The Use of LPC and Wavelet Transform for Influenza Disease Modeling

Khaled Daqrouq * and Mohammed Ajour

Department of Electrical and Computer Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia; majour@kau.edu.sa

* Correspondence: haleddaq@gmail.com; Tel.: +966-5-66-980400; Fax: +966-5-695-2686

Abstract: In this paper, we investigated the modeling of the pathological features of the influenza disease on the human speech. The presented work is novel research based on a real database and a new combination of previously used methods, discrete wavelet transform (DWT) and linear prediction coding (LPC). Three verification system experiments, Normal/Influenza, Smokers/Influenza, and Normal/Smokers, were studied. For testing the proposed pathological system, several classification scores were calculated for the recorded database, from which we can see that the proposed method achieved very high scores, particularly for the Normal with Influenza verification system. The performance of the proposed system was also compared with other published recognition systems. The experiments of these schemes show that the proposed method is superior.

Keywords: LPC; wavelet transform; speech; influenza disease modeling

1. Introduction

As humans, we interact with each other in many ways, such as talking, writing, by signs, and so on, the most commonly used of which is talking. Human speech is a sound; the sound is a mechanical wave of pressure and displacement. The human speech is formed with the help of many organs, such as the lungs, larynx, throat, nose, mouth, sinuses, and other secondary organs. The air exhaled through these organs, which comes from the lungs through the vocal folds inside the larynx, provides energy for the vocal folds to produce the sound. The other organs then help to formulate the sound into speech [1–4].

Pathology is the science of causes and effects of diseases. The development of this field depends on the development of other related sciences, and technology will produce new diagnosing methods and more accurate disease diagnoses [5–10]. The human vocal tract, as with any human organ, is affected by many factors, including aging, diseases, or congenital flaw. These factors affect their functionality with different levels of effects, and somehow change the sound these organs produce. Every pathological (disease) factor may have specific changes in the human sound, modeling these changes will enable us to detect some diseases by the sound of the patient [11–14].

Many types of research have been applied in voice pathology diagnostics, such as characterization methods for the detection of multiple voice disorders: neurological, functional, and laryngeal diseases. In this research, many methods are tested, such as a Mel-frequency cepstral coefficient (MFCC) and a Gaussian classifier for modeling the pathology by speech. The study came up with a range of accuracies between 81% and 98% for different pathologies. Pathological voice recognition for vocal fold disease was investigated by Gaussian mixture model [14], where a range of accuracies between 83% and 98% were achieved.

Researchers have used different ideas to achieve better performance of feature extraction methods. In Tohan et al. [15], acoustic parameters using multi corpus optimization with a neural network

based support vector machine (SVM) classifier was proposed. Linear prediction coding (LPC), LPCC, MFCC, and PLP were used by Palo et al. [16] and Selvasraj et al. [17], and modified by Pao et al. [18] and Sato et al. [19]. Feature extraction methods based on wavelet transform were proposed in many works [20–22]. Wang et al. [23] investigated the use of a Fourier-based method for better feature extraction quality.

In this paper, wavelet transform is used in conjunction with LPC for speech pathology modeling and recognition. For classification, probabilistic neural network (PNN) is used. The verification task is conducted for investigation of the method quality.

2. The Method

In this paper, we investigated the pathological modeling of the influenza disease. Novel research based on a real database and a new combination of previously used methods, wavelet transform and LPC, was proposed. Wavelet transform is a very attractive approach for many researchers in the pattern recognition field and signal processing [13]. The main reason behind this is the possibility to decompose the signal into sub-signals of different bandpasses of frequency. The reason behind using the discrete wavelet transform (DWT) method is the advantage of decomposing the signal into different pass bands of frequency for each sub-signal. Thus, there is no frequency overlapping. It is highly effective in achieving a better feature extraction possibility. By utilizing the LPC for each sub-signal to be linked into one vector containing all LPC coefficients of these sub-signals, we get a feature extraction vector. The level of DWT was chosen empirically for the best performance. The determination of DWT of level 5 was adopted as a tradeoff between better performance and less dimensionality of the feature vector [24].

In this paper, we used the following steps:

- Decomposing the speech signal into DWT of level 5:

$$F = [D_1,\ D_2,\ D_3, D_4,\ D_5, A_5] \tag{1}$$

where,

$$
\begin{aligned}
D_1 &= d_{11}, d_{12}, \dots, d_{1N/2} \\
D_2 &= d_{21}, d_{22}, \dots, d_{2N/4} \\
D_4 &= d_{31}, d_{32}, \dots, d_{3N/8} \\
D_4 &= d_{41}, d_{42}, \dots, d_{4N/16} \\
D_5 &= d_{51}, d_{52}, \dots, d_{5N/32} \\
A_5 &= d_{51}, d_{52}, \dots, d_{5N/32}
\end{aligned}
\tag{2}
$$

where D_1 is the detail DWT sub-signal of level l that represented the high frequency part of the original signal $x(t)$, and was calculated as a convolution of the signal with the basis $\psi_{1,k}(t)$ generated from the mother wavelet function. For the DWT of level $j = 1$:

$$D_1 = x(t) * \psi_{1,k}(t) \tag{3}$$

where,

$$\psi_{1,k}(t) = 2^{-\frac{1}{2}}\psi(2^{-1}t - k) \tag{4}$$

and $k = 1, 2, \dots, N/2$; N is the number of the data record; and A_J is the approximated DWT sub-signal of the last level L that represented the low frequency part of the original signal $x(t)$, and was calculated as a convolution of the signal with the basis $\varphi_{J,k}(t)$ generated from the father wavelet function [13]:

$$A_J = x(t) * \varphi_{J,k}(t) \tag{5}$$

For the DWT of level $j = J$:

$$\psi_{J,k}(t) = 2^{-\frac{J}{2}} \varphi(2^{-J}t - k) \tag{6}$$

- Calculating LPC of each sub-signal of 30 coefficients as follows:

$$lpc_l = \text{LPC}(d_l) \tag{7}$$

where *lpc* calculated the coefficients of a forward linear predictor by minimizing the prediction error in the least squares sense. The *lpc* was applied in the speech coding and filter design. The Matlab function LPC found the coefficients of a pth-order linear predictor. LPC predicted the present value based on the past samples as follows: $-a(2) \times (n - 1) - a(3) \times (n - 2) - \ldots -a(p + 1) \times (n - p)$, where *p* is the order and a is [1 $a(2)$ ... $a(p + 1)$]. The determination of the number of LPC coefficients (exactly as in the DWT level) was adopted as a tradeoff between better performance and less dimensionality of the feature vector [13].

- The feature extraction vector is

$$FV = [lpc_1, lpc_2, lpc_3, lpc_4, lpc_5, lpc_{a5}] \tag{8}$$

The choice of the wavelet mother function type is crucial and is dependent on the intended application. In our study, we investigated many wavelet functions and their corresponding recognition rate. Based on our investigation, we chose to use the wavelet function type Daubechies five (also known as db5) on the basis that it yielded the best recognition rate. Therefore, it was considered for rest of our investigation.

For classification, a probabilistic neural network was proposed. Ganchev proposed PNN with Mel-frequency cepstral coefficients for classification [25]. Although, there does exist numerous modified versions of the original PNN that are more economical and offer better performance. For simplicity of the exposition, we implemented the original PNN for the classification task. The proposed algorithm was of the following construction:

$$\text{Net} = \text{PNN (X, R, SP)} \tag{9}$$

where X is the matrix of 180X8 input speech feature vectors (pattern) of 180 LPC coefficients; R is the target class vector, R = [1, 2, 3, . . . , 8]; and SP is the spread. The number 180 came from the idea that we calculated six DWT sub-signals: D1, D2, D3, D4, D5, and A5. Then, we calculated 30 LPC coefficients for each sub-signal. After that, they were linked together in one feature vector of 180 coefficients in length. The number, "eight", came from the training signal, "four", that was used for the training of each type at each verification system. Four signals of each type were recorded separately for training tasks along with the recorded testing database. For example, for classification of normal/influenza, four of each type were trained. If the system recognized the testing signal as one of the first four that belonged to normal cases, it was detected as the normal type, but if it recognized the testing signal as one of the second four, it was detected as the influenza type.

3. Results and Discussion

For testing of the proposed system, a suitable data set was recorded. For this purpose, 101 persons were involved in the dataset recording. Each participant had to record his speech with the same microphone in a quiet university room. The sampling frequency was 8 KHz. Twenty-two of the total people engaged in the recording had influenza disease in different levels, 48 smoking persons, and 41 normal persons, without influenza disease (normal). All participants were males aged between 20 and 27 years old. Four additional signals for each type were recorded for different participants. The four speech signals were utilized for the training process; therefore, all other recorded signals were

used as testing signals. One Arabic sentence was recorded by all participants, which was composed of four words, and means, "peace, mercy, and blessings of God".

The authors modeled all signals, the training and testing signals, by wavelet and LPC, and then trained the four additional speech signals for each type by PNN in Matlab. The models of testing signals were compared with the trained models by the simulation process offered in Matlab using the "sim" function. The verification system was adopted in the proposed study to investigate the accuracy of the presented approach, which used two types of the speech signal in the system; for example, Influenza/Normal. Influenza was determined as a positive class, and normal was determined as a negative class. Moreover, four parameters were crucial for the analysis of the results. These parameters were true positive (TP), false positive (FP), true negative (TN), and false negative (FN). The parameters mentioned were used to calculate several important specifications that are mentioned below.

Figure 1 presents the speech signal for two normal people, two people with influenza disease, and two smokers. The feature vectors by the proposed method and spectrograms calculated for the feature vectors also are illustrated. By observing the feature vectors, as well as the spectrogram, we can easy recognize the normal case and the influenza case, as well as the smoker case, using these methods.
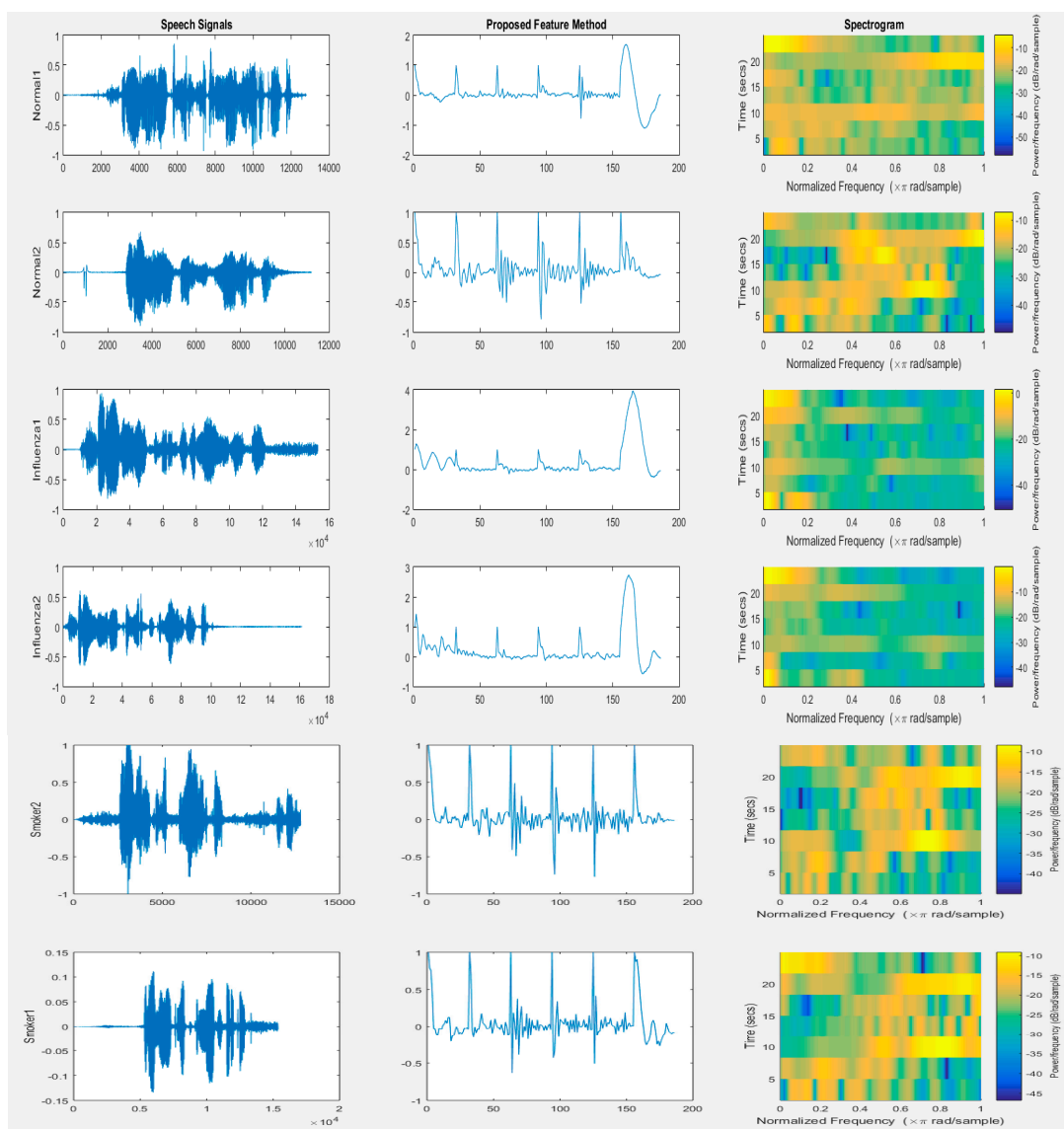

Figure 1. The speech signal for two normal people, two people with influenza disease, and two smokers; the feature vectors by the proposed method; and the spectrograms calculated for the feature vectors.

The recognition rate was calculated as an average of three measures:

$$\text{(1)} \quad \text{Sensitivity (S)} = \frac{\text{TP}}{\text{TP+FN}} \times 100\%$$

$$\text{(2)} \quad \text{specificity (P)} = \frac{\text{TN}}{\text{TN+FP}} \times 100\% \tag{10}$$

$$\text{(3)} \quad \text{positive predictivity (PP)} = \frac{\text{TP}}{\text{TP+FP}} \times 100$$

In Table 1, three experiments are studied, namely the Normal/Influenza, Smokers/Influenza, and Normal/Smokers verification systems. The recognition rate was calculated as an average of the three measures S, P, and PP for each verification system.

Table 1. Results of recognition rate for proposed method. TN—true negative; TP—true positive; FN—false negative; FN—false negative.

Negative/Positive	TN/TP	FN/FP	Recognition Rate
Normal/Influenza	37/22	0/4	92.60%
Smokers/Influenza	36/22	0/13	78.78%
Normal/Smokers	36/38	10/5	85.11%

Where TN is a true negative results number when the system identifies the testing signal as the normal case. TP is a true positive results number when the system confirms the tested signal as the influenza case, and in the case of the tested signal, really belongs to the influenza case. In other words, the criteria for determining the TP, TN, FN, and FP was based on categorization of the result of the testing samples into these four categories, such that if the tested sample was originally an influenza type for the Normal/Influenza verification system and the classifier recognized it as an influenza type, it was categorized as a true positive. However, if the classifier mistakenly recognized it as normal, it was categorized as a false negative.

If the tested sample was originally a normal type and the classifier recognized it as a normal type, it was categorized as a true negative. However, if the classifier mistakenly recognized it as an influenza type, it was categorized as a false positive. Then, the TP, TN, FN, and FP were the number of these cases for each category.

Other measures were also calculated for testing the method, such as the following measures:

- False rejection error (FRR), which is based on the false negative (FN) and true positive (TP) parameters as follows:

$$\text{FRR} = \frac{\text{FN}}{(\text{TP} + \text{FN})} \tag{11}$$

- False acceptance error (FAR), which is based on false positive (FP) and true negative (TN) parameters as follows:

$$\text{FAR} = \frac{\text{FP}}{(\text{FP} + \text{TN})} \tag{12}$$

- Accuracy (AC)

$$\text{AC} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{FN} + \text{FP} + \text{TN})} \tag{13}$$

- Efficiency (EF)

$$\text{EF} = 100 * \left[1 - \frac{\text{FN}}{(\text{TP} + \text{FN})}\right] \tag{14}$$

For testing the proposed pathological system, all of the above classification scores were calculated for the recorded database for different Negative/Positive systems. The FAR, FRR, S, P, PP, AC, and EF scores were calculated (see Table 2), and we can see that the proposed method achieved very high scores, particularly for Normal with Influenza.

Table 2. The results of different recognition statistical parameters for proposed method. FAR—false acceptance error; FRR—false rejection error; S—sensitivity; P—specificity; PP—positive predictivity; AC—accuracy; EF—efficiency.

Negative/Positive	FAR	FRR	S	P	PP	AC	EF
Normal/Influenza	9.76	0	100	90.62	84.24	93.65	100
Influenza/Smokers	26.53	0	100	73.47	62.86	81.69	100
Normal/Smokers	12.20	20.83	79.17	87.80	88.37	83.15	79.16

The value based on the recognition rate of the proposed systems was also compared with other published recognition systems, such as power spectrum density (PSD) [12], LPC [13], and MFCC [12]. Table 3 tabulates the comparative results obtained on the recorded database. The experiments of these schemes show that the proposed method is superior.

Table 3. Results of recognition rate for a different method for comparison. LPC—linear prediction coding; MFCC—Mel-frequency cepstral coefficient.

Normal/Influenza	TN/TP	FN/FP	Recognition Rate
PSD	33/18	4/8	77.18%
LPC	33/12	10/8	64.85%
MFCC	36/17	5/5	81.00%
Proposed Method	37/22	0/4	92.60%

4. Conclusions

In this paper, we investigate the pathological modeling of the influenza disease using speech signal via testing over a recorded database. Novel research, based on original samples recorded in university circumstances, was suggested. A new combination of previously used methods, discrete wavelet transform and linear prediction coding (LPC), was proposed. Based on our investigation, we have chosen to use the wavelet function type Daubechies five (also known as db5) on the basis that it yields the best recognition rate. Three verification systems experiments, Normal/Influenza, Smokers/Influenza, and Normal/Smokers were studied. Many statistical scores were utilized for testing, such as sensitivity, specificity, false acceptance rate, false rejection rate, or efficiency. The performance of the proposed system was also compared with other published recognition systems. We can state that the proposed method achieved very high scores, particularly for the Normal with Influenza verification system.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ensign, J.G. Astronomy Picture of the Day. 19 August 2007. Available online: https://apod.nasa.gov/apod/ap070819.html (accessed on 26 June 2015).
2. Olson, H.F. *Music, Physics and Engineering*; Dover: New York, NY, USA, 1967; 249p, ISBN 9780486217697.
3. *The American Heritage Dictionary of the English Language*, 4th ed.; Houghton Mifflin Company: Boston, MA, USA, 2000.
4. How the Voice Works. Available online: http://www.entnet.org/content/how-voice-works (accessed on 3 June 2018).

5. Definition of Path in English. Oxford English Dictionary. OED. 12 October 2013.

6. Robbins, S. *Robbins and Cotran Pathologic Basis of Disease*, 8th ed.; Saunders/Elsevier: Philadelphia, PA, USA, 2010; ISBN 978-1-4160-3121-5.

7. Long, E. *History of Pathology*; Dover: New York, NY, USA, 1965; ISBN 0-486-61342-9.

8. Arcolani, G. *Commentary on the Chapter Nine of the Book of Medicine Dedicated to Mansur—Commentaria in Nonum Librum Rasis ad Regem Almansorem*; World Digital Library; Luca-Antonio Giunta: Venice, Italy, 1542. (In Latin)

9. King, L. *Transformations in American Medicine: From Benjamin Rush to William Osler*; Johns Hopkins UP: Baltimore, MD, USA, 1991; ISBN 0-8018-4057-0.

10. Machevsky, A.; Wick, M.R. Evidence-based Medicine, Medical Decision Analysis, and Pathology. *Hum. Pathol.* **2004**, *35*, 1179–1188. [CrossRef]

11. Rothstein, W.G. Pathology: The Evolution of a Specialty in American Medicine. *Med. Care* **1979**, *17*, 975–988. [CrossRef] [PubMed]

12. Orozco-Arroyave, J.R.; Belalcazar-Bolaños, E.A.; Arias-Londoño, J.D.; Vargas-Bonilla, J.F.; Skodda, S.; Rusz, J.; Daqrouq, K.; Hönig, F.; Nöth, E. Characterization Methods for the Detection of Multiple Voice Disorders: Neurological, Functional, and Laryngeal Diseases. *IEEE J. Biomed. Health Inform.* **2015**, 1820–1828. [CrossRef] [PubMed]

13. Daqrouq, K.; Al Azzawi, K.Y. Average framing linear prediction coding with wavelet transform for text independent speaker identification system. *Comput. Electr. Eng.* **2012**, *38*, 1467–1479. [CrossRef]

14. Pravena, D.; Dhivya, S. Pathological voice recognition for vocal fold Disease. *Int. J. Comput. Appl.* **2012**. [CrossRef]

15. Tahon, M.; Devillers, L. Towards a small set of robust acoustic features for emotion recognition: Challenges. *IEEE/ACM Trans. Audio Speech Lang. Pocess.* **2016**, *24*, 16–28. [CrossRef]

16. Palo, H.K.; Mohanty, M.N.; Chandra, M. Use of different features for emotion recognition using MLP network. *Adv. Intell. Syst. Comput.* **2015**, *332*, 7–15.

17. Selvaraj, M.; Bhuvana, R.; Padmaja, S. Human speech emotion recognition. *Int. J. Eng. Technol.* **2016**, *8*, 311–323.

18. Pao, T.-L.; Chen, Y.-T.; Yenand, J.-H.; Lia, W.-Y. Detecting emotions in Mandarin speech. *Comput. Linguist. Chin. Lang. Process.* **2005**, *10*, 347–362.

19. Sato, N.; Obuchi, Y. Emotion recognition using MFCCs. *Inf. Media Technol.* **2007**, *2*, 835–848.

20. Sharma, A.; Shrotriya, M.C.; Farooq, O.; Abbasi, Z.A. Hybrid wavelet based LPC features for Hindi speech recognition. *Int. J. Inf. Commun. Technol.* **2008**, *1*, 373–381. [CrossRef]

21. Wang, K.; An, N. Speech emotion recognition based on wavelet packet coefficient model. In Proceedings of the 9th International Symposium on Chinese Spoken Language Processing (ISCSLP), Singapore, 12–14 September 2014; pp. 478–482.

22. Kumar, P.; Chandra, M. Hybrid of wavelet and MFCC features for speaker verification. In Proceedings of the World Congress on Information and Communication Technologies, Mumbai, India, 11–14 December 2011; pp. 1150–1154.

23. Wang, K.; An, N.; Li, B.N.; Zhang, Y. Speech emotion recognition using Fourier parameters. *IEEE Trans. Affect. Comput.* **2015**, *6*, 69–75. [CrossRef]

24. Daqrouq, K.; Sweidan, H.; Balamesh, A.; Ajour, M. Off-Line Handwritten Signature Recognition by Wavelet Entropy and Neural Network. *Entropy* **2017**, *19*, 252. [CrossRef]

25. Daqrouq, K.; Tutunji, T.A. Speaker identification using vowels features through a combined method of formants, wavelets, and neural network classifiers. *Appl. Soft Comput.* **2015**, *27*, 231–239. [CrossRef]