

Controlling vehicle functions with natural body language

Dr. Alexander van Laack¹, Oliver Kirsch², Gert-Dieter Tuzar³, Judy Blessing⁴

Design Experience Europe, Visteon Innovation & Technology GmbH¹

Advanced Technologies, Visteon Electronics Germany GmbH²

Design Experience Europe, Visteon Electronics Germany GmbH³

Market & Trends Research Europe, Visteon Innovation & Technology GmbH⁴

Abstract

As vehicles become more context sensitive with sensors providing detailed information about the exterior, understanding the context and state of the driver becomes crucial as well. This paper presents an approach to develop a system to recognize human gestures in the automotive cockpit. It provides an overview of the proposed technology and discusses the advantages and disadvantages of certain systems. Qualitative user research studies highlight recommendations for future implementations.

1 Introduction

Social communication is more than just the spoken word. Non-verbal communication is a way for human beings to express meanings and the fact that people move their hands while talking is a phenomenon which can be observed across cultures and ages. Even babies start gesturing long before they can articulate a word (Goldwin-Meadow 1999). Although some may argue that the communicative value of gestures is low, or even none existent (Krauss 1991), we observe that even blind people use gestures when communicating, although they have never seen themselves doing a gesture and although the listener they are talking to, may be blind as well (Iverson 1998).

In the automotive context, OEMs and suppliers work together on building smarter vehicles that are not only able to understand the context of the driving situation but also the state of the driver. Multi-modal interaction systems are playing a key role in future automotive cockpits and identifying human gestures is a first step of driver monitoring.

This functionality can be provided by 2D/3D interior camera systems and related computer software solutions. Especially 3D sensing technology promises a new modality to interact in a

more “natural” and pleasant way (intuitive HMI) with the interior and infotainment system of the vehicle because the exact position of the objects in 3D space is known.

2 Interaction Technology

2.1 State-of-the-Art

Looking to camera based solutions, several principles and technologies are available on the market and in the research domain to record 3D depth images. Triangulation is probably the most common and well-known principle that enables the acquisition of 3D depth measurements with reasonable distance resolution. It can be either a passive system based on a Stereo Vision (SV) camera system using two 2D cameras, or an active system using one single 2D camera together with a projector that beams a light pattern into the scene. The drawback of active triangulation systems is that they have difficulties with shadows in the scene area. Stereo Vision systems have difficulties with scene content that has no or only rare contrast since the principle of SV based triangulation relies on identification of the same features in both images.

An alternative depth measurement method that Visteon uses in one of its technology vehicles is called the “Time-of-Flight” (ToF) principle. A camera system that uses this technology can provide an amplitude image and a depth image of the scene (within the camera field of view) with a high frame rate and without moving parts inside the camera. The amplitude image looks similar to an image known from a 2D camera. That means that well reflecting objects appear brighter in that image than less reflecting objects. The distance image provides distance information for each sensor pixel of the scene in front of the camera.

The essential components of a ToF camera system are the ToF sensor, an objective lens and an (IR) illumination light source that emits RF modulated light.

2.2 Advanced Development

Visteon started to investigate the ToF technology several years ago in the Advance Development Department to build up know-how and to understand the performance and maturity of the technology. Beside the technology investigation inside the lab, a proprietary hardware and software solution has been developed that enables recognition of hand gesture in mid-air and further use cases that can be linked to certain hand poses or hand and finger position in 3D space.

The right column of Figure 1 shows the system partitioning of a ToF camera system and provides high-level information about the different software blocks in our current system.

The column on the left provides an overview of the system architecture with a high-level flow chart on how the gesture recognition concept works.

This OEM independent development phase was necessary to reach the following goals:

- Find out the potential and limitation of the current ToF technology and investigate ways to reduce current limitations
- Build up a good understanding of feasible use cases that can work robustly in a car environment by implementing and testing them; and executing improvement loops based on the findings
- Investigate the necessary hardware (HW) and software (SW) components of such a system

These activities brought us into a position to build up our proprietary know-how in this field, demonstrate this know-how to OEMs and to give recommendations about which uses cases (per our experience and knowledge) can work robustly in a car environment. These achievements also reduce risks - either by offering a solution for co-innovation projects or RFQs for serial implementation.

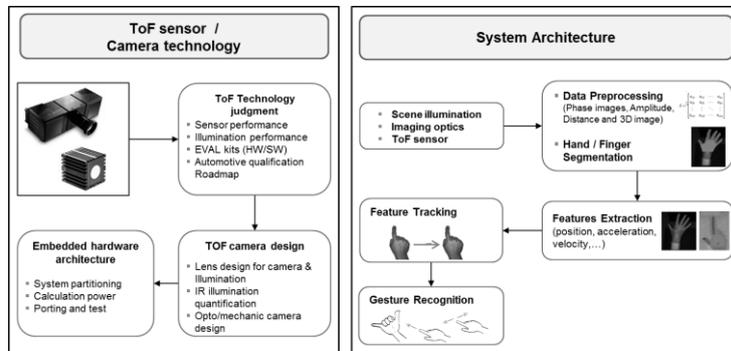


Figure 1: Visteon's ToF camera technology and system architecture

2.3 Vehicle integration

During the advanced development activities one ToF camera system was integrated into a drivable vehicle to investigate use cases for this new input modality that can be used, for example, for driver information and infotainment systems to reduce driver distraction and enhance the user experience.

The implementation in the car demonstrates the current status of the functionalities that have been achieved so far – considering a camera position that is located in the overhead console (see Figure 2 and Figure 3).

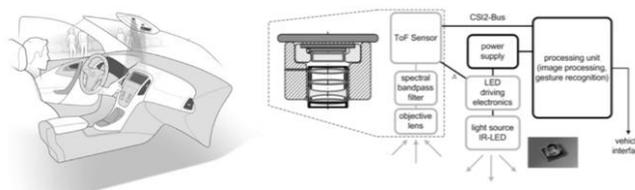


Figure 2: ToF Camera setup in Visteon's Time-of-Flight technology Vehicle

This car is used to demonstrate the possibilities of ToF technology to OEMs. In parallel, the vehicle is also used continuously as the main development platform to implement new use cases and to improve our software algorithms in order to enable a robust hand/arm segmentation and tracking under real environmental and build-in conditions.

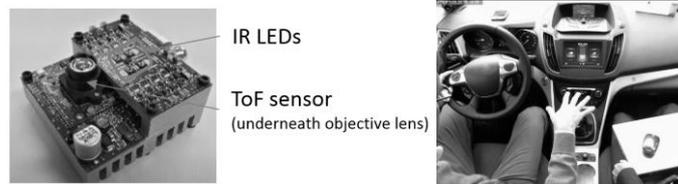


Figure 3: Camera head of ToF prototype camera (left), hand gesture interaction within the car (right)

2.4 Investigation of use cases beyond simple 3D gestures

Because the ToF camera system provides - besides the amplitude image - also depth information, it can be used to do more than just detecting simple gestures (like a swipe or approach in a short range). The camera sees the details of the scene (i.e. number of visible fingers and pointing direction) and the exact position in 3D space. Therefore, the target is really to realize a more natural human machine interaction (HMI) within a large interaction space.



Figure 4: Examples of Hand Gestures and hand poses for HMI interaction

Uses cases such as hand pose detection, touch detection on display, and driver/co-driver hand distinction are already integrated and show-cased in the technology vehicle to demonstrate the potential of the technology and enabling new ways for interacting with the vehicle. (See also Figure 4)

3 Consumer Insight Studies

3.1 Usability Research Project

Prior to the ToF implementation in the car, a qualitative pilot study was conducted to identify users' expectations on gesture control in order to control specific car functions that could be implemented. Figure 5 depicts six use cases with their highest acceptance for gesture control. After the ToF technology was successfully implemented in the headliner of the test vehicle, a qualitative user research clinic was conducted to identify the ease of use and the efficiency of the developed system. In this approach, subjective user data and objective data were collected

and evaluated. As this study was conducted with a small sample size of only 11 participants, it was not intended to present statistically significant research data at this point, but rather give an indication on how to approach the topic of in-vehicle gesture interaction.

The clinic was structured in 60 minute interviews applying a “think aloud” technique to gather data. Prior to testing participants were shown an instruction video and they were introduced to general gestures. It was ensured that the selected participants were driving more than 12.000 km per year in a new vehicle. As the gesture system was installed in a Ford C-Max, the conventional Ford C-Max controls were used as baseline for this study.

3.1.1 Time to goal

During a “time to goal” approach participants were asked to perform pre-defined tasks with regular interaction modalities, such as switches, to compare with gesture interaction. The time to fulfill each task is measured and gives an indication about the efficiency of the two systems for each particular use case.

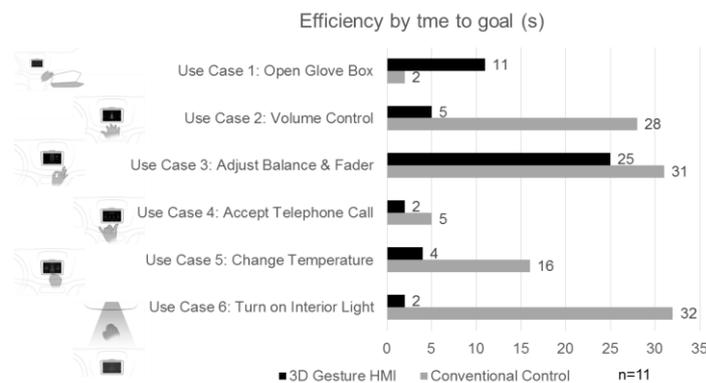


Figure 5: "Time to Goal" results for different use cases

Figure 5 visualizes the results of the six use cases presented during this clinic. In 5 out of 6 use cases the “time to goal” for gesture interaction is shorter than using conventional controls. During use case 1, the user opened the glove box by approaching the glove box with the hand. The gesture camera recognizes the approach and opens it automatically. Unfortunately, there was a learning curve to get the gesture right to the point where the system could understand it correctly. This resulted in a longer interaction time to reach the goal. In contrast the conventional way to open the glove box was done without any mistakes by 100% of the participants.

3.1.2 Task success rate

How successful each task was completed is visualized in Figure 6. It becomes obvious that there are significant differences between the use cases and, in some cases, the conventional control created more mistakes than the gesture interaction.

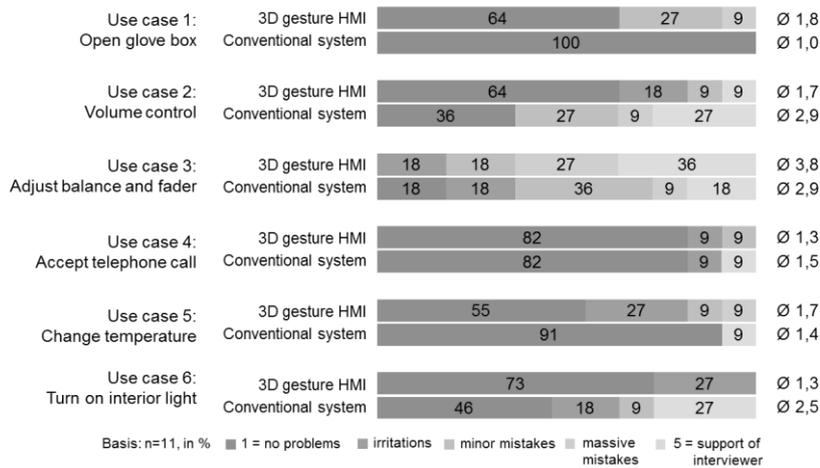


Figure 6: Results on success (objective)

3.1.3 User experience

In addition to the task completion time and measuring the success rate, we investigated the subjective perception of each use case. Participants were asked to rate their experience for each use case and each modality on a scale from 1, being very good, to 5, being very poor.

The results presented in Figure 7 show that for 4 out of 6 use cases, the gesture interaction was on average rated better than the conventional system. In general, we found that concepts of simple and “magical” use cases, which are providing added value, are rated highest. Therefore, turning on the light with a simple gesture was one of the most compelling use cases for the participants.

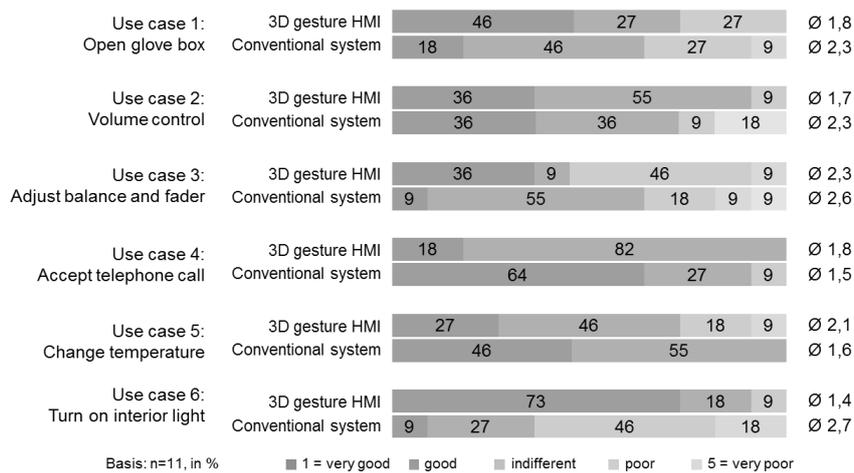


Figure 7: Results on user experience (subjective on concept)

3.2 Consumer Research Clinic

In addition to the first usability investigations, a consumer research clinic was conducted to test how the ToF gesture interaction fits to the ideal consumer experience. In total 45 consumers, 69% males and 31% females, participated in this research clinic. The majority of participants were D-segment drivers with 67%, 30% E-segment drivers, and 4% were driving a B-Segment car. About 40% were between the ages of 33 and 45, 27% were between 18 and 32 years old, and 33% were between 46 and 65 years of age.

3.2.4 Image of gesture interaction

To evaluate the image of gesture interaction, participants were asked what comes to their mind when they think about operating functions in the car using gestures. Concerns about gesture interaction prevailed from the start, which can be explained by the fact that this is still a very new and not yet widely experienced technology. Also, gesture interaction is also not common in current consumer electronic (CE) devices.

After unprompted discussion, participants were shown a video of the previously introduced gesture use cases, which were demonstrated in the test vehicle. Interviewers asked the consumers if and how they think differently about gesture interaction after having seen the video. The results were sobering as the negative feelings were not reduced. People were sceptic about the technology's precision and that it might cause more distraction.

3.2.5 Controlling functions with gesture

After the initial perception of gesture interaction was gauged, consumers were asked to control the functions inside the vehicle. Similarly to the usability research, it was measured to which extent participants could interact with the vehicle functions without further support.

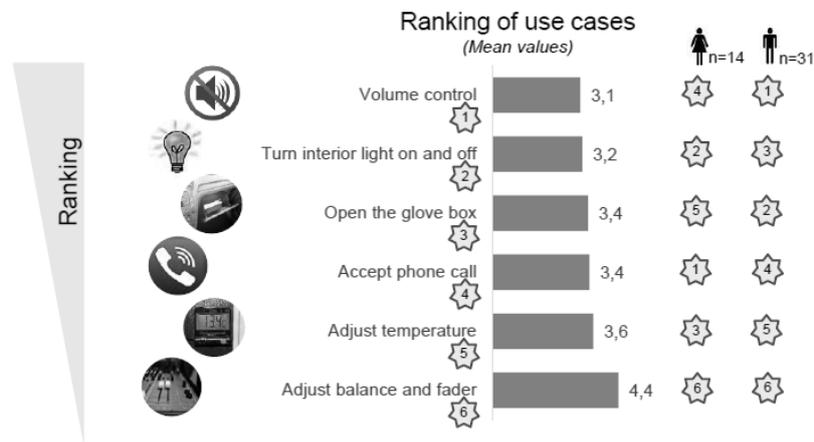


Figure 8: Ranking of use case preferences (overall, mean values, female ranking, and male ranking)

In spite of the first skepticism, most participants did not have significant problems operating vehicle functions with gestures. Comparable to the results of the first clinic, turning on the interior light and accepting phone calls were the easiest use cases for consumers. Overall gesture interaction found acceptance with about half of the respondents, with only very little differences between the various use cases as shown in Figure 8.

4 Conclusion

In this paper, we investigated the implementation and application of a gesture recognition system in a vehicle. Based on extensive state-of-the-art research, the time of flight (ToF) technology was identified as the most suitable to detect a high diversity of spatial and on-surface gestures with high accuracy inside the vehicle. To address the demands of OEMs, a new ToF camera was developed and implemented in a test vehicle.

The current implementation of the ToF system offers a high enough resolution to identify gestures with a high level of details. Due to a fairly large field of view of the camera, it can detect both the driver and the passenger's hand and, therefore, understand who is trying to interact at which point in time. This offers a first step to a contextual system, which can react on active as well as passive gestures.

The conducted user research showed that gesture control can be more enjoyable than conventional interaction. However, it also became obvious that not all interactions with the vehicle should be substituted by gestures. Gesture control should only be offered for dedicated interactive functions and not for safety relevant functions. Best experiences are established when focusing on natural and simple gestures with a feedback for each interaction step.

Additional cultural differences should also be considered when a system like this is introduced to the market. They do not only play a significant part in the judgmental evaluation but also a relevant part in gesture interaction. (van Laack 2014)

References

- Goldin-Meadow, S. (1999). *The role of gesture in communication and thinking*. In: *Trends in Cognitive Sciences*, Vol. 3, No. 11, November 1999, p. 419
- Iverson, J.M. & Goldin-Meadow, S. (1998). *Why people gesture as they speak*. In: *Nature*, Vol. 396, p. 228
- Krauss, R.M. & Morrel-Samuels, P., Colasante, C. (1991). *Do conversational hand gestures communicate?* In: *Journal of Personality and Social Psychology*. Vol. 61, pp. 743–754
- van Laack, A.W. (2014). *Measurement of Sensory and Cultural Influences on Haptic Quality Perception of Vehicle Interiors*. Aachen: van Laack GmbH Buchverlag