# Free-hand human-machine interaction in vehicles

Tomaž Čegovnik and Jaka Sodnik
University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia
tomaz.cegovnik@fe.uni-lj.si
jaka.sodnik@fe.uni-lj.si

***Abstract* - This paper focuses on alternative interaction techniques and user interfaces for in-vehicle information systems (IVIS). Human-machine interaction in vehicles needs to introduce new intuitive and more natural interaction approaches, which would reduce driver's distraction and increase safety. We present a prototype of a gesture recognition system intended for in-vehicle free-hand interaction. Our solution is based on a Leap Motion Controller. We report also on a short user study with the proposed system. Test subjects performed a set of tasks and reported on their experience with the system through a user experience questionnaire. The study reveals that free-hand interaction is attractive and stimulating, but it still suffers from various technological issues, particularly with the efficiency and robustness of free-hand gestures recognition techniques.**

## I. INTRODUCTION

Today's vehicles provide an increasing number of new functionalities that enhance the safety and driving performance of drivers or raise their level of comfort. In addition to a variety of passive and active safety systems, information systems have also become increasingly common, enabling modern communication mechanism and luxury facilities [1]. The most common examples are navigation systems, multimedia devices and connectivity services. With every generation of vehicles, the range of these functionalities increases and requires new techniques for human-machine interaction (HMI). Due to the growing functional complexity and mostly restriction to tactile input and visual output, many user interfaces show very poor usability. In addition these systems require long learning periods, which often increases the potential of errors and user frustration [2][3].

However, the primary task in vehicles remains the driving process itself, which demands a certain amount of visual and cognitive attention. Secondary tasks, such as controlling a complex infotainment system, can distract the driver from controlling the car and focusing on the traffic. Since inattention proved to be a major cause of many car accidents, it is reasonable to search for interaction options, which cause less driver distraction in both cognitive and visual domains. Interaction with IVISs is most commonly enabled through buttons and rotary knobs attached to different parts of the vehicles' steering wheel or dashboard [4]. Other input possibilities are speech and interaction via touch interfaces. All interactions, except speech, require eye contact and lead to a visual distraction of the driver.

HMI in vehicles needs to introduce intuitive and natural interaction approaches such as multimodal interaction interfaces. Multimodality defines an interaction form in HMI in which more modalities (i.e. communication channels) are used simultaneously. Input and output processes are combined in a coordinated manner [5]. Currently available conventional interfaces can be upgraded by simultaneous use of speech, gestures and touch for input and displays, speech non-speech sounds and haptic feedback for output. Theoretical foundation for this principle is the unique characteristic of human working memory which seems to be working in fully functional separable components for different human senses [6]. Should a component be overwhelmed (e.g. visual working memory while driving a vehicle) other components (e.g. gestures and speech) can be used instead. By using more components simultaneously, the processing capacity of the working memory is optimized and the visual working load is relived, which leads to less visual distraction and errors.

Such intuitive and natural interaction systems address also the problem of increasing amount of complex functionalities since this approach shortens drivers' learning and adaptation period. Namely, the learning process represents even higher distraction for the primary task as it derives a huge amount of visual and mental attention to a new user interface. During the learning period, driver's brain builds a mental representation of the interaction system which is a mental mirror image of the real system. We call a system intuitive, if the mental representations are already present and no or very little learning period is needed. Speech, movements and natural gestures are such examples since they are widely used in common every day's situations and interactions.

In this paper, we are focusing on free-hand gesture interaction as a big potential technique to enhance intuitiveness of interaction with in-vehicle devices [7]. It poses numerous advantages over tactile and touch modalities, since it requires lower amount of visual load, reduces driving errors and increases the level of user acceptability. It follows the intuitive interface approach as it uses natural hand gestures where no learning effort is needed. As a part of a multimodal interaction system (e.g. in addition to visual and auditory modalities), it allows the driver to keep his eyes on the road and improve his or her safety.

## II. Related work

A lot of in-vehicle interaction research focuses on minimizing driver workload and reducing distraction caused through visual attention of such systems [8]. Different interaction techniques were compared focusing on their effects on driver performance and eye glance behavior. The results showed that gesture interaction could reduce eye glances on simple secondary tasks although gesture interaction is not fully attention free [9].

A research of the BMW Group shows that using a gesture vocabulary reduced to meaningful gestures can be recognized using simple state of the art hardware. Their system is able to distinguish 17 different hand gestures, which can be used to provide skipping between radio stations and navigating. It also distinguishes six different head gestures, which can be used for simple yes/no decisions (e.g. accepting and denying incoming calls) [3].

Other research states that two techniques provide “low attention” interaction (i.e. speech recognition and gesture-based interaction) opposed to touch-based and tactile-based techniques [9]. The latter suffer from a number of inherent limitations in terms of reducing visual attention [10]. Speech recognition on the other hand, can also be very cognitive demanding and therefore impractical and flawed for in-vehicle interaction [11][12] although it seems very useful as it provides hands-free and visual-free interaction.

Since gesture-based interaction could provide a suitable alternative to other types of interaction, research focuses on developing gesture recognition systems. Almost all developed interfaces for gesture recognition are vision based [13][14] since there are numerous advantages of such approach. Firstly, the camera could serve a multi-purpose analyzer of also other activities and not just hand-gestures. Secondly, it offers flexibility where the gestures can be performed and allows location customization. Thirdly, there are advantages in terms of cost and simplicity of installation. On the other hand, the major challenge of vision-based gesture recognition systems is how to generalize a great number of different users and possible variation of gestures. The algorithms must be resistant also to varying in-vehicle illumination conditions [1].

## III. Gestures

Firstly, we need to define gestures since they have numerous definitions, depending on a specific research field. Physically, gestures are movements of individual limbs. Often the term refers also to facial expressions, gaze tracking, head movements and whole body postures [4]. The primary goal of various movements of body parts is augmenting the basic verbal interpersonal communication by exchanging also expressions. Gestures, as a mean of communication between persons, are used and identified subconsciously and their semantics is often hidden or unknown.

In this paper, we primarily refer to gestures as conscious and intentional movements of a selected hand and arm to communicate information with a system. Gestures can be identified and described by two well-defined attributes: temporal seclusion and movement trajectory [3].

In order to recognize various gestures, we have to understand their primary structure. Although gestures look like one continuous movement, they consist of three phases [4].

- Preparation: The hand is brought to the position to start the stroke.
- Stroke: Main part of the gesture which defines its meaning.
- Retraction: The hand is brought back to the resting position.

The stroke phase is essential to the gesture, but in order to identify more gestures merged into one continuous sequence, we need to take into account also the first and last phase.

According to Geiger gestures in HMI can be divided into seven categories [10].

- Mimic gestures: imitating an object (e.g. pick up a phone),
- Schematic gestures: special kind of mimic, standardized symbols witch need to be learned,
- Kinemimic gestures: imitating a direction,
- Symbolic gestures: imitating an abstract feature like an emotion, feeling or thought, e.g. thumb up for yes.
- Deictic gestures: pointing towards the intended destination.
- Technical gestures: used by experts in their working field where no other communication is possible (e.g. diving).
- Encoded gestures: language of technical gestures.

A literature review and analysis of a research published in [15] led to an overall classification of hand gestures based second controls (see Figure 1).

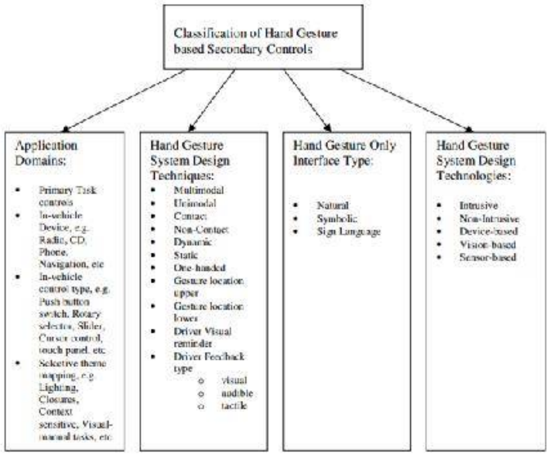


Figure 1: The diagram shows the organization and the categories used for classifying gestures [15]

## IV. Gesture recognition

There are two different technology bases for gesture recognition (i.e. video-based systems and sensor-based systems). Video-based recognition is realized by cameras and requires constant illumination. The various lighting conditions in vehicles present a problem for robust image processing, therefore a constant near-infrared lightning source is needed and a daylight filter to compensate the varying conditions. Sensor-based gesture recognition depends on distant measurement sensors, and works better than the video-based gesture recognition since the background is farther away from the sensor and can be blocked out [4].

However, gesture recognition requires light independent techniques. A motion based entropy technique has to be applied to the near infrared imaging. Human skin has the characteristic of high reflectance of infrared radiation and in majority of cases the hand is the brightest object in the scene [3]. An effective segmentation algorithm is required to segment a hand shape from the background, such as entropy motion segmentation or approaches based on restricted coulomb energy [14].

In-vehicle HMI requires a hand to be displayed in a 2D image field visible to the camera and restricted by the camera's view cone. If the human body is beyond the effective interaction space, the hands may not be fully displayed in the image and the gesture recognition would not be complete. In order to address this problem and to design better gesture interfaces, several researches have defined the regions within the vehicle's cockpit for optimal performance of gesture recognition systems [16].

## V. Leap Motion Controller

The Leap Motion Controller is a small cost effective and accessible USB device, which recognizes all types of hand and finger movements and measures their position and velocity [17]. It illuminates the space over the controller with three infra-red LED light sources and captures the hands with two cameras (see Figure 2). The captured stereo-image is processed with a segmentation algorithm resulting in a data structure, which contains precise position of each finger at every moment. The Leap Motion's API processes this data and gives precise velocities of each finger and hand, and also combines movement patterns into gesture frames. It recognizes four different gesture types:

- Circle - A finger drawing a circle.
- Swipe - A long, linear movement of a hand and its fingers.
- Key Tap - A tapping movement by a finger as if tapping a keyboard key.
- Screen Tap - A tapping movement by the finger as if tapping a vertical computer screen.

The controller's field of view is an inverted pyramid centered on the device [18]. The effective range of the controller extends from approximately 25 to 600 millimeters above the device. The main limitation of the controller's performance is the low and inconsistent sampling frequency (i.e. its mean value is less than 40Hz).

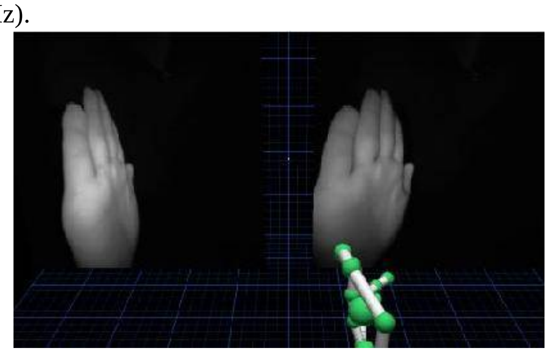

Figure 2: Hand illuminated by infrared LED diodes and captured by the two cameras of the Leap Motion controller (i.e. the bottom right part of the figure [21]).

## VI. Free-hand gesture interaction with IVIS

We used the Leap controller's gesture models to develop a simple free-hand gesture interface for IVIS. Our IVIS simulated the majority of common functionalities related to navigation (e.g., traffic reports, navigation assistance), entertainment (e.g., audio, video, communication) and vehicle control (e.g., air conditioning, system information, cruise control) [19]. All features were accessible through a hierarchical menu structure. The top-most level of the structure was called the "Main menu" and each level of the individual sub-menu consisted of up to eight options. At each level, the user could freely navigate between all the available options, select one of them and enter the corresponding sub-menu or exit and return to the previous menu.

The output of the interface was a simple graphical menu displayed on a small dashboard screen (Figure 3). Four different commands were required to control this menu: up, down, confirm and return. Up and down were moving the selecting cursor up and down on the list while confirm and return entered or left a submenu. Four free-hand gestures were bound to these four commands (see Figure 4). A "circle" gestures drawn with a finger were used for up and down commands – a clockwise circle for moving down and counterclockwise for moving up. A finger "tap" gesture was used for selecting an item and a "swipe to the left" gesture was used for returning to the previous menu.

Figure 3: The study setup: the Leap motion controller is located below the hand and the IVIS display is located next to the screen of the simulator.


Figure 4: Gestures controlling the IVIS.

## VII. Preliminary user study with the proposed system

We performed a short user study to evaluate the usability of the proposed free-hand gesture interaction system. The goal was primarily to get some user feedback on gesture interaction and to assess its acceptance and efficiency.

16 subjects (9 female and 7 male) participated in the user experience evaluation. They had to perform three different tasks with the system:

- set a temperature to the selected value,
- set a radio receiver to the specific station or
- check the vehicle's status (e.g. battery state, fuel level or errors).

The experiment was performed in provisionary driving simulator with stable light conditions. After performing the task, each subject filled out the User Experience Questionnaire (UEQ) [20]. The UEQ is intended to be a user-driven assessment of a system quality and usability. It consists of 26 bipolar items rated on a seven-point Likert scale. The UEQ allows the experience to be rated using the following six subscales: attractiveness, perspicuity, efficiency, dependability, stimulation and novelty of the display technique.

## VIII. Results and discussion

Results of our user study show high interest of test subject for this new interaction technique. Figure 5 compares mean values for all six UEQ categories. The gesture interface has very positive values in the categories novelty, stimulation and attractiveness. We believe since the technology is new for the subjects, they are attracted to it and highly stimulated to use it.

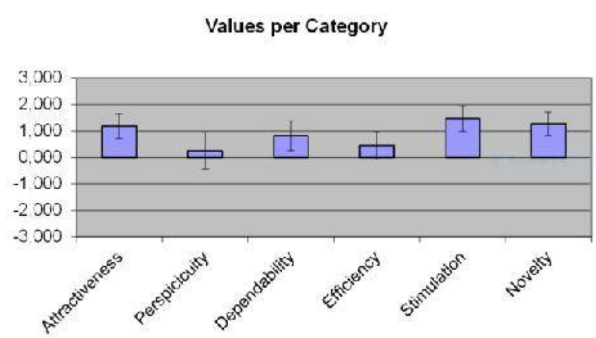


Figure 5: Result values per category of the UEQ [21].

On the other hand, mid-value results in the category dependability indicate that test subjects are not entirely convinced they could rely on this type of interface in real vehicles. The poorest results, but still positive, are in the categories perspicuity and efficiency. The latter reflects the problems of the proposed interface to clearly and robustly detect gestures performed by variety of test subjects. Despite using only simple and intuitive gestures, their correct execution depended strongly on individual's performance and varied a lot between different users [21].

In the course of the study, we also noticed that several users did not clearly understand the exact form for performing individual gestures. The biggest problem seemed to present the execution of a circular gesture, which should be only a simple circular movement of one finger. Instead, some users performed big circles with their entire arm which often escaped the controller's field of view and could not be successfully detected. We believe this is the main contributor to the poor result in the UEQ efficiency category.

## IX. Conclusion

Nowadays the increasing range of functionalities included in in-vehicle interaction requires new interaction techniques. Conventional techniques often result in decreased driving performance since they require a huge amount of driver's visual attention and visual workload. Increased visual workload or even overload causes high distraction and great deficit of focus on the road and traffic situations. In this paper, we proposed a possibility to lower the visual workload in vehicles by using new and very intuitive interaction technique. We explored the opportunity of using gestures as interface input - they can be used intuitively because they are already part of interpersonal communication.

Our experimental results show that there is an attraction and interest among people for using this type of interaction. However, there is still a clear need to improve the detection performance of such systems. Majority of the development and research in this area try to recognize gestures through vision-based technology, i.e. capturing the hand with cameras and illuminating the vision field with an infra-red light source, which tries to nullify the restrictions due to variable lighting condition in vehicles. The appropriate mathematical segmentation algorithms need to be applied to extract usable data from the captured images.

An example of such technology is already in production in the BMW 7 series for triggering some basic phone functions and controlling the volume. Other car manufacturers (Jaguar, Mercedes, VW, Nissan) announced to implement gesture control systems at latest in 2018.

Although the free-hand gesture interface is designed to be intuitive and use intuitive gestures, each person executes gestures slightly different and the interface needs to adapt to these differences. Furthermore, additional studies need to take into consideration other modalities and other combinations of interaction techniques in order to develop highly usable, adaptable and robust configurations.

## References


[1] Ohn-Bar, E.; Trivedi, M.M., "Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision-Based Approach and Evaluations," in *Intelligent Transportation Systems, IEEE* Transactions on, vol.15, no.6, pp.2368-2377, Dec. 2014.

[2] Tönnis, M.; Broy, V.; Klinker, G., "A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers" in *Proceedings of the 2006*

[3] Althoff, F.; Lindl, R.; Walchshäusl, L., "Robust Multimodal Hand- and Head Gesture Recognition for Controlling Automotive Infotainment Systems", In: VDI-Tagung: Der Fahrer im 21. Jahrhundert, 2005.

[4] Zeiser, J., "Use of Freehand Gestures for Interactions in Automobiles: An Application-Based View", In: *Ubiquitous Computing, University of Munich*, 2011.

[5] Sharon O., "Ten myths of multimodal interaction" In *Commun. ACM* 42, 11, pp. 74-81 1999.

[6] Ablaßmeier, M., "Multimodales, kontextadaptives Informations-management im Automobil", Dissertation, Technische Universität München, 2009.

[7] Loehmann, S.; Diwischek, L.; Schröer, B., „The User Experience of Freehand Gestures", In: *User Experience in Cars Workshop at INTERACT 2011 - 13th IFIP TC13 Conference on Human-Computer Interaction*, pp. 54-58, 2011.

[8] M. Geiger, M. Zobl, K. Bengler in M. Lang, "Intermodal differences in distraction effects while controlling automotive user interfaces," 2001.

[9] Majlund Bach, K.; Gregers Jaeger, M.; Skov, M. B., „You Can Touch, but You Can't Look: Interacting with In-Vechicle Systems" in *CHI 2008 Proceedings*, pp.1139-1148, 2008.

[10] Geiger, M., „Berührungslose Bedienung von Infotainment-Systemen im Fahrzeug", Dr.-Ing. Dissertation, Technische Universität München.

[11] Gellatly, A. W., „The Use of Speech Recognition Technology in Automotive Applications", Faculty of the Virginia Polytechnic Institute and State, 1997.

[12] Sezgin, T. M.; Davies, I.; Robinson, P., "Multimodal Inference for Driver-Vecicle Interaction", In: *Proceedings of the 2009 Inter-national Conference on Multimodal Interfaces*, 2009.

[13] McAllister, G.; McKenna, S. J.; Ricketts I. W., "Towards a non-contact driver-vehicle interface", in *IEEE Intelligent Transportation Systems*, pp.58-63, 2000.

[14] Dong, G.; Yan Y., Xie, M., "Vision-Based Hand Gesture Recognition for Human-Vehicle Interaction" in *5th Int. Conf. Control, Automation, Robotics and Vision*, Singapore, 1998, pp. 151–155.

[15] Niedermaier, F. B., "Entwicklung und Beertung eines Rapid-Prototyping Ansatzes zur multimodalen Mensch-Maschine-Interaktion im Kraftfahrzeug", Dr.-Ing. Dissertation, Technische Universität München.

[16] Riener, A.; Ferscha, A.; Bachmair, F.; Hagmüller, P.; Lemme, A.; Muttenthaler, D.; Pühringer, D.; Rogner, H.; Tappe, A. & Weger, F. (2013), Standardization of the in-car gesture

interaction space., *in* Jacques M. B. Terken, ed., 'AutomotiveUI' , ACM, , pp. 14-21.

[17] Leap Motion Controller https://www.leapmotion.com/, 1/2016.

[18] J. Guna, G. Jakus, M. Pogačnik, S. Tomažič in J. Sodnik, “An analysis ofthe precision and reliability of the leap motion sensor and its suitability forstatic and dynamic tracking,”, In: *Sensors*, vol. 14, no. 2, pp. 3702, 2014.

[19] Jakus, G., Dicke, C., Sodnik, J., “A user study of auditory, head-up and multi-modal displays in vehicles”. In: Applied ergonomics, vol. 46, 2015, pp. 184-192.

[20] B. Laugwitz, T. Held and M. Schrepp, "Construction and Evaluation of a User Experience Questionnaire", *Lecture Notes in Computer Science*, pp. 63-76.

[21] Čegovnik, T., “Prostoročna interakcija človek-stroj v vozilih”, In Master Thesis, University of Ljubljana, Faculty of Electrical engineering, 2015.