International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol. 10, No. 5 (2017), pp. 45-56
http://dx.doi.org/10.14257/ijsip.2017.10.5.04

Effective Speech Enhancement Algorithm for Mobile Communication Using Cascading of Frequency and Time Domain Algorithms

Harish Chander[1*], Balwinder Singh[2] and Ravinder Khanna[3]

[1]PhD Scholar, Inder Kumar Gujral Punjab Technical University, Jallandhar, India
[2]Centre for Development of Advanced Computing, Mohali, India
[3]Department of Electronics & Communication Engineering, Maharishi Markandeshwar University, Sadopur, India
[1]harishrajni@rediffmail.com, [2]balwinder@cdac.in, [3]ravikh2006@gmail.com

Abstract

In this paper we present a speech enhancement algorithm which is developed by cascading the time and spectral domain speech enhancement algorithms. For time domain we have used Kalman filter and for spectral domain we have used "Improved Minima Controlled Recursive Averaging" algorithm presented by Israel Cohen. Both of these algorithms give better results in their respective domains. The cascaded algorithm presented in this paper is tested under various types of real world noises that are generally experienced by a mobile user. The performance of the cascaded algorithms is evaluated using three widely used speech quality objective parameters, the Signal-to-Noise Ratio (SNR), Segmental SNR and Perceptual Evaluation of Speech Quality (PESQ). The simulation is performed in MatLab signal processing tool box. Comparative study of experimental results proves that there is substantial improvement in SNR, Segmental SNR and PESQ of the enhanced speech.

Keywords*: Speech Enhancement; Kalman Filter; Spectral Domain; Time Domain; SNR; PESQ*

1. Introduction

Speech communication has been the prime and effective way of communication ever since the beginning of human creation. With the advent of mobile phones the possibility of communicating with nears and dears at any time and from any place have become a reality. The mobile users now expect smooth communication even at situations which are highly noisy. Some of these situations are, inside a room with exhaust fan on, inside a running car, at railway platform during announcement and train arrival/departure, party places, multiple people talking (multitalker babble), on road with heavy traffic *etc.* Researchers, in the past, have developed several algorithms to overcome this problem. Most of these algorithms are based on processing the speech signal either in spectral domain or in temporal domain [1]. So far no algorithm is able to improve the speech quality as well as the intelligibility for all types of stationary and non-stationary noises and under all sorts of SNR conditions [2]. In this paper we have explored a combination of spectral and temporal algorithms and have achieved comparatively better results under all sorts of noises and SNR environments.

[*] Corresponding Author

ISSN: 2005-4254 IJSIP

1.1. Spectral Speech Enhancement Algorithms

Boll [3] presented one of the first algorithms for speech enhancement using spectral subtraction. Drawback of this algorithm was that a negative residual noise called musical noise was introduced. Later on Berouti and others [4] tried to remove the musical noise to some extent but was not successful under low SNR conditions. The MMSE based algorithms presented by Ephraim and Malah [5]-[6] became backbone of all modern spectral domain speech enhancement algorithms. These algorithms too failed due to non availability of efficient method to estimate a-priori SNR, as needed by these algorithms. Wolf and Godsill [7] tried to resolve the issue of estimation of a-priori noise estimation problem by presenting computationally efficient algorithms. Israel Cohen [8] in his work, 'Improved Minima Controlled Recursive Averaging', has suggested a better method for noise estimation, based on averaging past spectral power values using a time-varying frequency dependent smoothing parameter adjusted by signal presence probability. In this paper we have used the approach presented by Cohen for cascading with temporal domain algorithm.

1.2. Temporal Speech Enhancement Algorithms

In the temporal speech processing the corrupted speech signal is directly filtered using one of the time domain filters. Frazier and others [9] used comb filtering by exploiting the periodicity and pitch period of the voiced signal. Lim and Oppenheim [10]-[11] used Weiner filtering and linear predictive coding which is based on autoregressive model of speech. Kalman filter in speech enhancement was explored by various researchers such as Whipple and Basu [12], Sorqvist and others [13], Goh and others [14], Wu and others [15], Kybic [16] and Popescu and Zeljkovic [17]. Kalman filter, which has advantages over others, has been used in the present cascaded algorithm because it uses finite data sets and can adapt to both the stationary and non stationary speech signals and noises.

2. Improved Minima Controlled Recursive Averaging (IMCRA) Algorithm

One of the important aspects of an efficient speech enhancement algorithm is accurate estimation of noise signal and speech presence interval detection. In IMCRA the noise is estimated by averaging past spectral power values using a smoothing parameter which is controlled by the minimum values of the smoothing parameter adjusted by the speech presence probability in sub bands. The detection of speech presence is carried out in two stage iteration. In the first iteration speech presence periods are estimated roughly and in the second iteration stronger speech components are eliminated thus ensuring minimum tracking during speech presence [18]-[19]. Speech presence probability is controlled more in speech absence period and very less during speech presence periods. The working of the complete IMCRA algorithm is explained below:

Let $x(t)$ and $d(t)$ represent pure speech and noise signals in time domain. Since noise is considered to be additive in nature, the resultant noise corrupted speech signal is represented by $y(t) = x(t) + d(t)$. After sampling, it is represented as $y(n) = x(n) + d(n)$, where n is the sampling instant. The noise mixed speech signal $y(n)$ is divided into short time overlapping frames using a window function and then transformed into frequency domain using short time Fourier transform (STFT) [20]. In STFT form, $y(n)$ is represented as $Y(k, l) = X(k, l) + D(k, l)$, where k and l are the frequency bin and frame index respectively. The *a posteriori* and *a priori* SNRs are denoted by $\gamma(k, l)$ [21] and $\tilde{\xi}(k, l)$ [19] and are determined using equations (1) and (2) respectively.

$$\gamma(k,l) \triangleq \frac{|Y(k,l)|^2}{\lambda_d(k,l)} \tag{1}$$

Where $|Y(k, l)|^2$ is the power spectrum of noisy speech and $\lambda_d(k, l)$ is estimation of noise spectrum

$$\tilde{\xi}(k,l) = \alpha G_{H_1}^2(k,l-1)\gamma(k,l-1) + (1-\alpha)\max\{\gamma(k,l)-1,0\} \tag{2}$$

Where α is taken as the weighting factor that controls the trade-off between noise reduction and speech distortion [5], [22] and G_{H_1} is the conditional spectral gain function of the Log Spectral Amplitude estimator, when speech is definitely present and is determined with the help of the following expression.

$$G_{H_1}(k,l) \triangleq \frac{\tilde{\xi}(k,l)}{1+\tilde{\xi}(k,l)}\exp\left(\frac{1}{2}\int_{v(k,l)}^{\infty}\frac{e^{-t}}{t}dt\right) \tag{3}$$

Where $v(k,l) \triangleq \gamma(k,l)\tilde{\xi}(k,l)/\{1+\tilde{\xi}(k,l)\}$

1st iteration smoothed power spectrum of the speech signal in frequency and time domain are denoted by $S_f(k, l)$ and $S(k, l)$ respectively and are obtained using following equations.

$$S_f(k,l) = \sum_{i=-w}^{w} b(i)|Y(k-i),l)|^2 \tag{4}$$

$$S(k,l) = \alpha_s S(k,l-1) + (1-\alpha_s)\, S_f(k,l) \tag{5}$$

Where α_s is the smoothing parameter whose value is typically set to 0.9 and b is the Henning window.

Past minimum values $S_{min}(k, l)$ of $S(k, l)$ within a finite window length of D are stored separately. Rough estimation of the speech presence probability $I(k, l)$ is found using following relation.

$$I(k,l) = \begin{cases} 1, if\ \gamma_{min}(k,l) < \gamma_0\ and\ \zeta(k,l) < \zeta_0 \\ 0, otherwise \end{cases} \tag{6}$$

Where $\gamma_{min}(k,l) \triangleq \frac{|Y(k,l)|^2}{B_{min}S_{min}(k,l)}$, and $\zeta(k,l) \triangleq \frac{S(k,l)}{B_{min}S_{min}(k,l)}$. The values of γ_0 and ζ_0 are set to 4.6 and 1.67 respectively.

2nd iteration smoothed power spectrum of the speech signal in frequency and time domain are denoted by $\tilde{S}_f(k, l)$ and $\tilde{S}(k, l)$ respectively and are obtained using following equations.

$$\tilde{S}_f(k,l) = \begin{cases} \frac{\sum_{i=-w}^{w} b(i)I(k-i,l)|Y(k-i,l)|^2}{\sum_{i=-w}^{w} b(i)I(k-i,l)}, if\ \sum_{i=-w}^{w} I(k-i,l) \neq 0 \\ \tilde{S}_f(k,l-1), \quad otherwise \end{cases} \tag{7}$$

$$\tilde{S}(k,l) = \alpha_s\tilde{S}(k,l-1) + (1-\alpha_s)\tilde{S}_f(k,l) \tag{8}$$

Minimum tracking of the smoothed power spectrum in the second iteration is stored as $\tilde{S}_{min}(k, l)$. Speech absence probability is denoted by $\hat{q}(k, l)$ and is estimated as below:

$$\hat{q}(k,l) = \begin{cases} 1, & \text{if } \tilde{\gamma}_{min}(k,l) \leq 1 \\ (\gamma_1 - \frac{\tilde{\gamma}_{min}(k,l)}{\gamma_1 - 1}, & \text{if } 1 < \tilde{\gamma}_{min}(k,l) < \gamma_1 and\ \tilde{\zeta}(k,l) < \zeta_0 \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

Where $\tilde{\gamma}_{min}(k,l) \triangleq \frac{|Y(k,l)|^2}{B_{min}\tilde{S}_{min}(k,l)}$, and $\tilde{\zeta}(k,l) \triangleq \frac{S(k,l)}{B_{min}\tilde{S}_{min}(k,l)}$, γ_1 is the threshold, typically set to 3, B_{min} represents the bias of a minimum noise estimate whose value depends on window length D and smoothing parameter α_s. Typically the value of B_{min} is set to 1.66.

Final speech presence probability is denoted by $p(k, l)$ and is estimated as below:

$$p(k,l) = \left\{1 + \frac{\hat{q}(k,l)}{1-\hat{q}(k,l)}(1 + \tilde{\xi}(k,l)\exp(-v(k,l))\right\}^{-1} \tag{10}$$

Finally, the recursive averaging of the noise spectrum is obtained using following expression.

$$\bar{\lambda}_d(k,l+1) = \beta[\tilde{\alpha}_d(k,l)\bar{\lambda}_d(k,l) + \{1 - \tilde{\alpha}_d(k,l)\}|Y(k,l)|^2] \tag{11}$$

Where $\tilde{\alpha}_d(k,l) \triangleq \alpha_d + (1 - \alpha_d)p(k,l)$ is time varying frequency dependent smoothing parameter which is adjusted by the speech presence probability and β is the bias compensating parameter in the absence of speech and is typically set to 1.47.

3. Kalman Filtering

Kalman filter is one of the finest time domain filters that provide optimum recursive solution using least square method [16]. Kalman filter works on the principle of prediction and correction with feedback control. For noisy speech signal $y(n) = x(n) + d(n)$, where $x(n)$ and $d(n)$ are clean speech and noise signals respectively, the prediction and correction equations of Kalman filter are represented as below.

$$x_{n|n-1} = A_{n,n-1} . x_{n-1} \tag{12}$$

$$R_{e,n|n-1} = A_{n,n-1} . R_{e,n-1} . A^T_{n,n-1} + u . R_w . u^T \tag{13}$$

$$R_{e,n} = R_{e,n|n-1} - R_{e,n|n-1} . C^T . F_n^{-1} . C . R_{e,n|n-1} \tag{14}$$

$$x_n = x_{n|n-1} + R_{e,n|n-1} C^T . F_n^{-1}(y_n - C . x_{n|n-1}) \tag{15}$$

$$F_n = C . R_{e,n|n-1} . C^T + R_v \tag{16}$$

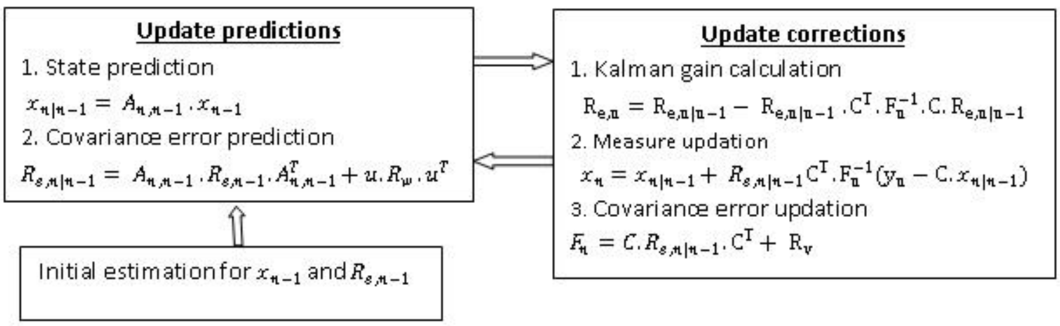


Figure 1. Working of Kalman Algorithm

Where A is mxm matrix that relates the present state with the previous state, R_w and R_v represent covariance matrices of the perturbation of the process and measure respectively, $R_{e,n}$ is the Kalman gain chosen such that it minimizes the covariance error of the next state. C is nxm matrix that relates the present state with the measure Y_n. The complete algorithm is described with the help of block diagram in Figure1.

4. Proposed Cascading of IMCRA and Kalman Filtering

It has already been established that time domain and frequency domain speech enhancement algorithms alone do not provide complete enhancement of the noisy speech which can improve speech quality as well as speech intelligibility under all types of noise environments and SNR conditions. In this paper we propose an algorithm which is cascading of IMCRA and Kalman filtering. We have tried two way cascading, *i.e*. IMCRA-Kalman and Kalman-IMCRA. In both the algorithms we have achieved better results than the individual IMCRA and Kalman algorithms. Figure 2 shows the block diagram representation of the cascading algorithms, where (a) represents first IMCRA followed by Kalman and (b) represents first Kalman followed by IMMCRA.

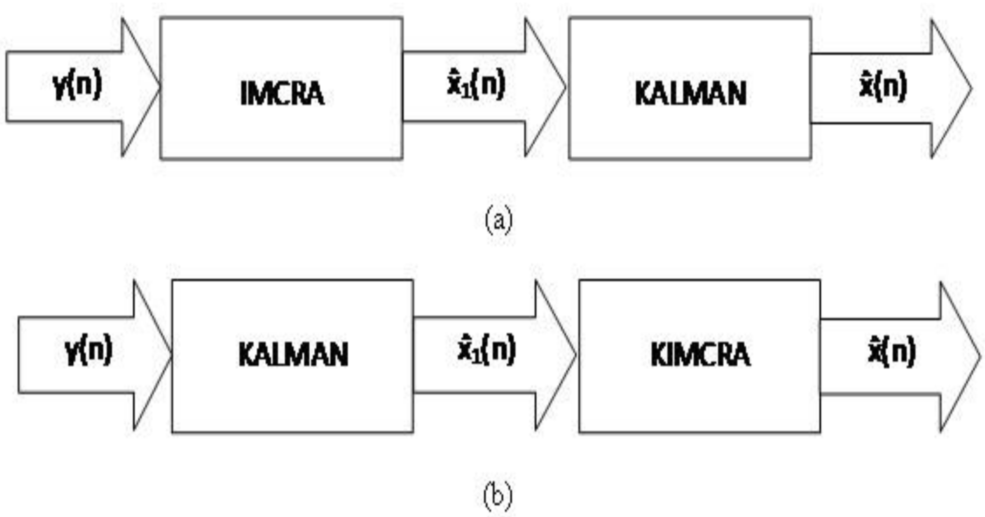


Figure 2. Block Diagram Representation of Spectral and Time Domain Cascading Algorithms, (a) IMCRA-KALMA, (b) KALMAN-IMCRA

5. Performance Evaluation

Performance of all the algorithms is evaluated using MATLAB tool. The pure speech signals are taken from TIMIT database [23]. Four sentences, two from male and two from female speakers, as depicted in Table 1, have been used to simulate the algorithms. To present real time analysis, the noise signals, used in the analysis are real time noises recorded using a Nokia mobile phone under typical Indian noise environments as depicted in Table 2. The speech signal is sampled at the rate of 16 kHz which is mixed with noises at global SNRs of -5dB, 0dB, 5dB and 10dB. For IMCRA algorithm, the noise degraded speech signal is divided into overlapping samples of 32ms (512 samples) each using Hamming windows of 512 samples. Figure 3 (a) shows pure speech signal of sentence sp01, “the birch canoe slid on the smooth planks”. The speech sp01 degraded with multitalker babble noise ns01 at (-5) dB SNR is shown in Figure 3 (b). The speech sp01 degraded with noise ns01 at -5dB SNR is enhanced using Kalman, IMCRA, Kalman-

IMCRA cascaded and IMCRA-Kalman cascaded algorithms. The Kalman and IMCRA enhanced signals are shown in Figure 4 (a) and (b) respectively. The cascaded enhanced signals of Kalman-IMCRA and IMCRA-Kalman are shown in Figure 5 (a) and (b) respectively.

Table 1. List of Pure Speech Sentences Used

S. No.	Speaker	Sentence
sp01	Male-X	The birch canoe slid on the smooth planks
sp02	Male-Y	We find joy in the simplest things
sp03	Female-X	The friendly gang left the drug store
sp04	Female-Y	Let us all join as we sing the last chorus

Table 2. Types of Noise Signals Used

S. No.	Type of Noise
ns01	Multitalker babble noise
ns02	Railway platform train arrival
ns03	Car inside with window closed
ns04	Exhaust fan noise
ns05	Street noise in running auto rickshaw

The performance of the four algorithms used in this paper is measured using three widely used speech quality measures [24], Global SNR (SNR_{Glo}), Segmental SNR (SNR_{Seg}) and Perceptual Evaluation of Speech Quality (PESQ) which are defined by:

$$SNR_{Glo}\ (dB) = 10log_{10}\left(\frac{\sum_n |X(n)|^2}{\sum_n |X(n) - \hat{X}(n)|^2}\right) \tag{17}$$

$$SNR_{Seg}\ (dB) = \frac{10}{L}\sum_{l\in L} \frac{\sum_k |X(k,l)|^2}{\sum_k |X(k,l) - \hat{X}(k,l)|^2} \tag{18}$$

$$PESQ = a_0 - a_1 \cdot D_{ind} - a_2 \cdot A_{ind} \tag{19}$$

Where L is the number of frames, $a_0 = 4.5$, $a_1 = 0.1$, and $a_2 = 0.0309$, D_{ind} is the average disturbance and A_{ind} is the average asymmetrical disturbance [25]. For measuring the segmental SNR the frames with SNR values of less than -10dB and greater than 35dB are discarded.

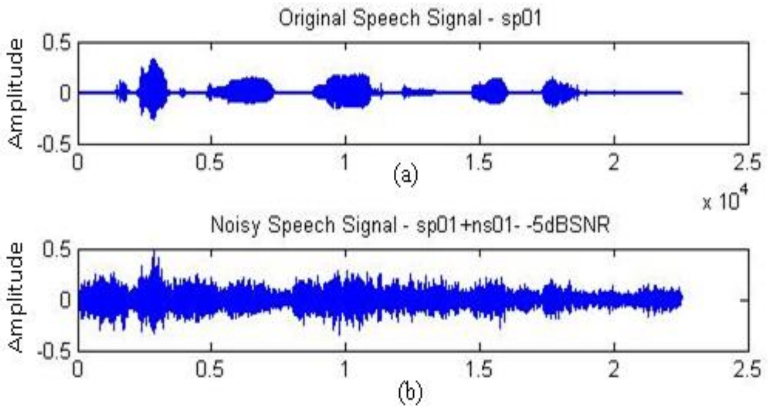


Figure 3. Signal Representation, (a) Pure Speech Signal of Sentence sp01, (b) Speech, sp01, Mixed with Multitalker Babble Noise ns01at -5dB SNR

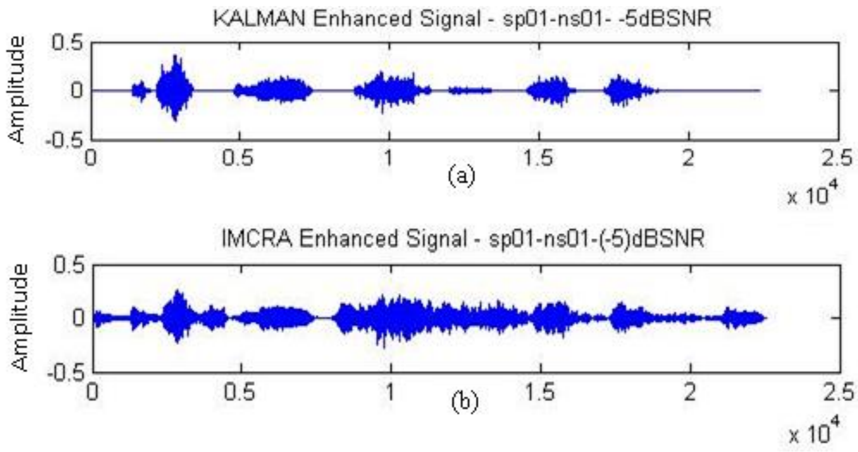


Figure 4. Enhanced Signal Representation of Sentence sp01 Mixed with Noise ns01 at -5dB SNR (a) Kalman Enhanced Signal, (b) IMCRA Enhanced Signal

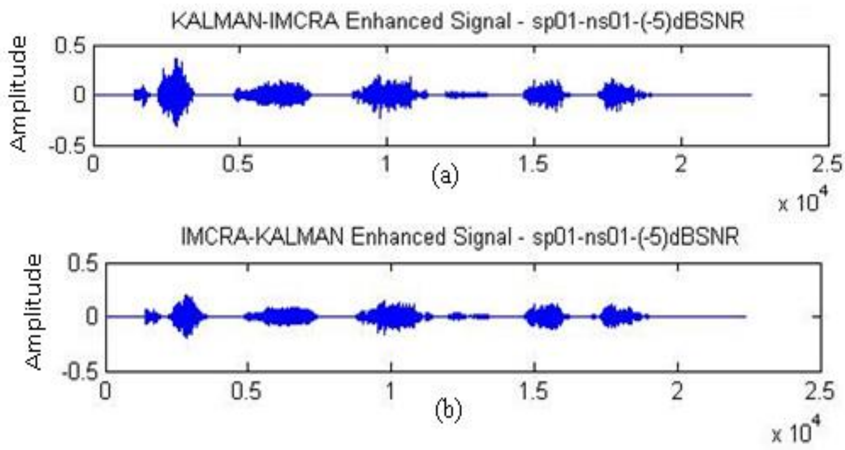


Figure 5. Cascaded Enhanced Signal Representation of Sentence Sp01 Mixed with Noise ns01 at -5dB SNR, (a) IMCRA-Kalman Enhanced Signal, (b) Kalman-IMCRA Enhanced Signal

6. Results and Conclusion

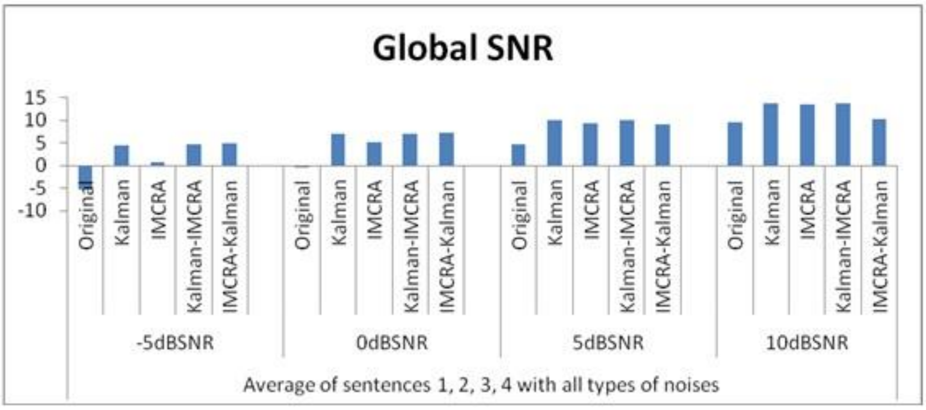
Table 3 shows the average of each performance measures of four speech signals sp01 to sp04 degraded with five different noises ns01 to ns05 at -5dB, 0dB, 5dB and 10dB SNR_{Glo} values, processed using four different algorithms. The table clearly indicates that SNR_{Glo} and SNR_{Seg} parameters of the enhanced speech have been improved using cascaded algorithms. For low SNR conditions of -5dB and 0dB, IMCRA-Kalman performs better whereas for higher SNR conditions of 5dB and 10dB, Kalman-IMCRA performs better. PESQ is better for IMCRA-Kalman algorithm under all types of SNR conditions except for -5dB SNR for which it is marginally less than Kalman algorithm.

Table 3. Average SNR_{Glo}, SNR_{Seg} and PESQ Measures of Four Speech Signals Degraded with Five Types of Noises ns01 to ns05

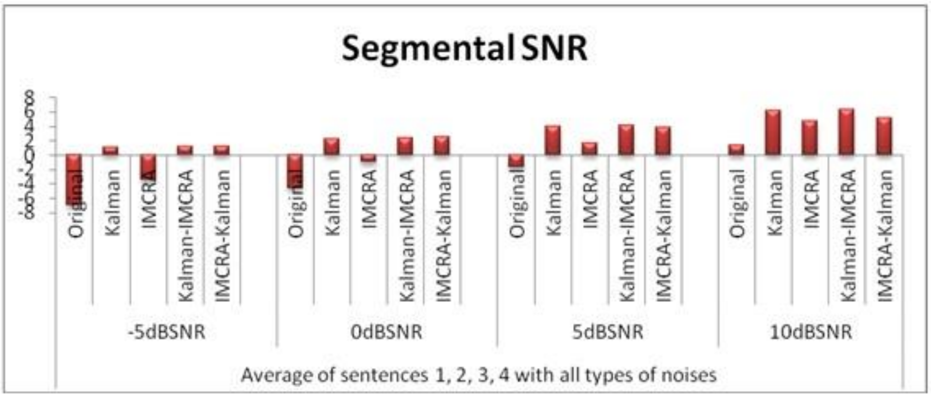
Parameter	Original	Kalman	IMCRA	Kalman-IMCRA	IMCRA-Kalman
-5dB SNR					
SNR_{Glo}	-5.359	4.503	0.794	4.561	4.943
SNR_{Seg}	-6.893	1.098	-3.374	1.180	1.280
PESQ	1.478	2.024	1.593	1.948	1.990
0dB SNR					
SNR_{Glo}	-0.358	6.994	5.078	7.061	7.151
SNR_{Seg}	-4.568	2.293	-1.010	2.373	2.507
PESQ	1.797	2.215	1.986	2.129	2.289
5dB SNR					
SNR_{Glo}	4.640	10.049	9.376	10.132	9.020
SNR_{Seg}	-1.744	4.003	1.717	4.090	3.865
PESQ	2.064	2.466	2.342	2.384	2.597
10dB SNR					
SNR_{Glo}	9.640	13.658	13.608	13.759	10.252
SNR_{Seg}	1.387	6.237	4.707	6.332	5.114
PESQ	2.374	2.759	2.689	2.696	2.871

Figure 6 shows graphical representation of average Global SNR, Segmental SNR and PESQ of four sentences degraded with all five types of noises processed using Kalman and IMCRA alone and cascaded algorithms. Figure 7 (a) represents the spectrograms of pure speech signal sp01 and Figure 7 (b) represents the noisy speech signal sp01 mixed with noise ns01 at -5dB SNR. The spectrograms of Kalman enhanced, IMCRA enhanced, Kalman-IMCRA cascaded enhanced and IMCRA-Kalman cascaded enhanced signals of speech sp01 mixed with noise ns01 at -5dB SNR are represented in Figure 8 (a) to (d) respectively.

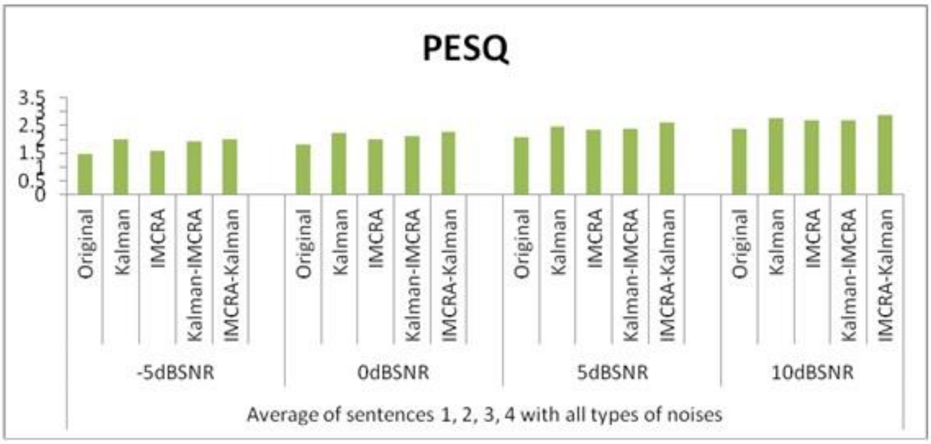
From the comparative analysis of objective parameters as given in Table 3 and the comparative view of spectrograms as shown in Figure 7 and 8, it is concluded that degraded speech enhanced using cascaded algorithms provide better speech quality and intelligence than individual frequency domain algorithm the IMCRA and time domain algorithm the Kalman filter under all types of noise environments. Out of the two cascaded algorithms, the IMCRA-Kalman algorithm is superior.



(a)

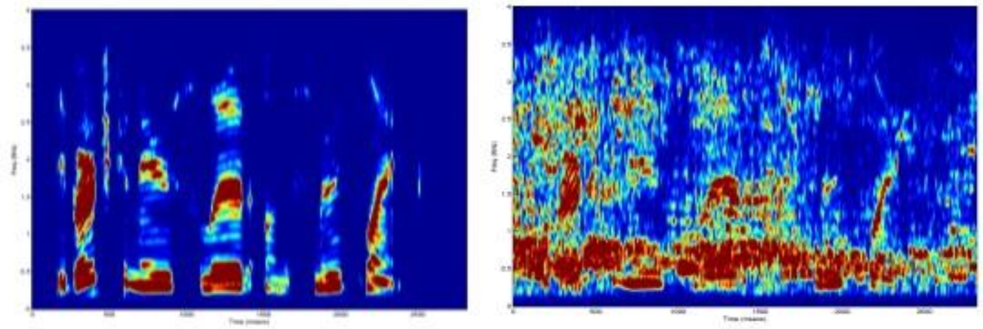


(b)



(c)

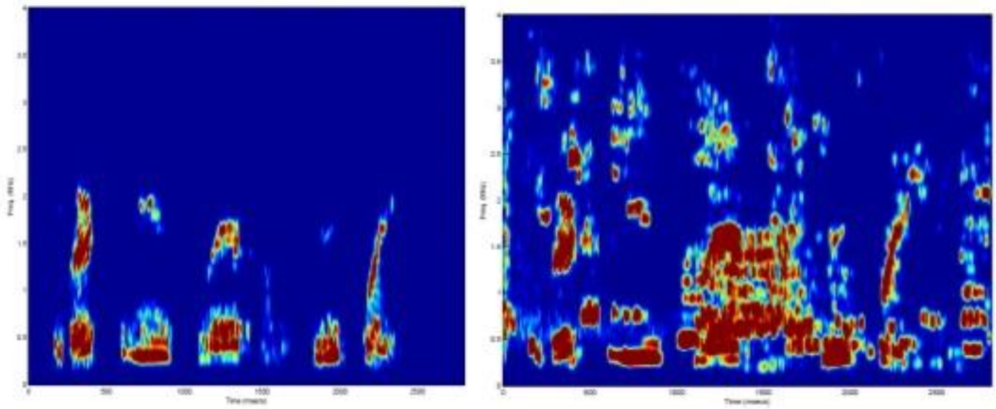
Figure 6. Graphical Representation of Comparative Performance of Different Algorithms, Taking Average of all the Four Sentences Mixed with all the Five Types of Noises Separately, (a) Global SNR (SNR_{Glo}), (b) Segmental SNR (SNR_{Seg}), (c) Perceptual Evaluation of Speech Quality (PESQ)



(a) Pure speech signal sp01

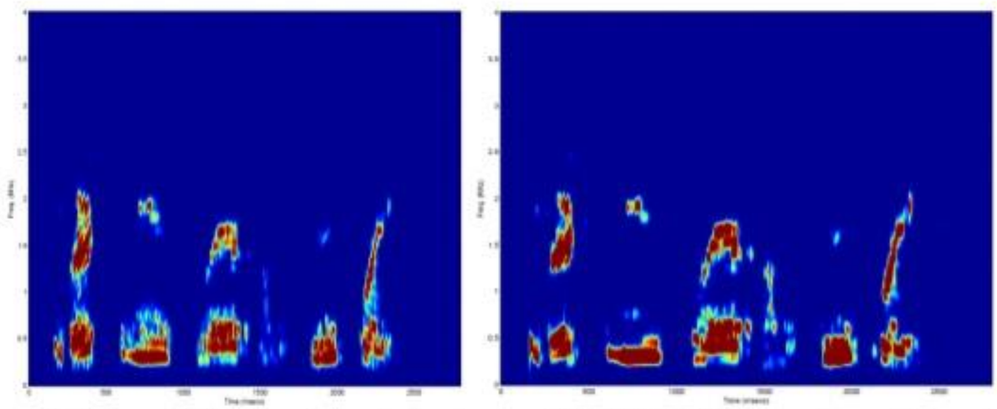
(b) Noisy signal sp01 mixed with noise ns01 at (-5)dB SNR

Figure 7. Spectrogram Representation of Unprocessed Speech Signal (a) Pure Speech Signal sp01, (b) Noisy Speech Signal sp01 Mixed with Noise ns01 at -5dB SNR



(a) Kalman enhanced signal

(b) IMCRA enhanced signal



(c) Kalman-IMCRA enhanced signal

(d) IMCRA-Kalman enhanced signal

Figure 8. Spectrogram Representation of Processed Speech Signal sp01 Mixed with Noise ns01 at -5dB SNR (a) Kalman Enhanced, (b) IMCRA Enhanced, (c) Kalman-IMCRA Enhanced (d), IMCRA-Kalman Enhanced

References

[1] Yariv Ephraim and Israel Cohen. "Recent Advancements in Speech Enhancement", **(2004)**, Accessed May 20, 2014, https://pdfs.semanticscholar.org /8ae7/8c10b671e8e029bb25f702d2701fc 5c228e5.pdf.

[2] Philipos C. Loizou and Gibak Kim, "Reasons why Current Speech-Enhancement Algorithms do not Improve Speech Intelligibility and Suggested Solutions," IEEE Transactions on Audio, Speech, and Language Processing, vol. 19, no.1, **(2011)**, pp. 47-56.

[3] Boll S. F., "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE trans. Acoust., Speech, Signal Processing, vol. 27, **(1979)**, pp. 113-120.

[4] Berouti M., Schwartz R. and Makhoul J., "Enhancement of Speech Corrupted by Acoustic Noise," IEEE Proc. ICASSP, **(1979)**, pp. 208-211.

[5] Ephraim Y. and Malah D., "Speech Enhancement using a Minimum-Mean Square Error Short Time Spectral Amplitiude Estimator," IEEE Trams. Acoust., Speech Signal Processing, ASSP-32, no. 6, December **(1984)**, pp. 1109-1121.

[6] Ephraim, Y. and Malah, D., "Speech Enhancement using a Minimum-Mean Square Error Log Spectral Amplitiude Estimator," IEEE Trams. Acoust., Speech Signal Processing, ASSP-33, April (1985), pp. 443-445.

[7] Wolfe, P. J. and Godsill, S. J., "Efficient Alternatives to the Ephraim and Malah Suppression Rule for Audio Signal Enhancement," Eurasip Journal on Applied Signal Processing, vol. 10, **(2003)**, pp. 1043-1051.

[8] Israel Cohen, "Noise Spectrum Estimation in Adverse Environments: Improved Minima Controlled Recursive Averaging," IEEE Transactions on Speech and Audio Processing, vol. 11, issue 5, **(2003)**, pp. 466-475.

[9] Ronald H. Fraziert, Siamak Samsamtt, Louis D. Braida and Alan V. Oppenheim, "Enhancement of Speech by Adaptive Filtering," Proc. IEEE Int. Conf. Acoustic, Speech and Signal Processing, Kobe, Japan, **(1976)**, pp. 251-253.

[10] Lim, J. S., Oppenheim A. V. and Braida, L. D., "Evaluation of an Adaptive Comb Filtering Method for Enhancing Speech Degraded by White Noise Addition," IEEE Trans. on Acoust., Speech Signal Processing, vol. 26, **(1978)**, pp. 354-358.

[11] Lim, J. S. and Oppenheim, A. V., "All Pole Modeling of Degraded Speech," IEEE Trans. on Acoust., Speech Signal Processing, vol. 26, **(1978)**, pp. 197-210.

[12] Whipple, K. K. and Basu, A., "A Speech Enhancement Method Based on Kalman Filtering," Proc. ICASSP, **(1987)**, pp. 177-180.

[13] Sorqvist, P., Handel, P. and Ottersten, B., "Kalman Filtering for Low Distortion Speech Enhancement in Mobile Communication," ICASSP, **(1997)**, no. 2, pp. 1219-1222.

[14] Goh, Z., Tan, K. C. and Tan, B. T. G., "Kalman Filtering Speech Enhancement Method on a Voiced-Unvoiced Speech Model," IEEE Trans. on Speech and Audio Processing, vol. 7, **(1999)**, pp. 510-524.

[15] Wen-Rong Wu, Po-Chen Chen, Hwai-Tsu Chang and Chun-Hung Kuo, "Frame-based Subband Kalman Filtering for Speech Enhancement," Proceedings of Fourth International Conference on Signal Processing, **(1998)**, vol. 1, pp. 682-685.

[16] Kybic, B. J., "Kalman Filtering and Speech Enhancement," Diploma dissertation, Ecole Polytechnique De Lausanne, **(1998)**.

[17] Popescu, D. and Zeljkovic, I., "Kalman Filtering of Colored Noise for Speech Enhancement," Proceedings ICASSP, **(1998)**, vol. 2, pp. 997-1000.

[18] I. Cohen and B. Berdigo, "Spectral Enhancement by Tracking Speech Presence Probability in Subbands," Proc. IEEE Workshop on Hands Free Speech Communication, HSC2001, Kypto, Japan, April **(2001)**, pp. 95-98.

[19] I. Cohen and B. Berdigo, "Speech Enhancement for Non-stationary Noise Environments," Signal Processing, vol. 81, no. 11, November **(2001)**, pp. 2403-2418.

[20] L. R. Rabiner and R. W. Schafer, "Introduction to Digital Speech Processing," Foundations and Trends in Signal Processing, **(2007)**, vol. 1, nos. 1–2, pp. 1-194.

[21] R. J. McAulay and M. L. Malpass, "Speech Enhancement Using a Soft Decision Noise Suppression Filter," IEEE Trans. Acoustics, speech and signal Processing, vol ASSP-28, no. 2, April **(1980)**, pp. 137-145.

[22] O. Cappe, "Elimination of the Musical Noise Phenomenon with the Ephram and Malah Noise Suppressor," IEEE Trans. Speech and Audio Processing, vol. 2, no. 2, April **(1994)**, pp. 345-349.

[23] John S. Garofolo, Lori F. Lamel, William M. Fisher, Jonathan G. Fiscus, David S. Pallett and Nancy L. Dahlgren, "Getting started With the DARPA TIMIT CD-ROM: An Acoustic Phonetic Continuous Speech Database," National Institute of Standards and Technology (NIST), Gaithersburg, Maryland, **(1993)**.

[24] S. Quackenbush, T. Berwell and M. Clements, "Objective Measures of Speech Quality," Englewood Cliffs:NJ Prentice-Hall, **(1988)**.

[25] ITU-T P.862, Perceptual Evaluation of Speech Quality (PESQ), and Objective Method for End-to- End Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs. ITU-T

Authors

Harish Chander, Born in 1968, has graduated in Electronics and Communication Engineering from Institution of Engineers, India, in 1994. He has completed his Masters in Digital Systems from MNNIT, Allahabad, India, in 2002. He has served the Indian Air Force as electronics & telecommunication engineer for 20 years. For the last 10 years he has been serving various academic organizations at different positions and presently he is working as Controller of Examinations with Institution of Electronics & Telecommunication Engineers, India. He is pursuing his PhD from I K Gujral Punjab Technical University, Jallandhar, India. His areas of research are signal processing, speech processing and wireless communication.

Balwinder Singh, has graduated in Electronics and Communication Engineering from National Institute of Technology, Jallandhar, India, in 2002. He has completed his Masters in Microelectronics from Punjab University, Chandigarh, India, in 2004 and PhD from Guru Nanak Dev University, Amritsar, India, in 2014. He is presently working as Senior Engineer & Coordinator ACS Division, Centre for Development of Advanced Computing (C-DAC), Mohali, India. His Research Interest are Low power VLSI Design and Testing, Digital IP cores and analog modules, Sensor and MEMS design and modeling, FPGA based embedded systems and Image Processing for Embedded systems.

Ravinder Khanna, has graduated in Electrical Engineering from Indian Institute of Technology (IIT), Dehli in 1970 and has completed his Masters and Ph.D in Electronics and Communication Engineering from the same Institute in 1981 and 1990 respectively. He worked as an Electronics Engineer in Indian Defense Forces for 24 Years where he was involved in teaching, research and project management of some of the high tech weapon systems. Since 1996 he has full time Switched to academics. He has worked in many premiere technical institutes in India and abroad. Currently he is Professor and Dean Research with Maharishi Markandeshwar University, Sadopur, Haryana, India. He is active in the general area of Computer Networks, Image Processing and Natural Language Processing.