Sergey Subbotin
Editor

Computer Modeling and Intelligent Systems

Proceedings of The Third International Workshop
CMIS-2020

Zaporizhzhia, Ukraine
April 27- May 1, 2020

CEUR-WS.org/Vol-2608/preface.pdf

Subbotin S., (Ed.): The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020). Zaporizhzhia, Ukraine, April 27 - May 1, 2020, CEUR-WS.org, online

This volume represents the proceedings of the Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020), held in Zaporizhzhia, Ukraine, in April-May, 2020. It comprises 83 contributed papers that were carefully peer-reviewed and selected from 177 submissions.

Copyright © 2020 for the individual papers by the papers' authors.
Copying permitted only for private and academic purposes.
This volume is published and copyrighted by its editors.

Programme Committee

Chair

Prof. **Sergey Subbotin**, National University "Zaporizhzhia Polytechnic", Ukraine

Members

Prof. **Yevgeniy Bodyanskiy**, Kharkiv National University of Radio Electronics, Ukraine
Prof. **Martin Drahansky,** Brno University of Technology, Czech Republic
Prof. **Eugene Fedorov,** Cherkasy State Technological University, Ukraine
Prof. **Iurii Krak**, Taras Shevchenko National University of Kyiv, Ukraine
Prof. **Vitaly Levashenko**, University of Zilina, Slovak Republic
Prof. **David Luengo**, Universidad Politecnica de Madrid, Spain
Prof. **Volodymyr Lytvynenko**, Kherson National Technical University, Ukraine
Prof. **Nataliya Pankratova,** National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine
Prof. **Anatoliy Sachenko**, Ternopil National Economic University, Ukraine
Prof. **Natalya Shakhovska**, Lviv Polytechnic National University, Ukraine
Prof. **Serhiy Shtovba,** Vinnytsia National Technical University, Ukraine
Prof. **Vitaliy Snytyuk,** Taras Shevchenko National University of Kyiv, Ukraine
Prof. **Carsten Wolff**, Dortmund University of Applied Sciences and Arts, Germany
Prof. **Karsten Henke**, Ilmenau University of Technology, Germany
Prof. **Elena Zaitseva**, University of Zilina, Slovak Republic
Dr. **Peter Arras**, Katholieke Universiteit Leuven, Belgium
Dr. **Ivan Izonin**, Lviv Polytechnic National University, Ukraine
Dr. **Igor Korobiichuk,** Warsaw University of Technology, Poland
Dr. **Anzhelika Parkhomenko**, National University "Zaporizhzhia Polytechnic", Ukraine
Dr. **Andrii Oliinyk,** National University "Zaporizhzhia Polytechnic", Ukraine
Dr. **Galyna Tabunshchyk**, National University "Zaporizhzhia Polytechnic", Ukraine
Dr. **Heinz-Dietrich Wuttke**, Ilmenau University of Technology, Germany

Local Organizational Committee

Prof. **Sergey Subbotin**, National University "Zaporizhzhia Polytechnic", Ukraine
MSc **Serhii Leoschenko**, National University "Zaporizhzhia Polytechnic", Ukraine
Maxim Andreev, National University "Zaporizhzhia Polytechnic", Ukraine

CEUR-WS.org/Vol-2608/organization.pdf

CEUR Workshop Proceedings
http://ceur-ws.org
ISSN 1613-0073

Vol-2608
urn:nbn:de:0074-2608-1

Copyright © 2020 for the individual papers by the papers' authors. Copyright © 2020 for the volume as a collection by its editors. This volume and its papers are published under the Creative Commons License Attribution 4.0 International (CC BY 4.0).

CMIS-2020
Computer Modeling and Intelligent Systems

Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020)

Zaporizhzhia, Ukraine, April 27-May 1, 2020.

Edited by

Sergey Subbotin *

* National University "Zaporizhzhia Polytechnic", Faculty of Computer Science and Technologies, 69063, Zaporizhzhia, Ukraine

Table of Contents

Session 1: Information Technologies for Decision Making, Optimization and Control

Session 2: Information Systems and Data Processing

Session 3: Advanced Info-Communicational Technologies

Session 4: Information Technologies for Computer Modeling

Session 5: Computing in Applications

Session 6: Intelligent Information Technologies

2020-05-05: submitted by Sergey Subbotin metadata incl. bibliographic data published under Creative Commons CC0
2020-05-19: published on CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073) |valid HTML5|

Speech Vocal Clustering Methods

Eugene Fedorov[1[0000-0003-3841-7373]], Tetyana Utkina[1[0000-0002-6614-4133]], Kostiantyn Rudakov[1[0000-0003-0000-6077]], Andriy Lukashenko[2[0000-0002-6016-1899]], Ihor Zubko[1[0000-0002-3318-3347]], Maryna Chychuzhko[1[0000-0001-5329-7897]], Valentyna Lukashenko[1[0000-0002-6749-9040]]

[1] Cherkasy State Technological University, Cherkasy, Shevchenko blvd., 460, 18006, Ukraine
{t.utkina, ckc, k.rudakov, i.zubko, m.chychuzhko}@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 clustering efficiency of vocal speech sounds is considered. Centroid and medoid clustering methods that use normalized distances which increases the efficiency of clustering, are proposed. Characteristics and quality criteria based on them are suggested for these methods. These methods have been investigated on TIMIT database and are intended for intelligent biometric identification systems.

Keywords: vocal speech sound, centroid, medoid, clustering, inter-cluster distance, intra-cluster distance, compression ratio.

CEUR-WS.org/Vol-2608/paper79.pdf

1 Introduction

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

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

To increase the efficiency of analysis, storage and transmission of a speech signal, knowledge of the structure of vocal speech sound is required, for which the following clustering methods are used.

Traditional methods (without the use of artificial neural networks and metaheuristics) [3-6] search for a solution faster than metaheuristic methods, some methods do not require setting of the number of clusters, but perform only a directed search for a solution and some methods require setting of the number of clusters or additional parameters.

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Neural network methods [7-10] perform a search for a solution faster than metaheuristic methods, but perform only a directed search for a solution and require exact setting of the number of clusters.

Metaheuristic methods [11-14] perform a random search for a solution, but the search for a solution is longer than that of other methods, and require exact setting of the number of clusters.

Thus, only some traditional methods do not require an exact setting of the number of clusters, but instead require setting of additional parameters.

The aim of the work is to increase the clustering efficiency of vocal speech sounds by providing an adaptive number of clusters, a minimum number of set parameters and parallel information processing.

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

1. to develop a centroid clustering method;
2. to create medoid clustering methods;
3. to determine the characteristics and quality criterion of the clustering method;
4. to conduct a numerical study of the proposed clustering methods.

2 Problem statement

The problem of increasing the clustering efficiency of vocal speech sound comes down to the problem of finding such vector of parameters Θ^* that satisfies the criterion $F = SSWB + \frac{1}{C} \to \min\limits_{\Theta}$, where $SSWB$ is the ratio of the sums of average intra-cluster and inter-cluster distances, C is the compression ratio of speech sound.

3 Literature review

Traditional clustering methods include:

1. Partition-based (partitioning-based) or center-based methods

   In this case, a cluster is a set of objects, each of which is closer to the center of this cluster than to the center of any other cluster. The center of the cluster is usually a centroid (average value of the coordinates of all objects in the cluster) or a medoid (an object of the cluster, the average difference of which from other objects in the cluster is minimal). These methods consider intersecting areas with a high density of objects as different clusters, for example, K-means [15], PAM (k-medoids) [16], FCM [17], ISODATA [18] methods. The advantage of these methods consists in a quick search for a solution; some of them do not require setting of the number of clusters and have poor sensitivity to noise or random emissions. The disadvantage of these methods consists in the presence of only a directed search for a solution; a problem with clusters of different shapes, sizes and densities; some of them require setting of the number of clusters and are sensitive to noise or random emissions or require setting of additional parameters.

2. Model mixture or distribution-based or model-based methods

   In this case, a cluster is described by the probability density function. These methods should be used when clusters are of different sizes, and the set of objects of all clusters can be described by a mixture of distribution densities, for example, EM method [19]. The advantage of these methods consists in a quick search for a solution; determination of the shape and size of each cluster by the corresponding distribution density. The disadvantage of these methods consists in the presence of only a directed search for a solution; requirement for precise setting of the number of clusters; sensitivity to noise or random emissions; the difficulty of choosing the appropriate distribution density.
3. Density-based methods

   In this case, a cluster is an area with a high density of objects, which is separated by areas with a low density of objects from areas with a high density of objects. These methods consider intersecting areas with a high density of objects as a single cluster. These methods should be used when noise and random emissions are present, clusters have different shapes and sizes, and the number of clusters is unknown. Methods that define clusters (for example, DBSCAN method [20]) and those that provide visualization of clusters (for example, OPTICS method [21]) are distinguished. The advantage of these methods consists in the absence of a requirement to set the number of clusters; the presence of various shapes and sizes of a cluster; poor sensitivity to noise or random emissions. The disadvantage of these methods consists in the presence of only a directed search for a solution; a problem with clusters of heterogeneous density; slow search for a solution.
4. Hierarchal methods

   These methods provide visualization of a cluster tree called a dendrogram. On a dendrogram, pairs of clusters that are joined (in the case of agglomerative methods) or obtained as a result of separation (in the case of divisive methods) are connected by a U-shaped arc, the height of which corresponds to the distance between clusters. By the method of constructing a dendrogram, these methods are divided into agglomerative or bottom up ones – each object is considered as a singleton cluster, after which a stepwise combination of pairs of the closest clusters is performed (for example, centroid communication, Ward, single communication, full communication, group average methods [22, 23]) and divisive or top down ones – all objects are considered as one cluster, and at each step one of the constructed clusters is divided into a couple of clusters (for example, DIANA, DISMEA methods [22, 23]). The advantage of these methods consists in the absence of a requirement to set the number of clusters; visibility; some of them don't strongly tend to clusters of a certain shape and size and have poor sensitivity to noise or random emissions. The disadvantage of these methods consists in the presence of only a directed search for a solution; slow search for a solution; some of them strongly tend to clusters of a certain shape and size, are sensitive to noise or random emissions, tend to discard clusters of high power.

Usually, the methods listed above either require setting of the number of clusters or have a slow search for a solution or require setting of additional parameters, which leads to a decrease in the clustering efficiency.

Therefore, the urgent task is to increase the clustering efficiency of vocal speech sounds by providing an adaptive number of clusters, a minimum number of set parameters and parallel information processing.

4 Method clustering of vocal speech sound

4.1 Centroid clustering of vocal speech sound based on minimum distance method

Centroid clustering of vocal speech sound based on the author's minimum distance method includes the following steps:

1. Set a lot of samples of vocal speech sound $S = \{s_i(n)\}$, $i \in \overline{1,I}, n \in \overline{1,N}$, which are in a single amplitude-time window, where I is the number of samples, *N* is the length of the sample. Set the number of quantization levels of a speech signal L (for an 8-bit sound sample $L = 256$). Set normalized threshold ε, $0 < \varepsilon < 1$. Set the number of clusters $K = 0$.
2. Calculate normalized squared distance between each pair of sound samples

$$D_{ij} = \frac{\left\| s_i - s_j \right\|^2}{NL^2}, \ i \in \overline{1,I}, \ j \in \overline{1,I}.$$

3. Calculate the distance between each sound sample and a lot of sound samples

$$d_i = \sum_{j=1}^{I} D_{ij}, \ i \in \overline{1,I}.$$

4. Determine the number of sound sample with a minimum distance

$$i^* = \arg\min_i d_i, \ i \in \overline{1,I}.$$

5. Set a sound sample with a minimum distance as the new cluster center, i.e. $m_{K+1} = s_{i^*}$, set the number of sound samples in the new cluster per unit, i.e. $a_{K+1} = 1$, increase the number of clusters, i.e. $K = K + 1$.
6. Set the sound sample number $i = 1$.
7. If $i^* = i$, then go to step 13.
8. Calculate normalized squared distance between the i^{th} sound sample and cluster centers

$$D_k = \frac{\left\| s_i - m_k \right\|^2}{NL^2}, \ k \in \overline{1,K}.$$

9. Calculate the smallest normalized squared distance between the i^{th} sound sample and cluster centers

$$d^* = \min_k D_k \,,\ k \in \overline{1,K} \,.$$

10. Determine the number of the cluster with a minimum distance

$$k^* = \arg\min_k D_k \,,\ k \in \overline{1,K} \,.$$

11. If $d^* \le \varepsilon$, then calculate the new center of the $k^{*\,\text{th}}$ cluster, i.e. $m_{k^*} = \dfrac{a_{k^*} m_{k^*} + s_i}{a_{k^*} + 1}$, increase the number of sound samples in the $k^{*\,\text{th}}$ cluster, i.e. $a_{k^*} = a_{k^*} + 1$.
12. If $d^* > \varepsilon$, then set the i^{th} sample as the new cluster center, i.e. $m_{K+1} = s_i$, set the number of sound samples in the new cluster per unit $a_{K+1} = 1$, increase the number of clusters, i.e. $K = K + 1$.
13. If $i < I$, then go to the new sample, i.e. $i = i + 1$, go to step 7.

The method results in an adaptive set of centroids.

4.2 Medoid clustering of vocal speech sound based on minimax distance method

Medoid clustering of vocal speech sound based on minimax distance method, unlike the traditional version, preliminarily determines the number of a sound sample with a minimum distance for a non-random choice of the first cluster center and includes the following steps:

1. Set a lot of samples of vocal speech sound $S = \{s_i(n)\}$, $i \in \overline{1,I}, n \in \overline{1,N}$, which are in a single amplitude-time window, where I is the number of samples, N is the length of the sample. Set the number of clusters $c = 0$.
2. Calculate the squared distance between each pair of sound samples

$$D_{ij} = \left\| s_i - s_j \right\|^2 ,\ i \in \overline{1,I} \,,\ j \in \overline{1,I} \,.$$

3. Calculate the distance between each sound sample and a lot of sound samples

$$d_i = \sum_{j=1}^{I} D_{ij} \,,\ i \in \overline{1,I} \,.$$

4. Determine the number of sound sample with a minimum distance

$$i^* = \arg\min_i d_i \,,\ i \in \overline{1,I} \,.$$

5. Set a sound sample with a minimum distance as the new cluster center, i.e. $\phi_{c+1}(n) = s_{i^*}(n)$, $n \in \overline{1,N}$, increase the number of clusters, i.e. $c = c+1$.
6. Calculate the squared distance between each sound sample and each cluster center

$$D_{ik} = \|s_i - \phi_k\|^2, \ i \in \overline{1,I}, \ k \in \overline{1,c}.$$

7. Calculate a minimax squared distance between sound samples and cluster centers

$$d^* = \max_i \min_k D_{ik}, \ i \in \overline{1,I}, \ k \in \overline{1,c}.$$

8. Determine the number of sound sample with a minimax squared distance

$$i^* = \arg\max_i \min_k D_{ik}, \ i \in \overline{1,I}, \ k \in \overline{1,c}.$$

9. If $c = 1$, then set a sound sample with a minimax squared distance as the new cluster center, i.e. $\phi_{c+1}(n) = s_{i^*}(n)$, $n \in \overline{1,N}$, increase the number of clusters, i.e. $c = c+1$, go to step 6.
10. Calculate the average squared distance between cluster centers

$$\varepsilon = \frac{2}{c^2 - c} \sum_{i=1}^{c} \sum_{j=i}^{c} \|\phi_i - \phi_j\|^2.$$

11. Verification of the termination condition.

If $d^* > \frac{\varepsilon}{2}$, then set a sound sample with a minimax squared distance as the new cluster center, i.e. $\phi_{c+1}(n) = s_{i^*}(n)$, $n \in \overline{1,N}$, increase the number of clusters, i.e. $c = c+1$, go to step 6.

12. Calculate the distance between each sound sample and each cluster center

$$D_{ik} = \|s_i - \phi_k\|, \ i \in \overline{1,I}, \ k \in \overline{1,c}.$$

13. Determine for each sound sample the cluster center closest to it

$$u_i = \arg\min_k D_{ik}, \ i \in \overline{1,I}, \ k \in \overline{1,c}.$$

The method results in an adaptive set of medoids.

4.3 Medoid clustering of vocal speech sound based on subtractive clustering method

Medoid clustering of vocal speech sound based on subtractive clustering method, unlike the traditional version, uses the normalization of squared distances for calculating the potentials of sound samples to specify normalized standard deviations and includes the following steps:

1. Set a lot of samples of vocal speech sound $S = \{s_i(n)\}$, $i \in \overline{1,I}, n \in \overline{1,N}$, which are in a single amplitude-time window, where I is the number of samples, N is the length of the sample. Set the number of quantization levels of a speech signal L (for an 8-bit sound sample $L = 256$). Set the threshold to stop the method δ, $0 < \delta < 1$. Set standard deviations σ_a and σ_b, $\sigma_b > \sigma_a$, $0 < \sigma_a < 1$, $0 < \sigma_b < 1$. Set the number of clusters $c = 0$.
2. Calculate the potential of each sound sample

$$P(i) = \sum_{j=1}^{I} e^{-\frac{4\|s_i - s_j\|^2}{\sigma_a^2 N L^2}}, \; i \in \overline{1,I}.$$

3. Determine the number of sound sample with a maximum potential

$$i^* = \arg\max_i P(i), \; i \in \overline{1,I}.$$

4. Set the potential $P(i^*)$ as the new cluster potential, i.e. $\overline{P}(c+1) = P(i^*)$.
5. Calculate the new potential of each sound sample

$$P(i) = P(i) - \overline{P}(c+1) e^{-\frac{4\|s_i - \phi_{c+1}\|^2}{\sigma_b^2 N L^2}}, \; i \in \overline{1,I}.$$

6. Verification of the termination condition.

   If $\dfrac{\overline{P}(c+1)}{\overline{P}(1)} \geq \delta$, then set a sound sample with the highest potential as the new clus-

ter center, i.e. $\phi_{c+1}(n) = s_{i^*}(n)$, $n \in \overline{1,N}$, increase the number of clusters, i.e. $c = c+1$, go to step 3.

7. Calculate the distance between each sound sample and each cluster center

$$D_{ik} = \|s_i - \phi_k\|, \; i \in \overline{1,I}, \; k \in \overline{1,c}.$$

8. Determine for each normalized sound sample the cluster center closest to it

$$u_i = \arg\min_k D_{ik}, \; i \in \overline{1,I}, \; k \in \overline{1,c}.$$

The method results in an adaptive set of medoids.

4.4 Medoid clustering of vocal speech sound based on minimum average distance method

Medoid clustering of vocal speech sound based on the author's method of minimum average distance includes the following steps:

1. Set a lot of samples of vocal speech sound $S = \{s_i(n)\}$, $i \in \overline{1, I}, n \in \overline{1, N}$, which are in a single amplitude-time window, where I is the number of samples, N is the length of the sample. Set the number of quantization levels of a speech signal L (for an 8-bit sound sample $L = 256$). Set the radius of the neighborhood of sound samples ε, $0 < \varepsilon < 1$. Set the number of clusters $c = 0$. Set a lot of samples of speech sound that haven't fallen into existing clusters, $\tilde{S} = \{s_i(n)\}$.
2. Calculate normalized squared distance between each pair of sound samples
$$D_{ij} = \frac{\left\| s_i - s_j \right\|^2}{NL^2}, \ i \in \overline{1, I}, \ j \in \overline{1, I}.$$
3. Calculate the distance between each sound sample and a lot of sound samples
$$d_i = \sum_{j=1}^{I} D_{ij}, \ i \in \overline{1, I}.$$
4. Determine the neighborhood of each sound sample
$$U_{i,\varepsilon} = \{ j \mid D_{ij} < \varepsilon, j \in \overline{1, I} \}, \ i \in \overline{1, I}.$$
5. Determine the neighborhood of the sound sample with a minimum average distance
$$i^* = \arg\min_{i \in \tilde{S}} \frac{d_i}{|\tilde{S}|},$$
$$U^* = U_{i^*,\varepsilon}.$$
6. Calculate the new distance between each sound sample and a lot of sound samples that haven't fallen into existing clusters
$$d_i = \begin{cases} 0, & i \in U^* \\ d_i - \sum_k D_{ik}, & i \notin U^* \end{cases}, \ k \in U_{i,\varepsilon} \bigcap U^*, \ i \in \tilde{S}.$$
7. Determine the new neighborhood for each sound sample that hasn't fallen into existing clusters.

$$U_{i,\varepsilon} = \begin{cases} \varnothing, & i \in U^* \\ U_{i,\varepsilon} \setminus U^*, & i \notin U^* \end{cases}, \ i \in \tilde{S} \ .$$

8. Set a sound sample with a minimum average distance as the new cluster center, i.e. $\phi_{c+1}(n) = s_{i^*}(n)$, $n \in \overline{1,N}$, increase the number of clusters, i.e. $c = c+1$.
9. Verification of the termination condition.

   If $\tilde{S} \setminus U^* \neq \varnothing$, then $\tilde{S} = \tilde{S} \setminus U^*$, go to step 5.
10. Calculate the distance between each sound sample and each cluster center (codebook vector)

$$D_{ik} = \| s_i - \phi_k \| , \ i \in \overline{1,I} , \ k \in \overline{1,c} \, .$$

11. Determine for each sound sample the cluster center closest to it

$$u_i = \arg\min_k D_{ik} , \ i \in \overline{1,I} , \ k \in \overline{1,c} \, .$$

The method results in an adaptive set of medoids.

5 Determination of characteristics and quality criterion for clustering methods of vocal speech sound

To evaluate clustering methods, the following characteristics are used in the work:

1. The sum of average intra-cluster distances:

– in the case of centroid clustering methods

$$SSW = \sum_{k=1}^{K} \left(\frac{\sum_{i=1}^{I} \chi_{A_k}(s_i) \| s_i - m_k \|^2}{\sum_{i=1}^{I} \chi_{A_k}(s_i)} \right),$$

– in the case of medoid clustering methods

$$SSW = \sum_{k=1}^{c} \left(\frac{\sum_{i=1}^{I} \chi_{A_k}(s_i) \| s_i - \phi_k \|^2}{\sum_{i=1}^{I} \chi_{A_k}(s_i)} \right),$$

$$\chi_{A_k}(s_i) = \begin{cases} 1, & s_i \in A_k \\ 0, & s_i \notin A_k \end{cases},$$

where A_k is a cluster.

2. The sum of inter-cluster distances:

– in the case of centroid clustering methods

$$SSB = \sum_{k=1}^{K} \|m_k - m\|^2 ,$$

$$m = \frac{\sum_{i=1}^{I} s_i}{I}$$

– in the case of medoid clustering methods

$$SSB = \sum_{k=1}^{c} \|\phi_k - \phi_1\|^2 .$$

3. The ratio of the sums of average intra-cluster and inter-cluster distances

$$SSWB = \frac{SSW}{SSB} .$$

In addition, to assess the quality of clustering methods, the following characteristics – compression ratio for speech sound – is proposed in the work:

$$C = \frac{I}{c} ,$$

where I is the number of samples of vocal speech sound, c – the number of clusters.

The following criterion for the quality of clustering, which means choosing such a value Θ that minimizes the sum of the inverse of compression ratio and the ratio of the sums of average intra-cluster and inter-cluster distances, is formulated in the paper:

$$F = SSWB + \frac{1}{C} \to \min_{\Theta} . \tag{1}$$

For centroid clustering of vocal speech sound based on minimum distance method $\Theta = \varepsilon$. For medoid clustering of vocal speech sound based on subtractive clustering method $\Theta = (\sigma_a, \sigma_b, \delta)$. For medoid clustering of vocal speech sound based on minimum average distance method $\Theta = \varepsilon$.

6 Experiments and results

Numerical experiments were carried out using notebook Intel Core i5 8th Gen, MATLAB package, CUDA technology of parallel information processing on the Ge-

Force 920M graphics card with the number of threads in the block N^s = 1024. In this case, the most time-consuming (computational complexity $O(I^2)$, where I is the number of samples of vocal speech sound) step 2 of all four proposed clustering methods was parallelized. This made it possible to speed up the search for a solution.

For speech signals containing vocal sounds, the sampling frequency f_d = 8 kHz and the number of quantization levels L = 256 were set. The length of the sample of vocal speech sound $N = 256$.

For modified method of subtractive clustering, the following fixed parameter values δ =0.01, $\sigma_b = 1.25\sigma_a$ were set.

The results of a numerical study of the proposed clustering methods for vocal sounds of people from TIMIT database are presented in Table 1.

The result presented in Table 1 shows that the author's method of minimum average distance provides the smallest F value, calculated according to (1). The compression ratio is approximately 7.5, i.e. the number of stored samples is reduced by about 7.5 times.

Based on the experiments, the following conclusions can be drawn.

The author's method of minimum distance should be used only when centroids are required, and not medoids, since it has the largest F and the lowest compression ratio.

Table 1. Results of a numerical study of vector quantization methods

Characteristics of clustering methods	Author's method of minimum distance	Modified minimax distance method	Modified subtractive clustering method	Author's method of minimum average distance
Compression ratio	4.5429	10.7500	7.5834	7.5834
F	0.3157	0.2712	0.2420	0.2400
Number of parameters	1	0	3	1

The modified method of minimax distance performs a coarser adjustment of the number of clusters than the modified method of subtractive clustering and the author's method of minimum average distance, because it does not use parameters. On the other hand, an operator does not need to set any parameters, the values of which are established empirically.

The modified method of subtractive clustering performs finer adjustment of the number of clusters than the modified method of minimax distance, because it uses parameters. It requires more complex setting than the author's method of minimum average distance, because it uses three parameters, the values of which are established empirically.

The author's method of minimum average distance performs finer adjustment of the number of clusters than the modified method of minimax distance, because it uses parameters. It requires simpler adjustment than the modified method of subtractive clustering, because it uses only one parameter, the value of which is established empirically.

7 Conclusions

The article considers the problem of increasing the clustering efficiency of vocal speech sounds. The following clustering methods are proposed – the author's method of minimum distance, the author's method of minimum average distance, the modified method of minimax distance (unlike the traditional version, it preliminarily determines the number of a sound sample with a minimum distance for non-random choice of the first cluster center), the modified method of subtractive clustering (unlike the traditional version, it uses squared distances normalization to set normal standard deviations in calculating the potentials of sound samples). Characteristics and quality criteria based on them are proposed for these methods. The proposed methods allow to increase the clustering efficiency of vocal speech sounds by providing an adaptive number of clusters, a minimum number of specified parameters and parallel information processing. The proposed methods are intended for software implementation on GPU using CUDA technology, which speeds up the process of finding a solution. Software that implements the proposed methods has been developed and researched on TIMIT database. The conducted experiments have confirmed the operability of the developed software and allow to recommend it for use in practice when solving problems of biometric identification of a person. 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. Procedia Engineering. 38, 3122–3126 (2012).
2. Weychan, R., Marciniak, T., Dabrowski, A.: Implementation aspects of speaker recognition using Python language and Raspberry Pi platform. 2015 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA). (2015). doi: 10.1109/SPA.2015.7365153
3. Brusco, M.J., Shireman, E., Steinley, D.: A comparison of latent class, K-means, and K-median methods for clustering dichotomous data. Psychological Methods. 22, 563–580 (2017). doi: 10.1037/met0000095
4. Lenarczyk, P., Piotrowski, Z.: Speaker recognition system based on GMM multivariate probability distributions built-in a digital watermarking token. Przeglad Elektrotechniczny. 89(2), 59–63 (2013).
5. Ram, A., Jalal, S., Jalal, A.S., Kumar, M.: A Density Based Algorithm for Discovering Density Varied Clusters in Large Spatial Databases. International Journal of Computer Applications. 3, 1–4 (2010). doi 10.5120/739-1038
6. Krishnamurthy, A., Balakrishnan, S., Xu, M., Singh, A.: Efficient Active Algorithms for Hierarchical Clustering. CoRR. abs/1206.4672, (2012).
7. Haykin, S.S.: Neural networks and learning machines. Pearson, Delhi (2016).

8. Du, K.-L., Swamy, M.N.S.: Neural Networks and Statistical Learning. (2014). doi: 10.1007/978-1-4471-5571-3
9. Sivanandam, S.N., Sumathi, S., Deepa, S.N.: Introduction to neural networks using MATLAB 6.0. Tata McGraw-Hill Education, New Delhi (2010).
10. Larin, V.J., Fedorov, E.E.: Combination of PNN network and DTW method for identification of reserved words, used in aviation during radio negotiation. Radioelectronics and Communications Systems. 57, 362–368 (2014). doi: 10.3103/s0735272714080044
11. Das, S., Abraham, A., Konar, A.: Metaheuristic clustering. Springer Berlin, Berlin (2009).
12. Saemi, B., Hosseinabadi, A.A.R., Kardgar, M., Balas, V.E., Ebadi, H.: Nature Inspired Partitioning Clustering Algorithms: A Review and Analysis. Soft Computing Applications Advances in Intelligent Systems and Computing. 96–116 (2017). doi: 10.1007/978-3-319-62524-9_9
13. Brownlee, J.: Clever algorithms: nature-inspired programming recipes. Creative Commons, Melbourne (2011).
14. Subbotin, S., Oliinyk, A., Levashenko, V., Zaitseva, E. Diagnostic rule mining based on artificial immune system for a case of uneven distribution of classes in sample. Communications. 3. 3-11. (2016).
15. Steinley, D., Brusco, M.J.: Initializing K-means Batch Clustering: A Critical Evaluation of Several Techniques. Journal of Classification. 24, 99–121 (2007). doi: 10.1007/s00357-007-0003-0
16. Kaufman, L., Rousseeuw, P.J.: Finding groups in data: an introduction to cluster analysis. Wiley-Interscience, Hoboken, NJ (2005).
17. Wu, K.-L., Yu, J., Yang, M.-S.: A novel fuzzy clustering algorithm based on a fuzzy scatter matrix with optimality tests. Pattern Recognition Letters. 26, 639–652 (2005). doi: 10.1016/j.patrec.2004.09.016
18. Memarsadeghi, N., Mount, D.M., Netanyahu, N.S., Moigne, J.L.: A Fast Implementation Of The Isodata Clustering Algorithm. International Journal of Computational Geometry & Applications. 17, 71–103 (2007). doi: 10.1142/S0218195907002252.
19. Hastie, T., Friedman, J., Tisbshirani, R.: The Elements of statistical learning: data mining, inference, and prediction. Springer, New York (2017).
20. Borah, B., Bhattacharyya, D.: DDSC : A Density Differentiated Spatial Clustering Technique. Journal of Computers. 3, (2008). doi: 10.4304/jcp.3.2.72-79
21. Liu, P., Zhou, D., Wu, N.: VDBSCAN: Varied Density Based Spatial Clustering of Applications with Noise. 2007 International Conference on Service Systems and Service Management. (2007). doi: 10.1109/ICSSSM.2007.4280175
22. Aggarwal, C.C., Reddy, C.K.: Data Clustering: Algorithms and Applications. Chapman and Hall/CRC, Boca Raton, FL (2018).
23. Gan, G., Ma, C., Wu, J.: Data clustering: theory, algorithms, and applications. Siam, Philadelphia (2007).