

Exploring the proper statistical proxy to distinguish random structures

Mei Xie(解梅), Fei Xie(谢飞), Baoyu Song(宋宝玉), Qiaoyu Guo(郭桥雨), and Xuechen Jiao(焦学琛)
Citation: Chin. Phys. B, 2025, 34 (**5**): 058902. DOI: 10.1088/1674-1056/adbd2b
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

WT-FCTGN: A wavelet-enhanced fully connected time-gated neural network for complex noisy traffic flow modeling

Zhifang Liao(廖志芳), Ke Sun(孙轲), Wenlong Liu(刘文龙), Zhiwu Yu(余志武), Chengguang Liu(刘承光), and Yucheng Song(宋禹成)
Chin. Phys. B, 2024, 33 (**7**): 078901. DOI: 10.1088/1674-1056/ad3349

Speed limit effect during lane change in a two-lane lattice model under V2X environment

Can Jin(金灿), Guang-Han Peng(彭光含), and Fang-Yan Nie(聂方彦)
Chin. Phys. B, 2024, 33 (**3**): 038902. DOI: 10.1088/1674-1056/ad0bf6

Essential proteins identification method based on four-order distances and subcellular localization information

Pengli Lu(卢鹏丽), Yu Zhong(钟雨), and Peishi Yang(杨培实)
Chin. Phys. B, 2024, 33 (**1**): 018903. DOI: 10.1088/1674-1056/acd7ca

Self-similarity of complex networks under centrality-based node removal strategy

Dan Chen(陈单), Defu Cai(蔡德福), and Housheng Su(苏厚胜)
Chin. Phys. B, 2023, 32 (**9**): 098903. DOI: 10.1088/1674-1056/acd3e2

Percolation transitions in edge-coupled interdependent networks with directed dependency links

Yan-Li Gao(高彦丽), Hai-Bo Yu(于海波), Jie Zhou(周杰), Yin-Zuo Zhou(周银座), and Shi-Ming Chen(陈世明)
Chin. Phys. B, 2023, 32 (**9**): 098902. DOI: 10.1088/1674-1056/acd685

Exploring the proper statistical proxy to distinguish random structures

Mei Xie(解梅), Fei Xie(谢飞), Baoyu Song(宋宝玉), Qiaoyu Guo(郭桥雨), and Xuechen Jiao(焦学琛)[†]

National Synchrotron Radiation Laboratory, University of Science and Technology of China, Hefei 230029, *China*

(Received 11 November 2024; revised manuscript received 11 November 2024; accepted manuscript online 6 March 2025)

The canonical description of structures comprises two aspects: (1) basic structural elements and (2) arrangement pattern between those elements. This tidy description has been very successful and facilitates the development of structural physics tremendously, enabling the classification, comparison and analysis of an extremely wide range of structures, including crystals, quasi-crystals, liquid crystals, semi-crystalline materials and so on. However, it has been gradually realized that many novel materials and devices exhibit random structures in which either basic elements or arrangement patterns may not exist. With the rapid development of modern advanced materials, this type of apparently random structure pops up frequently, leaving researchers struggling with how to describe, classify and quantitatively compare them. This paper proposes the utilization of statistical characteristics as the major indicators for the description of apparently random structures. Specifically, we have explored many statistical properties, including power spectral density, histograms, structural complexity, entropic complexity, autocorrelation, etc., and found that autocorrelation may serve as a promising statistical proxy to distinguish similar-looking random structures. We discuss eight atomic force microscope images of random structures, demonstrating that autocorrelation can be used to distinguish them. In addition, 14 more diverse datasets are used to support this conclusion, including atomic force microscopy images of polymers and non-polymers, transmission electron microscopy images of nanocomposite layers and scanning electron microscopy images of non-polymers.

Keywords: complexity, complex structure, autocorrelation

PACS: 89.75.–k, 89.75.Fb, 89.75.Kd

DOI: 10.1088/1674-1056/adbd2b **CSTR:** 32038.14.CPB.adbd2b

1. Introduction

Structure, as a basic concept in physics, exists everywhere in our daily lives and scientific research. Photos taken by a camera, movies seen on TV, electron microscopy images recorded in the laboratory and x-ray diffraction curves measured by a diffractometer are all embodiments of structures. The advance of all these subjects relies on an understanding of structure. For instance, the efficiency of image compression depends on the simplicity of the structural model adopted;[1–7] the accuracy of automatic recognition of human faces from a security camera depends on the robustness of the feature extraction method;[8–10] and the applicability of the description of crystallization behavior depends on the physical model applied.[11–17] Pioneering researchers have established a classical paradigm for the description of structure, that is, any structure can be decoupled into (1) basic structural elements and (2) the arrangement pattern adopted by the elements.[18] The former has characteristics such as shape, size, interface and so on, while the latter can be described mathematically as the summation of repeating pattern and random spatial perturbation.[19] Although this classical paradigm has helped us make tremendous advances in fields such as solid physics,[20] semiconductor physics.[21] and even soft matter physics,[22] it must be recognized that it does not satisfactorily cover many novel phenomena. For instance, the atomic arrangement in ordered materials can be mathematically described as $\Delta \boldsymbol{r}_{i,j} = \boldsymbol{r}_j - \boldsymbol{r}_i = \Delta \boldsymbol{r}_{i,j}|_{\text{ideal}} + \boldsymbol{\delta r}(\boldsymbol{r}_j - r_i)$, where $\boldsymbol{\delta r} = 0$ indicates a perfect crystal and $\boldsymbol{\delta r} \neq 0$ corresponds to disorder.[23] This type of canonical expression gradually encounters more difficulties as disorder is increased (see Fig. S1 for an illustrative example). Therefore, more practical mathematical/physical models for structures with random features are urgently needed.

Apparently random (AR) structures can be seen as the most predominant examples challenging the classical paradigm for structure description. More specifically, atomic force microscopy (AFM) images of high-precision x-ray mirrors,[24] electron microscopy images of amorphous polymers,[25,26] Markov random images generated by simulation,[27] and so on, are all AR structures in which neither apparent structural elements nor dominant arrangement patterns are observed. These examples can be further appreciated in Fig. S2. Ironically, although we can immediately realize that two AR structures are different from each other, we can hardly differentiate the two structures in a concrete manner. Taking the images in Figs. S2(e) and S2(f) as examples, we can easily tell that these two images are different but are hardly capable of describing the difference using quantitative parameters numerically. This difficulty prevents us from

[†]Corresponding author. E-mail: xjiao@ustc.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

further analyzing the structure and discovering the underlying physics. To conquer this dilemma, we devote this paper to exploring the statistical approaches that capture essential structural variability. Herein, both mathematical and physical ideas are adopted as statistical descriptors for structure, including histograms, power spectral density (PSD), structural complexity, entropic complexity, autocorrelation and so on. We hope that this work will serve as an effort to build a new paradigm toward scrutinizing AR structures in a quantitative manner.

2. Results

We use published AFM images of polymers as the model system.[28] As displayed in Fig. 1, all eight images appear quite similar. Although we can easily tell they are different from each other, it is hard for us to describe the differences in a numerical or analytical manner. The very first statistical description we attempted to use is based on a mathematical idea — a histogram of pixel intensity in an image representing the structure of interest. From this perspective, an image is treated as a realization of a spatial stochastic process defined on the R^2 domain. A histogram is the simplest probability model, expressing the frequency of appearance of the various levels of individual pixel intensity without much account for the spatial correlation between pixels. Histograms of individual pixel intensity are calculated and shown in Fig. 2. Roughly speaking, all histograms exhibit non-uniform distribution, with the heaviest weight located at low-level intensity and extending the full intensity level toward the higher level. Further analysis based on multi-peak fitting can be implemented; this finds that all histograms can be decoupled into a multitude of lognormal distributions as exemplified by the peak-fitting results found in Fig. S3, with an additional summary denoted

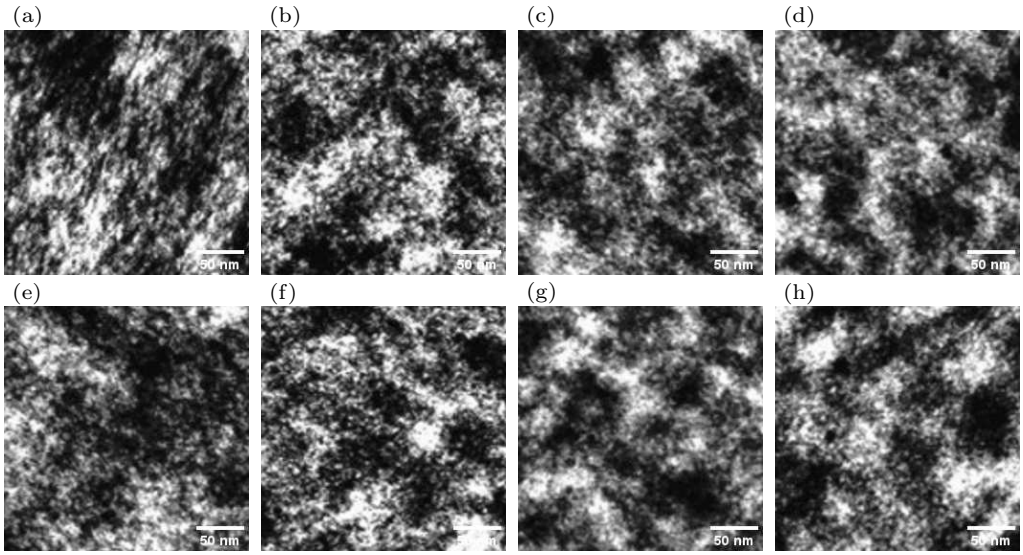


Fig. 1. AFM images of polymers (256 nm×256 nm), represented by 8-bit grayscale. Figures reproduced with permission from Ref. [28].

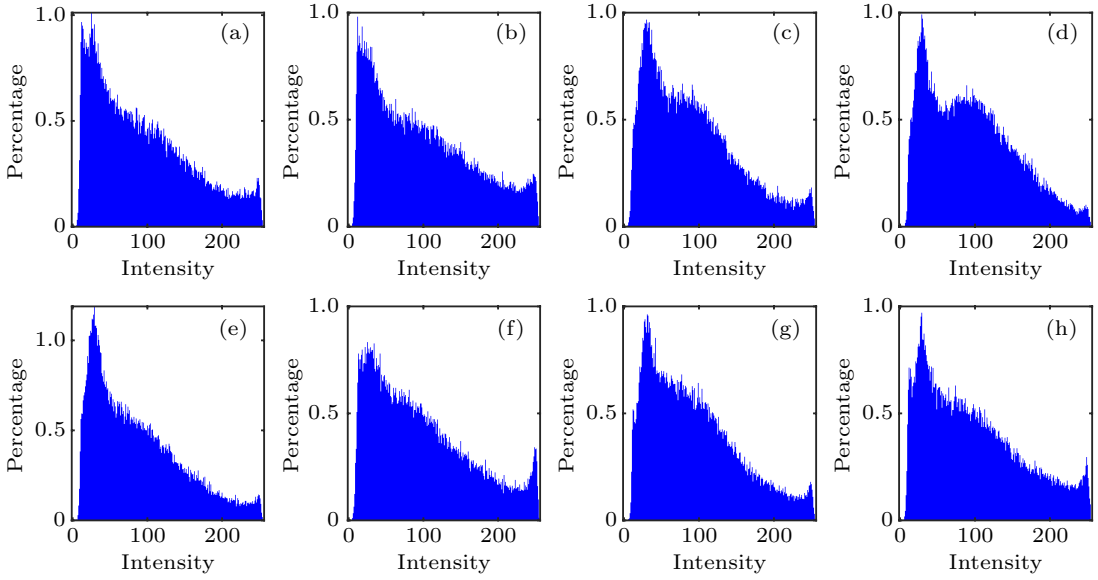


Fig. 2. Normalized histograms of the investigated sample images.

in Table S1. The table provides very detailed information regarding the distribution of histograms and allows us to differentiate all images based on the shape position, shape width of each distribution and relative ratio among them. However, we have to recognize that these achievements are only suitable for accounting for the intensity distribution of each individual pixel, without any ability to interrogate the spatial correlation between each pixel.

The above notion leads us to consider another question: although a histogram is a statistical characteristic feature for images, does it determine the appearance of an image in a sufficient and necessary relation? To answer this, we further investigate in greater detail the dependence of histograms on the appearance of an image. We take the first image in the model system as an example and re-create several copies of it by various means. Figure 3 denotes each type of conversion. It can be judged by eye that all versions of the first image appear the same with no discernible difference while their histograms vastly differ. To give this observation more generality, we also treat another image of a human portrait (structure depicting a human face) and the same trend holds. More details can be seen in Fig. S4. This suggests that the histogram itself is not sufficient as a statistical descriptor for apparently random structures.

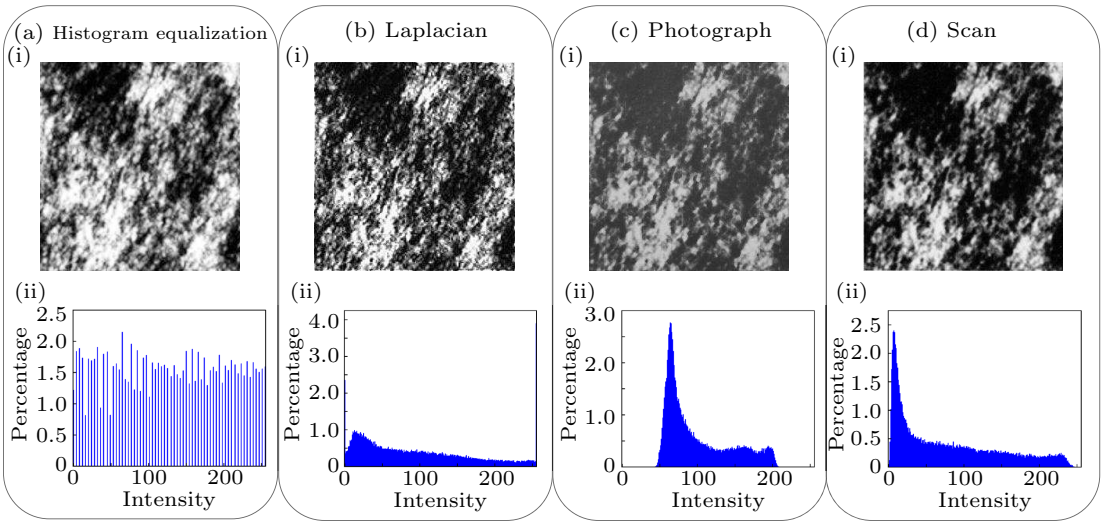


Fig. 3. Four copies of the first image in Fig. 1, including (a) histogram equalization and (b) a Laplacian, which is usually used to detect edges and textures in images and can be used as part of image enhancement. For two-dimensional images, the Laplacian is usually expressed as the following convolution kernel: [0, 1, 0; 1, −4, 1; 0, 1, 0]. (c) Photograph is obtained by printing out the image and taking pictures with a mobile phone. (d) Scanning is obtained by printing out the image and scanning it into an electronic version.

In order to overcome the drawbacks of the above histogram method, we further analyze jump histograms between adjacent pixels. Let us define an intensity jump as the difference between gray levels of two adjacent pixels along a line $s_j(i) = s(i) = f_j(i+1) - f_j(i)$, where $f_j(i)$ denotes the intensity level of pixel located at row j and column i. The variable $s(i)$ generally varies between –255 and +255. The intensity jump histograms are shown in Fig. 4. All distributions have a well-marked maximum for $s = 0$, which expresses that the zero jump is the most probable and pixel intensity changes gradually. Their shapes obey the pseudo-Voigt function with well-defined symmetry and fast decrease toward two sides of the distribution. These profiles are modeled by the pseudo-Voigt function and the standard deviations σ are parameterized in Table S2. The small value of σ corresponds to an image with a smooth transition, while the opposite holds for large values. In our case, σ varies between 1.635 and 14.55. The intensity jump histogram is basically the same for all structures, which renders this method incapable of discriminating the spatial distribution, although this method intrinsically takes the spatial correlation between adjacent pixels into account. This seemingly awkward observation is actually reasonable: AR structures are formed by near-neighbor interactions, which only generate a Markov process. It should be mentioned that a Markov process is a typical random process where the structure loses its spatial correlation immediately at a distance of more than one pixel and thereafter renders the jump histogram pseudo-Voigt-like.[27] Note that the above jump histogram only accounts for a horizontal jump, while a vertical jump histogram is delineated in Fig. S5. The above experiments confirm that the role of a histogram may be superficial and the shape of the distribution is only distantly linked to the significant elements of the AR structure.

We further explore the PSD of structures, which can be quantified by Fourier transformation. Defined as $P(w) = \int_{R^2} C(x)\, \mathrm{e}^{-\mathrm{i}wx}\, \mathrm{d}x$, where x denotes the difference between any two pixel locations, PSD is a tool capable of describing spatial relationships across the whole image globally. A PSD spectrum measures the relative contribution of different Fourier components to the image as a whole. Colloquially,

it describes the proportion of large-scale structure and its rate of propagation into fine details in an image. From Fig. 5(a), all PSD curves exhibit similar trends, with the amplitude monotonically declining with spatial frequency. On top of this, two barely seen curvatures exist at spatial frequencies of around 1 and 10 pixel^{-1}, suggesting the existence of aggregated domains in these AR structures. Also, the amplitudes of the PSDs of all samples fall at similar rates as a function of spatial frequency. To determine the precise slopes of each spectrum, we fit them with functions of the form

$$\text{amplitude of PSD} \propto \text{spatial frequency}^{-\alpha}.$$

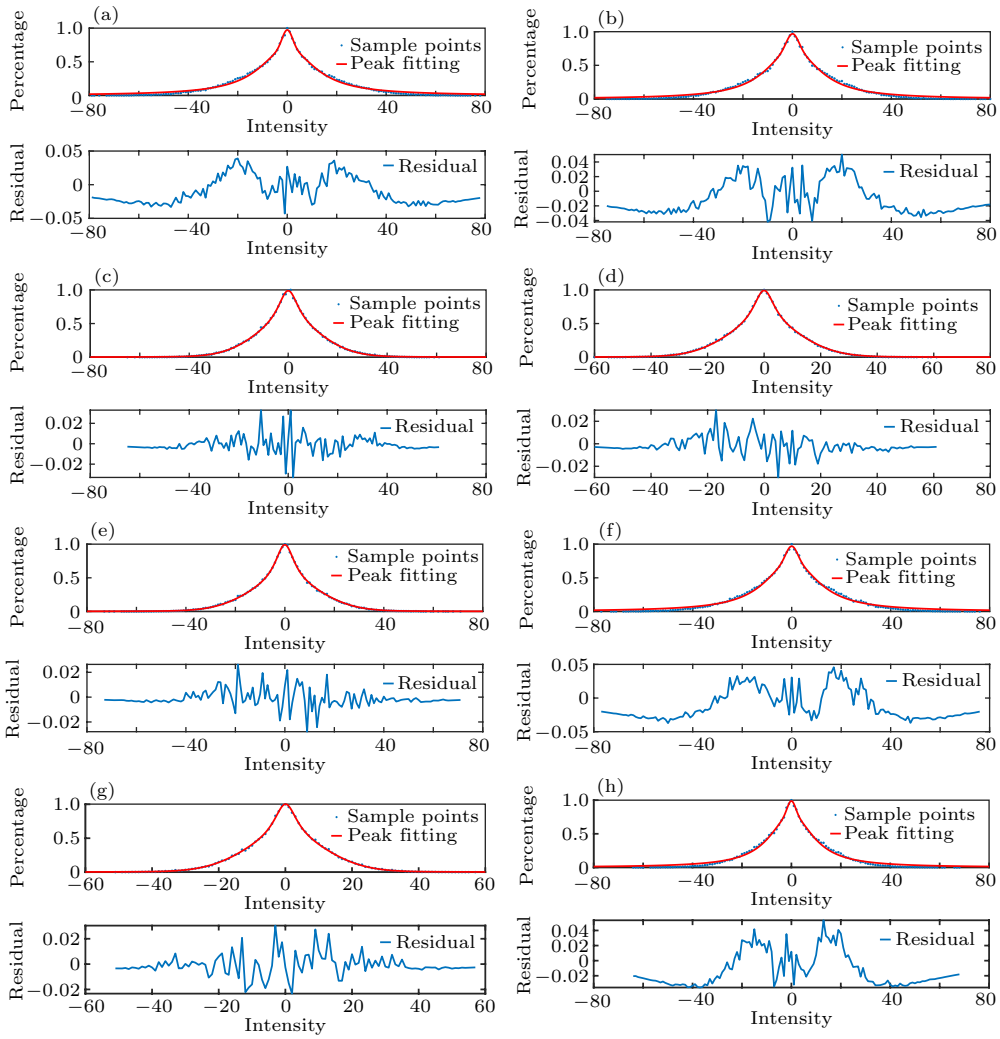


Fig. 4. Horizontal intensity jump histograms of the investigated sample images.

This should be represented as a straight line with a slope of $-\alpha$. The best-fitting regression was used and the fitted slope values are delineated in the legend of Fig. 5(a). The values range from -2.47 to -2.94. Compared with the typical value of -2 demonstrated by fractal-like structures,[29,30] the PSD slopes exhibited here indicate that the AR structures possess a greater proportion of large-scale structure and faster rate of propagation into fine details than fractal structures. It should be mentioned that a similar study on natural scenes, including both landscapes and human portraits, also gave PSD slopes lower than -2.[8] This statistical regularity obtained in PSD curves may not serve as a sensitive identifier for distinguishing AR structures considering the non-discernible features among all structures investigated here.

Judging from the appearance of all the images shown in Fig. 1, the local features are all distributed randomly and irregularly. Historically, the meaning of "order", "regular", "interest" and "random" are intertwined without universally quantitative criteria and clear distinction between them took much effort before consistency was reached.[31–34] This makes us curious to explore entropic complexity, Kolmogorov complexity and structural complexity in the hope of differentiating these images using these parameters. Although all these parameters are used to quantify the degree of irregularity of structures, they possess distinct physical meanings. We start with the investigation of entropic complexity, quantified by Shannon entropy.[35,36] This parameter is suitable for distinguishing states between regular and disordered. The mathe-

matical definition is as follows:

$$H = -\sum_{x} p_x \log p_x, \tag{1}$$

where p_x is the probability distribution of pixel intensity x. The resultant entropic complexities of all structures in this study are tabulated in Table 1. It should be mentioned that the maximum entropic complexity occurs for a system with an equal number distribution of each pixel intensity, namely, white noise, for which the value reaches 8.[37] Table 1 tells us that all structures display an entropic complexity ranging from 7.6 to 7.8, very close to the maximum value. This observation brings up two important implications: (1) due to the entropic complexities being close to 8, all structures of interest give the impression that they are random; (2) again due to their entropic complexities all being less than 8, these structures are apparently random, but not actually random. The second point corresponds to the above observation regarding the PSD curves, namely that slight curvature exists for each AR structure, indicating the existence of aggregation.

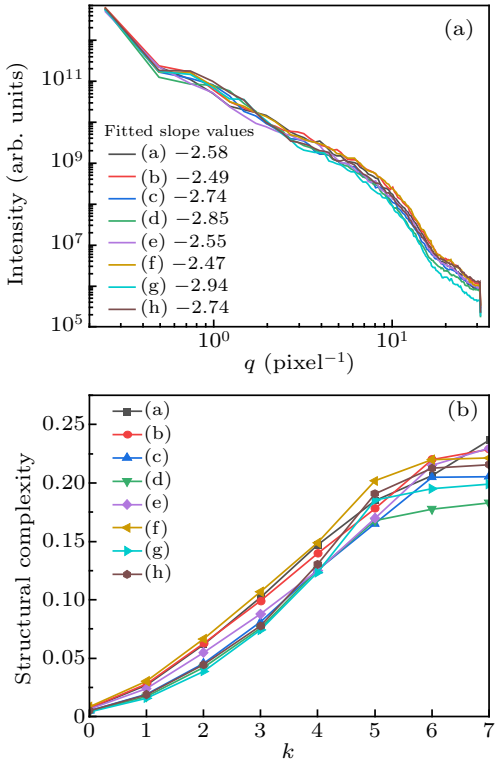


Fig. 5. (a) The power spectral density of the model system. (b) The structural complexity of the investigated sample images based on coarse-graining steps k.

The second parameter to be explored is Kolmogorov complexity, which is defined as the length of the shortest computer program that outputs the information depicting the given structure.[38] This concept is usually difficult to quantify, but the calculation becomes practical in the case of images. Herein, we utilize the compression ratio, defined as the file size of the compressed image divided by the size of the original file. The compression ratio is obtained and is also tabulated in Table 1. It should be mentioned that white noise is characterized by a value of 1 in terms of Kolmogorov complexity due to the fact that no spatial correlation can be established and thereafter no meaningful compression algorithm can take effect. In comparison, the values of AR structures span from 0.73 to 0.78, all of which are less than 1. This clearly indicates that the AR structures are not actually random since certain regularity exists and enables the compression algorithm to take effect. The non-unity results of Kolmogorov complexity, together with a value of entropic complexity of less than 8, indicate that there definitely exist perceptual structures in AR structures.

The above-discussed entropic complexity and Kolmogorov complexity only quantify the states of order, disorder and randomness. To further investigate the states of regularity and interestingness, we move forward to structural complexity. Structural complexity is a relatively novel concept compared with the other two types of complexity: it quantifies the dissimilarity of a given structure under different length scales. The development of structural complexity by researchers took several detours. Before the 1980s, structural complexity was regarded as a similar concept to entropy to quantify the degree of order/randomness. Later on, people realized that structural complexity behaves differently from entropic complexity, and should be seen as a separate concept to describe a structure's intrinsic properties.[39,40] After intensive studies, the consistency was reached that structural complexity describes the status of how far away from regularity a pattern is or its "interestingness", rather than the degree of ordering/randomness. An illustrative notion is the prototypical experiment of making coffee, where the structural complexity rises and falls during the whole process of mixing, substantially different from the behavior of monotonically increasing entropy.[41]

Structural complexity is defined as an inverse function of the degree of overlap between neighboring renormalized layers. In intuitive perception, the higher degree of overlapping mentioned above correlates with less complexity and suggests a "simple" structure. To quantify structural complexity generically, a structure is seen as a function $f(x)$ defined on a given domain D. A coarse-graining step at scale λ is applied to the original $f(x)$ and the resulting structure is denoted by $f_\lambda(x)$. The next level coarse-graining step at scale $\lambda + \mathrm{d}\lambda$ is further applied and denoted by $f_{\lambda+\mathrm{d}\lambda}(x)$. If there exists a difference between $f_\lambda(x)$ and $f_{\lambda+\mathrm{d}\lambda}(x)$, it is considered that scale λ contributes to complexity. Thus, the contribution from level λ is written as

$$C_\lambda = \left[\left\langle f_\lambda(x) | f_{\lambda+\mathrm{d}\lambda}(x) \right\rangle - \frac{1}{2} \left(\left\langle f_\lambda(x) | f_\lambda(x) \right\rangle + \left\langle f_{\lambda+\mathrm{d}\lambda}(x) | f_{\lambda+\mathrm{d}\lambda}(x) \right\rangle \right) \right], \tag{2}$$

where $\langle f(x)|g(x)\rangle = \int_D \mathrm{d}x f(x)g(x)$ is a non-normalized overlap of two patterns.

The total complexity can be calculated by summing up C_λ over all scales as

$$C = \Sigma_\lambda C_\lambda . \tag{3}$$

Figure 5(b) plots the partial structural complexity of all AR structures as a function of the coarse-graining step. It increases monotonically in a similar manner for all samples, suggesting that structural complexity may not serve as a sensitive prober for differentiating AR structures. The values of the total structural complexity are also summarized in Table 1.

Table 1. The entropic, Kolmogorov and structural complexities of the investigated sample images.

Complexity	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Entropic	7.73	7.79	7.68	7.69	7.62	7.77	7.68	7.77
Kolmogorov	0.7763	0.7746	0.7491	0.7433	0.7489	0.7772	0.7360	0.7531
Structural	0.2368	0.2288	0.2054	0.1832	0.2298	0.2213	0.1991	0.2159

We further resort to the autocorrelation function as another statistical characteristic to check its sensitivity as a meaningful descriptor of AR structures. Historically, image correlation analysis has been a powerful technique for detecting the arrangement pattern for weakly ordered materials, such as fluorophores in bioscience,[42,43] and we aim to explore its generality toward AR structures in a wider range of applications.

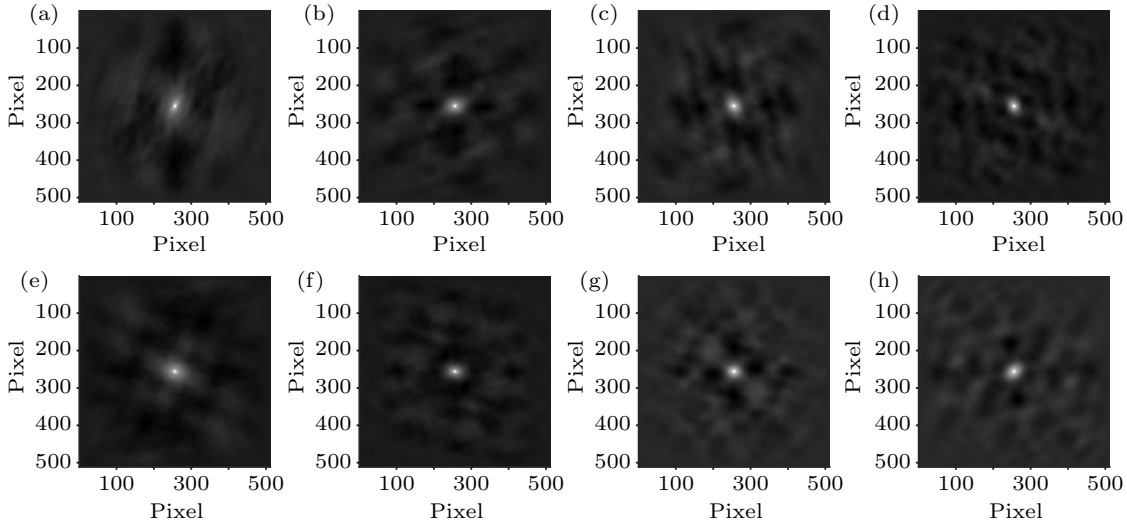


Fig. 6. Two-dimensional autocorrelation functions of the investigated sample images.

From Fig. 6, the two-dimensional autocorrelation functions are distinct from each other with easy-to-describe features. For instance, sample (a) displays four-fold symmetry azimuthally, while sample (b) shows eight-fold symmetry; sample (a) displays a single maximum radially at approximately 240 pixels, while sample (b) shows multiple maxima oscillating radially. Unlike the original structures, which are difficult to describe analytically, all these autocorrelation characteristics are easy to depict quantitatively. This implies that it may encode detailed and easy-to-describe information about structures. It should be noted that although autocorrelation is mathematically related to PSD by Fourier transformation, its sensitivity and applicability for the detection of statistical characteristics of random structures may be quite different. An autocorrelation function is logically equivalent to comparing all possible pixel pairs and reporting the likelihood of a spatial correlation as a function of the distance and direction of separation. This is a statistical method accounting for the given structure globally in real space, rather than reciprocal space like PSD. To make the comparison of autocorrelation more convenient and quantitative, we rebuilt the original autocorrelation functions into a two-dimensional pattern, where the x-axis depicts the angular distribution and the y-axis depicts the radial distribution. As shown in Fig. 7, comparison and analysis become even easier with the aid of the rebuilt patterns. For instance, the autocorrelation at a radius between 50 and 100 pixels behaves differently for each sample, where samples (a), (b) and (e) exhibit four periods of oscillation, and sample (f) barely shows any features around this range. Again, the autocorrelation behavior differs at the radius range from 150 to 200 pixels, where sample (f) exhibits the strongest fluctuation and sample (e) the least fluctuation. Furthermore, we have invented novel criteria to quantify the goodness of these statistical descriptors. The value of goodness is calculated by the

following equation:

$$G = \frac{|\overline{(I_1 - I_2)}|}{|\bar{I}_1|}, \tag{4}$$

where G is the value of goodness and I_1 and I_2 are the intensity values of various statistical descriptors for different samples, respectively. From Table S4, we have calculated the goodness values of various statistical descriptors for samples (a) and (c). In addition, we also calculated the goodness values of scanning electron microscope and transmission electron microscope images, as shown in Tables S7 and S11. From the above, it is seen that autocorrelation may be the rational candidate for a statistical proxy to distinguish random images. More diverse datasets can be found in the supporting information.

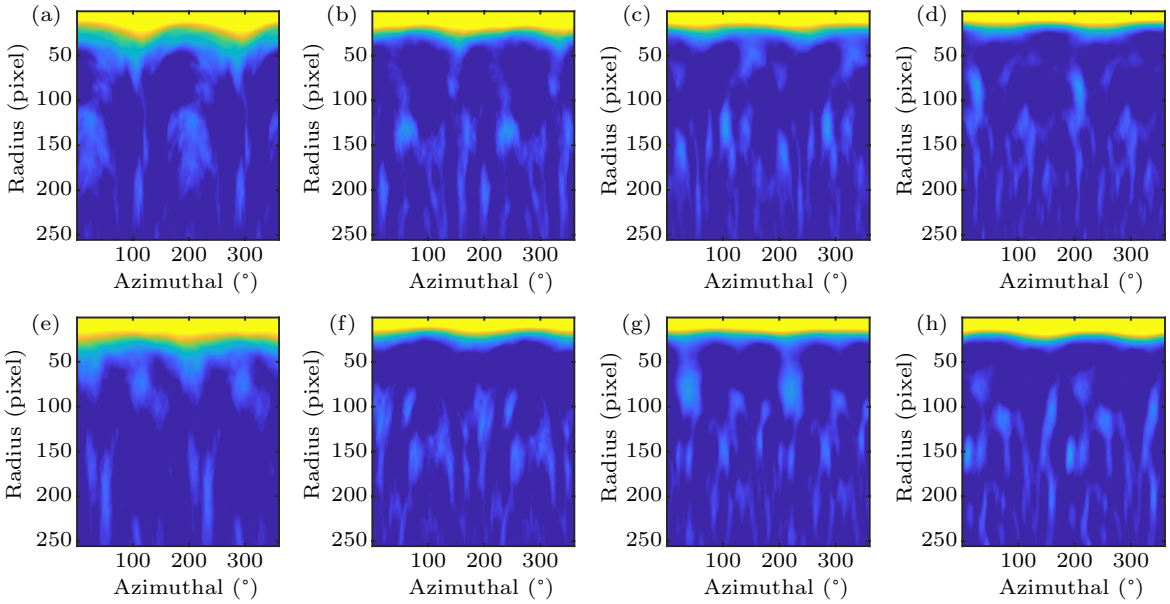


Fig. 7. Two-dimensional pattern diagrams of the autocorrelation function.

3. Discussion

Histograms are the most commonly used spatial domain characteristic for structural image analysis; they quantify the intensity distribution of image pixels. From the above observation regarding histograms of individual pixel intensity, the occurrence of intensity spans the whole scale from 0 to 255, although with a non-uniform distribution. This behavior is similar to natural scenes taken with other low-resolution equipment (such as cameras), including both landscape and human portraits. This type of continuous distribution of pixel intensity is quite different from that of certain human-made objects (e.g., a chessboard with only two discrete pixel intensities), and suggests that the origin of features for AR structures may be connected with those from natural scenes. A step further from the histogram, the intensity jump histogram overcomes the disadvantages of the former but is not capable of exploring spatial correlation. A jump histogram considers the intensity variation of adjacent pixels (depending on specific requirements, nearest-neighbor or further-neighbor are possible). The trend seen from a jump histogram is quite different from that of a histogram. Herein, all AR structures display mostly the same jump histogram, namely a pseudo-Voigt distribution. This finding is interesting in the sense that the AR structures captured by AFM possess similar jump histogram behaviors to natural scenes. This may imply certain connections between microscopic structures and macroscopic structures, although the underlying physics is unknown. Anyway, the lack of sensitivity means that jump histograms are not suitable for differentiated AR structures.

Complexity is an interesting concept and deserves more scrutiny. This work is intended to find a sensitive and quantitative descriptor, helping differentiate, compare and analyze AR structures that were hard to implement formerly. Intuitively, AR structures usually give the impression of being “complex” due to the fact that they are difficult to describe and the included features are difficult to recognize. We therefore resort to complexity for our answer in the hope that it may provide some promising inspiration. However, as its name suggests, the concept of complexity itself is “complex”. The definition of complexity varies, depending on the context in which it is applied and the researchers’ subjective choice. There is nothing wrong with complexity being context-dependent and a bit subjective. Yet, we decided to explore three types of objective complexity in the hope of obtaining a more “observer-independent” descriptor.

The three definitions of complexity — entropic complexity, Kolmogorov complexity and structural complexity — have intrinsic connections, but with clear differences. Entropic

complexity derives from Shannon's information theory and treats the pixel values as the state possessing useful information. This definition was utilized to understand the complexity of natural scenes and establish the relationship that most natural scenes with discernible features (such as landscapes and portraits) display a value less than that of white noise, which possesses the highest entropic complexity and highest potential for holding information. Although it is useful for understanding natural scenes it lacks the consideration of spatial correlation, thus making it less suitable for human-made structures. A convincing instance is a chessboard with equally distributed black-and-white blocks. Imaging two chessboards with black and white blocks of identical total areas but with different block sizes clearly shows that these two are different in terms of how "complex" they are, although the entropic complexity is identical.

Kolmogorov complexity is similar to entropic complexity, but with a comprehensive consideration of both intensity and spatial correlation. In this paper, we calculate Kolmogorov complexity based on a compression algorithm, which looks for redundant intensity modes and repeating spatial patterns and achieves compression based on the redundancy and repeatability of the original files. This enables Kolmogorov complexity not only to quantify the state of "order" or "randomness" of natural scenes but also of human-made structures. A typical example is crystals and glass. The well-arranged behavior of crystals is quite ordered and delivers a lower compression ratio. In contrast, the randomly arranged behavior of glass does not leave much opportunity for the algorithm to carry on effective compression. Very interestingly, we calculated both entropic complexity and Kolmogorov complexity of all AR structures in this study and found mostly the same trend for both. As indicated in Fig. S6, higher entropic complexity corresponds to a higher compression ratio, in other words, higher Kolmogorov complexity.

Structural complexity is a bit different from the other two. The former focus on characterizing randomness, while structural complexity focuses on irregularity and structural non-triviality. Despite the relevance between structural randomness and non-triviality, they are totally different ideas. Again, taking the coffee-making experiment as an example, after pouring milk into the coffee, the whole cup starts in a state where the milk and coffee are completely delaminated with two well-separated layers. At this moment, no randomness develops. With time, milk molecules move into the coffee and intermixing happens. After a long enough time, the milk and coffee are completely mixed. Now the randomness reaches a maximum. From the perspective of entropic complexity and Kolmogorov complexity, the complexity starts at a very low level and develops monotonically into a high-level state at the end of mixing. However, if seen from the perspective of structural complexity, the development of complexity is not monotonic. Specifically, its value starts low and also ends low, due to both cases only displaying trivial features (actually no discernible features can be seen), while it reaches a maximum in the middle of the mixing procedure. Coming back to the original aim of this paper: do any of these three types of complexity serve as a suitable proxy for AR structures? Unfortunately, none of them brings perceptive trends corresponding to the appearance of structures.

The trend observed from the PSD curves is a bit of a surprise for us. Historically, researchers have found the structural investigation of disordered polymers to be very difficult due to the fact that these materials possess neither structural elements nor arrangement patterns, namely, AR structure. To conquer this, people have resorted to the scattering method, which is the physical counterpart to the mathematical tool of Fourier transform.[22] It has to be pointed out that utilization of the scattering method has been tremendously successful in terms of exploring polymer structures. This big success leads us to believe that PSD analysis could be a sensitive proxy to differentiate AR structures. However, the difference in PSD curves in this work is neither strong enough nor clear enough to serve this purpose. To minimize the ambiguity, we reaffirm that we by no means imply that PSD analysis cannot serve as a meaningful statistical method for AR structure investigation. The PSD curves of all eight samples in this paper are of course different from each other and can deliver important distinctions using in-depth analysis, such as Porod analysis, Guinier analysis and other higher-moment analyses. However, we must note that the appearance of PSD curves is indeed similar and does not illustrate distinct differences, as can be easily noted by an untrained person. This means that this method is not a proper proxy to differentiate AR structures due to its lack of sensitivity and processing difficulties.

Regarding autocorrelation functions, several observations deserve further scrutiny. The first is that the applicability of PSD and autocorrelation functions is quite different for apparently random structures. It is well-known that Fourier transformation and autocorrelation calculations can be inter-translated and connected by the following equation:

$$g_{ii}(a,b) = \frac{F^{-1}\{F\left[i\left(x,y\right)\right]^2\}}{NMi^2} - 1, \tag{5}$$

where $F(x,y)$ is the Fourier transform of structure image $i(x,y)$, N and M are the dimensions of image $i(x,y)$ and $g_{ii}(a,b)$ is the autocorrelation function. Even with the above-mentioned mathematical connection their sensitivity to differentiating structural distribution is quite different, suggesting that they cannot replace each other. This is actually consistent

with the historical development of structural analysis: scattering methods have been utilized for structural analysis for nearly 100 years, while the first work utilizing autocorrelation analysis only dates back to 2004.[44] As indicated by the above sections, the autocorrelation function is capable of providing guidelines regarding structural arrangement in terms of both radial and azimuthal directions. With the rebuilt version of the autocorrelation function, the analysis can be easily implemented in a quantitative manner. For an even deeper understanding of the spatial arrangement of AR structures, we can further reduce the two-dimensional rebuilt autocorrelation functions into one-dimensional curves, either along a radial or azimuthal direction, at any given spatial coordinate. Figures S7 and S8 delineate the detailed one-dimensional curves. From Fig. S7, the autocorrelation variation at a radius of 60 pixels as a function of angular change displays distinguishable trends between different samples. For instance, samples (a) and (b) can be easily identified by the observation that (a) exhibits two-cycle fluctuation while (b) exhibits four-cycle fluctuation. Figure S8 shows another autocorrelation variation at a radius of 190 pixels. More autocorrelation variation curves at different radii facilitate the applicability of this method. Take the comparison between samples (b) and (e) as an example. The autocorrelation variations at a single radius of 60 pixels do not deliver enough distinguishable features. An extra comparison at another radius of 190 pixels clearly differentiates these two samples.

4. Conclusion

In this work we note that structural physics is a unique and important subfield of modern physics. Although it does not directly investigate the mechanism of how functionality occurs, it helps deepen our understanding of what structure is and how its description can be improved, eventually promoting the optimization of functionality by improved structure–functionality relationships. We start with the historical consideration of structural physics, where the dominant paradigm is based on "structural elements + their arrangement patterns". Despite its tremendous success, we humbly believe that this paradigm may not be general enough to cover all phenomena. Herein, the typical counter-example is an AR structure that lacks either a dominant structural element or a clear arrangement pattern. It is worth noting that more AR structures pop up in modern society including, but not limited to, disordered polymers, biomaterials and ultra-precision surfaces. We propose that on top of the canonical paradigm, novel methodology based on statistics should be utilized to better describe AR structures. With our limited effort, some of the commonly known statistical characteristics have been explored and we point out that the autocorrelation function may be a promising proxy for descriptors of AR structures.

Supplementary material

See the supplementary material for: (1) summary of structures with varying degrees of ordering, (2) three types of apparently random structures, (3) the fitting results of the normalized histogram of the investigated sample images, (4) four copies of a human portrait, (5) the fitting results of the vertical intensity jump histogram of the investigated sample images, (6) the entropic, Kolmogorov and structural complexities of the investigated sample images, (7) one-dimensional curves of the autocorrelation function at a radius of 60 pixels, (8) one-dimensional curves of the autocorrelation function at a radius of 190 pixels. And there are more diverse datasets. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Data availability statement

The data that support the findings of this study are openly available in Science Data Bank at the following link https://doi.org/10.57760/sciencedb.j00113.00210

Acknowledgements

This work is supported by the School Important Direction Project Cultivation Fund and Key Fund Project for Youth Innovation (Grant Nos. WK2310000101, YD2310002006, and BJ2310000055) and Anhui Provincial Natural Science Foundation (Grant No. 2308085MA19).

References

[1] Gidas B 1989 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** 164
[2] Lindeberg T 1998 *International Journal of Computer Vision* **30** 79
[3] Buccigrossi R W and Simoncelli E P 1999 *IEEE Transactions on Image Processing* **8** 1688
[4] Jinggang H and Mumford D 1999 *IEEE Xplore* **1** 541
[5] Srivastava A, Lee A B, Simoncelli E P and Zhu S C 2003 *Journal of Mathematical Imaging and Vision* **18** 17
[6] Munn B and Gong P L 2018 *Phys. Rev. Lett.* **121** 058101
[7] Sinno Z, Caramanis C and Bovik A C 2018 *IEEE Transactions on Image Processing.* **27** 3194
[8] Graham D J and Field D J 2008 *Perception* **37** 1341
[9] Koch M, Denzler J and Redies C 2010 *Plos One* **5** 8
[10] Foreman S, Giedt J, Meurice Y and Unmuth-Yockey J 2018 *Phys. Rev. E* **98** 052129
[11] Crist B and Cohen J B 1979 *Journal of Polymer Science Part B-Polymer Physics* **17** 1001
[12] Wilke W 1983 *Colloid and Polymer Science* **261** 656
[13] Hindeleh A M and Hosemann R 1988 *J. Phys. C-Solid State Phys.* **21** 4155
[14] Somashekar R and Somashekarappa H 1997 *Journal of Applied Crystallography* **30** 147
[15] Haji-Akbari A, Engel M, Keys A S, Zheng X Y, Petschek R G, Palffy-Muhoray P and Glotzer S C 2009 *Nature* **462** 773
[16] Keen D A and Goodwin A L 2015 *Nature* **521** 303

[17] Carrasquilla J and Melko R G 2017 *Nat. Phys.* **13** 431
[18] Oganov A R, Pickard C J, Zhu Q and Needs R J 2019 *Nature Reviews Materials* **4** 331
[19] LeSar R 2014 *Annual Review of Condensed Matter Physics* **5** 375
[20] Sólyom J 2007 *Fundamentals of the Physics of Solids* (Heidelberg: Springer Berlin) 13
[21] Böer K W and Pohl U W 2020 *Semiconductor Physics* (Cham: Springer) 59
[22] Jiao X C, Ye L and Ade H 2017 *Advanced Energy Materials* **7** 1700084
[23] Rivnay J, Noriega R, Kline R J, Salleo A and Toney M F 2011 *Phys. Rev. B* **84** 045203
[24] Keil P, Lützenkirchen-Hecht D, Novikov D V, Hahn U and Frahm R 2001 *Nuclear Instruments and Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment* **467** 275
[25] Liu F, Gu Y, Shen X B, Ferdous S, Wang H W and Russell T P 2013 *Progress in Polymer Science* **38** 1990
[26] Fradler C, Rath T, Dunst S, Letofsky-Papst I, Saf R, Kunert B, Hofer F, Resel R and Trimmel G 2014 *Solar Energy Materials and Solar Cells* **124** 117
[27] Hernández-Lemus E 2021 *Frontiers in Physics* **9** 641859
[28] Lee J Y, Song K W, Ku J R, Sung T H and Moon D K 2011 *Solar Energy Materials and Solar Cells* **95** 3377
[29] Bies A J, Boydston C R, Taylor R P and Sereno M E 2016 *Journal of Environmental Psychology* **47** 155
[30] Knill D C, Field D and Kersten D 1990 *J. Opt. Soc. Am. A* **7** 1113
[31] Adami C 2002 *Bioessays* **24** 1085
[32] Stephens G J, Mora T, Tkacik G and Bialek W 2013 *Phys. Rev. Lett.* **110** 018701
[33] Bolzani V D 2013 *Planta Medica* **79** 10
[34] Crutchfield J P and Young K 1989 *Phys. Rev. Lett.* **63** 105
[35] LopezRuiz R, Mancini H L and Calbet X 1995 *Phys. Lett. A* **209** 321
[36] Cha S H 2007 *International Journal of Mathematical Models and Methods in Applied Sciences* **1** 300–307
[37] Zanette D H 2018 *Plos One* **13** 11
[38] Wolpert D H and Macready W 2007 *Complexity* **12** 77
[39] Wolpert D H and Macready W G 1997 *IEEE Transactions on Evolutionary Computation* **1** 67
[40] Bagrov A A, Iakovlev I A, Iliasov A A, Katsnelson M I and Mazurenko V V 2020 *Proc. Natl. Acad. Sci. USA* **117** 30241
[41] Aaronson S, Carroll S and Ouellette L 2014 *arXiv: Statistical Mechanics* CALT-68-2927
[42] Kolin D L and Wiseman P W 2007 *Cell Biochemistry and Biophysics* **49** 141
[43] Robertson C and George S C 2012 *Journal of Biomedical Optics* **17** 080801
[44] Rubin D M 2004 *Journal of Sedimentary Research* **74** 160