Explicitness of Societal Influence

How computer vision applications change our everyday life

Essay by André-Marcel Hellmund, Karlsruhe Institute of Technology
E-Mail: andre-marcel.hellmund@kit.edu

Introduction

The camera obscura is assumed to be one of the earliest image recording devices in human history[21]. Its first naming dates back to the 4th century B.C., however it took more than 2000 years until cameras have been widely available for everyday use. Initially, expensive and exclusive cameras were used by professional photographers and technical enthusiasts, but the technological advances leading to mass production and device miniaturization (e.g. in smartphones) resulted in a massive spread. The still most frequent and important use of cameras is to capture and retain (un)forgettable moments – Aristotle must have been a far-seeing man and must have solely been thinking about Selfies as *the* use-case of cameras. Thank you for this!

However, the camera is not purely a memorizing device, it is a sensor device – and indeed a highly informative and powerful one. Given the set of today's available sensor devices, the camera is by far distance the device with the highest information contained in the measurements. Which other sensor device is capable of emitting more than ten million single measurements in a second with a size not larger than a usual pocket book? LIDAR scanners? Maybe[14]. But the setup and maintenance costs are not comparable to commonly used camera devices. Though, cameras will be used as the one and only sensor for all upcoming automation and control applications dealing with any perception task? Obviously not, but cameras in combination with other existing sensor devices play a key and dominant role in future, visionary multi-modal life-changing products like humanoid robotics or autonomously driving vehicles. In the following, the focus will be set on driver assistance functions and autonomously driving vehicles as one recently hyped application of computer vision research.

Motivation for Assisted and Autonomous Driving Functions

While the computer vision research is continuously ongoing in various application areas, e.g. visual testing of manufacturing plants or automated person tracking, autonomously driving vehicles and driver assistance systems have recently gained a high momentum in research, industry and press media. Intelligence-enhanced vehicles have shown in an explicit manner how multi-modal computer vision applications fundamentally change our everyday way of driving. Apart from convenience, autonomously driving vehicles and driver assistance systems address three highly important, societal discussion aspects: the security, ecology and psychological arousal of driving. According to a statistical report by the World Health Organization in 2010[15], approximately 1.24 million people died in traffic all over the world in a single year with India and China together reporting about half a million traffic deaths. The trend shows that the numbers are generally increasing due to higher traffic densities in the upcoming years[5]. A side-effect of the higher traffic densities and the higher number of traffic accidents are the environmental pollution caused by the large amount of congestions. An American study investigated and reported in this regard that 20% of

CO2 emissions could be saved by reduced traffic congestions[3]. As an exemplary calculation: considering a predicted traffic-caused CO2 emission of approximately 200 million tons in Germany in 2020[20], the savings could be tremendous approximately 4% of Germanys total CO2 emission. Further reductions could be achieved by constant velocity driving, e.g. by car/truck platooning. A second side effect is a perceivably higher drivers' psychological stress, e.g. stop and go driving in congestions or stress caused by potentially missing an important meeting[9, 13].

Pathetically envisioning the world wide traffic in 10 years, a high coverage of advanced driver assistance systems and automated driving functions[6] will have a thorough impact by saving many lives and it will additionally contribute to a healthier planet[1].

Sensors for Assisted and Autonomous Driving Functions

Autonomously driving vehicles are equipped with a multitude of different sensor systems which may be divided into two categories. The first category covers in-vehicle sensors that are required to control the vehicle, e.g. sensors to measure the steering angle, wheel movements, translation and orientational accelerations as well as engine and chassis internals. The second category includes sensors for higher-level driving functionalities like environment perception, vehicle localization and drive planning, e.g. RADAR (short-, middle- and long-ranging), LIDAR, real-time kinematics (RTK), GNSS, and, of course, cameras. Cameras have herein been applied in many configurations, e.g. single-view[7], dual-view[7], multi-view and catadioptric[18].

Requirements for Assisted and Autonomous Driving Functions

Advanced driver assistance systems are deployed in vehicles to support the driver in difficult and complex driving scenarios, e.g. driving in crowded urban areas or at night. The today's available driver assistance systems are the predecessor of autonomously driving vehicles and many of those assistance functions (or at least parts of them) will be re-used for the automated driving functions. While the setting is identical, e.g. structured roads, the requirements are somewhat different for both kinds of systems. Driver assistance systems support the driver and do not claim to operate in every situation – and do not have to according to Vienna Convention on Road Traffic. Autonomously driving vehicles in turn must be robust and reliable in every scenario including night, rain, snow and dust – failures in operation might have a catastrophic effect. This requires the use of redundantly designed sensor platforms that complement the weaknesses of other sensor devices, e.g. a LIDAR system or infrared camera complement cameras at night. Thus, the applied algorithms have to operate in multi-modal mode and fuse the measurements from a multitude of sensors simultaneously.

The list of available driver assistance systems (e.g. lane keeping assistant[16], lane departure warning[22], pedestrian recognition assistant[17], driver surveillance systems[6]) and autonomous driving demonstrations (e.g. Bertha Benz Memorial Drive[23]) that deploy and rely on the application of one or many

[1] A critical counterpart may legitimately argue that a better way of reducing death and CO2 pollution statistics is to reduce or avoid driving at all. The author fully agrees to this, but likewise believes that vehicles are meanwhile a strong indicator for freedom and flexibility (sometimes purely an indicator of luxury). And, this flexibility paired with a cheaper access to vehicles all over the world will unfortunately not decrease the number of vehicles entering our roads every day.

camera systems along with other input modalities is long and discussing the details of those systems and demonstrations is out of the scope of this essay.

Digital Maps for Autonomously Driving Vehicles

To cope with the high requirements for automated driving functions, recent approaches relied on a priori generated, highly accurate, digital road maps as world knowledge, e.g. to assist the localization or planning of the vehicle[12, 11, 19, 2, 8]. To generate and equally important to update the digital maps, algorithms are required that fuse recorded data from a multitude of sensors in an automated way, e.g. cameras[12, 11, 19], LIDAR[10] and GNSS data[4]. The map content differs depending on the application of the map, e.g. loosely connected point features[12, 11], lane markings[19], lane boundaries and traffic signs or planning information like permitted driving corridors and valid traffic rules, e.g. yielding in complex intersection scenarios. Cameras may be used in this context to detect and extract:

- lane markings, lane boundaries (e.g. curbs), road paintings (e.g. crosswalks)
- road lane details (e.g. number, widths and curvature)
- drivable free space, driving directions and corridors (e.g. via the detection of other traffic participants or pedestrians)
- traffic signs and their interpretation (e.g. speed limitations or navigational hints)

The information extracted by cameras may further be enhanced or re-validated by LIDAR sensors, while GNSS sensors provide global reference data to fix the obtained data in a global coordinate system.

Vision: Collaborative Map Generation

One of the main challenges of the digital road map approach is that accurate and up-to-date map material must be available of all roads to achieve a high operating range of driver assistance functions and automated driving functions. Compared to the approach of Agarwal to build a 3D model of high-attraction buildings in Rome by uncalibrated tourist photos[1], an ambitious approach of multi-modal map generation is to generate road maps collaboratively using uncalibrated image sequences and low-cost GNSS trajectories from a high number of smartphone recordings to extract the road information as described above. Beside the application for automated driving functions, such digital maps may be used for advanced navigational systems that warn before crossings or if pedestrians are crossing the street in complex driving scenarios and thus increase the security even before first fully automated driving functions are in production.

To conclude, autonomously driving vehicles strongly deploy the results of many years of intensive computer vision research, but there is still much work to do to leave the prognosticated societal influence and enter the world of safe, efficient and stress-free driving.

References

[1] Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S., and Szeliski, R. Building Rome in a Day. *Commun. ACM 54*, 10 (Oct. 2011), 105–112.

[2] Badino, H., Huber, D., and Kanade, T. Visual Topometric Localization. In *IEEE Intelligent Vehicles Symposium (IV)* (June 2011), pp. 794–799.

[3] Barth, M., and Boriboonsomsin, K. Real-World CO2 Impacts of Traffic Congestion. Working Papers, University of California Transportation Center, 2008. Available Online: `http://www.uctc.net/papers/846.pdf`, Accessed On: June, 13th 2014.

[4] Betaille, D., Toledo-Moreo, R., and Laneurit, J. Making an Enhanced Map for Lane Location Based Services. In *IEEE Conference on Intelligent Transportation Systems* (October 2008), pp. 711–716.

[5] Bundesamt, S. Statistisches Jahrbuch 2013. Tech. rep., October 2013. Language: German, Available Online: `https://www.destatis.de/DE/Publikationen/StatistischesJahrbuch/StatistischesJahrbuch2013.pdf`, Accessed On: June, 13th 2014.

[6] Chang, T., and Chen, Y. Driver Fatigue Surveillance via Eye Detection. In *IEEE International Conference on Intelligent Transportation Systems (ITSC)* (Oct 2014), pp. 366–371.

[7] Dang, T., Kammel, S., Duchow, C., Hummel, B., and Stiller, C. Path Planning for Autonomous Driving Based on Stereoscopic and Monoscopic Vision Cues. In *IEEE Internation Conference on Multisensor Fusion and Integration for Intelligent Systems* (September 2006), pp. 191–196.

[8] Deusch, H., Nuss, D., Konrad, P., Konrad, M., Fritzsche, M., and Dietmayer, K. Improving Localization in Digital Maps with Grid Maps. In *IEEE International Conference on Intelligent Transportation Systems* (October 2013), pp. 1522–1527.

[9] Hennessy, D. A., and Wiesenthal, D. L. Traffic congestion, driver stress, and driver aggression. *Aggressive Behaviour 45* (1999), 409–423. Available Online: `http://www.researchgate.net/publication/229863510_Traffic_congestion_driver_stress_and_driver_aggression`, Accessed On: June 13th 2014.

[10] Joshi, A., and James, M. Generation of Accurate Lane-Level Maps from Coarse Prior Maps and Lidar. *IEEE Intelligent Transportation Systems Magazine 7*, 1 (Spring 2015), 19–29.

[11] Lategahn, H., Schreiber, M., Ziegler, J., and Stiller, C. Urban Localization with Camera and Inertial Measurement Unit. In *IEEE Intelligent Vehicles Symposium (IV)* (June 2013), pp. 719–724.

[12] Lategahn, H., and Stiller, C. City GPS using Stereo Vision. In *IEEE International Conference on Vehicular Electronics and Safety (ICVES), 2012* (July 2012), pp. 1–6.

[13] Levy, J. I., Buonocore, J. J., and von Stackelberg, K. Evaluation of the public health impacts of traffic congestion: a health risk assessment. *Environmental Health 9* (2010). Available Online: `http://www.ehjournal.net/content/pdf/1476-069X-9-65.pdf`, Accessed On: June, 13th 2014.

[14] Lidar, V. Velodyne LIDAR HDL-64E. `http://velodynelidar.com/lidar/hdlproducts/hdl64e.aspx`.

[15] Organization, W. H. Global Health Observatory Data - Number of Road Traffic Deaths. http://www.who.int/gho/road_safety/mortality/traffic_deaths_number/en/.

[16] Risack, R., Mohler, N., and Enkelmann, W. A Video-Based Lane Keeping Assistant. In *Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE* (2000), pp. 356–361.

[17] Roncancio, H., Hernandes, A., and Becker, M. Vision-Based System for Pedestrian Recognition using a Tuned SVM Classifier. In *Engineering Applications (WEA), 2012 Workshop on* (May 2012), pp. 1–6.

[18] Schönbein, M., Rapp, H., and Lauer, M. Panoramic 3D Reconstruction with Three Catadioptric Cameras. In *Intelligent Autonomous Systems 12*, vol. 193 of *Advances in Intelligent Systems and Computing*. Springer Berlin Heidelberg, 2013, pp. 345–353.

[19] Schreiber, M., Knöppel, C., and Franke, U. LaneLoc: Lane Marking based Localization using Highly Accurate Maps. In *IEEE Intelligent Vehicles Symposium (IV)* (June 2013), pp. 449–454.

[20] Umweltbundesamt. CO2-Emissionsminderung im Verkehr in Deutschland. Tech. rep., May 2010. Language: German, Available Online: http://www.umweltbundesamt.de/sites/default/files/medien/461/publikationen/k3773.pdf, Accessed On: June, 13th 2014.

[21] Wright, T. *Visual Impact: Culture and the Meaning of Images*. 2008.

[22] Yu, B., Zhang, W., and Cai, Y. A Lane Departure Warning System Based on Machine Vision. In *Workshop on Computational Intelligence and Industrial Application* (Dec 2008), vol. 1, pp. 197–201.

[23] Ziegler, J., Bender, P., Schreiber, M., Lategahn, H., Strauss, T., Stiller, C., Dang, T., Franke, U., Appenrodt, N., Keller, C., Kaus, E., Herrtwich, R., Rabe, C., Pfeiffer, D., Lindner, F., Stein, F., Erbs, F., Enzweiler, M., Knöppel, C., Hipp, J., Haueis, M., Trepte, M., Brenk, C., Tamke, A., Ghanaat, M., Braun, M., Joos, A., Fritz, H., Mock, H., Hein, M., and Zeeb, E. Making Bertha Drive – An Autonomous Journey on a Historic Route. *IEEE Intelligent Transportation Systems Magazine 6*, 2 (2014), 8–20.