Method for Parametric Identification of Gaussian Mixture Model Based on Clonal Selection Algorithm

Eugene Fedorov[1[0000-0003-3841-7373]], Valentyna Lukashenko[1[0000-0002-6749-9040]], Tetyana Utkina[1[0000-0002-6614-4133]], Andriy Lukashenko[2[0000-0002-6016-1899]], Kostiantyn Rudakov[1[0000-0003-0000-6077]]

[1] Cherkasy State Technological University, Cherkasy, Shevchenko blvd., 460, 18006, Ukraine
{ckc, k.rudakov, t.utkina}@chdtu.edu.ua, fedorovee75@ukr.net
[2] E. O. Paton Electric Welding Institute, Kyiv, Bozhenko str., 11, 03680, Ukraine
ineks-kiev@ukr.net

Abstract. The problem of increasing the efficiency of parametric identification of Gaussian mixture model (GMM) is considered. A method for identifying GMM parameters based on clonal selection algorithm with preliminary formation of a learning set taking into account the structure of vocal sounds, which increases the likelihood of speaker recognition, is proposed. The parameter identification method is intended for software implementation on the GPU using the CUDA technology, which speeds up the process of parametric identification. This method has been studied on the TIMIT database and serves for intelligent systems of biometric personal identification.

Keywords: quasi-periodic signal, speaker recognition, gaussian mixture model, learning set formation, parametric identification, clonal selection algorithm.

1 Introduction

Automated biometric identification of a person means making decisions based on acoustic and visual information, which improves the quality of recognition of the person under study. Unlike the traditional approach [1], computer biometric identification speeds up and improves the accuracy of the recognition process, which is especially critical in the conditions of limited time.

A special class of biometric identification of a person is formed by methods based on the analysis of acoustic information [2].

The well-known voice identification methods, such as dynamic programming [3-4], analyze the entire signal, which increases the likelihood of recognition, but require entire storage of learning signals and long comparison of the analyzed signal with all learning signals. Other well-known methods, such as vector quantization [5-6], neural network method [7-8] and decision tree [9], store only generalized characteristics of learning signals and carry out a quick classification of signal components, but analyze these components without interconnecting them, which reduces the recognition probability. Network methods, including those that use Gaussian mixture models (GMM)

[10-13], which quickly analyze the entire signal and store only generalized characteristics of learning signals, are a compromise between the above groups of methods. At the same time, GMM-based methods use the identification of its parameters based on local search, which reduces the probability of recognition.

The aim of the work is to increase the efficiency of parametric identification method of Gaussian mixture model (GMM) due to a metaheuristic clonal selection algorithm, with preliminary formation of a learning set.

To achieve this goal, it is necessary to solve the following tasks:

1. to develop a method for learning set formation;
2. to create a method of GMM parametric identification;
3. to develop an algorithm of GMM parametric identification;
4. to conduct a numerical study of the proposed method of parametric identification.

2 Formal problem statement

Let a learning set be defined for a particular speaker.

Then the problem of increasing the probability of speaker recognition by Gaussian mixture model (GMM) is presented as the problem of finding such a vector of parameters for this model that satisfies the criterion.

3 Literature review

Biometric identification methods based on dynamic programming [3-4] compare the signal representing the recognizable speaker with all signals representing the well-known speakers. The advantage of these methods lies in their ability to analyze the entire signal and compare signals of different lengths, which increases the probability of recognition. The disadvantage of these methods consists in the entire storage of signals representing the well-known speakers and the duration of the procedure for comparison of the recognizable signal with all available signals.

Biometric identification methods based on vector quantization [5-6] form the codebook vectors as averaged components of the signals representing the well-known speakers and compare components of the signal representing the recognizable speaker with these vectors. The advantage of these methods lies in the storage of only the codebook vectors and quick comparison of the recognizable signal components with these vectors. The disadvantage of these methods consists in the ability to analyze only single components of the signal without interconnecting them, which reduces the likelihood of recognition.

Biometric identification methods based on artificial neural networks [7-8] identify the neural network parameters, which store information about signal components representing the well-known speakers and recognize signal components representing the recognizable speaker through the neural network. The advantage of these methods lies in the storage of only the neural network parameters and quick classification of the recognizable signal components. The disadvantage of these methods consists in

the identification of neural network parameters based on local search and the ability to analyze only single components of the signal without interconnecting them, which reduces the likelihood of recognition.

Biometric identification methods based on decision trees [9] automatically identify threshold values (for example, formants) at decision tree nodes representing speech characteristics of the well-known speakers and compare threshold values of the signal component representing the recognizable speaker with threshold values of the decision tree. The advantage of these methods lies in the storage of only threshold values and quick classification of the recognizable signal components. The disadvantage of these methods consists in the complexity and subjectivity of threshold values formation and the ability to analyze only single components of the signal without interconnecting them, which reduces the likelihood of recognition.

Biometric identification methods based on Gaussian mixture models (GMMs) [10-13] identify GMM parameters, which store information about signal components representing the well-known speakers, and recognize the signal representing the recognizable speaker by means of these GMMs. The advantage of these methods lies in the ability to quickly recognize the entire signal and store only GMM parameters. The disadvantage of these methods consists in the identification of GMM parameters based on local search, which reduces the likelihood of recognition.

A common feature of all methods listed above consists in the following: they form a learning set of signal patterns and analyze a signal without taking into account the structure of vocal sounds, which leads to a decrease in the probability of recognition.

Therefore, the increase of the efficiency of the method of GMM parametric identification with preliminary formation of a learning set, taking into account the structure of vocal sounds, is an urgent task.

4 Method of learning set formation

The method of learning set formation includes the following steps:

- partitioning of a quasi-periodic signal into discrete patterns;
- shifting of discrete patterns in time and amplitude;
- interpolation of discrete patterns;
- shifting and scaling of continuous patterns in time;
- shifting and scaling of continuous patterns in amplitude;
- sampling of continuous patterns.

1. Partitioning of a quasi-periodic signal into discrete patterns

Let's define a finite set of discrete patterns of a quasi-periodic signal, described by a family of integer bounded finite discrete functions $X = \{x_i \mid i \in \{1,...,I\}\}$, in the form:

$$x_i(n) = \begin{cases} f(n), & n \in \{N_i^{\min},...,N_i^{\max}\} \\ 0, & n \notin \{N_i^{\min},...,N_i^{\max}\} \end{cases}, \; i \in \{1,...,I\},$$

$$A_i^{\min} = \min_n f(n),\ n \in \{N_i^{\min},...,N_i^{\max}\},\ i \in \{1,...,I\},$$

$$A_i^{\max} = \max_n f(n),\ n \in \{N_i^{\min},...,N_i^{\max}\},\ i \in \{1,...,I\},$$

where $A_i^{\min}, A_i^{\max}$ – the minimum and maximum values of the function x_i on the compact $\{N_i^{\min},...,N_i^{\max}\}$.

2. Shifting of discrete patterns in time and amplitude

Let's define a finite set of discrete patterns shifted in time and amplitude, which are described by a finite family of integer bounded finite discrete functions $X^S = \{x_i^S \mid i \in \{1,...,I\}\}$, in the form:

$$x_i^S(n) = \begin{cases} x_i(n + N_i^{\min}) - A_i^{\min}, & n \in \{0,...,N_i\} \\ 0, & n \notin \{0,...,N_i\} \end{cases},\ i \in \{1,...,I\},$$

$$N_i = N_i^{\max} - N_i^{\min},\ A_i = A_i^{\max} - A_i^{\min}.$$

3. Interpolation of discrete patterns

A linear interpolation, which requires the least computational complexity, has been chosen in the article. Let's define a finite set of continuous patterns, obtained as a result of linear interpolation and described by a finite family of real-valued bounded finite continuous functions $\Psi = \{\psi_i \mid i \in \{1,...,I\}\}$, in the form:

$$\forall t \in [-\Delta t, T_i + \Delta t]\ \ \psi_i(t) =$$

$$= \sum_{n=-1}^{N_i} \chi_{(t_n,t_{n+1})}(t)\left(x_i^S(n) + \frac{x_i^S(n+1) - x_i^S(n)}{\Delta t}(t - t_n) \right) + \sum_{n=-1}^{N_i+1} \chi_{\{t_n\}}(t) x_i^S(n),$$

$$i \in \{1,...,I\},\ \forall t \notin [-\Delta t, T_i + \Delta t]\ \ \psi_i(t) = 0,\ T_i = N_i \Delta t,\ t_n = n\Delta t,$$

$$\chi_B(t) = \begin{cases} 1, & t \in B \\ 0, & t \notin B \end{cases} \text{ – indicator function,}$$

where Δt – the quantization step in time.

4. Shifting and scaling of continuous patterns in time

Let's define a finite set of shifted and scaled in time continuous patterns, described by a finite family of real-valued bounded finite continuous functions $\Psi^S = \{\psi_i^S \mid i \in \{1,...,I\}\}$, in the form:

$$\forall t \in [\breve{T}^{\min}, \breve{T}^{\max}] \;\; \psi_i^s(t) = \psi_i \left(T_i \frac{t - \breve{T}^{\min}}{\breve{T}^{\max} - \breve{T}^{\min}} \right),$$

$$\forall t \notin [\breve{T}^{\min}, \breve{T}^{\max}] \;\; \psi_i^s(t) = 0,$$

where $[\breve{T}^{\min}, \breve{T}^{\max}]$ – the compact, set and single for all patterns.

The article proposes to define $\breve{T}^{\min}, \breve{T}^{\max}$ as follows:

$$\breve{T}^{\min} = \Delta t,$$

$$\breve{T}^{\max} = round\left(\frac{f_d}{f_{\min}} \right) \Delta t,$$

where f_d – the sampling frequency in Hz (for vocal sounds 8 kHz is enough), $f_{\min}$ – the minimum frequency of frequency range of speech sound (for vocal sounds 50 Hz is enough), $round()$ – the function, rounding the number to the nearest integer.

5. Shifting and scaling of continuous patterns in amplitude

Let's define a finite set of continuous patterns shifted and scaled in amplitude, which are described by a finite family of real-valued bounded finite continuous functions $\Psi^{ss} = \{\psi_i^{ss} \mid i \in \{1,...,I\}\}$, in the form:

$$\forall t \in [\breve{T}^{\min}, \breve{T}^{\max}] \;\; \psi_i^{ss}(t) = \breve{A}^{\min} + \frac{\breve{A}^{\max} - \breve{A}^{\min}}{\tilde{A}_i^{\max} - \tilde{A}_i^{\min}} \psi_i^s(t),$$

$$\forall t \notin [\breve{T}^{\min}, \breve{T}^{\max}] \;\; \psi_i^{ss}(t) = 0,$$

$$\tilde{A}_i^{\max} = \max_t \psi_i^s(t),$$

$$\tilde{A}_i^{\min} = \min_t \psi_i^s(t),$$

$$t \in [\breve{T}^{\min}, \breve{T}^{\max}],$$

where $\tilde{A}_i^{\min}, \tilde{A}_i^{\max}$ – the minimum and maximum values of the function ψ_i^s on the compact $[\breve{T}^{\min}, \breve{T}^{\max}]$; $\breve{A}^{\min}, \breve{A}^{\max}$ – the minimum and maximum values, set and single for all patterns, on a given compact $[\breve{T}^{\min}, \breve{T}^{\max}]$.

The article proposes to define $\breve{A}^{\min}, \breve{A}^{\max}$ as follows: $\breve{A}^{\min} = 0$, $\breve{A}^{\max} = 2^b - 1$, where b – the number of level quantization bits.

6. Sampling of continuous patterns

Let's define a finite set of discrete patterns, obtained from continuous ones by sampling in time and described by a finite family of integral bounded finite discrete functions $S = \{s_i \mid i \in \{1,...,I\}\}$, in the form:

$$s_i(n) = round(\psi_i^{ss}(n\Delta t)),\ n \in \{\breve{N}^{\min},...,\breve{N}^{\max}\},$$

$$\breve{N}^{\min} = \breve{T}^{\min} / \Delta t,\ \breve{N}^{\max} = \breve{T}^{\max} / \Delta t,\ N = \breve{N}^{\max} - \breve{N}^{\min}.$$

Each received discrete sample is considered as a feature vector and is located in a single amplitude-time window.

5 Method for parametric identification of Gaussian mixture model

Each GMM associated with a particular speaker is defined on the basis of the total probability formula as follows:

$$P(s) = \sum_{k=1}^{K} P(k)\, p(s \mid k),$$

$$p(s \mid k) = \frac{1}{\sqrt{(2\pi)^N \det C_k}} \exp\left(-\frac{1}{2}(s-m_k)^T C_k^{-1}(s-m_k)\right),$$

where $P(k)$ – a priori unconditional probabilities (weights of mixture components), $p(s \mid k)$ – densities of multi-dimensional Gaussian distribution, $m_k = (m_{k1},...,m_{kN})$ – vector of mathematical expectations, C_k – covariance matrix, K – the number of mixture components.

In the work, for GMM parametric identification the target function, which means the choice of such values of parameter vector Θ that deliver likelihood functions to the maximum logarithm, is chosen:

$$F = \ln P(S \mid \Theta) = \sum_{i=1}^{I} \ln \sum_{k=1}^{K} P(k)\, p(s_i \mid k) \to \max_{\Theta},$$

where $\Theta = ((P(1),...,P(K)),(m_1,...,m_K),(C_1,...,C_K))$ – GMM parameters vector.

Since the traditional EM algorithm used for GMM parametric identification implements only local search, so currently for GMM parametric identification are actively applied evolutionary, swarm and immune metaheuristics, which use the parameter vector Θ as an individual of the population [14-16]. However, due to the large dimensionality of the covariance matrix C, only a diagonal matrix C is used in these metaheuristics, which reduces the likelihood of identification. Therefore, in this work, the joint probability vector $(P(s_1,1),...,P(s_i,k),...,P(s_I,K))$, where

$P(s_i, k) = P(k)\,p(s_i \mid k)$, is used as an individual of the population, which allows to work with the off-diagonal matrix C.

GMM parametric identification is based on clonal selection algorithm proposed in [17-23] and is presented in the following form:

1. The number of mixture components K, the mutation parameter β, the maximum number of iterations $\tau^{\max}$ are set.
2. The intermediate population $U = \{u\}$ of the power $\mu + \lambda + \gamma$ is created, each antibody of the population being represented as $u = (u_{11}, ..., u_{IK})$.

   Each element u_{ik} is defined as $u_{ik} = rand()$, where $rand()$ – a function that returns a uniformly distributed random number in the range $[0,1]$.

   A finite set of discrete patterns S is given.
3. The iteration number $\tau = 1$ is set, the maximum value of the target function $b^{old} = 0$.
4. A posteriori conditional probabilities for each antibody u are calculated:

$$P(k \mid s_i) = \frac{u_{ik}}{\sum_{z=1}^{K} u_{iz}},\ k \in \{1,.,K\},\ i \in \{1,...,I\}.$$

5. A priori unconditional probabilities (weights of the mixture components) for each antibody u are calculated $P(k) = \frac{1}{I}\sum_{i=1}^{I} P(k \mid s_i)$, $k \in \{1,.,K\}$.
6. Expectation vectors for each antibody u are calculated:

$$m_k = \frac{\sum_{i=1}^{I} P(k \mid s_i) s_i}{\sum_{i=1}^{I} P(k \mid s_i)},\ k \in \{1,.,K\}.$$

7. Covariance matrices for each antibody u are calculated:

$$C_k = \frac{\sum_{i=1}^{I} P(k \mid s_i)(s_i - m_k)^T (s_i - m_k)}{\sum_{i=1}^{I} P(k \mid s_i)},\ k \in \{1,.,K\}.$$

8. The densities of multidimensional Gaussian distribution for each antibody u are calculated:

$$p(s_i \mid k) = \frac{1}{\sqrt{(2\pi)^N \det C_k}} \exp\left(-\frac{1}{2}(s_i - m_k)^T C_k^{-1} (s_i - m_k)\right),\ k \in \{1,.,K\},\ i \in \{1,...,I\}.$$

9. The total probabilities for each antibody u are calculated $P(s_i) = \sum_{k=1}^{K} P(k)\, p(s_i \mid k)$.
10. The value of the target function for each antibody u is calculated:

$$F(u) = \sum_{i=1}^{I} \ln P(s_i).$$

11. The intermediate population $U = \{u\}$ is ordered in descending of the target function value.
12. The current population $H = \{h\}$ of the power μ is created by means of reproduction and replacement operators, with each antibody of this population being represented as $h = (h_{11}, ..., h_{IK})$.

    As antibodies of the current population $H = \{h\}$, the first μ (the best in the target function) antibodies of the intermediate population $U = \{u\}$ are taken. This corresponds to the reduction operator with a selection scheme that provides the search directionality (the best antibodies are preserved). Replacement (randomly generated) antibodies can be among the best antibodies. This corresponds to the replacement operator.
13. The value of the target function of the first antibody of the current population $H = \{h\}$ is set by the minimum value of the target function a.
14. The value of the target function of the last antibody of the current population $H = \{h\}$ is set by the maximum value of the target function b.
15. If $|b - b^{old}| > \varepsilon$ and $\tau < \tau^{\max}$, then $\tau = \tau + 1$, $b^{old} = b$, going to 16, otherwise parametric identification is complete.
16. The affinity for each antibody h is calculated.

    The affinity determines the proximity of the current antibody to the best one and is calculated based on the utility function in the form:

$$\Phi(h) = \frac{F(h) - a}{b - a}.$$

    If $\Phi(h) = 1$, then the *i*-th antibody is the best one.

    If $\Phi(h) = 0$, then the *i*-th antibody is the worst one.
17. The mutation probability for each antibody h is calculated $p(h) = \exp\left(-\beta\Phi(h)\right)$.

    The larger β, the less likely is the mutation probability.
18. A set of clones $C = \{c\}$ of the power λ is created by means of a cloning operator, each clone of this set being represented as $c = (c_{11}, ..., c_{IK})$.

    The cloning operator plays a role similar to the reproduction operator of genetic algorithm and is applied to the current population $H = \{h\}$. The number of clones for each antibody h is defined as λ/μ.

19. A set of mutated clones $\widehat{C} = \{\widehat{c}\}$ of the power λ is formed by means of the proposed mutation operator, each mutated clone of this set being represented in the form $\widehat{c} = (\widehat{c}_{11},...,\widehat{c}_{IK})$.

    The mutation operator allows to obtain new antibodies with sharply different properties from antibody clones.

    Each element $\widehat{c}_{ik}$ is defined as:

$$r_{ik} = rand(),\ \widehat{c}_{ik} = \begin{cases} c_{ik}, & r_{ik} \geq p(h) \\ round(), & r_{ik} < p(h) \end{cases},\ k \in \{1,.,K\},\ i \in \{1,...,I\}.$$

    The features of the proposed variant of the mutation operator are the following: it does not require the use of binary potential solutions, i.e. it doesn't need to convert probabilistic potential solutions into binary ones before the mutation and to convert binary potential solutions into probabilistic ones after the mutation, which reduces the computational complexity of the mutation operator and allows to quickly find a solution.

20. A set of replacement antibodies $D = \{d\}$ of the power γ is created, each antibody of the set is represented as $d = (d_{11},...,d_{IK})$.

    Each element d_{ik} is defined as $d_{ik} = rand()$.

21. The intermediate population $U = \{u\}$ of the power $\mu + \lambda + \gamma$ is formed.

    The antibodies of the current population $H = \{h\}$ and a set of mutated clones $\widehat{C} = \{\widehat{c}\}$ are taken as the first $\mu + \lambda$ antibodies of the intermediate population $U = \{u\}$. The antibodies of a set of the replacement antibodies $D = \{d\}$ are taken as the last γ antibodies of the intermediate population $U = \{u\}$. Going to 4.

    As a result of the method, GMM parameters will be defined.

6 Creation of an algorithm for parametric identification of Gaussian mixture model based on clonal selection algorithm

The algorithm for GMM parametric identification based on clonal selection algorithm, which is intended for implementation on the GPU using the CUDA technology, is presented in Fig. 1. This block diagram functions as follows.

Step 1 – Set the number of mixture components K, the mutation parameter β, the maximum number of iterations $\tau^{\max}$.

Step 2 – Set the intermediate population U and a finite set of discrete patterns S.

Step 3 – Set the iteration number $\tau = 1$, the maximum value of the target function $b^{old} = 0$.

Step 4 – Calculate the total probabilities $P(s_i)$ for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot (\mu + \lambda + \gamma)$ one-dimensional

blocks. In each block, based on the reduction, a sum from K elements of the form u_{ik} is calculated.

Step 5 – Calculate a posteriori conditional probabilities for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot K \cdot (\mu + \lambda + \gamma) / N^S$ one-dimensional blocks, where N^S is the number of threads in the block. Each thread computes $P(k \mid s_i) = u_{ik} / P(s_i)$.

Step 6 – Calculate the sum of a posteriori conditional probabilities q_k for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma)$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $P(k \mid s_i)$ is calculated.

Step 7 – Calculate a priori unconditional probabilities (weights of the mixture components) $P(k)$ for each antibody u using $K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma) / N^S$ one-dimensional blocks. Each thread computes $P(k) = q_k / I$.

Step 8 – Calculate expectation vectors m_k for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma)$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $P(k \mid s_i) s_i / q_k$ is calculated.

Step 9 – Calculate the covariance matrices C_k for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $K \cdot (\mu + \lambda + \gamma)$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $P(k \mid s_i)(s_i - m_k)^T (s_i - m_k) / q_k$ is calculated.

Step 10 – Calculate the densities of Gaussian multidimensional distribution $p(s_i \mid k)$ for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot K \cdot (\mu + \lambda + \gamma) / N^S$ one-dimensional blocks. Each thread computes:

$$p(s_i \mid k) = \frac{1}{\sqrt{(2\pi)^N \det C_k}} \exp\left(-\frac{1}{2} (s_i - m_k)^T C_k^{-1} (s_i - m_k) \right).$$

Step 11 – Calculate the total probabilities $P(s_i)$ for each antibody u using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot (\mu + \lambda + \gamma)$ one-dimensional blocks. In each block, based on the reduction, a sum from K elements of the form $P(k) p(s_i \mid k)$ is calculated.

Step 12 – Calculate the target function values $F(u)$ for each antibody u using $I \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $\mu + \lambda + \gamma$ two-dimensional blocks. In each block, based on the reduction, a sum from I elements of the form $\ln P(s_i)$ is calculated.

Step 13 – Order the target functions by the value based on parallel sorting by merging into the intermediate population U using $\mu+\lambda+\gamma$ threads that are grouped into one one-dimensional block.

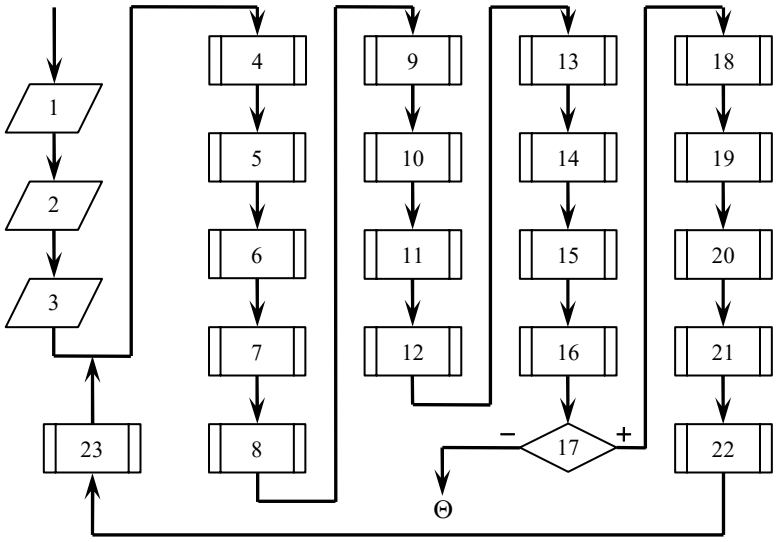

Fig. 1. Block diagram of the algorithm for parametric identification of Gaussian mixture model based on clonal selection algorithm.

Step 14 – Create the current population $H=\{h\}$ of the power μ by means of element-by-element μ copying of the first antibodies of the intermediate population $U=\{u\}$ using $I\cdot K\cdot\mu$ threads that are grouped into $I\cdot K\cdot\mu/N^s$ one-dimensional blocks.

Step 15 – Set the value of the target function of the first antibody of the current population $H=\{h\}$ by the minimum value of the target function a.

Step 16 – Set the value of the target function of the last antibody of the current population $H=\{h\}$ by the maximum value of the target function b.

Step 17 – If $|b-b^{old}|>\varepsilon$ and $\tau<\tau^{\max}$, then $\tau=\tau+1$, $b^{old}=b$, going to 18, otherwise parametric identification is completed.

Step 18 – Calculate the affinity for each antibody h using $\mu+\lambda+\gamma$ threads that are grouped into one one-dimensional block. Each thread computes:

$$\Phi(h)=\frac{F(h)-a}{b-a}.$$

Step 19 – Calculate the mutation probability for each antibody h using λ threads that are grouped into one one-dimensional block $p(h)=\exp\left(-\beta\Phi(h)\right)$.

Step 20 – Create a set of clones $C = \{c\}$ of the power λ by means of element-by-element copying of the antibodies of the current population $H = \{h\}$ using $I \cdot K \cdot \lambda$ threads that are grouped into $I \cdot K \cdot \lambda / N^S$ one-dimensional blocks.

Step 21 – Form a set of mutated clones $\widehat{C} = \{\hat{c}\}$ of the power λ by means of the proposed mutation operator, using $I \cdot K \cdot \lambda$ threads that are grouped into $I \cdot K \cdot \lambda / N^S$ one-dimensional blocks. Each thread computes:

$$r_{ik} = rand(),\ \hat{c}_{ik} = \begin{cases} c_{ik}, & r_{ik} \geq p(h) \\ round() & r_{ik} < p(h) \end{cases}.$$

Step 22 – Create a set of random antibodies $D = \{d\}$ of the power γ using $I \cdot K \cdot \gamma$ threads that are grouped into $I \cdot K \cdot \gamma / N^S$ one-dimensional blocks. Each thread computes $d_{ik} = rand()$.

Step 23 – Generate the intermediate population $U = \{u\}$ of the power $\mu + \lambda + \gamma$ by means of element-by-element copying of the antibodies of the current population $H = \{h\}$, a set of mutated clones $\widehat{C} = \{\hat{c}\}$, a set of replacement antibodies $D = \{d\}$, using $I \cdot K \cdot (\mu + \lambda + \gamma)$ threads that are grouped into $I \cdot K \cdot (\mu + \lambda + \gamma) / N^S$ one-dimensional blocks. Going to 4.

7 Experiments and results

Numerical experiments were carried out using the CUDA technology of parallel processing of information on the GeForce 920M video card with the number of threads in the block $N^S = 1024$. Parametric identification was carried out for each of the 100 GMMs, corresponding to 100 speakers. In the work it has been accepted that the power of a set of patterns of vocal speech sounds $I = 1024$, the power of the current population $\mu = 20$, the power of a set of clones $\lambda = 1000$, the power of a set of random antibodies $\gamma = 0.2, \mu = 4$, the mutation parameter $\beta = 2.3$, the maximum number of iterations $\tau^{\max} = 1000$, the parameter $\varepsilon = 10^{-6}$.

The dependence of the probability of incorrect recognition of the speaker on the number of components of the mixture is shown in Fig.2. This dependence shows that the probability of incorrect recognition of the speaker decreases with increasing number of components of the mixture.

Table 1 presents the speaker's recognition probabilities obtained from the TIMIT database using a neural network based on radial basis functions (RBFNN), trained on the basis of error correction with the MFCC feature system, Gaussian mixture model (GMM), trained on the basis of EM-algorithm with the MFCC feature system, Gaussian mixture model, trained on the basis of clonal selection algorithm, with features obtained by means of the proposed method of learning set formation.

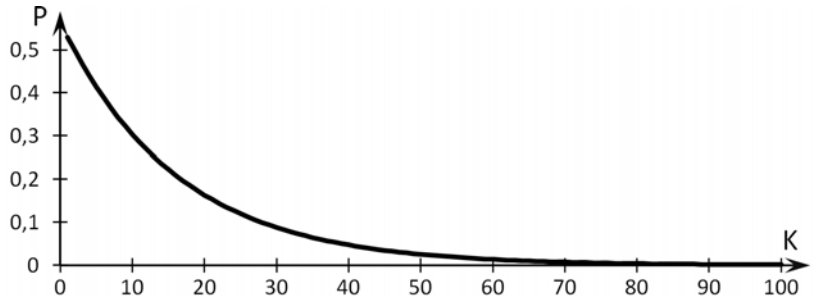

Fig. 2. The dependence of the probability of incorrect recognition of the speaker on the number of components of the mixture.

According to Table 1, the best results are given by GMM, trained on the basis of clonal selection algorithm with features obtained by means of the proposed method of learning set formation.

Table 1. Speaker recognition probability.

Model + learning algorithm + feature system	Probability of recognition
RBFNN + error correction + MFCC	0.81
GMM + EM-algorithm + MFCC	0.92
GMM + proposed method + proposed feature system	0.98

This is due to the fact that error correction and the EM algorithm only perform a local search, which increases the likelihood of falling into a local extremum, and learning set formation based on the MFCC is done without taking into account the structure of vocal sound.

8 Conclusion

The article deals with the problem of increasing the efficiency of parametric identification of Gaussian mixture model (GMM). The method of learning set formation, which uses shifting, scaling, interpolation and sampling of signal patterns that correspond to quasi-periods to locate them in a single amplitude-time window, which allows to take into account the quasi-periodic signal structure and increase the probability of speaker recognition, has been improved. The method of GMM parametric identification, which is based on clonal selection algorithm that allows to reduce the likelihood of falling into a local extremum, has reached further development. The proposed method allows probabilistic individuals of the population (potential solutions) in the mutation operator, which accelerates parametric identification (it is not necessary to perform the operations of converting real individuals into binary ones and vice versa), and uses not the parameter vector, but the vector of joint probabilities as an individual of the population, which allows to work with a non-diagonal covariance

matrix and increases the likelihood of speaker recognition. The algorithm of GMM parametric identification, which is intended for software implementation on the GPU using the CUDA technology that speeds up the GMM learning process, has been created. The software that implements the proposed algorithm has been developed and investigated on the TIMIT database. The experiments have confirmed the operability of the developed software and allow to recommend it for use in practice. Prospects for further research are to test the proposed methods on a wider set of test databases.

References

1. Singh, N., Khan, R. A., Shree, R.: Applications of Speaker Recognition. In: Rajesh, R., Ganesh, K. (eds.). International Conference on Modelling Optimization and Computing, vol. 38, pp. 3122–3126. Elsevier Procedia Engineering (2012) doi: 10.1016/j.proeng.2012.06.363
2. Campbell, J. P.: Speaker Recognition: a tutorial. Proceedings of the IEEE 85(9), 1437–1462 (1997) doi: 10.1109/5.628714
3. Togneri, R., Pullela, D.: An Overview of Speaker Identification: Accuracy and Robustness Issues. IEEE Circuits And Systems Magazine 11(2), 23–61 (2011) doi: 10.1109/MCAS.2011.941079
4. Beigi, H.: Fundamentals of Speaker Recognition, Springer, New York (2011) doi: 10.1007/978-0-387-77592-0
5. Reynolds, D. A.: An Overview of Automatic Speaker Recognition Technology. In: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 4072–4075. IEEE, Orlando, FL, USA (2002) doi: 10.1109/ICASSP.2002.5745552
6. Kinnunen, T., Li, H.: An Overview of Text-Independent Speaker Recognition: From Features to Supervectors. Speech Communication 52(1), 12-40 (2010) doi: 10.1016/j.specom.2009.08.009
7. Reynolds, D. A., Rose, R. C.: Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models. IEEE Transactions on Speech and Audio Processing 3(1), 72–83 (1995) doi: 10.1109/89.365379
8. Zeng, F.-Z., Zhou, H.: Speaker Recognition based on a Novel Hybrid Algorithm. In: 25th International Conference on Parallel Computational Fluid Dynamics, vol. 61, pp. 220–226. Elsevier Procedia Engineering (2013) doi: 10.1016/j.proeng.2013.08.007
9. Jeyalakshmi, C., Krishnamurthi, V., Revathi, A.: Speech Recognition of Deaf and Hard of Hearing People Using Hybrid Neural Network. In: 2010 2nd International Conference on Mechanical and Electronics Engineering, pp. 83–87. IEEE, Kyoto, Japan (2010) doi: 10.1109/ICMEE.2010.5558589
10. Rabiner, L. R., Jang, B. H.: Fundamentals of speech recognition, Prentice Hall, NJ, USA (1993)
11. Nayana, P. K., Mathew, D., Thomas, A.: Comparison of Text Independent Speaker Identification Systems Using GMM and i-Vector Methods. In: 7th International Conference on Advances in Computing & Communications, pp. 47–54. Elsevier

Procedia Engineering, Cochin, India (2017) doi: doi.org/10.1016/j.procs.2017.09.075
12. Chauhan, V., Dwivedi, Sh., Karale, P., Potdar, S. M.: Speech to text converter using Gaussian Mixture Model (GMM), International Research Journal of Engineering and Technology (IRJET) 3(5), 160–164 (2016)
13. Reynolds, D. A.: Automatic Speaker Recognition Using Gaussian Mixture Speaker Models. The Lincoln Laboratory Journal 8(2), 173–195 (1995)
14. Lin, L., Wang, Sh.: A New Genetically Optimized GMM for Speaker . In: Published in 6th World Congress on Intelligent Control and Automation, pp. 10235–10239 (2006) doi: 10.1109/WCICA.2006.1714005
15. Zablotskiy, S., Pitakrat, T., Zablotskaya, K., Minker, W.:GMM Parameter Estimation by Means of EM and Genetic Algorithms. In: Human-Computer Interaction. Design and Development Approaches: 14th international conference on Human-computer interaction: design and development approaches, pp. 527–536. Proceeding, Orlando, FL (2011) doi: 10.1007/978-3-642-21602-2_57
16. Saeidi, R., Mohammadi, H. R. S., Ganchev, T., Rodman, R. D.: Particle Swarm Optimization for Sorted Adapted Gaussian Mixture Models. IEEE Transactions on Audio, Speech, and Language Processing 17(2), 344–353 (2009) doi: 10.1109/TASL.2008.2010278
17. de Castro, L. N., von Zuben, F. J.: The Clonal Selection Algorithm with Engineering Applications. In: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '00), Workshop on Artificial Immune Systems and Their Applications, pp. 36-39. Las Vegas, Nv (2000)
18. de Castro, L. N., von Zuben, F. J.: Learning and Optimization Using Clonal Selection Principle. IEEE Transactions on Evolutionary Computation 6(3), 239–251 (2002) doi: 10.1109/TEVC.2002.1011539
19. Babayigit, B. A., Guney, K., Akdagli, A.: Clonal Selection Algorithm for Array Pattern Nulling by Controlling the Positions of Selected Elements. Progress in Electromagnetics Research B 6, 257–266 (2008) doi: 10.2528/PIERB08031218
20. White, J. A., Garrett, S. M.: Improved Pattern Recognition with Artificial Clonal Selection? In: Timmis J., Bentley P. J., Hart E. (eds). 2nd International Conference on Artificial Immune Sysyems, vol. 2787, pp. 181–193. Springer, Berlin (2003) doi: 10.1007/978-3-540-45192-1_18
21. Alba, E., Nakib, A., Siarry, P.: Metaheuristics for Dynamic Optimization, Springer-Verlag, Berlin (2013)
22. Du, K.-L., Swamy, M. N. S.: Search and Optimization by Metaheuristics. Techniques and Algorithms Inspired by Nature, Springer, Birkhäuser Basel (2016) doi: 10.1007/978-3-319-41192-7
23. Brownlee, J.: Clever Algorithms: Nature-Inspired Programming Recipes, Melbourne, Brownlee (2012)