FACTA UNIVERSITATIS
Series: **Automatic Control and Robotics** Vol. 20, Nº 2, 2021, pp. 99 - 110
https://doi.org/10.22190/FUACR210321008J

Regular Paper

TWO-DIMENSIONAL GMM-BASED CLUSTERING IN THE PRESENCE OF QUANTIZATION NOISE

UDC ((621.391:519.226)+004.8)

Aleksandra Jovanović, Zoran Perić

University of Niš, Faculty of Electronic Engineering,
Department of Telecommunications, Republic of Serbia

Abstract. *In this paper, unlike to the commonly considered clustering, wherein data attributes are accurately presented, it is researched how successful clustering can be performed when data attributes are represented with smaller accuracy, i.e. by using the small number of bits. In particular, the effect of data attributes quantization on the two-dimensional two-component Gaussian mixture model (GMM)-based clustering by using expectation–maximization (EM) algorithm is analyzed. An independent quantization of data attributes by using uniform quantizers with the support limits adjusted to the minimal and maximal attribute values is assumed. The analysis makes it possible to determine the number of bits for data presentation that provides the accurate clustering. These findings can be useful in clustering wherein before being grouped the data have to be represented with a finite small number of bits due to their transmission through the bandwidth-limited channel.*

Key words: *Unsupervised learning, clustering, Gaussian mixture model, expectation-maximization algorithm, quantization noise*

1. INTRODUCTION

The probabilistic theory represents a powerful tool for understanding phenomena and description of problems in statistical analysis, signal processing, detection, clustering, classification and in many areas of artificial intelligence. A very important probabilistic model widely used in data mining, pattern recognition, machine learning and statistical analysis is a Gaussian mixture model (GMM) that models the presence of normally distributed populations/clusters within larger population/cluster by defining marginal distribution as a weighted sum of Gaussian densities [1]-[3]. Thus, the marginal GMM

Received March 21, 2021 / Accepted July 05, 2021
Corresponding author: Aleksandra Jovanović
University of Niš, Faculty of Electronic Engineering, Department of Telecommunications, Aleksandra Medvedeva 14, 18000 Niš, Republic of Serbia
E-mail: aleksandra.jovanovic@elfak.ni.ac.rs

distribution is characterized with following parameters: the mixture weights (prior probabilities of clusters), the means and variances/covariances of clusters. Then, the GMM-based clustering is unsupervised learning that finds the unknown parameters of marginal GMM distribution and responsibilities for each data. One elegant and powerful method for GMM-based clustering is the expectation–maximization (EM) algorithm that from unlabeled data iteratively estimates the model parameters starting from some initial values. Each of iterations consists of an expectation (E) step, which determines responsibilities for each data, i.e assigns data to clusters with some probabilities given the current estimated GMM parameters, and a maximization (M) step, which re-estimates the parameters by maximizing likelihood function, under the assumption that clustering in the E step is correct. The iterative process continues through the E and M steps successively until convergence [1]-[7]. To summarize, the EM algorithm actually groups data with respect to their similarities whereby neither the knowledge about the statistical characteristics of data set nor the appropriate number of clusters are known. With the EM algorithm, the system is learning from unlabeled data. Thereby, data are assigned to clusters with some probabilities called responsibilities. Because of that, the EM algorithm represents an unsupervised learning technique that enables soft clustering.

In this paper we extend analysis from [8], where the effect of uniform data quantization on one-dimensional two-component GMM-based clustering by means of the EM algorithm was considered. Actually, in this paper we focus on the EM algorithm for clustering data modeled with two-dimensional two-component GMM in the presence of quantization noise. This analysis and analysis from [8] are significant for clustering problems wherein before being grouped by using the EM algorithm the data attributes are quantized by using a small number of bits due to their transmission through the bandwidth-limited channel. The importance of study on quantization noise influence on clustering and classification was also recognized in [9] and [10].

This paper is organized as follows. In Section 2, the quantizers for data attributes are designed and the GMM-based clustering by using the EM algorithm is described. In Section 3, the results of MATLAB simulation of data clustering in the presence of quantization noise are presented and discussed. The conclusions are given in Section 4.

2. EM ALGORITHM IN THE PRESENCE OF QUANTIZATION NOISE

To model complex densities, Gaussian mixture distribution is used defined as linear combination of K Gaussian probability density functions (pdfs) [1], [4], [6]:

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k N(\mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\ , \tag{1}$$

where $N(\mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the d-dimensional Gaussian pdf

$$N(\mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{|\boldsymbol{\Sigma}_k|}} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k) \right\}, \quad k = 1, \ldots, K\ , \tag{2}$$

so that $\mathbf{x}$ is the vector consisting of d variables, $\boldsymbol{\mu}_k$ is a d-dimensional vector denoting the mean values of corresponding variables, while $d \times d$ covariance matrix $\boldsymbol{\Sigma}_k$ -represents a measure of how changes in one variable are related with changes in another. Parameter π_k represents a mixture weight, so it holds:

$$\sum_{k=1}^{K} \pi_k = 1 . \tag{3}$$

Equation (1) can also be viewed as a marginal distribution $p(\mathbf{x})$ obtained by summing the joint distribution of observation **x** and the discrete latent variable **z,** $P(\mathbf{z})p(\mathbf{x}|\mathbf{z})$, over all possible states of **z** [1]**. Z** is a K-dimensional binary random variable $\mathbf{z}= [z_1, \ldots, z_K]$ having 1-of-K representation in which a particular element z_k is equal to 1 with probability $P(z_k = 1) = \pi_k$ and all other elements are equal to 0. This discrete latent variable can be interpreted as defining assignment of data point to specific component of the mixture. Therefore, the latent variable **z** has K states, $\mathbf{z}_k$, $k = 1,2,\ldots,K$, whereby $z_k = 1$ defines state $\mathbf{z}_k$ with marginal distribution [1]:

$$P(\mathbf{z}_k) = \prod_{l=1}^{K} \pi_l^{z_l} = P(z_k = 1) = \pi_k , \quad k = 1,\ldots, K , \tag{4}$$

while the conditional pdf of **x** given a particular value of **z** is the Gaussian pdf [1]

$$p(\mathbf{x}|\mathbf{z}_k) = \prod_{l=1}^{K} N(\mathbf{x}; \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)^{z_l} = N(\mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) , \tag{5}$$

so that it holds [1]

$$p(\mathbf{x}) = \sum_{k=1}^{K} P(\mathbf{z}_k) p(\mathbf{x}|\mathbf{z}_k) = \sum_{k=1}^{K} \pi_k N(\mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) . \tag{6}$$

The equality (6) points out that complex marginal distribution over the observed variable **x**, $p(\mathbf{x})$ can be expressed in terms of more tractable joint distribution over observed and latent variables.

One quantity of the Gaussian mixture distribution important for GMM-based clustering is the conditional probability of **z** given **x,** $P(\mathbf{z}_k|\mathbf{x}) = P(z_k = 1|\mathbf{x})$, $k = 1,..,K$, here denoted with $r_k(\mathbf{x})$. This conditional probability can be determined by using the Bayes' theorem [1]

$$\begin{aligned} r_k(\mathbf{x}) = P(\mathbf{z}_k|\mathbf{x}) &= \frac{P(\mathbf{z}_k) p(\mathbf{x}|\mathbf{z}_k)}{p(\mathbf{x})} \\ &= \frac{\pi_k N(\mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{l=1}^{K} \pi_l N(\mathbf{x}; \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)} , \quad k = 1,\ldots, K \end{aligned} . \tag{7}$$

Since the discrete latent variable defines assignment of data point to a certain component of the mixture, one can conclude that the mixture weight π_k can be viewed as the prior probability of $\mathbf{z}_k$, while $r_k(\mathbf{x})$ is the corresponding posterior probability after receiving **x**. This posterior probability is commonly referred as the responsibility, i.e. the probability that data **x** belongs to the kth cluster. By determining these responsibilities for each data, the clustering is done.

The clustering of data modeled with GMM is based on determining GMM parameters $\{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}_{k=1}^{K}$ and responsibilities $r_k(\mathbf{x}_n)$, $n = 1,2,\ldots,N$, $k = 1,2,\ldots, K$, where N is the size of data set subjected to the clustering. In this paper we assume that before being grouped the data are transmitted through the bandwidth-limited channel, because of which the data attributes are represented with smaller accuracy, i.e. by using the finite small number of bits. More precisely, we assume that the data attributes are quantized with the simplest quantization technique by utilizing uniform quantizers characterized with the cells of the same length and with the output levels in the midpoints of the cells, except for possibly the first and the last cells [11]-[16]. We opt for this kind of quantization because the GMM parameters are not known and the optimization of quantization cannot be performed as it was possible in [17]. Besides, since the correlation between data attributes is also not known, we decide to process independently attributes of two-dimensional data, whereby quantizers can be somewhat adjusted by setting different support limits and the number of levels. The support limit is a very important parameter in uniform quantization where amplitude being quantized is unbounded [12]-[22], as it is the case with the GMM distribution. Namely, with support limit decreasing the quantization error and thus the quantization noise for amplitudes within support limits decrease. But, on the other hand, the quantization error for amplitudes outside the support limits increases which can cause the significant increase of overall quantization noise. Therefore, it is preferable to optimize the support limit subject to average quantization error [12]-[22]. In our case it is not possible since the expression for mean square error is the function of the pdf having parameters which are unknown and which actually should be determined by the EM algorithm. From this reason we determine support limits by finding minimal and maximal values for each attribute of unlabeled data. This leads us to asymmetric uniform scalar quantizers $\mathrm{USQ}^{\mathrm{I}}$ and $\mathrm{USQ}^{\mathrm{II}}$ with supports $[x^{\mathrm{I}}_{\mathrm{min}}, x^{\mathrm{I}}_{\mathrm{max}}]$ and $[x^{\mathrm{II}}_{\mathrm{min}}, x^{\mathrm{II}}_{\mathrm{max}}]$, respectively. If we assume L^{I} and L^{II} for the number of levels of $\mathrm{USQ}^{\mathrm{I}}$ and $\mathrm{USQ}^{\mathrm{II}}$, respectively, then the decision and output levels for these quantizers are defined by equations:

$$\begin{aligned} x_i^{\mathrm{j}} &= x_{\mathrm{min}}^{\mathrm{j}} + (i-1)\frac{x_{\mathrm{max}}^{\mathrm{j}} - x_{\mathrm{min}}^{\mathrm{j}}}{L^j}, \qquad i = 1,2,\ldots,L^j+1, \quad j = \mathrm{I},\mathrm{II} \\ y_i^{\mathrm{j}} &= x_{\mathrm{min}}^{\mathrm{j}} + \left(i-\frac{1}{2}\right)\frac{x_{\mathrm{max}}^{\mathrm{j}} - x_{\mathrm{min}}^{\mathrm{j}}}{L^j}, \qquad i = 1,2,\ldots,L^j, \quad j = \mathrm{I},\mathrm{II} \end{aligned}, \tag{8}$$

where $\Delta^j = \dfrac{B^j}{L^j} = \dfrac{x_{\mathrm{max}}^{\mathrm{j}} - x_{\mathrm{min}}^{\mathrm{j}}}{L^j}$, $j = \mathrm{I},\mathrm{II}$ is the cell length or the uniform step size for USQ^j, $j = \mathrm{I}, \mathrm{II}$. Then, quantization of data attributes x_1 and x_2 should be performed in the following manner

$$\begin{aligned} \hat{x}_1 &= Q^{\mathrm{I}}(x_1) = x_{\mathrm{min}}^{\mathrm{I}} + \left(\left\lfloor \frac{x_1 - x_{\mathrm{min}}^{\mathrm{I}}}{\Delta^{\mathrm{I}}} \right\rfloor + \frac{1}{2}\right) \\ \hat{x}_2 &= Q^{\mathrm{II}}(x_2) = x_{\mathrm{min}}^{\mathrm{II}} + \left(\left\lfloor \frac{x_2 - x_{\mathrm{min}}^{\mathrm{II}}}{\Delta^{\mathrm{II}}} \right\rfloor + \frac{1}{2}\right) \end{aligned}, \tag{9}$$

where $\lfloor\cdot\rfloor$ denotes rounding to the nearest lower value. The number of bits used to transmit attributes is:

$$\log_2 L^{\mathrm{I}} + \log_2 L^{\mathrm{II}} = \log_2 (L^{\mathrm{I}} L^{\mathrm{II}}), \tag{10}$$

which results in quantizer rate, i.e the number of bits per dimension:

$$R = \frac{\log_2 (L^{\mathrm{I}} L^{\mathrm{II}})}{2} = \log_2 \sqrt{L^{\mathrm{I}} L^{\mathrm{II}}} \; . \tag{11}$$

As we have already stated, an elegant and powerful method for the GMM-based clustering is the EM algorithm. The EM algorithm formulation for the case wherein data attributes are firstly quantized is given below:

1) Initialize the means $\boldsymbol{\mu}_k$, covariances $\boldsymbol{\Sigma}_k$ and mixture weights π_k, $k = 1,\dots, K$.
2) E (expectation) step: evaluate the responsibilities for quantized data set $\hat{\mathbf{x}}_n = [\hat{x}_{n1} \quad \hat{x}_{n2}]$, $n = 1,2,\dots, N$ assuming the current parameters of GMM

$$r_k(\hat{\mathbf{x}}_n) = r_{nk} = \frac{\pi_k N(\hat{\mathbf{x}}_n; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{l=1}^{K} \pi_l N(\hat{\mathbf{x}}_n; \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)}, \quad n = 1,\dots N, k = 1,\dots, K \; . \tag{12}$$

3) M (maximization) step: re-estimate the GMM parameters using the current responsibilities

$$\begin{aligned} \mu_k &= \frac{\sum_{n=1}^{N} r_k(\hat{\mathbf{x}}_n)\hat{\mathbf{x}}_n}{\sum_{n=1}^{N} r_k(\hat{\mathbf{x}}_n)} \\ \Sigma_k &= \frac{\sum_{n=1}^{N} r_k(\hat{\mathbf{x}}_n)(\hat{\mathbf{x}}_n - \boldsymbol{\mu}_k)(\hat{\mathbf{x}}_n - \boldsymbol{\mu}_k)^T}{\sum_{n=1}^{N} r_k(\hat{\mathbf{x}}_n)}, \quad k = 1,\dots, K \; . \\ \pi_k &= \frac{\sum_{n=1}^{N} r_k(\hat{\mathbf{x}}_n)}{N} \end{aligned} \tag{13}$$

4) Check for convergence of the parameters. If the convergence criterion is not satisfied, return to step 2.

Finally, after the EM algorithm finishing, the cluster value for each point from quantized data set can be determined by selecting k that gives the maximal responsibility for this point:

$$c(\hat{\mathbf{x}}_n) = \arg\max_k r_k(\hat{\mathbf{x}}_n), \quad n = 1, 2, \cdots, N \; . \tag{14}$$

To estimate the influence of attribute quantization on the results of clustering, we also perform clustering of unquantized data

$$c(\mathbf{x}_n) = \arg\max_k r_k(\mathbf{x}_n), \quad n = 1, 2, \cdots, N \tag{15}$$

with the aim to compare the obtained results. As in [8], we define the similarity index as a ratio between the number of matched cluster values and the total number of data points

$$\gamma = \frac{\sum_{n=1}^{N}[1-c(\hat{\mathbf{x}}_n)\oplus c(\mathbf{x}_n)]}{N} = 1-\frac{\sum_{n=1}^{N}c(\hat{\mathbf{x}}_n)\oplus c(\mathbf{x}_n)}{N}, \qquad (16)$$

where $\oplus$ denotes modulo 2 summation. As γ has a value closer to 1, the similarity with the regular clustering is greater, that is quantization noise less degrades the GMM-based clustering by utilizing the EM algorithm.

3. Simulation Results

Within this research, by using MATLAB we generate a data set having two-dimensional Gaussian mixture distribution composed from two components ($d = 2$, $K = 2$) and consisting of 10^5 data points ($N = 10^5$). For a given R, we quantize the attributes of generated data points in accordance with eq. (9), and after that we cluster quantized data set by utilizing the EM algorithm. We present the obtained results for two different data sets. Actually, we generate data sets with parameters specified in Table 1. The main difference between these data sets is that the components within the first set are more distinctly separate than they are in the second set. The generated data sets we use as unlabeled and we further quantize and process them by using the EM algorithm. The obtained results we present in Figs. 1, 2 and 3 for the first set, and in Figs. 4, 5 and 6 for the second set. In all figures, we also present the pdf and clusters obtained by performing clustering of unquantized data set. The figures show that with the increase of R, i.e. with the increase of $L^{\mathrm{I}} = L^{\mathrm{II}} = L$, the results of the EM clustering are better matched with the ones of EM clustering in unquantized attribute case. The increase of matching with L is also verified with the similarity index increasing. It is expected since higher L causes smaller quantization error and provides more accurate presentation of data attributes enabling better estimation of the GMM parameters and more accurate clustering. However, we can notice that the degree of matching measured through the similarity index does not depend only on the quantizer rate R, but also on the distinction between clusters. Thus, in the first case, where the clusters are on larger distance, the rate of 2bits/dimension is satisfactory for successful data clustering since it is only important to determine the data category, not the adequate approximation for the values of the data attributes (see Fig. 1b and Fig. 2b). On the other hand, in the second case, where the clusters are on smaller distance, the rate of 2bits/dimension does not provide the successful clustering (compare Fig. 4b with Fig. 4a, and also Fig. 5b with Fig. 5a). In this case, the rate of 3 bits/dimension and higher enables the successful clustering. To avoid rate increasing when clusters overlap, we are going to decrease the support limit values in our future research. We expect that the narrowing of the support region will enable better quantization of data near the cluster boundary, and thus better differentiation and grouping of these data.

Table 1 Data set parameters.

	μ_1	Σ_1	π_1	μ_2	Σ_2	π_2
The first data set	[1 2]	$\begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}$	0.5	[-3 -5]	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	0.5
The second data set	[1 2]	$\begin{bmatrix} 3 & 0.2 \\ 0.2 & 2 \end{bmatrix}$	0.7	[-1 -2]	$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$	0.3

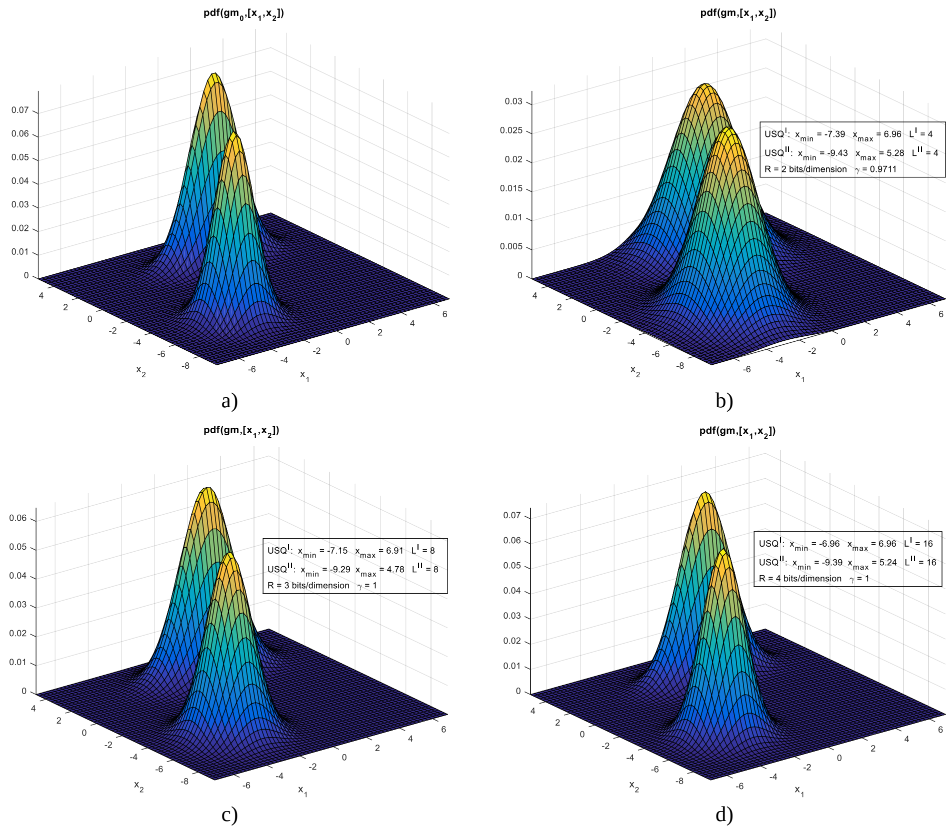

Fig. 1 Probability density function estimated for the first data set and for: a) unquantized attributes, b) $R = 2$ bits/dimension, c) $R = 3$ bits/dimension and d) $R = 4$ bits/dimension

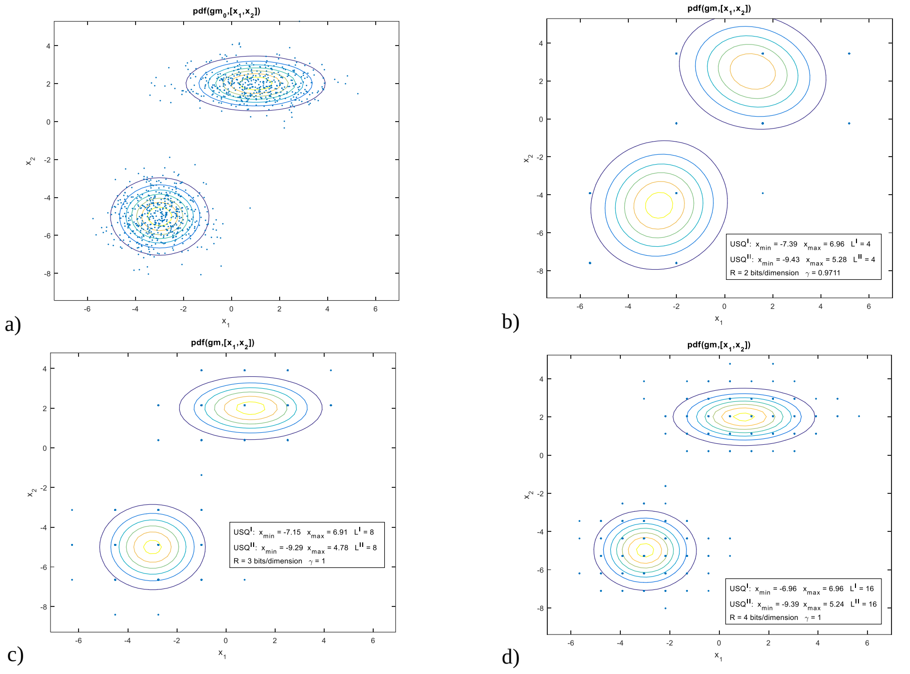

Fig. 2 The estimated probability density contours for the first data set and different R

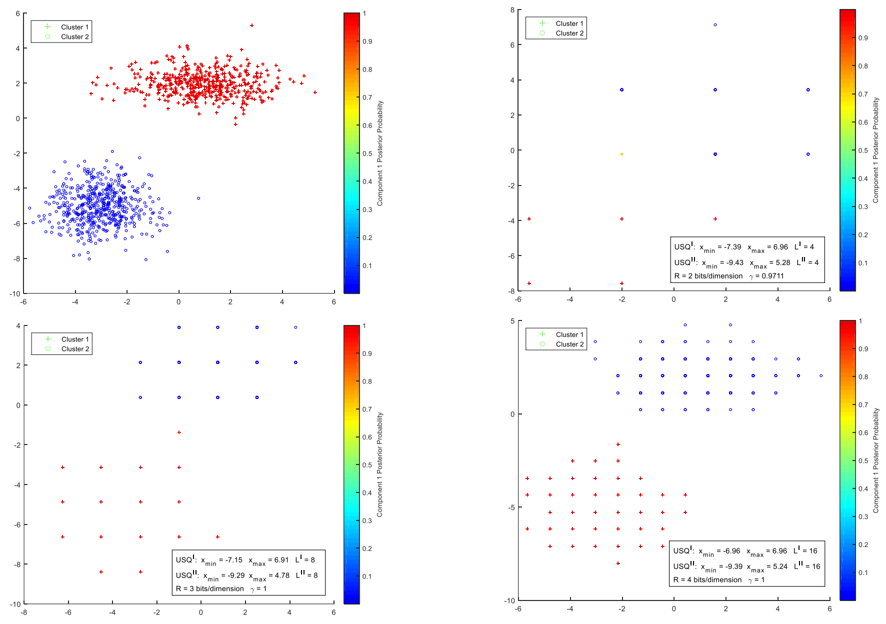

Fig. 3 Partition of the first data set into clusters for different R

In Figs. 3 and 6 the data points are assigned to one of two components of the estimated mixture distribution forming in such way two clusters. Data points are assigned to clusters based on the estimated posterior probabilities that a point is from a certain component. The highest posterior probability of data point determines the data cluster. In Figs. 3 and 6, for each data point a posteriori probability of data point belonging to cluster 1 is provided. Figures show that with the quantizer rate increasing the quantized data sets and cluster boundaries approach to original data sets and their cluster boundaries. In Fig. 3 there is no overlap between clusters, while in Fig. 6 the clusters overlap, whereby their peaks in Fig. 5c and Fig. 5d are distinct opposite to situation in Fig. 5b. This points out that the second data set can reasonably be divided into two clusters for rates 3 and 4 bits/dimension, while it is not possible for rate 2 bits/dimension.

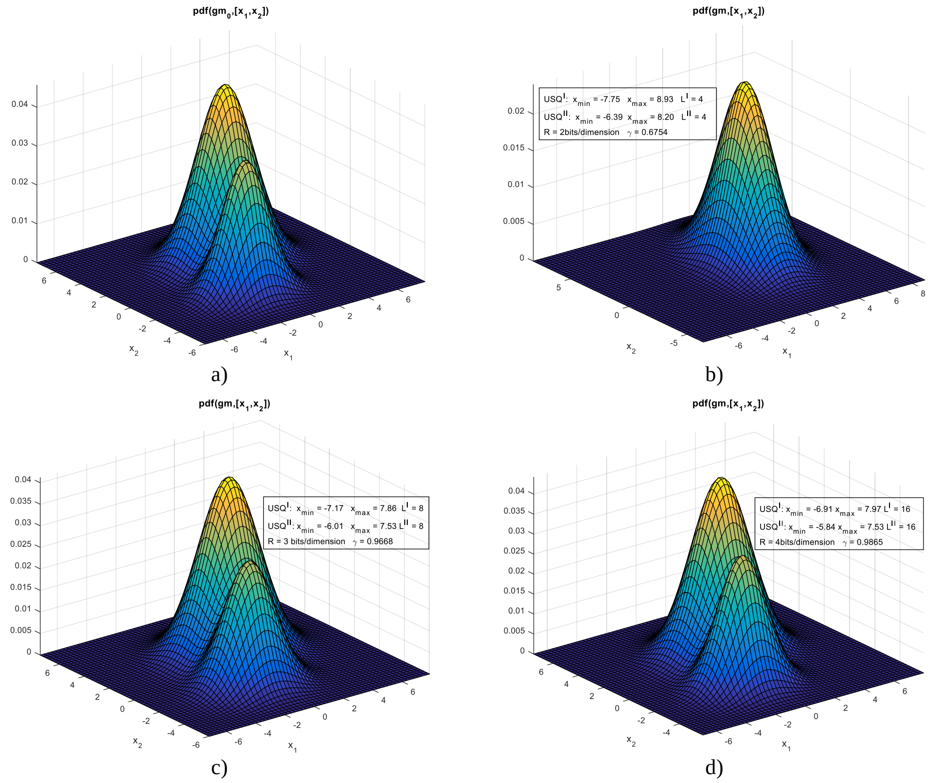

Fig. 4 Probability density functions estimated for the second data set: a) unquantized attributes, b) R = 2 bits/dimension, c) R = 3 bits/dimension and d) R = 4 bits/dimension

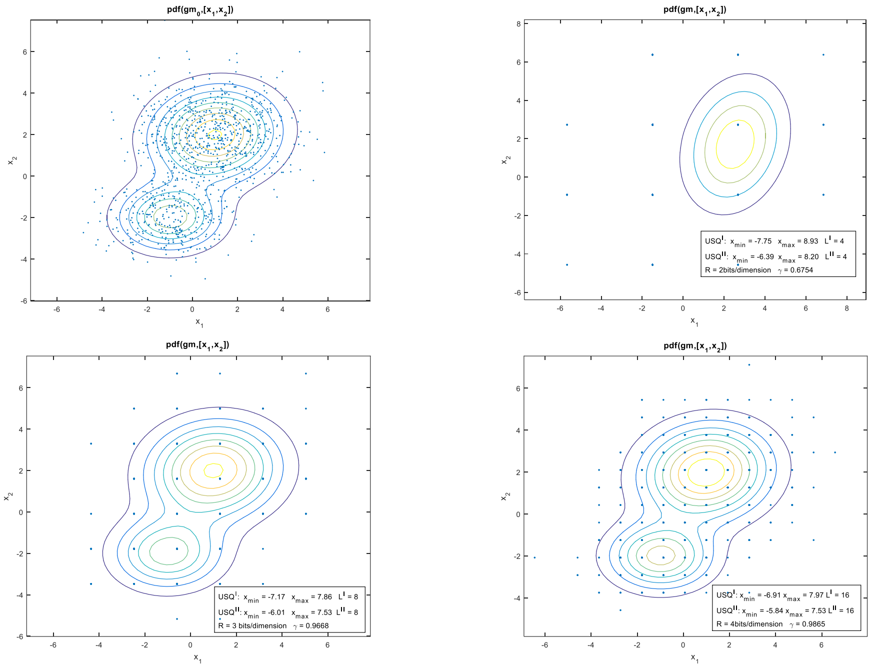

Fig. 5 The estimated probability density contours for the second data set and different R

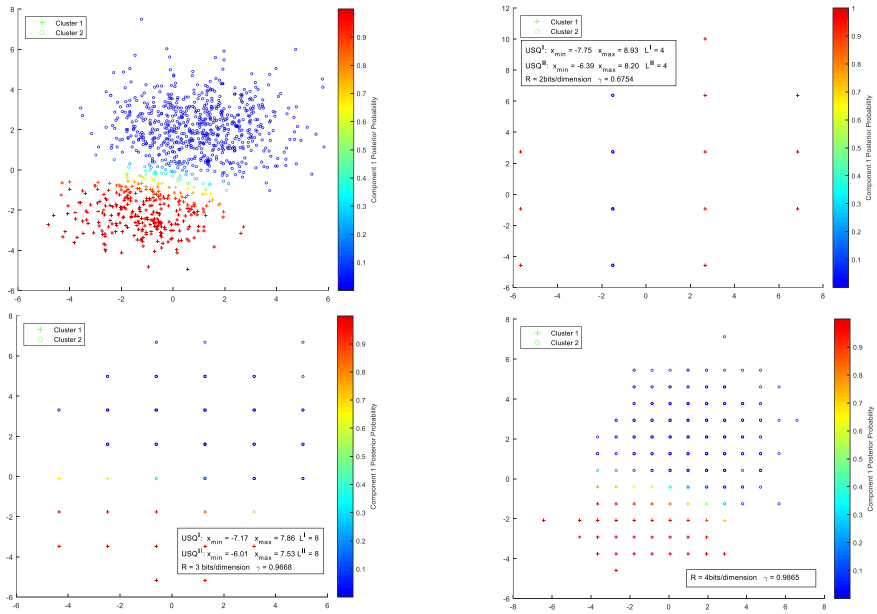

Fig. 6 Partition of the second data set into clusters for different R

From Figs. 3 and 6 we can notice that the number of possible data values is far smaller than in the case without quantization, indicating the significant compression in the number of bits used for data presentation and transmission. Actually, due to independent quantization of data attributes the possible data values are constrained to $L^2 = 2^{2R}$ uniformly distributed values. However, Figs. 3 and 6 show that the number of possible data values is even smaller than L^2. This is because the data whose probabilities of attributes are very small have a negligible probability of occurrence and therefore do not actually appear. This observation points out that there is a space for further compression in the required number of bits for data presentation. In order to achieve this it is necessary to perform joint quantization of data attributes. The design of joint attribute quantization of data modeled with two-dimensional GMM whose parameters are unknown is a complex issue left for future research.

4. Conclusion

In this paper we studied the influence of data attribute quantization on two-dimensional GMM-based clustering by using the EM algorithm. We considered independent quantization of attributes by using uniform quantizers with the support limits adjusted to the minimal and maximal values of corresponding attribute. With the aim to provide the accurate clustering we determined the necessary number of bits for data presentation. These findings are especially important for overlapping clusters. The presented analysis is useful in situation wherein before being grouped the data have to be represented with a lower accuracy due to their transmission through the bandwidth-limited channels. In our future research, we are going to optimize values for support limits, as well as the bit allocation between USQs used to quantize attributes of data modeled with two-dimensional GMM.

Acknowledgement: *This research was supported by the Science Fund of the Republic of Serbia, 6527104, AI-Com-in-AI.*

References

[1] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[2] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, 2016.

[3] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, *Data Mining: Practical Machine Learning Tools and Technique*s, 4th ed., Morgan Kaufmann Series in Data Management Systems, 2016.

[4] A. P. Dempster, N. M. Laird, D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," Journal of the Royal Statistical Society, vol. 39, no. 1, pp. 1–38, 1977. [Online]. Available: https://web.mit.edu/6.435/www/Dempster77.pdf

[5] X. Lin, X. Yang, Y. Li, "A deep clustering algorithm based on Gaussian mixture model," ISAI 2019, Journal of Physics: Conference Series, 1302, 2019. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1302/3/032012/pdf

[6] M. S. Yang, C. Y. Lai, C. Y. Lin, "A robust EM clustering algorithm for Gaussian mixture models," Pattern Recognition, vol. 45, no. 11, pp. 3950–3961, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0031320312002117

[7] A. Cedric, L. A. John, V. Michael, "On convergence problems of the EM algorithm for finite Gaussian mixtures," in *Proceedings of 11th European Symposium on Artifical Neural Networks*, Bruges, Belgium, pp. 99–106, 2003. [Online]. Available: http://www0.cs.ucl.ac.uk/staff/c.archambeau/publ/esann_ca03.pdf

[8] A. Jovanović, Z. Perić, D. Aleksić, J. Nikolić, "The effect of uniform data quantization on GMM-based clustering by means of EM algorithm", *20th International Symposium INFOTEH-JAHORINA*, VRT-2.7 (45), March 17-19, 2021, Jahorina, RS, B&H.
[9] I. H. Tseng, O. Verscheure, D. S. Turaga, U. V. Chaudahari, "Quantization for adapted GMM-based speaker verification," in *Proceedings of IEEE ICASSP*, Toulouse, France, May 2006. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.332.4539&rep=rep1&type=pdf
[10] I. H. Tseng, O. Verscheure, D. S. Turaga, U. V. Chaudhari, "Optimized one-bit quantization for adapted GMM-based speaker verification," in *Proceedings of INTERSPEECH*, Antwerp, Belgium, August 27–31, pp.786-789, 2007. [Online]. Available: https://www.isca-speech.org/archive/archive_papers/ interspeech_2007/i07_0786.pdf
[11] N. S. Jayant, P. Noll, *Digital Coding of Waveforms*, Englewood Cliffs, NJ: Prentice-Hall, 1984.
[12] D. Hui, D. L. Neuhoff, "Asymptotic analysis of optimal fixed-rate uniform scalar quantization," IEEE Transaction on Information Theory, vol. 47, no. 3, pp. 957–977, 2001. [Online]. Available: https://www.researchgate.net/publication/3080367_Asymptotic_analysis_of_optimal_fixed-rate_uniform_scalar_quantization
[13] S. Na, D. Neuhoff, "Monotonicity of step sizes of MSE-optimal symmetric uniform scalar quantizers," IEEE Transaction on Information Theory, vol. 65, no. 3, pp. 1782−1792, 2019.
[14] S. Na, D. Neuhoff, "On the convexity of the MSE distortion of symmetric uniform scalar quantization," IEEE Transaction on Information Theory, vol. 64, no.4, pp. 2626−2638, 2018.
[15] Z. Eskić, Z. Perić, J. Nikolić, "A method of designing adaptive uniform quantizer for LPC coefficient quantization," Przeglad Elektrotechniczny, vol. 87, no. 7, pp. 245–248, 2011. [Online]. Available: https://www.researchgate.net/publication/228520367_A_method_of_designing_an_adaptive_uniform_quantizer_for_LPC_coefficients_quantization
[16] A. Jovanović, Z. Perić, J. Nikolić, "Iterative algorithm for designing asymptotically optimal uniform scalar quantization of the one-sided Rayleigh density," IET Communications, pp. 1–7, February 2021. [Online]. Available: https://doi.org/10.1049/cmu2.12114
[17] Z. Perić, J. Lukić, J. Nikolić, D. Denić, "Application of mean-square approximation for piecewise linear optimal compander design for Gaussian source and Gaussian mixture model," Information Technology and Control, vol. 42, no. 3, pp. 277–285, 2013. [Online]. Available: https://doi.org/10.5755/ j01.itc.42.3.4349
[18] A. Ž. Jovanović, Z. H. Perić, "Geometric piecewise uniform lattice vector quantization of the memoryless Gaussian source," Information Sciences,vol. 181, no. 14, pp. 3043–3053, 2011.
[19] A. Ž. Jovanović, Z. H. Perić, J. R. Nikolić, "An efficient iterative algorithm for designing an asymptotically optimal modified unrestricted uniform polar quantization of bivariate Gaussian random variables," Digital Signal Processing, vol. 88, pp. 197–206, May 2019.
[20] A. Ž. Jovanović, Z. H. Perić, J. R. Nikolić, M. R. Dinčić, "Asymptotic analysis and design of restricted uniform polar quantizer for Gaussian sources," Digital Signal Processing, vol. 49, pp. 24–32, February 2016.
[21] Z. Perić, J. Nikolić, "Asymptotic analysis of switched uniform polar quantization for memoryless Gaussian source," IEEE Signal Processing Letters, vol. 20, no. 1, pp. 75–78, 2013.
[22] Z. H. Perić, M. D. Petković, J. R. Nikolić, A. Ž. Jovanović, "Support region estimation of the product polar companded quantizer for Gaussian source," Signal Processing, vol. 143, pp. 140-145, 2018.