EURASIP Journal on Advances in Signal Processing

a SpringerOpen Journal

RESEARCH
Open Access

Throat polyp detection based on compressed big data of voice with support vector machine algorithm

Wei Wang[1*], Zhangliang Chen[2], Jiasong Mu[1] and Tingting Han[1]

Abstract

Classification in large-scale data is a key problem in big data domain. The theory of compressive sensing enables the recovery of a sparse signal from a small set of linear, random projections which provides a compressive classification method operating directly on the compressed data without reconstructing for big data. In this paper, we collected the compressed vowel /a:/ and /i:/ voice signals using compressive sensing for throat polyp detection. The throat polyp prediction procedure based on wavelet packet transform and support vector machine intelligent algorithm was deduced. The experiments for throat polyp prediction with the proposed classification algorithm were carried out. The results showed that the correct rate of prediction was stable under different number of samples and different random measurement matrices.

Keywords: Big data; Compressive sensing; Throat polyp detection; Support vector machine

1 Introduction

Throat polyps are small fleshy growths which form on the vocal cords, usually as a result of overuse. They are mainly caused by straining or overusing the voice, for example, public speaking. Professional singers, sports/ fitness coaches, or actors are all prone to developing throat polyps. The most common symptoms of throat polyps are a hoarse or deeper voice or a breathy sounding voice similar to laryngitis.

The traditional methods of throat polyp diagnosis are indirect laryngoscope, video-laryngoscope, and strobo-scope light [1]. These methods need special instrument and depend on the experience of the pathologists. Usually, the patients will feel uncomfortable pain. Due to the fact that voice change of the patients is the most common symptoms of throat polyps, it would be desirable if the throat polyps could be detected based on the patient voices. In [1], Zhong et al. tried to detect throat polyps based on patient voices. Two fuzzy classifiers and a Bayesian classifier were designed for throat polyp

detection based on patient vowel voices /a:/ and /i:/. The experimental results showed that an interval type 2 fuzzy classifier performed the better. In this paper, we will use the compressive sensing and support vector machine (SVM) algorithm to detect the throat polyps with patient vowel voices /a:/ and /i:/ while reducing the burden of voice data collection and storage.

Compressive sensing (or compressed sampling (CS)) theory demonstrated that a high-dimensional signal can be projected into a low-dimensional space with a random measurement matrix when the signal was sparse or compressible which was proposed by Donoho and Candès in 2006 [2,3]. Then, the original signal can be recovered from the low-dimensional information with solving an optimization problem. The provable success of CS for signal reconstruction motivated that the low-dimensional signal contained the main features of the original signal. Thus, the universality of CS theory can be leveraged in the hypothesis testing problem and mitigate the complexity of data computing [4].

The hypothesis testing in compressed domain can not only reduce the pressure of data storage and transmission but also overcome the large amount of data calculation. In [5,6], Budhaditya used the compressed sensor network data for anomaly detection based on spectrum

* Correspondence: wangweivip@tju.edu.cn
[1]College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
Full list of author information is available at the end of the article

Springer

theory method and obtained satisfactory detection results in the light of residual analysis of compressed data. In [7,8], random projection in conjunction with principal component analysis (PCA) was implemented for anomaly detection in compressed domain, and an application of this proposed methodology to detect IP-level volume anomalies in computer network traffic suggested a high relevance to practical problems. In [9], an anomaly detection criterion based on wavelet packet transform and statistic process control theory in compressed domain was used for through wall human detection. The experimental results showed that the proposed algorithm could effectively detect the existence of human being through compressed signals.

Because of the advantage of compressed classification in big data based on compressive sensing comparative with classification in original data [10-15], a throat polyp detection algorithm based on compressive sensing and support vector machine is proposed in this paper. The remainder of this paper is organized as follows: In Section 2, the compressive sensing theory will be introduced. Throat polyp detection procedure based on CS and SVM will be deduced in Section 3. Experimental results of throat polyp detection will be shown in Section 4. Section 5 is the conclusion and discussion.

2 Background on compressive sensing

Compressive sensing states that the signal often contains some type of structure that enables intelligent representation and processing which builds on a core tenet of signal processing and information theory [16].

Suppose that an observer makes measurements of a signal x, it can be expressed as follows:

$$x = \Phi\theta, \tag{1}$$

where $\theta \in R^N$ is the expansion coefficient vector under the orthonormal basis Φ. If θ has only $K \leq N$ nonzero coefficients, we can say that signal x is K-sparse.

The surprising result of CS is that a length-N signal that is K-sparse in some basis can be recovered exactly/approximately from a nonadaptive linear projection of the signal onto a random basis. In matrix notation, it can be described as follows [17-21]:

$$y = \Psi x, \tag{2}$$

where y is an $M \times 1$ column vector and Ψ is an $M \times N$ random matrix. The appeal of CS is that we only need to collect $M = O(K \log(N/K))$ random measurements to recover the signal x by solving the following l_0-norm-constrained optimization problem:

$$\widehat{\theta} = \arg\min\|\theta\|_0 \ \ s. \ t. \ \ y = \Psi\Phi\theta, \tag{3}$$

where the $\|\theta\|_0$ norm counts the number of nonzero components of θ. However, solving Equation 3 was both numerically unstable and NP-complete. Instead of solving the l_0 minimization problem, nonadaptive CS theory seeks to solve the 'closest possible' tractable minimization problem, i.e., the l_1 minimization:

$$\widehat{\theta} = \arg\min\|\theta\|_1 \ \ s. \ t. \ \ y = \Psi\Phi\theta. \tag{4}$$

Although $M < N$, the recovery of the signal x from the measurements y become possible and practical under the additional assumption of signal sparsity or compressibility. The provable success of CS for signal reconstruction can indicate that the collected low-dimensional measurements contained the main features of the original signal. Therefore, it provides us a novel procedure for hypothesis testing of big data which can be carried out in the compressed domain.

3 Throat polyp detection procedure

The most common symptom and the first to typically appear in the throat polyp patients is a general roughness or hoarseness of the voice, which may or may not be accompanied by a sore throat or a full feeling in the throat. In other words, the frequency components of the same voice such as vowel voices will be varied when a person suffers throat polyps. Therefore, we can detect the throat polyps by analyzing the frequency component of the voice signal.

3.1 Acquire the frequency component by WPT

Fourier transform (FT) is the conventional signal frequency spectrum analysis tool which is a global transform and has low-frequency resolution. Due to its shortage in recognizing the tiny change of the frequency spectrum of FT, wavelet packet transform (WPT) has become the widest implement in the field of signal frequency analysis.

WPT is one extension of the wavelet transform (WT) which provides a complete level-by-level decomposition. It can enable the extraction of features from signals which combine stationary and nonstationary characteristics with an arbitrary time-frequency resolution [22].

In this paper, we extract the features of vowel voices /a:/ and /i:/ to detect the throat polyps. According to the principle of WPT, the vowel voice signal $x(t)$ is decomposed into j levels of decomposition, and the node signals are reconstructed as $x_j^i(t)$. Then, it can be expressed as follows:

$$x(t) = \sum_{i=1}^{2^j} x_j^i(t). \tag{5}$$
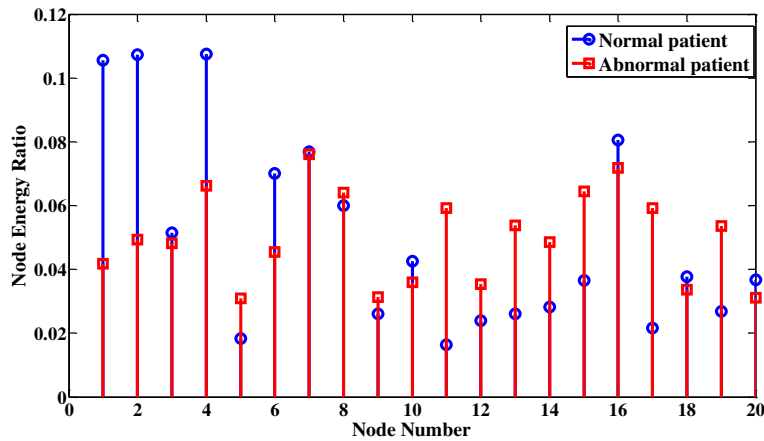
The node signal energy can be defined E_j^i as

Figure 1 Node energy ratio of compressed vowel /a:/ voice signal for a normal and abnormal patient.

$$E_j^i = \int_{-\infty}^{\infty} x_j^i(t)^2 dt = \sum x_j^i(t)^2. \qquad (6)$$

On the basis of WPT, Equation 6 illustrates that the node signal energy E_j^i stores the energy of a specific time-frequency window. In other words, E_j^i indicates the proportion of corresponding frequency component in the original signal. Thus, according to the principle described above, the throat polyp detection can be achieved by investigating the changing trend of E_j^i.

In order to eliminate the influence of volume for throat polyp detection, we define ΔE_j^i as the node signal energy ratio in the total signal energy further:

$$\Delta E_j^i = \frac{E_j^i}{\sum\limits_{i=1}^{m} E_j^i}, \qquad (7)$$

where m denotes the first dominant nodes which contained the main energy of the signal. It can eliminate the noise effect on the tiny energy node and improve the detection accuracy.

3.2 Throat polyp detection procedure with SVM

Support vector machines (SVM) are a popular machine learning method for classification, regression, and other learning tasks. In this method, one maps the data into a higher dimensional input space and one constructs an optimal separating hyper plane in this space [23,24].

Given a training set of N data points (y_i, x_i), $i = 1, 2, 3, \ldots, N$ where $x_i \in R^n$ and $y_i \in \{1, -1\}$. The classifier is constructed as follows. One assumes that

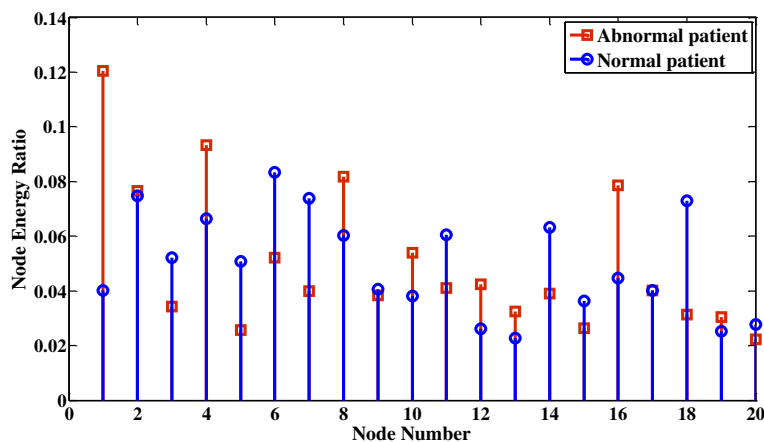$$\omega^T \varphi(x_i) + b \geq 1, \quad if \quad y_i = 1$$

Figure 2 Node energy ratio of compressed vowel /i:/ voice signal for a normal and abnormal patient.
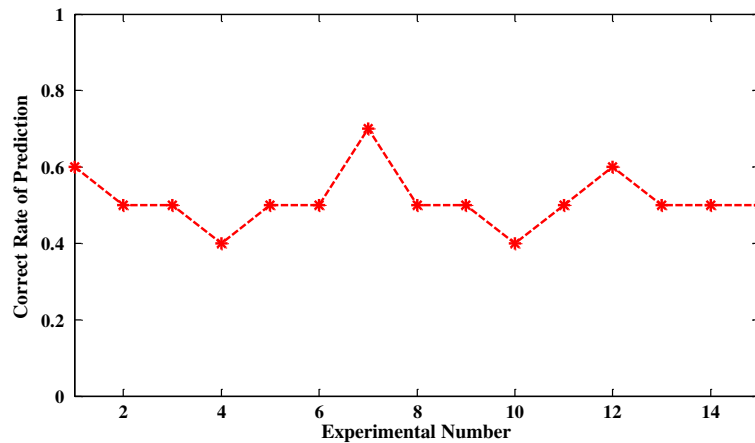
Figure 3 Correct rate of throat polyp prediction under different random measurement matrices.

$$\omega^T \varphi(x_i) + b \le 1, \quad if \quad y_i = -1 \tag{8}$$

which is equivalent to

$$y_i\left[\omega^T \varphi(x_i) + b\right] \ge 1, \quad i = 1, 2, ..., N \tag{9}$$

where ϕ is a nonlinear function which maps the input space into a higher dimensional space. However, the function (9) is not explicitly constructed. In order to obtain the separating hyper plane in the higher dimensional space, variables ξ_i are introduced to solve the following primal optimization problem

$$\min_{\omega, b, \xi} \quad \frac{1}{2}\omega^T \omega + C\sum_{I=1}^{N}\xi_i$$

$$subject \ to \quad y_i\left(\omega^T \varphi(x_i) + b\right) \ge 1 - \xi_i. \tag{10}$$

Through the training data, we can obtain the support vectors and kernel parameters in the model for prediction.

As we know, a continuous speech usually consists of vowels and voiced consonants. The vocal cord does not vibrate when producing voiced consonants which come from the vibration of the lips and teeth. It will bring up interference in the afterward steps because of useless signal collection. Meanwhile, multi-vowel in one speech sample will result in an aliasing in spectrum map of the sample. Therefore, in this paper, we only use the vowels /a:/ and /i:/ to detect the throat polyps of the patients.

In this paper, we acquired the vowel voice signal based on compressive sensing and extract the features constructed by frequency components of compressed signals. Lastly, the SVM method is used to obtain the classification model for throat polyp detection. The procedure is depicted in Algorithm 1.
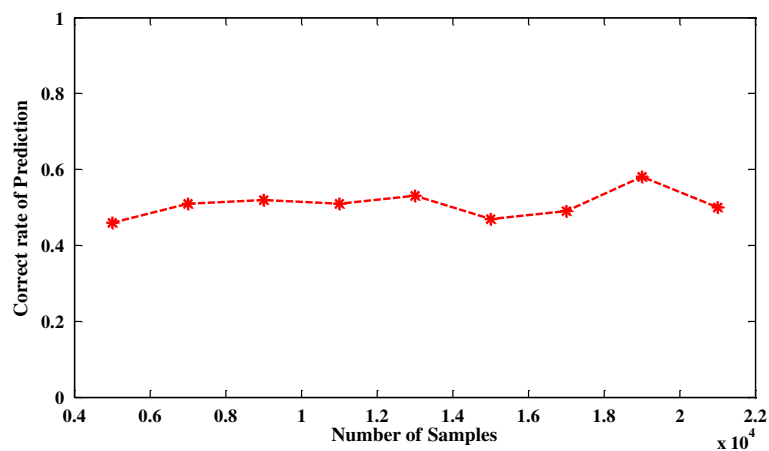

Figure 4 Correct rate of throat polyp prediction under different number of samples.

Algorithm 1 Throat polyp detection algorithm

1. Collect voice signals xa_i of vowel /a:/ and xi_i of vowel /i:/ of a patient, $i = 1, 2, 3, ..., N$, that is, we totally obtain vowel voice signals of N patients.

2. According to the following equations, we gain compressed signals of vowel voice xa_i, xi_i. Actually, we can obtain the compressed signals in the sampling process.

$$ya_i = \Psi xa_i, \quad yi_i = \Psi xi_i, \quad \Psi \in R^{p \times q}$$

3. Perform WPT of compressed signals and extract the features $\{Ea_k, Ei_k\}, k = 1, 2, ..., m$ of each patient based on Equations 5 to 7.

4. Features $\{y, Ea_k, Ei_k\}$ of N patients are the training data for SVM to obtain the classification model, where $y = 1$ if the patient is normal without throat polyps, else $y = 0$.

5. Collect the vowel voice signals of new patient and achieve the features $\{ENa_k, ENi_k\}$ according to steps 1 to 3. Then, detect the existence of throat polyps based on the predict model acquired in step 4.

4 Experimental results and analysis

In the experiments, vowel /a:/ and /i:/ voice signals of 26 patients were collected, among which 13 patients have throat polyps and 13 patients did not have throat polyps. The Gaussian random measurement matrix was used for compressed data obtained. The compressed voice signals were decomposed by eight-layer wavelet packet with 'db10' wavelet, and the first 20 node signals were used to extract the features. The C-SVM program proposed by Dr. Lin was used for setting up the classification model and throat polyp prediction [23].

In the first experiment, we used 7,000 samples in the original vowel /a:/ and /i:/ of 26 patients, respectively, to construct the features. The compression ratio is 50%, and the features of eight normal patients without throat polyps and eight abnormal patients with throat polyps were used for training to establish a classification model. The other features of ten patients were used to test the performance of the proposed algorithm (Algorithm 1).

Figures 1 and 2 showed the node energy ratio of compressed vowel /a:/ and /i:/ signals of a normal patient without throat polyps and an abnormal patient with throat polyps. It can be seen that the frequency components of two kinds of patients are different and the low-frequency component in vowel voice signals changed more obviously when the patient has throat polyps. In other words, the frequency component of patients would vary when he or she suffers throat polyps. Thus, frequency component energy of vowel voice signals could be used as the features for throat polyp detection.

Figure 3 showed the prediction results of throat polyp patients under different random measurement matrices based on the proposed algorithm (Algorithm 1). We can see that the correct rate of prediction is about 50% with small fluctuations. It indicates that the features used for test and training were similar although they were obtained under different measurement matrices. Meanwhile, we repute that the low correct rate of prediction and small fluctuations were caused by the few training samples.

In the second experiment, we used different samples in the original vowel /a:/ and /i:/ of 26 patients, respectively, to construct the features with the same measurement matrix. The compression ratio, the training data, and test data were the same with the first experiment. The results were shown in Figure 4, while the correct rate of prediction has a mean value of ten predictions, respectively, for each number of samples. It can be seen that the correct rate of prediction was about 50% at different number of samples, while the fluctuations also were caused by the few training date and could not construct a high-accuracy prediction model. Meanwhile, the results demonstrated that our classifier is able to detect throat patient with a small number of samples.

5 Conclusions

Big data refers to large, diverse, complex, longitudinal, and distributed data sets. Some core technologies are needed to solve the problem in big data such as classification technology. Compressive sensing theory provided a new approach for big data classification which overwhelmed the limitation of Nyquist sampling theory and could sample and compress data simultaneously.

In this paper, we used the compressive sensing theory to acquire the compressed vowel voice signals for throat

polyp detection. The frequency component energy ratios of compressed data obtained by wavelet packet transform were used as features. Then, the support vector machine intelligent algorithm was used to detect the existence of throat polyps. The experimental results showed that the performance of prediction was stable, but the correct rate of prediction is low, due to the few samples of patient cases.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China. [2]College of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China.

References
1. Z Zhong, Z Chen, Q Liang, S Xiao, Throat polyps detection based on patient voices. Lecture Notes in Electrical Engineering **202**, 531–539 (2012)
2. DL Donoho, Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)
3. E Candès, Compressive sampling, in *Proceedings of International Congress of Mathematicians* (Madrid, 2006), pp. 1433–1452
4. J Haupt, R Castro, R Nowak, G Fudge, A Yeh, Compressive sampling for signal classification. Fortieth Asilomar Conference on Signals, Systems and Computer **10**, 1430–1434 (2006)
5. S Budhaditya, D-S Pham, M Lazarescu, S Venkatesh, Effective Anomaly Detection in Sensor Networks Data Streams, in *2009 Ninth IEEE International Conference on Data Mining* (Miami, 2009), pp. 722–727
6. D-S Pham, S Venkatesh, M Lazarescu, S Budhaditya, Anomaly detection in large-scale data stream networks. Data Mining and Knowledge Discovery (2012). doi:10.1007/s10618-012-0297-3
7. Q Ding, ED Kolaczyk, A compressed PCA subspace method for anomaly detection in high-dimensional data. IEEE Transactions on Information Theory **59**(11), 7419–7433 (2013)
8. DS Pham, B Saha, M Lazarescu, S Venkatesh, *Scalable Network-Wide Anomaly Detection Using Compressed Data* (Curtin University of Technology, Perth, 2009)
9. W Wang, D Lu, X Zhou, B Zhang, J Mu, On-line anomaly detection in big data based on compressive sensing. Lecture Notes in Electrical Engineering **212** (2013)
10. Q Liang, Situation understanding based on heterogeneous sensor networks and human-inspired favor weak fuzzy logic system. IEEE Systems Journal **5**(2), 156–163 (2011)
11. Q Liang, Biologically-inspired target recognition in radar sensor networks. EURASIP Journal on Wireless Communications and Networking **2010**, 523435 (2010)
12. Q Liang, X Cheng, S Samn, NEW: network-enabled electronic warfare for target recognition. IEEE Trans on Aerospace and Electronic Systems **46**(2), 558–568 (2010)
13. Q Liang, X Cheng, Underwater acoustic sensor networks: target size detection and performance analysis. Ad Hoc Networks Journal (Elsevier) **7**(4), 803–808 (2009)
14. Q Liang, X Cheng, KUPS: knowledge-based ubiquitous and persistent sensor networks for threat assessment. IEEE Transactions on Aerospace and Electronic Systems **44**(3), 1060–1069 (2008)
15. Q Liang, Automatic target recognition using waveform diversity in radar sensor networks. Pattern Recognition Letters (Elsevier) **29**(2), 377–381 (2008)
16. MA Davenport, MF Duarte, MB Wakin, JN Laska, D Takhar, K Kelly, R Baraniuk, The smashed filter for compressive classification and target recognition, in *Proceedings of the SPIE Computational Imaging V* (San Jose, 2007), p. 1
17. L Xu, Q Liang, X Cheng, D Chen, Compressive sensing in distributed radar sensor networks using pulse compression waveforms. EURASIP Journal of Wireless Communications and Networking (2013). doi:1186/1687-1499-2013-36
18. L Xu, Q Liang, Zero correlation zone sequence pair sets for MIMO radar. IEEE Trans on Aerospace and Electronic Systems **48**(3), 2100–2113 (2012)
19. L Xu, Q Liang, Orthogonal pulse compression codes for MIMO radar system, in *IEEE Globecom* (Miami, 2010)
20. L Xu, Q Liang, Waveform design and optimization in radar sensor network, in *IEEE Globecom* (Miami, 2010)
21. Q Liang, JM Mendel, Design interval type-2 fuzzy logic systems using SVD-QR method: rule reduction. Int. J. Intell. Syst. **15**(10), 939–957 (2000)
22. W Wang, X Zhou, B Zhang, J Mu, Anomaly detection in big data from UWB radars, Wiley Security and Communication Networks. doi:10.1002/sec.745,2003
23. CC Chang, CJ Lin, LIBSVM: a library for support vector machines. ACT Transactions on Intelligent Systems and Technology **2**(27), 1–27 (2011)
24. JAK Suykens, J Vandewalle, Least squares support vector machine classifiers. Neural Processing Letters **9**, 293–300 (1999)