JOURNAL OF PURE AND APPLIED MICROBIOLOGY, April 2013. Vol. **7**(Spl. Edn.), p. 263-268

Bioradar Non-air Conducted Speech Enhancement based on Adaptive Wavelet Packet Entropy

Sheng Li, Huijun Xue, Guohua Lu, Ying Tian, Teng Ma, Yang Zhang, Teng Jiao, Jianqi Wang* and Xijing Jing

School of Biomedical Engineering, Fourth Military Medical University, Xi'an, Shaanxi - 710 032, China.

(Received: 03 March 2013; accepted: 14 April 2013)

A novel bioradar is developed in our lab to obtain speech signal from human beings. Due to the special attribute of the millimeter wave bioradar, this speech detecting method may provide some exciting possibilities of wide applications. However, the resulting speech is of less intelligible and poor audibility since the present of the combined and colored additive noise. This paper, therefore, investigates the problem of the bioradar speech enhancement by corporates the wavelet packet entropy based voice/unvoiced radar speech adaptive detection method in a wavelet packet adaptation speech enhancement process. The results from both acoustic and listening evaluations suggest that the background noise can be reduced efficiently while the distortion of bioradar speech remains acceptable, suggesting that the proposed algorithm achieved a better performance of noise reduction over other traditional speech enhancement algorithms.

Key words: Bioradar, speech, wavelet packet, entropy.

Bioradar is an electric system which can be used to detect and evaluate the movements of living things, mainly human beings. Radar waves can be emitted by the system and the reflected signal was modulated by the movements of the body surface, such as breathing and heartbeats, even internal organs (Xiao *et al*., 2007). Bioradar has been used for many years to solve many medical and engineering problems (Staderini, 2002), among them are: sleep medicine (somnology), functional diagnostics, bedside monitoring, disaster medicine (A. Lazaro and Villarino, 2010; Byrne *et al*., 2010), antiterrorist operations, space medicine, pharmacology and zoo-psychology (Lazaro *et al*., 2009).[1]

Speech is also important information for human beings. However, there are few previous studies majoring on the speech detection by using bioradar. Other researches with respect to the bioradar speech concentrated on the non-acoustic sensors (Holzrichter *et al*., 1998; Jiao *et al*., 2010) and the measurement of speech articulator motions, such as vocal tract measurements and glottal excitation (Hu and Raj, 2005), but not on the bioradar speech itself. Therefore, there is a need to explore this new speech acquisition method (as well as the corresponding speech enhancement algorithm) to extend the existing speech acquisition method.

Although bioradar offers an exciting method for speech acquisition, the radar speech itself has several serious shortcomings, including artificial quality, reduced intelligibility, and poor audibility. This is not only because of some combined harmonics of the electromagnetic wave and electro-circuit noises are present in the detected speech, but also because of channel noise and ambient noise (Bellomo *et al*., 2010). Therefore, speech enhancement is a challenging topic for radar speech research.

* To whom all correspondence should be addressed.
Tel: +86-029-84774843;
E-mail: sheng@mail.xjtu.edu.cn

Previous studies have reported some radar speech enhancement methods. Li *et al.* (Li *et al.*, 2008a) proposed a multiband spectral subtraction approach that takes into account the fact that colored noise affects the speech spectrum differently at various frequencies. Although the speech quality was improved by this algorithm, it suffered from an annoying artifact called "musical noise" (Liu *et al.*, 2006), which is caused by narrowband tonal components appearing somewhat periodically in each frame and occurring at random frequencies in voice or silence regions. They also explored other methods focused on masking the musical noise using psychoacoustic models (Li *et al.*, 2008b). Results obtained by using these algorithms show that there is a need for further improvement in the radar speech enhancement algorithm.

The wavelet transforms (WT), which can be easily obtained by filtering a signal with multi-resolution filter banks (Guo and Wen, 2011), has been applied to various research areas, including signal and image denoising, compression, detection, and pattern recognition (Polivka *et al.*, 2011). Recently, WT have been applied in denoising signals on the basis of the threshold of the wavelet coefficients, where the wavelet threshold (shrinking) introduced by Donoho, D.L. *et al.* (Donoho and Johnstone, 1994) is a simple but powerful denoising technique based on the threshold of the wavelet coefficients. Previous studies have also reported the application of wavelet shrinking for speech enhancement (DL, 1995); however, it is not possible to separate the signal from noises by a simple threshold because applying a uniform threshold to all wavelet coefficients would remove some speech components while suppressing additional noise, especially for the colored noise corrupted signal and some deteriorated speech conditions.

Entropy is defined as a measure of uncertainty of information in a statistical description of a system (Gray, 1990), and the spectral entropy is a measure of how concentrated or widespread the Fourier power spectrum of a signal is. In this study, a time frequency description of bioradar speech, as described by the wavelet packet coefficient, is used to calculate the entropy, which forms the wavelet packet spectral entropy. By its very definition, wavelet packet entropy is considerably sensitive both to the time frequency distribution and to the uncertainty of information, therefore, this novel tool may have very useful characteristics with regard to speech section detection.

Therefore, the purpose of this investigation is to improve bioradar non-air conducted speech technology, especially in electronic environments and very low SNR condition (SNR < 10 dB). The purposes of this study are as follows: (1) to introduce this novel bioradar speech acquisition method; (2) to enhance the radar speech quality by using wavelet packet entropy and wavelet packet denoising method; (3) to evaluate the quality of enhanced radar speech in comparison to speech enhanced by the traditional spectral subtract algorithm.

METHOD

Bioradar speech acquisition system

The schematic diagram of this bioradar non-air conducted speech acquisition system is shown in Figure 1. A phase-locked oscillator generates a very stable millimeter wave (MMW) at 8 mm wavelength (34.5 GHz) with an output power of 100 mW. This output is fed into both the transmitting circuit and the receiving circuit. In the transmitting circuit, the MMW is up-converted to 35.5 GHz by mixed with a 1 GHz crystal oscillator, this wave is fed through a power attenuator before reaching the transmitting antenna frequency. For the receiving circuit, the reflected wave is amplified by a low-noise amplifier after received by the receiving antenna. The transmitting and receiving antennas are both parabolic antennas with a diameter of 300 mm, the estimated beam width is 9º×9º, and the maximum antenna gain is 38.5 dB at 35.5 GHz. The amplified wave is down converted with the 34.5 GHz phase-locked oscillator frequency, and then mixed with 1 GHz crystal oscillator frequency after amplified by an intermediate frequency amplifier. The mixer output provides the speech signal from the body, which is amplified by a signal processor and is then passed through an A/D converter before reaching a computer for further processing. All the signals were sampled at a frequency of 10000 Hz. As shown in Figure 1, the advantage of this kind of radar component layout is that it employs two-step indirect-conversion transceiver, so that to mitigate the severe DC offset problem and the associated 1/f noise at baseband, that occurs normally in the direct-conversion receivers.

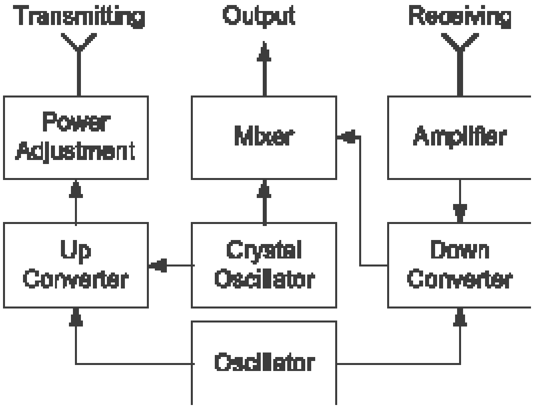

Fig. 1. Schematic diagram of the bioradar speech-detection system

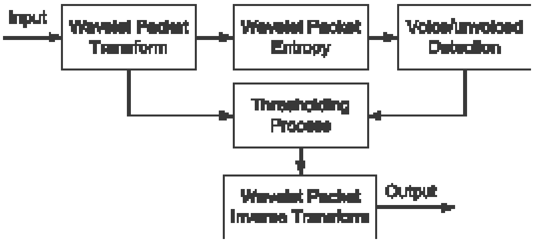

Fig. 2. The proposed adaptive wavelet packet entropy speech enhancement algorithm

The wavelet packet analysis (WPA), which is based on the wavelet transform, can offer a large range of possibilities for signal analysis (Ghanbari *et al*., 2006). If $y(n)$, the noisy speech, consists of the clean speech signal $s(n)$ and the uncorrelated additive noise signal $d(n)$, then:

$$y(n) = s(n) + d(n) \quad ...(1)$$

For a given level, the wavelet packets transform (WPT) decomposes the noisy signal $y(n)$ into 2^i subbands, with the corresponding wavelet coefficient sets as:

$$w_{j,m}^{i} = WP\{s(n), i\} \quad n=1,\ldots N. \quad ...(2)$$

where $w_{j,m}^{i}$ denotes the *m* th coefficient of the *j* th subband for the *i* th level, and m=1,…., N/2^i, k=1,…..2^i. In this study, *i* is set to 6.

The enhanced speech is synthesized with the inverse transformation of the processed wavelet packet coefficients:

$$\hat{s}(n) = WP^{-1}\{\hat{w}_{j,m}^{i}, i\} \quad ...(3)$$

where $\hat{s}(n)$ is the enhanced radar speech, and $\hat{w}_{j,m}^{i}$ is the updated wavelet packet coefficient which is calculated by the algorithm stated below. The schematic of the proposed algorithm is shown in Figure 2.

Wavelet packet entropy

The subband wavelet packet entropy is defined in terms of the relative wavelet energy of the wavelet coefficients (Blanco *et al*., 1998). The energy for each subband *j* and level *i* can be calculated as:

$$E_j^i = \sum_m |w_{j,m}^i|^2 \quad ...(4)$$

The total energy of the wavelet packet coefficients will then be given by:

$$E_{total}^i = \sum_j |w_j^i|^2 \quad ...(5)$$

and the probability distribution for each level can be defined as:

$$p_j^i = \frac{E_j^i}{E_{total}^i} \quad ...(6)$$

Since, following the definition of entropy given by Shannon (1948) (Shannon, 1948), the subband wavelet packet entropy is defined by using the probability distribution associated with scale level *i* (for further details see (Blanco *et al*., 1998) and (Rosso *et al*., 2001)), we have:

$$H(i) = -\sum_j p_j^i \log p_j^i \quad ...(7)$$

Two adaptive wavelet packet entropy thresholds are selected to detect the onset and offset of radar speech. The speech onset threshold is T_s and the offset threshold is T_n-T_s. is defined by adding a fixed value to a past mean wavelet packet entropy value T_m-T_m is calculated over the previous *t* ms (five frames). The speech offset (noise) threshold is calculated by adding another fixed value to . When H(i) (in Eq. 7) exceeds T_s, speech onset is detected and speech offset is detected when H(i) drops below T_m. Therefore, the wavelet packet entropy thresholds can be dynamically adjusted, in this study for bioradar speech, E_s and E_n are set at the constant values of 1.7 and 1.3, respectively.

RESULTS AND DISCUSSIONS

In order to analyze the time frequency distribution of the enhanced speech, in particular the structure of its residual noise, the speech spectrogram is used in this study as it can give more accurate information about residual noise and speech distortion than the corresponding time waveforms. Also, in order to evaluate the performance of this speech enhancement algorithm, other traditional enhancement algorithms are performed for comparison purpose. They are spectral subtraction (Boll, 1979) and noise estimation algorithm (Rangachari and Loizou, 2006). Figure 3 (a) is the spectrogram of the original radar speech. It can be seen from the figure that the original radar speech signal was permeated with noise, which are electromagnetic noise and circuit noise produced by radar system, as stated before. Two other speech enhancement algorithms were also performed for comparison purpose, they are: the traditional spectral subtraction algorithm (Boll, 1979) and noise estimation algorithm (Rangachari and Loizou, 2006). Figure 3 (b) shows the results of bioradar speech enhancement by using spectral subtraction. It can be seen from the figure that the component of noise was reduced efficiently, but there was a lot of noise energy exists in speech signal. It can be seen from the figure that the high frequency component of the enhanced speech was even stronger than the original one, suggest that the traditional spectral subtraction algorithm can even “create” some noise, which is the “music noise”. Figure 3 (c) shows the results of radar

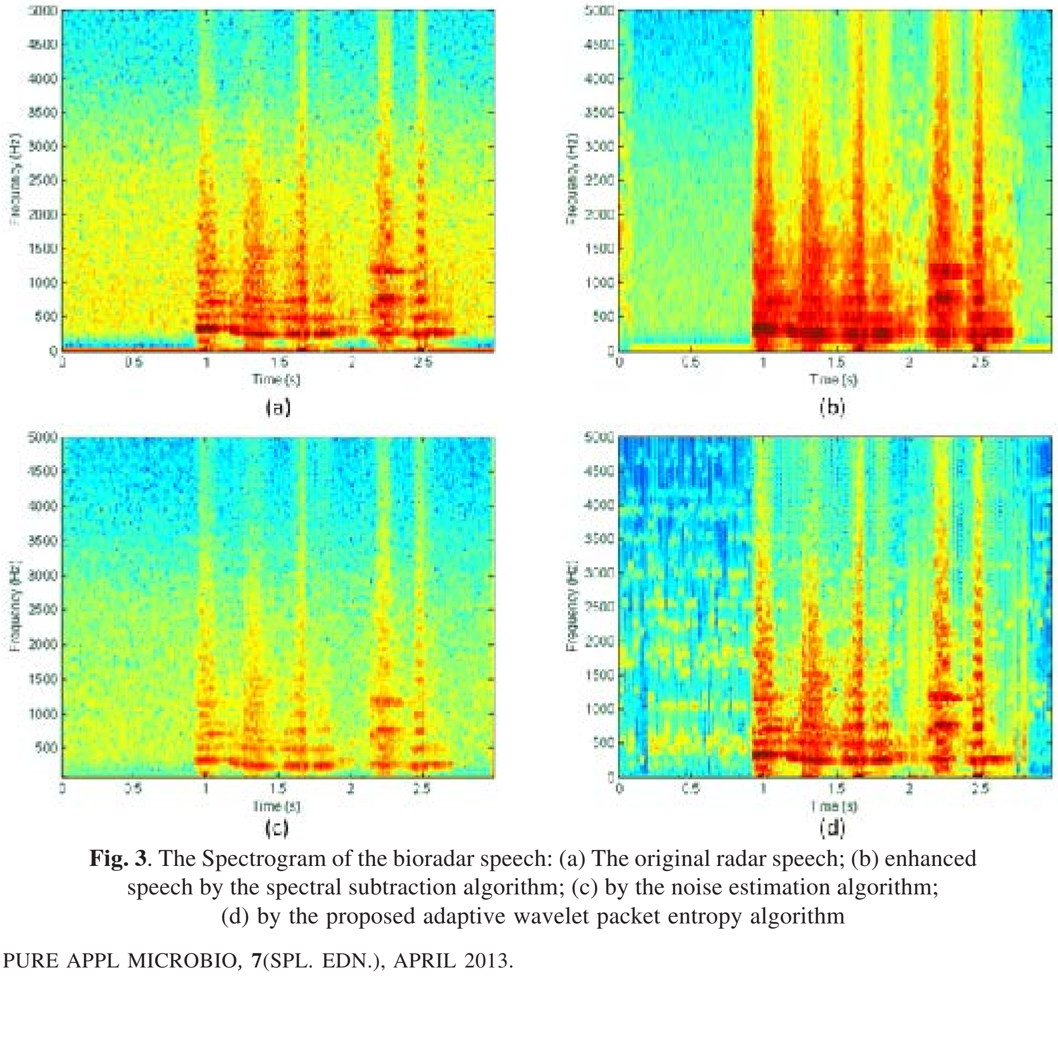

Fig. 3. The Spectrogram of the bioradar speech: (a) The original radar speech; (b) enhanced speech by the spectral subtraction algorithm; (c) by the noise estimation algorithm; (d) by the proposed adaptive wavelet packet entropy algorithm

speech enhancement by using noise estimation algorithm. Compare to original speech signal, it can be seen from the figure that the algorithm can reduce the radar speech noises, however, there is still too much remnant noise in the enhanced speech, especially in the frequency section in which the noise is concentrated, suggesting that the noise reduction is not satisfactory. Figure 3 (d) shows the results of radar speech enhancement by using proposed adaptive wavelet packet entropy denoising algorithm. Compare to two methods stated before, noise component was almost removed, and the speech signal was reserved. Also, there was no new noise component produced, especially in non-speech section. These results suggest that the proposed algorithm achieves a better reduction of the whole-frequency noise than traditional spectral subtraction methods. Informal listening tests also indicated that the speech enhanced with the proposed auditory masking algorithm is more pleasant, the residual noise is better reduced, and with minimal, if any, speech distortion.

The proposed algorithm also displays strong flexibility to adapt to complicated speech environments, because the wavelet packet entropy threshold can be adequately adjusted in the time domain in speech frequency band to fit different or complex speech environments. This makes it possible to get better speech quality via speech enhancement under some rigorous speech environments.

As a single channel speech enhancement method, the algorithm proposed in this paper can be applied for the enhancement of nonacoustic sensor speech using available electronics. For example, a bioradar wave conducted speech enhancing system, into which this algorithm is embedded, can be developed. With the help of digital signal processing (DSP) technology, the speech enhancement function can be realized with a microprocessor and implanted into a radar-telephone, radar-microphone, or other electronic equipment. Different enhancement algorithms, suitable for different noise conditions, can be selected by a switch. With the development of efficient enhancement methods, the quality of nonacoustic sensor speech will be vastly improved and will provide better perception.

CONCLUSION

A novel non-air conducted speech acquisition method is introduced in this study by means of bioradar. Because of the special features of the bioradar, this speech detecting method can provide some exciting possibilities for a wide range of applications. In order to increase the quality of the radar speech, this study also proposed an adaptive wavelet packet entropy speech enhancement algorithm. Comparing to other traditional speech enhancement method, the proposed wavelet packet entropy algorithm can remove noise efficiently, as well as prevent speech from quality deterioration.

ACKNOWLEDGEMENTS

This work was supported by the State Key Program of National Natural Science of China (Grant No. 60927003), the National Natural Science Foundation of China (Grant No. 61271102), and the Natural Science Foundation of Shannxi Province, China (Grant No. 2009K08-39). We also want to thank the participants from the E.N.T. Department, the Xi Jing Hospital, the Fourth Military Medical University, for helping with data acquisition and analysis.

REFERENCES

1. Brook J, Barry O., The Source of Abnormal Returns from Strategic Alliance Announcements. *Pac. Basin Financ. J.*, 2005; **13**(2): 145-61.
2. Bellomo L, Pioch S, Saillard M, Spano E., Time reversal experiments in the microwave range: description of the radar and results. Progress In *Electromagnetics Research*. 2010; **104**: 427-448.
3. Blanco S, Figliola A, Quiroga R. Rosso Q, O A, Serrano, E., Time-frequency analysis of electroencephalogram series. III. Wavelet packets and information cost function. *Phys. Rev*. 1998; **57**: 932 - 940.
4. Boll, S F., Suppression of acoustic noise in speech using spectral subtraction. IEEE Transactions on Acoustics, *Speech and Signal Processing*. 1979; **27**: 113-120.
5. Byrne D, M O'Halloran, M G, Jones E., Data independent radar beamforming algorithms for breast cancer detection. Progress In *Electromagnetics Research*. 2010; **107**: 331-348.
6. DL, D., Denoising by soft thresholding. IEEE

Trans Inform Theory. 1995; **41**: 613-627.
7. Donoho D L, Johnstone, J. M., Ideal spatial adaptation by wavelet shrinkage. *Biometrika.* 1994; **81**: 425-455.
8. Ghanbari Y, Reza M, Karami-Mollaei., A new approach for speech enhancement based on the adaptive thresholding of the wavelet packets. *Speech Communication*. 2006; **48**: 927-940.
9. Gray R M., Entropy and Information Theory (Springer, New York), 1990.
10. Guo B, Wen, G., Periodic Time-Varying Noise in Current-Commutating Cmos Mixers. Progress *In Electromagnetics Research*. 2011; **117**: 283-298.
11. Holzrichter J F, Burnett G C, Ng L C., Speech articulator measurements using low power EM-wave sensors. *J. Acoust. Soc. Am*. 1998; **103**: 622-625.
12. Hu R, Raj B., A robust voice activity detector using an acoustic Doppler radar. IEEE Workshop on Automatic Speech Recognition and Understanding. 2005; **27**: 319 - 324.
13. Jiao M, Lu G, Jing X, Li S., Li Y, Wang J., A Novel Radar Sensor for the Non-Contact Detection of Speech Signals. *Sensors.* 2010; **10**: 4622-4633.
14. Lazaro A, D G, Villarino R., Analysis of vital signs monitoring using an ir-UWB radar. Progress In Electromagnetics Research. 2010; **100**: 265-284.
15. Lazaro A. Girbau D, Villarino R., Wavelet-based breast tumor localization technique using a UWB radar. Progress In Electromagnetics Research. 2009; **98**: 75-95.
16. Li S, Wang J, Niu M, Jing X., The enhancement of millimeter wave conduct speech based on perceptual weighting. Progress In Electromagnetics Research. 2008a; **9**: 199-214.
17. Li S, Wang J, Niu M, Liu T, XiJing J., Millimeter wave conduct speech enhancement based on auditory masking properties. *Microwave and Optical Technology Letters.* 2008b; **50**: 2109 - 2114.
18. Liu H J, Zhao Q, Wan M X, Wang S P., Enhancement of electrolarynx speech based on auditory masking. *IEEE Trans. Biomed. Eng.* 2006; **53**: 865-874.
19. Polivka J, Fiala P, Machac J., Microwave noise field behaves like white light. *Progress In Electromagnetics Research.* 2011; **111**: 311-330.
20. Rangachari S, Loizou P C., A noise-estimation algorithm for highly non-stationary environments. *Speech Communication*. 2006; 220-231.
21. Rosso O A, Blanco S, Yordanova J, Kolev V, Figliola A, SchÜrmann M, Basar E., Wavelet entropy : a new tool for analysis of short duration brain electrical signals. *Journal of neuroscience methods*. 2001; **105**: 65-75.
22. Shannon C E., A mathematical theory of communication. *Bell System Technical Journal.* 1948; **27**: 379-423 and 623-656.
23. Staderini E M., UWB Radars in Medicine. IEEE *Aerospace and Electronic Systems Magazine.* 2002; **1**: 13-18.
25. Xiao Y, Li C, Lin J C., A portable noncontact heartbeat and respiration monitoring system using 5-GHz radar. *IEEE Sens. J*. 2007; **7**: 1042-1043.