Proceedings CLME2017/VCEM
8º Congresso Luso-Moçambicano de Engenharia / V Congresso de Engenharia de Moçambique
Maputo, 4-8 Setembro 2017; Ed: J.F. Silva Gomes *et al*.; Publ: INEGI/FEUP (2017)

ARTIGO REF: 6594

EVALUATION OF GROUND-TRUTH SYSTEM FOR IN CAR HUMAN BODY POSE RECOGNITION

João Borges Silva[1], Vítor Coelho[2], Estevão Silva[2], Adriano Tavares[1], João Monteiro[1], Jaime Fonseca[1], José A. Mendes[1(*)]
[1]Universidade do Minho, Depart. Engª Eletrónica Industrial (DEI) - Guimarães, Portugal
[2]Bosch Car Multimédia, Braga, Portugal
[(*)]*Email:* jose.mendes@dei.uminho.pt

ABSTRACT

With the current market trend of automated driving (AD) cars, the human occupants will dedicate their time with other tasks that are still to be considered, creating a new paradigm for future in-car human interaction and safety. With this new trend for future development, we can point out a transversal technological need, and that is Occupant Monitoring. Human body posture recognition of the occupants inside a car is very important, where relevant algorithms outputs must be evaluated/benchmarked with a ground-truth system. This ground-truth concept is related to the measurements made in cartography, where the remotely captured information (e.g. by satellites) are validated with real measurements made on the ground.

In car human body pose recognition algorithms, based on data captured by ToF (time-of-flight) cameras are to be deployed in future cars. These cameras produce an image with depth data, and come from a technological concept named LIDAR (Light Detection and Ranging), [Radke 2013]. This technology doesn't measure the time of flight, but it measures the phase angle of each pixel, reducing the range and pixel accuracy (centimeter accuracy). In the end, it is possible to have a 2D frame with depth data for each pixel, where the depth information can be highly immune to lighting conditions. There are several algorithms developed in this area, that use both depth or RGB-D (red green blue depth) data, from which it can be identified different approaches: discriminative approaches where the pose is inferred from a trained dataset of body poses (Figure 1); generative approaches that try to estimate and predict the correct joint positioning by taking time into consideration (Figure 1); and hybrid approaches that try to create a synergy from both.

Fig. 1 - Body pose recognition: (Left) discriminative[Shotton et al. 2013], (Right) generative[Ye & Yang 2016]

The validation of in car human body pose recognition algorithms, requires the adoption of a ground-truth system that measures the orientation of several axis of the human body (at least 14 joints) and possibly its position, that doesn't suffer from occlusions created by the car interior, and also must provide higher performance than the algorithms. Currently, there are several technologies that can create a good ground-truth system, each with its own pros and cons. Inertial sensors (MEMS), [TEA 2016][Xsens 2016], allow the user to obtain the 3DoF orientation of the sensor, which can relate to the body joint, it haves a (1/10)° accuracy and doesn't need field of vision, unfortunately it can suffer from drift and magnetic sensitivity.

Vision based systems can also be a good alternative, such as, marker based systems (Vicon Mocap), [Vicon 2016], and markerless, [OrganicMotion 2016]. Vicon is a very mature and used system for ground-truth, but it suffers from high setup time and label occlusion problems (requires constant field of view). Markerless solutions do not require a lot of setup time but the accuracy is a lot worse, and also needs full field of view of the body. Electromagnetic sensors, [Polhemus 2016], achieves 6DoF (orientation and position) with the highest resolution, but are very sensitive to metals and electronics.

The main objective of this work is the creation of a toolchain (Figure 2) to evaluate the potential use of an inertial solution as a markerless and occlusion proof ground-truth system, which is a requirement for in car operation. Two inertial solutions are used concurrently with Vicon Mocap as a reference, and are evaluated with specific Key Performance Indicators (KPI) in mind, Angular and Cartesian joint deviation. There will be two specific procedures, where the KPI will be evaluated, the procedures are: Full body evaluation (FBE), and Individual segments evaluation (ISE) where we avoid the kinematic forwarding error, and only evaluate the specific joint error.

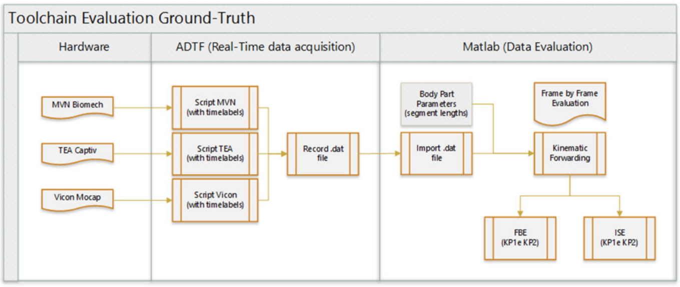

Fig. 2 - Ground-truth evaluation toolchain

The selected inertial suit, in conjunction with the ToF camera, will allow us to create a real dataset, that stores both RGB-D data with synced ground-truth data. This can be later on used for human body pose recognition algorithmic evaluation/benchmark.

REFERENCES

[1]-OrganicMotion, 2016. OrganicMotion @ www.organicmotion.com. Available at: http://www.organicmotion.com/.

[2]-Polhemus, 2016. electromagnetics @ polhemus.com. Available at: http://polhemus.com/applications/electromagnetics.

[3]-Radke, R.J., 2013. COMPUTER VISION FOR VISUAL EFFECTS C. U. Press, editions.

[4]-Shotton, J. et al., 2013. Real-time human pose recognition in parts from single depth images. Studies in Computational Intelligence, 411, pp.119-135. Available at: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/BodyPartRecognition.pdf.

[5]-TEA, 2016. captiv-motion @ teaergo.com. Available at: http://teaergo.com/site/en/products/manufacturers/tea/captiv-motion.

[6]-Vicon, 2016. Motion Capture. Available at: www.vicon.com/what-is-motion-capture.

[7]-Xsens, 2016. MVN BIOMECH @ www.xsens.com. Available at: https://www.xsens.com/products/mvn-biomech/.

[8]-Ye, M. & Yang, R., 2016. Real-time Simultaneous Pose and Shape Estimation for Articulated Objects Using a Single Depth Camera., 38(8), pp.1517-1532.