

A Novel Method for Identifying the Keyless Authentication Entry System using Mobile for Auto Mobiles (CAR)

V. Murali Krishna

Assoc.Prof, Eswar College of Engineering,
Narasarao Pet, Guntur Dist, A.P, India.

Y. Mallikarjuna Reddy

Phd, Principal, VVIT, Namburu,
Guntur Dist, A.P, India,

ABSTRACT

Mobile technology plays a vital role in modern era. In the emerging trend of speech technologies, speech/non-speech detection is an unsolved problem, which affects numerous speech related applications. One of the major reasons for these thefts is the unauthenticated access to the car doors and engine. Recently the Remote keyless system emerged for users to operate their cars without keys. For identifying the speech signal, Automatic speech recognition is used. It identifies the signal from the particular person with the help of Voice Active Detector (VAD). It is widely used within the field of speech communication for achieving high speech coding efficiency and low-bit rate transmission. These identify the signal and prioritize the unwanted noise signal using Hidden Markov Model (HMM). In our system, we propose a novel way for keyless entry enabled through mobile phone. The user can dial a specific number connected to the car and can give voice commands. The system built in the car receives the voice commands, goes for voice recognition & authentication. By employing Mel Frequency Cepstral Coefficients (MFCC) processor, the speech will be recognized. After positive authentication, it obeys the command of the car owner. The normal operation based on voice commands are Car Door Lock/Unlock, Engine ON/OFF etc.

Keywords

Mobile Technology, Automatic Speech Recognition (ASR), Hidden Markov Model (HMM), Mel Frequency Cepstral Coefficients (MFCC) processor.

1. INTRODUCTION

Modern automobiles are no longer mere mechanical devices; they are pervasively monitored and controlled by dozens of digital computers coordinated via internal vehicular networks. While this transformation has driven major advancements in efficiency and safety, it has also introduced a range of new potential risks[1]. The term remote keyless system (RKS), or more commonly just as keyless entry, refers to a lock that uses an electronic remote control as a key which is activated by a handheld device or automatically by proximity. The remote keyless system is widely used in automobile industry to lock or unlock the automobile's door, trunk, and start the ignition. A remote keyless system can include both a remote keyless entry system (RKES) and a remote keyless ignition system (RKI). It consists of a handheld radio key fob which is attached on a keychain and a radio transceiver which is located at the automobile [2]. Recently most vehicle services are implemented with sophisticated electronic devices to improve safety and convenience. Vehicle manufactures have introduced a passive keyless entry system (PKES), which is enabling for users to open their vehicle remotely [3]. Active a key initiates the authentication process and transmits a message. For instance, in a Remote Keyless Entry system for a car, the key (remote

control) becomes active when the owner pushes a button, while the system in the car wakes up frequently (every 20ms) to receive commands from the key [17]

Biometric technology is fast gaining popularity as means of security measures to reduce cases of fraud and theft due to its use of physical characteristics and traits for the identification of individuals [4]. Biometrics is the science and technology of measuring and analyzing biological data. In information technology, biometrics refers to technologies that measure and analyze human body characteristics, such as DNA, fingerprints, eye retinas and irises, voice patterns, facial patterns and hand measurements, for authentication purposes. Currently Voice biometric systems are mostly used for telephony-based applications. Speech recognition has not only concerned the meaning of spoken words, but also such non-verbal information as language, gender, age and emotion [6]. Speech biometrics has been an active area of research for several years now but continues to hold interest because of the practical utility of being able to identify a speaker based on voice in several applications. Many voice based automated self help systems need the speaker to identify themselves before being serviced, for example banks and financial institutions [10].

Mobile phones are playing an important role in changing the socio-economic landscapes of emerging economies like India. A proper voice-based user authentication will help in many new mobile based applications including mobile-commerce and banking [7]. A voice-based access control system is a performance biometric which offers an ability to provide positive verification of identity from an individual's voice characteristics to access secure locations (e.g. office, laboratory, home etc..) [9]. Voice Active Detector is a technique used in speech processing in which the presence or absence of human speech is detected. It can facilitate speech processing and can also be used to deactivate some processes during non-speech section of a speech session [5]. Voice verification is used for government, healthcare, call centers, electronic commerce, financial services, customer authentication for service calls, house arrest, and probation-related authentication. A powerful voice/talk recognition system can provide many benefits for paralysis patients. Talk recognition is suitable for wheelchairs, home appliances, systems with phones and emergency systems operating by giving voice command [8]

Wireless Home Automation System (WHAS) is to provide those with special needs with a system that can respond to voice commands and control the on/off status of electrical devices, such as lamps, fans, television etc, in the home. The system should be reasonably cheap, easy to configure, and easy to run [12]. Automatic speech recognition (ASR) has been an active research area for several decades, but in spite of the enormous efforts, the performance of current ASR systems is far from the performance achieved by humans: error rates are often one

order of magnitude apart. Most state-of-the-art ASR systems make use of the acoustic signal only and ignore visual speech cues [11]. A sound source localization technique which tracks humans using a laser range sensor and it extends the application to speech recognition and demonstrate logging of recognized commands from multiple speakers in multiple locations and report the association about the tracking and sound sources in a static situation [16].

In real-world operation, the introduction of public keys is done in the field where user device is not necessarily under the control of OEM and a live online connection to OEM may not be possible. PKC is computationally very expensive, especially when aiming for high security levels [19]. More fine-grained access control commonly involves the use of access control matrices which specify permissions indexed by authenticated subject and object (resource), or equivalently, per-object access control lists specifying privileges associated with authenticated users [20]. The voice recognition system that is relevant to text needs users to pronounce according to the stated contents, and then everybody's speech model is built up accurately. Because identification also need users to pronounce according to the stated contents, the effect is very good [13]

When a voice signal is detected, the speech recognizer is triggered to process the signal. The detection of voice signals can be performed by traditional voice activity detection (VAD) method. Although it is effective for isolating voice signals from noises, traditional VAD methods cannot differentiate the voice signals of a legitimate user from others [14]. It is to improve the authentication accuracy in noisy environments by using bone-conducted (BC) speech instead of air-conducted (AC) speech. BC speech is inherently noise-free when a fine BC microphone is used to obtain the speaker's speech. [15]. The advantages of such automatic services are clear -much higher capacity compared to human-operated services with hundreds or thousands of phone calls being processed simultaneously. In fact, the focus of speaker recognition research over the years has been evaluated and few techniques are processed for identifying the keyless cars as well as it have been used for numerous applications [18]

2. RELATED WORKS

Numerous researches have been made on the area of voice recognition used in mobile for analyzing the keyless authentication using filtering technique. In this section, a brief review of some imperative contributions from the existing literature is presented.

Voice Activity Detectors (VADs) are extensively used in a number of applications including speech coding, speech enhancement, and speech recognition. An accurate VAD extracts a set of discriminative speech features from the noisy speech and makes the decision in terms of well-defined rule. J. Ramirez et al [26] have discussed about the major challenges in robust speech detection. They have utilized three efficient VAD techniques and those techniques have achieved high speech/non-speech discrimination accuracy as well as improved the performance of speech recognition systems working in noisy environments. The assessment of those techniques had shown the experiments most often conducted to compare VADs: (1) speech/non-speech discrimination analysis, (2) the receiver operating characteristic curves, and (3) speech recognition system tests.

Tanzeem Choudhury et al [25] have discussed that activity aware systems have enthused innovative user interfaces and new applications in smart environments, surveillance, emergency response, and military missions. Systems that

identify human activities from body wore sensors can further assist the healthcare applications namely, fitness monitoring, eldercare support, long term precautionary, chronic care, and cognitive assistance. Wearable systems have the benefit of being with the user continually. For instance, a fitness application could make use of real-time activity information to support users to perform opportunistic activities. In addition, such activity recognition systems were widely accepted by the general public because they were normally easy to turn-off or eliminate.

Al-Hadda [21] has devised an algorithm for noise removal by means of recursive least square (RLS) and pattern recognition by means of fusion technique of Dynamic Time Warping (DTW) and Hidden Markov Model (HMM). He argued that speech signals were often disturbed by background noise and the changes in signal characteristics were fast. Those issues were mainly significant for robust speech recognition. Robustness was a major issue in speech recognition. The proposed algorithm was tested on speech samples from the Malay dataset. The fusion technique were employed to combine the pattern recognition outputs of DTW and HMM. Moreover, refinement normalization was applied by using weight mean vector to achieve better performance. 94% accuracy was obtained on pattern recognition using Hybrid HMM and DTW compared to 80.5% using DTW and 90.7% using HMM, individually. The accuracy of the proposed algorithm was further improved to 98% by employing the RLS adaptive noise removal technique.

Keith Vertanen and Per Ola Kristensson [23] have investigated how to recognize and correct the voice web search queries. They have described their dataset of web search queries and revealed how it was employed to enhance the recognition accuracy. They have also proved that using a search specific vocabulary with automatically generated pronunciations was better than using a vocabulary limited to a fixed pronunciation dictionary. A formative user study was carried out to examine the recognition and correction aspects of voice search in a mobile context. During the user study, they identified that despite a word error rate of 48%, users were capable of speaking and correcting the search queries in about 18 sec. Users did such process while walking around using a mobile touch-screen device.

Matthew M. Loper et al [24] have discussed that mobile peer-to-peer interaction between humans and robots can be made via structured light-based depth sensing with standard perception algorithms. They supposed that the use of recent promising devices for depth-based imaging can allow robot perception of non-verbal cues in human movement despite lighting and minor terrain variations. Moreover, an integrated robotic system was proposed capable of person following and responding to verbal and non-verbal commands under varying lighting conditions and uneven field. The capability of their system for peer-to-peer HRI was shown through two experiments in indoor and outdoor environments.

Oytun Turk et al [22] have discussed that voice conversion allows creation of a desired speaker's voice from audio recordings of another speaker. They have concentrated on a music application and explained the initial steps towards creating voices of music celebrities by means of traditional voice conversion methods. Particularly, rap singing conversions from English to Spanish were carried out via parallel training material in English. Voice conversion based on weighted code book mapping with two diverse alignment techniques and temporal smoothing of the transformation filter was performed. The first aligner has utilized a HMM trained for each source

recording to force-align the corresponding target recording. The second aligner has utilized speaker-independent HMMs trained from a large number of speakers. In addition, smoothing was done to reduce discontinuities and to enhance performance. The results of subjective evaluations have shown that the performance of both aligners was equivalently well. The smoothing approach has improved both similarity to target singer and quality significantly despite the alignment techniques.

Josef Rajnona and Petr Pollak [27] have investigated Automatic Speech Recognition (ASR) systems for use within the noisy environment and suggested best configuration under different noisy conditions. The performance of standard parameterization methods was evaluated in terms of robustness against background noise. It was done for Mel-frequency cepstral coefficients (MFCC), Perceptual linear predictive (PLP) coefficients, and their modified forms combined the main blocks of PLP and MFCC. Then, the function of modified methods containing frequency domain noise suppression and VAD was evaluated. The aforesaid methods were tested by means of signals in real noisy environment within Czech digit recognition task and AURORA databases. Eventually, they have studied the role of certain VAD selective training and MLLR adaptation of acoustic models for different signal features.

3. SPEECH RECOGNITION

Speech Recognition technology can translate spoken words into text. Some SR systems use "training" where an individual speaker reads sections of text into the SR system. These systems analyze the person's specific voice and use it to fine tune the recognition of that person's speech, resulting in more accurate transcription. Recent progress in Automatic Speech Recognition (ASR) technology has enabled the development and deployment of more sophisticated and more accurate speech recognition applications. This progress, combined with an explosion in the capabilities and use of wireless and mobile communication and computing terminal devices, makes it feasible for ASR to become a common feature and service for current and future portable terminal devices and in mobile or wireless networks. This mobile ASR capability can be applied both as a user interface to the terminal device as well as a data Input/output modality between the user and the remote application.

Automatic speech recognition is among other things useful in situations where an operator is inputting data to a computer in parallel with using his hands for other tasks, e.g., when sorting goods or inspecting products. Another application is the voice control of a word processor, while inputting the text manually.

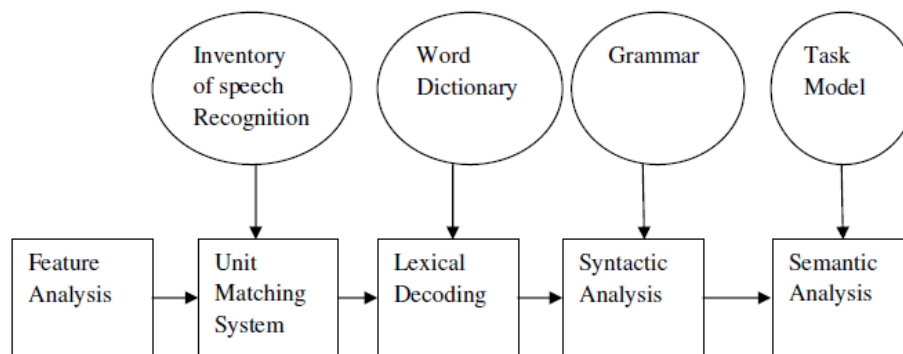


Figure 1. Block Diagram of speech recognition

3.1. Mobile Network Speech Recognition

The recognizer is implemented in a location remote to the user so the speech signal has to be transmitted from the user's terminal to the recognition server through a wireless link. The most evident contrast of this modality with respect to fixed network recognition (e.g., telephone-based recognition) is the wireless channel transmission medium. The implications of this basic difference are twofold: the need for reduction of the signal's bit rate through source coding techniques and the effects of the wireless transmission channel on the reconstructed signal (i.e., transmission errors, data dropouts, interference noise etc.). However, having a recognizer residing in a central server enables larger and more powerful computers to perform recognition, permitting more sophisticated and elaborate ASR applications (e.g., dialog-based systems which currently incorporate parsers, natural language modules, speech synthesis and database queries) than those possible on terminal devices.

Speech biometrics is a mechanism of verifying the identity of the speaker based on their voice characteristics. All speech biometric systems are learning systems, namely they have to be

trained for a particular speaker in order for them to be able to identify or recognize the speaker at a later date. Essentially there is a training phase which requires the speaker to register with the system by providing his voice samples and there is a test phase where the system identifies the speaker based on the voice samples.

3.2. A Mathematical Formulation of the Speech Recognition Problem

The problem of automatic speech recognition consists of the task of transcribing the associated text of an utterance. This task is performed given an observation of the features of a realization of the utterance. The most successful and widely used approach to speech recognition is of a statistical nature and is based on Hidden Markov Models.

Let a be the observed set of feature vectors computed by the front-end processor of the recognizer from the speech signal. Let this sequence of feature vectors be denoted by $a_1, a_2, a_3, \dots, a_m$, which, without loss of generality can be

thought of as symbols taken from a possibly very large alphabet Ψ

$$a_1, a_2, a_3, \dots, a_m \quad a_i \in \Psi \quad (1)$$

Let the utterance W produced by a speaker be a sequence of words belonging to a fixed given vocabulary \mathcal{V} :

$$w_1, w_2, w_3, \dots, w_m \quad w_i \in \mathcal{V} \quad (2)$$

If $p(w|A)$ denotes the probability that the string of words W was produced given the acoustic observation A , then a speech recognition system will have the objective of finding the most likely word string \hat{w} given the observed acoustic evidence A

$$\hat{w} = \arg \max_w p(w|A) \quad (3)$$

This sort of approach (called direct modeling), implies the need to have an inventory of all the possible acoustic realizations A of every single utterance, which is infeasible. To make the

Problem tractable we base our recognition in terms of, $p(A|w)$ which can be modeled more easily than the case above following a source-channel model and phonetic modeling of the string sequences A

$$\hat{w} = \arg \max_w P(w) \frac{P(w)P(A|w)}{P(A)} \quad (4)$$

The term $P(A)$ is the average probability of observing A and is not a function of the uttered string, so it can be eliminated from Eq (4). The problem now reduces to the maximization of the term $P(w)P(A|w)$. The term $P(w)$ represents the probability of the string W , regardless of acoustic observations. This term can possibly be computed considering the semantic, syntactic, grammatic and pragmatic likelihood of the candidate string W . The term $P(A|W)$ is evaluated using a set of acoustic models that describes the likelihood of the observed feature sequence for every candidate hypothesis. Modern speech recognizers perform this acoustic modeling making use of a phonetic representation of the words in the system's dictionary, and a collection of Hidden Markov Models associated to each phonetic unit of the task's language.

3.2.1. Mobile Phones

As A Strategy: Mobile phones are playing an important role in changing the socio-economic landscapes of emerging economies like India. A proper voice-based user authentication will help in many new mobile based applications including mobile-commerce and banking. Mobile technologies and social media are invaluable tools for businesses of all sizes. The AT&T study shows that these tools are great for saving resources, reaching out to customers, and empowering your workforce. These mobile technologies mostly used to safeguard the belongings of the public. It is widely used in various applications like cars, banks, atm, hospitals, etc. These strategies are well used by using a server is connected to one terminal, which receives the signal when a mobile makes a call. These phenomenal methods were used by various techniques.

3.3. Voice Activity Detection (VAD) Methods

When a voice signal is detected, the speech recognizer is triggered to process the signal. The detection of voice signals can be performed by traditional voice activity detection (VAD) methods. This approach, however, raises several concerns. These are used in real world environments where much of the speech around the devices is not directed to them. Although it is effective for isolating voice signals from noises, traditional VAD methods cannot differentiate the voice signals of a legitimate user from others. This causes the speech recognizer to be frequently activated to perform unnecessary tasks. It is desirable to prevent the frequent activation of the system on small mobile devices with limited power supply. The system should be activated only when the voice signal from a legitimate user is detected. Furthermore, due to their computational cost, the full-fledged speech recognizers are unsuited to devices with limited computing power. In summary, to work effectively on a device with limited power supply and computing power, the voice trigger system must have a small computational cost and be able to detect only the keywords uttered by legitimate users without a fully featured speech recognizer utterance of about one second, without a full-fledged speech recognizer. When an unregistered keyword is uttered or the voice of an impostor is detected, the voice trigger system should reject the signal. Thus, a voice trigger system may consist of two components, i.e., keyword recognition and speaker recognition

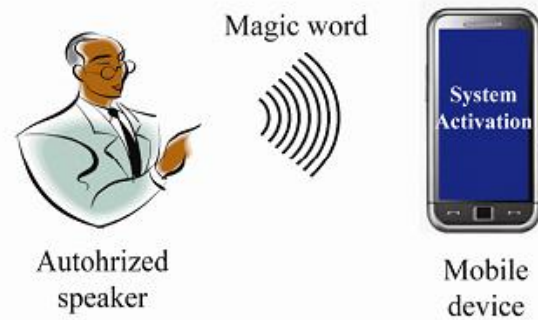


Figure 2. A voice trigger system for mobile devices.

4. MEL FREQUENCY CEPSTRAL COEFFICIENTS (MFCC) PROCESSOR

A block diagram of the structure of an MFCC processor is given in Figure 2. The speech input is recorded at a sampling rate of 22050Hz. This sampling frequency is chosen to minimize the effects of aliasing in the analog-to-digital conversion process. Figure 2 shows the block diagram of an MFCC processor.

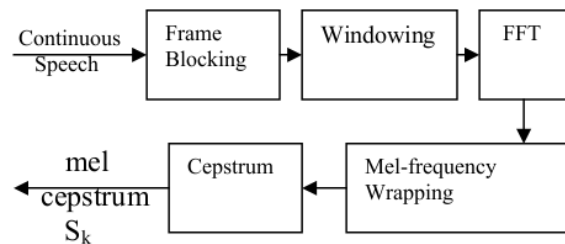


Figure 3. Block diagram of the MFCC processor

4.1. Mel-Frequency Wrapping

The speech signal consists of tones with different frequencies. For each tone with an actual frequency, f , measured in Hz, a subjective pitch is measured on the 'Mel' scale. The mel-frequency scale is linear frequency spacing below 1000Hz and a logarithmic spacing above 1000Hz. As a reference point, the pitch of a 1 kHz tone, 40dB above the perceptual hearing threshold, is defined as 1000 Mels. Therefore we can use the following formula to compute the Mels for a given frequency f in Hz

$$Mel(f) = 2595 \log_{10}(1 + f/700) \quad (5)$$

The log Mel spectrum has to be converted back to time. The result is called the Mel frequency cepstrum coefficients (MFCCs). The cepstral representation of the speech spectrum provides a good representation of the local spectral properties of the signal for the given frame analysis. Because the Mel spectrum coefficients are real numbers (and so are their logarithms), they may be converted to the time domain using the Discrete Cosine Transform (DCT). The MFCCs may be calculated using this equation

$$C_n^{\%} = \sum_{k=1}^k (\log S_k^{\%}) \left[n(k - \frac{1}{2}) \frac{\pi}{k} \right], \quad (6)$$

$$N=1, 2, \dots, k$$

The number of Mel cepstrum coefficient is typically chosen as some value. The first component is chosen as a value. The first component is \log excluded from the DCT since it represents the mean value of the input signal which carries little speaker specific information. By applying the procedure described above, for each speech frame of about 30 ms with overlap, a set of Mel-frequency cepstrum coefficients is computed. This set of coefficients is called an acoustic vector. These acoustic vectors can be used to represent and recognize the voice characteristic of the speaker. Therefore each input utterance is transformed into a sequence of acoustic vectors.

5. HARDWARE IMPLEMENTATION

In the hardware design, there are several factors to take into consideration such as lifespan of the key fob's battery, key fob size, system cost, customer-acceptable protocol execution time, radio coverage range, etc. We adopt a 32 bit low-cost ARM embedded Processor (LPC 2128) for both the car transceiver and the key fob. The ARM processor which is shown in Figure. 4 operates on LPC 2128 boot loader.

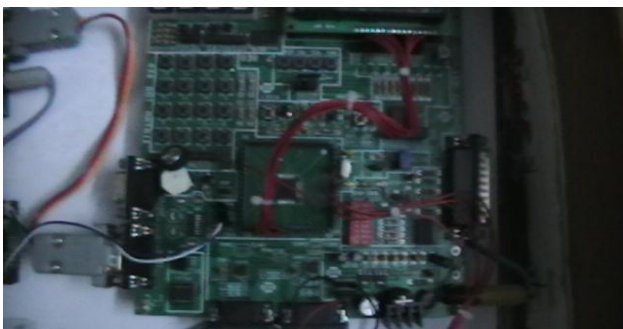


Figure 4. ARM Processor used in Proposed methodology

LPC2128 boot loader is 12K long and starts at address 0x7FFF D000, although it is remapped. So after booting I halted the device, looked at an ARM code disassembly at this address,

assigned the PC to this address and started stepping. There is no reason to assume this is not the remapped start address. It will operate with 4 relays. Here we use IR software and Keil C software for implementation. Figure. 5 describes the GSM used in the processor



Figure 5. GSM used in Processor

UNL 2003

The ULN2003 is a monolithic high voltage and high current Darlington transistor arrays. It consists of seven NPN darlington ton pairs that features high-voltage outputs with common-cathode clamp diode for switching inductive loads. The collector-current rating of a single darlington ton pair is 500mA. The darling ton pairs may be paralleled for higher current capability. Applications include relay drivers, hammer drivers, lamp drivers, display drivers (LED gas discharge), line drivers, and logic buffers. The ULN2003 has a 2.7kΩ series base resistor for each darling ton pair for operation directly with TTL or 5V CMOS devices. The entire hardware design is shown in Figure. 6.



Figure 6. Entire Hardware design

In our system, a novel way for keyless entry enabled through mobile phone was designed based on these specific hardware and software specifications. When the user could dial a specific number connected to the car and it gave voice commands. The system built in the car receives the voice commands, goes for voice recognition & authentication. After positive authentication, it obeys the command of the car owner. The normal operation based on voice commands are Car Door Lock/Unlock, Engine ON/OFF etc.

Performance Measure

The voice data were partitioned into a sequence of 25 millisecond frames with a 10 millisecond advance. A Hamming window was applied to each frame. 12 dimensional mel-frequency cepstral coefficients (MFCCs), log energy, and their first and second order time derivatives were used as the feature vector. The equal error rate (EER), which is the error rate measured when the false alarm rate and the false rejection rate are the same, was used as the performance measure. Here the performance of the proposed system is evaluated by using the following measures

Log likelihood

We first calculate the log likelihood, S_m to verify the registered keyword and user:

$$S_m = \log P(x_{1:T} | r; m) \quad (7)$$

where $x_{1:T}$ and m represents the input voice data.

Equal Error Rate (EER)

The error rate of a verification system when the operating threshold for the accept/reject decision is adjusted such that the probability of false acceptance and that of false rejection become equal. The following graph shows the comparison of proposed Mobile Based Keyless Authentication and Entry System (MBKAES) and existing systems [28] with respect to EER. Here the EER is less for the proposed system when compared to others.

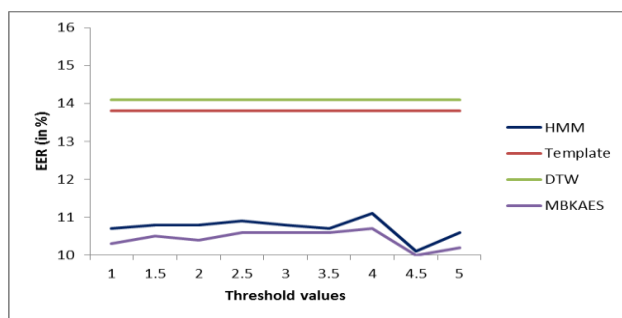


Figure 7. EER comparison

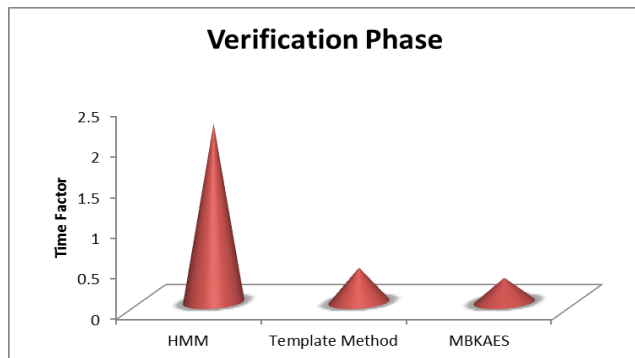


Figure 8. comparison of execution time

The verification decision was made by the system. It takes roughly real-time to register a keyword and a user, and much less than real-time for recognition, using the template based method. In contrast, it takes almost 4.5 times real-time for registration, and two times real-time for recognition, in the HMM based method. The template based method is much faster than the HMM based method [28]. The proposed method is better than the other two methods mentioned in the graph Figure. 8

Table 1. Comparison On Execution Time

Verification Phase	
Methods	Execution Time
HMM	2.2
Template	0.425
Proposed MBKAES System	0.3

Mention the name of the Method :

Table II shows the results of the EER of all the three approaches. The two step decision procedure decreased the EER by 2.0% in the HMM based method and 3.6% in the template based method compared to the one step decision scheme. In addition, the HMM based method decreased the recognition error rate by 27.8% compared to the template based method (from 13.2% to 9.5%). Thus, the GMM based method can be used as a more reliable voice trigger for home appliances and digital devices, while the template based method has a fast voice trigger at the expense of accuracy. Our proposed method is more reliable than other 2 methods which is described in the following table.

Table 2: performance measure

Decision Methods	One step	Two Step	Improvement
HMM	13.7	13.2	3.6
Template	9.7	9.5	2.0
Proposed MBKAES System	8.1	7.9	2.0

6. CONCLUSIONS

An important problem to determine the presence of speech periods in a given signal is examined. Voice Active Detector is used to minimize the data of the corrupted signals. The study reveals the combination of Mel frequency and Hidden Markov Model (HMM) gives the better performance. HMM is used to improve the efficiency and precision of the segmentation to deal with crosstalk, laughter, and uncharacteristic speech sounds. A more effective normalization algorithm can be adopted on extracted parametric representations of the acoustic signal, which would improve the identification rate further. The MFCC technique has been applied for speaker identification. Mel scale is also less vulnerable to the changes of speaker's vocal cord in course of time. Finally MFCC is used to implement a robust parametric representation for speaker identification.

7. REFERENCES

- [1] Karl Koscher, Alexei Czeskis, Franziska Roesner, Shwetak Patel, and Tadayoshi Kohno, " Experimental Security Analysis of a Modern Automobile," IEEE Symposium on Security and Privacy (SP), pp. 447- 462, 2010.
- [2] Xiao Ni, Weiren Shi and Victor Foo Siang Fook." AES Security Protocol Implementation for Automobile Remote Keyless System," IEEE Conference on Vehicular Technology, pp. 2526 - 2529, Apr 2007.
- [3] Hisashi Oguma, Naoki Nobata, Kazunari Nawa, Tsutomu Mizota and Mitsuru Shinagawa," Passive Keyless Entry System for Long Term Operation, IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, pp. 1-3, Jun 2011.
- [4] Rozeha A. Rashid, Nur Hija Mahalin, Mohd Adib Sarijari and Ahmad Aizuddin Abdul Aziz," Security System Using Biometric Technology: Design and Implementation of Voice Recognition System (VRS)," In Proc. of the International Conference on Computer and Communication Engineering, Jul 2008

- [5] Dr.V.Radha, Ms. Vimala.C and Ms. M.Krishnaveni," An Efficient Voice Activity Detection Method for Automatic Tamil Speech Recognition System using VADSOHN Algorithm," IEEE International Conference on Communication Control and Computing Technologie (ICCCCT), pp. 597- 602, Dec 2010.
- [6] Wei Li, Dong-Ju Kim, Chul-Hwan Kim and Kwang-Seok Hong," Voice-Based Recognition System for Non-Semantics Information by Language and Gender," Third International Symposium on Electronic Commerce and Security, pp. 84-88, 2010.
- [7] Amitava Das, Ohil K. Manyam , Makarand Tapaswi and Veeresh Taranalli," Multilingual spoken-password based user authentication in emerging economies using cellular phone networks," IEEE Spoken Language Technology Workshop, pp. 4-8, 2008.
- [8] Nihat Ozturk and Ulvi Unozkan," Microprocessor Based Voice Recognition System Realization," International Conference on Application of Information and Communication Technologies (AICT), pp.1-3, Aug 2010.
- [9] Syazilawati Mohamed and Wahyudi Martono," Design of Fusion Classifiers for Voice-Based Access Control System of Building Security," World Congress on Computer Science and Information Engineering, 2009.
- [10] Sunil Kumar Kopparapu," A Robust Speech Biometric System for Vehicle Access," International conference on Vehicular Electronics and Safety (ICVES), pp. 174 - 177, Nov 2009.
- [11] Stéphane Dupont and Juergen Luettin," Audio-Visual Speech Modeling for Continuous Speech Recognition," IEEE Transactions on multimedia, Vol. 2, No. 3, pp. 141-151, Sep 2000.
- [12] Humaid Al Shueili, Gourab Sen Gupta and Subhas Mukhopadhyay," Voice Recognition Based Wireless Home Automation System," International Conference on Mechatronics, May 2011.
- [13] .Bo Cui and Tongze Xue," Design and Realization of an Intelligent Access Control System Based on Voice Recognition," International Colloquium on Computing, Communication, Control, and Management, pp. 229-232, 2009.
- [14] Hyeopwoo Lee, Sukmoon Chang, Dongsuk Yook and Yongserk Kim," A Voice Trigger System using Keyword and Speaker Recognition for Mobile Devices," IEEE Transactions on Consumer Electronics, Vol. 55, No. 14, pp. 2377 - 2384, Nov 2009.
- [15] Naoto Yamasaki and Tetsuya Shimamura," Accuracy Improvement of Speaker Authentication in Noisy Environments Using Bone-Conducted Speech," IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 197 - 200 2010.
- [16] Simon Thompson, Yoko Sasak and Satoshi Kagami," Associating Spoken Commands with Multiple Human Users in a Dynamic Environment," International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pp. 207- 212, Dec 2008.
- [17] Feng Zhu, Matt W. Mutka and Lionel M. Ni," Private Entity Authentication for Pervasive Computing Environments," International Journal of Network Security, Vol.14, No.2, pp. 86-100, Mar. 2012.
- [18] Tomi Kinnunen and Haizhou Li," An Overview of Text-Independent Speaker Recognition: from Features to Super vectors," Speech Communication, Vol. 52, No. 1, pp. 12-40. Jan 2010.
- [19] Alexander Klimm, Benjamin Glas, Matthias Wachs, Sebastian Vogel, Klaus D. Muller Glaser and Jurgen Becker," A Security Scheme for Dependable Key Insertion in Mobile Embedded Devices," International Journal of Reconfigurable Computing, Vol. 2011, pp. 1-19, 2011.
- [20] P.C. van Oorschot," Message Authentication by Integrity with Public Corroboration," In Proc. of the 2005 workshop on New security paradigms, 2005
- [21] S.A.R. Al-Haddad, S.A. Samad, A. Hussain, K.A. Ishak and A.O.A. Noor," Robust Speech Recognition Using Fusion Techniques and Adaptive Filtering," American Journal of Applied Sciences, Vol. 6, No.2, pp. 290-295, 2009.
- [22] Oytun Türk, Osman Büyük, Ali Haznedaroglu and Levent M. Arslan," Application of Voice Conversion for Cross-Language Rap Singing Transformation," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3597 - 3600, Apr 2009.
- [23] Keith Vertanen, Per Ola Kristensson," Recognition and Correction of Voice Web Search Queries," In Proc. of the International Conference on Spoken Language Processing, 2009.
- [24] Matthew Maverick Loper, Nathan P. Koenig, Sonia Chernova, Chris V. Jones, Odest Chadwicke Jenkins. Mobile human-robot teaming with environmental tolerance. In Proc. Of HRI, pp.157-164, 2009.
- [25] Tanzeem Choudhury, Gaetano Borriello and Dirk Haehnel," The Mobile Sensing Platform: An Embedded Activities Recognition System," PERVASIVE computing, pp. 32-41, Jun 2008.
- [26] J. Ramírez, J. M. Górriz and J. C. Segura," Voice Activity Detection. Fundamentals and Speech Recognition System Robustness, System, pp. 1-22, 2007.
- [27] Josef Rajnoha, Petr Pollák, ASR systems in Noisy Environment: Analysis and Solutions for Increasing Noise Robustness," Radio Engineering, Vol. 20, NO. 1, Apr 2011.
- [28] Hyeopwoo Lee; Sukmoon Chang; Dongsuk Yook; Yongserk Kim; "A voice trigger system using keyword and speaker recognition for mobile devices" IEEE Transactions on Consumer Electronics, 2009 IEEE.