WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Industrial Applications of Automatic Speech Recognition Systems

Ramakant Tiwari

PG Student, Dr. Shakuntala Misra National Rehabilitation university

ABSTRACT

Contemporary technological developments are constructing vital pathways to the future, bolstered by the proactive and effective utilization of technology to enhance human life. Speech signal processing, encompassing automatic speech recognition, synthetic speech, and natural language processing, is emerging as a pivotal force driving advancements across business, industry, and personal computing convenience. Moreover, it serves as a conduit for delving into the intricate workings of the human brain. The evolution of speech recognition technology over the past five decades has catalyzed a plethora of industrial applications. However, the current landscape merely offers a glimpse of the expansive potential for speech and voice interface technology to supplant keyboards with microphones, thereby facilitating seamless human-machine interactions with increasingly intelligent systems. This paper endeavors to furnish a comprehensive survey of speech recognition technology gleaned from existing literature, amalgamating insights garnered from the analysis of diverse research and development endeavors. Furthermore, it delineates the contemporary applications of speech recognition across various real-world domains, with a focus on medical, industrial robotics, forensic, defense, and aviation sectors.

Keywords: Artificial Intelligence, Automatic Speech Recognition, Human-Machine Interface, Voice User Interface

I. INTRODUCTION

Human perception of the external world is primarily facilitated by the five fundamental senses, with visual and auditory senses holding paramount importance. Speech, historically revered as a cornerstone of human communication predating the advent of writing, occupies a central position among these senses. Given its innate significance in human interaction, speech holds immense potential as a mode of communication with computers. Notably, its capability for remote communication renders it favorable in scenarios where manual engagement is impractical.

The realm of signal processing technology is currently undergoing rapid evolution, spurred by its applications in automation. Recent strides in signal processing, coupled with advancements in communication systems and Artificial Intelligence (AI), find extensive application in industrial automation and biomedical domains. These advancements have significantly propelled the field of speech signal processing, resulting in remarkable achievements. Speech signal processing finds extensive research and application prospects in diverse fields such as security, medicine, and automation, among others. Key domains within speech signal processing encompass speech recognition, speech coding, speech synthesis, and speech enhancement.

The burgeoning wearable electronics industry has introduced innovative applications, such as wearable electronic devices embedded with Automatic Speech Recognition (ASR) systems, GPS, and other communication technologies, catering to the needs of individuals with disabilities. Leveraging speech recognition facilitates tasks ranging from browsing websites to dictation applications and command execution, including e-learning instructions.

Voice User Interfaces (VUIs) leverage speech technology to furnish users with access to information, facilitate transactions, and streamline communication processes. ASR serves as the backbone of VUIs, enabling users to bypass traditional manual input methods like keyboards and mice. This feature proves particularly beneficial for individuals with severe physical or neuromotor disabilities. VUIs engender human-computer dialogues tailored to predefined speech inputs pertinent to specific task domains. ASR, esteemed for its customized approach, finds extensive applications across diverse sectors including security, medicine, and automation. The versatile nature of ASR systems has propelled their integration into mainstream research, owing to their adaptability to various applications. Speech recognition and understanding systems have thus emerged as pivotal components in contemporary research endeavors, owing to their adeptness in addressing specific application requirements.

II. Methodology of ASR Systems

The primary objective of speech recognition is to identify the sequence of speech sounds that best corresponds to the input speech through the utilization of pattern recognition technology. Conversely, speaker recognition focuses on automatically identifying the speaker rather than deciphering the content of the speech. A fundamental ASR system operates in two primary modes: training and recognition, as illustrated in Figure 1. Robust features of speech utterances are extracted during both training and recognition modes, encapsulating information in both the time and frequency domains. During the

training phase, the template generation module leverages potential variations in utterances of the same word to establish templates of feature vectors for different words. Subsequently, in the recognition phase, feature vectors are extracted and compared with stored templates for words to identify the best match for recognizing the word

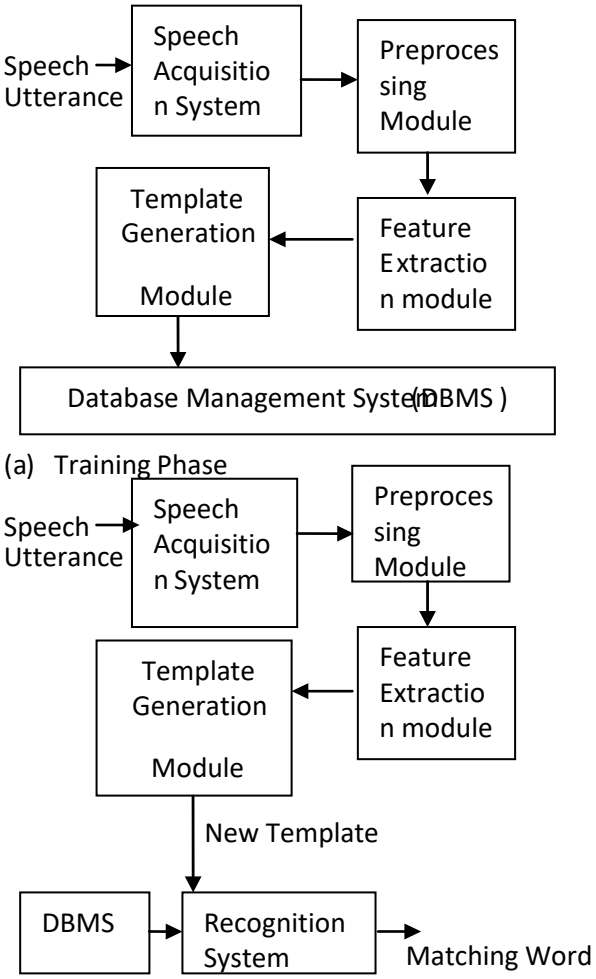

Fig.1 Block Diagram of ASR system (a) Training Phase (b) Recognition Phase

ASR integrated with embedded technology is increasingly gaining popularity, with embedded ASR systems classified into software, hardware, and combined hardware-software co-design approaches.

The notable highlights of ASR systems that render them preferable across various application domains include:

- Accessibility: Users do not require specialized skills, such as typing, to operate speech recognition systems. Speech is a natural skill for most individuals, cultivated from an early age.
- Efficiency: Speech-based communication is significantly faster compared to other modalities such as typing or writing. Users can communicate verbally up to ten times faster than writing manually.
- Multimodality: ASR systems enable users to engage in speech-based communication while concurrently performing other activities involving their hands, legs, eyes, or ears, enhancing multitasking capabilities.
- Affordability: The input methods for automatic speech recognition systems, primarily microphones and telephones, are economically viable and widely accessible.

III. Applications of ASR systems

The journey of automatic speech recognition research commenced in the early fifties, focusing on extracting significant features from acoustic data and classifying and recognizing them using methodologies rooted in pattern recognition. Subsequently, in the seventies, artificial intelligence technologies were harnessed for designing speech understanding systems. A primary motivation for research in this domain lies in the enticing prospect of gaining deeper insights into the complex mechanisms underlying human perception of spoken language, as well as characterizing speech sounds in terms of

physically detectable features in the brain. Such research holds significant utility in psychology and the development of hearing aids for individuals with disabilities. Additionally, research in speech understanding serves as a fertile ground for investigating complex parallel processing systems capable of modeling human perception.

Furthermore, research in automatic speech recognition is fueled by its industrial applications, aimed at simplifying communication between humans and machines. By circumventing intermediary keystrokes and handwritten steps, ASR systems enable multitasking, allowing users to interact with machines while concurrently engaging their hands or eyes in other functions. Moreover, speech input inherently offers higher speed compared to other methods.

thus driving the adoption of automatic speech recognition systems across various industrial sectors, ranging from defense and medical to consumer products. These systems not only prove more cost-effective but also exhibit greater efficiency in real-world applications requiring command and control with minimal human effort. The advancement of computational capabilities of processors has further facilitated the cost-effective deployment of ASR systems, rendering them invaluable tools for industries.

Presently, ASR systems find diverse applications across various domains, including:

- Medical Assistance
- Industrial Robotics
- Forensic and Law Enforcement
- Defense & Aviation
- Telecommunications Industry
- Home Automation and Security Access Control
- Information Technology and Consumer Electronics

The subsequent subsections provide a brief overview of the developments in each of these areas, highlighting the significant contributions and advancements enabled by ASR technology.

A. Medical Assistance

Speech recognition systems in medical applications, as described in the literature, are primarily utilized for reporting, data entry, and in situations where the hands of medical personnel are occupied. Key parameters governing the applicability of speech recognition technology in the medical field include accuracy, efficiency, and reliability. Speech can serve as input for instrument control during surgical procedures, addressing the need to keep hands free and enhancing ease of use.

Recent advancements in ASR systems have focused on large vocabulary, speaker-independent recognition. In the medical realm, these innovations have been integrated into dictation systems for generating reports in areas such as radiology, pathology, and endoscopy. Physicians can dictate directly to the computer, eliminating the need for transcription services and reducing time consumption. Moreover, adoption of this technology enables surgeons to control medical instruments, such as cameras, with speech commands during procedures, requiring only a few commands for effective control. In scenarios where traditional methods of interacting with computers are impractical, such as in microscopy, ASR technology is proving invaluable by reducing response time and skilled manpower requirements for medical and clinical purposes.

ASR techniques have also been employed to aid patients with speech and voice disorders, including dysglossia and dysphonia. Evaluation of voice and speech disorders in conditions like head and neck cancer, which may impair speech and voice, is facilitated by ASR systems. These systems enable fast and inexpensive assessment of speech and voice, augmenting traditional perceptual evaluation methods. ASR, particularly with Hidden Markov Model (HMM) integration, enables precise evaluation of speech and voice, crucial for scientific investigations into therapy outcomes, surgical procedures, and non-surgical therapies. Additionally, ASR methods can predict the emotional state of patients, though increased age has been shown to negatively influence automatic speech recognition.

ASR systems are also evolving to serve as virtual therapists, aiding in the validation of patients' performance in communication disorders such as aphasia. Aphasia, an acquired communication disorder affecting speech and language functionalities, can be addressed through frequent and intense speech therapy sessions. ASR systems enable word verification, as demonstrated by the VITHEA system developed at the Spoken Language Systems Lab in Portugal, facilitating online therapy for aphasic patients with significant success.

B. Industrial Robotics

Advancements in powerful yet affordable microprocessors and enhanced algorithms are propelling the proliferation of commercial applications in various domains, including computer command, consumer electronics, data entry, speech-to-text conversion, telephone systems, and voice verification. Notably, robust speaker-independent recognition systems for command and navigation in personal computers have become readily accessible. Voice input computers find extensive utility in industrial inspection applications, where hands and eyes are occupied with concurrent tasks, enabling direct data input into computers without the need for manual keystrokes or transcription.

Fig.2 in the worldwide annual supply of robots.

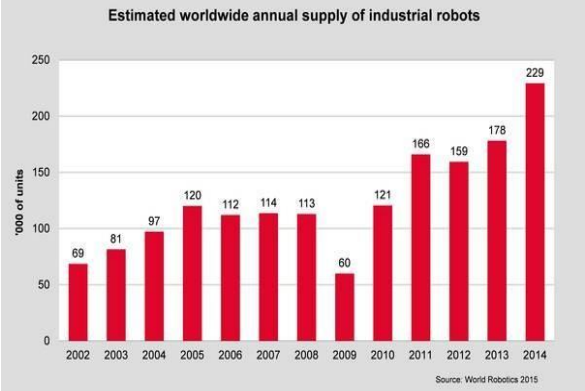

Fig.2 Estimated Worldwide annual shipments of industrial robots [10]

Speech recognition technology is witnessing successful integration into industrial robotics, garnering widespread acceptance across diverse industries. In an era where the demand for superior products at reduced costs drives manufacturing endeavors, autonomous systems requiring minimal operator intervention emerge as crucial enablers of efficiency. This trend is exemplified by the development of Human-Machine Interfaces (HMI) for wheelchair navigation, particularly catering to individuals with severe motor handicaps or muscle degeneration. By enhancing wheelchair navigation autonomy and incorporating intuitive voice or joystick instructions, such as those employed in the Robchair, individuals can achieve greater social independence, navigating dynamically in environments alongside humans.

C. Forensic and Law Enforcement

The voice of an individual can play a pivotal role in forensic applications, with increasing interest observed in the development of automatic forensic voice comparison systems. The Expert Working Group for Forensic Speech and Audio Analysis (FSAAWG) within the European Network of Forensic Science Institutes (ENFSI), notably the Netherlands Forensic Institute (NFI), has spearheaded collective research in this domain. Forensic speaker recognition (FSR) entails the process of determining the suspected speaker, while forensic automatic speaker recognition (FASR) refers to the adaptation of automatic speaker recognition methods for forensic applications. FASR finds utility in judicial and law enforcement contexts, facilitating the identification of speech samples from suspected speakers. This process relies on the similarity between speaker-dependent features extracted from the questioned recording (trace) and those extracted from the speech of the suspect. The diagram below illustrates the process for biometric evidence (E) utilizing suspect R:

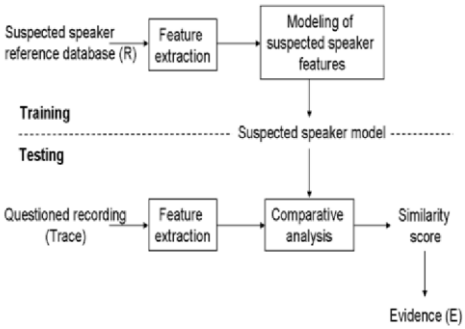

Fig. 3 FASR system determining biometric evidences [11]

In forensic or investigative speaker comparison, speech utterances were previously evaluated by humans; however, the advent of FASR machines has streamlined this process for use in court or investigations.

D. Defence & Aviation

The utilization of speech recognition in aviation is rapidly advancing, particularly in the design of future flight decks. Ongoing enhancements in computer hardware and software are rendering voice control applications on the flight deck more practical, flexible, and reliable. Presently, several military applications leverage Automatic Speech Recognition (ASR) to augment the Human-Machine Interface (HMI), thereby reducing pilot workload in advanced aircraft.

Boeing is integrating ASR into the new Joint Strike Fighter, while the Eurofighter Typhoon is also incorporating this capability into its aircraft. Notably, while there have been notable applications of speech recognition in the cockpit of military aircraft, achieving the requisite reliability and pilot acceptance to enable deployment in the civilian market remains a pivotal milestone.

In the Joint Strike Fighter developed by Boeing and BAe Systems, continuous speech recognition is being integrated into the cockpit to enable selected cockpit controls to be operated solely through voice commands. This speech recognition system aids pilots by minimizing distractions associated with

manual tasks, allowing them to focus on critical aspects of the mission. The supplier of the speech recognition system for this aircraft is ITT Industries' Voxware (formerly VERBEX) voice recognition system.

In the mid-1990s, NASA conducted experiments utilizing speech recognition and voice control in aircraft. The speech recognition system employed for this study was an ITT VRS-1290 speaker-dependent system, featuring a small vocabulary of 54 words. Despite the nascent stage of development, the ASR system achieved an accuracy rate of 95% when tested on 12 pilots under various load and acceleration conditions, garnering general pilot acceptance. However, as the system was speaker-dependent, pilots had to train the speech recognizer to their unique voice patterns prior to use. This training was conducted at ground-based, personal computer support stations to create a voice template for each pilot, which was then transferred to the aircraft via a data loader prior to flight.

E. Telecommunications Industry

The telecommunications landscape, along with speech recognition technology, is undergoing rapid transformation and evolution. Early applications of speech recognition have demonstrated varying degrees of success.

In the early 1990s, speech recognition was introduced into the telecommunications network for two primary reasons: to reduce costs through the automation of attendant functions and to introduce new revenue-generating services that were previously impractical due to attendant-related expenses. Telecommunications services aimed at cost reduction include the automation of operator services, directory assistance, and voice dialing, among others.

Conversely, telecommunications services developed to generate new revenue streams encompass voice banking services, voice prompters, directory assistance, call completion, information services, customer care, computer-telephony integration, and voice dictation. Furthermore, in communication functions, voice control can facilitate tuning radio frequencies by verbally specifying the desired frequency.

The future holds promising prospects for enhanced performance in speech recognition technology, with increased robustness to variations in speakers and background noise. This trajectory is poised to culminate in the provision of reliable, robust voice interfaces for all telecommunications services, making them universally accessible.

F. Home Automation and Security Access Control

The automation industry is witnessing the integration of Automatic Speech Recognition Systems across various sectors, including households, automobiles, and industries. The objective of designing speaker identification-based automation is to facilitate efficient and convenient integration and inter-operation. ASR systems have been implemented in various vehicles to control driving actions. Speech-controlled tasks in vehicles hold the potential to enhance driver safety by reducing the visual (drivers' eyes-off-road time) and manual demands associated with current in-vehicle tasks.

Speech, as a biometric feature, can also be utilized for access control using ASR systems. It offers an unconventional and more secure means of permitting entry without the need to remember passwords, identity numbers, lock combinations, or the use of keys, magnetic cards, or any other easily pilferable device. However, it's important to note that valid access may be denied due to mimicry, accidental changes, emotional or physical stress, etc. Acoustical systems offer significant advantages, particularly in scenarios where the perception of visual information is impaired (e.g., darkness, excessive light, vibration).

Individuals with severe speech and motor impairments, such as those resulting from cerebral palsy, face challenges in moving independently and controlling home electric devices. While computers offer potential assistance to people with disabilities, the standard human-machine interface (e.g., keyboard and mouse) remains inaccessible to this population. Consequently, home automation houses have been developed to apply new technologies in real environments, such as the Welfare Techno Houses of Tamura et al., the Intelligent Sweet Home of Park et al., and the Smart House of West et al. Interfaces based on gestures or voices have gained traction for home automation purposes.Powered wheelchairs represent another avenue for unique mobility for individuals with disabilities and the elderly with motor impairments.

G. Information Technology and Consumer Electronics

The advancements in computational capabilities and the widespread availability of software for speech recognition have significantly benefited consumer products such as laptop computers and personal digital assistants (PDAs). Voice recording and voice annotation functionalities are increasingly being integrated into voice-powered navigation systems.

The diverse array of applications in the fields of information technology and consumer electronics can be classified and discussed as follows:

Conversational Personal Assistance for Web Content Accessibility:

Speech recognition support facilitates internet access for individuals with mobility impairments, visual impairments, and senior citizens. This conversational assistance enables users to control computers via voice commands, access web content, and dictate text using speech recognition software technologies. Various speech recognition software solutions, such as Dragon Naturally Speaking and the built-in Speech Recognition support in Windows operating systems, make web content accessible to users. Commercial software products from companies like Nuance Communications, IBM, Intel, and NVIDIA (GPU) offer flexibility for customized work and efficient operation. Speech-enabled web services, utilizing tags like SALT (Speech Application Language Tag), enable users to access web content through speech commands.

Multimedia Application:

Commercial applications of Automatic Speech Recognition (ASR) are emerging in data centers and portable devices. Many-core computing platforms like the NVIDIA GTX480 present significant opportunities for ASR applications to operate substantially faster when efficiently implemented on GPUs compared to traditional sequential implementations on CPUs. The rapid increase in highly parallel computation from servers to workstations to laptops and portable devices is driving the demand for adapting business and consumer applications to specific usage scenarios, with speech processing technology playing a vital role in this transition.

Mobile Devices:

The proliferation of mobile devices such as smartphones, internet tablets, and cell phones has been exponential in recent years. Sales of smartphones have witnessed significant growth, leading to increased demand for speech recognition technology in mobile applications. Various models of speech recognition with mobile devices, including embedded speech recognition, cloud-based speech recognition, distributed speech recognition, and shared speech recognition based on user adaptation, have been proposed by researchers. Mobile devices require high processing power and/or continuous network connections for optimal performance of speech recognition systems. Challenges faced by mobile devices include limited storage space, variable microphones, low processor power, small cache memory, variable environments due to mobility, and power requirements.

Search Engine: Speech technology has gained prominence due to its ease of access, enabling anytime, anywhere access to information. Speech input can be utilized for searching and accessing information, with companies like Google, Apple, and Microsoft offering speech-based search engines and virtual assistant services such as Google Now, Siri (Speech Interpretation and Recognition Interface), and Cortana. Multimodal platforms like smartphones provide powerful interface capabilities, enhancing user experience and accessibility.

IV. Education and Learning through ASR

Self-learning is undergoing a rapid transformation from traditional book-based learning to electronic learning (e-learning), mobile learning (m-learning), or ubiquitous learning (u-learning). Speech recognition, coupled with mobile devices, has significantly facilitated this shift, thereby playing a crucial role in shaping self-learning experiences. Automatic speech recognition technology holds substantial potential for social impact on self-learning, particularly for users with low literacy levels.

Speech applications hold particular importance for individuals with limited literacy skills. Speech recognition-supported games on mobile devices, for example, can aid in improving literacy levels, particularly in developing countries. Research indicates that practicing recalling and vocalizing words to express intended meanings can enhance word reading skills. Further exploration is necessary to understand how vocabulary training can be effectively delivered through non-visual user interfaces. It may be feasible to provide only phonological, rather than orthographic, hints over such interfaces.

As technology continues to evolve, there are boundless opportunities to leverage automatic speech recognition technology to enhance educational and learning experiences, particularly for individuals with diverse learning needs and preferences. Continued research and innovation in this field hold the potential to revolutionize self-learning practices and foster greater accessibility and inclusivity in education.

Automated speech recognition (ASR) technology represents a significant advancement in touch-free human-computer interaction, addressing the challenge of data entry and retrieval beyond traditional keyboard and video display methods. The demand for alternative interaction methods with computers is on the rise, highlighting the importance of ASR technology in various industrial applications.

This paper has reviewed key industrial application areas of ASR systems. Medical facilities, for instance, are adopting ASR systems for generating medical reports with large vocabularies, particularly in emergency rooms. In the telecommunications sector, applications such as operator services and interactive voice response systems utilize ASR technology, contributing to its growth.

While the availability of new products drives the growth of ASR in industrial applications, factors like cost, integration, and user acceptance also play crucial roles. Vendors are actively collaborating with users to address human factors issues that impact product usability and application. The introduction of inexpensive digital signal processing chips for speech computing applications is poised to significantly impact the adoption of speech interface technologies in price-sensitive mass market applications.

Despite advancements, challenges remain, particularly in the identification of individuals based on their voice. Factors such as recording environment, speaker age, health condition, and emotional state can influence the recognition and identification process. In forensic applications, where accuracy and reliability are paramount, caution is warranted due to the variability of conditions and factors affecting speech samples.

In conclusion, advancements in ASR technology hold promise for various industrial applications, but continued efforts are needed to address challenges and enhance performance, particularly in critical areas like forensic speaker recognition. Striving for the highest technological standards is imperative, especially in applications that directly impact people's lives.

VI. Future directions

The future of automatic speech recognition (ASR) systems lies in aligning their performance, nature, and scope with human auditory processing systems. While ASR technology has advanced significantly in recent years, it has yet to achieve the level of capability and reliability inherent in direct human communication. Bridging this gap requires addressing the limitations of current technology comprehensively.

Research efforts are focused on improving ASR systems by leveraging advancements in signal processing, computational capabilities, and specialized hardware. Developments in areas such as parallel processing, high-accuracy real-time system implementation using very-large-scale integration (VLSI) hardware, and field-programmable gate array (FPGA) techniques aim to enhance the accuracy and user-friendliness of ASR technology.

The proliferation of high-performance mobile devices and cloud computing has propelled mobile computing to the forefront. Speech recognition, as a natural input modality, offers a novel approach to web searching and interaction. With vast data resources, computational power, and ongoing innovation, there is an opportunity to integrate speech recognition seamlessly into mobile devices, making it as ubiquitous and useful as any other input method.

In conclusion, the future of ASR technology holds promise for more natural and intuitive human-computer interaction, driven by advancements in both hardware and software. By addressing current limitations and leveraging emerging technologies, ASR systems have the potential to revolutionize the way we interact with technology in the years to come.

1. Silke Lotterbach, Matthias Peissner, "Voice User Interfaces in Industrial Environments", available at http://subs.emis.de/LNI/Proceedings/Procee dings68/GI-Proceedings.68-121.pdf
2. Furui, Sadaoki. "Digital Speech Processing: Synthesis, and Recognition." 2nd ed., CRC Press, 2018.
3. Octavian Cheng, Waleed Abdulla, Zoran Salcic, "Hardware–Software Codesign of Automatic Speech Recognition System for Embedded Real-Time Applications", IEEE Transactions on Industrial Electronics, Vol. 58, No.3, March 2011
4. Michael A. Grasso, "Automated Speech Recognition in Medical Applications", available at https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUK Ewj64pb0pafLAhVDjo4KHfDuAewQFggf MAA&url=http%3A%2F%2Fciteseerx.ist.p su.edu%2Fviewdoc%2Fdownload%3Fdoi%3D10.1.1.76.9240%26rep%3Drep1%26type%3Dpdf&usg=AFQjCNHH51m5EfB3kIhJp dxwSKhxTFyGsg&sig2=h0MHTadcfE9F2tVWwWvtQ
5. Andreas Maier, Tino Haderlein, "Automatic Speech Recognition Systems for the Evaluation of Voice and Speech Disorders in Head and Neck Cancer", EURASIP Journal on Audio, Speech, and Music Processing, Volume 2010, Article ID 92695
6. Anna Pompili, Alberto Abad, "VITHEA: On-line therapy for aphasic patients exploiting automatic speech recognition", available at www.inescid.pt/pt/indicadores/Ficheiros/8416.pdf
7. Oberteuffer, J. A. "Commercial Applications of Speech Interface Technology: An Industry at the Threshold." *Proceedings of the National Academy of Sciences of the USA*, vol. 92, 1995.
8. Gabriel Pires and Urbano Nunes, "A Wheelchair Steered through Voice Commands and Assisted by a Reactive Fuzzy-Logic Controller", Journal of Intelligent and Robotic Systems, 34, 2002
9. Industrial robot statistics 2015, available at http://www.ifr.org/industrial-robots/statistics
10. Andrzej Drygajlo, "Automatic Speaker Recognition for Forensic Case Assessment and Interpretation", Forensic Speaker Recognition, DOI 10.1007/978-1-46140263-3_2, Springer Science+Business Media, LLC, 2012
11. Perry, William J., and Andrew W. Campbell. "Challenges in Forensic Speaker Recognition." *IEEE Signal Processing Magazine*, vol. 26, no. 4, July 2009.
12. Douglas W. Beeks, "Speech Recognition and Synthesis", Chapter 8, the avionics handbook, 2001 by CRC Press LLC
13. Lawrence R. Rabiner, "Applications Of Speech Recognition In The Area Of Telecommunications", IEEE, 1997, available at: https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUK EwjU9ZG8qafLAhVQco4KHWnRBkYQFg gcMAA&url=http%3A%2F%2Fwww.ece.ucsb.edu%2FFaculty%2FRabiner%2Fece259%2FReprints%2Fapplications%2520of%2520speec h%2520recognition%2520to%2520telecom.pdf&usg=AFQjCNFrIiTQTAtvlemflu 3qEqSTBoD16g&sig2=VIRQgXqoiIfVc2w gnYMufQ
14. Jike Chong, Ekaterina Gonina, Dorothea Kolossa, Steffen Zeiler, Kurt Keutzer, "An Automatic Speech Recognition Application Framework for Highly Parallel Implementations on the GPU", Technical Report No. UCB/EECS-2012-47 at Electrical Engineering and Computer Sciences University of California at Berkeley, April 26, 2012
15. Jonathan Foote, "An overview of audio information retrieval", Multimedia Systems 7, Springer-Verlag, 1999
16. Anuj Kumar, Anuj Tewari, "Rethinking Speech Recognition on Mobile Devices", IUI4DR, California, USA, February 13, 2011
17. Johan Schalkwyk, Doug Beeferman, Fran_coise Beaufays, Bill Byrne, Ciprian Chelba, Mike Cohen, Maryam Garret, Brian Strope, "Google Search by Voice: A case study", available at https://www.google.co.in/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUK Ewij6Necq6fLAhUBcY4KHZwFBgQFgg.