JPE

Combining HMM with A Genetic Algorithm for Fault Diagnosis of Photovoltaic Inverters

Hong Zheng*,**, Ruoyin Wang†, Wencheng Xu*, Yifan Wang*, and Wen Zhu *

†,*School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, China
**Jiangsu Key Laboratory of Drive and Intelligent Control for Electric Vehicle, Zhenjiang, China.

Abstract

The traditional fault diagnosis method for photovoltaic (PV) inverters has had a difficult time meeting the requirements of the current complex systems. The main weakness lies in the study of nonlinear systems, but the diagnosis time is also long, and the accuracy is low. To solve these problems, we use a hidden Markov model (HMM) that has unique advantages in its training model and recognition for diagnosing faults. However, the initial value of the HMM has a great influence on the model, and it is possible to achieve a local minimum in the training process. Therefore, we use a genetic algorithm to optimize the initial value and achieve global optimization. In this paper, the HMM is combined with the genetic algorithm (GHMM) for PV inverter fault diagnosis. We first use Matlab to implement the genetic algorithm and determine the optimal HMM initial value, and then we use the Baum-Welch algorithm for iterative training, and finally, we use the Viterbi algorithm for fault identification. The experimental results show that the correct PV inverter fault recognition rate by HMM is about 10% higher than that of traditional methods. Using GHMM, the correct recognition rate is further increased by approximately 13%, and the diagnosis time is greatly reduced. Therefore, it is faster and more accurate to use GHMM in diagnosing PV inverter faults.

Key words: Fault diagnosis, Genetic algorithm, Hidden Markov model (HMM), Photovoltaic (PV) inverter

I. INTRODUCTION

Due to the increasingly serious environmental situation and the growing scarcity of resources, the development and utilization of solar energy has gradually developed into the focus of the world's energy strategies, and photovoltaic power generation technology is the most common and most valuable of these. The development of photovoltaic power generation control technology is more and more large and complex, and the increasing degree of automation increases the probability of failure of the system. The control system of photovoltaic power generation systems consists of a high power inverter, which, when the inverter has faults that are not timely diagnosed and repaired, will cause economic losses and security risks which cannot be undone; thus, the research on fault diagnosis technology of photovoltaic (PV) inverters is particularly critical. The current diagnostic method is based on a fast sampling circuit that raises the alarm when a fault occurs. The running state of the collected data is processed by a microprocessor to judge whether the system is in trouble, and simultaneously send out the corresponding alarm signal. However, this method requires a long time and cannot accurately generate the alarm. With the increasing demand for reliability and safety of photovoltaic power generation systems, it is an urgent problem to diagnose and locate faults in time. The traditional fault diagnosis method primarily studies the linear system. However, most of the systems are nonlinear in practical applications.

At present, the common fault diagnosis methods of PV inverters can be roughly divided into the following categories: (1) Knowledge-based fault diagnosis methods[1], which are based on the experience and knowledge that has already been mastered, including neural networks, SVMs (support vector machines), and expert systems. R. L de Araujo Ribeiro et al. proposed an inverter fault diagnosis method based on a multi neural network structure[2]. H. Keskes and A. Braham used the SVM and pitch synchronous wavelet transform to complete fault diagnosis of a motor system[3]. D.J. Chen and Y.Z. Ye used a multi neural network algorithm to diagnose the fault of a three level inverter[4]. G.S. Hu, J. Xie and F.F. Zhu proposed the classification of power quality disturbances using wavelets and fuzzy support vector machines[5]. (2) Signal processing-based methods, such as Fourier transform and wavelet theory. S. Xu, W.X. Huang et al. used a method for fault diagnosis of a six phase permanent magnet fault tolerant motor system based on fast Fourier analysis[6]. A. Bouzida et al. used the wavelet transform to perform the fault mode identification in industrial induction machines[7]. M. Pineda-Sanchez et al. used a method for diagnosis of induction motor faults in the fractional Fourier domain[8]. Y. Yin, J. Yang et al. used wavelet packet and Fourier analysis fault diagnosis of rolling bearings[9]. These methods do not need mathematical models, are easy to use, and have high diagnostic efficiency, but the parameters need to be set according to expert experience; this problem reduces the usefulness of these methods. (3) Other methods of fault diagnosis include state estimation and bond graph theory[10]. Generally, the main problems in the fault diagnosis of PV inverters can be summarized in four aspects: (1) At present, there is much theoretical research but few practical applications. (2) There are many off-line diagnostic systems but fewer online diagnostic systems. (3) There is a failure to make full use of the potential information within the system. (4) The diagnostic system is closed and needs advanced programming technology.

The object under research in this paper is the diode clamped PV inverter. In this paper, we study the fault diagnosis of the main circuit topology, combine HMM with a genetic algorithm to detect the faults of the inverter, and compare it with the neural network and SVM methods. HMM is a model based on statistical analysis[11], which has been successfully applied in the field of speech recognition and has achieved great success. Because of its powerful ability in pattern recognition, in recent years many scholars have also applied it in other fields, for example, in character recognition, face recognition, behavior recognition, and ECG recognition. In the field of fault diagnosis, there have been many achievements, such as when C.L. Zhang, X. Yue et al. used HMM to diagnose the failure of rotating machinery[12]. H. Ocak and K.A. Loparo, used HMM to monitor and diagnose the fault of bearings[13]. X. Yue proposed a complex condition fault diagnosis technology based on HMM[14]. HMM is a method based on statistical pattern recognition theory; it can deal with dynamic processes well and makes full use of potential information within the system. It can monitor and diagnose the dynamic process during system operation. However, there is a fatal disadvantage in HMM, that is, the initial value of B in the HMM has a great influence on the model, and it is possible to achieve a local minimum value in the training process, so the correct recognition rate is reduced. The genetic algorithm[15] is a simulation of the survival of the fittest nature of the evolution of the phenomenon, where the search space is mapped to a genetic space, and the possible solution is transformed into a vector (chromosome), where each element of the vector is called a gene. By calculating the fitness value of each chromosome, we select the best chromosome and get the optimal solution. Therefore, one of the most important characteristics of the genetic algorithm is global search. Therefore, we can optimize the initial value of B by the genetic algorithm. Therefore, it is of great significance to combine HMM with the genetic algorithm for fault diagnosis of PV inverters.

II. Analysis of Fault Types

Before diagnosing a fault of the PV inverter, it is necessary to generalize and summarize the possible failure modes of the main electric power. In this paper, a three-level neutral-point-clamped (NPC) PV inverter is chosen as the research object. The problem of PV inverters is the failure of the control system, which is mainly caused by the failure of power switching devices (known in this paper as IGBTs). For the study of power devices, the failure of the PV inverter can be roughly divided into the following categories:

(1) Base drive fault of any IGBT;
(2) Short circuit fault of any IGBT;
(3) Intermittent fault of any IGBT;
(4) Open circuit fault of several IGBTs in the same phase;
(5) Short circuit fault of several IGBTs in the same phase;
(6) Several IGBT faults occurring in a cross phase.

In these faults, open circuit faults and short circuit faults are the most common. Between the open circuit and short circuit faults, there is no exact distinction. Shortly after the short circuit fault occurs, it will be converted into an open circuit fault, and because the duration of the short circuit fault is extremely short, it is almost impossible to detect in real time. In an open circuit fault, the abnormal output voltage will cause serious damage to the load, so we must increase the detection of this kind of fault in practical applications. In view of this, this paper mainly studies the open circuit mode of power devices (the above numbers 1, 4 and 6). To simplify the problem, this paper only considers faults in a single phase, which can be divided into two types: single-IGBT faults and multi-IGBT faults. Two power devices having a fault at the same time is more common in multi-IGBT faults, but the fault of more than two devices is not significant because the inverter system is already unable to run.

The topology of the PV inverter is shown in Fig. 1. A total of 12 power switching devices (known in this paper as IGBTs). An open circuit fault will cause the distortion of the output current of the inverter, and the total harmonic rate will increase, which will not meet the requirements of the grid connection. It will cause more serious accidents if not handled in time. It can be seen from Fig. 1 that the open circuit fault mode of the system can be divided into the

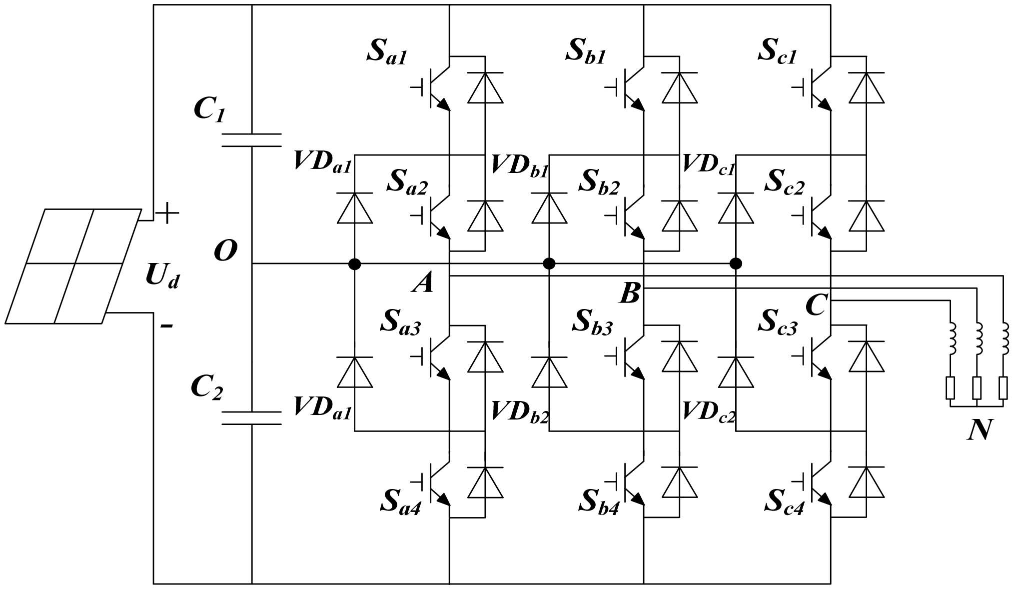

Fig. 1. Topology of Neural-Point-Clamped photovoltaic inverter

TABLE I
FAULT CODE

Number	Open circuit	Number	Open circuit
1	S_{a1}	6	S_{a1} and S_{a3}
2	S_{a2}	7	S_{a1} and S_{a4}
3	S_{a3}	8	S_{a2} and S_{a3}
4	S_{a4}	9	S_{a2} and S_{a4}
5	S_{a1} and S_{a2}	10	S_{a3} and S_{a4}

following two cases:

The first type: single IGBT fault, S_{ij} , $i = a,b,c$, $j = 1,2,3,4$

The second type: two simultaneous IGBTs fault in the same bridge arm, of which there are 18 scenarios.

This amounts to a total of 30 failure modes. Due to the three-phase symmetry of the NPC photovoltaic inverter, this paper mainly studies the typical fault mode which can reflect the whole fault, namely, 10 kinds of fault modes in single phase. These 10 fault states are coded in Table I.

III. OPTIMIZATION OF THE HMM BY THE GENETIC ALGORITHM

In this paper, we first use a genetic algorithm to optimize the HMM (GHMM) and then use GHMM for fault diagnosis of the PV inverter.

A. The basic theory and algorithm of HMM

HMM is the Hidden Markov Model and is an extension of the analysis of Markov chains. Its application is an important achievement in the field of speech recognition since the 1980s. Because the actual problem is more complex than that described by Markov chains, the HMM introduces a probabilistic statistical model and uses the probability density function to calculate the output probability of the speech parameter to the HMM models. In searching for the best state sequence, the identification results are found by the criterion of the maximum posterior probability. An HMM has five basic elements that are represented as a five element array, $\{N, M, \pi, A, B\}$:

(1) *N*: the number of states in the model. We express the *N* states as $\theta_1, \theta_2, \cdots, \theta_N$, and the state at time *t* as q_t , $q_t \in (\theta_1, \theta_2, \cdots \theta_N)$;

(2) *M*: the number of distinct observation symbols per state. We express the *M* states *v* as $v_1, v_2 \cdots v_M$, and the observable symbols at time *t* as O_t , $O_t \in (v_1, v_2, \cdots, v_M)$;

(3) π : the initial state probability distribution $\pi = (\pi_1, \pi_2, \cdots, \pi_N)$, where $\pi_i = P(q_t = \theta_i)$, $1 \le i \le N$;

(4) *A*: the state transition probability matrix, $A = \{a_{ij}\}$, where $a_{ij} = P(q_{t+1} = \theta_j, q_t = \theta_i)$, $1 \le i,\ j \le N$;

(5) *B*: the observation probability matrix, $B = \{b_j(k)\}$, where $b_j(k) = P(o_t = v_k, q_t = \theta_j)$, $1 \le j \le N, 1 \le k \le M$.

For convenience, we use the simplified form (π, A, B). Thus, an HMM can be expressed as $\lambda = (\pi, A, B)$.

There are three basic algorithms in HMM; they are, respectively, (1) the Forward–Backward algorithm, (2) the Viterbi algorithm, and (3) the Baum–Welch algorithm. The Forward–Backward algorithm is the most basic and must be used by the other two algorithms. In this paper, we introduce the Backward algorithm, and Forward is similar with it. We define a backward variable $\beta_t(i)$:

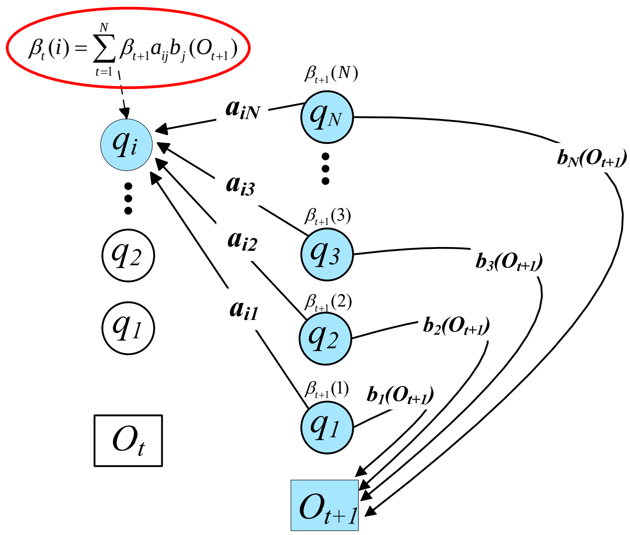

Fig. 2. Procedure of Backward algorithm.

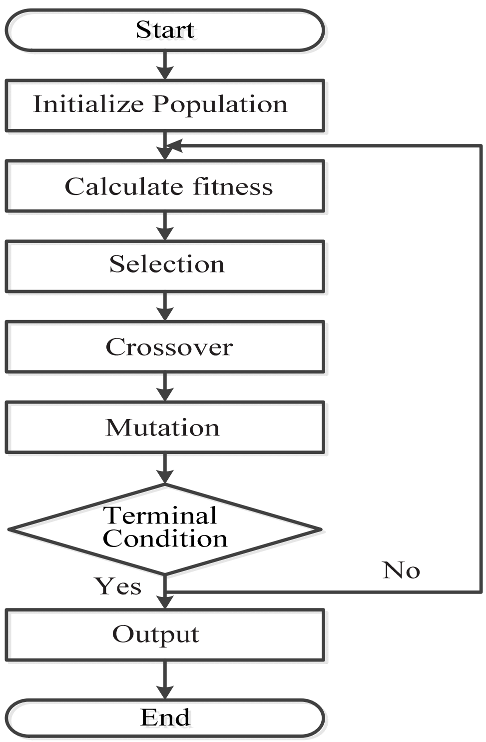

Fig. 3. Procedure of genetic algorithm.

$$\beta_t(i)=P(O_{t+1},O_{t+2},\cdots O_T;q_t=S_i\,|\,\lambda) \qquad (1)$$

Fig. 2. shows the process. Then we can calculate the backward variables of all hidden states at each time point(shown in the red circle), $\beta_t(i)$. To calculate the probability of observation sequence O at time t,we must add them all:

$$P(O|\lambda)=\sum_{i=1}^{N}\beta_t(i) \qquad (2)$$

B. Genetic algorithm

As a problem-solving strategy, the genetic algorithm is a programming technique that mimics biological evolution. The main feature is the information exchange between the group searching strategy and the individual in the group. The genetic algorithm adaptively controls the search process to obtain the optimal solution, so it is a type of global optimization search algorithm which is completely different from the traditional HMM algorithm. First, the genetic algorithm maps the problem space into the space of chromosomes. This process is called coding, and it uses the evaluation function as a basis for genetic manipulation. The evaluation function is also called the fitness function; its value is closely related to the search problem. The algorithm is initialized to form a group of chromosomes that is composed of a number of individuals. Then, the group is updated by the appropriate operator. Genetic operators include selection, crossover and mutation. The genetic algorithm is a search method using a randomized technique, but it is not the same as the general random search algorithm. The genetic algorithm has a clear search direction, which makes it much more efficient. The steps of the genetic algorithm are shown in Fig. 3.

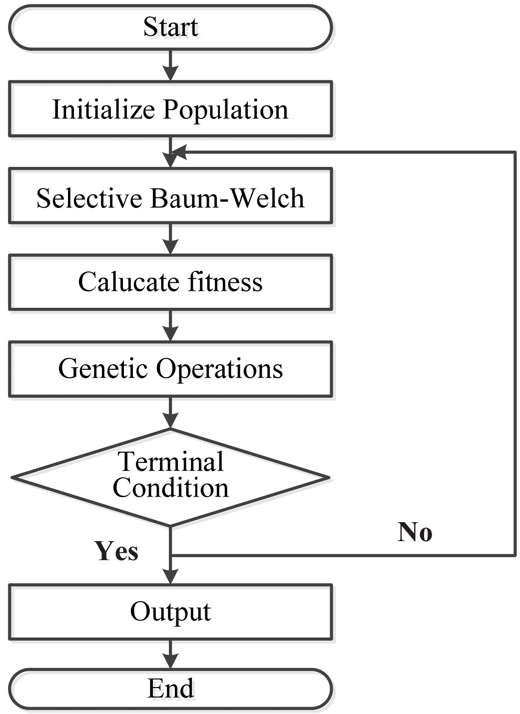

Fig. 4. Procedure of GHMM.

C. Application technology of genetic algorithm in HMM

The realization of fault diagnosis consists of two parts: training and recognition. In an HMM, the Baum–Welch algorithm is the basic training algorithm, but it has a fatal flaw: the final solution depends on the selection of the initial value, so it is easy to fall into a local optimum, which affects the recognition rate of the system. We need to set the initial

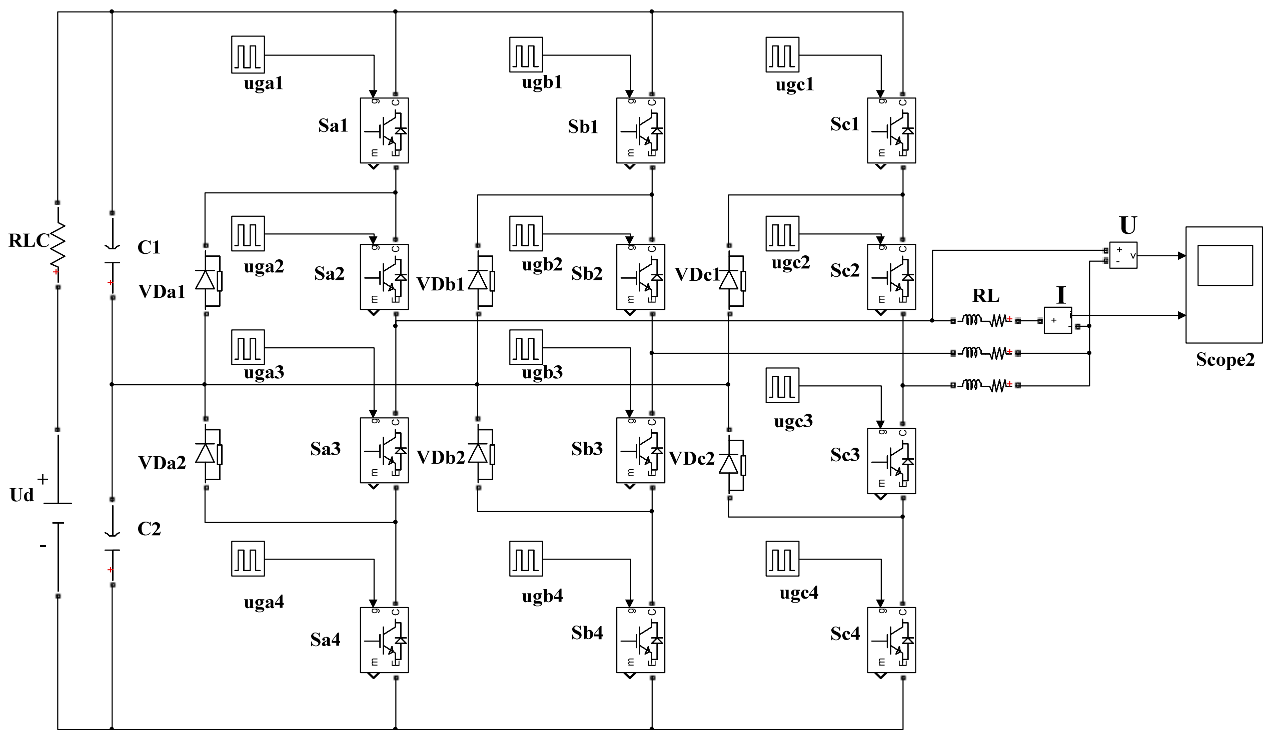

Fig. 5. Simulation model of the Three-Level NPC inverter.

value of the parameters of each group when we use the Baum–Welch algorithm. If the initial value is not set appropriately, it will take too much iteration to achieve convergence, and the algorithm may converge to a locally optimal solution rather than a globally optimal solution. Therefore, the initial value selection is a very important problem, in particular, the initial value of *B*. In this paper, we use the genetic algorithm to optimize the initial value, and then use the Baum–Welch algorithm to find the optimal model. We use the genetic algorithm to improve HMM, and the procedure is shown in Fig. 4.

The application of the genetic algorithm in HMM mainly includes the coding parameter, designing the fitness function and operator, and setting the control parameters. The key steps in GHMM are as follows:

1) Parameter coding: The general genetic algorithm does not directly address the problem of spatial parameters, but the mapping of the genetic algorithm space composed of genes according to a certain structure of the chromosome, this process is called encoding.

2) Fitness function: Evaluating the genetic algorithm is done exclusively through optimizing the fitness function..

3) Genetic operations in GHMM: The genetic operators are composed of selection, crossover and mutation. Crossover crossover is the core operator of genetic algorithm.

4) Setting control parameters: The control parameters of the general genetic algorithm include setting the population, crossover probability, mutation probability, etc.

a) Length of the coding string L: when the binary code is used to represent the individual, the selection of the length *L* of the code string is related to the accuracy of the solution.

b) Population size N: the number of individuals in a group is called the size of the population. A reasonable value for *N* depends on the specific circumstances.

c) Crossover probability P_c: the crossover probability P_c controls the frequency of the crossover operation. Generally, useful crossover probabilities take values only from 0.25 to 0.9.

IV. Fault Diagnosis Technique Based on GHMM

In actual operation, we usually cannot understand all the fault characteristics, so to diagnose how the fault occurred in the inverter, we must first collect the fault information obtained corresponding to various fault modes, so it is necessary for us to analyze all relevant information in the inverter through simulation with MATLAB software. In this paper, SIMULINK is used to model the NPC PV inverter, and the various fault models are simulated in the model. The simulation model is shown in Fig. 5. The information obtained is used for the fault diagnosis of the system. Sampling is done to measure the 10 different states and involves measuring the 2 characteristic variables of each state, that is, the voltage *U* and current *I*. Therefore, we collect the output voltage *U* and current *I* in the 10 fault states by simulation. The fault diagnosis of the PV inverter is mainly divided into two parts: the training model and fault identification.

A. Training model

(1) Put U and I, which are obtained by sampling, into the model as the observation sequence O, namely, $O = [U, I]$.

(2) Establish an HMM, left-right model, determine the initial value of the model, that is, the state number, the initial state probability, and the state transition matrix. Based on the circuit model, the implicit state is set to 4. The initial state probability is set to $\pi = [1\ 0\ 0\ 0]$, and the state transition matrix A is set to:

$$A = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

(3) Use the genetic algorithm to find the optimal initial value of B. The steps here are as follows:

(a) Coded representation: In a hidden Markov model, due to the left-right model, state transitions can only be transferred from the low state to the high state. As the initial system can only be in a low state, it is not difficult to launch the initial state probability vector as follows:

$$\pi_i = \begin{cases} 1, & i = 1 \\ 0, & i \neq 1 \end{cases} \quad 1 \le i \le N \tag{4}$$

Because the selection of the initial value of B has the greatest impact on the model, we optimize the initial value of B. First, the initial value of B is coded; the binary coding method is used in this paper. The binary encoding method is one of the most commonly used methods of encoding in genetic algorithm, using {0,1}. The range of the initial value of B is from [0, 1], the length of the parameter is 64, so it can produce 2^{64} different codes in total. The corresponding relationship is as follows:

$$000000\cdots000000=0 \longrightarrow 0$$

$$000000\cdots000001=1 \longrightarrow 0+\gamma$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$111111\cdots111111=2^{64}-1 \longrightarrow 1$$

The accuracy of binary coding is:

$$\gamma = \frac{1}{2^{64}-1} \tag{5}$$

Assume an individual's code,

$$X : a_{64}a_{63}a_{62}\cdots a_2a_1 \tag{6}$$

The corresponding decoding formula is:

$$x = \left(\sum_{i=1}^{64} a_i \cdot 2^{i-1}\right) \cdot \frac{1}{2^{64}-1} \tag{7}$$

The corresponding decoding formula is:

$$\sum_{k=1}^{m} b_{jk} = 1 \ , 1 \le j \le N \ , 1 \le k \le M \tag{8}$$

(b) Fitness function: The fitness function can reflect the pros and cons of each chromosome. The traditional algorithm takes $P(O|\lambda)$ as the optimization target. The chromosome that has the largest $P(O|\lambda)$ is the best chromosome. According to the above, we assume that Q is a state sequence. The forward and backward algorithms calculate the sum of all state sequences' $P(Q,O|\lambda)$, but the Viterbi algorithm calculates the $P(Q',O|\lambda)$ of the optimal path Q' . In terms of fault treatment, the range of $P(Q,O|\lambda)$ is very large, and the maximum value of $P(Q',O|\lambda)$ is the only component that has the absolute advantage in all $P(Q,O|\lambda)$. Therefore, we often use $P(Q',O|\lambda)$ to replace $\Sigma P(Q,O|\lambda)$. The recognition algorithm of this paper is the Viterbi algorithm, so we take the $P(Q',O|\lambda)$ as the optimization target.

The fitness of individuals is represented by the log-likelihood of each training sample.

$$f(\lambda) = \ln(P(O^{(k)} | \lambda)) \tag{9}$$

$O^{(k)}$ means the kth observation sequence, $P(O^{(k)}|\lambda)$ is obtained by the Viterbi algorithm.

(c) Design of genetic operators: The genetic operators include the crossover operator and the mutation operator. In a certain sense, the crossover operator is equivalent to a local search operation, which produces two offspring near the parent generation. The mutation operator allows the individual to jump out of the current local search area, so the combination of the two exactly reflects the optimization of genetic algorithm. In this paper, we use multi-point crossovers and multi-point mutations.

(d) Termination criterion: In this paper, the maximum algorithmic iteration is set to 100.

In the experiment, the performance of the genetic algorithm has a great relationship with the setting of the control parameters, and the optimal performance of the algorithm often requires optimal parameter setting. Therefore, we must have an optimizing search on P_c and P_m. In this paper, we use the C language to implement the above genetic algorithm, from which the optimal initial value of B is obtained.

(4) Use the Baum-Welch algorithm to iterate the initial parameters until the parameters converge to the set range.

First we define a variable:

$$\xi_t(i,j) = P(q_t = \theta_i, q_{t+1} = \theta_j \mid O, \lambda) \tag{10}$$

This expresses, under the conditions of a given model λ and observation sequence O, the probability of state θ_i at time t and state θ_j at time t+1.

$$\xi_t(i,j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O|\lambda)} = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)} \tag{11}$$

Define $\gamma_t(i)$ as the conditional probability of the state θ_i at time t.

$$\gamma_t(i)=\sum_{j=1}^{N}\xi_t(i,j) \tag{12}$$

Sum $\gamma_t(i)$ in t (from 1 to *T*-1). This will yield a quantity that can be interpreted as the number of transfers from the state θ_i.

The method for estimating parameters of a hidden Markov model includes the following:

The revaluation formula of π :

$$\overline{\pi_i}=\gamma_1(i)=\frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^{N}\alpha_t(i)\beta_t(i)} \tag{13}$$

The revaluation formula of a_{ij} :

$$\overline{A_{ij}}=\frac{\sum_{i=1}^{T-1}\xi_t(i,j)}{\sum_{i=1}^{T-1}\gamma_t(i)} =\frac{\sum_{i=1}^{T-1}\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)/P(O\mid\lambda)}{\sum_{i=1}^{T-1}\alpha_t(i)\beta_t(i)/P(O\mid\lambda)} \tag{14}$$

The revaluation formula of $b_j(k)$:

$$\overline{B_j}(k)=\frac{\sum_{t=1}^{T}\gamma_t(j)_{O_t=v_k}}{\sum_{t=1}^{T}\gamma_t(j)} =\frac{\sum_{t=1,\,O_t=v_k}^{T}\alpha_t(i)\beta_t(i)/P(O\mid\lambda)}{\sum_{t=1}^{T}\alpha_t(i)\beta_t(i)/P(O\mid\lambda)} \tag{15}$$

In our study, we use multiple training samples to train the models. Because during model training, the output may differ under different control models or even under the same control but the application is not the same. Therefore, the problem of model generalization should be considered when training model to make GHMM usable in practical application. We must consider the common features of multiple datasets; that is, we need to use multiple training samples to train a model for fault diagnosis. For fault diagnosis, the output must be collected under the same fault condition as that for the training samples.

First we define *H* training samples, that is observation sequences $O^{(1)},O^{(2)},O^{(3)},\cdots,O^{(H)}$. They may be relevant or statistically independent, therefore:

$$\begin{aligned}
P(O\mid\lambda)&=P(O^{(1)}\mid\lambda)P(O^{(2)}\mid O^{(1)},\lambda)\cdots P(O^{(H)}\mid O^{(H-1)}\cdots O^{(1)},\lambda)\\
P(O\mid\lambda)&=P(O^{(2)}\mid\lambda)P(O^{(3)}\mid O^{(2)},\lambda)\cdots P(O^{(1)}\mid O^{(H)}\cdots O^{(2)},\lambda)\\
&\vdots\\
P(O\mid\lambda)&=P(O^{(H)}\mid\lambda)P(O^{(1)}\mid O^{(H)},\lambda)\cdots P(O^{(H-1)}\mid O^{(H)}\cdots O^{(1)},\lambda)
\end{aligned} \tag{16}$$

Then introduce weight coefficient ω :

$$\begin{aligned}
\omega_1&=\frac{1}{H}P(O^{(2)}\mid O^{(1)},\lambda)\cdots P(O^{(H)}\mid O^{(H-1)}\cdots O^{(1)},\lambda)\\
\omega_2&=\frac{1}{H}P(O^{(3)}\mid O^{(2)},\lambda)\cdots P(O^{(1)}\mid O^{(H)}\cdots O^{(2)},\lambda)\\
&\vdots\\
\omega_H&=\frac{1}{H}P(O^{(1)}\mid O^{(H)},\lambda)\cdots P(O^{(H-1)}\mid O^{(H)}\cdots O^{(1)},\lambda)
\end{aligned} \tag{17}$$

In formulas (14) and (15), we use the following formula to express $P(O\mid\lambda)$:

$$P(O\mid\lambda)=\sum_{h=1}^{H}\omega_h P(O^{(h)}\mid\lambda) \tag{18}$$

The initial model is defined as $\lambda=(A,B,\pi)$, and the revaluation model is defined as $\overline{\lambda}=(\overline{A},\overline{B},\overline{\pi})$. The Backward algorithm is used to calculate $P(O|\overline{\lambda})$, such that $P(O|\overline{\lambda})>P(O|\lambda)$. These calculations are then repeated using $\overline{\lambda}$ instead of λ until convergence.

B. Fault diagnosis

After the training models are completed, we use the Viterbi algorithm to identify the diagnosis. The Viterbi algorithm is a method based on dynamic programming to find a single optimal state sequence which has the largest $P(Q\mid O,\lambda)$.

First, we define a variable $\delta_t(i)$:

$$\delta_t(i)=\max_{q_1,q_2,\cdots,q_{t-1}}P\,[q_1,q_2,\cdots,q_t=S_i,O_1,O_2,\cdots O_t\mid\lambda] \tag{19}$$

This variable is at time *t*, along a path to reach the state S_i, and generates the maximum probability of the observed $\{O_1, O_2, ..., O_t\}$.

(1) Initialization:

$$\delta_t(i)=\pi_i b_i(O_1),\ 1\le i\le N,\ \psi_1(i)=0 \tag{20}$$

(2)Iterative calculation:

$$\delta_t(j)=\max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}]b_j(O_t)\ ,\ 2\le t\le T,1\le j\le N \tag{21}$$

$$\psi_t(j)=\arg\max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}]\ ,\ 2\le t\le T,1\le j\le N \tag{22}$$

(3) Final calculation:

$$P=\max_{1\le i\le N}[\delta_T(i)] \tag{23}$$

Then, we can identify the diagnosis that has the largest *P*.

C. Training and diagnostic results in simulation:

With the convergence error set to 1×10^{-5}, the training curves for the 10 fault states are shown in Fig. 6.

The number of iterative steps required by the simulation under different fault models is shown in Table II. Each fault corresponds to a GHMM. Therefore, we can see that the training of the GHMM requires fewer iteration steps and has a fast convergence rate.

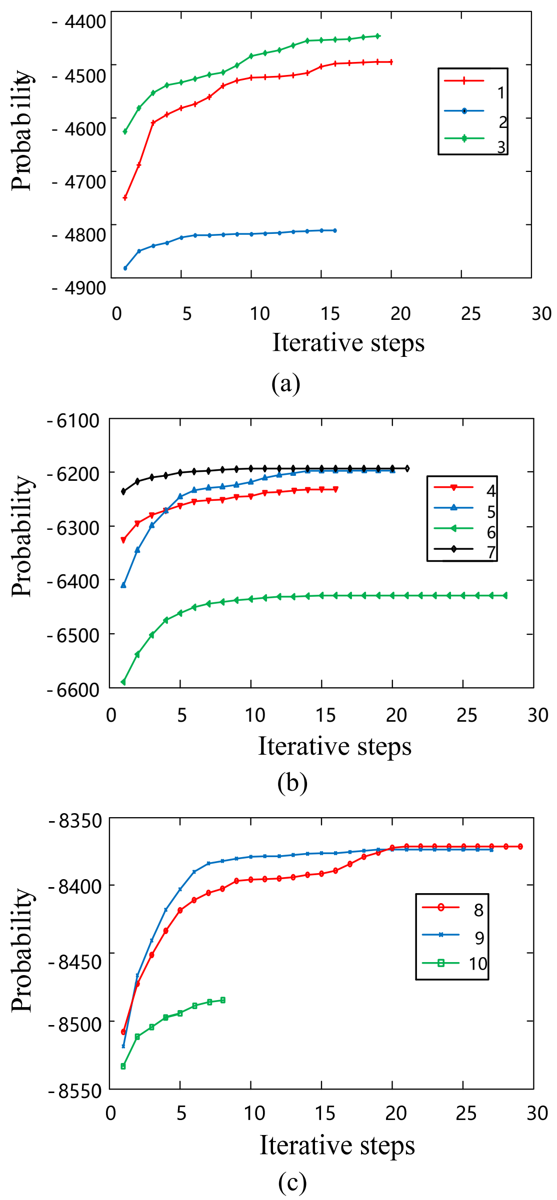

Fig. 6. (a) Fault numbers 1~3, (b) fault numbers 4~7 and (c) fault numbers 8~10

Then, we use the trained GHMM model to identify the measured signal and calculate $P(Q|O,\lambda_i)$, $i\in(1,\cdots,10)$. The model corresponding to the maximum value of $P(Q|O,\lambda_i)$ is the result. We take the test samples of 10 kinds of faults into each model to calculate the $P(Q|O,\lambda_i)$ and find the largest one. Table III and Table IV show the outputs of each model. The red number is the largest value in each column. We find that the model number corresponds to the fault number accurately.

The test result is the model corresponding to the maximum of $P(Q|O,\lambda_i)$, $i\in(1,\cdots,10)$. Therefore, the model trained by GHMM can accurately identify the fault.

To verify the validity of the circuit state identification

TABLE II
ITERATIVE STEPS

Fault number	Iterative steps	Fault number	Iterative steps
1	20	6	28
2	16	7	22
3	19	8	29
4	16	9	27
5	20	10	8

TABLE III
OUTPUTS OF MODELS

Model number \ Fault number	1	2	3	4	5
Number 1	**13.01**	-487.66	-86.12	-275.01	-205.51
Number 2	-169.32	**4.13**	-421.23	-1452.12	-210.12
Number 3	-3620.15	-41.76	**-1.55**	-542.13	-254.17
Number 4	-4402.29	-660.89	-43.12	**-5.42**	-3468.75
Number 5	-1542.14	-5015.06	-452.18	-2013.45	**-2.74**
Number 6	-263.16	-3123.02	-212.09	-2111.71	-2419.89
Number 7	-473.45	-743.11	-311.45	-3710.14	-273.33
Number 8	-907.66	-763.82	-4211.74	-112.43	-172.02
Number 9	-3167.78	-513.69	-76.54	-542.27	-4475.67
Number 10	-616.01	-122.53	-816.01	-101.75	-163.08
Test result	NO.1	NO.2	NO.3	NO.4	NO.5

TABLE IV
OUTPUTS OF MODELS

Model number \ Fault number	6	7	8	9	10
Number 1	24.45	-347.12	-86.59	-721.45	-1.79
Number 2	-143.75	-632.12	-711.42	-1010.46	-501.46
Number 3	-2109.34	-444.17	-501.92	-57.63	-1790.84
Number 4	-301.43	-500.58	-1420.46	-427.22	-314.78
Number 5	12.47	-2175.53	-817.62	-2936.34	-3075.44
Number 6	**102.01**	-142.35	-287.97	-1101.86	-200.13
Number 7	-743.85	**-0.18**	-2157.42	-340.17	-347.48
Number 8	43.47	-793.19	**-5.64**	-56.86	-1920.81
Number 9	-1435.02	-1032.75	-267.37	**-15.44**	-19.09
Number 10	-39.47	-843.52	-142.09	-410.14	**27.01**
Test result	NO.6	NO.7	NO.8	NO.9	NO.10

using GHMM, we carried out many experiments in this paper and selected test samples in different states, for a total of 1000 groups (there are 100 test samples in each fault state).

Table V shows the diagnostic results in simulation. The overall recognition has 98% accuracy, which is a good recognition performance.

In this paper, we also use an HMM without the genetic algorithm to diagnose the PV inverter fault. The diagnostic results are shown in Table V. The overall recognition of this algorithm has 86.3% accuracy. Therefore, we find that the GHMM can realize the optimization of the initial value of B

TABLE V
DIAGNOSTIC RESULTS IN SIMULATION

Fault number	Times of diagnosis	GHMM		HMM	
		Correct times	Correct rate	Correct times	Correct rate
1	100	100	100%	84	84%
2	100	100	100%	91	91%
3	100	100	100%	83	83%
4	100	96	96%	93	93%
5	100	94	94%	91	91%
6	100	100	100%	89	89%
7	100	100	100%	75	75%
8	100	95	95%	88	88%
9	100	95	95%	84	84%
10	100	100	100%	85	85%

and effectively improve the recognition rate.

To verify the advantages of using GHMM in diagnosing PV inverter faults, we compare it with other traditional pattern recognition methods. In previous studies, traditional pattern recognition methods have been based on neural networks and SVMs (support vector machines). A neural network is an information processing system for simulating the biological nervous system, which can imitate the human brain in learning, memory, recognition and many other functions. Since the development of neural networks, there have been many structural models and algorithms, of which the back propagation (BP) neural network is the most widely used. The BP neural network is a feed-forward network. The parameters of a BP neural network are shown in Table VI.

We used a BP neural network to identify the 10 kinds of faults in our NPC PV inverter model. The results of the simulation are shown in Table VII. In this method, the choice of the number of hidden layer nodes considerably affects the results. In theory, the higher the number of hidden layer nodes is, the higher the correct recognition rate is. Fault diagnosis using the BP neural network method has the following disadvantages [16]. First, the method falls into the local minimum, possesses a slow convergence rate, and produces oscillation. Second, no explicit formulas and theories are available as a guide to determine the numbers of hidden layers and nodes; these are usually calculated based on experience. Therefore, this algorithm is flawed, and this is why the BP method used in this study has a long training time and a very low correct rate.

SVM is a new generation of learning algorithm based on statistical learning theory. SVM is based on the principle of structural risk minimization. There are 3 problems in using SVM: (1) the selection of the kernel function; (2) the selection of the kernel function parameter and error cost coefficient C; (3) generalization of SVM for multi class problem identification. In this paper, we choose a radial basis function: $K(x,y)=\exp(-\gamma\|x-y\|^2)$, $\gamma=0.0082334$, and the error cost coefficient C=128. Then, we use a one-to-one voting strategy. The parameters are shown in Table VIII.

TABLE VI
BP NEURAL NETWORK PARAMETERS

Parameter	Value
Learning rate	0.02
Maximum number of iterations	1000
Target error of training	0.0001
Neuron transfer function in the hidden layer	logsig
Neuron transfer function in the output layer	trainlm

TABLE VII
BP NEURAL NETWORK RESULT

Average number of iterations	422
Average training time	30.01s
Correct recognition rate	71.02%

TABLE VIII
SVM PARAMETERS

Parameter	Value
Kernel function	radial basis
Parameter γ	0.0082334
Error cost coefficient C	128
Promotion strategy	one to one vote

TABLE IX
SVM RESULTS

Average number of iterations	657
Average training time	75.46s
Correct recognition rate	75.66%

The results of the simulation are shown in Table IX. In this method, the selection of the kernel function and error cost coefficient C significantly influences fault identification. The selection of these two parameters differs in different systems. When using SVM for fault diagnosis, satisfactory accuracy cannot be obtained if the method is not optimized [17]-[18].

Therefore, this method is also flawed; it is unstable and has poor robustness. This method has lower recognition accuracy than GHMM.

To study the advantages and feasibility of applying GHMM to fault diagnosis of PV inverters, we compared and analyzed the above four different methods. Fig. 7 shows the simulation data obtained by different methods.

The following conclusions can be obtained from the experimental data shown in Fig. 7. From Fig 7(a), the average number of steps of the GHMM and HMM iterations are much less than that of the other two methods. The neural network and SVM models need to be iterated several times to converge to the set value. From Fig. 7(b), regarding the training time, SVM takes the longest time, followed by the neural network, and while GHMM and HMM require almost the same time, the training time of HMM is the shortest. From Fig. 7(c), the correct recognition rate of the GHMM is

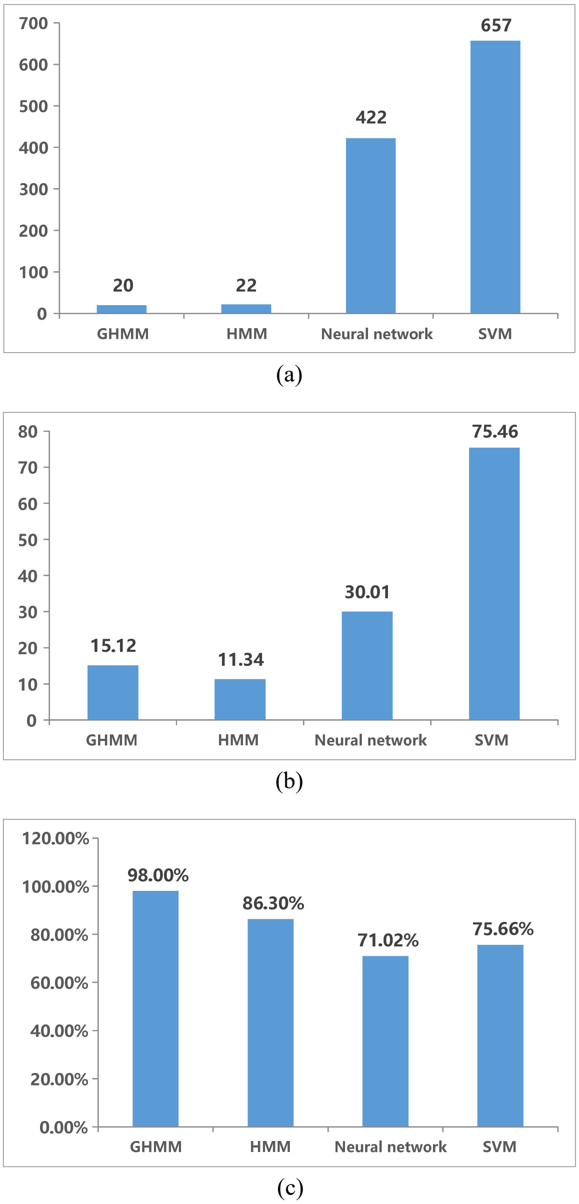

Fig. 7. (a) Average iterative steps, (b) average training time(s) and (c) correct recognition rate.

significantly better than those of the other three methods.

Comparing the GHMM and HMM, the number of iteration steps and training time are almost the same, but the recognition rate of the former is obviously higher than that of the latter, reaching 98%. Therefore, fault diagnosis based on the GHMM method is feasible and more advantageous.

V. Experimental Verification

We studied the fault diagnosis of a photovoltaic inverter. The NPC inverter was selected as the research object, and its topology is shown in Fig. 1. Ten open-circuit faults are shown in Table 1. The experimental setup is shown in Fig. 8.

The experimental setup was developed using DSP TMS320F2812 to generate the PWM pulse command of the

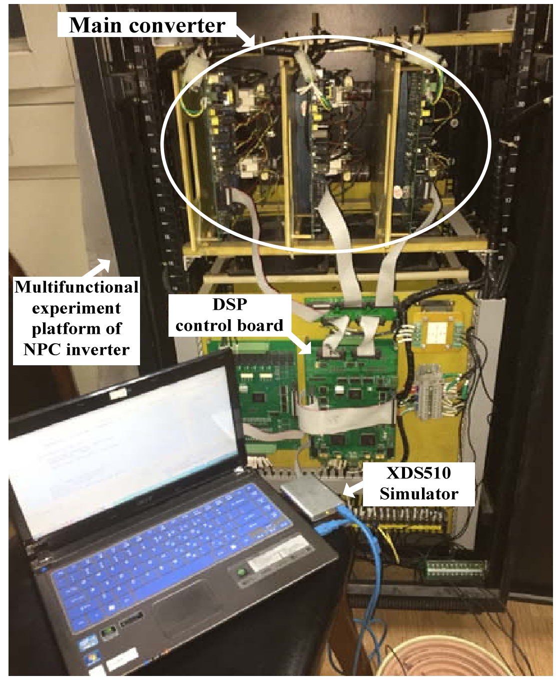

Fig. 8. Experimental setup

TABLE X
SYSTEM PARAMETERS

Parameter	Value
U_d	560 V
C_1,C_2	2200 μF
R-L load	15 Ω/2 mH

NPC inverter. The system parameters are shown in Table X. An open-circuit fault was generated by disconnecting the gate signal of the IGBT module.

Owing to the symmetry of the NPC inverter, we selected phase A for the experimental analysis, and the other phases were similar.

We measured the output voltage U and current I in each fault state with an oscilloscope, and the experimental waveform are shown in Fig. 9. We can see the output voltage U and current I differed in each fault state, and because of this difference, output voltage U and current I can be used as the characteristic value, and each model can be trained by GHMM independently. The output voltage and current data of each fault state were received by the PC and DSP control board and subsequently read and displayed by MATLAB. Observation sequence O, which consisted of the sampled U and I, was the input of the model training with GHMM. The MATLAB software was used to read and process the sampled data for model training and fault diagnosis. In our experiments, the algorithms were realized with the MATLAB program. First, we used the genetic algorithm to obtain the optimal initial value of B. Second, the Baum–Welch algorithm was used for the iterative training of the fault models. Finally, the Viterbi algorithm was used to identify

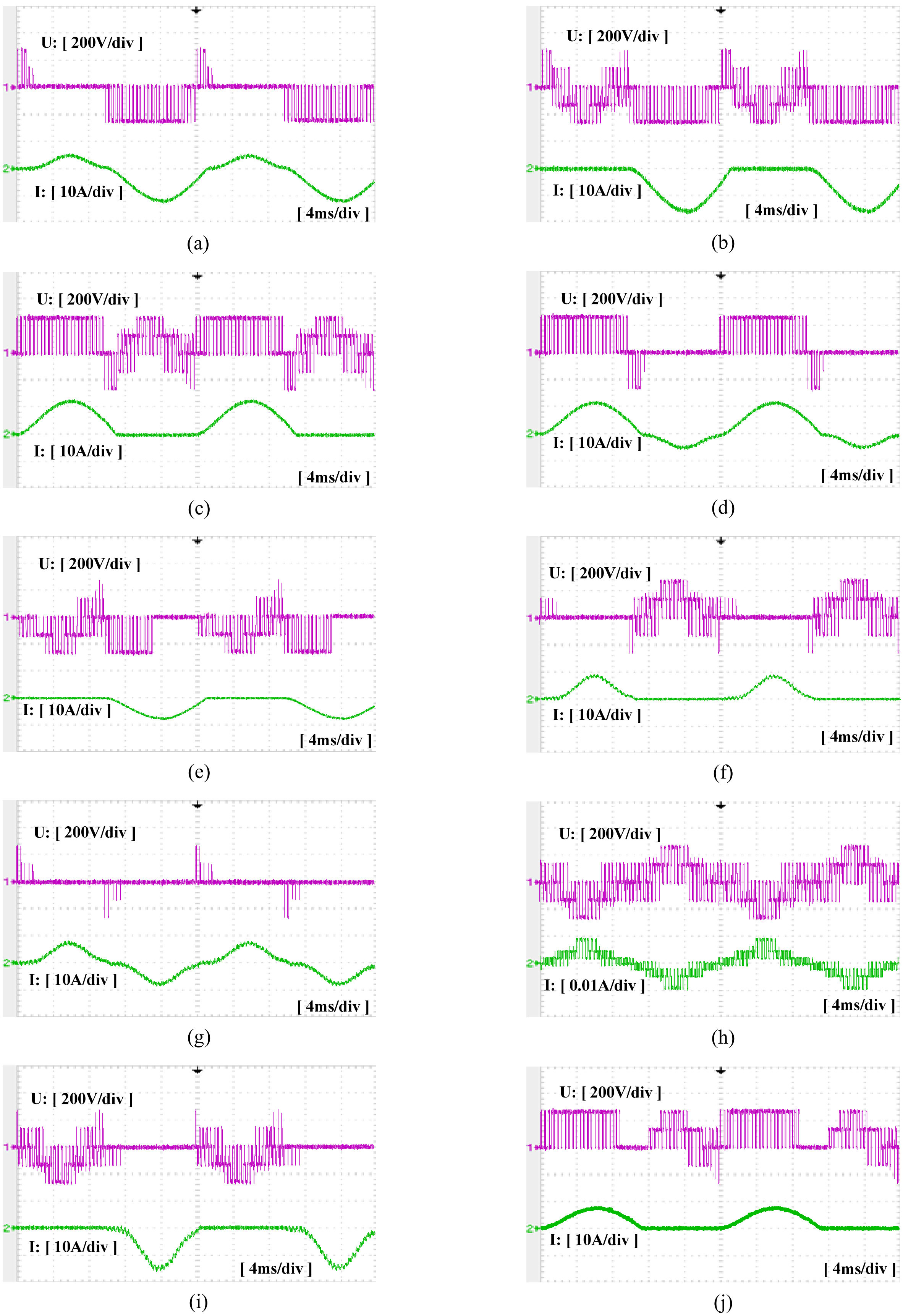

Fig. 9. Experimental waveforms in each fault state.(a) Fault number 1, (b) fault number 2, (c) fault number 3, (d) fault number 4, (e) fault number 5, (f) fault number 6, (g) fault number 7, (h) fault number 8, (i) fault number 9, (j) fault number 10.

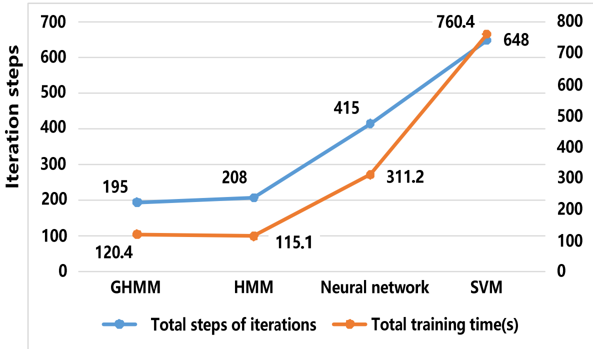

Fig. 10. Iteration steps and training time

TABLE XI
DIAGNOSTIC RESULTS IN EXPERIMENT

Diagnosis method / Fault number	GHMM	HMM	Neural network	SVM
1	100%	84%	69%	79%
2	100%	91%	80%	76%
3	100%	84%	58%	65%
4	96%	92%	75%	80%
5	95%	90%	81%	83%
6	100%	89%	65%	70%
7	100%	83%	71%	73%
8	96%	85%	80%	71%
9	96%	85%	72%	84%
10	100%	86%	76%	80%

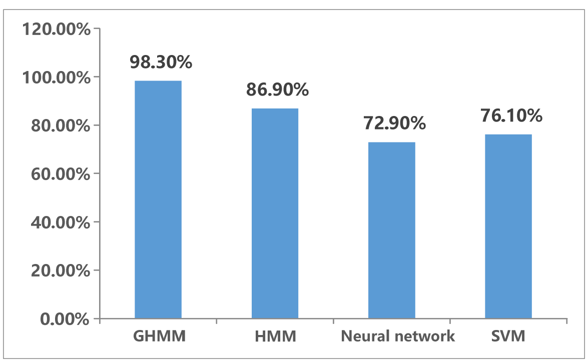

Fig. 11. Correct recognition rate in experiment

faults. Model training and fault diagnosis were completed on a PC. The two other traditional methods, namely, BP neural network and SVM, were also implemented with the MATLAB software.

The number of iteration steps and training time in each method are shown in Fig. 10.

After training all the models, we conducted the diagnostic experiment. We performed multiple diagnoses for each fault to verify the effectiveness of inverter fault diagnosis by using GHMM and eliminate contingency. We selected 100 samples in each fault condition, and a total of 10 × 100 = 1000 test samples were used to carry out fault diagnosis. The experimental results are shown in Table XI.

Fig. 11 shows the comparison of the total correct recognition rates of the four methods in experiment. The experimental results indicated that the simulation results were correct: it was faster and more accurate to use GHMM in diagnosing inverter faults.

Thus, the feasible and advantageous of the GHMM have been proven.

VI. CONCLUSIONS

As a method based on statistical pattern recognition theory， HMM can handle the dynamic process well. Compared with the traditional fault diagnosis method, HMM can monitor and diagnose the faults in the dynamic process of the system, and determine the faults in time. The classical training algorithm (Baum-Welch) in HMM has a fatal flaw, that is, the final solution depends on the initial value, so it is often only a local optimum. Due to the genetic algorithm's use of a global search based on the population, the chance of obtaining the globally optimal solution is much larger.

In this paper, the GHMM is introduced for fault diagnosis of PV inverters. We first use the genetic algorithm to search for the optimal initial value of *B*, and then use Baum-Welch algorithm to train the model. Finally, the Viterbi algorithm is used to identify the fault, and the improved fault recognition rate is obtained. The simulation and experimental results show that it is feasible and effective to use GHMM to diagnose faults of PV inverters. Compared with a traditional HMM, the recognition rate of GHMM is much higher. At the same time, regarding the number of iterations, the training time and the correct recognition rate, GHMM is significantly better than the neural network and SVM models. As a global optimization method, GHMM can deal with dynamic processes very well, and this is especially useful because the working mode of the PV inverter circuit is nonlinear. Therefore, it is of great theoretical and practical value to combine the genetic algorithm with an HMM for PV inverter fault diagnosis.

ACKNOWLEDGMENT

This work is supported in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

REFERENCES

[1] S.Peuget, S. Courtine, and J. P. Rognon, "Fault detection and isolation on a PWM inverter by knowledge-based model," *Industry Applications IEEE Transactions on* , Vol. 34, No. 6, pp. 1318-1326, Nov. 1997.

[2] R.L. de Araujo Ribeiro, C.B. Jacobina, et al. "Fault detection of open-switch damage in voltage-fed PWM

motor drive systems," *IEEE Transactions on Power Electronics*, Vol. 18, No. 2, pp. 587-593, Mar. 2003.

[3] H. Keskes and A. Braham, "DAG SVM and pitch synchronous wavelet transform for induction motor diagnosis," in *Proc. Iet International Conference on Power Electronics, Machines and Drives* , pp. 0166-0166, Apr. 2014.

[4] D.J. Chen and Y.Z. Ye, "Open circuit fault diagnosis method for three level inverter based on multi neural network," *Transactions of China Electrotechnical Society*, Vol. 28, No. 6, pp. 120-126, Jun. 2013.

[5] G.S. Hu, J. Xie, F.F. Zhu, "Classification of power quality disturbances using wavelet and fuzzy support vector machines," in *Proc. International Conference on Machine Learning and Cybernetics* IEEE, pp. 3981-3984, Vol. 7, Aug. 2005.

[6] S. Xu, W.X. Huang, et al, "A novel six phase permanent magnet fault tolerant motor system and its fault diagnosis method," in *Proc. Annual meeting of China Power Supply Society*, May. 2009.

[7] A. Bouzida, et al, "Fault Diagnosis in Industrial Induction Machines Through Discrete Wavelet Transform," *IEEE Transactions on Industrial Electronics*, pp. 370-377, vol.58, No. 9, pp. 4385-4395, Sep. 2011.

[8] M. Pineda-Sanchez, M. Riear-Guasp, et al, "Diagnosis of Induction Motor Faults in the Fractional Fourier Domain," *Instrumentation & Measurement IEEE Transactions on,* Vol. 59, No, 8, pp. 2065-2075, Aug. 2010.

[9] Y. Yin, J. Yang, et al, "Fault Diagnosis of Rolling Bearing Based on Wavelet Packet and Fourier Analysis," *in Proc. Computational Aspects of Social Networks,* pp. 703-706, Sep. 2010.

[10] Z.M. Wu, "Analog circuit fault diagnosis based on information fusion and extreme learning machine," M.S. Thesis, Hunan University, China, 2011.

[11] Y.P. Bao, J. Zheng, et al, "Speech recognition system based on HMM and genetic neural network," *Computer engineering and Science,* Vol. 33, No. 4, pp. 139-144, Apr. 2011.

[12] C.L. Zhang, X. Yue, et al, "Fault Diagnosis of Rotating Machinery Base o`Energy Moment and HMM," *Key Engineering Materials,* Vol. 455, pp. 558-564, Dec. 2010.

[13] H. Ocak, K.A. Loparo, "A new bearing fault detection and diagnosis scheme based on hidden Markov modeling of vibration signals." in *Proc. Acoustics, Speech, and Signal Processing,* Vol. 5, pp. 3141-3144, May. 2001.

[14] X. Yue, "Research on fault diagnosis of complex condition based on HMM," M.S. Thesis, South China University of Technology, China, 2012.

[15] R.F. Han, *Principle and application of genetic algorithm,* Weapon Industry Press, 2010.

[16] J.J. Fu, "Study on the Fault Diagnosis System of Active Neutral Point Clamped Three Level Inverter Based on Neural Network," M.S. Thesis, China Mining University, China, 2016.

[17] F,F. Xie, "Fault Diagnosis Method Based on Support Vector Machine," M.S. Thesis, Hunan University, China, 2006.

[18] A. Widodo, B.S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical Systems & Signal Processing,* Vol. 21, pp. 2560-2574, Aug. 2007.

Hong Zheng was born in Anhui, China, in 1965. He received his B.S and PH.D degrees in Power Electronics and Drives Engineering from Jiangsu University, Jiangsu, China, in 1987 and 2011. From 1987 to 2001, he was a design engineer working in Xi'an Power Electronics Research Institute (PERI), China. Since 2009, he has been working as a professor in School of Electrical and Information Engineering Jiangsu University. His current research interests include power electronic converters, custom power technology and power quality control technology for Smart Grid, distributed generation and energy storage technology, etc.

Ruoyin Wang was born in China, in 1991. He received his B.S. degree from Jiangsu University of Science and Technology, Zhenjiang China, in 2013. He is presently working toward his M.S. degree in Electrical Engineering from Jiangsu University, Zhenjiang, China. His current research interests include power electronics, photovoltaic inverters and electric drives.

Wencheng Xu was born in China, in 1990. He received his B.S. degree from Applied Technology College of Soochow University, Suzhou China, in 2014. He is presently working toward his M.S degree in Electrical Engineering from Jiangsu University, Zhenjiang, China. His current research interests include power electronics and power quality control.

Yifan Wang was born in China, in 1991. He received his B.S. degree from Jiangsu University in 2015. He is presently working toward his M.S. degree in Electrical Engineering from Jiangsu University, Zhenjiang, China. His current research interests include power electronics and DC-DC converters.

Wen Zhu was born in China, in 1992. He received his B.S. degree from Jingjiang College of Jiangsu University in 2015. He is presently working toward his M.S. degree in Electrical Engineering from Jiangsu University, Zhenjiang, China. His current research interests include power electronics and AC-DC.