Designing Smart Cities: Security Issues

Young Im Cho

HAL Id: hal-01551737
https://inria.hal.science/hal-01551737

Submitted on 30 Jun 2017

Designing Smart Cities : Security Issues

Young Im Cho

Dept. of Computer Science, The University of Suwon, Korea

ycho@suwon.ac.kr

Abstract. Smart city means an intelligent city called post-ubiquitous city. However, to be a smart city, there are some issues to consider carefully. The first is the security issues having platform and intelligent surveillance function and analysis function as well as inference. The second is the governance issues between government and cities which are developing. And the third is the service issues to be realized in the cities. In this paper, we are mainly concentrated on the first security issue. Here we propose a speech recognition technology in emenrgency situation for secure cities. For the emergency detection in general CCTVs(closed circuit television) environment of our daily life, the monitoring by only images through CCTVs information occurs some problems especially in emergency state. Therefore for detecting emergency state dynamically through CCTVs as well as resolving some problems, we propose a detection and recognition method for emergency and non-emergency speech by Gaussian Mixture Models(GMM). The proposed method determines whether input speech is emergency or non-emergency speech by global GMM firstly. If this is an emergency speech, then local GMM is performed secondly to classify the type of emergency speech. The proposed method is tested and verified by emergency and non-emergency speeches in various environmental conditions. Also, we discuss about the platform issues having analysis and inference function of big data in smart city.

Keywords: smart city, security, big data, speech recognition, platform

1   Introduction

In 1988, Mark Weiser introduced a new phrase to describe a new paradigm of computing. That phrase was ubiquitous computing and it refers to 'smart' and networked devices embedded within our environments. Now we are living in the era of ubiquitous computing. Ubiquitous computing is roughly the opposite of virtual reality**.** Usually, virtual reality puts people inside a computer-generated world, but ubiquitous computing forces the computer to live out here in the world with people. Virtual reality is primarily a horse power problem, but ubiquitous computing is a very difficult integration of human factors, computer science, engineering, and social sciences. Ubiquitous computing is a post-desktop model of human-computer interaction in which information processing has been thoroughly integrated into

everyday objects and activities. The core technology of ubiquitous computing is an autonomic collaboration model for ubiquitous fusion services in ubiquitous computing environment. To do some action in ubiquitous computing, many elements coordinate to complete such action. Ubiquitous networking in ubiquitous computing is completed by the collaboration of many types of networks, city technologies, city sociology and governance[1,2,3].

Ubiquitous city is a city with a virtual overlay on top of its physical construct that allows people to interact with their environment everywhere and in real time for their convenient better quality life.

Nowadays, the pervasiveness of smart phones, tablets and laptops has led to instant communication via the web, and that has profoundly changed human behavior. We make a blog, tweet and share status messages to a global audience with little cost. Information and social or otherwise is overabundant and cheap. The internet or the ubiquitous platform has been with us for a while and all of that is really obvious.

In short, smart city is a ubiquitous city having high intelligent capability with a lot of smart device and software. Nowadays smart city us more popular than ubiquitous city. Smart city is not a 20th century city with a few computers and screens thrown at it. It is not a just conventional city with its citizens carrying mobile phones. So, it needs a different planning or designing concept.

Actually, there are a lot of different typed-data to be analyzed especially in smart city. Because many smart devices and software produce a lot of data with or without our intention. So, it is very important to deal with security strategy for abundant data in smart city.

In this paper, we will survey and discuss some important security issues when we designing smart cities at this paper.

2 The concept of smart city

In computerized generation, industrial society paradigm was popular. That means that making some products by industry was so important at that time. After that as time goes by, in 'e' generation and 'u' generation, information society paradigm is more popular. That means that information is very important thing[3].

Main keywords in new paradigm are the power of people, collective intelligence and new ecosystem. The power of people use mobile device and social media increasing the power of people in life. Collective intelligence is the creation of value by collective intelligence as well as collective power. New ecosystem is the creation of additional value by new ecosystem using open as well as sharing.

Actually, the goal of ubiquitous computing is to make the safety, security and peace of mind society. The key technology to be a ubiquitous society is ubiquitous sensor network, in short USN. There are many application areas that use USN, such as home control, health medical care, welfare, hospital, school or education, traffic, culture, sports, environments, tracking of produce, anti-disaster, city security, etc.

Nowadays smart society is dawning. If we disregard this changes or needs, we will degenerate from information ages. Smart paradigm is just an extension of ubiquitous paradigm, so called post-ubiquitous age.

There are four types of viewpoints in ubiquitous computing. In the aspect of computing paradigm, the technology is changed from static community to automatic community. In the aspect of technical principle, the technology is changed from situation aware to self growing. However, in the aspect of application domain, the technology is changed from healthcare to environment preservation. The fourth aspect concerns the application area, the technology is changed from home or building to society[1,3].

Smart city can be identified (and ranked) along six main axes or dimensions[4,5]. These axes are smart economy, smart mobility, smart environment, smart people, smart living, and smart governance. These six axes connect with traditional regional and neoclassical theories of urban growth and development. In particular, the axes are based -respectively- on theories of regional competitiveness, transport and ICT economics, natural resources, human and social capital, quality of life, and participation of citizens in the governance of cities.

A city can be defined as 'smart' when investments in human and social capital and traditional and modern ICT communication infrastructure fuel sustainable economic development and a high quality of life, with a wise management of natural resources, through participatory governance.

Why the smart city is necessary? It is for maximizing the utilization and functionality of space as well as increasing the value of the city. Also, it is for providing a better quality of life. For that linking the city internationally is very important. Nowadays the environment of our life is very important, so reducing many kinds of pollution is another goal of smart city.

Among many application areas of ubiquitous, smart city is the constructed city by ubiquitous technologies and paradigm. Usually many ubiquitous services and hybrid technologies using RFID, USN, IPv6, sensing devices, monitoring, auto-control, real-time management etc. make smart city.

However, there are three important issues to be s smart city as of now. Firstly, the most important issue is the research about the infra structure of smart city such as platform, security, service scenario etc. The second issue is the research about the paradigm of smart city such as role play between government and local government to perform smart city etc. The third issue is the research about the consulting of smart city such as the best service model according to many types of organs, and business model, and so on.

In short, smart city is a combination of safety and security for the sake of peaceful society. Now many services of ICT mobile devices are provided with personalized manners. For completion of real ubiquitous space through smart city, the standard model and some platforms should be a prerequisite condition for success.

In this paper, our research is mainly focused on the first issue such as security issue and platform in designing smart city.

3 Security issues

3.1 Speech recognition issue

In smart city, how to implement secure city is important issue. To give a smart capability in smart city, it is usually used CCTV(closed circuit television) for

monitoring the city. However, many CCTVs catch or sense a lot of image information from the defined area rather than speech sound. However, even though they sense speech sound, it is hard to manage the emergency state or situation. For example, if somebody is in some emergency situation, he or she shouts as loud as possible. In this case, because the CCTVs only catch the images from the restricted or defined area, in case of the emergency speech or sound is a far from the CCTVs, it is hard to manage the emergency situation quickly. But if the sound is not an emergency sound, it is no problem. Only emergency sound is to be managed as soon as possible for secure city. To realize this concept is necessary for secure smart cities.
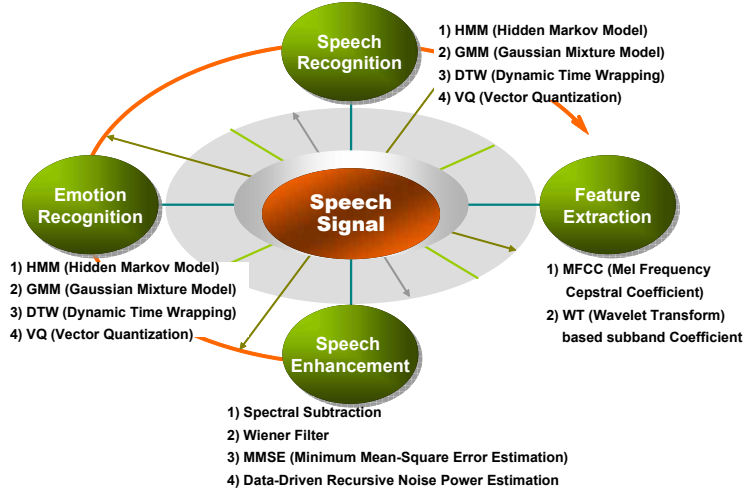
However, the difficult for speech recognition is the background noise. As the noise is the main cause for decreasing the performance, the place or environment is very important in speech recognition[6]. Here, we propose a simulated classification method for emergency and non-emergency speech.

For the purpose that, we use GMM(Gaussian Mixture Model). GMM is among the most statistically mature method for clustering (though they are also used intensively for density estimation). We introduce the concept of clustering, and see how one form of clustering in which we assume that individual data points are generated by first choosing one of a set of multivariate Gaussians and then sampling from them can be a well-defined computational operation. We see how to learn such a thing from data, and we discover that an optimization approach. This optimization method is called Expectation Maximization (EM). We spend some time giving a few high level explanations and demonstrations of EM, which turns out to be valuable for many other algorithms beyond Gaussian Mixture[7,8,9].

Therefore the proposed method determines whether input speech is emergency or non-emergency speech by global GMM. If the sensing speech is classified as an emergency speech, then local GMM is performed to recognize the type of emergency speech. By the simulation result, the proposed method is well classified whether the speech has noise or not([Fig.3]).

One of the key factors in the speech recognition is quite different from the controlled environment of the speech laboratory. However, the surrounding noises are a particularly difficult problem in the real speech recognition. The difference in the controlled environment and the real environment in speech recognition comes into play in three distinct processes: signal process, feature space process, and model process. Of these three processes, the difference is most evident in the signal process[10]. Here, the noise in the speech data after the signal process is filtered by a novel digital filtering system. A FIR(Finite Impulse Response) filter [6,10] is first used to separate the speech region and the noise region, and then a Wiener filter is used to improve the overall speech recognition ([Fig.1]).

Generally, the speech recognition system is configured by six stages. In stage 1, voice data are inputted by converting the audio signals into the electrical digital signals. In stage 2, the voice signals are separated from the surrounding noises. In stage 3, useful traits in speech recognition are extracted by using a speech recognition model. In stage 4, a standard speech pattern database is formed by speech recognition training. In stage 5, the new voice data are compared to the standard speech pattern a base, and the closest match is searched. In the final stage of 6, the matched result is put to use through the user interface.

Speech Recognition

1) HMM (Hidden Markov Model)
2) GMM (Gaussian Mixture Model)
3) DTW (Dynamic Time Wrapping)
4) VQ (Vector Quantization)

Emotion Recognition

1) HMM (Hidden Markov Model)
2) GMM (Gaussian Mixture Model)
3) DTW (Dynamic Time Wrapping)
4) VQ (Vector Quantization)

Speech Signal

Feature Extraction

1) MFCC (Mel Frequency Cepstral Coefficient)
2) WT (Wavelet Transform) based subband Coefficient

Speech Enhancement

1) Spectral Subtraction
2) Wiener Filter
3) MMSE (Minimum Mean-Square Error Estimation)
4) Data-Driven Recursive Noise Power Estimation

Fig.1. Speech recognition summary

In the preprocessing (noise elimination) stage of 2, analog audio signals from a CCTVs or sensors are digitized and then fed to the digital filter. The digital filter, which is widely used and proven, selects the pass-band and filter out the stop-band.

Depending on the presence of feedback processes, the digital filter is divided into IIR(Infinite Impulse Response) and FIR filters. The latter is known to be less error-prone. For noise elimination, the Wiener and Kalman filters[4] are widely used. In emergency situations that require accurate interpretation of a rather brief voice data, the Wiener filter is usually preferred. The general model-based Wiener filtering process can be expressed as follows:

$$\hat{s}(t) = g(t) * (s(t) + n(t)) \tag{1}$$

Where, $\hat{s}(t)$ is the speech to be recognized, $s(t)$ is the speech data containing noise, $n(t)$ is the noise, and $g(t)$ is the Wiener filter. In Eq. (1), $s(t)$ is being sought. In it, an estimate of $n(t)$ is derived from $s(t)$, and then the approximate value of $s(t)$ is obtained by using $n(t)$. In order to achieve a better approximation of $s(t)$, the GMM as expressed below in Eq.(2) is used. It expresses mathematically the general characteristic of speech data.

$$P(s) = \sum_{k}^{K} p(k) N(s; \mu_K; \sum k) \tag{2}$$

Based on Eq. (2), the model-based Wiener filter is designed per following steps: In the inputted current frame, the noise region is determined by a statistically-based VAD. In the noise region found, the noise model is renewed to the previous value. In

the preprocess-WF block, a temporally noise-free clean speech is estimated using the decision-directed Wiener filter. Using the estimated values from the previous step, the Gaussian post probabilities of the GMM are calculated. In the final, the probabilities are used to estimate the noise-free clean speech. The estimated noise-free speech and the noise model are used to design the final Wiener filter. The current frame is processed using the Wiener filter designed, and the noise-free clean speech is obtained. Then the above five steps are repeated for the next frame.

Because of the reason of the performance, we propose a fast recognition filtering method for emergency detection. The basic concept is to selectively use the audio signal being transmitted from the CCTVs'. That is, from the transmitted signal, only the audio energy spectrum that is relevant to the speech is to be selected, digitized, and saved for further analysis. A high-performance FIR Wiener filer can be used to digitally filter out the unwanted portion of the audio signal, prior to actual speech recognition.
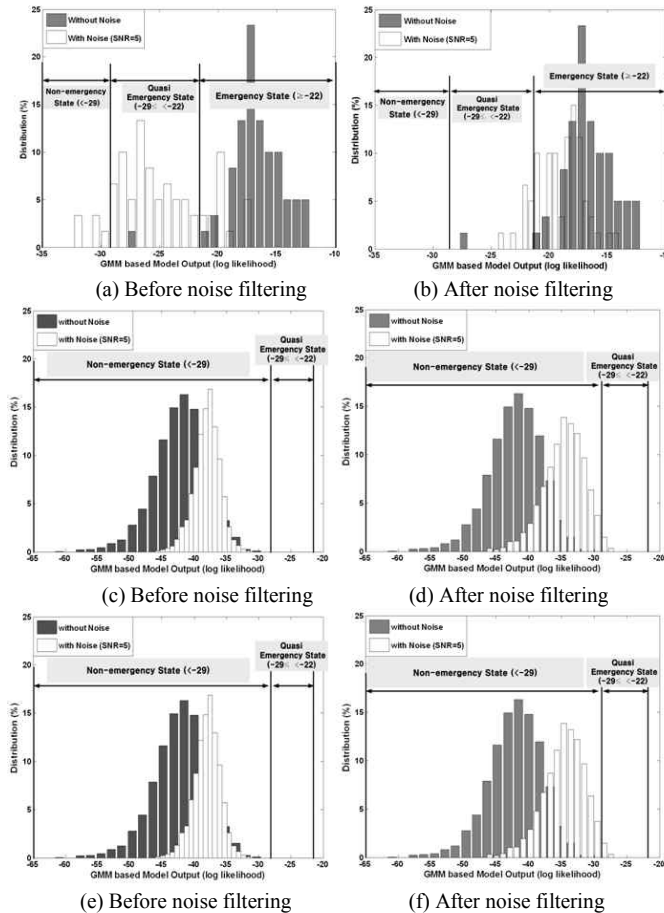
Fig.2 is the proposed detection and recognition algorithm of emergency speech. After speech sound is detected, noise filtering and speech detection algorithms are applied. And then, through the feature extraction and similarity method by GMM we can decide whether the sound is emergency sound or not. After determination, we can apply the sound to emergency models in DB, and then finally we decide the sound and display on the dedicated screen.
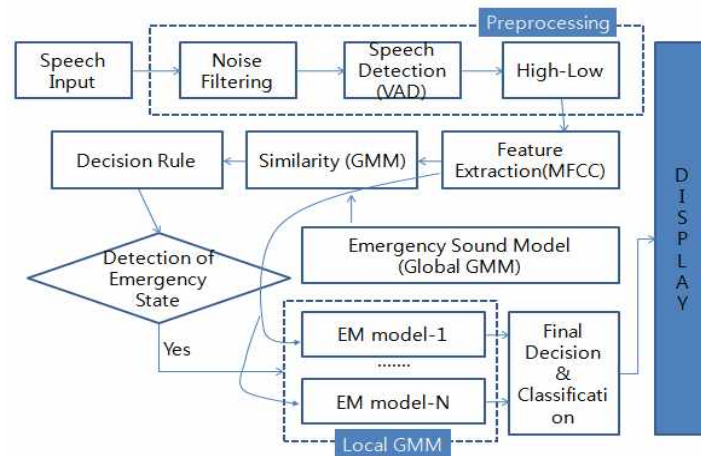
Fig.2. The proposed detection and recognition process for emergency speech

In this paper we construct the global GMM and local GMM. Global GMM is for detection whether the sound is emergency sound or not. Global GMM is constructed based on the feature extraction of emergency sound.   If the sound is in emergency sound, the type or the speech of the emergency sound is determined in local GMM process.

Finally, Fig.3 is shown the results of our method.

(a) Before noise filtering   (b) After noise filtering

(c) Before noise filtering   (d) After noise filtering

(e) Before noise filtering   (f) After noise filtering

Fig.3. The estimation results of GMM (upper: in case of white noise, middle : in case of having general noise, lower :in case of having automobile noise)

3.2 Platform issue

Smart city has to have the following functions such as instrumented, interconnected and intelligent functions. To realize those functions, smart city has to have a standard protocol(platform) to manage many kinds of activities. Platform is the structure and law for easy transaction among groups. Usually the platform of a smart city is as important as services. Platform is a system structure for information service of special area. The role of platform is for communication between side A and side B. For the purpose of that, they have to have platform having architecture and some rule for protocol. If not, it is hard to communicate each other as well as share some information. The goal of smart city is for collaboration. To reach this goal, the platform should prepare first of all.

Fig.4 is a smart city platform which is developed by Korea government at 2009[11] firstly. But it is updated as of now. In Fig.4, the number 1 is for infra interface, number 2 is for external service interfaces. The number 3 is for core utility services, so security functions are included here. The number 4 is for workflow management, so situation awareness or event management functions are implemented here. The number 5 is for information hub, and finally number 6 is for integrated databases. From the platform, we can realize that the number 3 and 4 modules have lots of important or core functions to be a smart city.

Now intelligent surveillance technologies are developed in the platform. So, the framework of platform is like Fig.5. Smart platform should have analysis capability of many kinds of data or situation. So inference capability should include here.
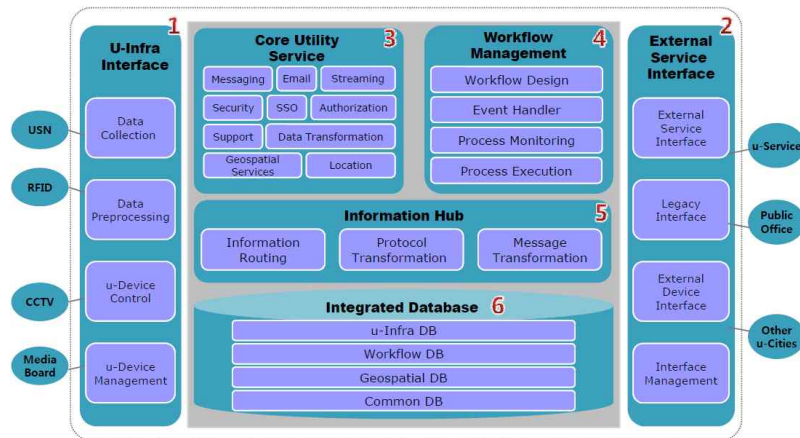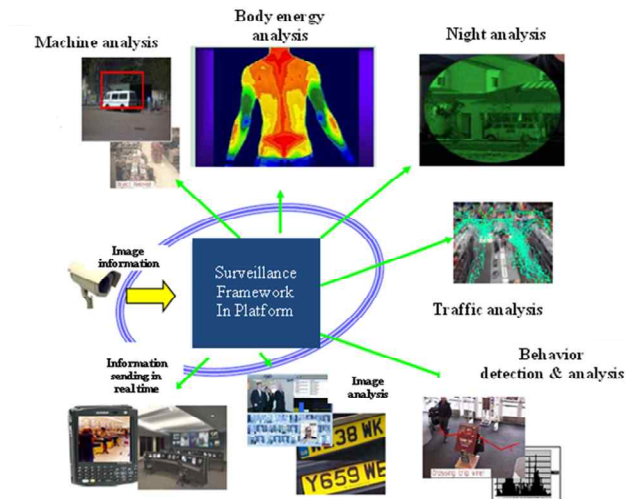

Fig.4. Smart city platform

**Fig.5.** Surveillance framework in platform

In information technology, the 'big data' is a loosely-defined term used to describe data sets so large and complex that they become awkward to work with using on-hand database management tools. Difficulties include capture, storage, search, sharing, analysis, and visualization. The trend to larger data sets is due to the additional information derivable from analysis of a single large set of related data, as compared to separate smaller sets with the same total amount of data[12].

Data from citizens, systems, and general things are the single most scalable resources available to smart city stakeholders today. This big data is constantly captured through sensors and from open data sources. More and more data services for city officials, utility services, and citizens become available, which allows efficient access and use of big data, a necessary requirement for smart cities.

Advances in digital sensors, communications, computation, and storage have created huge collections of data, capturing information of value to business, science, government, and society[13]. To many, the term 'big data' refers to algorithms and software programs that help companies or researchers make discoveries and unearth trends by allowing them to visualize and analyze information better. But it has another meaning too. Big data literally means big data, dizzying amounts of customer records, sound recordings, images, text messages, Facebook comments and technical information that has to be stored, retrieved and understood in its proper context to be any good to anyone.

Historically, data analytics software hasn't had the capability to take a large data set and use all of it—or at least most of it—to compile a complete analysis for a query. Instead, it has relied on representative samplings, or subsets, of the information to render these reports, even though analyzing more information produces more accurate results.

Our approach is changing with the emergence of new big data analytics engines, such as Apache Hadoop, LexisNexis'HPCC Systems. These new platforms are causing 'the disappearing role of summarization'. So, we propose a comprehensive personalized multi-agent based information retrieval platform to support the big data analysis and process in Fig.6.
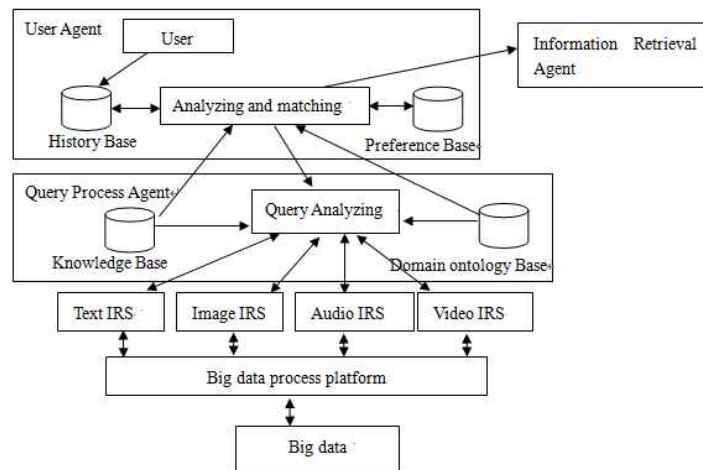
Fig.6. Comprehensive personalized information retrieval platform

There are two main functions of user agent: first, it interacts with users and accepts the users' retrieval query and returns the retrieval results back to users; second, it analyses the received user's query and matching the preference base of each user, then returns the estimated proposals back to the user for choosing the best one and finally sends the refined query to the Information Retrieval Agent. The main function of the Query Process Agent is to analyze which domain the query may belong to according to the domain ontology base and knowledge base. For different type of data, we us Map/Reduce function to analyze and process data. Map/Reduce computing model, reasoning algorithm was designed to process mass data. Query was also transformed into Finite Semantic Graph, and semantic matched with full graph. Different kinds of data should use different Map/Reduce function, and find the optimal Map/Reduce function.

4   Conclusion

Urban performance currently depends not only on the city's endowment of hard infrastructure, but also, and increasingly so, on the availability and quality of knowledge communication and social infrastructure. The latter form of capital is decisive for urban competitiveness. It is against this background that the concept of the 'smart city' has been introduced as a strategic device to encompass modern urban production factors in a common framework and to highlight the growing importance of ICT technologies, social and environmental capital in profiling the competitiveness of cities. The significance of these two assets - social and environmental capital - itself goes a long way to distinguish smart cities from their more technology-laden counterparts, drawing a clear line between them and what goes under the name of either digital or intelligent cities.

In this paper, we deal with the speech recognition issue firstly. This is for the recognition of emergency or non-emergency sound to detect the emergency situation in smart city. Unlike the controlled environment where a speech recognition system can easily filter out the extraneous noises, it is rather difficult in the real environment where sensors, such as CCTVs, collects abundant noises from various human, mechanical, and natural sources. The success of speech recognition in the real environment thus depends critically on how well these noises are filtered. Just as important, the processing time for noise filtering needs to be reduced, as time is the most critical element in emergency situations. Thus, effective noise filtering combined with fast processing time is considered to be the essence of speech recognition. Towards these goals, an improved speech recognition system is proposed in this work. The system has the FIR and Wiener filters as the key elements and effectively filters out the extraneous noises and produces clean noise-free speech data in a reasonable time. Here, we proposed the classification method between emergency sound and non-emergency sound by GMM. This method can do that if some sound is in emergency sound, the sound can be recognized by local GMM algorithm. From the simulation results, we can conclude that the recognition ratio is more that 80%, so that this method can apply in real situation.

And secondly we discuss with platform issue having the functions for not only

analysis but also inference. Big data is another inference issue. We discuss the structure of big data platform here.

REFERENCES

[1] Young Im Cho, "U-Society and U-Convergence Paradigm", 2008 Asia Pacific Woman Leaders Forum for Science & Technology, pp.183-204, September 3, 2008.

[2] Manuel Roman et al, A Middleware Infrastructure for Active Spaces, IEEE pervasive computing, 2002.

[3] Young Im Cho, "Practical Approaches for Ubiquitous Convergence System", Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International symposium on Advanced Intelligent Systems, September 17-21, 2008.

[4] Caragliu, A, Del Bo, C. & Nijkamp, P, "Smart cities in Europe", Serie Research Memoranda 0048 (VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics). http://ideas.repec.org/p/dgr/vuarem/2009-48.html. 2009.

[5] Giffinger, Rudolf; Christian Fertner, Hans Kramar, Robert Kalasek, Nataša Pichler-Milanovic, Evert Meijers, "Smart cities – Ranking of European medium-sized cities". http://www.smart-cities.eu/. Vienna: Centre of Regional Science. http://www.smart-cities.eu/download/smart_cities_final_report.pdf, Retrieved 2009-11-11, 2009.

[6] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction", IEEE Trans., ASSP, vol. 37, no. 2, pp. 113-120, 1979.

[7] Doclo, S., Rong Dong, Klasen, T.J., Wouters, J., Haykin, S., Moonen, M., "Extension of the multi-channel Wiener filter with ICTD cues for noise reduction in binaural hearing aids", Applications of Signal Processing to Audio and Acoustics, vol. 16, no. 16, pp 70-73, 2005.

[8] J. S. Erkelens and Richard Heusdens, "Tracking of nonstationary noise based on data-driven recursive noise power estimation," IEEE Trans. Audio, Speech and Language Processing , vol. 16, no. 6, pp. 1112-1123. 2008.

[9] Ethem Alpayd, "Soft vector quantization and the EM algorithm", Neural Networks, vol. 11, pp. 467-477, Issue 3, 1998.

[10] P. Dhanalakshmi., S. Palanivel, V. Ramalingam, "Classification of audio signals using AANN and GMM," Applied soft computing, vol. 11, no. 1, pp. 716-723, 2011.

[11] Yoo Jin Soo, U-City platform, Iinternational conference on U-City, August 27-28, 2009.

[12] White, Tom. Hadoop: The Definitive Guide. 1st Edition, O'Reilly Media, 2009.

[13] Amanda Spink a., Bernard J., Jansen b,, "A study of results overlap and uniqueness among major Web search engines", Information Processing and Management, 42, pp.1379–1391, 2006.