ISSN 0735-2727, Radioelectronics and Communications Systems, 2016, Vol. 59, No. 11, pp. 502–509. © Allerton Press, Inc., 2016.
Original Russian Text © A. Prodeus V.S. Didkovskyi, 2016, published in Izvestiya Vysshikh Uchebnykh Zavedenii, Radioelektronika, 2016, Vol. 59, No. 11, pp. 37–46.

Objective Estimation of the Quality of Radical Noise Suppression Algorithms

A. Prodeus* and V. S. Didkovskyi**

National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, Ukraine
**ORCID: 0000-0001-7640-0850, e-mail: aprodeus@gmail.com*
***ORCID: 0000-0002-0807-822X, e-mail: vs.didkovskyi@aae.kpi.ua*

Received in final form March 4, 2016

Abstract—There are compared six noise suppression algorithms with application of objective factors of the speech signal quality, and also with application of through quality factor of the system of automated speech recognition in form of speech recognition accuracy. It is shown that radical noise suppression algorithms are worse than traditional noise suppression algorithms by both restored speech quality and speech recognition accuracy due to essential signal distortion.

DOI: 10.3103/S0735272716110042

INTRODUCTION

One of the way of provision of acceptable speech quality in communication systems and in devices for compensation of the lack of hearing and also for increase of robustness of the system of automate speech recognition (ASR) in case of presence of the noise interferences is application of the noise suppression algorithms [1–4].

Following algorithms can be considered as traditional noise suppression ones: SpecSub (Spectral Subtraction) algorithm, Wiener filtering algorithm, filtering with minimization of mean square error of short-time amplitude spectrum MMSE (minimal mean square error) and filtering with minimization of the mean square error of the logarithm of estimation of short-time amplitude spectrum LogMMSE (logarithmic minimal mean square error) [1–4]. In 2004–2006 there are proposed two new noise suppression algorithms: algorithm of two-step noise suppression—TSNR (Wiener two-step noise reduction) and algorithm of noise suppression with harmonics regeneration Wiener-HRNR (Wiener harmonic regeneration noise reduction) [5, 6]. These algorithms provide enough low level of the rest noise that allows to classify them as radical noise suppression algorithms.

The authors' estimation results of Wiener-TSNR and Wiener-HRNR algorithms quality are positive. In [5] it is shown that applying Wiener-TSNR algorithm in ASR systems in conditions of moderate data consistency in learning and testing modes we achieve 9% decrease of recognition errors of "change" type and 22% decrease of recognition errors of "insert" type. In [6] there are represented the results of new algorithms quality estimation using such objective measures of the speech quality as "cepstral distance" and "segment signal to noise ratio", there are represented the results of subjective estimation of the speech quality using comparative scale of mean expert evaluation CMOS (comparative mean opinion score) and it is shown that application of Wiener-HRNR algorithm allows to achieve more qualitive speech sounding in compare to Wiener-TSNR algorithm.

But hearing the signals restored by means of Wiener-TSNR and Wiener-HRNR algorithms notices the speech signal distortion. We should note that science community which is usually attentive to promising innovation does not respond to appearance of Wiener-TSNR and Wiener-HRNR algorithms practically. Therefore, the purpose of the paper is to carry out additional verification of Wiener-TSNR and Wiener-HRNR algorithms efficiency using more wide amount of objective factors in compare to [5, 6]. Another purpose of the paper is to find the way of weakening of the negative influence of Wiener-TSNR and Wiener-HRNR algorithms on the speech signals.

REFERENCES

1. J. Benesty, J. Chen, Y. Huang, I. Cohen, *Noise Reduction in Speech Processing* (Springer-Verlag, Berlin–Heidelberg, 2009).
2. Y. Ephraim, D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process*. **32**, No. 6, 1109 (Dec. 1984), DOI: 10.1109/TASSP.1984.1164453.
3. Y. Ephraim, D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.* **33**, No. 2, 443 (Apr. 1985), DOI: 10.1109/TASSP.1985.1164550.
4. M. Brooks, *VOICEBOX: Speech Processing Toolbox for MATLAB*, Imperial College London, Electrical Engineering Department, http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html.
5. C. Plapous, C. Marro, L. Mauuary, P. Scalart, "A two-step noise reduction technique," *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, ICASSP, 17–21 May 2004 (IEEE, 2004), Vol. 1, pp. 289–292, DOI: 10.1109/ICASSP.2004.1325979.
6. C. Plapous, C. Marro, P. Scalart, "Improved signal-to-noise ratio estimation for speech enhancement," *IEEE Trans. Audio, Speech, Language Process.* **14**, No. 6, 2098 (Nov. 2006), DOI: 10.1109/TASL.2006.872621.
7. Y. Hu, Philipos C. Loizou, "Evaluation of objective quality measures for speech enhancement," *IEEE Trans. Audio, Speech, Language Process.* **16**, No. 1, 229 (Jan. 2008), DOI: 10.1109/TASL. 2007.911054.
8. Jianfen Ma, Yi Hu, Philipos C. Loizou, "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions," *J. Acoust. Soc. Am.* **125**, No. 5, 3387 (May 2009), DOI: 10.1121/1.3097493.
9. S. Young, G. Evermann, M. Gales, et al., *The HTK Book* (University Engineering Department, Cambridge, 2009).
10. D. Ellis, PLP and RASTA (and MFCC, and inversion) in Matlab, http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/.
11. Recommendation P.862 (2001) Amendment 2 (11/05), http://www.itu.int/rec/T-REC-P.862-200511-I!Amd2/en.
12. A. Prodeus, "Calculation of Speech Recognition Quality Factor WB-PESQ in MATLAB," *Proc. of 14th Int. Theoretical and Practical Conf. on New Network Technologies in Ukraine*, 17–19 Sept. 2012, Partenit (Science issue Vesnik UNIIS, Kyiv, 2012), pp. 70–76.