Contactless Interaction for Automotive Applications

Thomas Kopinski, Stefan Geisler, Uwe Handmann

Computer Science Institute, Hochschule Ruhr West, University of Applied Sciences

Abstract

Touchless gestures are getting more and more popular since Microsoft introduced the Kinect. They are not only interesting for entertainment and games, but also advantageous for in-car applications, as the driver can keep the hand closer to the steering wheel and displays can be mounted closer to the line of visibility at positions normally not in the reach zone. Since the technology used in consumer electronics is generally not suitable for automotive applications an alternative approach including hard- and software aspects is presented in this paper. A second aspect is the identification of the driver with biometric methods. The base idea is to also work contactless and omit additional technology which drivers need to carry with them, but in many cases they don't or share it between different persons (like the key).

1 Introduction

One of the fundamental decisions related to the human-machine interaction (HMI) base concept while designing the cockpit of a car is between touchscreen systems and highly mounted displays combined with control buttons and in most cases turn push controller. The systems include functionality for radio, multimedia, telephony and other connectivity as well as navigation. While manufacturers like Volkswagen or Ford emphasize the advantage of touchscreens concerning the intuitiveness of direct control of functions with interaction concepts well known from smartphones, nowadays also including gestures, others like Audi or BMW accent the display position close to the windscreen and therewith the road which leads to lower distraction and reduced length of gaze diversion. Examples are shown in figure 1.

Developing a combined solution is not easy as the position of the highly mounted display is not in the reach zone of the driver. So the optimal position in the line of visibility is not fully compatible with touching the surface of the display.

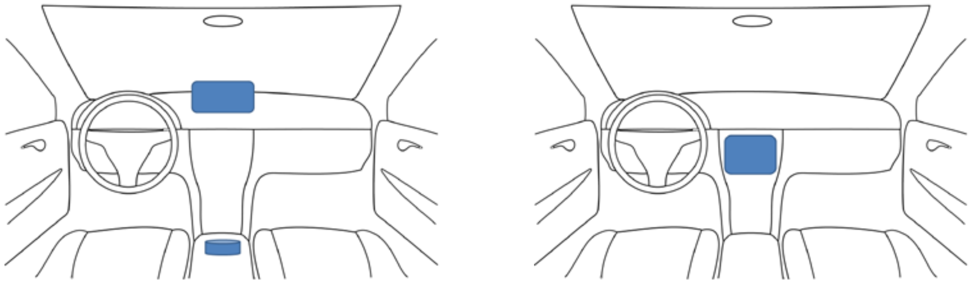

Figure 1: Typical cockpit styles: left:highly mounted display with controller, right: touchscreen

Against this background different research activities at different research institutes and OEMs have been started for touch-less gesture control within the last years. This technology is widely known from Microsoft's Kinect especially for games and recently introduced for TV remote control from Samsung and now brought to automotive applications. An overview of these activities as well as a first insight into an own research project is given in section 2 resp. 3. A reliable gesture recognition can result in a more flexible cockpit design with lower number of hard keys and flexibility for designing the cockpit surfaces.

Touch gestures and many other system settings, e.g. radio presets, seat position, mirror settings, parameters for driver assistance systems like time distance gap in ACC or sensitivity of a lane departure warning system are driver dependent. Recognizing the driver without explicit input is comfortable. One broadly discussed technical solution is the driver identification with the key. However especially in family scenarios not only the car itself is shared, but also the key. Hence for this approach, a change in human behaviour would be required. In section 4 an approach which does not require any additional artefacts and is based on a biometric system is outlined.

The concepts described in this paper should be considered as a contribution to a more natural interaction within the vehicle cockpit. After the driver hast taken place on the seat he/she is directly recognized by the biometric system. Any kind of explicit interaction in this case would be an unnatural step in the flow of the vehicle start scenario. The disadvantages of artefacts have been discussed before. The proposed integration of a biometric system neither needs artefacts or explicit interaction, nor is any input from the driver required.

2 Touchless gestures for Automotive Applications

This section should give an overview of a selected number of projects and prototypes related to touchless gesture recognition especially for automotive applications.

In (Akyol et al., 2000) a concept for gesture control of in-car systems is described. The system developed from BMW and the RWTH Aachen works with near infrared light and a modified CCD camera and a Maximum-Likelihood classification algorithm. A set of 20 static hand gestures is defined as well as (for a better recognition rate) a subset of 6. For 6 hand gestures the recognition rate is approx. 98%. Problems occur under certain lighting conditions.

The most important question before developing a new technology for a new field of application is, if the usage makes sense there. In 2001 a study was published by Martin Zobl at the Munich University of Technology (Zobl et al. 2001) followed by the technical development of a system (Zobl et al. 2003). Generally speaking the authors distinguish two types of gestures: Gestures that don’t have to be learned (e.g. pointing or waving) and those which have to be learned and are in some cases culture-dependent (mimics, e.g. virtual phone). For latter reason and to allow personalizing gestures the system outlined in section 3 shall be a learning system so that users can teach their own gestures and is not limited to an OEM provided set.

(Kollorz et al. 2008) in collaboration with the University of Nürnberg and Audi developed a system using a time-of-flight camera (a camera which is able to measure the distance to an object) for gesture recognition. The classification algorithm is based on k-nearest-neighbors. 12 different static hand gestures could be classified with a rate of 94.61% on a standard PC.

A combination of hand and head gestures was developed at BMW (Althoff et al. 2005). The authors conclude that head gestures have the greatest potential for yes/no questions while hand gestures are proposed for skipping between audio tracks or radio stations and to navigate in a map. Their multimodal approach also includes voice control.

In (Alpern, M., & Minardo, K. 2003) a gesture based system in combination with a head-up display for secondary tasks is described and tested. Their results present that the system helps to reduce distraction, although users also claim low precision when pointing.

The concepts already made their way from research to public prototype: e.g. Audi at the CES 2012, Hyundai at the Detroit Auto Show 2013.

A broader overview can be found in (Pickering et al. 2007).

3 A multimodal hand-gesture recognition system

We combine the effectiveness and precision of a multimodal approach with modern time-of-flight technology in order to develop a robust system for recognizing static as well as dynamic hand-gestures under noisy conditions within the frame of real-time applicability.

While many approaches make use of the Kinect system (Wu et al. 2012, Keskin et al. 2011), our sensors deliver depth-data of the environment even during intense sunlight conditions while also remaining small in size. The latter remains a crucial argument for being able to realize our system within the automotive environment.

The presented approach displays a gesture recognition system in the field of automotive HMI which is realized by multiple time-of-flight camera modules combined with a microprocessor unit. This setup allows for full operationality during alternating weather conditions and day and night change while maintaining real-time applicability for a reasonable amount of hand gestures.

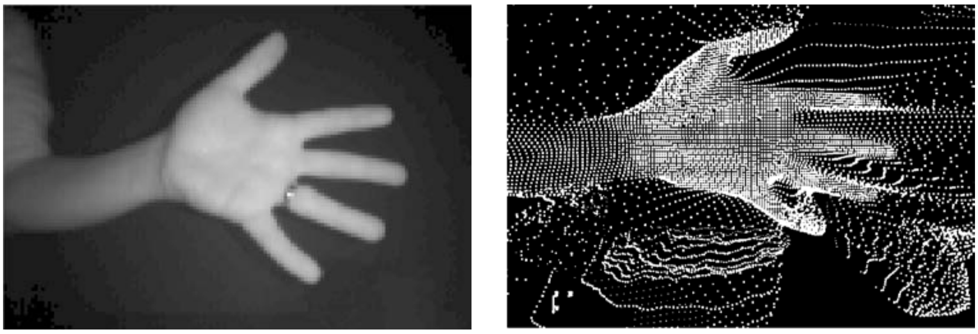

Figure 2:Intensity image from the PMD ToF-sensor (left) and the visualization of the point cloud image acquired from the ToF-camera (right)

In order to develop a system to be fully applicable in such an environment, the technical prerequisites to be considered are demanding. Given the described setting a set of static (cf. Fig. 2) and dynamic gestures has to be defined. To this end we describe static hand gestures as contrasting to dynamic since the position of the hand does not change in a significant manner over time whereas dynamic hand gestures are understood as a movement or change of the hand as a whole in any of the three possible dimensions in addition to the 24 degrees of freedom contained in the joints of the hand (Lee et al. 1995).

Given these demanding conditions in the presented environment the desired functionality can only be achieved by effectively processing large amounts of data. The presented camera system works with an illumination wavelength of 850nm and a resolution of 165*120 px which on the one hand makes it robustly applicable especially in challenging daylight situations and on the other hand optimized for close distances as the mentioned car-setting.

In order to achieve real-time applicability, the given pointclouds are processed by algorithms optimized to extract features from human hand gestures. These descriptors are used to distinguish between a given set of gestures. They therefore need to balance the trade-off between high robustness as well as fast computability.

To achieve the described functionality a high recognition rate for a reasonable amount of hand gestures is desirable. To this end we realize the system within a multi-sensor environment. The additional camera-system allows for an improved setup as we now can scan the desired area, within which the gestures are to be recognized in a significantly improved precise manner.

Moreover, the machine learning algorithms applied to this problem mutually support their classifications leading to an enhanced functioning system.

With respect to applying this system in the desired environment, the presented setup is realized in an embedded system. The data processing is done by multiple microprocessors which, combined with the small camera-technology, allows for the complete system to be incorporated in an automotive setting.

4 Biometric System

An additional challenge to be solved in the field of automotive applications is the personalized interaction with people. To start an interaction process between a device inside a car and a person, one important information is the knowledge about the interacting partner's identity and whether the interacting partner is present or not. This means, the application must be able to detect and be finally able to identify persons.

As already discussed in the introduction there are certain disadvantages in currently available solutions. Either an interaction step is required to select the driver from a menu or an artifact like the key needs to be bound permanently to a single person. An alternative are biometric systems which can use a camera to detect and identify the driver. Accurate identification of specific individuals has to be done by analyzing the individual features of each person.

In general people could be detected by various kinds of sensors (Arras et al. 2009, Wichert et al. 2004, Wiegert et al. 2004, Yang et al. 2002). A typical feature set that allows for a distinct identification of a specific person is often extracted from the facial image acquired by a camera. This feature-set is stored in a database to allow the time-independent recognition of different persons by comparing given feature-sets. The recognition process typically compares individual feature-sets that are extracted from an acquired image against feature-sets of known people that are stored in a given database. For example, feature-sets, also known as biometric templates, could be fingerprint features such as minutiae or facial features such as texture descriptors.

In order to decide if a recognized individual is contained in the database, a one-to-many comparison has to be performed. Since there is no claimed identity available for the person in front of the system, the given template has to be compared against all known templates stored in the database, the so-called gallery (Gehlen et al. 2001, Philips et al. 2007, Zhao et al. 2003). The decision whether a person is known to the system or not is made on the most similar gallery element found, which needs to exceed a threshold. If no template of a given person is stored in the gallery, identification fails, because no template in the database is similar to the acquired template.

If the system recognizes that an unknown person sits on the driver seat the new person can choose to create a new profile. In this case the image and the feature vector are stored in the system database together with the settings made by this new user. The next time this person enters the car again he/she will be recognized and the settings will be loaded from the database to meet the user's preferences. Changes made by the user during the drive will be stored automatically to his/her profile so that it will be directly available for the next drive.

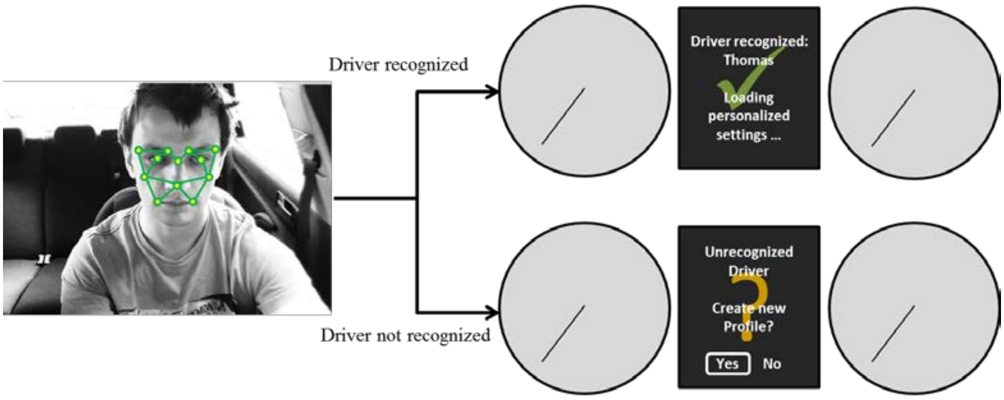

Figure 3: Driver identification with biometric face recognition algorithms (schematic diagram)

Personalized interaction with people in an in-car scenario allows for the opportunity to adapt the assistance system, the infotainment system, or to be able to activate specifically required features for an interaction partner.

In the field of automotive applications a face-based recognition process permits a contactless interaction between the application and the user. Additionally the necessary sensors for a face-recognition process could be used for other automotive applications, e.g. in-car eye-tracking (Green 2002). For this reason a face-recognition process is preferred (figure 3).

5 Conclusion and Outlook

We present the design of a system for touchless gestures which considers improving existing approaches by extending the setup into a multi-sensor environment. We expect to exceed existing results in terms of recognition rate and robustness. Allowing the user to define own gestures – static as well as dynamic – we yield a gesture recognition system which is diverse in terms of applicability. Combining this setup with a biometric system gives us the possibility of individualizing e.g. a set of gestures depending on the current user.

One of the challenges to overcome is the design of the designated system. Using multiple cameras to increase the recognition rate and stabilize the detection is not only more expensive than the compared approaches but also requires further overhead in terms of design and space management.

These factors are already being considered during the early stage of development as we focus our approach for instance on recognizing hand gestures in close range and moreover validate our implemented algorithms by various setups concerning different angles and alignments of the cameras. Especially the latter factors are of major importance and therefore are planned and tested extensively.

In addition to the technical validation a user research is planned for the new approach. This includes interaction concepts for daily use but also for the learning phase. Furthermore the question shall be addressed, what kind of displays are to be controlled by such a system and moreover what a suitable camera position might be to achieve a high recognition rate as well as convenient interaction.

For personalization and driver identification a biometric approach is proposed as this works without any artefacts. More precisely a camera-based system working with face recognition gives the advantage that no driver interaction is required and additional, camera-based functions can be realized based on the same hardware.

References

Akyol, S., Canzler, U. Canzler, Bengler, K., & Hahn, W. (2000). Gestensteuerung für Fahrzeugbordsysteme. *In Mustererkennung 2000, 22. DAGM-Symposium*, Springer-Verlag, London, UK, pp. 139-146.

Alpern, M., & Minardo, K. (2003). Developing a car gesture interface for use as a secondary task. In *CHI'03 extended abstracts on Human factors in computing systems.* pp. 932-933.

Althoff, F., Lindl, R., Walchshausl, L., & Hoch, S. (2005). Robust multimodal hand-and head gesture recognition for controlling automotive infotainment systems. VDI BERICHTE, 1919, 187.

Arras, K.O. & Martinez-Mozos, O. (eds.) (2009). *ICRA 2009 Workshop Proceedings of People Detection and Tracking*

Gehlen, S., Rinne, M. & Werner, M. (2001): *Hierarchical Graph-Matching, European Patent 01118536.0*

Green, P. (2002). *Where do drivers look while driving (and for how long)? In R.E. Dewar & P.L.Olson (Eds), Human factors in traffic safety. Tucson, AZ: Lawyers and Judges Publishing,* pp. 77-110.

C. Keskin, F. Kirac, Y. E. Kara & L. Akarun (2011). Real time hand pose estimation using depth sensors. In *Computer Vision Workshops (ICCV Workshops),* pp. 1228-1234.

Kollorz, E., Penne, J., Hornegger, J. & Barke, A. (2008). Gesture recognition with a Time of Flight camera. In *Int. J. Intell. Syst. Technol. Appl. 5, 3/4 (November 2008),* pp. 334-343.

J. Lee & T. Kunii (1995). *Model-based analysis of hand posture. IEEE Computer Graphics and Applications,* 15:77-86.

Pickering, C. A., Burnham, K. J., & Richardson, M. J. (2007). A research study of hand gesture recognition technologies and applications for human vehicle interaction. In *Automotive Electronics, 3rd Institution of Engineering and Technology Conference on Automotive Electronics*, pp. 1-15.

Phillips, P.J. et al.(2007): *FRVT 2006 and ICE 2006 Large-Scale Results, National Institute of Standards and Technology Gaithersburg, MD 20899, NISTIR 7408* (March 2007)

Wichert, G., Handmann, U., et al. (2004): The Robotic Bar – An Integrated Demonstration of a Robotic Assistant. In *Advances in Human-Robot Interaction. Springer Tracts in Advanced Robotics (STAR), vol. 14. Springer.*

Wiegand, S., Igel, C. & Handmann, U. (2004): Evolutionary Multi-Objective Optimization of Neural Networks for Face Detection. *International Journal of Computational Intelligence and Applications 4(3),* 237–253.

Shen Wu, Feng Jiang, Debin Zhao, Shaohui Liu & Wen Gao (2012). *Viewpoint-independent hand gesture recognition system.* IEEE Visual Communications and Image Processing (VCIP), pp. 1-5.

Yang, M.-H., Kriegman, D.J. & Ahuja, N (2002).: Detecting Faces in Images: A Survey. *IEEE Transactions On Pattern Analysis And Machine Intelligence 24(1)*

Zhao, W., Chellappa, R., Rosenfeld, A. & Philips, P.J. (2003): Face Recognition: A Literature Survey. *ACM Computing Surveys 35(4),* pp. 399–458.

Zobl, M., Geiger, M., Bengler, K., & Lang, M. (2001). A usability study on hand gesture controlled operation of in-car devices. In *Abridged Proceedings, HCI 2001, 9th Int. Conference on Human Machine Interaction*, pp. 166–168.

Zobl, M., Geiger, M., Schuller, B., Lang, M. & Rigoll, G. (2003). A real-time system for hand gesture controlled operation of in-car devices. In *Proceedings of the 2003 International Conference on Multimedia and Expo - Volume 3 (ICME '03) - Volume 03 (ICME '03)*, pp. 541-544.

Contact

Thomas Kopinski, E-Mail: thomas.kopinski@hs-ruhrwest.de
Prof. Dr. Stefan Geisler, E-Mail: stefan.geisler@hs-ruhrwest.de
Prof. Dr.-Ing. Uwe Handmann, E-Mail: uwe.handmann@hs-ruhrwest.de