www.jmss.mui.ac.ir

Are Speech Attractor Models Useful in Diagnosing Vocal Fold Pathologies?

Sir,

Development of non-invasive methods in diagnosing different diseases can lead to improvement of prevention and care programs. The speech is an easily accessible signal, which clearly represents the characteristics of larynx and vocal folds. Therefore, application of some proper machine learning algorithms (e.g., feature extraction and classification methods) on a small part of a recorded speech signal may help in diagnosing vocal fold diseases such as paralysis, edema, nodules, and polyp.[1-3] Generally, transforming the input signal into the set of features is called feature extraction. If these features are accurately extracted, it is expected that the feature set will capture the relevant pathological information of speech signal to predict the diagnosis.

Conventionally, some acoustic features of speech signal like pitch frequency, shimmer (amplitude perturbation), and jitter (pitch perturbation) are used to distinguish between normal and pathological cases.[4,5] On the other hand, there are some experimental evidences that proof the existence of chaotic behavior in speech production system (e.g., turbulence airflow) not considered in the conventional and mentioned feature extraction methods.[6,7] For example, some of recent researches have considered chaotic characteristics of speech signal such as correlation dimension, the largest Lyapunov exponent, approximate entropy, fractal dimension, and Ziv–Lempel complexity.[8,9]

One of the best domains to represent chaotic properties of different biological signals is the phase space domain.[7,10] Takens have introduced delay coordinate embedding theorem to reconstruct a signal in the phase space domain. This theorem shows that a one-dimensional signal (e.g., a recorded speech signal) can be embedded and reconstructed as a set of points in a high dimensional space so-called reconstructed phase space (RPS) topologically equivalent to its original system.[10-12] Often, these points show a trajectory in the RPS, which is called an "attractor." The true dynamic of signals generated by different systems can be exhibited in the RPS. So, the proposed method is based upon modeling the trajectory of pathological signal as it is captured in the RPS.

Based on above mentioned points, we hypothesize that not only modeling of pathological voice as a speech trajectory or speech attractor in the RPS is suitable for detection of vocal fold pathologies, but also their obtained results will be comparable to conventional classification methods. Hence, for each normal or pathological voice such as paralysis, edema, nodules, and polyp, a specific speech attractor model will be constructed using a parametric and probabilistic model e.g., Gaussian Mixture model (GMM) in the RPS.[11,12] It learns the probability distribution of the attractor in the RPS. One of the powerful characteristics of the GMM is its ability to form smooth approximations of attractors. Utilizing the GMM-based attractor models learned for each class of pathological signals, a set of probability scores such as likelihood can be computed for each unknown test signal. Finally, behavior of vocal fold for unknown test signal will be predicted in a non-invasive procedure by comparing the computed probability scores using a naive Bayesian maximum likelihood classifier. Surely, experimental researches are needed to validate our hypothesis.

Yasser Shekofteh, Shahriar Gharibzadeh, Farshad Almasganj

Department of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran

E-mail: y_shekofteh@aut.ac.ir

Submission: 10-07-2012

Accepted: 28-04-2013

REFERENCES

1. Khalilarjmandi M, Pooyan M, Mikaili M, Vali M, Moqarehzadeh A. Identification of voice disorders using long-time features and support vector machine with different feature reduction methods. J Voice 2011;25:275-89.
2. Erfaniansaeedi N, Almasganj F, Torabinejad F. Support vector wavelet adaptation for pathological voice assessment. Comput Biol Med 2011;41:822-8.
3. Khalilarjmandi M, Pooyan M. An optimum algorithm in pathological voice quality assessment using wavelet-packet-based features, linear discriminant analysis and support vector machine. Biomed Signal Process Control 2012;7:3-19.
4. Awan SN, Frenkel ML. Improvements in estimating the harmonics-to-noise ratio of the voice. J Voice 1994;8:255-62.
5. Moran RJ, Reilly RB, Chazal P, Lacy PD. Telephony-based voice pathology assessment using automated speech analysis. IEEE Trans Biomed Eng 2006;53:468-77.
6. Banbrook M, McLaughlin S, Mann I. Speech characterization and synthesis by nonlinear methods. IEEE Trans Speech Audio Process 1999;7:1-17.

7. Kokkinos I, Maragos P. Nonlinear speech analysis using models for chaotic systems. IEEE Trans Speech Audio Process 2005;13: 1098-109.
8. Jiang JJ, Zhang Y. Nonlinear dynamic analysis of speech from pathological subjects. Electronics Lett 2002;38:294-5.
9. Vaziri G, Almasganj F, Behroozmand R. Pathological assessment of patients' speech signals using nonlinear dynamical analysis. Comput Biol Med 2010;40:54-63.
10. Kantz H, Schreiber T. Nonlinear time series analysis. Cambridge, England: Cambridge University Press; 1997.
11. Shekofteh Y, Almasganj F. Feature extraction based on speech attractors in the reconstructed phase space for automatic speech recognition systems. ETRI J 2013;35:100-8.
12. Povinelli RJ, Johnson MT, Lindgren AC, Ye J. Time series classification using Gaussian mixture models of reconstructed phase spaces. IEEE Trans Knowl Data Eng 2004;16:779-83.

How to cite this article: Shekofteh Y, Gharibzadeh S, Farshad Almasganj F. Are speech attractor models useful in diagnosing vocal fold pathologies? J Med Sign Sens 2012;3:185-6

Source of Support: Nil, **Conflict of Interest:** None declared