Analysis of the AE signal parameters during granite compression process based on the K-means clustering method

Feng Xu\textsuperscript{1,2,3}, Jiaming Xiang\textsuperscript{3}, Yao Li\textsuperscript{4}, Huibo Wen\textsuperscript{1,2}, Zhongshuai Liu\textsuperscript{1,2}, Xiaofei Ma\textsuperscript{1,2}, Guangqin Yang\textsuperscript{1,2}

\textsuperscript{1}China Merchants Chongqing Communications Research & Design Institute Co., Ltd, Chongqing, China
\textsuperscript{2}National local joint engineering laboratory for road engineering and disaster prevention and reduction technology in Mountainous Areas, Chongqing, China
\textsuperscript{3}School of Civil Engineering, Chongqing Jiaotong University, Chongqing, China
\textsuperscript{4}Highway Bureau of Tibet Autonomous Region, Tibet, China

Abstract. The average frequency (AF), rise-time divided by amplitude (RA), and peak frequency of acoustic emission (AE) generated during the granite uniaxial and triaxial compression failure process are taken as the main factors of the type of sample crack. Further, AE signals can be characterized based on the K-means clustering method. Results show that there is a large difference between the AE signals generated by uniaxial and triaxial compression. The former has more diversity in the category, e.g., better clustering effect and uniaxial or triaxial compression. AE signal parameters can be divided into four categories based on the K-means clustering method. Acoustic signals A, B, C, D in the two compression modes are substantially similar and uniaxial compression mainly generates tensile crack before the peak stress, and shear crack appears after the peak stress. During the triaxial compression process, tensile and shear cracks are both generated before the peak stress, but shear crack is no longer present after the peak stress.

Keywords: acoustic emission; hit; K-means clustering; crack type

1. Introduction
When the rock is deformed under stress, microcracks appear inside it. Acoustic emission (AE) explains stress waves caused by sudden release of some energy during the generation, expansion, closure, and coalescence of microcracks [1-2]. AE technology has been widely employed in geotechnical laboratory experiments and engineering sites because of its promising outcomes [3-15].

Scholars around the world have conducted many research works on AE and explored the evolution laws of AE events, energy, amplitude, and frequency during uniaxial, triaxial, tensile, shear, and hydraulic fracturing tests. However, most scholars have taken into account a single AE parameter for analysis. Jiang [4], Li [5], Xu [6], and a few other scholars studied the relationship between the amount of AE events, the rate, and the time on different rocks under uniaxial or triaxial test. Moradian [7] studied the change in AE count and energy over time for granite under different shear conditions. Furthermore, Hou [8] and Song [9] studied the number of AE events during hydraulic fracturing. Different AE signals are triggered by various failure mechanisms during the stress process of rock specimens. In this regard, Gu [10] studied the microstructure and fractal characteristics of shale during...
the failure of triaxial compression using rise-time divided by amplitude (RA). Moreover, Dimitrios [11] and Farhidzadeh [12] divided the cracks into shear cracks and opening cracks based on RA and AF in the process of cement block failure. Also, Ji [13] investigated the change of the peak frequency of the AE signals during the granite failure process, where the results showed the distribution of peak frequency increased from 1–2 main frequency bands at the initial stage of loading to 5 main frequency bands at failure.

Although a few researchers have investigated crack characteristics based on one or two parameters, the classification is according to certain experiences [10–11]. The boundary of particular classifications has no theoretical basis. The classification of cracks is related to RA, AF, peak frequency, and so on. In order to consider the impacts of RA, AF, and peak frequency on the classification of cracks in a comprehensive manner, this study classifies the cracks by the K-means clustering method during the uniaxial and triaxial compression of granite and investigates the evolution law of the AE signals during the granite failure process.

2. K-means clustering method

2.1. Selection and standardization of AE parameters

According to the results of previous research works [1-13], five AE parameters including rise time, amplitude, count, duration, and peak frequency are the basic parameters of classification. It is worth to underline that parameters cannot be analyzed directly because of the different dimensions and orders of magnitude of different AE parameters. Further, the parameter with a greater value will play a highlighted role in the K-means clustering analysis. Therefore, in order to secure reliability of the results, the original parameter data needs to be normalized [16]. The five AE parameter data are normalized as follows:

\[ x'_i = \frac{x_i - \bar{x}}{\sigma} \]

\[ \sigma = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2} \]

Where, \( x_i \) and \( x'_i \) are initial and normalized parameters, and \( \bar{x} \) is average of the parameters.

2.2. The similarity measure principle

A similarity measure principle needs to be applied between signal parameters before the AE signals are divided into different clusters. The Euclidean distance, between two points in m-dimensional space, is employed as the similarity measure principle. The parameters of AF, RA, and peak frequency are considered as the three dimensions of space in this study. The similarity spatial distance of two AE behaviors is:

\[ d(x_i, x_j) = \sqrt{\sum_{m=1}^{3} (x_{im} - x_{jm})^2} \]

Where, \( d(x_i, x_j) \) represents Euclidean distance between the independence AE signals \( x_i \) and \( x_j \); \( x_{im} \) and \( x_{jm} \) represent values of the AE parameters.

2.3. Calculation principle of the K-means clustering method

The K-means algorithm was proposed by MacQueen [17] in 1967. It has been widely utilized in the classical clustering algorithms in scientific researches and industrial applications. Its core idea is to divide n vector objects into k clusters so that the distance from the data points in each cluster to the center of the cluster to become the smallest. The process is as follows:
(1) Assume that all AE signals have \( k \) clusters, and the initial cluster centers given to \( k \) clusters randomly or according to a certain rule. The initial clustering center of \( k \) clusters is \( c_i \), where the value of \( i \) is 1 to \( k \).

(2) Calculate the Euclidean distance from the \( n \) vectors in space to each cluster center \( x \) in turn. The vectors will be assigned to the cluster with the closest Euclidean distance to the cluster center, respectively.

(3) The mean of each category is re-used as the center of each cluster \( c_i \) after all vectors are assigned.

(4) Repeat steps 2 and 3 until the position of the cluster center converges. Then, the clustering process ends.

### 2.4. The value of \( k \)

Accuracy of the K-means algorithm depends on the selection of the number of clusters. The selection of the number of clusters directly influences the rationality and reliability of vector partitioning. In order to reduce the impact of choosing different clustering numbers on the results of classification, this research employs different clustering numbers \( k \), where \( k \) is between 2 and 10. Considering the contour value as a criterion, the number of clusters, selected in the end, is determined. Silhouette coefficient \( s \) can be expressed as:

\[
s = \frac{1}{n} \sum_{i=1}^{n} \frac{[\min(b(i,k) - a(i))] - \max[a(i), \min(b(i,k))]}{\max[a(i), \min(b(i,k))]}\]

where, \( b(i,k) \) is average distance, which represents the average distance between vector \( i \) and the points of other clusters \( k \); and, \( a(i) \) represents average distance between vector \( i \) and other points in the same cluster. The larger the value of \( s \), the more divided classes, i.e. the better the clustering effect. The value of \( k \) corresponding to the maximum value of \( s \) is selected as the number of clusters, which indicates that the cluster classification effect is the best.

### 3. Testing equipment and experimental design

The rock triaxial test machine (Model: XTR01-01) and AE monitoring system supplied by American Physical Acoustics Company (PAC) were utilized in the study. It can be used in a laboratory triaxial test, which overcomes the problems of incomplete signal reception or serious attenuation caused by placing the probe in the outdoor triaxial test. The physical map of the modified probe is shown in Figure 1.

The test specimens were drilled on the same granite without obvious defect and cylinder with dimensions \( \Phi 100 \times 200 \). Both ends are guaranteed to be smooth and flat, and the parallelism fulfills the requirements of the rock mechanics test (within \( \pm 0.02 \) mm) to avoid uneven stress. After measuring its mass and volume, the average density is calculated as 2.61 g/cm\(^3\). By employing the displacement loading control method with a loading rate of 0.08 mm/min, the specimens were subjected to uniaxial and triaxial tests, respectively. The confining pressure is constant at 2 MPa during the triaxial test. The test schematic is shown in Figure 2.

![Figure 1. High-pressure resistant AE probe](image-url)
1. Indenter; 2. Granite specimen; 3. High-pressure resistant AE probe
4. Amplifier; 5 Industrial PC

Figure 2. Schematic diagram of the test setup

4. Analysis and comparison of test results

4.1. Analysis of uniaxial test results

During the uniaxial compression of granite, more than 25,000 AE hits were received. According to the five parameters of rising time, amplitude, counting, duration, and peak frequency, we classify them using AF, RA, and peak frequency as clustering variables after sorting. Matlab K-means clustering method is utilized to select various clustering numbers k and calculate the corresponding value of the silhouette coefficient s. The s–k curve calculated by the K-means clustering method for the AE signals during uniaxial compression is shown in Figure 3. As can be seen from the figure, when k = 4, the maximum value of s is reached at 0.7648. This, in turn, indicates that the classification effect is the best when the AE signals are divided into four categories. Hence, the AE signals are divided into four categories during uniaxial compression test.

Figure 4 shows that the AE signals are divided into A, B, C, and D categories during uniaxial compression. Figure 4(a) shows the distribution of the AF-RA relationship of AE clusters, where the RAs of categories A, B, and D are mainly at 0–2 ms/v, while the RA of category C has a clear boundary with categories A, B, and D, which is mainly at 2–15 ms/v. According to previous studies [10-12], category C has the characteristics of high RA, low AF, and obvious shear crack features. Categories A, B, and D are characterized with high AF, low RA, and obvious tensile crack features. Among them, AF and RA values of the categories A and D are in the same range. Therefore, it is difficult to distinguish between them. AF values of the categories A and D are in the range of 20–350 kHz, and their RA values are in the range of 0–2 ms/v, while AF values of the category B are extremely high and its RA values approach zero. Figure 4(b) shows the RA-peak frequency distribution of the AE clustering. Classification of the four categories A, B, C, and D is clear and visible. Category A is in the low RA value ranges and low peak frequency range. Category D is in the range of low RA values and high peak frequency. The peak frequency range of category B is widely distributed at 100–300 kHz, but the RA value is close to zero. Category C has a frequency range of 100–300 kHz, but its RA values can reach 3–15 ms/v at 100–200 kHz, while RA values are 3–8 ms/v at 200–300 kHz. Figure 4(c) is the AF-peak frequency distribution of the AE clustering. Hence, the four categories A, B, C, and D are clearly classified with their particular characteristics described above.

Figure 5 demonstrates the stress-time curve and the change process of different types of the AE signals during uniaxial compression. In this research, this process is divided into three stages. The initial stage (0–800 s) is a quiet period, and the number of impacts is almost negligible compared with the second stage. The second stage (801–1050 s) is a growth period, where A and B dominate with a small number of D, and finally, there are very few C. The number of categories A, B, and D has increased
significantly over time, and category C has appeared at the end. It reveals that before the peak stress is reached, mainly the tensile type categories A and B and a small amount of category D are generated, and no shear category C appears. The peak frequency distribution is relatively scattered at this stage, and there is no obvious concentration area. The last stage (1051 s-end) is the deceleration period, where time category A disappears and categories B and C are the main types. This indicates that after reaching the peak stress, the tensile category A with high peak frequency disappears, the tensile category B with low peak frequency continues to exist and the shear category C starts to appear.

Figure 3. S-k curve under uniaxial compression
Figure 4. Clustering of uniaxial compressed AE signals
4.2. Analysis of triaxial test results

In the process of granite triaxial compression, more than 46,000 AE activities were received, which was more than that of the process of uniaxial compression. The K-means clustering method, which was the same classification method as uniaxial compressed AE signals, was adopted for classification. Figure 6 demonstrates the s-k curve of the AE signals obtained by the K-means clustering method in the process of triaxial compression. According to Figure 6, when k is four, the maximum value of s is reached at 0.7209, which means the classification effect is the best with four categories of the AE signals. Figure 7 demonstrates that the AE signals in the process of triaxial compression can be divided into four categories A, B, C, and D. Figure 7(a) depicts distribution of the AF-RA relationship of the AE clustering, where the RAs of categories A, B, and D are mainly in the range of 0–2.5 ms/v. Category C, with an obvious boundary between categories A, B, and D, is mainly in the range of 2.5–15 ms/v. Further, category C has obvious characteristics of shear cracks, while categories A, B, and D have obvious characteristics of tensile cracks. Among them, the AF and RA values of categories A and D are in the same range. Their AF values are in the range of 20–300 kHz, and RA values are in the range of 0–2.5 ms/v. Hence, they are difficult to distinguish. The AF values of category B are extremely high and its RA values approach zero. Figure 7(b) shows RA-peak frequency distribution of the AE clustering. Classification of the four categories is clear and visible. Category A is in low RA and low peak frequency range. Category D is in low RA and high peak frequency range, and the peak frequency range of category B has a wide range of 100–350 kHz, but its RA values approach zero. The peak frequency of category C is in the range of 100–300 kHz, and its RA is 3–20 ms/v. Figure 7(c) shows AF-peak frequency distribution of the AE clustering. The four categories are clearly classified and characteristics of each category have been described.

Figure 8 demonstrates the stress-time curve and changes of different categories of the AE signals during triaxial compression. This process is divided into three stages. The initial stage (0-1100 s) is a quiet period, and the number of hits is almost negligible compared to the second stage. The second stage (1101-1350 s) is a growth period. Furthermore, A and D are the main categories, there are a small number of category C, and finally, a small number of category B appears. The number of categories A and D increase significantly with the time increment. It proves that before the peak stress is reached, main categories A, B, D with tensile properties and a small amount of category C with shear properties appear. At this stage, all four categories of the AE signal clustering are generated. The last stage (1351 s-end) is the deceleration period, when category C disappears, categories A and B are dominant, and category D has a certain amount. It proves that after reaching the peak stress, the shear category C disappears, while the tensile categories A, B, and D at each peak frequency band continue to be generated.
Figure 6. S-k curve under triaxial compression
4.3. Test results analysis
The AE signals were clustered based on uniaxial and triaxial tests and the characteristics of the AE signals in the stress-time process were obtained through comparison.
Using the K-means clustering method for classification, value of k and s changes differently during the uniaxial and triaxial compression tests. When uniaxial compression is adopted and different k values are used, value of s differs greatly, and the range is 0.284. Under the influence of confining pressure in the triaxial test, for different k values, value of s is similar, and the range is only 0.137. It shows that the AE signal categories are more widely separated from each other in the uniaxial compression process. This implies that the clustering effect is better. In the triaxial compression process, damage of the triaxial compression process may be progressive due to the confining pressure. The generated
signal is relatively balanced, which is different from the non-uniformity of the uniaxial compression fracture form at different stages.

Comparing Figure 4(a) and Figure 7(a) reveals that the AE signals can be divided into tensile categories A, B, and D and shear category C in both uniaxial and triaxial compression tests. Among them, categories A and D have RA and AF values of the same range, which is difficult to distinguish. The RA of shear category C in uniaxial compression is more concentrated, while it is more dispersed in triaxial compression. This indicates that in triaxial compression test, RA range of this signal is wider, and the content is richer during the shear process. Comparing Figure 4(b) and Figure 7(b) demonstrates that both uniaxial and triaxial compression have little influence on peak frequency, i.e., all are in the range of 100–300 kHz.

Although the signals generated by uniaxial and triaxial compression are all classified into A, B, C, and D categories, the signals generated at different stages in the entire compression process are very different. In uniaxial compression, significant acoustic signals appear at around 800 s; while in triaxial compression, significant acoustic signals appear at around 1100 s. This was because of the limited radial deformation of the confined pressure-bound specimen, which caused the cracking and expansion of microcracks to be restricted. Before the peak stress is reached, A and B with tensile cracks are the main categories, and a small amount of category D with tensile cracks appears. However, category C with tensile cracks appears before the peak stress is reached. In triaxial compression, cracks of A, B, C, and D all appear before the peak stress is reached, but the shear crack of C is always less. After the peak stress reached, shear crack of C disappears. This is precisely the opposite of the generation and disappearance law of C cracks during uniaxial compression. According to the analysis of the gained results, the main reason behind this is that after the confining pressure is applied, the original shear cracks of C, caused by the lateral expansion, are severely restricted so that during the triaxial compression process the tensile cracks of A, B, and D are the main types with a small amount of crack C.

5. Conclusion

(1) The AE signals generated by uniaxial and triaxial compression processes are considerably different. In the uniaxial compression process, the categories of AE signals are separated more widely, which means better clustering effect is yielded.

(2) Regardless of uniaxial or triaxial compression, the generated AE signals can be divided into four categories according to the K-means clustering method, and the parameter distribution ranges of the four acoustic signals categories are roughly similar.

(3) In uniaxial compression, tensile cracks are the main crack type before the peak stress reached, and shear cracks appear after the peak stress is reached. During triaxial compression, tensile and shear cracks appear before the peak stress is reached, and shear cracks no longer appear after the peak stress is reached.

Acknowledgment

The authors would like to acknowledge the financial support of the National Key Research and Development Program of China (2018YFC1504903), and the Science and Technology Project of Jiangxi Provincial Department of Transportation (2020H0026, 2020H0052).

References

[1] JI Hongguang, Research and application of acoustic emission properties of concrete materials [M].Beijing: Coal industry press, 2004

[2] Qin Siqing, Introduction to rock acoustic emission technology [M].Xi’an: Southwest Jiaotong University press, 1993
[3] ZHANG Ru, XIE He-ping, LIU Jian-feng, et al. Experimental study on AE characteristics of rock failure under uniaxial multilevel loadings [J]. Chinese Journal of Rock Mechanics and Engineering, 2006, 25(12): 2584–2588.

[4] JIANG Deyi, CHEN Jie, REN Song, et al. Experimental study of strain rate effect and AE characteristics of salt rock under uniaxial compression [J]. Chinese Journal of Rock Mechanics and Engineering, 2012 (2).

[5] Yin Guangzi, Qin Hu, Huang Gun, Experimental study of characteristics of seepage and AE of gas-filled coal under different stress paths [J]. Chinese Journal of Rock Mechanics and Engineering, 2013, 32(7): 1315-1320.

[6] Xu, Feng, Yang, et al. Effect of bedding planes on wave velocity and AE characteristics of the Longmaxi shale in China [J]. Arabian Journal of Geosciences, 2017, 10(6):141.

[7] Moradian Z A, Ballivy G, Rivard P, et al. Evaluating damage during shear tests of rock joints using AEs[J]. International Journal of Rock Mechanics and Mining Sciences, 2010, 47(4): 590-598.

[8] HOU Bing, CHEN Mian, TAN Peng, et al. Monitoring of hydraulic fracture network by AE method in simulated tri-axial fracturing system of shale gas reservoirs [J]. Journal of China University of Petroleum, 2015, 39(1): 66-71.

[9] Song I, Suh M, Won K S, et al. A laboratory study of hydraulic fracturing breakdown pressure in table rock sandstone [J]. Geosciences Journal, 2001, 5(3): 263-271.

[10] GU Yilei, WANG Zepeng, LI Qingmiao, et al. Laboratory study on RA value fractal feature of shale AE under conventional triaxial compression [J]. Journal of Chongqing University, 2018.

[11] Aggelis D G., Classification of cracking mode in concrete by AE parameters [J]. Mechanics Research Communications, 2011, 38(3): 153-157.

[12] Farhidzadeh A, Mpalaskas AC, Matikas T E, et al. Fracture mode identification in cementitious materials using supervised pattern recognition of AE features [J]. Construction and Building Materials, 2014, 67: 129-138.

[13] JI Hongguang, LU Xiang, characteristics of AE and rock fracture precursors of granite under conventional triaxial compression [J]. Chinese Journal of Rock Mechanics and Engineering, 2015, 34(4): 694–702.

[14] Hardy H R., Geotechnical field applications of AE/MS techniques at the Pennsylvania State University: a historical review [J]. NDT & E International, 1994, 27(4): 191-200.

[15] Zimmer V L, Sitar N., Detection and location of rock falls using seismic and infrasound sensors [J]. Engineering Geology, 2015, 193: 49-60.

[16] Ma Liping, Standardization of statistical data -- dimensionless methods [J]. Beijing statistics, 2000, 34-35.

[17] Kang Haisong, Damage pattern recognition of thermal barrier coatings based on acoustic emission signal analysis [D]. Xiangtan: Xiangtan University, 2014.