. Papanikolopoulos, "Driver activity monitoring through supervised and unsupervised learning," in *Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005*. IEEE, 2005, pp. 580–585.

[25] E. Ohn-Bar and M. Trivedi, "In-vehicle hand activity recognition using integration of regions," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*. IEEE, 2013, pp. 1034–1039.

[26] "leap motion, official us website," https://www.leapmotion.com/, accessed: 2016-09-14.

[27] E. Ohn-Bar, C. Tran, and M. Trivedi, "Hand gesture-based visual user interface for infotainment," in *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 2012, pp. 111–115.

[28] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.

[29] A. W. Fitzgibbon, R. B. Fisher *et al.*, "A buyer's guide to conic fitting," *DAI Research paper*, 1996.

[30] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 21, no. 5, pp. 476–480, 1999.

[31] "fitting an ellipse to a set of data points," http://nicky.vanforeest.com/misc/fitEllipse/fitEllipse.html, accessed: 2016-09-14.

[32] A. Sgroi, K. W. Bowyer, and P. J. Flynn, "The prediction of old and young subjects from iris texture," in *2013 International Conference on Biometrics (ICB)*. IEEE, 2013, pp. 1–5.

[33] A. Zaim, M. Quweider, J. Scargle, J. Iglesias, and R. Tang, "A robust and accurate segmentation of iris images using optimal partitioning," in *18th International Conference on Pattern Recognition (ICPR'06)*, vol. 4. IEEE, 2006, pp. 578–581.

[34] T. Simon, H. Joo, I. Matthews, and Y. Sheikh, "Hand keypoint detection in single images using multiview bootstrapping," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, 2017.