Rescue Method Based on V2X Communication and Human Pose Estimation

Azamat Zarkeshev[1]*, Csaba Csiszár[1]

[1] Department of Transport Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics, H-1521 Budapest, P.O.B. 91, Hungary
* Corresponding author, e-mail: zarkeshev.azamat@mail.bme.hu

Abstract

Advances in science and technologies not only improve our workflow and increase the quality of life but also change the way emergency rescue operates. The eCall system is the evidence of such a new evolutionary step. However, studies are not available about 'smart' rescue methods of independent human (neither pedestrian nor a driver) who turned out to be on the roadside for one or another reason and need first aid. The research questions were how to increase the quality and speed of the rescue process by using Vehicle-to-Everything (V2X) communication and the human pose estimation function and how to increase the reliability of transmitted information. In the article, we presented an overview about the current state-of-the-art in human pose estimation technology, discussed current methods of rescue, concept, and disadvantages of the current approach, and finally proposed our own method of rescue based on blending of vehicular communication networks with advancements in human pose estimation function. The research novelty in a scientific sense is a concept of a new information management method, where Autonomous Vehicles (AVs) act as witnesses itself without any human intervention. An example of SOS Packet Format has been designed. We proposed a novel view on the future ambulance.

Keywords

autonomous vehicles, ambulance service, intelligent transportation systems, human pose estimation

1 Introduction

Autonomous driving is a rapidly developing and promising technology. In the past five years, it became a real game changer in the automotive industry. However, it is still considered as a technology of the future [1] and nobody precisely knows how the final product will look like. Autonomous transportation system is a wide area to explore in terms of traffic distribution problem [2], vehicle to infrastructure communication and vehicle to human interaction. A lot of current studies focused on Autonomous Vehicle's (AV) collision avoidance system. The main issue of such researches is how to improve pedestrian recognition and tracking for cases where AVs are operated in mixed traffic with pedestrians in order to avoid collision with them. However, studies are not available about detection of independent human (neither pedestrian nor a driver) who turned out to be on the roadside for one or another reason and need first aid.

In our study, we propose a novel method of interaction of AVs and ambulance service, where AVs serve as witnesses

of an incident with human health and alert the ambulance center. This method can significantly decrease rescue time applying Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) communications and human pose estimation. In other words, a new incident notification paradigm is provided by this method.

The objective of the article is to elaborate new solution of 'saving human life' via faster communication with ambulance service than regular method using cellular phone. In this study, the transportation infrastructure is considered as a fully automated network within 'Smart city'. One of the key factors of the proposed method is the human pose estimation function, which helps to recognize if a person falls down, typically through illness or injury. Studies focused on technologies that help to recognize if the person has a strange position of the body which is not peculiar to a healthy person were reviewed in Section 2. The remainder of the paper is structured as follows. In Section 3, the overview of the existing rescue methods is

provided. The main features of the new method of rescue are determined and summarized in Section 4. The paper is completed by the concluding remarks, including further research directions.

2 Human pose estimation technologies

Currently the 'Eyes of AVs' are LiDAR (Light Detection and Ranging), RADAR (Radio Detection and Ranging) and cameras which are required and applied for safe and reliable movement of the vehicle, but not used for human pose estimation purposes [3].

In our case, pose estimation is an important function to understand the state of a person in a public urban environment. For instance, unusual postures such as crouching or lying could be a sign of emergency. In this concern, our proposed method requires installation of sensors on AVs for human pose identification. State of the art technologies in this field which may be implemented in our method are described below.

Human pose estimation is one of the difficult tasks in computer vision because of large degrees of freedom of human structure and frequent occlusions. Various methods [4, 5] have been proposed for reliable estimation. Image-based methods for joint position estimation [6, 7] and part segmentation [8] have been proposed due to recent advances in deep learning techniques.

Depth image-based pose estimation has an advantage in terms of privacy and insensitivity to illumination conditions. A human pose estimation method using depth-based features with a random forest classifier has been also developed [9]. Although, the method shows a nice performance, its applicability to unusual poses and heavily occluded situations is rather limited. A method of generating human depth images with pixel-wise body part labels by a combination of computer graphics and motion capture techniques has been proposed [10, 11]. It has been shown that a deep neural network trained by using generated images can recognize well a variety of human poses in real scenes. Human joint estimation and behavior recognition algorithms using depth information only were also presented [12]. In the case of the human joint estimation algorithm, a method that integrates a geodesic graph and a support vector machine based human joint estimator has been proposed.

The other group of researchers presents the first real-time method to capture the full global 3D skeletal pose of a human in a stable, temporally consistent manner using a single RGB camera, see Fig. 1. The method combines a

Fig. 1 The real-time 3D pose estimation method provides a natural motion interface, e.g. for sport games (source: [13])

new convolutional neural network (CNN) based on pose regressor with kinematic skeleton fitting. A real-time kinematic skeleton fitting method uses the CNN output to yield temporally stable 3D global pose reconstructions on the basis of a coherent kinematic skeleton. This makes an approach the first monocular RGB method usable in real-time applications such as 3D character control - thus far, the only monocular methods for such applications employed specialized RGB-D cameras. The method's accuracy is quantitatively equal with the best offline 3D monocular RGB pose estimation methods. It has been stated that the approach is more broadly applicable than RGB-D solutions, i.e., it works for outdoor scenes, community videos, and low quality commodity RGB cameras [13].

As we can see analysis of human behavior through visual information has been a highly active research topic in the computer vision community. Some authors proposed their own solutions on this topic, while some of them [14] reviewed the main published research on the use of depth imagery for analyzing human activity. Many existing studies focus on body part detection and pose estimation. A growing research area addresses the recognition of human actions.

The aim of the literature review was not to choose an exact technology of human pose estimation being suitable for our proposed method described in Section 4 but to emphasize successfulness of researches on this topic. Considering the fact that this research direction is rapidly

and continuously developing, human pose estimation function can be implemented by sensors and cameras installed on the stage of AVs manufacturing.

3 Existing methods of rescue

In case of emergency, an ambulance should be dispatched to the scene without delay. In order to provide appropriate first aid or to improve the chance of survival, it is desirable to reduce the response time of rescue teams. Thus, once the incident has occurred, it is crucial to quickly summon the ambulance.

Fig. 2 shows the method of incident notification used nowadays, where a witness summoning for an ambulance to the incident site by means of cellular phone. The witness gives information about the location of the victim and a description of the circumstances. Once the ambulance service is notified, the rescue effort is started.

This method is the same as the conventional rescue method in case of car accidents. However, in the case of car accidents, it was improved due to an implementation of the crash notification systems such as eCall shown in Fig. 3.

When an accident occurs, a call is made to a Public Safety Answering Point (PSAP) in order to send information about the accident and to ask for help. This system acts based on two key features [15]:

- Send a Minimum Set of Data (MSD): Standardized set of data, containing information about the accident, such as time, accurate location and driving direction (using a GNSS) and description of the vehicle.
- Immediate set-up of a voice call between the vehicle passengers and the PSAP: The PSAP operator takes the necessary measures, forwarding the information to the appropriate entities.
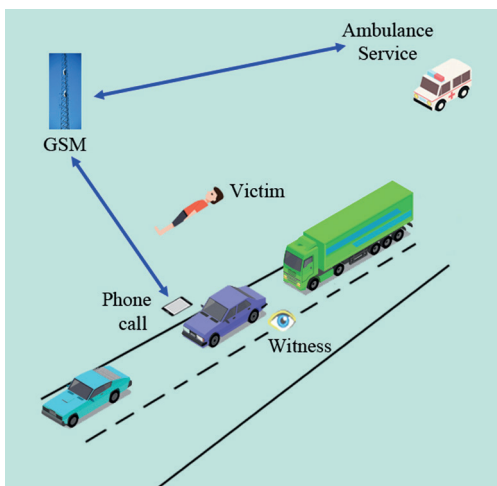
The eCall system can be triggered either manually or automatically. To manual activation, a passenger of the vehicle pushes the emergency button. The vehicle triggers the automatic activation based on the data gathered from the sensors (e.g. airbags). The information collected is used to generate an eCall MSD. eCall [16] is one of the most important road safety efforts made under the European Union's eSafety initiative. It is mandatory since 2018 for passenger cars and light goods vehicles registered in the European Union.

4 Novel method of rescue

The eCall system is an example that advances in science and technologies are changing the way emergency rescue operates. A faster and more efficient rescue will increase the chances of survival and recovery for injured victims. Therefore, we propose a novel method of rescue of sick or injured people on the roadside based on "technologies of the future" explained in Subsection 4.1.

4.1 Method description

Implementation of the novel method is possible due to qualitative information management and the following technologies:

- V2I interaction carries out communication between vehicles and shore infrastructure, which can be a cellular or wireless antenna [17]. In some cases, V2I communication serves as a base by different integrated projects and subprojects [18].
- V2V communication that uses dedicated short-range communication (DSRC), the standard established by the International Organization for Standardization (ISO) and the Federal Communications Commission
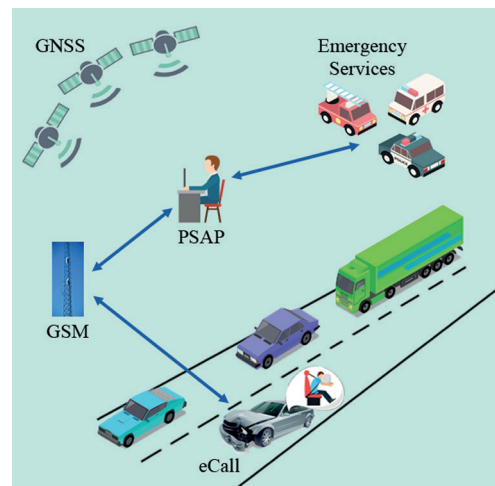

Fig. 2 Conventional method of rescue using just cellular phone

Fig. 3 Method of eCall service

(FCC). There is an IEEE standard, namely the 802.11 standard family for these communication purposes. The 802.11p is dedicated for vehicular communication. The diapason is about 300 m (approximately 10 seconds) at highway speeds. V2V comprises a wireless network where automobiles send messages to each other with information about what they are doing, what is happening around them and what they will be doing [19].

- Human pose estimation technology which may serve as a tool for emergency situation sensing.

If a person is lying on the ground for some continuous time or has a crouching posture for quite a long time, it is a rather significant reason to worry. Using human pose estimation function, AVs identify such a person on the roadside by special equipment (sensors, cameras). By using vehicular communications, cars acting as witnesses send alert and other important information about the incident to nearby vehicles and to the nearest wireless base stations called Roadside Units (RSUs). After that, the gathered information is delivered to the PSAP using cellular or wired connection. PSAP estimates received data about the victim and the appropriate rescue resources before summoning ambulance service. The PSAP then takes actions such as notifying ambulance service and dispatching ambulance vehicles to the incident location. To define that human behaves strangely due to real sickness and it is not a misunderstanding, the PSAP awaits several signals from several AVs. Thus, if the first AV sends an alert of an incident and after some time the next AV confirms this information by sending the same alert, this situation can be evaluated as an emergency. Although the time gap between two consequent vehicles could be different because of AV's unpredictable movement, PSAP will be on the alert after the first received signal. In addition, vehicular networks can allow faster notification of any incident occurring on the roadside, since sensing and propagation of incident information are done on-the-spot in real-time via multi-hop V2V communications. Real-time information collected from multiple connected vehicles generally leads to better outcomes due to AVs are able to check each other's data. Besides, PSAP dispatchers receive data about a victim made by AV's camera in order to analyze the situation precisely and identify whether he/she is unconscious or not, etc. Technology should be compatible with the air interfaces under existing implementations or standardizations and signaling protocol. V2I might use any of the 3G/4G currently available cellular technologies and V2V communication might be compatible with the 802.11p standard [20]. The concept of the proposed method is shown in Fig. 4.

4.2 Information management

The information shown in Table 1 is essential to be managed in ambulance service.

The injuries might be coded using the Abbreviated Injury Scale (AIS). It is an anatomical scoring system that provides an accurate ranking of the severity of injury. In the AIS scale, injuries are ranked on a scale from 1 to 6, with 1, 5, and 6 representing minor, severe, and unsurvivable injury, respectively [21].

Fig. 5 illustrates an example of SOS Information Packet Format that includes all the above-mentioned information.

With knowledge of information from SOS Packet, the work efficiency of paramedics and physicians can be improved in a significant way. The first step of treatment can be initiated by retrieving information about the victim status.
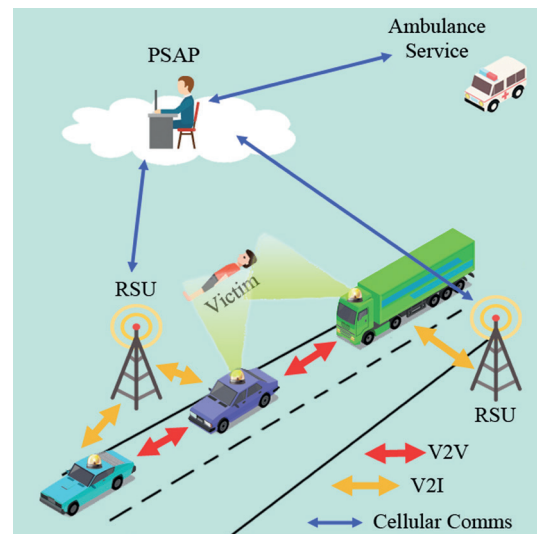

Fig. 4 Proposed method

Table 1 Information to be sent after victim identification

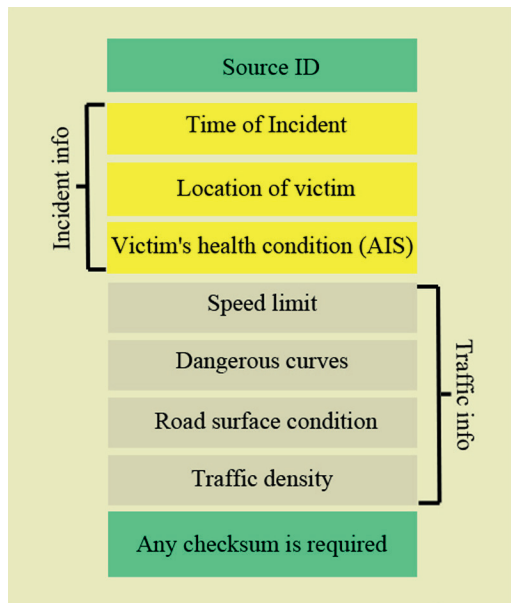| Information	Purpose
Timestamp	The exact time when the victim was identified by an AV
Geographical position	The exact location of the victim
Severity of the health condition (as far as the vehicle can evaluate it or diagnose any anomalies)	Parameters about a victim's health condition, such as conscious or unconscious, bleeding or not bleeding, bone injuries, etc.
Traffic information	Traffic condition (e.g. traffic congestion, speed limits, weather forecast, etc.) in order to facilitate ambulance service to drive to the incident site

Fig. 5 SOS Information Packet Format example

Fig. 6 Current method versus the novel method in terms of time loss

All this important information should be sent to the PSAP through the RSUs by each AV that became a witness of an incident. Thereafter, these data shall be processed by dispatchers and their workstations to decide the resources needed to correctly take care of and manage the incident.

4.3 Comparison of the conventional and novel method of rescue

Table 2 represents a comparison of the traditional and novel method of rescue according to several aspects.

A comparison of the traditional and novel method in terms of time loss is shown in Fig. 6.

Considering the above-mentioned aspects, it can be summarized that the main advantages of the new method compared to the existing one are the eliminations of socio-
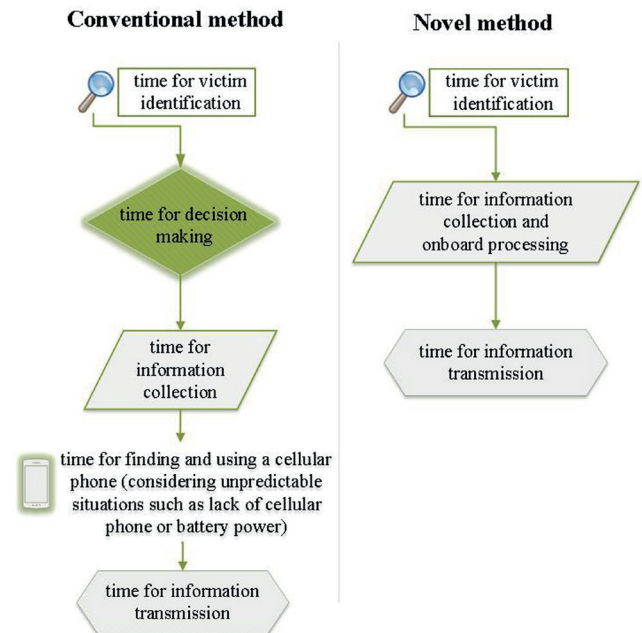
psychological parameter and the human factor. The validity of this statement can be approved by a couple of examples happened in real life. In 2004, on south-east London, bleeding and an unconscious 25-year-old female was lying next to the curb while 12 drivers driving round the woman without stopping to help [22]. Fig. 7 represents the footage was captured from a bus camera, whose driver stopped to aid the woman.

The real-life experiment [23] conducted by traffic police in 2017 in Brest (Belarus) not only reflects the disappointing statistics of most drivers' indifference to another's misfortune but also provide real measurements enabling to perform an average time calculation for victim identification in case of the traditional method of rescue. The police

Table 2 Current method versus the novel method according to the main aspects

Aspects	Conventional method	Novel method
Socio-psychological	Human behavior is unpredictable; it is unknown whether witnesses initiate alert to the ambulance service or will ignore the incident	AVs do not have a socio-psychological parameter; AVs initiate alert to the ambulance service every time when it is needed
Human factor	A human may forget or may not know the emergency number. He/she may not have a cellular phone or may have a lack of battery power or even get confused and transmit inaccurate information	Victim identification, data collection, and data transmission are performed automatically by several different AVs; AVs check each other's data
Visual perception	Any problems with human vision (e.g. poor eyesight) might be an obstacle for human identification. A human may also miss information, which may be vital for the rescue operation	Identification of human depends on the technical capabilities of the equipment (sensors, cameras) and on the environmental circumstances (e.g. dust, fog, surface conditions, etc.)
Weather	Harsh weather might be an obstacle for human identification	Identification of human depends on the technical capabilities of the equipment (sensors, cameras) and on the environmental circumstances (e.g. dust, fog, surface conditions, etc.)

Fig. 7 An injured person on the roadside in London (source: [22])

Fig. 8 Cars passing by a mannequin lying on the roadside (source: [23])

members put a mannequin very similar to a real person on the roadside (Fig. 8) and began to monitor how many cars would stop to help him.

Already in the first minutes of the experiment, 4 cars stopped near the 'man' lying on the sideline. But over the next 30 minutes, almost no one stopped. In just 1 hour of the experiment, 10 drivers stopped. More than 100 cars drove past the victim without stopping to help. Only 2 calls during this time arrived on emergency service number with information about a person lying on the road. The traffic police also reported that in almost one month 11 people died on this road, most of them pedestrians and cyclists.

The outputs of the experiment represent that the rescue process which includes human factor has stochastic character. However, considering the fact that in the novel method every single car will identify the victim continuously, and using the current average victim identification time for a standard method, comparative rate of victim identification time between conventional and proposed method may be calculated by the following formula:

$$C_r = \frac{((t_e / d_s) - (t_e / AV))}{(t_e / AV)} \times 100\% \quad \approx 733\%, \quad (1)$$

where,

t_e - experiment time = 60 min.,
d_s - drivers who actually stopped (+ 2 calls) = 12,
AV - drivers who could potentially stop = 100.

Thus, the process of victim identification using the novel method could be faster than the conventional method on more than 700 % or approximately 4.4 minutes. It may essentially decrease the time of the ambulance service call which in turn leads to an increase in chances to save a human life.

Some studies [24, 25] confirm that ambulance service delays influence death rating during a rescue mission. For instance, according to a survey of 2017, approximately 20 % of the world's population died because of ambulance late coming to the incident side [25]. At the same time, these studies try to solve this problem by proposing improvements in the ambulance service itself using Ambulance-to-Vehicle communications or by using transmitter and receiver. But our method emphasizes the high importance of the timely call of ambulance service that precedes the arrival of an ambulance vehicle. It is a time-sensitive process when even the 1-second delay may become a big threat to human life and even cause death.

Thus the fewer time-consuming acts are a presence in the ambulance calling process the faster Ambulance Service could be notified about the incident. As a result, the victim will receive faster first aid.

5 Conclusions

In the future, the current incident notification paradigm will change with the introduction of vehicular networks. Intelligent Transportation Systems are capable of improving the timeliness and responsiveness of roadside ambulance services due to a combination of V2V, V2I communications, and human pose estimation technology.

The main contribution of the paper is the concept of a new rescue method based on advanced information management. Our results have high relevance for innovation and development. The key findings are as follows:

- The reliability of the rescue process is enhanced due to the elimination of socio-psychological parameter and human factor.
- The reliability of managed information is enhanced due to AVs that are able to check each other's data.
- The rescue process is accelerated by eliminating such steps as decision making and using a cellular phone manually.

Lessons learnt: as the integration of information systems is an enormous and complex task, not only logically but also physically, future research is expected to proceed in several directions. Our future plans include:

- Elaboration of a method for dispatching ambulance vehicles to the incident location considering several relevant data (e.g. real-time traffic information).

- Investigation of the economic consequences of the proposed method (e.g. cost of equipment installation, benefits, etc.)
- Continue the investigation into this new area of inquiry by detailed study each of the novel method stages.

Acknowledgment

References

[1] Tettamanti, T., Varga, I., Szalay, Z. "Impacts of Autonomous Cars from a Traffic Engineering Perspective", Periodica Polytechnica Transportation Engineering, 44(4), pp. 244–250, 2016.
https://doi.org/10.3311/PPtr.9464

[2] Pauer, G., Török, Á. "Static System Optimum of Linear Traffic Distribution Problem Assuming an Intelligent and Autonomous Transportation System", Periodica Polytechnica Transportation Engineering, 47(1), pp. 64–67, 2019.
https://doi.org/10.3311/PPtr.11548

[3] Kala, R. "2 - Basics of Autonomous Vehicles", In: On-Road Intelligent Vehicles, Motion Planning for Intelligent Transportation Systems, Butterworth-Heinemann Ltd., Oxford, UK, 2016, pp. 11–35.
https://doi.org/10.1016/B978-0-12-803729-4.00002-7

[4] Moeslund, T. B., Hilton, A., Krüger, V. "A survey of advances in vision-based human motion capture and analysis", Computer Vision and Image Understanding, 104(2–3), pp. 90–126, 2006.
https://doi.org/10.1016/j.cviu.2006.08.002

[5] Liu, Z., Zhu, J., Bu, J., Chen, C. "A survey of human pose estimation: The body parts parsing based methods", Journal of Visual Communication and Image Representation, 32, pp. 10–19, 2015.
https://doi.org/10.1016/j.jvcir.2015.06.013

[6] Toshev, A., Szegedy, C. "DeepPose: Human Pose Estimation via Deep Neural Networks", In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 1653–1660.
https://doi.org/10.1109/CVPR.2014.214

[7] Cao, Z., Simon, T., Wei, S. E., Sheikh, Y. "Realtime Multi-person 2d Pose Estimation Using Part Affinity Fields", In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 1302–1310.
https://doi.org/10.1109/CVPR.2017.143

[8] Oliveira, G. L., Valada, A., Bollen, C., Burgard, W., Brox, T. "Deep learning for human part discovery in images", In: 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 2016, pp. 1634–1641.
https://doi.org/10.1109/ICRA.2016.7487304

[9] Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M., Moore, R. "Real-time human pose recognition in parts from single depth images", Communications of the ACM, 56(1), pp. 116–124, 2013.
https://doi.org/10.1145/2398356.2398381

[10] Nishi, K., Demura, M., Miura, J., Oishi, S. "Use of Thermal Point Cloud for Thermal Comfort Measurement and Human Pose Estimation in Robotic Monitoring", In: 2017 IEEE International Conference on Computer Vision Workshop (ICCVW), Venice, Italy, 2017, pp. 1416–1423.
https://doi.org/10.1109/ICCVW.2017.168

[11] Nishi, K., Miura, J. "Generation of human depth images with body part labels for complex human pose recognition", Pattern Recognition, 71, pp. 402–413. 2017.
https://doi.org/10.1016/j.patcog.2017.06.006

[12] Kim, H., Lee, S., Kim, Y., Lee, S., Lee, D., Ju, J., Myung, H. "Weighted joint-based human behavior recognition algorithm using only depth information for low-cost intelligent video-surveillance system", Expert Systems with Applications, 45, pp. 131–141, 2016.
https://doi.org/10.1016/j.eswa.2015.09.035

[13] Mehta, D., Sridhar, S., Sotnychenko, O., Rhodin, H., Shafiei, M., Seidel, H. P., Xu, W., Casas, D., Theobalt, C. "VNect: Real-time 3D human pose estimation with a single RGB camera", Journal ACM Transactions on Graphics, 36(4), 2017.
https://doi.org/10.1145/3072959.3073596

[14] Chen, L., Wei, H., Ferryman, J. "A survey of human motion analysis using depth imagery", Pattern Recognition Letters, 34(15), pp. 1995–2006, 2013.
https://doi.org/10.1016/j.patrec.2013.02.006

[15] Cabo, M., Fernandes, F., Pereira, T., Fonseca, B., Paredes, H. "Universal Access to eCall System", Procedia Computer Science, 27, pp. 104–112, 2014.
https://doi.org/10.1016/j.procs.2014.02.013

[16] European Union "An automated eCall system", [online] Available at: https://europa.eu/investeu/projects/automated-ecall-system_en [Accessed: 31 March 2018]

[17] Meneguette, R. I., De Grande, R. E., Loureiro, A. A. F. "Vehicle-to-Infrastructure Communication", In: Intelligent Transport System in Smart Cities, Urban Computing, Springer, Cham, Switzerland, 2018, pp. 53–77.
https://doi.org/10.1007/978-3-319-93332-0_4

[18] Krausz, N., Lovas, T., Barsi, Á. "Radio Frequency Identification in Supporting Traffic Safety", Periodica Polytechnica Civil Engineering, 61(4), pp. 727–731, 2017.
https://doi.org/10.3311/PPci.10114

[19] Martinez, F. J., Toh, C. K., Cano, J. C., Calafate, C. T., Manzoni, P. "Emergency Services in Future Intelligent Transportation Systems Based on Vehicular Communication Networks", IEEE Intelligent Transportation Systems Magazine, 2(2), pp. 6–20, 2010.
https://doi.org/10.1109/MITS.2010.938166

[20] Murillo, E. I., Poveda, H. E., Jo, K. H., Hernández, D. C. "Evaluation of IEEE 802.11n and IEEE 802.11p Based on Vehicle to Vehicle Communications", In: 2018 11th International Conference on Human System Interaction (HSI), Gdansk, Poland, 2018, pp. 491–497.
https://doi.org/10.1109/HSI.2018.8431337

[21] Copes, W. S., Sacco, W. J., Champion, H. R., Lawnick, M. M., Gann, D. S., Gennarelli, T., MacKenzie, E., Schwaitzberg, S. "Progress in characterizing anatomic injury", The Journal of Trauma, 30(10), pp. 1200–1207, 1990.
https://doi.org/10.1097/00005373-199010000-00003

[22] BBC NEWS "Would you help an injured person?", [online] Available at: http://news.bbc.co.uk/2/hi/talking_point/3701848.stm [Accessed: 02 October 2004]

[23] MEDIA BREST.by "В Бресте провели эксперимент. Сколько водителей остановились помочь лежащему на обочине человеку?" (An experiment was conducted in Brest. How many drivers stopped to help a person lying on the sidelines?), [online] Available at: https://mediabrest.by/news/sobytia/v-breste-proveli-eksperiment-skolko-voditeley-ostanovilis-pomoch-lezhaschemu-na-obochine-cheloveku [Accessed: 03 October 2017] (in Russian)

[24] Suthaputchakun, C., Cao, Y. "Rescue Mission Enhancement through Ambulance-to-Vehicle Communications", In: 2018 28th International Telecommunication Networks and Applications Conference (ITNAC), Sydney, NSW, Australia, 2018, pp. 1–6.
https://doi.org/10.1109/ATNAC.2018.8615129

[25] Dhianeswar, R., Naveen Kumar, T., Kishore, N., Ashwinn, K., Sumathi, S. "Enhanced Ambulance Service Using Transmitter and Receiver", In: 2018 IEEE 3rd International Conference on Communication and Information Systems (ICCIS), 2018, pp. 154–157.
https://doi.org/10.1109/ICOMIS.2018.8645050