Primi Isolated Words Spectrogram Classification by Support Vector Machine Based on Immune Genetic Algorithm

Hua-zhen DONG, Wen-lin PAN*, Hua YANG and Mei-jun FU

School of Mathematics & Computer Science, Yunnan Minzu University, Kunming, Yunnan, China, 650500

*Corresponding author

Keywords: Primi isolated words spectrogram, Support Vector Machine (SVM), Immune Genetic Algorithm (IGA), Binary feature.

Abstract. We propose a method for Primi isolated words spectrogram classification by support vector machine based on immune genetic algorithm (SVM-IGA). Firstly, time-frequency spectrograph of Primi isolated words is generated by Short Time Fourier Transform (STFT). Secondly, binary feature is extracted by binarization spectrogram. Thirdly, spectrogram classification is realized by IGA-SVM. The experimental results show that the predictive accuracy rate of Primi isolated words spectrogram classification was 88~91%. Compared with the speech signal classification, the spectrogram classification by SVM-IGA is better.

Introduction

From 2015 onwards, the National Language Committee decided to start the Chinese Language Resource Protection Project. The project collected and recorded minority language corpus by using modern technology, however, a large amount of language annotation is wrong. Therefore, we need to speech recognition for corpus, so that find out the speech of error-tagging. Primi language [1] is an endangered minority language that has no text. In this paper, Primi isolated words as the research object, studied its acoustic characteristics resorting to phonological diagram, and ultimately achieved Primi isolated words classification.

In essence, isolated words classification still belongs to the field of pattern recognition [2], which can be summarized as "Feature Extraction → Isolated Words Pattern Recognition", isolated words pattern recognition is the core issue in Primi isolated words classification. At present, the main algorithms for pattern recognition are Artificial Neural Networks (ANN), k Nearest Neighbor Classifier (KNNC) [3, 4], support Vector Machine (SVM) etc. KNNC and ANN requires a large amount of sample data to train the model. Because ANN is based on the empirical risk minimization, there is a risk of overfitting. However, SVM maps non-separable samples into a high-dimensional linear separable feature space by using kernel functions, and then it transfers high-dimensionality into a quadratic programming problem based on structural risk minimization [5, 6]. Finally, the global optimal solution is obtained by convex optimization. Compared with the speech signal classification, spectrogram classification without considering voiceless consonants and segmental division influence on isolated words. In addition, it is more intuitive to use phonological spectrum study Primi language acoustic characteristics.

Reasonable choices of penalty parameters $c$ and kernel function parameters $g$ of SVM will directly affect classification performance. Recently, a variety of modern intelligent optimization methods have been widely used in parameters optimization process, such as Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Simulated Annealing (SA) and so on. PSO has characters of simple, optimization efficiency and strong robustness when dealing with simple problems, but easily getting into local minimum when dealing with complex problems. The parallel search ability of GA algorithm is excellent, but exists premature phenomenon. SA has excellent serial search ability, but needs sufficient iterations to guarantee global convergence [7-9]. Immune Genetic Algorithm (IGA) is a novel random search and optimization method based on natural
selection and biological genetic mechanism, which integrates the biological immune system’s antigen recognition, antibody diversity, concentration control and other mechanisms by considers the cross between antibodies to find the optimal solution [10]. Therefore, IGA preserves the advantage of global parallel search of GA, and avoids the disadvantage of prematurity phenomenon.

In order to improve the classification accuracy and generalization ability of SVM, this paper fully analyzes the performance of kernel functions, and presents Primi isolated words spectrogram classification using support vector machine optimized by immune genetic algorithm (IGA-SVM).

**Support Vector Machine (SVM)**

With a known training set: $x_i \in R^n, y_i \in \{-1, 1\}$, where $x_i (i=1, \cdots, l)$ is the eigenvector, select the appropriate parameters $C$ and kernel function $K(x_i, y_j)$ for SVM, the optimization model is:

$$
\begin{align}
\max & \sum_{i=1}^{l} a_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} a_i a_j y_i y_j K(x_i, y_j) \\
st & \sum_{i=1}^{l} a_i y_i = 0, 0 \leq a_i \leq C, i=1, 2, \cdots, l 
\end{align}
$$

which is a convex quadratic programming problem. Hence, there exists the optimal solution. Solving the above (1), we get optimal solution: $\alpha^* = (\alpha_1^*, \cdots, \alpha_l^*)^T$, where $T$ represents transpose of matrix. Then take a positive component $\alpha_j^*$ from $a^*$ satisfying $0 < \alpha_j^* < C$, calculate the threshold from $b^* = y_i - \sum_{i=1}^{l} \alpha_i^* y_i K(x_i - x_j)$, and finally decision function is:

$$f(x) = \text{sgn}(\sum_{i=1}^{l} \alpha_i^* y_i K(x, x_i) + b^*)$$

SVM maps non-separable samples into a high-dimensional linear separable feature space by using non-linear transformation. However, non-linear mapping is generally difficult to get, SVM converts nonlinear transformation into inner product by introducing kernel function. The common kernel functions are shown below:

(1) Linear function

$$K(X, Y) = K(X, Y)$$

(2) Polynomial function

$$K(X, Y) = (X \cdot Y + c)^d$$

(3) Gaussian radial basis function

$$K(X, Y) = \exp(-\gamma \|X - Y\|^2)$$

(4) Sigmoid function

$$K(X, Y) = \tanh(\nu(X, Y) - c)$$

**Extracting Binary Feature from Primi Spectrum**

The binary feature is the pixel whose gray value is greater than or equal to the threshold value in the binary image, which is expressed as a specific object, whose gray value is 1; and the pixel whose gray level is less than the threshold value is expressed as the background area, and the gray value is 0. The main steps of extracting Primi isolated words spectrum binary feature as following:
(1) Selecting all the Primi isolated words speech signal from the corpus.
(2) Using energy entropy ratio method for endpoint detection [11].
(3) Using STFT to draw the spectrogram.
(4) Finding the threshold of binarization by the iterative optimal threshold method [12].
(5) Extracting the binary feature.

Comparison of Kernel Functions
For the same test set, the predicted classification accuracy of different kernel functions is shown in Table 1.

| Kernel function     | Accuracy (%) | SVM parameters options |
|---------------------|--------------|------------------------|
| Linear function     | 67.7462      | 't 0 -c 3'             |
| Polynomial function | 63.1434      | 't 1 -c 3 -d 3'        |
| Radial basis function| 82.2515      | 't 2 -c 3 -g 4'       |
| Sigmoid function    | 6.5712       | 't 3 -c 3 -g 4 -r 0'  |

The experimental results show that the accuracy of prediction classification is the highest with the radial basis function as the kernel function.

Select the Best Parameters by Grid Method
The work flow is shown in Figure 1.

In comparison experiment of kernel functions, the value of penalty parameters \( c \) and kernel function parameters \( g \) based on experience. In order to improve the classification accuracy, the aim of experiment 2 is to find the better parameters \( c \) and \( g \) based on grid method. The experimental results are shown in Figure 2, the range of \( c \) can narrow down to \([2,4]\), and the range of \( g \) can narrow down to \([-4,-3]\). Based on the above rough parameters selection, we can use SVMcgForClass procedure to search fine parameters selection, the experimental results are shown in the figure 2, when \( c = 5.278 \), \( g = 0.0625 \), the local optimal accuracy is 89%.
Select the Best Parameters by Immune Genetic Algorithm

In the sense of CV, using the grid search method can only find the local optimal solution. However, using immune genetic algorithm can find the global optimal solution without traversing all parameters points in the grid.

Immune genetic algorithm has the advantage of genetic algorithm, not only has the ability of random global parallel search, but also retains the high fitness antibody in the population. Because the immune operator is introduced, the immune genetic algorithm can effectively avoid the premature convergence problem of the traditional genetic algorithm. Therefore, immune genetic algorithm can be used for global optimization of SVM model parameters.

In IGA-SVM, IGA algorithm is used to optimize penalty factor $c$ and kernel function parameter $g$. First, penalty factor $c$ and kernel function parameters $g$ are used to construct the initial antibody vector: $X = [c, g]$. In order to minimize the error sum of squares between the actual output and the expected output of SVM, the fitness function $f(x_i)$ of SVM training sample classification accuracy $E(x_i)$ is defined: $E(x_i): f(x_i) = E(x_i)$. Where the $f(x_i)$ is $i$-th antibody fitness, $E(x_i)$ is $i$-th antibody corresponding classification accuracy. The flowchart of the IGA-SVM is shown in Figure 3.
The results of Primi isolated words spectrogram classification based on IGA-SVM are shown in Figure 4, it can be seen from the Figure 4, when $c = 4.5948$ and $g = 0.0625$, the classification accuracy can be achieved 90%.

**Conclusion**

Two sets of contrasting experiments were made using a radial basis function as kernel function. Optimal parameters $c = 4.5948$, $g = 0.0625$ and classification accuracy 89% are obtained based on grid method, optimal parameters $c = 3.8972$, $g = 0.06485$ and classification accuracy 91% are obtained based on IGA-SVM. Therefore, immune genetic algorithm is the most effective method for Primi isolated words spectrogram classification.

**References**

[1] Shaozun Lu, Primi language Ji Zhi [M], Beijing: Nationalities Press, 1983.
[2] Tartakovsky, D.M. and S. Broyda, Vapnik V.N, The nature of statistical learning theory, 2000.
[3] Kaminskyj I, Czaszejko T, Automatic Recognition of Isolated Monophonic Musical Instrument Sounds using kNNC, 2005, 24 (2-3) 199-221.
[4] S.J. Wang, Chunyan Xu, X.X. Pan, An Dong, X. Chen, et al., Single Figure Syllable Modeling Based on Neural Network for Continuous Speech Recognition. Acta Electronica Sinica, 2005, 33 (10) 1883-1885.
[5] Ben-Hur A., Noble W.S., Frontiers Kernel methods for predicting protein-protein interactions, Bioinformatics 2005, 21 (Suppl) 1-38.

[6] Raghavendra. N.S, Deka PC, Support vector machine applications in the field of hydrology: A review. Appl Soft Comput 2014, 19 (Supplement C) 372-386.

[7] Taylor, C.E., Book Review: Adaptation in Natural and Artificial Systems: An introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, Complex Adaptive Systems, John H. Holland, 2004, 69 (1).

[8] Steinbrunn M., Moerkotte G, Kemper A, Heuristic and randomized optimization for the join ordering problem [J], VIdb Journal, 1997, 6 (3) 191-208.

[9] Eberhart J.K.R., Particle swarm optimization [C]: Neural Networks, 1995, Proceedings, IEEE International Conference on, IEEE, 1995, pp. 1942-1948.

[10] Luo J.W., Wang T., Motif discovery using an immune genetic algorithm, Journal of Theoretical Biology, 2010, 264 (2) 319-325.

[11] Song Zhiyong, Application of MATLAB in Speech Signal Analysis and Synthesis. 2013: Beijing University of Aeronautics and Astronautics Press.

[12] Shaikh S.H., Maiti A.K., Chaki N., A new image binarization method using iterative partitioning, Machine Vision & Applications, 2013, 24(2) 337-350.