

# Similarity Evaluation of 3D Gray Rock Image Using Pattern Density Classification Function

Xiaohai He, Zhengji Li, Qizhi Teng\*, Linbo Qing and Xiaohong Wu

College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China

\*Corresponding author

**Abstract**—Aiming at the problem that the existing 3D core similarity evaluation methods cannot effectively evaluate gray core images, we proposed a similarity evaluation algorithm based on Pattern Density Classification Function (PDCF). First of all, the 3D template is used to extract the texture patterns of 3D core images, and then the pattern density classification function is formed with the extracted patterns by adopting K-means algorithm. An adaptive method is used to find out the appropriate K value. Finally, the pattern density classification function is used to measure the texture similarity between 3D gray rock models.

In this paper, a comparative experiment of multiple groups of core images is carried out. Combined with the existing similarity evaluation algorithm for the binary image of 3D core, the 3D gray core model similarity characterization is realized from morphological and texture distribution.

**Keywords**—3D rock models; pattern density classification function; morphological similarity; texture similarity

## I. INTRODUCTION

Digital rock physics is a newly developed method based on imaging and digitizing of 3D pore and mineral structure of actual rock and numerically computing rock physical properties, such as permeability, elastic moduli and formation factor[1].

The studies on rock physics parameters and characterization of reservoirs enhances the understanding of their lifetime performance and management. Developing a comprehensive reservoir characterization function is an outstanding challenge that requires various tools for feature extraction, data analysis.

It is very difficult to characterize the rocks samples, because of the widely variety of morphological structure and texture distribution. Over the decades, the researchers have addressed in this area and put forward many methods to describe the similarity of binary 3D core image, such as porosity, autocorrelation function[2-4], linear-path[4] and so on. With these methods, the morphological distribution of binary image of rocks can be described. These methods are used to analyze the similarity of reservoir rocks and they are also used in the process of 3D core digital reconstruction methods[2].

### A. Autocorrelation Function

Autocorrelation function is also known as two-point probability function. It was introduced by Yeong[2] in 1997 as a constraint condition of 3D reconstruction of porous media and it was treated as a popular similarity evaluation method for

3D reservoir rocks model. The Autocorrelation function  $S_2^{(j)}(r_1, r_2)$  is defined as:

$$S_2^{(j)}(r_1, r_2) = \langle I^{(j)}(r_1) I^{(j)}(r_2) \rangle \quad (1)$$

Where  $r_1$  and  $r_2$  are two arbitrary points in the model, angular brackets denote an ensemble average, and the characteristic function  $I^{(j)}(r)$  can be calculated as:

$$I^{(j)}(r) = \begin{cases} 1, & r \in \text{phase } j \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

The quantity of  $S_2^{(j)}(r_1, r_2)$  can be interpreted as the probability of finding two points at positions  $r_1$  and  $r_2$  both in phase j. As it is related to distance between  $r_1$  and  $r_2$  only, the autocorrelation function can be expressed simply as  $S_2^{(j)}(r)$ , where  $r$  refers to the distance between two points.

### B. Lineal-path Function (LPF)

Another popular evaluation method for random heterogeneous materials is proposed by Lu[5] in 1992 is called as lineal-path function  $L^{(j)}(r)$ . The quantity of LPF is the probability that a line segment of length r is wholly in phase j and it is a more accurate measure of the characteristic length of phase j.

The lineal-path function is defined as:

$$L^j(r) = \overline{P(x, x + r)} \quad (3)$$

Where,

$$P(x, x + r) = \begin{cases} 1, & \vec{r} \in \text{phase } j \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

The LPF can reflect the size of the phase  $j$  and the connectivity in the heterogeneous porous materials.

However, these functions are used to morphology similarity measurement in many cases for two phase models[6-9], they cannot describe the texture difference for 3D reservoir rock models in gray level.

In this paper, we proposed an evaluation algorithm based on the classification of pattern density. The 3D template is used to extract the texture pattern of core images, and then the pattern density classification function (PDCF) is formed with the K-means algorithm. With the pattern density classification function, the texture similarity between 3D gray rock images can be measured. Combing with autocorrelation function and LPF, the similarity of reservoir rock models can be evaluated both in morphology and texture.

## II. EXACT REPRESENTATION OF THE PATTERN DENSITY CLASSIFICATION FUNCTION

It is assumed that  $V$  is a 3D gray image to be evaluated. A template with size is used to extract the pattern  $P$  from  $V$ , and the sampling center is  $V(X, Y, Z)$ .

$$P = V(X \pm x, Y \pm y, Z \pm z) \quad (5)$$

$$\text{Where } x, y, z \in \left[-\frac{L}{2}, \frac{L}{2}\right].$$

The pattern set  $P_{set}$  is generated by sampling the whole model  $V$  point by point.

$$P_{set} = \{P_i\}, i = 1, 2, \dots, M \quad (6)$$

Where  $M$  is the number of all patterns. We assume  $K_n$  is all possible patterns for  $P$  and  $n$  is the number of all possible patterns.  $N_{K_n}$  indicates the number of times  $K_n$  appears in the collection. The pattern density function(PDF)  $f(K_n)$  is defined as:

$$f(K_n) = N_{K_n} \quad (7)$$

The distribution of the texture pattern feature  $K_n$  on  $V$  can be reflected by PDF. During the calculation of PDF for 3D model, the number of  $K_n$  can be huge. In this case, according to the different core structure, texture complexity, template size and different size of core samples, the number of pattern will usually be above one hundred thousand orders of magnitude, which is not easy for similarity measurement. As this reason, we use the K mean clustering algorithm (KMEANS) to classify

the sampled pattern set according to texture similarity instead of PDF.

Random centroid  $\{u_1, u_2, \dots, u_k\}$  are selected for  $K$  cluster from the patter set . $K$  clusters are created by associating each atom in the  $P_{set}$  with the nearest mean according to formula 8.

$$C(i) = \arg \min_{1 \leq j \leq k} \|P_i - u_j\|^2, i = 1, 2, \dots, M \quad (8)$$

According to the atoms distribution of each cluster, the centroid of each of the  $K$  clusters is set as the new mean. The criterion function  $E$  is calculated as:

$$E = \sum_{i=1}^R \sum_{P_i \in C_j} |P_i - m_i|^2 \quad (9)$$

Where  $E$  is the sum of all the square errors of the atoms in the data set and  $m_i$  is the average value of all members in the cluster  $C_j$ . The center of each cluster is adjusted repeatedly and the cluster members are updated until the cluster center is no longer changed or the criterion function  $E$  is convergent of the maximum number of the iterations is reached.

The pattern density classification function (PDCF) can be calculated as:

$$G(x) = \frac{\text{crad}(C_j)}{M} \quad (10)$$

Where  $\text{crad}(C_j)$  is the number of atom in cluster  $j$  and  $M$  is the number of all patterns. Finally the PDCF  $G(x)$ , the centroid location of each cluster  $\text{Cent}_j$  and the mean distance  $D$  from point to every centroid within-cluster are obtained.

In this algorithm, the number of cluster and the size of template are two critical parameters. As shown in figure I, the mean distance from point to every centroid within-cluster decreases with the increasing of the number of clusters. When the value of  $K$  increases to a certain degree, the decline area of the  $D$  is slow.

In our approach, an adaptive method is used to find the appropriate  $K$  value. The initial value of  $K$  is set to 5 and the initial length of template is set to 10, then the value of  $K$  is increased by 5, and difference of the mean distance  $D$  is calculated. When the difference is smaller than  $\varepsilon$ , it is considered that the mean distance  $D$  descending trend has tended to be slow, and the value of current  $K$  is selected. The length of template  $L$  is increased by 2, if the difference is still large than  $\varepsilon$  when the value of  $K$  reaches 100. The whole producer is processed repeatedly until the difference is small than  $\varepsilon$  or the  $K$  and  $L$  reach the maximum value.

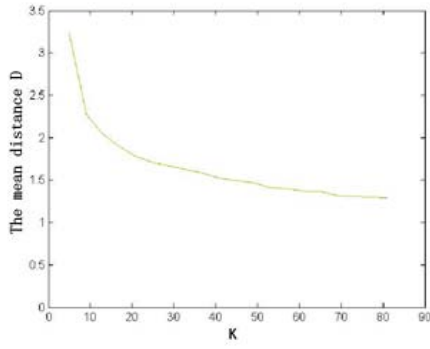


FIGURE I. THE RELATION BETWEEN THE MEAN DISTANCE D WITHIN-CLUSTER AND K

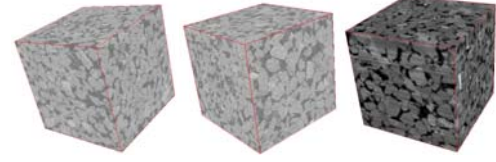
The procedure of calculating PDCF is shown in Table I:

TABLE I. THE PROCEDURE OF EVALUATING THE TEXTURE SIMILARITY FOR DIFFERENT ROCK SAMPLES WITH PDCF

1. Set  $\varepsilon = 0.05$
2. Extract the patterns  $P_{setA}$  from sample A
3. For  
[  $10 \leq L \leq \text{half size of side length of the rock sample}$  ]
4. Set  $K = 5$
5. Cluster the patterns  $P_{setA}$  and calculate the mean distance  $D$ , the centroid location of each cluster  $Cent_j$  with  $K$  clusters and the PDCF  $G(x)$
6.  $D_{min} = D, Cent_{min} = Cent_j, G_{min}(x) = G(x)$
7. For [  $10 \leq K \leq 100$  ]
8. Cluster the patterns  $P_{setA}$  with  $K$  clusters and calculate the mean distance  $D'$ , the centroid location of each cluster  $Cent'_j$  with  $K$  clusters and the PDCF  $G'(x)$
9. If (  $D' < D$  and  $D - D' \leq \varepsilon$  )
10.  $D_{min} = D', Cent_{min} = Cent'_j, G_{min}(x) = G'(x)$ ,  
Jump to 20
11. Else if (  $D' < D$  and  $D - D' > \varepsilon$  )
12.  $D = D', Cent_{min} = Cent', G_{min}(x) = G'(x)$
13. End 12
14. End 9
15.  $K = K + 5$
16. End 7
17.  $L = L + 2$
18. End 3
19.  $D_{min} = D', Cent_{min} = Cent'_j, G_{min}(x) = G'(x)$  , (if the constraint condition  $D' < D$  and  $D - D' \leq \varepsilon$  cannot be satisfied, the  $P_{setA}$  is clustered with  $L = \text{half size of side length of the rock sample}$  and  $K = 100$ )
20. Extract the patterns  $P_{setB}$  from sample B
21. Cluster the  $P_{setB}$  according to the centroid location of each cluster  $Cent_j$  and calculate the  $G(y)$
22. Evaluate the similarity between sample A and B with  $G_{min}(x)$  and  $G(y)$

### III. EXPERIMENT AND RESULT ANALYSIS

In order to verify the effectiveness of the algorithm, three sets of cores with size  $128 \times 128 \times 128$  are selected for comparison. The sample A and sample B are two adjacent blocks in the same homogeneous core, and the sample C is cut from another core image. Visually, sample A and sample B are similar in morphology and texture. Sample C has the similar morphology distribution but different in texture.



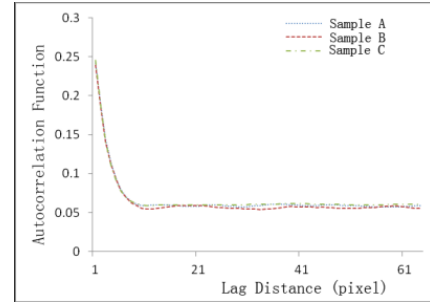
Sample A

Sample B

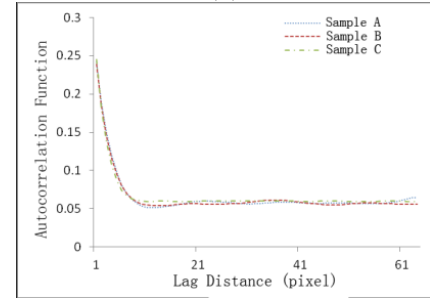
Sample C

FIGURE II. THE ILLUSTRATION OF THREE SAMPLES OF RESERVOIR ROCKS

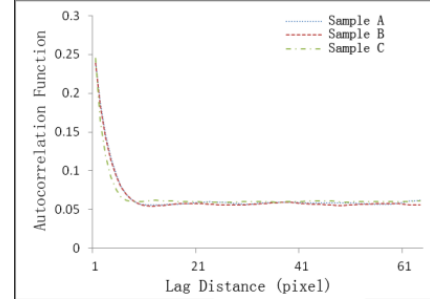
In the following experiment, the 3 models are transform into binary images and autocorrelation function, lineal-path function (LPF) are used to evaluate the morphological similarity and the pattern density classification function (PDCF) are used to evaluation the similarity of pattern distribution of texture between three samples directly. The result is shown in Figure III.



(A)



(B)



(C)

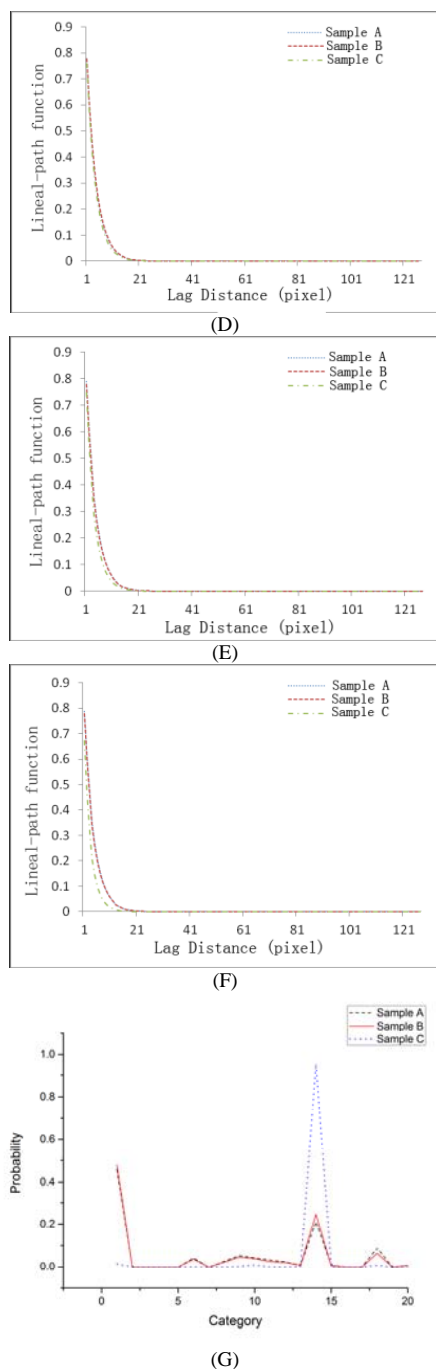


FIGURE III. COMPARISON OF AUTOCORRELATION FUNCTIONS, IN (A) THE X DIRECTION, (B) THE Y DIRECTION, AND (C) THE Z DIRECTION, LINEAL-PATH FUNCTION, IN (D) THE X DIRECTION, (E) THE Y DIRECTION, AND (F) THE Z DIRECTION, AND PATTERN DENSITY CLASSIFICATION FUNCTION (G)

As Figure III (A)-(F) shown, three samples have the similarity distribution on autocorrelation function and lineal-path function in x, y and z directions. It indicates that three samples are similar in morphological characteristics. With autocorrelation function and lineal-path function, we cannot distinguish the different between sample A, B and C.

According to (G) of figure III, the PDCF of sample A and sample B are very similar, while there is a large gap between PDCF of sample C and that of sample A(sample B).

The figure III reveals that three reservoir rocks are similar in morphological characteristics, while sample A and sample B have similar pattern distribution, but sample C has different pattern distribution to sample A and B. This conclusion is quite agreed with the prior knowledge.

In order to further verify the effectiveness and stability of the algorithm, 10 groups of homogeneous core samples are selected, and each group of core samples is composed of two 3D CT models which are the adjacent area of the same isotropic core with similar morphology and pattern distribution. These images are marked as: sample 1(a), sample 1(b), sample 2(a), sample 2(b) and so on. Then, we use the PDCF to calculate the similarity of pattern distribution of texture between the samples.

For convenience of evaluation, we calculated the mean square error of PDCF, and the result is presented in tabular form in Table II.

In Table II, each line is the mean square error of PDCF between sample (b) with other 10 sets of sample (a). The smaller of the difference indicates the more similar of samples in the pattern distribution of texture. As table II shown, the two samples which are from the same rock model are much similar than others.

TABLE II. THE MEAN SQUARE ERROR OF PDCF FOR 10 SETS OF RESERVOIR ROCKS

	Sample 1 (a)	Sample 2 (a)	Sample 3 (a)	Sample 4 (a)	Sample 5 (a)	Sample 6 (a)	Sample 7 (a)	Sample 8 (a)	Sample 9 (a)	Sample 10 (a)
Sample 1 (b)	<b>0.006621</b>	0.06753	0.036041	0.031458	0.055751	0.05473	0.135982	0.429998	0.499629	0.133445
Sample 2 (b)	0.083717	<b>0.003029</b>	0.023497	0.013298	0.270761	0.127632	0.411091	0.060531	0.228607	0.259259
Sample 3 (b)	0.04935	0.05209	<b>0.011562</b>	0.040549	0.179598	0.089381	0.3111	0.036109	0.488982	0.53065
Sample 4 (b)	0.041342	0.041752	0.027695	<b>0.003414</b>	0.189487	0.074355	0.310717	0.026394	0.25563	0.326633
Sample 5 (b)	0.071218	0.287161	0.179833	0.194669	<b>0.006333</b>	0.042207	0.021947	0.097204	0.650674	0.884639
Sample 6 (b)	0.004541	0.037861	0.030336	0.018843	0.010505	<b>0.002709</b>	0.012773	0.010722	0.383701	0.366348
Sample 7 (b)	0.145174	0.32654	0.265525	0.260605	0.033636	0.090322	<b>0.00284</b>	0.166194	0.23614	0.427065
Sample 8 (b)	0.001625	0.029112	0.014123	0.013204	0.058204	0.013345	0.14205	<b>0.0004</b>	0.025857	0.081643
Sample 9 (b)	0.367324	0.175331	0.311081	0.215047	0.582639	0.392537	0.697382	0.139994	<b>0.006076</b>	0.024013
Sample 10 (b)	0.552535	0.298655	0.474819	0.349208	0.812024	0.594515	0.956796	0.519526	0.041177	<b>0.003174</b>

#### IV. CONCLUSION

To accurately measure the pattern distribution features of reservoir rocks, the pattern density classification function is proposed in this paper. The 3D template is used to extract the patterns of 3D core gray images, and then the pattern density classification function (PDCF) is formed with the extracted patterns by adopting K-means algorithm and an adaptive method is used to find out the appropriate K value. Combining PDCF, autocorrelation function and lineal-path function, the similar of reservoir rocks can be measured in morphological and texture.

Several experiments are used to comprehensively test the proposed algorithm. The experiments demonstrate that the autocorrelation function and lineal-path function can reveal the morphological similarity of rocks samples and the pattern density classification function can describe the similarity of reservoir samples in texture distribution.

We believe that the algorithm proposed in this paper can be used for practical applications. The method also has potential applications in similarity compassion in the reconstruction of variety of porous media.

#### ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundations of China (No. 61372174) and (No. 61471248).

#### REFERENCES

- [1] Karimpouli S, Tahmasebi P. "Conditional reconstruction: An alternative strategy in digital rock physics", J. Geophysics, vol. 81(4), pp. 465-D477, 2016.
- [2] Yeong C L Y, Torquato S. "Reconstructing Random Media", J. Physical Review E, vol. 57(1), pp. 495-506, 1998.
- [3] Staraselski Y, Brahme A, Mishra R K, Inal K. "Reconstruction of the 3D representative volume element from the generalized two-point correlation function", J. Modelling and Simulation in Materials Science and Engineering, vol. 23(1), pp. 15007-15030, 2015.
- [4] A. Ghazavizadeh, N. Soltani, M. Baniassadi, F. Addiego, S. Ahzi, H. Garmestani, "Composition of two-point correlation functions of

subcomposites in heterogeneous materials", J. Mech. Mater. Vol. 51, pp. 88-96, 2012.

- [5] Lu B, Torquato S. "Lineal-Path Function for Random Heterogeneous Materials. I. Effect of Polydispersivity". Physical Review A, vol. 45(10), pp. 7292-7301, 1992.
- [6] PejmanTahmasebi, Farzam Javadpour, Muhammad Sahimi, Mohammad Piri, "Multiscale study for stochastic characterization of shale samples", J. Advances in Water Resources, pp. 89:91-103, 2016.
- [7] Karimpouli S, Tahmasebi P., "Conditional reconstruction: An alternative strategy in digital rock physics", J. Geophysics, vol. 81(4), pp. 465-477, 2016.
- [8] Gao Mingliang, Teng Qizhi, He Xiaohai, Zuo Chen, Li Zhengji. "Pattern density function for reconstruction of three-dimensional porous media from a single two-dimensional image", J. Physical Review E, vol. 93(1), pp. 012140, 2016.
- [9] Tahmasebi P, Sahimi M. "Enhancing multiple-point geostatistical modeling: 1. Graph theory and pattern adjustment", J. Water Resources Research, vol. 52, pp. 2074-2098, 2016.
- [10] Gerke Kirill, Karsanina Marina, Vasilyev R V, Mallants Dirk, "Improving pattern reconstruction using directional correlation functions", J. Epl, vol. 106, pp. 66002, 2014.