



24th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2013

Possibilities of using Speech Recognition Systems of Smart Terminal Devices in Traffic Environment

Sinisa Husnjak*, Dragan Perakovic, Ivan Jovovic

University of Zagreb, Faculty of Transport and Traffic Sciences, 10000 Zagreb, Republic of Croatia

Abstract

The ability to recognize human speech has always been an area of interest of people primarily because of the large range of applications in almost every segment of society. The development of science and technology made visible improvements in the capabilities and the quality of recognition of the human voice/speech using some kind of terminal devices. Speech recognition enables devices to adapt voice information in understandable form, which means complete identification and understanding of information. The primary purpose of recognizing of human speech is the ability to customize information (which are intended to people) for using with devices. The aim of voice recognition systems is to provide links and ease of communication between the device and the man and the development of further possible applications. Given the current trends in the use of smart terminal devices and technical and technological opportunities that they offer, it is inevitable to study the potential of these devices in the area of human voice recognition. The purpose of the paper is to show the possibilities of human voice recognition using smart terminal devices. Furthermore, all of this will be analyzed primarily in the transport(car) environment with emphasis on the potential advantages and disadvantages of such tools and principles.

© 2014 The Authors. Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).
Selection and peer-review under responsibility of DAAAM International Vienna

Keywords: Voice recognition; smart terminal devices; mobile applications; car user interface; driving safety

1. Introduction

Time spent in car drivers are trying to use for all the actions that can then be done in the car; such as listening to music, information services (radio news) and/or making telephone calls. Generally, drivers want to be informed and they want to have fun when they spend their time in the car. With the increasing use of mobile, portable and

* Corresponding author. Tel.: +385-1-2457945; fax: +385-1-2457918.
E-mail address: sinisa.husnjak@fpz.hr

navigation devices, increasingly emphasizes the desire to manage these devices while driving a car. This fact poses a serious security problem with regard to that the use of mobile devices while driving distracts the attention of the driver to his primary task - to drive a car safely. Speech recognition systems are one of the most desirable functions and features of the car. That way of communication represents an action that in a significant way does not affect the driver and does not interfere with it while driving. With the development of technology, and in particular the progress of hardware and software of the of smart terminal devices and the fact that it is precisely these devices almost always with the user, it is necessary to recognize the possibility of applying and upgrading devices with new functionalities in the form of improved quality of life and meeting the communications needs of people - with emphasis on the safety of their use.

2. Voice recognition systems

According to [1], voice recognition is the ability of a machine or program to identify words and phrases in spoken language and convert them to a machine-readable format.

Basic characteristic of rudimentary speech recognition software is that it has a limited vocabulary of words and phrases and may only identify these if they are spoken very clearly. More sophisticated software has the ability to accept natural speech. Speech recognition applications include voice search, call routing, voice dialing and speech-to-text.

2.1 Working principle

There are two types of speech recognition systems:

- Speech recognition systems based on speech recognition methods that operate directly on the terminal
- Speech recognition systems based on speech recognition methods that operate on a server

Those that use a server transmit the sound signal or sound feature values to the server, which runs a speech recognition engine and returns the voice/speech recognition result back to the terminal device.

Speech recognition systems that operate on the terminal device have a speech recognition engine on that terminal device. Speech recognition on a terminal device is restricted to relatively small vocabularies due to limitations in processing and power consumption. Also, they are not affected by communication conditions such as delay or being out of range. [2]

There are three major categories of services for which the majority of speech recognition systems are used for:

- Automated serving
- Call routing
- Value added services

The typical architecture of speech recognition system is shown in Fig. 1. At the first step human speech is processed to the segment that extracts the speech patterns by analyzing their frequency spectrum. Speech recognition system then compares the patterns of speech and aligns with pre-made acoustic and language model.

This determines the range of morphemes which faithfully reflects the mentioned voice samples. Acoustic model expresses the link between the characteristics of the speech and phonemes. Language model expresses similarity between created morpheme and previously created base of morphemes.

Precision and accuracy of speech recognition systems depends primarily on made language and acoustic models, ie what is the similarity of these models with speech samples that needs to be analyzed. Likewise, it is necessary to include a large selection of words and word variations when creating and refining existing language and acoustic models.

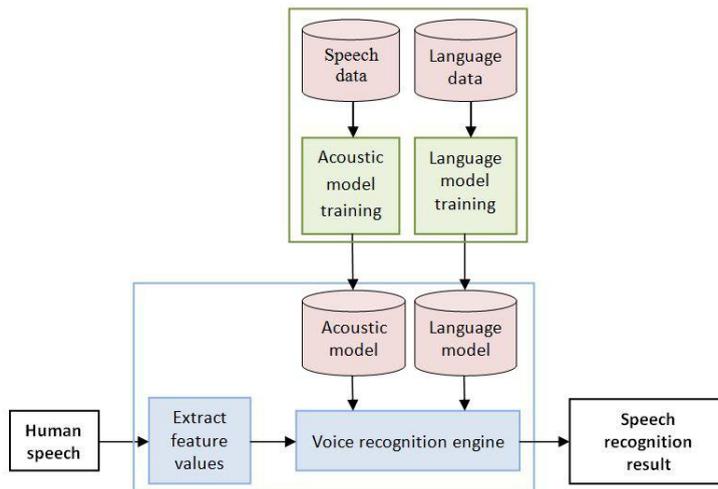


Fig. 1. The architecture of a typical voice recognition system, [2].

2.2 Possible interference

The most significant interferences to the good work of speech recognition system that can be identified are: individuality of human speech (gender, age), dialect, speaking rate, phonetic/prosodic context, background noise, microphone characteristics, directional characteristics of the speech source etc.

Fig. 2. presents the major categories of sources of interference and variations in human speech recognition. Speech recognition systems should be capable of robust, speaker-independent or speaker-adaptive, continuous speech recognition.

According to [3], certain categories of interference consist of the following potential sources of interference, shown in Table 1.

Table 1. Possible causes of acoustic variations in speech, [3].

INTERFERENCE CATEGORY	POTENTIAL INTERFERENCE
<i>Speaker</i>	a. Gender
	b. Voice quality
	c. Dialect
	d. Stress/emotions
	e. Speaking rate
<i>Microphone</i>	a. Electrical noise
	b. Distortion
	c. Directional characteristics
<i>Noise</i>	a. Other speakers
	b. Background noise
	c. Reverberations
<i>Channel</i>	a. Distortion
	b. Noise
	c. Echoes
	d. Dropouts

For voice recognition systems it is crucial to establish methods that are robust against voice variations due to individuality, the physical and psychological condition of the speaker, telephone sets, microphones, network characteristics, background noise, speaking styles etc.

Mobile environments are highly influenced by ambient noise, which may cause a significant deterioration of speech recognition performance, and there is a strong demand for the improvement of the performance. [4]

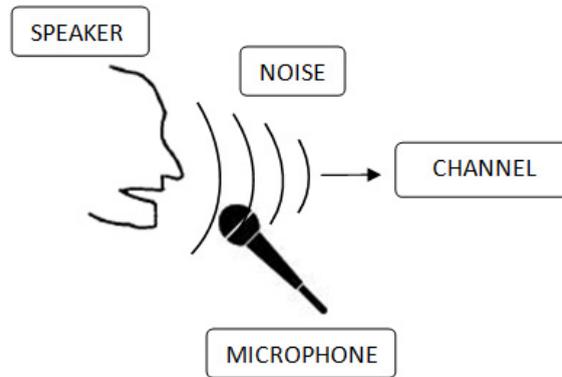


Fig. 2. Main causes of acoustic variations in speech, [3].

3. Functionality and usage of smartphone devices

Although there is no exact definition of what a smart terminal device (smartphone) is, it can be said that smartphone is a device that extends the capabilities of conventional mobile terminal device [5].

Additional features that are expected from smart terminal devices are not strictly defined and they change over time, which makes the definition of a smartphone harder. For example, several years ago the Global Positioning System (GPS) in the phone was considered as function possessed only by smartphones. Today this function have some terminal devices that are not considered as smart terminal devices [6].

Key features of the smart terminal device, according to [7], are: operating system (OS), applications, full QWERTY keyboard, permanent access to the Internet and ability to exchange messages.

3.1 Functionality of smartphone devices

Today's smartphones have touch screen that allows users intuitively usage of their mobile devices. Users have their smartphones always with them and they are often used as a substitute for GPS navigation, digital camera or video camera, media player, etc. Today's smartphones have more processing power and memory space than the personal computers from ten years ago. [8]

The fact is that features such as larger smartphone screens, better and stronger processing power, increased memory space, better connectivity and Internet access, quality of software, selection of the applications and great multimedia features of newer generation of smartphones represent almost a small computer, and affect on their mass use [9].

Those characteristics and everyday improvements of the device's hardware and software are only preconditions for the use of some new features and/or applications of smartphone devices. With the development of mobile operating systems such as iOS and Android OS the idea of using speech recognition system on smartphone devices have been created. Given that the benefits of such systems are immediately recognized, it is not surprising the fact that the number of applications that are based on the mentioned functionality has daily increase and usage.

Future characteristics and functionality of smartphones include:

- Hardware nanotechnology
- Smaller devices
- Flexible Smartphone Designs
- Augmented reality
- Artificial intelligence
- Applications improvement
- Higher CPU speed processing

3.2 Smartphone usage

One reason that smartphones and smartphone apps are so useful is that they can integrate intimately with our personal lives. Smartphones offer a platform for new kinds of applications and services and the devices are more and more integrated into the users' everyday lives.

Prediction of analyst firm Gartner indicate an almost linear increase in the use of smart terminal devices by the year 2015. According to their data, the average increase in the number of smart terminal devices is approximately 160 000 devices per year. [10]

Fig.3.. shows a forecast for global smartphone shipments from 2010 to 2017. According to the forecast, more than 1.1 billion smartphones will be shipped worldwide in 2014. By 2017, the number is expected to climb to over 1.5 billion. [11]

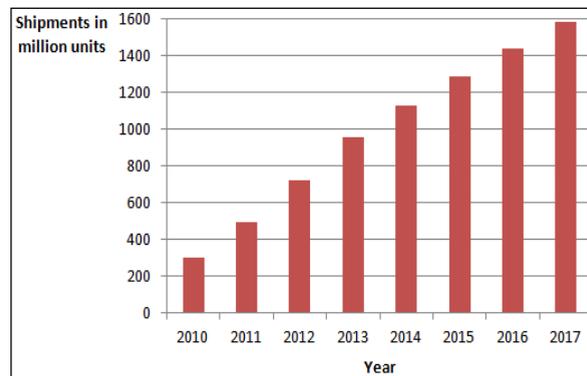


Fig. 3. Global smartphone shipments from 2010 to 2017 (in million units), [11].

Great use of smart terminal devices is also evident in their use in cars, while driving. Daily participation in traffic, no matter if being a driver, a passenger or traveling companion, affects the level of security, both ourselves and other road users. Mobile terminal devices are mostly used for talking, GPS navigation, and for writing or reading messages (SMS, MMS, E-mail).

Many studies and researches have been dedicated to these activities while driving a car. Observed activities, regardless from which point of view were studied, emphasizing the common conclusion - using mobile devices while driving is extremely harmful and significantly increases the possibility of an accident [12].

According to [12], young drivers while driving use mobile terminal devices without the use of accessories (handsfree, headset etc.) for activities that affect their perception of road traffic, as shown in Fig.4..

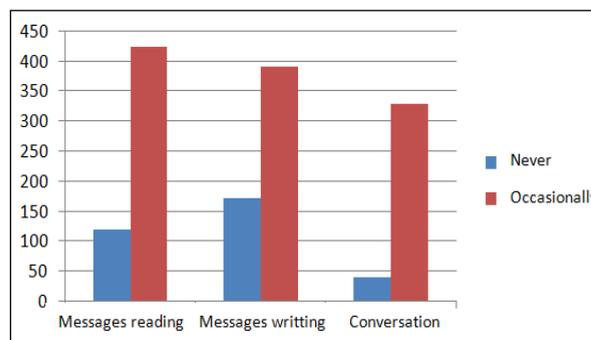


Fig.4.. Graphical presentation of the participation in driving when there was a usage of mobile terminal device, [12].

List of the actions that drivers most commonly use on mobile terminal devices while driving a car are shown in Table 2. The slightest action that is not directed towards driving a car can lead to driver distraction.

Table 2. The possibility of disturbing the driver by using a mobile terminal device when driving, [13].

ACTIVITY	POSSIBLE DRIVER DISTRACTION
<i>Messaging</i>	a. Need to look to decide if a message is important
	b. Unaware or distrust voice control/read out
	c. Check messages at stop lights
	d. Keep phone low, so moving eyes more
<i>Voice calls</i>	a. Need to pick up device to decide if call is important
	b. Bluetooth headsets are owned but not used because they are a pain to keep charged
	c. Phone's speaker max volume is too low
<i>Navigation system</i>	a. Don't use it enough to buy a dock/stand
	b. Need to pick it up to look at it
	c. Put it in dashboard blocking the instruments
<i>Multimedia</i>	a. Listening to streaming music from smartphone apps
	b. Awkward to control app while driving
	c. Speakers on phone are too quiet; want to connect to their car stereo speakers

4. Application possibilities

Availability of computer processing power and network connectivity in cars and mobile terminal devices has led to an explosion of available applications and services for users. One of the possible services when using mobile devices while driving in any case is to use the speech recognition function.

The automotive environment is one of the most challenging environments for speech recognition. In those cases it is important to reduce the visual and physical engagement of the driver due to the potential interference such as the car's occupants and their conversation, background music or similar background noises - the wind, the sound of wiper blades etc. For these and similar reasons, car and equipment manufacturers invest in improving and optimizing applications for speech recognition adapted to the specific environment of the car. Given the above, high quality microphones are installed, as well as technology that allows a reduction of noise. Applications are being improved using an acoustic environment that is specific for the car environment [14].

4.1 Smartphones and voice recognition in automotive environment

Speech is one of the natural modalities of human machine interaction [15]. Voice recognition capabilities are increasingly being developed and used in the automotive industry. The above is not surprising considering that the market competitiveness of modern cars largely depends on their technical features and innovations.

The development of speech recognition in the automotive industry is therefore based on four areas:

- Using a mobile terminal devices without using hands („Call office“)
- Voice directions when using navigation systems („How far is from the accident on the highway?“)
- Car on-board system interaction („Turn-on the radio and select the sports channel“)
- Voice control of car system

Smart terminal devices have become increasingly popular with the development of the hardware segment as well as new features that are the result of using an increasing number of sensors. One significant application of smartphones in any case is the possibility of voice recognition and processing of that information/commands.

Development of applications for modern smart terminal devices has great potential. Due to the specificity of individual mobile operating systems, diverse applications which allow at least some of the functions of speech recognition to a greater or lesser extent are developed. The aim of those solutions is to develop a software solution

that will provide as many functions which can use human speech, as the only (the most important) interface for input and data output.

Some of the currently available application solutions for voice recognition, which are used in smart terminal devices are listed in Table 3.

Table 3. Popular voice recognition applications, [16].

O.NR.	APPLICATION	MANUFACTURER
1	Voice Search	Google Inc
2	Vlingo Virtual Assistant	Vlingo Corporation
3	Iris. (alpha)	Dexetra
4	Skyvi	BlueTornado
5	Speaktoit Assistant	SpeaktoIt
6	AIVC	YourApp24
7	Car Home	Google Inc
8	Dragon Search	Nuance Communications
9	Voice Actions/Jeannie	Pannous
10	Everfriends	i-Free Innovations
11	Evi	True Knowledge Ltd
12	Andy - Siri for Android	74 Technologies
13	Edwin, Speech-to-Speech	Neureau
14	Dragon Go	Nuance Communications
15	Speak4it	AT&T Interactive R&D
16	Voice Assistant - Just use your voice	QuanticApps
17	Pocket Blonde	i-Free Innovations
18	EVA - Virtual Assistant	BulletProof
19	Ziplocal	Phone Directories Company/ZipLocal
20	Cluzee Your Personal Assistant	Tronton LLC

Table 3 shows a list of the most important smartphone applications for the recognition of human speech. The applications are arranged according to the number of downloads from the date of placement of applications on the applications market, ending with the situation in May of 2012. Since the stated list of applications of voice recognition is not final and is increasing almost every day, therefore there is evident potential and the need for these capabilities and systems.

Multinational companies such as Apple and Google have recognized the importance of such applications, and among the first was offering complete solutions. Siri application [17] in this case was a fundamental approach to the stated problem. She has been offered to a large population of users globally on easy way to share it integrated into the smart mobile terminal.

Possibilities of application of speech recognition with mobile terminal devices can be varied, while the most important use cases are listed in Fig.5. Specified use cases can be made in either of two modes, no matter if voice recognition operates directly on the terminal device or it is executed on the server.

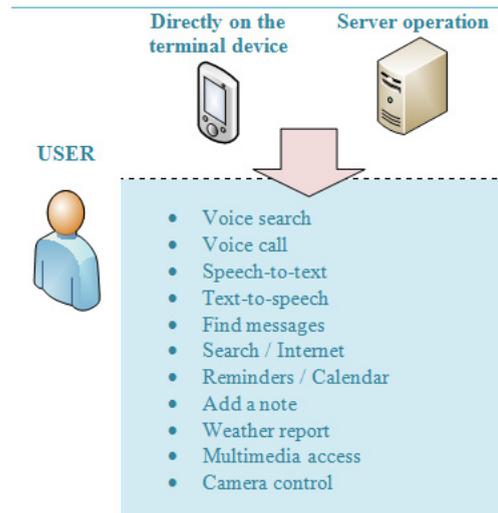


Fig.5. The most important use cases of speech recognition applications for smart terminal devices.

Fig.5. provides an insight into the possible usage scenarios of smart terminal devices when using them in traffic or car, using human speech. Given the above, it is possible to distinguish the following most important characteristics and the possibility of using smart terminal devices:

- Voice search – search information stored on the mobile terminal device
- Voice calling – voice dialing and calling the contact or number
- Search/Internet – finding the required information by using voice commands within a web browser of mobile terminal device
- Voice-to-Text convert – voice commands and messages sent by the speaker (user of mobile terminal device) are converted into text form and can be used to send messages in the form of text messages, e-mail messages etc.
- Voice reproduction – ability for easier reception and understanding of some of the messages (SMS, e-mail, instant messaging) in a way that message received in text format is automatically reproduced by using voice
- Messages finding – potential use of voice when searching information in a way that certain messages (SMS messages, reminders, calendar information, e-mail messages) filtered and used in accordance with the finding keywords
- Calendar/Reminders – add / delete / edit memos and obligations related to the calendar of the mobile terminal device by using voice
- Add/change notes – voice option of adding, deleting and editing notes of the user of mobile terminal device
- Waether report – weather forecast and weather monitoring application managed by voice
- Multimedia access – smart mobile terminal devices are multifunctional devices that provide numerous multimedia data (video files, music files, photo gallery, etc.) which can be easily accessed using the human voice

4.2 Smartphone–car integration

Touchscreen interfaces are widely used in modern technology, from mobile devices to in-car infotainment systems [18]. Speech recognition is also one of the fastest growing areas in the field of science and language engineering. With the development of new communication technologies, voice recognition systems appear as one of the great innovations of the interaction of human and machine retrofitting conversion of Text-to-Speech, Speech-to-Text and similar systems.

Due to the today's means of communication and the need for greater mobility of people it seems practical the use

of mobile devices while driving a car. Since the use of mobile phones while driving is prohibited by law and is necessary to use alternative methods that do not affect the driver and disturb him, and thus not to affect all other participants in traffic.

One such way is to integrate smart terminal devices in the car, or connectivity with built-in device in the car. Table 4. lists the potential advantages and disadvantages of such a system, but in any case it is necessary to keep in mind the relationship between price and quality of integration of smartphones in the car compared to the price and functionality that are obtained by already embedded automotive systems.

Table 4. Advantages and disadvantages of integrating smartphones into cars.

ADVANTAGES	DISADVANTAGES
a. High processing power	a. Legislation
b. App support	b. Small screen dimensions
c. Extensive data plans	
d. Support for fast data transfer (3G, LTE)	
e. Always with user	

The advantage of connecting smart terminal devices with the car on-board unit is seen primarily on the use of mobile devices while driving a car. Connectivity of smart terminal device with car's built-in device allows the driver to use only those functions of a mobile device that will not interfere with the driver. Benefits are also reflected in the number of applications that can be used while driving, and the use of the latest services. Thus provides a complete package of services, and one of the examples of such systems is the MyLink infotainment system. Fig.6. gives a review photos for the interface of MyLink infotainment system developed by General Motors Company (GM), and it supplied within its Chevrolet brand [19]. For this research, an outsourced MyLink system is implemented in the vehicle Chevrolet Cruze [20].



Fig.6. Tested infotainment MyLink system within a GM Chevrolet Cruze vehicle.

Typical architecture of connecting mobile terminal device with built-in car unit is visible in Fig.7.

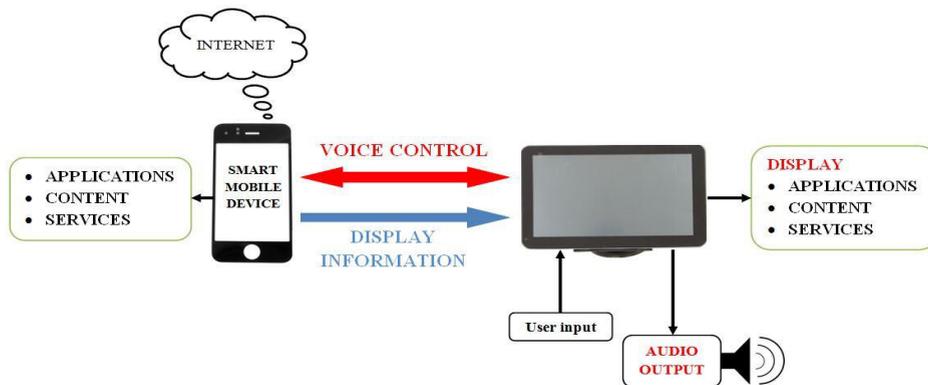


Fig.7. Architecture of integrating of smart mobile terminal device and car's built-in unit.

5. Conclusion

Development of hardware and software equipment has led to an increase in the functionality of smart mobile devices. One of the application possibilities and upgrades of such devices is presented in this paper – the ability of voice recognition. This research has shown how smart terminal devices represent a huge potential in traffic, primarily through the ability to manage, control and communicate exclusively using speech. Technical solutions include application packages which easily create additional value for smart terminal devices through the use of speech recognition. The largest contribution of the use of smart terminal devices is shown by the possibility of their integration with the built-in car system and using functionalities of smart terminal devices that will not interfere with the driver while driving. Thereby is changing the basic work mode of smart terminal devices and uses speech for commands and communication between device and human. This research shows application possibilities and a large number of functionalities that can be performed by using smart terminal devices and speech recognition in the car. The biggest advantage of that sort of communication is the freedom of the driver (driver's hands and eyes), and his greater focus on the driving process. New and better application possibilities of speech recognition on mobile terminal devices are being developed every day. This paper represents an overview of the functionality and usability of voice recognition in the traffic environment. Further research as a continuation of this paper will focus on a comparison of various application packages that enable speech recognition on smart terminal devices, primarily in traffic environment. Concretely, for future research with a comparative analysis of the above mentioned MyLink infotainment system and with it the current advanced systems (like Ford Sync), will be studied and their impact on the safety of traffic participants, compatibility with most existing models of smart mobile terminal devices and their availability and reliability through daily use in a real traffic environment.

References

- [1] Prather, M.: *SpeechRec 101: Understanding Speech Recognition Technology*, Colla Voice Consulting, San Francisco, CA, United States of America, 2012.
- [2] Iizuka, S., Tsujino, K., Oguri, S., Furakawa, H.: *Speech Recognition Technology and Applications for Improving Terminal Functionality and Service Usability*, Technology Reports, Japan, 2011.
- [3] Furui, S. *Automatic Speech Recognition and its Application to Information Extraction*, Tokyo Institute of Technology, Tokyo, Japan, 2000.
- [4] Furui, S.: *Noise-Robust Speech Recognition Technologies in Mobile Environments*, Furui Laboratory and Department of Computer Science, Graduate School of Information Science and Engineering, Tokyo Institute of Technology, Tokyo, Japan, 2007.
- [5] Peraković, D., Husnjak, S., Remenar, V.: *Research Of Security Threats In The Use Of Modern Terminal Devices // Annals of DAAAM for 2012 & Proceedings of the 23rd International DAAAM Symposium / Zadar : DAAAM International Vienna, TU Wien, Karlsplatz 13/311, A-1040, Vienna, Austria, 2012. 545-548.*
- [6] Centar Informacijske sigurnosti: *Programi za zaštitu pametnih telefona* (Information Security Center: *Programs to protect smartphones*), Zagreb, Republic of Croatia, 2011.
- [7] What makes a smartphone smart // http://cellphones.about.com/od/smartphonebasics/a/what_is_smart.htm
- [8] Šarić, S., Peraković, D., Husnjak, S.: *Analysis of the evolution of terminal devices in the use of sms service*, International Conference on Transport Science – ICTS 2012, Portorož, Republic of Slovenia, 2012.
- [9] Husnjak, S.: *Razvoj aplikacije za mobilne uređaje u funkciji podsjetnika temeljenog na lokaciji korisnika*, Sveučilište u Zagrebu, Fakultet prometnih znanosti, Diplomski rad, (*The development of applications for mobile devices in the function of reminders based on user location*, University of Zagreb, Faculty of transport and traffic sciences, Graduate thesis) Zagreb, Republic of Croatia, 2010.
- [10] Sales statistics of smart mobile devices by operating system globally by Gartner Inc. // <http://www.gartner.com/it/page.jsp?id=1622614>
- [11] Global smartphone shipments from 2010 to 2017 (in million units) by Statista Inc. statistics company // <http://www.statista.com/statistics/12865/forecast-for-sales-of-smartphones-worldwide>
- [12] Peraković, D., Jovović, I., Magušić, R. *An Influence Study of the Use of Modern Communication Equipment on Traffic Safety // 15th International Conference on Transport Science ICTS 2012, Portorož, Republic of Slovenia, 2012, PP. 1-10.*
- [13] Lanctot, R. C.: *The Connected Car – It's All About the HMI*, Strategy Analytics – Global Automotive Practice, Rome, Italy, 2012.
- [14] Neustein, A.: *Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics*, Springer Science & Business Media, 2010
- [15] Koolagudi, S. G., Rao, K. S.: *Emotion recognition from speech: a review*, International Journal of Speech Technology 15 (2012), 99-117
- [16] VisionMobile Ltd.: *Beyond Siri: the next frontier in User Interfaces*, London, United Kingdom, 2012.
- [17] Apple Siri application for iOS operating system // <http://www.apple.com/ios/siri/>
- [18] Pitts M. J., Skrypchuk L., Wellings T., Attridge A., Williams M. A., *Evaluating User Response to In-Car Haptic Feedback Touchscreens Using the Lane Change Test*, Advances in Human-Computer Interaction, vol. 2012 (2012), 13 pages.
- [19] General Motors Chevrolet infotainment system for vehicle - MyLink // <http://www.chevrolet.com/mylink-vehicle-technology.html>
- [20] Unikomerc-Automobili d.o.o. (Unikomerc-Cars Ltd.) Trade and Services, an authorized partner for the distribution and sale of General Motors Chevrolet vehicle in Republic of Croatia // <http://www.unikomerc-automobili.hr/>