DHMM Speech Recognition Algorithm Based on Immune Particle Swarm Vector Quantization

Aiping Ning[1,2], Xueying Zhang[1], and Wei Duan[1]

[1] College of Information Engineering, Taiyuan University of Technology, Taiyuan, China
tyzhangxy@163.com, dw861123@163.com
[2] College of Electronics and Information Engineering,
Taiyuan University of Science and Technology, Taiyuan, China
aiping_ning@126.com

Abstract. This article presents a novel Immune Particle Swarm Optimization (IPSO), which combines the artificial immune system methods like immunologic memory, immunologic selection and vaccination together, by making reference to the self adjusting mechanism derived from biological immune system. IPSO as a method of Vector Quantization applied to the Discrete Hidden Markov Model (DHMM) and proposes IPSO-DHMM speech recognition algorithm. Each particle represents a codebook in the algorithm. The experiments using IPSO vector quantization algorithm get optimal codebook. Finally it enters the DHMM speech recognition system to train and recognize. The experimental results show that the IPSO-DHMM speech recognition system has faster convergence, higher recognition ratio and better robustness than the PSO-DHMM algorithm.

Keywords: IPSO, DHMM, speech recognition, vector quantization.

1 Introduction

As one of the statistical models of speech signals, HMM has been widely applied to the field of speech recognition [1]. The DHMM isolated word speech recognition system involves with vector quantization technology and the key of vector quantization is codebook designing. LBG algorithm was proposed in the 1980s, which has been widely adopted as one of the best codebook designing methods, but it is very sensitive to the initial codebook and is easy to fall into local convergence [2]. Due to its fast convergence rate, simple operation, less parameters, parallel processing ability and other advantages, the PSO has been widely used to solve different kinds of complex optimization problem. Reference [3] proposes applying PSO to image vector quantization and it puts forward a vector quantization codebook designing method based on PSO. But the basic PSO learning mechanism is not strongly connected with vector quantization, thus it caused the unsatisfactory performance. Reference [4~5] proposed some improved PSO algorithm for vector quantization.

In the recent years, inspired by biological immune system and its function, people started to research artificial immune algorithm and have successfully applied it to clustering, anomaly detection, function optimization and other engineering problems

H. Deng et al. (Eds.): AICI 2011, Part III, LNAI 7004, pp. 420–427, 2011.

[6]. This article references the self adjusting mechanism of immune system and integrates the artificial immune system methods like immunologic memory, immunologic selection and vaccination into PSO. It presents a new IPSO. Every particle of the algorithm is an antibody. They helped the algorithm obtain a strong global convergence and improved the diversity of particle swarm. It combines IPSO and LBG together and presents IPSO vector quantization codebook design method. This article applies the method to DHMM speech recognition and proposes IPSO-DHMM speech recognition algorithm. The results show that the IPSO-DHMM speech recognition system has a higher recognition ratio.

2 IPSO Algorithm

PSO is an evolutionary approach introduced by [7]. This algorithm constitutes a metaphoric scheme of bird flocking and fish schooling. In the PSO algorithm, each particle represents a possible solution to the task at hand, and the particle swarm starts with the random initialization of a population of individuals in the space. All particles have fitness values. The particles fly through the *D*-dimensional problem space by learning from the best experiences of all the particles.

2.1 PSO Algorithm

A particle status on the search space is characterized by two factors: its position and velocity. The position vector and the velocity vector of the particle *i* in *D*-dimensional space can be indicated as $\mathbf{x}_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$ and $\mathbf{v}_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$ respectively. The best position encountered by each particle is called pbest ($\mathbf{p}_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$). The best position among all particles found so far at time *t* is called gbest ($\mathbf{p}_g = (p_{g1}, p_{g2}, \cdots, p_{gD})$).At each iteration *t*, the position $x_{id}(t)$ and the velocity $v_{id}(t)$ are updated for next iteration *t*+1 by the following equations.

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_1 (p_{id}(t) - x_{id}(t)) + c_2 r_2 (p_{gd}(t) - x_{id}(t)) \tag{1}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{2}$$

Where w is inertia weight; c_1 and c_2 are acceleration Constant; r_1 and r_2 are uniformly distributed random number between 0 and 1.

The value of $v_{ij}(t)$ can be clamped to the range $[v_{\min}, v_{\max}]$ to control excessive roaming of particles outside the search space. The best position of each particle is updated using equation (3) and the global best position found by using equation (4).

$$p_{id}(t+1) = \begin{cases} x_{id}(t+1) & if\ f(x_{id}(t+1)) < f(p_{id}(t)) \\ p_{id}(t) & otherwise \end{cases} \tag{3}$$

$$p_{gd}(t+1) = \arg\min_{i \in \{1,2,\cdots,N\}} f(p_{id}(t+1)) \tag{4}$$

Where $f(\cdot)$ denotes fitness function.

2.2 IPSO Algorithm

This article treats each particle as an antibody. By using immunologic memory, immunologic selection, vaccination and other processing mechanisms in PSO, it proposes IPSO algorithm.

Immunologic Memory: Immunologic memory means immune system treats the antibody produced by antigens as memory cells and stores them. When it encounters antigens again, immune system will activate memory cells to generate a lot of antibodies to eliminate them. In the article, the best particle P_g generated by iteration was stored as memory cells. When the fitness of some particles can't meet the requirement, they will be replaced by memory cells. Immunologic memory can speed up the search by using immunologic memory function.

Immune Regulation: During the process of particle swarm updating, we always want to store those particles with higher fitness. But when the density of the particle is too high, it can't guarantee the diversity of particles and it is easy to fall into local optimum situation. So the article uses the regulation mechanism of immune system, that is, the higher the antibody particle's density, the greater the inhibitory effect that the system imposes, thus reduces the probability of being chosen; on the contrast, the lower the antibody particle's density, the greater the promoting effect that the system provides, thus increases the probability of being chosen. This regulation mechanism ensures the diversity of particles.

Using equation (5) calculates the density of particle i.

$$D(x_i) = \frac{1}{\sum_{j=1}^{N+M} \left| f(x_i) - f(x_j) \right|}, i = 1,2,\cdots,N+M \tag{5}$$

The probability selection formula of particle density can be derived from formula (5):

$$P(x_i) = \frac{\frac{1}{D(x_i)}}{\sum_{i=1}^{N+M} \frac{1}{D(x_i)}} = \frac{\sum_{j=i}^{N+M} \left| f(x_i) - f(x_j) \right|}{\sum_{i=1}^{N+M} \sum_{j=1}^{N+M} \left| f(x_i) - f(x_j) \right|}, i = 1,2,\cdots,N+M \tag{6}$$

Where $f(x_i), i = 1,2,\cdots N+M$ is the fitness of particle *i*. We can know from formula (6) that the more the particles similar to particle *i*, the smaller the probability of particle *i* being chosen and vice versa.

Vaccination

Vaccination is an important aspect of applying immunologic memory function to medical application. This paper uses this idea in PSO. By using vaccine extracting, vaccination and immunologic selection to complete the search process, which improves the performance of the algorithm and inhibits the degradation. Vaccination means partly change the antibody genes in accordance with the vaccine. If the adaptive degree is not better than before, we'll keep the antibody before vaccination; while if it is better than before, we'll keep the antibody after vaccination.

Generally, vaccine production is according to problem analyzing and feature information. But sometimes the problem to be solved has no prior information, so we can't get the proper vaccine. This paper selects the best particle P_g derived from the particle swarm updating process as vaccine. Vaccination makes those best particles to be inherited, improves search ability and avoids degradation.

The steps of IPSO

The procedures of IPSO algorithm can be illustrated as follows:

Step1: Initialize swarm, randomly generate *N* particles, spatial dimension is *D*, Initialize a population of particles with random velocities and positions in the problem space.
Step2: Calculate the fitness of each particle, according to the equations (3) and (4) updating P_i and P_g, store P_g as immunologic memory particle;
Step3: Update the velocities and the positions of particles using the equations (1) and (2);
Step4: Randomly generate *M* particles to form new particle swarm;
Step5: Calculate density select probability of *N*+*M* particles according to formula (6) and choose *N* particles with higher probability into the next generation;
Step6: Vaccination: Randomly select a gene segment using P_g as vaccine and randomly take a certain number of particles according to a certain percentage. Replace the corresponding part of these particles with the vaccine gene segment;
Step7: Calculate the fitness of those replaced particles, if it is not better than their parents, cancel vaccination; if it is, keep these particles and form new particle swarm;
Step8: If it reaches the maximum iterative number, stop the process; if not, turn to step2.

3 DHMM Speech Recognition System

HMM speech recognition system needs to do preprocessing, feature extracting, vector quantizing, HMM model training and recognizing. HMM speech recognition system is shown in Figure 1.

The procedures of DHMM speech recognition algorithm based on IPSO vector quantization:

Signal preprocessing: Signal sampling, anti-aliasing filtering and other special processing.

Speech signal feature extracting: This article adopts zero-crossing rate and peak altitude method (ZCPA), which is 1024 dimensions speech feature vector.

Speech signal feature vector quantizing: After feature extracting, the data rate of feature vector sequence will be very high and it will affect the following processing. Therefore, coding method must be employed to compress data. Isolated word speech recognition system adopts DHMM, so the feature vector should be quantized. Vector

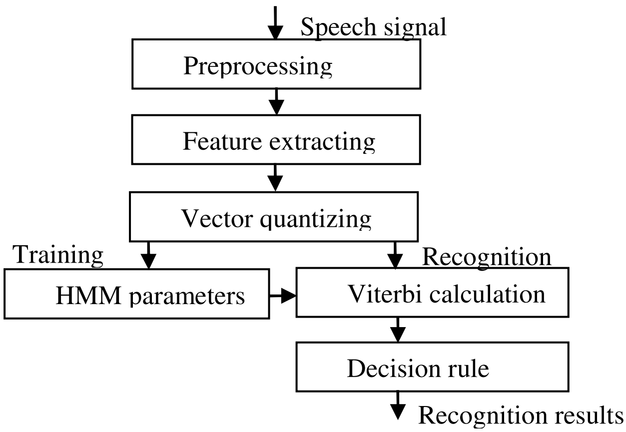

Fig. 1. Speech recognition system based on DHMM

quantization needs to generate codebook first. This article uses IPSO vector quantization to generate codebook. Every word's 1024 dimension feature vector will be quantized. Then these words are replaced by code-vector labels and become input signals of the next step. The specific process:

1) Swarm initialization. Pick out D vectors from the training vector set of speech database randomly to serve as a codebook, repeat N times to get N particles. Initialize particles velocity v_{ij} and position x_{ij} .

2) IPSO optimizes initial codebook. Carry out step 2 to step 7 of Immune Particle Swarm Optimization.
3) In accordance with the nearest-neighbor rule, reassure the cluster partition of each particle or each codebook and calculate the new cluster center to form new codebook.
4) If it reaches the maximum iterative number, stop the process; if not, turn to step2.
HMM model training: The code-vector label of word feature is the input sequence of HMM. The HMM is from left to right and no leaping, each word has 5 status. It adopts the classic Baum-Welch algorithm as training method, every word was trained into a set of model parameters ultimately.

Recognizing result decision: Use Viterbi algorithm to calculate the matching probability of each word and the model parameters are formed by the former step. The model with the biggest probability is the recognizing result.

4 Experiment Methods and Results

4.1 Experimental Setup

This article adopts C++ language to realize IPSO vector quantization HMM isolated word speech recognition system based on ZCPA speech feature.

Experimental data: Experimental data come from 16 persons, they read 10 words, 20 words, 30 words, and 50words under different Gaussian white noise (include 15dB、20dB、25dB、30dB and none) to form a speech database, each person read each word 3 times. 9 persons' voices were used as training database, other 7 persons' voice were used as recognizing database

Parameters Setting: The sampling rate of speech signal is 11.025 kHz. The parameters setting of IPSO algorithm: swarm size $N = 20$; new particles $M = 5$; particle dimension $D = 512$; maximum iterative number $t_{\max} = 10$; acceleration Constant $c_1 = c_2 = 1.8$; inertia weight W is updated using formula (7),

$$W = W_{\max} - t \times \frac{W_{\max} - W_{\min}}{t_{\max}} \tag{7}$$

$W_{\max}$ is the maximum value of inertia factor, value 1.0, $W_{\min}$ is the minimum value of inertia factor, value 0.4, t is iteration number, $t_{\max}$ is the maximum iteration number; fitness function f is calculated by the following equation,

$$f(i) = \sum_{j=1}^{J} \sum_{x \in \boldsymbol{\theta}_j} \sqrt{(\mathbf{x} - \mathbf{c}_j)^2} \quad i = 1,2,\cdots N \tag{8}$$

Where $\mathbf{c}_j$ is clustering center *jth*, $\boldsymbol{\theta}_j$ is the data set belonged to clustering center *jth*, clustering category number *J*. A smaller $f(i)$ indicates the codebook performance is better. We use 27 samples to train each word of the HMM model training.

Recognizing result is judged by the ratio of correctly recognized words and all the testing words.

4.2 Experimental Results and Analysis

Table 1 shows the recognition results in the experiment of speech recognition using IPSO-DHMM algorithm and PSO-DHMM algorithm on various SNR and various vocabulary cases.

We can see from the table 1 that DHMM speech recognition system based on the IPSO codebook design algorithm is better than that of the PSO in terms of recognition rate under different vocabulary size and different SNR. The IPSO has jumped out of the local optimum and improved the system performance. The recognition rates of the IPSO-DHMM are stable; the data distribution is more concentrated while the data distribution of PSO-DHMM is scattered. This shows that the IPSO-DHMM speech recognition system has a better robustness. But the increasing rate of recognition ratio are almost same under different SNR and the recognition ratio is best in no noise indicates, which shows that the anti-noise performance of the system need to be improved.

Table 1. The Comparision of Recognition Results Based on IPSO-DHMM and PSO-DHMM (%)

Vocabulary size	SNR(dB) / Vector Quantization	15	20	25	30	Clean	Average recognition rate
10	PSO	85.7	85.3	85.7	87.6	91.4	87.2
	IPSO	86.7	86.9	87.3	88.7	92.1	88.3
20	PSO	77.2	77.9	81.2	86.4	86.2	81.8
	IPSO	78.6	81.2	83.4	87.5	88.7	83.9
30	PSO	80.2	80.5	84.8	82.9	84.3	82.6
	IPSO	82.1	82.9	87.1	84.9	87.5	84.9
40	PSO	83.3	76.9	77.9	81.2	82.2	80.3
	IPSO	86.2	79.6	81.0	84.2	84.6	83.1
50	PSO	72.6	76.2	81.1	79.0	81.8	78.2
	IPSO	75.1	78.5	82.4	83.0	84.1	80.6
Average recognition rate	PSO	79.8	79.4	82.2	83.4	85.2	--
	IPSO	81.7	81.8	84.2	85.7	87.4	--

5 Conclusion

This paper proposes IPSO algorithm which combines the PSO global optimization ability and the immunologic information processing mechanism of immune system together. This algorithm is easy to realize, improves the ability of getting rid of the local-peak point of PSO and improves the convergence speed and accuracy of the algorithm evolutionary process and ensures the diversity of the particle swarm. IPSO is used to codebook design of vector quantization in the paper and proposes DHMM speech recognition algorithm based on IPSO vector quantization. Experimental results show that the IPSO-DHMM speech recognition system has higher recognition ratio than the PSO-DHMM speech recognition system, which objectively proves the effectiveness of the IPSO.

Acknowledgment. The project is sponsored by Taiyuan University of Science and Technology Research Foundation for Young Scholars（20093013）.

References

1. Vicente-Pena, J., Díaz-de-María, F., Bastiaan Kleijn, W.: The synergy between bounded-distance HMM and spectral subtraction for robust speech recognition. Speech Communication 52, 123–133 (2010)

2. Lib, I., Lin, T.-S., Liao, L.: Codebook Design for Vector Quantization with Most Dispersed Codewords in Initialization. Journal of Chinese Computer Systems 30, 780–783 (2009)
3. Chen, Q., Yang, J., Gou, J.: Image Compression Method Using Improved PSO Vector Quantization. In: Pro. of the International Conference on Advances in Natural Computation, Changsha, China, pp. 490–495 (2005)
4. Li, X.-J., Xu, L.-P., Yang, L.: An Improved Particle Swarm Optimization Algorithm for Vector Quantization. Patter Recognition and Artificial Inteligence 21, 285–289 (2008)
5. Ji, Z., Liao, H.-L., Xu, W.-H.: A Strategy of Particle-pair for Vector Quantization in Image Coding. Chinese Journal of Electronics 35, 1916–1920 (2007)
6. Hart, E., Timmis, J.: Application areas of AIS: The past, the present and the future. Applied Soft Computing 8, 191–201 (2008)
7. Kennedy, J., Eberhart, R.C., Shi, Y.: Swarm Intelligence. Morgan Kaufman Publisher, San Francisco (2001)