IOP Conf. Series: Materials Science and Engineering **335** (2018) 012060 doi:10.1088/1757-899X/335/1/012060

MMSE Estimator for Children's Speech with Car and Weather Noise

V Sayuthi*

Department of Information System, STMIK Indonesia Padang, Padang, Indonesia.

*vetho@stmikindonesia.ac.id

Abstract. Previous research mentioned that most people need and use vehicles for various purposes, in this recent time and future, as a means of traveling. Many ways can be done in a vehicle, such as for enjoying entertainment, and doing work, so vehicles not just only as a means of traveling. In this study, we will examine the children's speech from a girl in the vehicle that affected by noise disturbances from the sound source of car noise and the weather sound noise around it, in this case, the rainy weather noise. Vehicle sounds may be from car engine or car air conditioner. The minimum mean square error (MMSE) estimator is used as an attempt to obtain or detect the children's clear speech by representing simulation research as random process signal that factored by the autocorrelation of both the child's voice and the disturbance noise signal. This MMSE estimator can be considered as wiener filter as the clear sound are reconstructed again. We expected that the results of this study can help as the basis for development of entertainment or communication technology for passengers of vehicles in the future, particularly using MMSE estimators.

1. Introduction

Previous research mentioned that most people need and use vehicles for various purposes, in this recent time and future, as a means of traveling. Many ways can be done in a vehicle, such as for enjoying entertainment, and doing work, so vehicles not just only as a means of traveling [1].

This research focus on voice and speech activities which are done by human passengers in the vehicles, for example, speech activities of daily human doing or voice activities at work, where are done in the vehicles. To have speech activity services, the vehicles must need communication system to connect within and with outside environment.

Recent architecture of vehicle communication and mobile communication technology has served those activities. Femtocell mobile in the 5G communication cellular technology can make passenger to communicate within vehicle or to outside environment [2].

Besides that 5G networks, there are other technologies to realized it, some functions in these vehicles connect persons within it and with the outside of vehicles by "*connectivities*" known as v-2-x connectivities, include vehicle-to-sensor on-board (V2S), vehicle-to-vehicle (V2V), vehicle-to-road infrastructure (V2R), and vehicle-to-internet (V2I), all can be implemented by using wireless technology [3].

Some aspects of those technologies need better performances of wireless solution, frequency band, data rate, TX power, MAC protocol, and modulation. The application purposes of the technology must

be considered to work properly in v-2-x connectivities[3] and also in 5G cellular communication technology too.

The electronic computers had advanced in size and the performance, this makes the performance of communication technology had increased better, so it caused and helped the complexity and performance of speech technology increase and advance, and this trend is predicted will be continuous in the future. This system of communication technology needs the robustness of speech that will be significant to apply. Concern to have better performances of speech communication activities in vehicle by previous researches, we can conclude that the speech communication research area in vehicle consists of two main categories that must be considered, the speech processing and the speech communication technology, both are designed as speech communication system in the vehicles with vehicle based-centred, that can make human such as passenger or driver communicates each other within vehicle or outside environment, especially with voice interactive activities.

Many factors that must be concern in speech vehicle communication are robustness in environmental and internal vehicle sources noises, voice-interactive format dialogue and solutions, and the convenience and safety in mobile speech applications [4].

2. Human voice and Noise Sources

Many devices and speech processing for example in speech therapy for human voice or in communication, is intended for children user application, due to less complexity, recommended to above of six years old of age [6], and statistics said that for example in the Europe, percentages of number of passenger cars by age 5 to 10 years is around 25 % from 29 countries in Europe [7].In this study , we will examine the childern's speech from a girl of eight years old that satisfied [6] and [7] in the vehicle that affected by noise disturbances from the sound source of car noise and the weather sound noise around it, in this case, the rainy weather noise. Vehicle sounds may be from car engine or car air conditioner.

3. Car Noise

Noise in vehicle can happened in many version depends on many sources that appeared in the vehicle. Figure 2 explain some example that arise when noise within vehicle appears. [4].

We can divided the internal car noise by divided the categories of the sources of noise with specific and detail dividing, we can defined it, as follows [4];

1. First we define the type of vehicle, wether it is sedan, truck, or SUV.
2. Then if we choose sedan, which model of sedan vehicle that focused. The option can be Nissan Altima or Honda Civic
3. Then look out the condition of car we observed, wether it is stationary or in motion
4. Then we get into the car , it can be especially air conditioner , wether it is off or on
5. Any other significant real condition can be window status, or in other component in internal vehicle can be horn status.

4. Outside Environment and Weather Noise

Our research focus on vehicle that located in rainy weather condition. The geographical West Sumatra region, according to data has extreme daily rainfall values which ranges from 42.4 to 114.7 mm. These rainfall values are higher and more varied than the eastern region of Sumatra. That can be seen in figure 1 below, threshold of extreme rainfall in Sumatra island (1998-2012) [5].

This natural phenomena of rainy weather can make noise for vehicle, especially for human passengers those are doing speech activities in car or vehicles. It can be happened in the night or at noon time in West Sumatra region.

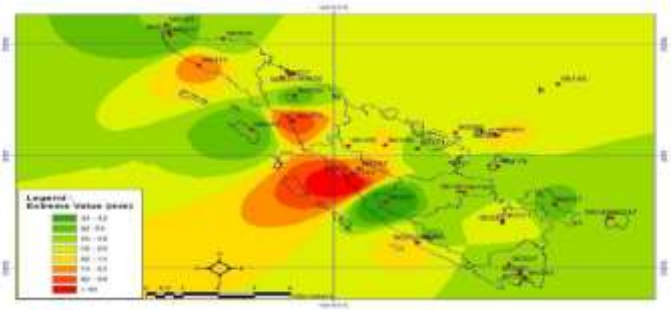

Figure 1. Threshold of extreme rainfall in Sumatra island (1998-2012) [5].

5. Speech Enhancement

Speech enhancement can be doing by three step, those are: 1. Data acquisition, 2. Analysis, computing, and method using of original signal detection, and 3. Removing, eliminating or reducing noise signal. [4], [11].

5.1. Audio Noise Reduction

Signal of speech and noise can be mixed in any area or environment for examples are in building, in outdoor street and within cars those are sources of noise which are located relatively near of person who speaks and hears the speeches.

If the signal of original speech, is wide-sense stationary, ergodic and uncorrelated with noise, it can be computed with wiener filter approach[10] and if the noise signal is white noise which has the constant magnitude spectrum along of frequency so it can be reduced with minimal error by using technique of Minimum Mean Square Error (MMSE) estimator which can be considered as a function of wiener filter [11] .

Let us denote original data speech signal as q_n, which is distorted by white noise p_n of noise sources, is received in our devices that hear it as o_n. So the equations

$$o_n = q_n + p_n \tag{1}$$

Wiener filter in this problem is Linear Time Invariant (LTI) System and Finite-duration Impulse Response (FIR) N-length Filter. The impulse response of this filter in not this interval $0 \leq n \leq N-1$ is $h(n) = 0$. [11]

By using Wiener filter approach of Simon Godsill works, until we have $J_{\min}$ and J_{actual} (Simon Godsill, 2015) [10], it can be followed as below equations, first is that we have autocorrelation function of noise p in discrete time k with σ_p^2 as,

$$r_{pp}[k] = \sigma_p^2 \, \delta[k] \tag{2}$$

Autocorrelation of noisy signal denote as $\mathrm{r}_{oo}[\mathrm{k}]$, while autocorrelation of desired signal is $\mathrm{r}_{qq}[\mathrm{k}]$, this value follows the equation

$$r_{oq}[k] = r_{qq}[k] \tag{3}$$

$\mathrm{r}_{\mathrm{pp}}[\mathrm{k}]$ can be measured by

$$r_{pp}[k] \approx \frac{1}{N} \sum_{n=0}^{N-1} o_n o_{n+k} \tag{4}$$

$$r_{qq}[k] = r_{oo}[k] - r_{pp}[k] \begin{cases} r_{oo}[k] & k \neq 0 \\ r_{oo}[0] - \sigma_p^2, & k = 0 \end{cases} \quad (5)$$

The wiener filter equation of h becomes

$$\boldsymbol{h} = \boldsymbol{R}_o{}^{-1}\boldsymbol{r}_{oq} \quad (6)$$

The theoretical minimum mean square error $J_{\min}$ of wiener filter equation of h is

$$J_{min} = r_{qq}[0] - \boldsymbol{r}_{oq}^T\boldsymbol{h} \quad (7)$$

and the actual mean square error J_{actual} of the true q_n from the wiener filter approach systems of h is

$$J_{actual} = \frac{1}{N}\sum_{n=0}^{N-1}(q_n - \hat{q}_n)^2 \quad (8)$$

6. Method and preparing data measurement of noised speech in vehicle

We can process data measurement by collected children voice in Indonesian language for the word of 'pohon' that equals to 'tree' in English. Then we collected data sounds of environment as shown in figure 2.

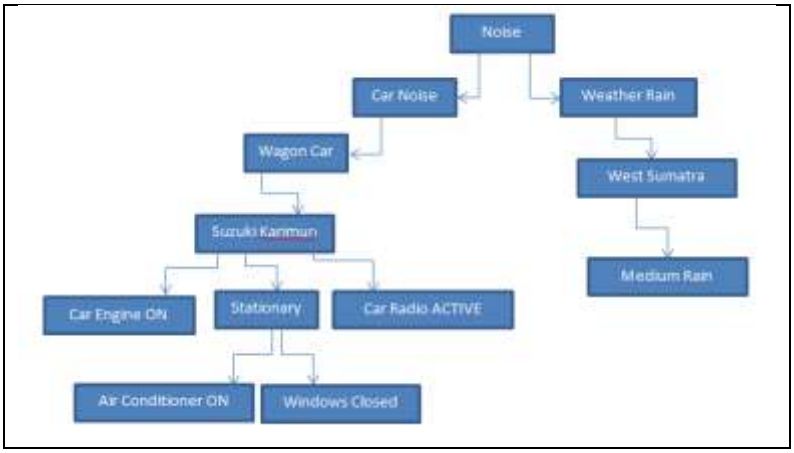

Figure 2. Car and weather noise tree structure. The leaves have most specific information [4].

7. Experimental Data Result of MMSE Enhancement Estimator

The minimum mean square error (MMSE) estimator is used as an attempt to obtain or detect the childern's clear speech by representing simulation research as random process signal that factored by the autocorrelation of both the child's voice and the disturbance noise signal. This MMSE estimator can be considered as wiener filter as the clear sound are reconstructed again.

With wiener filter of 14 taps we have the original speech signal, the noisy speech signal and the output as shown in figure 3,4,and 5. The quality result of our experiment is average in hearing of the reconstruction of children speech signal, because of the function of taps of wiener filter that values 14, and the method that we have done by conventional and basics or fundamentals of the wiener filter methods that mentioned in above explanation of our basics wiener filter methode.Result of experimentations are,

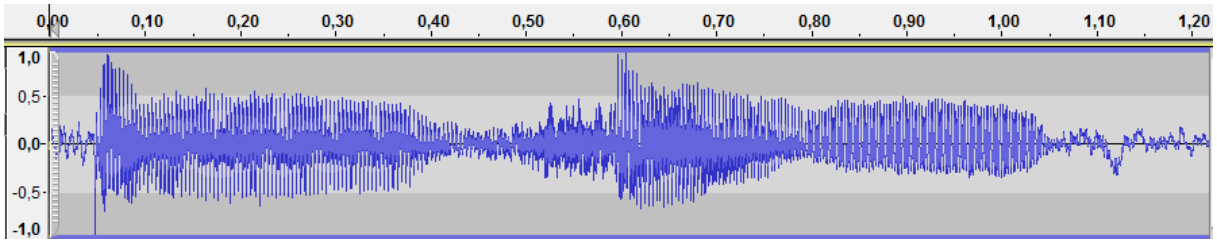

Figure 3. The original speech signal said 'pohon' in Indonesian language equals to 'tree' in English language.

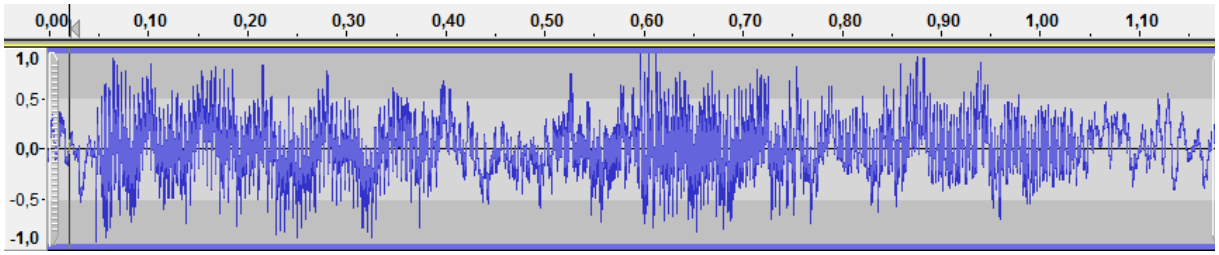

Figure 4. The noisy signal with impact of car and weather noises.

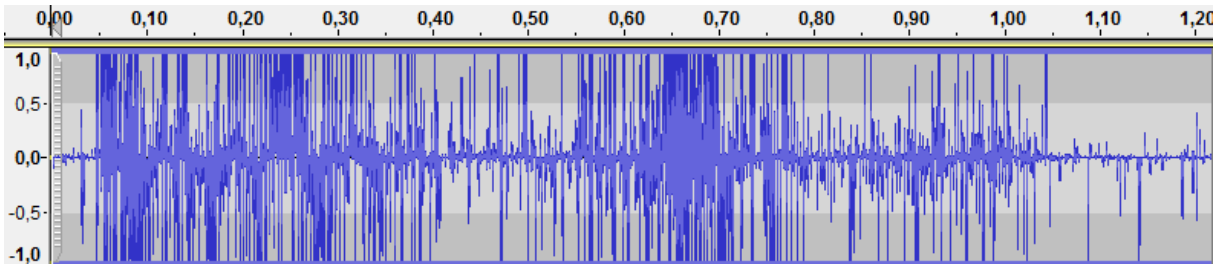

Figure 5. The output of wiener filter signal with taps 14 and autocorrelation scaling value of 0.0078.

8. Conclusion

As we expected that the results of this study can help as the basis for development of entertainment or communication technology for passengers of vehicles in the future, particularly using MMSE estimators. We have shown that with the conventional or basics or fundamentals of wiener filter we can contruct the original signal that affected by car and weather noise.With the limitation of the number of taps in wiener filter we must study in the next future of the number of other taps that may be improved our this final result of the conventional wiener filter to reconstruct or to bring out the original speech signal especially in a car.

With the used of conventional method of wiener filter it have some of difficulties for example in constructed or make formulas and algorithm in this wiener filter method, so hopefully in next future there are some ways to compute this wiener filter algorithms in the better form.

The estimator MMSE can detect speech voice especially children in destructed ways by vehicle sounds or various noises that will be happen may be from the car engine or car air conditioner noises. The techniques to improve the way for minimizing noises in these environments is needed in the future.

Acknowledgements
Our thanks to STMIK INDONESIA PADANG institute and AMAL BAKTI MUKMIN PADANG, foundation who help to succeed the research in contract number of 108/A.12/STMIK-I/2017.

References

[1] Anjan C, Joseph G, Matthew J, Haans-Werner K, and Philip W 2001 Telematics: Decision Time for Detroit *Business Strategy Review.* Vol 12 issue 2 p 21-38
[2] Cheng-Xiang Wang, Fourat Haider, Xiqi Gao and Xiao-Hu You, Yang Yang, Dongfeng Yuan, Hadi M. Aggoune, Harald Haas, Simon Fletcher, Erol Hepsaydir, 2014 5G WIRELESS COMMUNICATION SYSTEMS: PROSPECTS AND CHALLENGES Cellular Architecture and Key Technologies for 5G Wireless Communication Networks *IEEE Communications Magazine.* p 124
[3] Ning Lu, Nan Cheng, Ning Zhang, Xuemin (Sherman) Shen, Jon W. Mark, Connected Vehicles: Solutions and Challenges *IEEE Internet of Things Journal.* Vol 1 issue 4 p 1-4
[4] Nitish Krishnamurthy and John H.L. Hansen 2014 Car noise verification and applications *Interational Journal of Speech Technology.* Vol 17 issue 2 p 167
[5] S Supriyadi, R Hidayati, R Hidayat, and A Sopaheluwakan 2017 Mapping Extreme Rain Conditions in Sumatra by Influence Global Conditions *IOP Conference Series: Earth and Environmental Science.*
[6] Natalia Zharkova, Fiona E. Gibbon and Alice Lee 2016 Using Ultrasound Tongue Imaging to Identify Covert Contrasts in Children's Speech *Clinical Linguistics & Phonetics.* Vol 31 issue 1 p 32
[7] Eurostat 2017 Passenger Cars in the EU *Eurostat.*
[8] Sharmistha S. Gray, Daniel Willett, Jianhua Lu, Joel Pinto, Paul Maergner, and Nathan Bodenstab 2014 Child Automatic Speech Recognition for US English: Child Interaction with Living-Room-Electronic-Devices *Proceedings of Workshop on Child Computer Interaction.*
[9] Pardeep Sharma, Rajesh Mehra, Naveen Dubey 2015 *4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)*
[10] Simon Godsill, 2015, Random Processes - Random Processes, Optimal Filtering and Model-based Signal Processing. *Lecture Note, University of Cambridge, Signal Processing and Communications Laboratory, Department of Engineering.*
[11] Alan V. Oppenheim and George C. Verghese, 2010, Signals, Systems, and Inference. *Class Notes for 6.011: Introduction to Communication, Control and Signal Processing*.
[12] Saeed V. Vaseghi 2000, Advanced Digital Signal Processing and Noise Reduction , Second Edition. *John Wiley & Sons Ltd.*