Article

Gesture-Based User Interface for Vehicle On-Board System: A Questionnaire and Research Approach

Krzysztof Małecki * , **Adam Nowosielski** and **Mateusz Kowalicki**

Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, 52 Żołnierska Str., 71-210 Szczecin, Poland; anowosielski@wi.zut.edu.pl (A.N.); mfjpoland@gmail.com (M.K.)
* Correspondence: kmalecki@wi.zut.edu.pl; Tel.: +48-91-449-5661

check for updates

Abstract: Touchless interaction with electronic devices using gestures is gaining popularity and along with speech-based communication offers their users natural and intuitive control methods. Now, these interaction modes go beyond the entertainment industry and are successfully applied in real-life scenarios such as a car interior. In the paper, we analyse the potential of hand gesture interaction in the vehicle environment by physically challenged drivers. A survey conducted with potential users shows that the knowledge of gesture-based interaction and its practical use by people with disabilities is low. Based on these results we proposed a gesture-based system for vehicle on-board system. It has been developed on the available state-of-the-art solutions and investigated in terms of usability on a group of people with different physical limitations who drive a car on daily basis mostly using steering aid tools. The obtained results are compared with the performance of users without any disabilities.

Keywords: natural user interface (NUI); human computer interaction (HCI); gesture-based user interface; gesture recognition; people with physical disabilities; image processing

---

1. Introduction

Natural Interfaces allow human-device interaction through the translation of human intention into devices' control commands, analyzing the user's gestures, body movements or speech. This interaction mode arises from advancements of expert systems, artificial intelligence, gesture and speech recognition, dialog systems, semantic web, and natural language processing, bringing the concept of Intelligent Assistant. More and more often such solutions appear in novel cars and their task is to support drivers in the process of using the Advanced Driver Assistance Systems (ADAS) [1–4], including systems for detecting drivers' fatigue [5,6] and mobile applications for traffic improvement [7].

Conception of Natural User Interface (NUI) assumes free and natural interaction with electronic devices. The interface have to be imperceptible and should base on natural elements [8]. From the car interior perspective it is impossible to incorporate the whole body movements but gestures and speech are well suited for communication with an intelligent assistant or multimedia system. It must be stressed that these communication modes are not well suited for all possible cases. When a driver is listening to the radio or is travelling through a noisy surrounding the voice operated interface will inevitable fail. In many cases both hands are used simultaneously for maneuvering (steering wheel handling and gear shifting) and it is impossible to perform additional gestures.

In the World report on disability, WHO and the World Bank suggests that more than a billion people in the world experience some form of disability [9]. Even as many as one-fifth of them encounter serious difficulties in their daily lives [10]. To improve the every-day activities of people with disabilities new assistive devices and technologies are being introduced. Many of such citizens, living in developed countries, can drive cars on their own thanks to special technical solutions like manipulators or steering aid tools. Still, handling a button console, e.g., of a multimedia system, is out of a range of a handicapped person (due to a distance or form of disability). NUI can prove its value and potential here.

The aim of the paper is to analyse the potential of hand gesture interaction in the vehicle environment by both able-bodied and physically challenged drivers. Test scenarios were defined, consisting of sets of gestures assigned to control multimedia activities in an exemplary vehicle on-board system. Gesture recognition evaluation was conducted for 13 participants, both for the left and right hand. The research was preceded by a survey among participants to check their knowledge of NUI and its practical use. The exact results of the survey are presented in the Appendix A, Figure A1 and discussed in Section 4.1. It is worthy to notice that our experiment was not an invasive medical examination and did not disturb anyone's ethical standards.

The article is organized as follows. Section 2 presents related works. Research on hand operated interfaces is referred here together with its applications in vehicle environment. The main concepts of the interface are presented in Section 3 where the gesture acquisition, gesture recognition and gesture employment are addressed. In Section 4 a characteristic of research group is provided and scenarios of experiments are proposed. Results are presented in Section 5 and discussed in Section 6. The article ends with a summary.

2. Related Works

Vehicle interior designers use research results on potentially possible interaction areas, i.e., the best location for buttons and switches. At present, researchers' attention is focused on the problem of gesture interaction. Driven by latest advancements in pervasive/ubiquitous technology, gesture control using different techniques such as the infrared image analysis from "depth cameras", capacitive proximity sensing, etc., has finally found its way into the car. The principal technology allows to perform finger gestures in the gearshift area [11] or micro-gestures on the steering wheel [12]. It allows also to perform free multi-finger/hand gestures in the open space to control multimedia applications, or to interact with the navigation system [13].

Research work related to gestures includes: intuitiveness of gesture use [14], the quantity of gestures for efficient application [15], or the ergonomics of in-vehicles gestures [16,17]. In [18] authors proposed a classifier based on Approximate String Matching (ASM). Their approach achieves a high performance rates in gesture recognition. In [19] authors explores the use of Kinect body sensor and let users interact with the virtual scene roaming system in a natural way.

In follow of [20], activities carried out by the driver in the car can be divided into three categories: basic tasks, which are directly related to driving, secondary tasks-steering support activities, such as controlling direction indicators, and tertiary tasks-entertainment activities, such as radio control. Today's widespread connections and popular trends, e.g., in social media allow increasing variety of types of these tertiary tasks. Drivers can use gestures for secondary and tertiary tasks (Table 1), and can be distinguished by different criteria. Non-contact gestures were tested in the limited design space of the steering wheel and gear shifting area, but also as gestures in full motion in the air. They are most often made in the area including the rear-mirror, steering wheel and gearshift lever. The most commonly used solutions are based on the use of a depth camera directed towards the gear lever or a center console. Fingers (one or both hands) are used when using gear shifting or loosely resting a hand on it. However, it should be noted that people with disabilities use solutions such as hand gas and brake that restrict movement and tracking of the hands.

Table 1. Classification of in-car gestural interaction methods [21].

Property	Contact Type	Application Type	Body Input	Gesture Size
values	surface-based non-contact	secondary tasks tertiary tasks	1–5 fingers 1–2 hands	micro (finger, facial) macro (full hand)

Holding the gear lever creates problems when trying to make gestures. Grasping it usually means that three fingers from the middle finger to the little finger hold one side of gear lever firmly. This makes them unsuitable for making well-recognized gestures. It is proposed to use those fingers to press the

buttons to be attached to the side of the gear lever (one to three buttons) [21]. They can be used to activate gesture recognition and as modifier buttons to switch between recognition modes. It is worth noting, however, that disabled drivers can use a manual gear lever configured with a pneumatic clutch. This solution requires the use of a sensor located on the gear lever, which in turn makes it difficult to press these buttons. The micro-gestures themselves should be performed with the thumb, forefinger and their combinations. It has been recognized that the actual duration of practical use can be limited: using roads such as motorways requires longer driving periods without shifting gears. Drivers would not have to keep their hands close to the gear lever. It is believed, however, that drivers perceive this as a natural position when the arm rests on the center armrest and the hand on/or near the gear shift lever [21]. For gesture-based interactions, this can provide the driver with a more relaxed basic attitude than performing free-hand gestures in vertically higher interaction spaces, such as in front of the center console. Importantly, cars for people with disabilities are usually equipped with an automatic gear shifting, which excludes the need to change gears, and in turn such position is not natural.

It is worthy to notice that the recommendations of the general directions of traffic is necessity to use both hands at the wheel to drive if it is possible. However, the controlling of various on-board systems require drivers to take away one of their hands from the wheel to press a specific button or knob. In this context, making a gesture that does not require direct focus on the button that we want to use at the moment seems more natural. The certain gesture can be performed without distracting the driver's attention focused on observing the traffic situation.

Usage of the camera mounted inside a vehicle can also serve other purposes besides the natural user interface. For example, it allows to analyze additional hand activities or important objects in the car that are important for advanced driver assistance systems. In addition, it allows to specify the system user (driver or passenger), which can consequently be used for further penalisation [22].

It was found that the implementation of the gesture-based interfaces in the region bounded by a "triangle" of steering wheel, rear mirror, and gearshift is high acceptable by users [21]. Data evaluation reveals that most of the subjects perform gestures in that area with similar spatial extent, on average below 2 s. In addition to allowing the interface location and non-invasive interface to be customized, the system can lead to further innovative applications, e.g., for use outside of the vehicle. The use of cameras and a gesture-based interface can also generate some potential cost and security benefits. The placement of the camera itself is relatively easy and relatively cheap. The use of non-contact interfaces and their impact on the visual and mental load of the driver require thorough examination. For example, accurate coordination may be less needed when using a contact-less interface as opposed to using a touch screen, which can reduce the view on natural user interface [22].

Valuable information on the most desirable functions that drivers would like to use with gestures is provided by [22]. The authors conducted a survey of 67 participants of different ages, sex and technological similarities. In accordance with latest research of short-term and working memory, which notified that the mental storage capacity is about 5 gestures [23,24], participants were asked about their preferred 5 functions They would like to control using a gesture-based in-car interface. They were also asked which finger/hand gesture They would like to use for each of these functions, and participants had then to get into a parked car and perform the particular gestures in a static setting, i.e., without driving the car. By applying approval voting of 18 driving experts (car drivers with high yearly mileage), the 5 functions receiving the most votes were chosen. In a later improvement, the one or other gesture was added or replaced to the gestures selected before to make the whole setting more homogeneous and to reduce potential side effects such as non-recognition and poor designs of the gesture-to-function mapping to a minimum [25]. Furthermore, it is evident that free-hand gestures induce fatigue with prolonged usage [25]. On account of these, safety related tasks should not be solely assigned to gestures.

Majority of car manufacturers offer touchscreen multimedia control systems. In the latest solutions, these systems are based on two touch screens and voice assistant [26]. On premium vehicles, solutions based on the recognition of gestures can be found [27]. In the case of BMW, the Gesture Control system allows to accept or reject a phone call, increase or decrease the volume, and define a favorite function with an assignment to the victory gesture.

In recent years, wireless approaches of human interaction with devices are also considered, both in the context of gesture control [28] and human activities recognition [29,30]. The idea of wireless gesture control is to utilize the RFID signal phase changes to recognize different gestures. Gestures are performed in front of tags deployed in an environment. And then, the multipath of each tag's signal propagation is changed along with hand movements, which can be captured by the RFID phase information. The first examination brings promising results [28]. It seems that smart watches and wristbands, equipped with many sensors, will become another way of gesture control [31]. Solutions of this kind are supported by the deep-learning methods [32,33]. However, it should be noted that it is customary to wear smart watches and smart wristbands on the left hand, which makes the use of such equipment for drivers reasonable for vehicles with the steering wheel placed on the right side. This is an interesting research area, but the authors of this article are not currently concerned with it.

Gesture recognition technology is widely expected to be the next generation in-car user interface. A market study conducted in 2013 by IHS Automotive examined gesture-recognition technology and proximity sensing. According to IHS, the global market for automotive proximity and gesture recognition systems that allow motorists to control their infotainment systems with a simple wave of their hand will grow to more than 38 million units in 2023, up from about 700,000 in 2013. Automakers including Audi, BMW, Cadillac, Ford, GM, Hyundai, Kia, Lexus, Mercedes-Benz, Nissan, Toyota and Volkswagen are all in the process of implementing some form of gesture technology into their automobiles [34].

3. Proposed Approach of Gesture-Based Interaction Interface

While driving, a driver is obliged to perform maneuvers safely, paying attention to other road users. A safety critical actions include: acceleration, breaking and change of the movement direction (see Figure 1 for clarity). These actions are performed with dedicated controllers, i.e., pedals and steering wheel or in the case of motor impaired drivers with the use of special assistive tools. Many cars still are equipped with a manual gearbox and a clutch. The direction indicator is also recognized as an important mechanism crucial for driving. Some researchers treat such activities as secondary tasks-steering support activities [20]. Nevertheless, They constitute important tasks for safety reasons and we include them in one box.
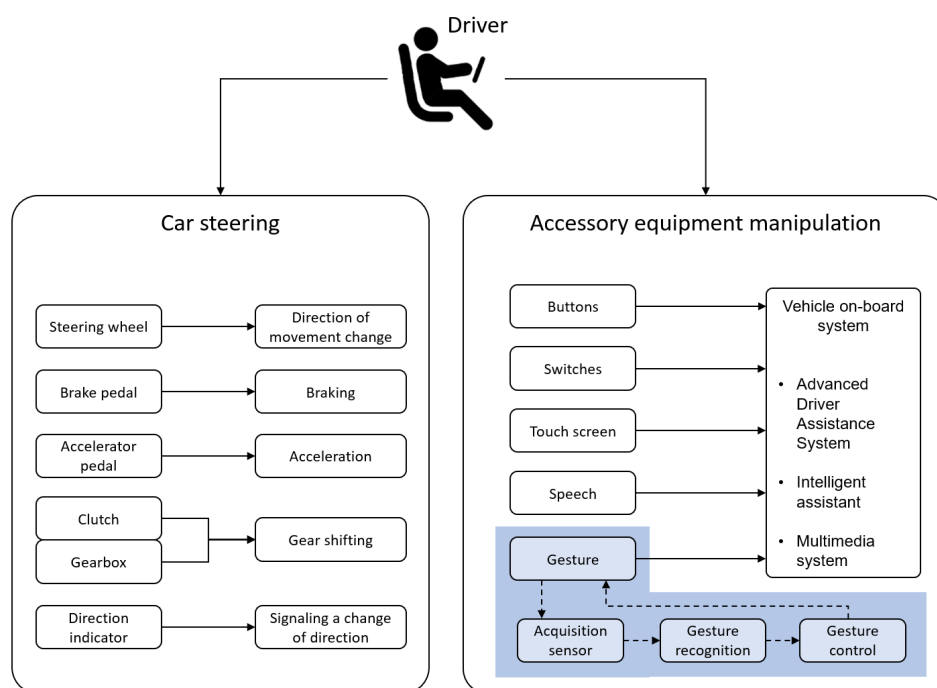

Figure 1. A general scheme of proposed approach.

On the right hand side of Figure 1 five steering mechanism are listed: buttons, switches, a touch screen, speech and gesture. These modes are used for manipulation of accessory car equipment supporting a driver (tertiary tasks). The gesture is highlighted in the diagram since it is the scope of our investigation. In modern solutions, a gesture-based system requires a hardware for acquisition, a dedicated engine for gesture recognition and a feedback for control. All of the components are analysed in the following subsections.

3.1. Gesture Acquisition

As presented in the "Related works" section gestures can be acquired using different technologies. Wide use of devices such as smartwatches, equipped with additional sensors (such as accelerometers or gyroscopes), can significantly facilitate the gesture recognition. There are also other wireless technologies (e.g., RFID), but we focus our attention on computer vision, emphasizing its non-invasive nature.

The interior of the car is a difficult environment because the lighting conditions are very different [6,35]. To recognize the actions performed by a driver the scene appearance should be acquired first. Typical capture devices, i.e., cameras that operate in the visible spectrum (electromagnetic waves from about 380 to 780 nm of wavelengths), can work properly with good lighting conditions. Despite the fact that often the parameters of such cameras are very high, in terms of resolution, sensitivity and dynamic range, They suffer from poor lighting condition encountered in road conditions. It is obvious that the driver's face while driving cannot be illuminated with any sort of additional light source and for this reason other acquisition mechanisms are employed. Insensitive to the light source, its power or type is thermal imagery technique. It registers the temperature of objects giving the thermal image of the scene. Despite the fact that this technique can be successfully used in human-computer interfaces (e.g., [36]) the price of thermal cameras is still a serious barrier.

In view of the above, the most appropriate solution seems to be the depth map technique. Here, the distance information from the scene objects to the camera is represented as pixels brightness. There are three major approaches to acquire depth information: stereo vision, structured-light, and time-of-flight. The first requires a stereoscopic camera or a pair of calibrated cameras. Due to the use of regular cameras, this approach is not suitable for use in a vehicle. In the structured light approach, an infrared (invisible for the human eye) light pattern is projected on the scene. Distorted on 3-D objects, it is registered by the IR camera offering, after geometric reconstruction, the 3D information about the observed scene. The time-of-flight approach is based on the measurements of the time required for a light signal to travel from the source to the scene and back to the sensor.

For our study we have employed the Intel RealSense sr300 camera (presented in Figure 2). The considered alternatives to this solution are MS Kinect and Leap Motion. The first and the second generation of Kinect hardware was excluded due to its dimensions. What is more, both versions are now considered obsolete and are not manufactured anymore. The latest release of Kinect (the Azure Kinect released this-2020-year) strictly base its operation on cloud computing, which exclude its usage on sparsely populated areas outside network coverage. The Leap Motion could have been a good alternative here for its small size and quite decent operation. However, the solution offers a limited number of movement patterns recognition (i.e., swipe, circle, key tap and screen tap).

In view of the above, the Intel RealSense provides the best suited solution for our research. It has a compact size suitable for the in-car use. The sensor includes both infrared and RGB colour cameras, IR laser projector and a microphone array. It captures a depth image of the scene in good and poor lighting conditions from a relatively close distances occurring in the car cabin. It offers 60 fps in 720 p resolution and 30 fps in 1080 p. The hardware load from today's perspective also seems to be undemanding. The sensor requires at least 6th Generation Intel® Core™ Processor and 4 GB of memory. With a price slightly exceeding $100 the component of this kind seems to be well suited for the widespread use.

We performed all experiments on the laptop equipped with 8th generation Intel® Core™ i5 Processor, 8 GB of memory and observed no delays during the operation.
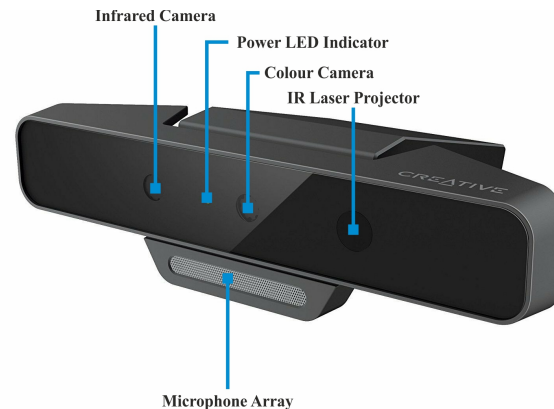
Figure 2. Intel RealSense sr300 sensor [own work based on https://us.creative.com/p/web-cameras/blasterx-senz3d].

3.2. Gesture Recognition

Two types of hand gestures are defined in the scientific literature, namely: static and dynamic [37]. Static gestures are sometimes called as postures while dynamic-as hand motions. A Static gesture denotes a specific orientation and position of hand in the space. The gesture lasts a specific period of time without any movement associated with it. A thumb up as an OK symbol is a good example here. A dynamic gesture requires a specific movement performed during a predefined time window. A hand moving forward in the air as if pressing a button or waving of hand are examples of dynamic gestures.

Figure 3 presents a set of established gestures we employed in our investigation. A gesture number one can be described as a movement of the hand, facing the sensor, to the left and immediately back to the starting position. Similar is the case with gesture number two, but the movement is to the right. Gestures number 3 and 4 mean respectively: a movement of the hand, facing the sensor, downwards and immediately back to the starting position and a movement of the hand, facing the sensor, upwards and immediately back to the starting position. These four gestures constitute for a typical swipe mechanism.

Figure 3. A set of gesture included in presented research. (Based on [38]).

The gesture number 5 means: move a hand close to the sensor and move it back in a way that imitates pressing a button. All the first 5 gestures presented in Figure 3 are typical dynamic gestures. The definition of gesture number 6 is: show a clenched hand towards the sensor. This is a static gesture similarly as the last one. The 7th gesture is defined as presentation of the hand with index and middle finger upwards with others hidden.

We have chosen for our investigations only those gestures that seems to be easy to perform, are natural, and engage small number of muscles. These are four directional swipe gestures, a push gesture and two static gestures (clenched fist and victory gesture). It must be noticed that many gestures are not well suited for the in-car environment. They require more space to be performed properly or engage too many muscles. The latter is especially important for people with physical limitations. For example, a "thumb down" gesture is performed with a closed hand and straightened thumb, but the rotary arm movement is required so that the thumb points down. Other gestures, like "spread fingers" might be extremely difficult for people with disabilities. We believe that our selected gestures are the easiest but for sure, in future, the user specific adaptation should be considered.

Gesture recognition is still an active research activity with many established solutions available for researchers. In our proposal we based the gesture recognition module on libraries provided with the Intel RealSense sr300. The hardware details of this solution has been provided in Section 3.1. The gesture recognition mechanisms are based here on skeletal and blob (silhouette) tracking. The most advanced mode is called "Full-hand" and offers tracking of 22 joints of the hand in 3D [38]. This mode requires the highest computational resources but offers the recognition of pre-defined gestures. The second is the "Extremities mode" and it offers tracking of 6 points: the hand's top-most, bottom-most, right-most, left-most, center and closest (to the sensor) points [38]. Other modes include "Cursor" and "Blob". They are reserved for the simplest cases when central position of hand or any object is expected to be tracked.

3.3. Gesture Employment

Having a recognized gesture one can employ it for steering or communication purposes. Since our study focuses on multimedia system control by persons with disabilities we connected each gesture to a specific function. Our prototype vehicle on-board system is operated using gestured presented in Figure 3 according to the scheme:

- g-1: previous song/station,
- g-2: next song/station,
- g-3: previous sound source,
- g-4: next sound source,
- g-5: turn on the multimedia system,
- g-6: turn off the multimedia system,
- g-7: a favourite function.

As can be observed the dynamic swipe gestures are used for navigation through sound source and songs or the radio stations. The system can be lunched with the in the air push gesture and switched off with the fist gesture. A sign of victory is reserved for a favourite function.

4. Evaluations

Evaluation of HCIs in general is user-based. During the experiments participants have different tasks constituting a set of actions to be performed. These actions are called scenarios. The correctness of action execution is verified and some performance measures are gathered.

In our experiments we investigated the ease of gesture-based steering by people with physical disabilities. Figure 4 presents experiment setup. For safety reasons, we decided not to conduct experiments in traffic. Operation of both hands have been evaluated and because we used cars with the steering wheel located on left-hand side, when tested the left-hand operation a driver was moved to the passenger seat. Depending on own preferences or degree of disability each user chose one of the two options for mounting the sensor: in the center dashboard (Figure 4 top) or near the armrest and gear shift (Figure 4 bottom). In the last case driver is able to perform gestures in horizontal plane, which is important for people with limited movements in the shoulder area. The sensor itself was attached with a double-sided tape. The exact location is denoted by the red circle.

In the examined version of the prototype the interface was not integrated with the car's console. We assume that the control would take place via the CAN bus. The application would send the information to the multimedia system which functions should be activated. Currently, the application gives the message in JSON format and this is used for validation. The dedicated application is launched on laptop in the car environment.

In the next section we briefly characterise our study group. Then, we provide description of the experiments.
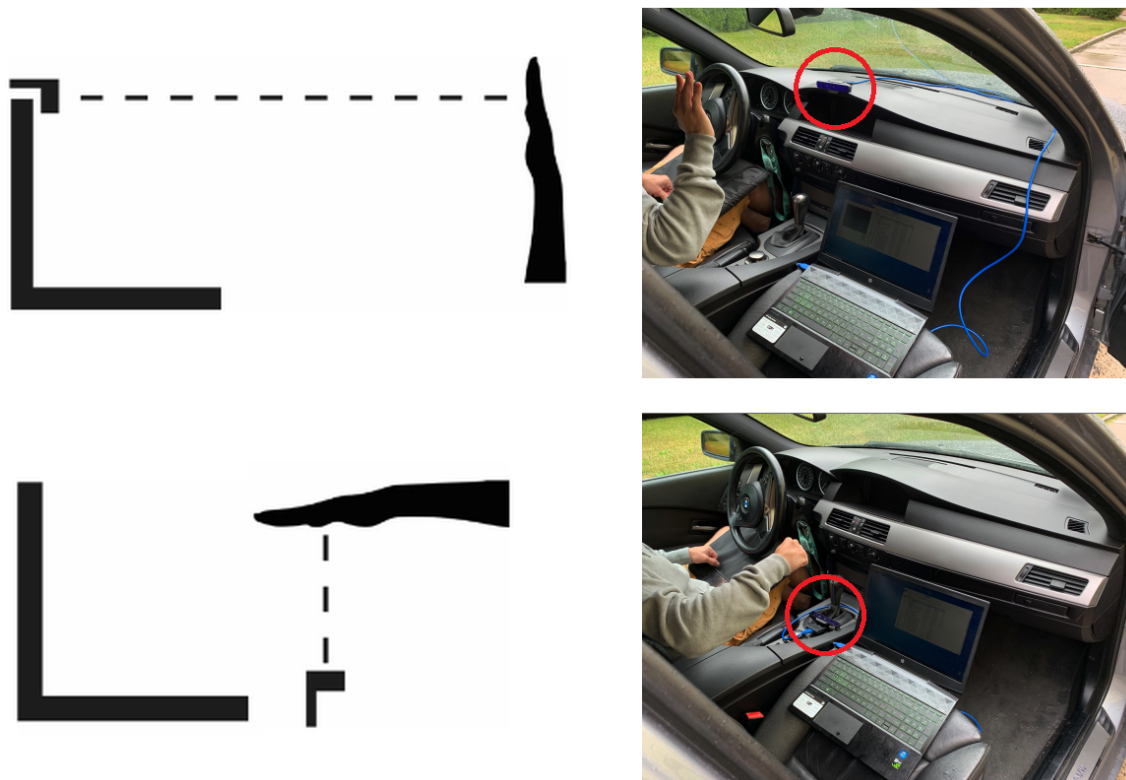
Figure 4. The sensor location: in the center dashboard (**top**), near the armrest and gear shift (**bottom**).

4.1. Characteristic of Research Group

Research study was conducted of two stages: (1) a NUI topic e-survey and (2) a set of gesture recognition experiments. A total of 40 voluntary participants took part in the first stage, including 23 fully physically fit people and 17 with mobility disabilities, regardless of gender. The survey results are presented in the Appendix A, Figure A1. The prepared survey included both questions about the determination of physical characteristics (e.g., degree of disability, type of disability) and determining knowledge about NUI. Most of the respondents (84%) were active vehicle drivers and 69% of all participants think that technology can significantly help decreasing the disability barrier during vehicle using. 74% of our respondents were severe disabled and 22% of them had a moderate disability. The remaining 4%, They were participants with a light disability. All of our respondents have any of steering aid tools in their cars, which means that their cars had to be specially adapted. Half of respondents knew what NUI means and They could identify popular NUI types (e.g., gesture-based recognition interface or voice-based interface). But, very important is, that most of them (52%) did not use NUI technology before. Thanks to this, these people could be invited for further research on GRI experiments.

Thus, a total of 13 voluntary participants took part in the second stage, including 4 fully physically fit people and 9 with mobility disabilities, regardless of gender (Table 2). It must be noticed that all disabled participants were physically independent and all of them had all upper and lower limbs. The age distribution of 13 people who took part in the research experiment is as follows: under 30 (u_01, u_02, u_03, u_08, u_11), 31–40 (u_04, u_09, u_10, u_12) and over 41 (u_05, u_06, u_07, u_13). Below, there is a brief description of people with disabilities.

U_02 suffers from cerebral palsy, four-limb paresis with a predominance of the left side. The left hand of the subject is spastic. Three weeks before testing, this hand was operated on for wrist flexor tenotomy. The user is not able to activate by this hand any of car switches in a conventional way. Right hand is functional, and the strength in both limbs according to the Lovet scale is 4/5. Research for the right hand did not show any differences from the test scenarios. The results are different for the

left hand. Some gestures are not recognized immediately and They need to be repeated several times. They also require additional focus of user to the movements performed. The left hand during the tests tired quickly, increased spasticity was observed. This could be due to the patient's illness, as well as the surgery three weeks before the test.

U_03 suffers from a (lat. myelomeningocele) in the lumbosacral segment, between the L5 and S1 vertebrae. He has paresis of lower limbs, so he uses the manual mechanism of gas and brake. He has a problem with reaching for the side window opening buttons and controlling the multimedia system and air conditioning (he controls them only when the vehicle is not moving). This is due to the location of the buttons. For this person, a more convenient location of the depth camera turned out to be place near the gear shift lever, due to having in his car a specialized equipment that requires constant use of the hand to control the gas and brake. The same condition occurs in u_04, ..., u_08.

U_09 suffers from a rare disease-type 4 of mucopolysaccharidoses, and because of this she moves with crutches every day. The biggest problem for this participant is getting in and out of the vehicle due to degeneration of the hip joints. She quickly got used to the gestures of the gesture-based system and it became intuitive for her. She preferred the location of the camera near the gear shift lever.

U_10 suffers from paralysis of the shoulder plexus and forearm of the right hand, which resulted in her paresis. Left hand fully functional. The participant has the biggest problem with controlling on-board devices, such as a multimedia system, air conditioning panel and wiper handle. The participant uses the steering wheel knob. Multimedia control is possible thanks to the multifunctional steering wheel. She was able to perform certain gestures with her right hand during the tests, but the system had a problem recognizing them. Directional gestures were slow. Static gestures in most cases were recognized except for the favorite gesture. Left hand tests were performed correctly. The participant said that such a solution in the right form would be very helpful for her. The participant's illness allows her to make gestures only in the horizontal plane.

Half of respondents preferred to have the sensor of gesture recognition interface located near the armrest and gear shift (to be able to perform horizontal gestures) and another half, preferred to have it in the center dashboard (Table 2 and Figure 4).

Table 2. Characteristics of research group.

User	Fully Physical Fit	With Disability	[M]ale [F]emale	Preference for Placing the Sensor in the Center Dashboard	Preference for Placing the Sensor Near the Armrest and Gear Lever
u_01	+		M		+
u_02		+	M	+	
u_03		+	M	+	
u_04		+	F		+
u_05		+	M	+	
u_06		+	M	+	
u_07		+	M	+	
u_08		+	M		+
u_09		+	F		+
u_10		+	F	+	
u_11	+		M		+
u_12	+		F		+
u_13	+		F		+

4.2. Scenarios

We have prepared six scenarios to be performed by each participant. Each scenario has been realised twice by each participant, using left and right hand to perform gestures. Each scenario assumed a typical interaction of vehicle on-board multimedia system. It contains a sequence of gestures from a set presented in Figure 3. Scenarios are as follow:

- S-1: g-5, g-3, g-2, g-4, g-1;
- S-2: g-4, g-1, g-5;
- S-3: g-5, g-7, g-6;

- S-4: g-2, g-7, g-4, g-4, g-4, g-3, g-3, g-3, g-3, g-5;
- S-5: g-4, g-2, g-2, g-2, g-1, g-1;
- S-6: g-5, g-6, g-5, g-6, g-5, g-6.

In the first scenario (S-1), a participant is expected to turn on the multimedia system with a push gesture (g-5), navigate to the previous sound source with a swipe down gesture (g-3), change the song to the next with a swipe right gesture (g-2), change the sound source with the swipe up gesture (g-4), and finally navigate to the previous song with a swipe left hand gesture (g-1).

All other scenarios (S-2, . . . , S-6) can be easily read in a similar way. We preferred to exploit different scenarios instead of repeating the same scenario few times. Such approach is more similar to the ordinary, everyday operation.

5. Results

The detailed results of conducted experiments are presented in Figures 5–10. Each figure presents two charts, respectively for the left and right hand, positioned on the corresponding side. The abscissa correspond to a specific user and the ordinate lists subsequent gestures in a given scenario. The green 'v' mark denotes a successful gesture and the red 'x'-accordingly a failure. The yellow triangle with an exclamation mark '!' represents an undesired or unintentional action. Usually, after the incorrect gesture, the user has undone the action, which is represented in the charts by a blue reload mark. For example, when accidentally a gesture for the next song was performed or recognized, user made a gesture for the previous song.

The ordinate which lists subsequent gestures in a given scenario, shows by default also the number of gestures required. When an additional or unrecognized gesture appeared, it is represented by an additional symbol between the expected gestures.

The accuracy of gesture recognition by particular scenarios is presented in Figures 11 and 12 shows the accuracy from the user's perspective. Both graphs depict the aggregated information accordingly for the left and right hand. The values have been calculated as the ratio of correct gestures (performed) to the total number of expected correct gestures. The percentages aggregate the detailed results presented in Figures 5–10.
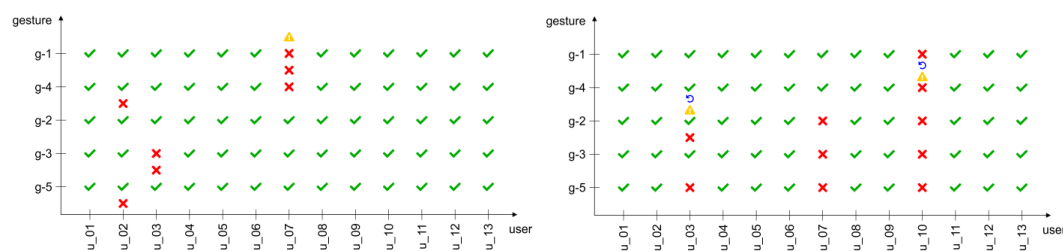

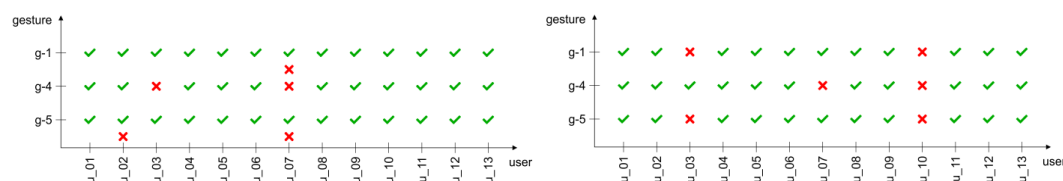**Figure 5.** Detailed results for the S-1 scenario.
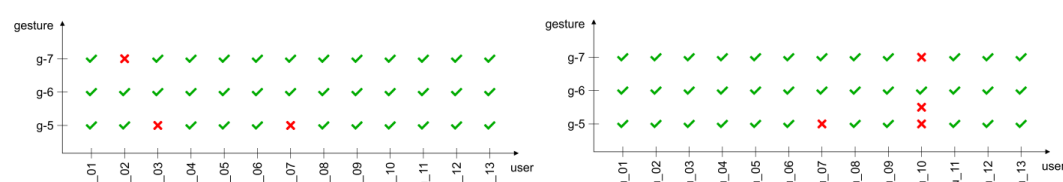

Figure 6. Detailed results for the S-2 scenario.

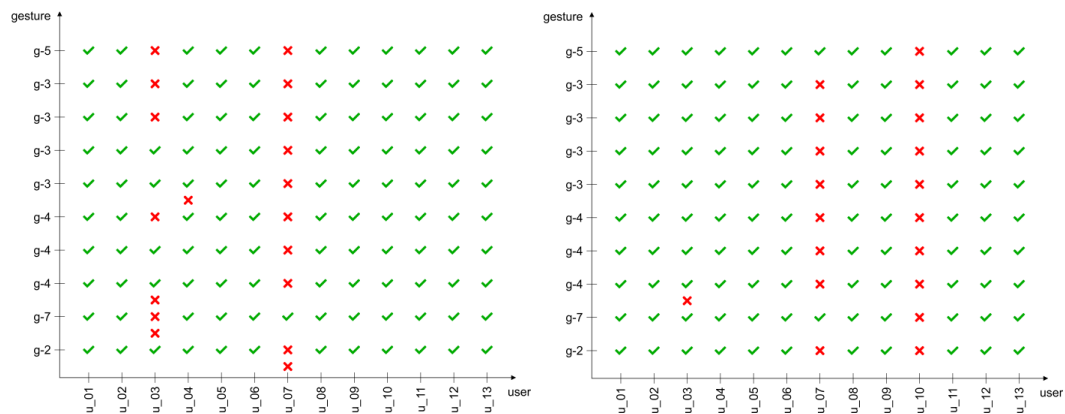**Figure 7.** Detailed results for the S-3 scenario.

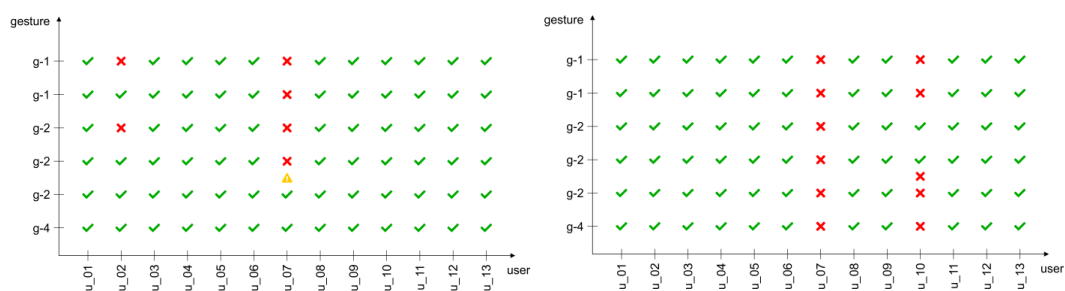**Figure 8.** Detailed results for the S-4 scenario.

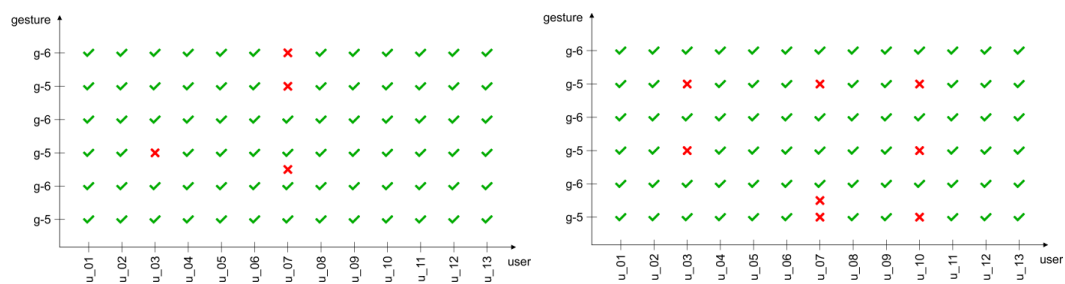**Figure 9.** Detailed results for the S-5 scenario.

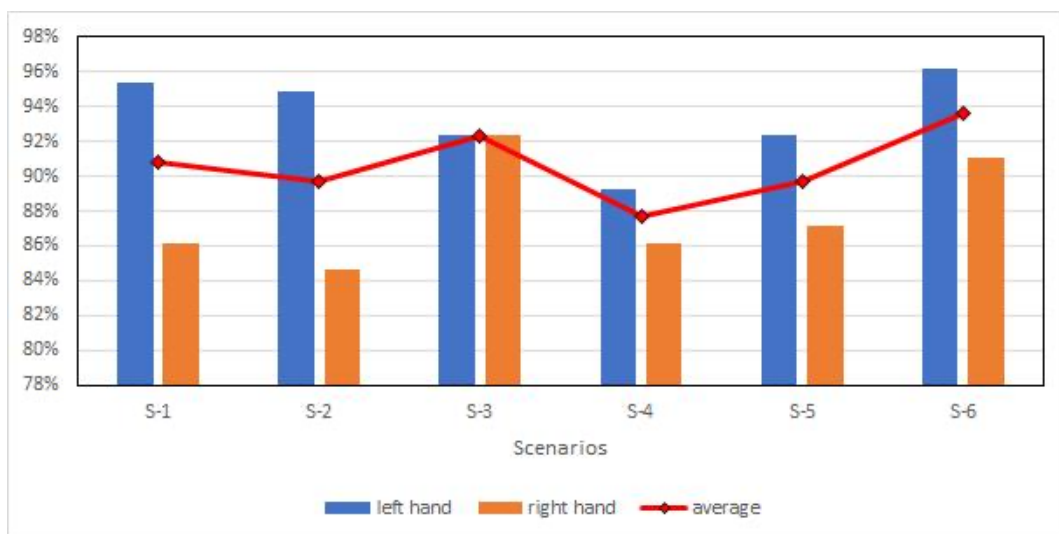**Figure 10.** Detailed results for the S-6 scenario.

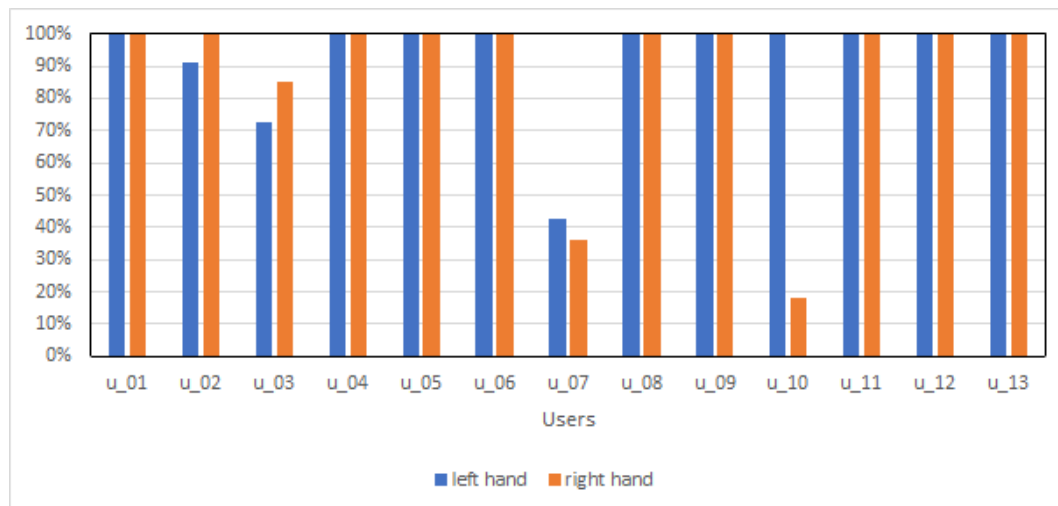**Figure 11.** Accuracy of gesture recognition by scenario.

Figure 12. Accuracy of gesture recognition by user.

6. Discussion

From the results presented in Section 5 it is evident that gesture-based interaction has potential for the in-vehicle systems operated by people with disabilities. From a group of thirteen participants as many as nine have performed all scenarios perfectly for both hands. Two, from the remaining four, performed perfectly all the actions but only for one hand. Participant u_02 has encounter some problems with the left hand and participant u_10 has failed with the right hand. Problems They have encountered during gesture steering are the result of their movement problems. Participant u_02 suffers from four-limb paresis with a predominance of the left side and participant u_10 suffers from brachial plexus and forearm palsy of the right hand. Relatively high results were obtained by user u_03. He suffers from myelomeningocele in the lumbosacral segment and has problems with opening the windows in the car. Nevertheless, he obtained the gesture interaction accuracy equal 73% for the left hand and 85% for the right hand. Gesture-based interaction prove to be too big a challenge for a participant u_07. The person suffered a spinal cord injury, is of the age of 59 and is skeptical of new technologies. It was very difficult for that person to get into gesture-based interaction and none of the test has been carried out completely correctly.

Analysis of individual scenarios shows that scenario S-6 was the least problematic with averaged score equal 93.59%. The most difficult, without distinction of the used hand, was scenario S-4 (87.69%). It must be noted, that this scenario was the longest one. Interestingly, the aggregated results from all experiments show that performance of the left hand was slightly better then that of the right hand (92.77% to 87.65%). Individually, scenario S-6 was the easiest for the left hand (96.15%) and S-3 for the right hand (92.31%). The worst result was registered for S-2 scenario operated with the right hand (84.62%).

7. Conclusions

In the article the touchless interaction with vehicle on-board systems is presented. In particular, the focus is on gesture-based interaction. The survey of 40 people was conducted collecting their opinions and their knowledge of the gesture control systems. It is important that the survey was attended by people who are both drivers and non-drivers, including physically challenged people.

A set of gestures was established. Later on, we developed six scenarios containing chosen gestures from a set and we conducted experimental studies with the participation of 13 volunteers. All experiments were carried out for each hand separately, giving the overall result of 73% for the left hand and 85% for the right hand.

Our study has significant practical importance. It shows that the location of the depth sensor is important for the comfort of the user (driver) of the gesture recognition system. Some of our

respondents indicated that the placement of the sensor in the dashboard causes rapid hand fatigue and thus discourages the use of the system. The problem was solved by sensor relocation to the space between the gearshift lever and the armrest.

As research shows, there is an interest in gesture-based systems, and further research is necessary to fully understand the needs of drivers. It is worthy to establish standard gestures recognized by various car multimedia system manufacturers.

Abbreviations

The following abbreviations are used in this manuscript:

ADAS   Advanced Driver Assistance Systems
HCI   Human-Computer Interface/Interaction
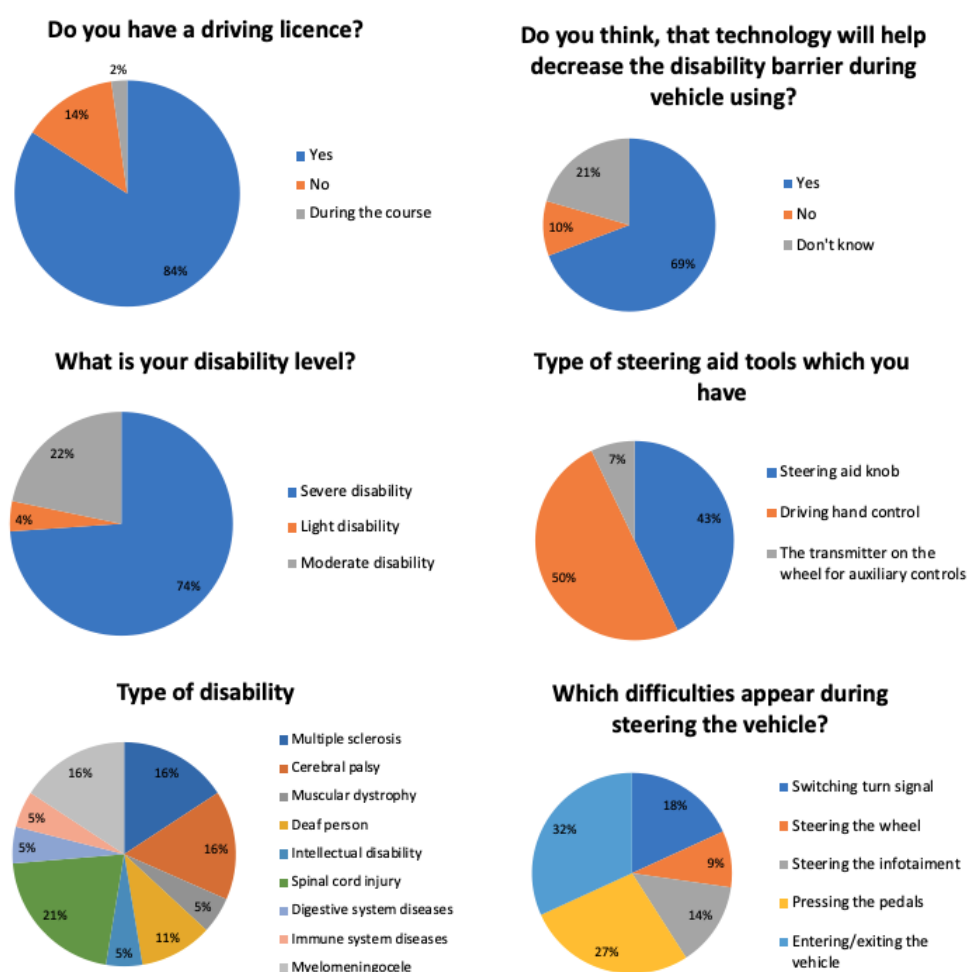NUI   Natural User Interface
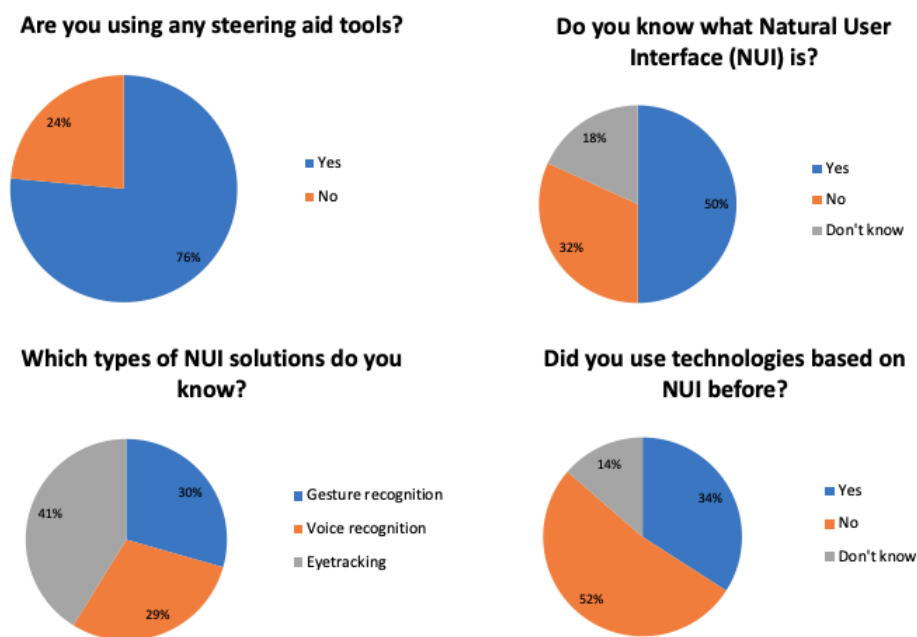
Appendix A

Figure A1. *Cont*.

Figure A1. Survey results.

References

1. Cyganek, B.; Gruszczyński, S. Hybrid computer vision system for drivers' eye recognition and fatigue monitoring. *Neurocomputing* **2014**, *126*, 78–94. [CrossRef]
2. Toledo, T.; Musicant, O.; Lotan, T. In-vehicle data recorders for monitoring and feedback on drivers' behavior. *Transp. Res. Part C Emerg. Technol.* **2008**, *16*, 320–331. [CrossRef]
3. Forczmański, P.; Małecki, K. Selected Aspects of Traffic Signs Recognition: Visual versus RFID Approach. In *Activities of Transport Telematics. Communications in Computer and Information Science*; Mikulski, J., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 395, pp. 268–274.
4. Nowosielski, A.; Małecki, K.; Forczmański, P.; Smoliński, A.; Krzywicki, K. Embedded Night-Vision System for Pedestrian Detection. *IEEE Sensors J.* **2020**, *20*, 9293–9304. [CrossRef]
5. Craye, C.; Rashwan, A.; Kamel, M.S.; Karray, F. A multi-modal driver fatigue and distraction assessment system. *Int. J. Intell. Transp. Syst. Res.* **2016**, *14*, 173–194. [CrossRef]
6. Małecki, K.; Nowosielski, A.; Forczmański, P. Multispectral Data Acquisition in the Assessment of Driver's Fatigue. In *Smart Solutions in Today's Transport. Communications in Computer and Information Science*; Mikulski, J., Ed.; Springer: Cham, Germany, 2017; Volume 715, pp. 320–332.
7. Iwan, S.; Małecki, K.; Stalmach, D. Utilization of Mobile Applications for the Improvement of Traffic Management Systems. In *Telematics-Support for Transport. Communications in Computer and Information Science*; Mikulski, J., Ed.; Springer: Berlin/Heidelberg, Germany, 2014; Volume 471, pp. 48–58.
8. Giorio, C.; Fascinari, M. *Kinect in Motion—Audio and Visual Tracking by Example*; Packt Publishing: Birmingham, UK, 2013.
9. World Health Organization. *The World Bank. World Report on Disability*; World Health Organization: Geneva, Switzerland, 2011.
10. Jensen, L.; Jordansen, I.K.; Sejer Jakobsen, H.; Spure, J.; Simonsen, J.; Nielsen, P. Usercommunities for innovation of assistive technology. Results from a demonstrations project about User-driven innovation. *Assist. Technol. Res. Ser.* **2011**, *29*, 987–994.
11. Riener, A.; Wintersberger, P. Natural, intuitive finger-based input as a means of in-vehicle information system operation. In Proceedings of the Auto-UI, Salzburg, Austria, 29 November–2 December 2011.
12. Fang, V.; Ainsworth, L. Designing and Rapid Prototyping a Gesture-Enabled Steering Wheel. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI'12), Portsmouth, NH, USA, 17–19 October 2012; pp. 53–56.

13. Ohn-Bar, E.; Tran, C.; Trivedi, M. Hand gesture-based visual user interface for infotainment. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Portsmouth, NH, USA, 17–19 October 2012; pp. 111–115.

14. Zobl, M.; Geiger, M.; Bengler, K.; Lang, M. A usability study on hand gesture controlled operation of in-car devices. In Proceedings of the International Conference on Human-Computer Interaction HCI2001, New Orleans, LA, USA, 5–10 August 2001.

15. Chung, C.; Rantanen, E. Gestural interaction with in-vehicle audio and climate controls. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Berlin, Germany, 27 September–1 October 2010; SAGE Publications: Los Angeles, CA, USA, 2010; Volume 54, pp. 1406–1410.

16. Riener, A. Gestural interaction in vehicular applications. *Computer* **2012**, *45*, 42–47. [CrossRef]

17. Yousaf, M.H.; Habib, H.A. Adjustable pedals by gesture recognition: A novel approach for user interface in automotive. *Middle-East J. Sci. Res.* **2012**, *11*, 1575–1583.

18. Alonso, D.G.; Teyseyre, A.; Soria, A.; Berdun, L. Hand gesture recognition in real world scenarios using approximate string matching. *Multimed. Tools Appl.* **2020**, *79*, 20773–20794. [CrossRef]

19. Li, F.; Feng, J.; Fu, M. Research on Natural Human-Computer Interaction in Virtual Roaming. *J. Phys. Conf. Ser.* **2020**, *1518*, 012022. [CrossRef]

20. Pickering, C.A.; Burnham, K.J.; Richardson, M.J. A research study of hand gesture recognition technologies and applications for human vehicle interaction. In Proceedings of the 3rd Institution of Engineering and Technology Conference on Automotive Electronics, Warwick, UK, 28–29 June 2007; pp. 1–15.

21. Lindemann, P.; Koelle, M.; Stockinger, T.; Diewald, S.; Kranz, M. A Survey on Gestural Interaction in the Car. In Proceedings of the 15th International Conference on Computer Aided Systems Theory, Las Palmas de Gran Canaria, Spain, 8–13 February 2015; pp. 246–247.

22. Riener, A.; Ferscha, A.; Bachmair, F.; Hagmüller, P.; Lemme, A.; Muttenthaler, D.; Pühringer, D.; Rogner, H.; Tappe, A.; Weger, F. Standardization of the in-car gesture interaction space. In Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Eindhoven, The Netherlands, 28–30 October 2013; pp. 14–21.

23. Cowan, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav. Brain Sci.* **2001**, *24*, 87–114. [CrossRef] [PubMed]

24. Kamiński, J.; Brzezicka, A.; Wróbel, A. Short-term memory capacity (7 ± 2) predicted by theta to gamma cycle length ratio. *Neurobiol. Learn. Mem.* **2011**, *95*, 19–23. [CrossRef] [PubMed]

25. Norman, D.A. Natural user interfaces are not natural. *Interactions* **2010**, *17*, 6–10. [CrossRef]

26. MobileGeeks—Mercedes Benz MBUX 2018 A Class Detailed Walk Through. Available online: https://www.mobilegeeks.com/review/mercedes-benz-mbux-2018-class-detailed-walk/?fbclid=IwAR3-OTuUyJNjvd1AzJLKd82Nw7ZjTEAd8XVpPchaMxG-eBvGmlpDbphNsJE (accessed on 27 July 2020).

27. BMWBLOG—OPINION: BMW Gesture Control Really Improves in-Car Operation. Available online: https://www.bmwblog.com/2019/11/08/opinion-bmw-gesture-control-really-improves-in-car-operation/ (accessed on 27 July 2020).

28. Zou, Y.; Xiao, J.; Han, J.; Wu, K.; Li, Y.; Ni, L.M. Grfid: A device-free rfid-based gesture recognition system. *IEEE Trans. Mob. Comput.* **2016**, *16*, 381–393. [CrossRef]

29. Pradhan, S.; Chai, E.; Sundaresan, K.; Qiu, L.; Khojastepour, M.A.; Rangarajan, S. Rio: A pervasive rfid-based touch gesture interface. In Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking, Salt Lake City, UT, USA, 9–13 October 2017; pp. 261–274.

30. Du, Y.; Lim, Y.; Tan, Y. A Novel Human Activity Recognition and Prediction in Smart Home Based on Interaction. *Sensors* **2019**, *19*, 4474. [CrossRef] [PubMed]

31. Jiang, S.; Lv, B.; Guo, W.; Zhang, C.; Wang, H.; Sheng, X.; Shull, P.B. Feasibility of wrist-worn, real-time hand, and surface gesture recognition via sEMG and IMU Sensing. *IEEE Trans. Ind. Inform.* **2017**, *14*, 3376–3385. [CrossRef]

32. Kasnesis, P.; Chatzigeorgiou, C.; Toumanidis, L.; Patrikakis, C.Z. Gesture-based incident reporting through smart watches. In Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kyoto, Japan, 11–15 March 2019; pp. 249–254.

33. Ahn, H.J.; Kim, J.S.; Shim, J.Y.; Kim, J.S. Hand gesture recognition for doors with neural network. In Proceedings of the International Conference on Research in Adaptive and Convergent Systems, Krakow, Poland, 20–23 September 2017; pp. 15–18.

34. Avnet—Gesture Recognition, Proximity Sensors Drive Advances in Automotive Infotainment. Available online: https://www.avnet.com/wps/portal/us/resources/article/gesture-recognition-proximity-sens ors-drive-advances-auto-infotainment/ (accessed on 27 July 2020).

35. Molchanov, P.; Gupta, S.; Kim, K.; Pulli, K. Multi-sensor system for driver's hand-gesture recognition. In Proceedings of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, Ljubljana, Slovenia, 4–8 May 2015; Volume 1, pp. 1–8.

36. Nowosielski, A.; Forczmański, P. Touchless typing with head movements captured in thermal spectrum. *Pattern Anal. Appl.* **2019**, *22*, 841–855. [CrossRef]

37. Rautaray, S.S.; Agrawal, A. Vision based hand gesture recognition for human computer interaction: A survey. *Artif. Intell. Rev.* **2015**, *43*, 1–54. [CrossRef]

38. Intel® RealSense™ SDK 2016 R2 Documentation. Available online: https://software.intel.com/sites/landi ngpage/realsense/camera-sdk/v1.1/documentation/html/index.html?doc_devguide_introduction.html (accessed on 27 August 2020).