J Membrane Biol (2011) 241:59–68
DOI 10.1007/s00232-011-9362-x

Computing Spatial Information from Fourier Coefficient Distributions

William F. Heinz · Jeffrey L. Werbin · Eaton Lattman · Jan H. Hoh

Received: 3 March 2011 / Accepted: 14 March 2011 / Published online: 5 May 2011

Abstract The spatial relationships between molecules can be quantified in terms of information. In the case of membranes, the spatial organization of molecules in a bilayer is closely related to biophysically and biologically important properties. Here, we present an approach to computing spatial information based on Fourier coefficient distributions. The Fourier transform (FT) of an image contains a complete description of the image, and the values of the FT coefficients are uniquely associated with that image. For an image where the distribution of pixels is uncorrelated, the FT coefficients are normally distributed and uncorrelated. Further, the probability distribution for the FT coefficients of such an image can readily be obtained by Parseval's theorem. We take advantage of these properties to compute the spatial information in an image by determining the probability of each coefficient (both real and imaginary parts) in the FT, then using the Shannon formalism to calculate information. By using the probability distribution obtained from Parseval's theorem, an effective distance from the uncorrelated or most uncertain case is obtained. The resulting quantity is an information computed in *k*-space (kSI). This approach provides a robust, facile and highly flexible framework for quantifying spatial information in images and other types of data (of arbitrary dimensions). The kSI metric is tested on a 2D Ising model, frequently used as a model for lipid bilayer; and the temperature-dependent phase transition is accurately determined from the spatial information in configurations of the system.

Keywords Ising model · Information theory · Membrane · Spatial organization

William F. Heinz and Jeffrey L. Werbin contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00232-011-9362-x) contains supplementary material, which is available to authorized users.

W. F. Heinz · J. L. Werbin · J. H. Hoh (✉)
Department of Physiology, Johns Hopkins School of Medicine, 725 N. Wolfe Streets, Baltimore, MD 21205, USA
e-mail: jhoh@jhmi.edu

E. Lattman
Department of Biophysics, Johns Hopkins University, Baltimore, MD, USA

Present Address:
E. Lattman
Hauptman-Woodward Medical Research Institute, Buffalo, NY, USA

Introduction

The amount of information in a system that varies in time or space, for example, is an interesting and useful quantity. Information theoretical approaches are employed to study questions in fields ranging from linguistics (Shannon 1951) to chemistry (Hummer et al. 1996) and cosmology (Bekenstein 1973). In biology and related sciences information theory has been of long-standing interest (Quastler 1953). Indeed, information would appear to be fundamental to widely studied problems such as cell–cell communication and signal transduction, as well as the genetic encoding and expression of proteins. However, while information

theory has been applied to these (Lenaerts et al. 2009; de Polavieja 2004; Gatlin 1966; Tkacik et al. 2008) and other problems including directional cell movement (Vergassola et al. 2007), protein structure (Fink and Ball 2001) and neuronal signal processing (Kimmel et al. 2007), to date it has had a modest impact in biology and biochemistry.

Our interest in this issue originates in a desire to understand spatial information in cellular microenvironments. For example, the extracellular matrix is a complex fibrous material composed of many different proteins that provides important signals that activate specific cellular functions (Geiger et al. 2001). We are attempting to understand how cells process spatial information in natural microenvironments, in part by using in vitro cell culture studies in which information is experimentally controlled (Werbin et al. 2007). We are also interested in the spatial relationships between molecules in membranes and how these might be connected to function. However, the lack of a suitable information metric for those studies prompted us to develop an information measure based on correlations between structures or events in a system. This new metric is general and useful for any type of data, and we present the underlying theory as well as several tests of the metric. Applications to biological problems will appear at a later time. Images are used as a vehicle for presentation, but the metric can be applied to any type of data for which a Fourier transform (FT) can be obtained.

We use "spatial information" in the conventional Shannon sense of observer-independent information. Here information is related to the likelihood of an event, which in turn is related to the reduction of uncertainty (Shannon 1948; Applebaum 1996). An uncertain outcome is one that is drawn from a large number of equivalent events from some ensemble. Thus, it is a relatively likely outcome with little information. In contrast, an outcome that is drawn from a smaller number of equivalent events is rare and, thus, reflects more information. A number of approaches to computing image information have been developed, many based on the well-known formalism established by Shannon in which

$$I_e = -k \log P_e \tag{1}$$

and

$$H = -k \sum_e P_e \log P_e \tag{2}$$

where P_e is the probability of the event e, I_e is the information associated with e and H is the informational entropy associated with all possible events. When $k = 1$ and the log is base 2, and the units are bits. Extending this to a collection of i elements, such as an image, I and H are both summed for all elements so

$$I_{\text{total}} = -\sum_i \log_2 P_{i,e} \tag{3}$$

and

$$H_{\text{total}} = -\sum_i \sum_e P_{i,e} \log_2 P_{i,e} \tag{4}$$

In the most direct implementations of information theory, P_e is computed from the statistical properties of the elements or events in the system. For example, in image analysis a common approach is to compute an "image entropy" (H_{RS}, subscript RS is for real space) by taking the value of each pixel as an event and using the normalized image histogram to determine P_e (Russ 2002). However, such an approach does not capture the spatial organization—that is, the spatial correlations between pixels—within the image. The histogram is essentially a measure of information based on composition, and all images with the same histogram and the same number of pixels will have the same image entropy, as well as the same information (I_{RS}), no matter the arrangement of the pixels.

A number of solutions to capture spatial components of the information in data have been developed. Shannon's original solution to dealing with contributions from correlations between elements in the data was to group elements into n-grams (Shannon 1948). The probabilities for the n-grams were then used to compute the information. While this can work well for small group sizes, it becomes increasingly difficult to employ as the n-gram size grows. A related approach was used for images to improve the histogram-based entropy measure of information; pairwise pixels or voxels were used to compute a quantity that incorporates a spatial component (Rueckert et al 2000). This approach was subsequently extended to additional dimensions (Russakoff et al. 2004). However, while the approach was shown to improve robustness of image registration, it is computationally intensive because the number of dimensions in the problem grows linearly with the radius of the neighborhood considered and the computational requirements scale with the number of pixels in the image raised to the number of dimensions. Thus, only neighborhoods of a few pixels in diameter were used. A different approach based on the distribution of curvatures in two- or three-dimensional (2D or 3D) objects to reflect shape information has also been proposed (Page et al. 2003). With this metric, simple objects with constant curvatures, such as planes and spheres, have a shape information of 0, while objects with changes in curvature have more shape information.

A measure of spatial information also emerges from considering the algorithmic or Kolmogorov complexity,

where the information is directly related to the size of the smallest set of instructions that completely describe that object (Li and Vitányi 2008). While it is not possible to compute the smallest set of instructions for an arbitrary image, the use of compression algorithms has become a practical implementation based on the same underlying idea (Bardera et al. 2006). The more compressible an image is, the less information it has.

Here, we propose a new approach to quantifying spatial information that is based on an analysis of the Fourier coefficients computed from an image. By using Parseval's theorem to determine the expected variances for the distributions of coefficients, a Shannon-like quantity can be readily computed from the FT of any image. Because this information metric is computed in *k*-space, we refer to it as *k*-space information (kSI). The term "*k*-space information" has also been used in the analysis of entropy associated with a particular pixel across a set of related images (Bilbao et al. 2004), an approach quite different from the one presented here. We validate kSI in part by comparing it to the Kolmogorov complexity as approximated by JPEG2000 (JP2K) compression (lossless implementation).

Theory and Computation

We begin with an image $f(x,y) = \{f_{xy}\}$, where x and y are integers that index a pixel value, f_{xy}. The discrete FT of $f(x,y)$ is $F(m,n)$, where $F(m,n) = \{F_{mn}\}$, and F_{mn} are complex numbers $F_{mn} = a_{mn} + ib_{mn}$. We note that the FT is an equivalent representation of the information in an image, and the inverse transform returns the original image. Further, any set of Fourier coefficients is uniquely associated with an image. In our approach, the probability distribution $P(e) = \{p_e\}$, where $\Sigma p_e = 1$, is the distribution of values for the Fourier coefficients, and the event, e, is a coefficient, F, having a particular value, $a + ib$. For a given coefficient, F_{mn}, $P_{mn}(a,b) = \{p_{ab}\}_{mn}$ is the probability distribution of the real and imaginary parts of the coefficients F_{mn} and $\Sigma\{p_{ab}\}_{mn} = 1$. For the case where the pixels in the original image uncorrelated (i.e. an uncorrelated image), we assume all coefficients are independent. Application of the central limit theorem determines $P_{mn}(a,b)$ and establishes that $P_{mn}(a,b) = P(a,b)$, $\{p_{ab}\}_{mn} = \{p_{ab}\}$ and that $\{p_{ab}\}$ obeys a normal distribution. Thus, the entropy of a coefficient can be written as $H_{mn} = H = -\Sigma p_{ab}\log p_{ab}$ and, therefore, $H_{\text{total}} = N \cdot H$ (where $N =$ the number of pixels in the image). Note that we employ the information theory convention that $p_{ab}\log p_{ab} = 0$ if $p_{ab} = 0$. The real and imaginary parts of the FT are also independent and, consequently, $P(a,b) = P(a)P(b)$ and $H = -\Sigma p_a p_b \log p_a p_b$, where $\Sigma p_a = 1$ and $\Sigma p_b = 1$. The distribution of Fourier coefficients of independent random variables obeys a normal distribution with a mean of 0, so the probability density function (pdf) is

$$p_a p_b = \frac{1}{2\pi\sigma_a\sigma_b} e^{-\left(\frac{a^2}{2\sigma_a^2}+\frac{b^2}{2\sigma_b^2}\right)} \tag{5}$$

$$= \frac{1}{2\pi\sigma_a^2} e^{-\left(\frac{a^2+b^2}{2\sigma_a^2}\right)} \tag{6}$$

and

$$\sigma_a^2 = \langle a_{mn}^2 \rangle = \frac{1}{2}\langle F_{mn}^2 \rangle \tag{7}$$

where σ_a^2 and σ_b^2 are the variances for the distributions of the real and imaginary parts of the Fourier coefficients (Drenth 1999). These variances can be calculated readily from the original real space image using Parseval's theorem, which relates the image to its FT by

$$\sum_{x,y} f_{xy}^2 = C\sum_{m,n} F_{mn}^2 = CN\langle F^2 \rangle = 2CN\sigma_a^2 \tag{8}$$

where C is a normalization factor that is dependent on the specific implementation of the FT, mostly commonly either 1 or N. In our programs C is equal to N. The underlying assumptions here were confirmed numerically (see supplemental information).

To determine the Fourier coefficient–based information, I_{kS}, or Fourier coefficient–based entropy (H_{kS}) of a particular image, we begin by computing the FT of the image and recording the real and imaginary parts of the coefficients (Fig. 1). We then use Parseval's theorem to compute the variances for the distributions of the real and imaginary parts of the coefficients that would be expected from all possible images with the same histogram as the image of interest. These variances are then used to compute the probabilities, P_a and P_b, for each of the coefficients in the FT of the image by integrating the pdf over a window from $c - 0.005\sigma$ to $c + 0.005\sigma$ (where c is either a or b) (see supplemental information). The probabilities are then used to compute I_{kS} by summing $-\log P$ over all coefficients. H_{kS} is computed by integrating the pdf in steps of 0.01σ.

Here, it is important to appreciate that the Fourier coefficients for many or even most images of interest are not independent and not normally distributed. Thus, what is being computed is the probability of a particular set of coefficients associated with the image of interest occurring in the FT of all possible images. Note that the H_{kS} can be computed from the image histogram alone, the computation does not require the coefficients from the FT and the H_{ks} of every coefficent is the same. Thus, any two images

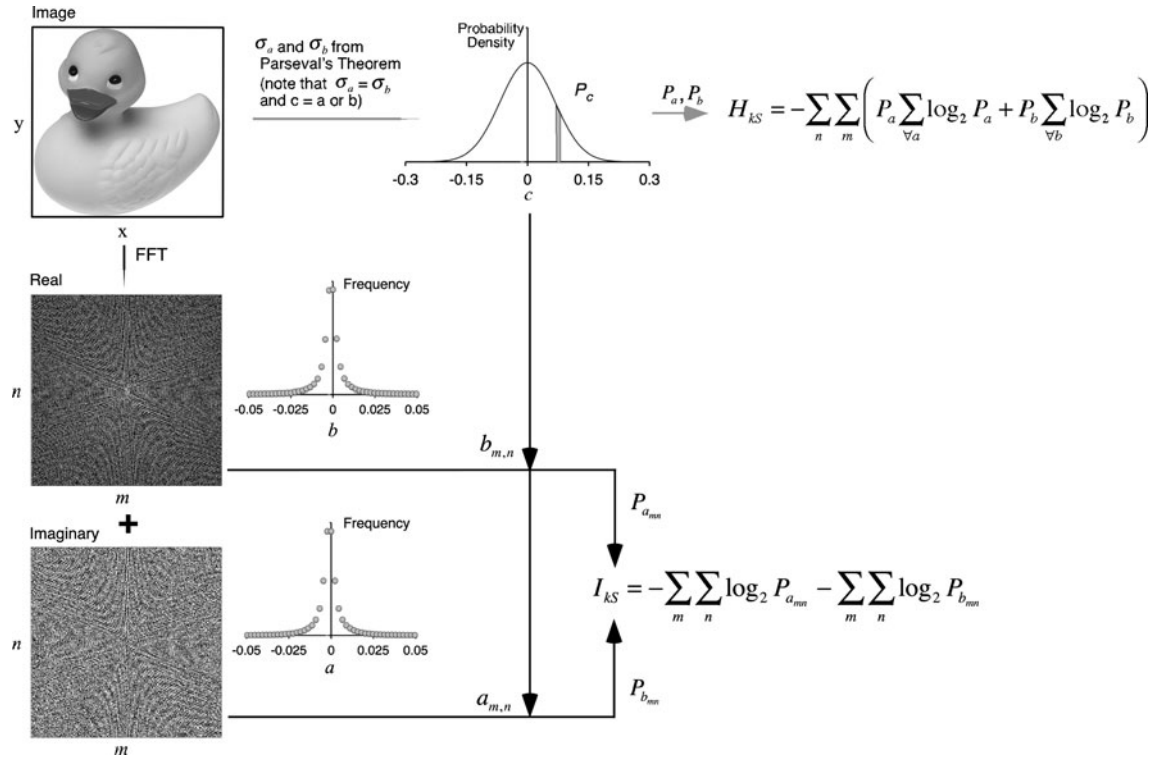


Fig. 1 Schematic representation of the approach for computing entropy and information in *k*-space. To compute the entropy, Parseval's theorem is first used to obtain the variance for the Fourier coefficients based on the values of the pixels in the image. The Fourier coefficients are normally distributed, and thus the variance provides a probability density function. This probability density function is then divided into bins of $\sigma/100$, and the entropy is computed by summing $P_c \log_2 P_c$ for a and b from -10σ to $+10\sigma$. Note that the $\forall c$ limit is used to indicate that the summation is being taken over all the bin values of a or b. To compute the spatial information, the real and imaginary parts of the Fourier coefficients for the image are determined. A probability for each coefficient is obtained from the probability density function above, and the information is computed by summing $\log_2 P$ for both parts

that share the same image histogram have the same H_{kS}. But I_{kS} depends on the specific set of coefficients in an image, making it suitable for quantifying the spatial information in the image. In addition to the I_{kS}, it is useful to think about the spatial information in a particular image relative to the greatest possible spatial information for an image with the same histogram. The H_{kS} sets an upper bound for the I_{kS}, and we define kSI as the information in an image relative to the entropy,

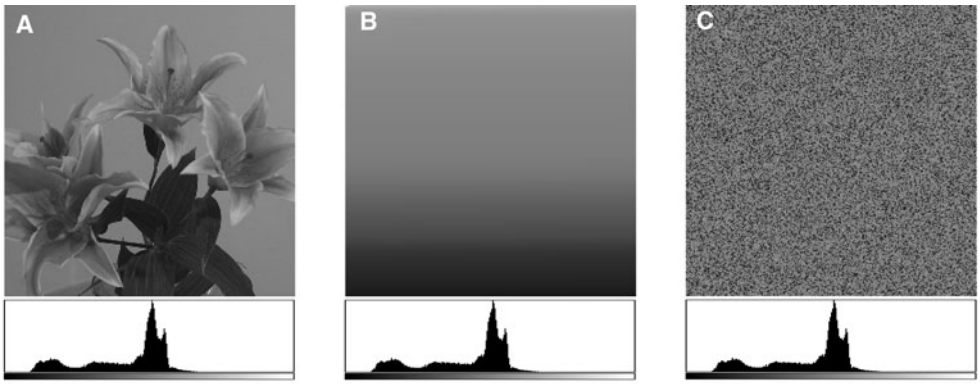
$$\begin{aligned} \mathrm{kSI} &\equiv H_{\mathrm{kS}} - I_{\mathrm{kS}} \\ &= -2N_{\mathrm{pixels}} \left(\sum_{\forall c} P_c \log_2 P_c \right) + \sum_{\forall c} \sum_{m,n} \log_2 P_{c_{mn}} \end{aligned} \quad (9)$$

Information is defined as a decrease or reduction of uncertainty (Shannon 1948). We take the H_{ks} to define the most common or uncertain distribution of Fourier coefficients. Thus, the kSI provides a measure of information by virtue of being related to how far from uncorrelated a particular collection of Fourier coefficients is.

Results

Fourier-Based Entropy and Information in 2D Gray-Scale Images

To characterize the kSI approach, we first examine 2D gray-scale images where each pixel has an eight-bit intensity value. We begin with three test images that share the same histogram but in which the pixels are organized differently: an image of a flower, an image where the pixels are organized from the lowest value, to the highest value and an image where the pixels is shuffled randomly in the *xy* plane (Fig. 2). Because these all have the same histogram, the H_{RS} and I_{RS} are identical. Similarly, H_{kS} provides an upper bound for I_{kS}, and it is identical for all three images. However, the I_{kS} depends on the specific set of Fourier coefficients in each image and, thus, varies significantly. The ordered image has the lowest I_{kS}, the randomly shuffled image has the highest I_{kS} and the original image has an intermediate I_{kS}. Subtracting the image I_{kS} from the H_{kS} gives the kSI for each



kSI	74,462	92,026	8
I_{kS}	1,064,670	1,047,106	1,139,124
H_{kS}	1,139,132	1,139,132	1,139,132
I_{RS}	432,832	432,832	432,832
H_{RS}	432,832	432,832	432,832
JP2K	2,202,560	82,944	4,021,568

Fig. 2 Information metrics for three test images with varying spatial organization. **a** Image of a flower. **b** Image with the same histogram as the flower, with the pixels arranged from highest to lowest values. **c** Image with the same histogram as the flower, with the pixels uncorrelated. The units are bits for all quantities

image, showing that the ordered image is furthest away from uncorrelated and the flower image is between uncorrelated and 'perfectly' ordered. A qualitatively similar result is obtained using JP2K compression, although the scale is inverted such that high complexity has large values and low complexity has low values.

Information Varies with "Disorder"

One of the most basic qualifications for a measure of spatial information is that it will vary in a predictable way with the amount of order in an image. To test the behavior of the information metric developed here on disorder/order, we use a series of images where the amount of spatial information is changed in a well-defined fashion. Starting with an image in which the pixels are perfectly ordered (as in Fig. 2b, except from a uniform histogram), a varying fraction (0–1.0) of the pixels is shuffled to random positions (Fig. 3a). For this data set, the I_{kS} increases monotonically with increasing amount of shuffle or disorder (Fig. 3b). Likewise, JP2K analysis shows a monotonic increase with increasing shuffle, although the rise is steeper at small shuffles and becomes less sensitive at higher shuffles. We have chosen to define kSI as the distance from a completely uncorrelated image as defined by the H_{kS}. Thus, the more order that is introduced into the image, the greater the information and the higher the kSI value (Fig. 3c). Note that the shuffled pixels are exchanged pairwise such that the histogram is maintained. Interestingly the I_{kS} and thus the kSI have a simple quadratic dependence on the shuffle fraction.

Precision Estimate

It is not possible to produce images with a known spatial information; therefore it is not possible to independently establish the accuracy of the kSI metric. However, it is possible to estimate the precision of the metric. Because the kSI introduces an arbitrary offset to the spatial information, we perform this analysis on the I_{kS}. We start with a gray-scale image and make a large number of images that are shuffled by a specific fraction. We then examine the distribution of I_{kS} values for these shuffled images (Fig. 4). These values are approximately normally distributed. The standard deviation of the distribution also depends on the amount of order in the image, and it is convenient to represent it as a coefficient of variation. Thus, for a 0.5 shuffle fraction the standard deviation is 0.045%, while for an 0.85 shuffle fraction it is 0.014%. The variance of the coefficient distribution provides an upper bound to the precision of the measurement; the precision cannot be any worse than that reflected by the variation that arises from repeated shuffles. Instead, some of the variance from the repeated shuffling will likely produce images that vary significantly in kSI, thus broadening the distribution beyond that defined by the precision. The JP2K metric is qualitatively similar to the kSI with a similar precision (0.054% standard deviation for a 0.5 shuffle, 0.040% for a 0.85 shuffle).

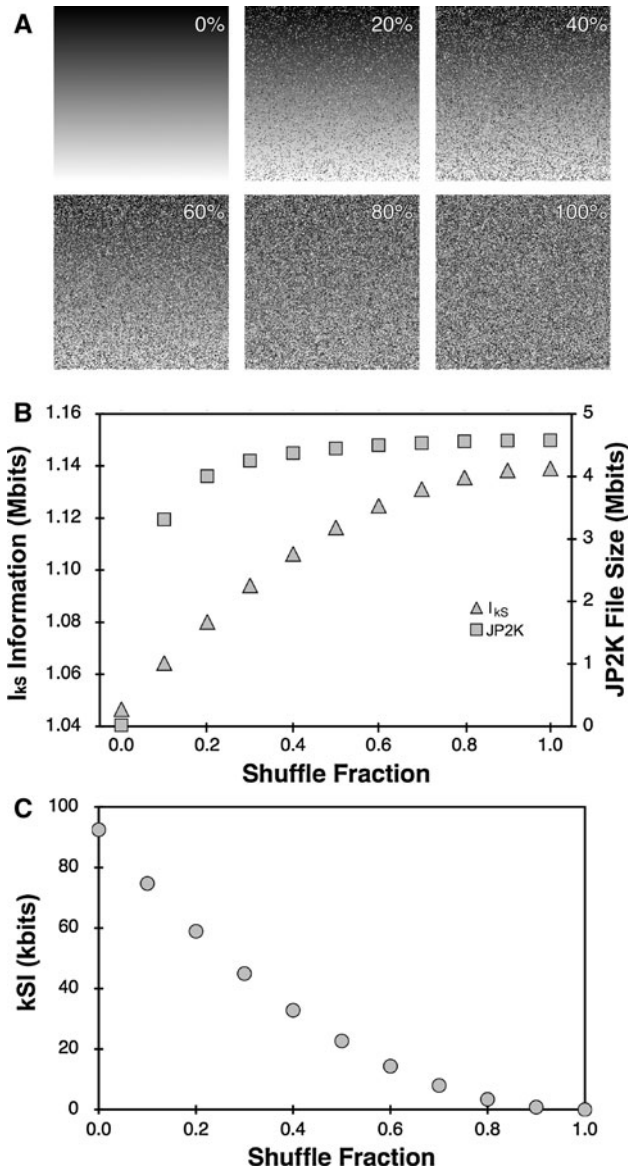


Fig. 3 Dependence of spatial information metrics on image order/disorder. **a** Starting with a 256 × 256 pixel eight-bit image composed of pixels arranged from highest to lowest (with a uniform histogram), the amount of order was varied by shuffling increasing fractions of pixels. The ordered image is selected because it has one of the most ordered arrangements possible. Although it may not have the absolute lowest I_{kS}, it represents a low-information bound. **b** For this shuffle series the I_{kS} increases monotonically from the most ordered to the most disordered image ($I_{kS} = -942.1318254x^2 + 20550.8962979x + 1027065.0735181$ and $R^2 > 0.9999999$, where x is the shuffle fraction). The JP2K compression also increases monotonically with increasing disorder. **c** The kSI decreases monotonically with increasing disorder in an image (kSI = $94213.1828343x^2 - 186666.3268458x + 92458.6869676$ and $R^2 > 0.9999999$)

One other effect on the precision arises from the fact that there is a small difference in the kSI that is computed for images with different rotations. When a rectangular image is rotated an integer multiple of 90 degrees, the spatial information is clearly conserved and one would

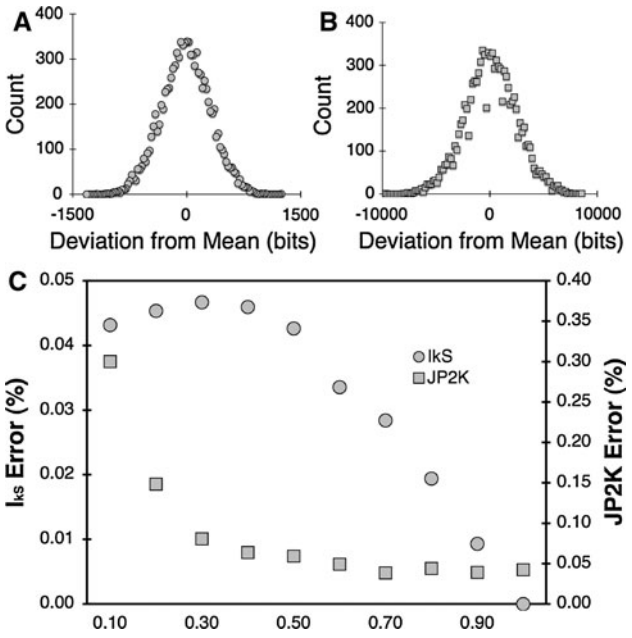


Fig. 4 Precision estimate for I_{kS}. **a** Distribution of the I_{kS} values for 50% shuffles plotted as a deviation from the mean (1,116,445 bits). Each of 10,000 uniform images with a uniform histogram was randomly shuffled to 50%, and the I_{kS} for each was computed. **b** Distribution of JP2K values for the same 10,000 images as in **a** (mean = 4,446,947 bits). **c** Precision of the I_{kS} and JP2K, represented by the %error (standard deviation/mean), as a function of shuffle fraction

expect the kSI for all such rotations to be identical. However, the kSI for rotations varies slightly (see supplemental information). This difference arises because of a shift of indices upon rotation of the image. The rotation-related error is small, and we typically ignore it. If needed, the index shift can be eliminated by padding the first row and the first column of an image with zeros (see supplemental information), although the padding itself introduces information into the image.

Composition Dependence of the Information

Two-dimensional images are typically intensity-based representations with variable histograms. For these types of images the I_{kS} and kSI are sensitive to the composition (i.e., histogram). This is an expected result from Parseval's theorem since the variance depends explicitly on the composition. The H_{kS}, on the other hand, is composition-independent. This somewhat counterintuitive result is related to the discrete computation of H_{kS}. When the integration window for the analysis is set to a fraction of the variance, the H_{kS} becomes identical for all images of the same size and the kSI for the 1.0 shuffled images approaches zero (irrespective of the histogram). Using a fixed integration window size would on the other hand produce a composition dependent value.

A Histogram-Independent Surface-Based Approach

One of our main interests here is to develop a metric that can be applied to a broad range of systems and that readily can be used to compare spatial information between systems. The histogram dependence described for the 2D gray-scale images above complicates comparisons. To solve this problem, we recast a 2D gray-scale image into a surface in a binary $n + 1$ dimensional structure, where the extra dimension is a binary representation of the intensity value at each position in the original structure (Fig. 5). For an eight-bit gray-scale image with N pixels, the additional dimension has 256 values, such that a gray-scale image is effectively recoded into a 3D surface. We adopt a nomenclature for the spatial information, (u:w)D, where u is the number of spatial dimensions and w is the number of additional dimensions that are used to encode nonspatial properties. Thus, the information and entropy computed for a 2D gray-scale image is (2:0)D; for the 3D representation these values are (2:1)D. Now, all images of the same size will have identical histograms: N values of 1, and 255 * N values of 0. Thus, all images of the same size have the same variance by Parseval's theorem, eliminating the histogram effect when comparing images. The approach is in effect similar to the symbolic vectorization of sequence data that has been used for DNA analysis (Jackson et al. 2000; Wang and Johnson 2002).

Aside from the histogram effects, the binary surface approach for the kSI metric qualitatively recapitulates all the results from the (2:0)D approach described above (Fig. 6). As with the (2:0)D case, the H_{kS} gives the largest possible value for I_{kS}, and the kSI decreases monotonically with increasing shuffle fraction. It should be noted that for the (2:1)D analysis we use an original image that is a flat plane where all the z values are the same and that the shuffling occurs in three dimensions. The shuffle series again shows a quadratic relationship between kSI and shuffle fraction. Interestingly, in this case the JP2K metric increases from the most ordered image to a 0.7 shuffle but then decreases slightly up to 1.0 shuffle. One notable difference between the (2:0)D and (2:1)D approaches is that the precision for the (2:1)D approach is significantly better. For example, for the 0.5 and 0.85 shuffled images, the (2:1)D standard deviation is $4 \times 10^{-9}\%$ and $7 \times 10^{-9}\%$, respectively. While these computations are more memory-intensive and require greater computational effort, we have found the (2:1)D approach for gray-scale images more useful than the (2:0)D.

Test of kSI Metric on an Ising Model

To provide a physical test of the kSI metric, we examined the phase transition in a 2D Ising model, for which there is an exact solution (Yang 1952). The Ising model and similar lattice models have been widely used to understand how lipids and other molecules are organized in bilayers (Scott 1988; Honerkamp-Smith et al. 2008). Recent findings suggest that Ising models not only help us to understand simple model membranes but also provide quantitative insight into the behavior of plasma membranes (Veatch et al. 2008; McConnell 2008). Here, we use the classic Ising ferromagnet. For this analysis energies and kSIs were computed for configurations at a series of temperatures (Fig. 7a, b). The energies for the system show the phase transition at a normalized temperature of 2.26 ± 0.01. From the kSI-based analysis the transition temperature was also 2.26 ± 0.01 (see supplemental information). These values agree to within error with the known transition

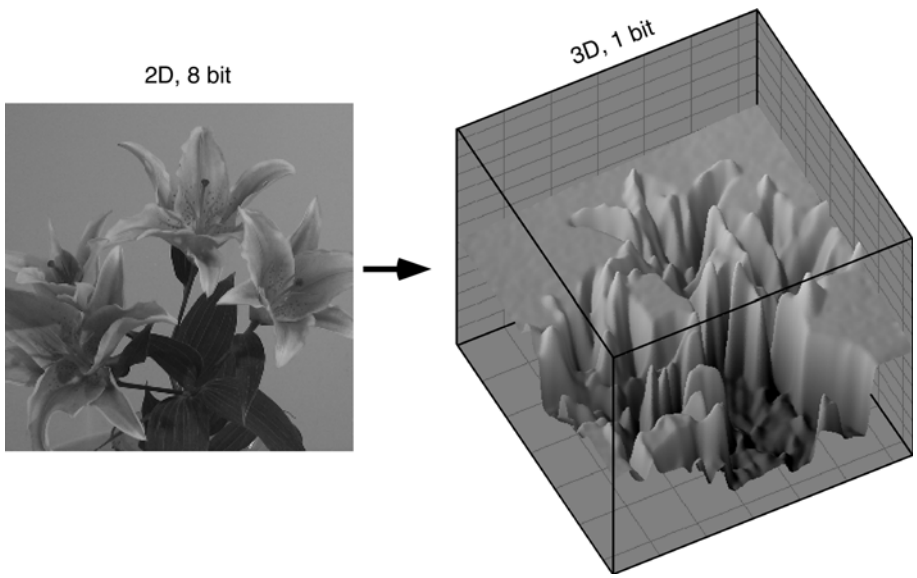


Fig. 5 Illustration of the mapping of a 2D gray-scale image in which (2:0)D metrics are computed to a 3D binary surface in which (2:1)D metrics are computed. The gray scale and shading in the 3D representation are for presentation purposes only

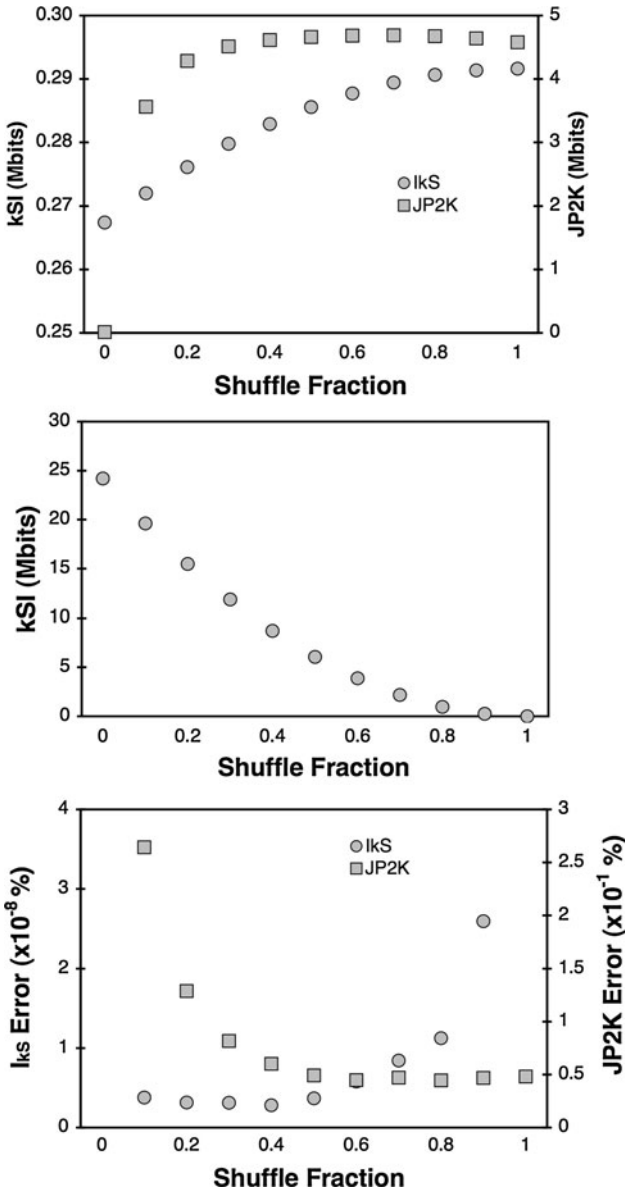


Fig. 6 (2:1)D spatial information analysis of a 3D shuffle series. This shuffle series begins with a 256 × 256 × 256 voxel one-bit image where the value for $z = 0$ at all xy points is 1 and all other values are 0. The shuffling is then carried out by randomly selecting an xy point and setting a random z value (voxel) to 1. **a** (2:1)D I_{kS} computed from the 3D (one-bit) representation and JP2K computed from the 2D (eight-bit) representation as a function of shuffle fraction ($I_{kS} = -242336.26457x^2 + 5329091.80210x + 262319053.10302$ and $R^2 > 0.99999$). The JP2K value is the minimum for the most ordered image but reaches the maximum at a shuffle fraction of ~ 0.7. **b** The kSI decreases monotonically with increasing disorder in an image. Also, the kSI fits the quadratic expression $kSI = 242035.31665967x^2 - 5325011.46447971x + 29290264.54958770$ and $R^2 > 0.99999999$). **c** Precision of the (2:1)D kSI and JP2K, represented by the %error (standard deviation/mean), as a function of shuffle fraction

temperature of 2.27 (Yang 1952). Thus, the phase transition of the Ising ferromagnet can be accurately determined from the spatial distribution of spins alone. We note that the energy and information in this analysis are close to linearly related over a large fraction of the temperature range, including the temperatures around the phase transition, but deviate from each other substantially at the highest temperatures.

The Ising model simulation can also be examined at different system sizes. We computed the information for several different system sizes in the range 100 × 100 and 400 × 400 pixels and found, as expected, that the kSI is strongly dependent on the size of the system. Likewise, the total k-space entropy depends on the size of the system. However, when normalized to the system size, both the kSI and entropy are scale-invariant. Further, the different-sized systems all produce the correct phase transition temperature.

Discussion

The main result presented here is an approach based on the FT to represent the spatial information content of a 2D image. Unlike approaches based on the Shannon formalism that use the distribution of pixel values as the probability distribution, this calculation takes into account the spatial arrangements in an image. The kSI metric provides a measure of how correlated the positions of pixels are relative to the limit where all the pixel positions are uncorrelated (as defined by the H_{kS}). It in effect describes an image by how rare it is relative to all possible images with the same histogram. The utility of this metric is illustrated by showing that it is sensitive to the phase transition in a 2D Ising model based only on the spatial distribution of spins.

Extending the kSI analysis of a 2D intensity-based data set to a 3D binary representation makes the metric histogram independent and improves the precision the kSI measurements, thus increasing the sensitivity significantly. Further, it enables direct comparison of kSI numbers for any images that are equal in size. Additional dimensions can also be used to encode other nonspatial information such as composition. Thus, our ongoing work with the kSI metric has been almost exclusively with the (2:1)D approach.

Contributions from Different Spatial Frequencies

We assume equal contributions to the kSI from the different spatial terms in the FT. However, it is likely that there will be reasons to use different weights, and specific weights might be employed to capture information at specific frequencies or combinations of frequencies. Such filtering of the contributions to the final sum, including setting some values to zero, is technically straightforward.

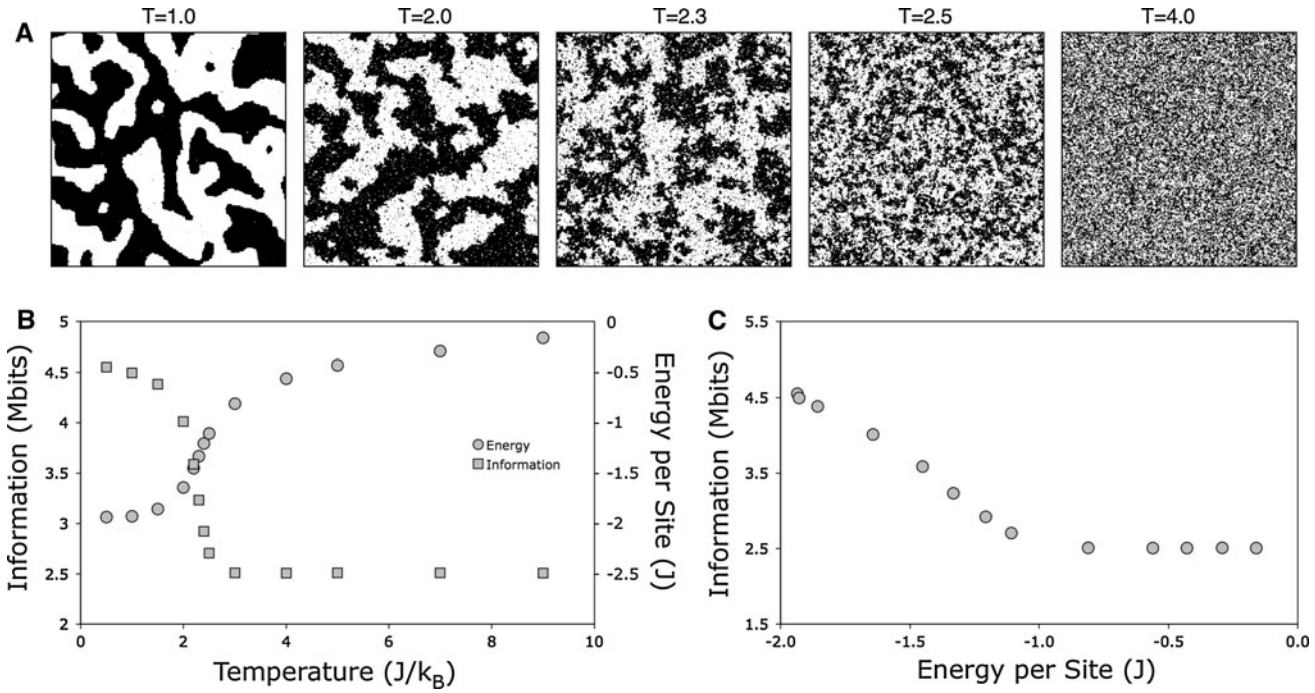


Fig. 7 Information-based analysis of the phase transition in a 2D Ising ferromagnet. **a** Representative snapshots of system configuration at different temperatures. The simulations were performed using Ising 1.1 (written by Dr. D. Schroeder). Configurations were generated by first randomizing the spins and then running the simulation for 5×10^7 steps for a 400×400 spin system (with periodic boundaries). The energy for each configuration was computed. The configurations were also converted to eight-bit gray scale TIFF files, and the kSI was computed for each. **b** The kSI shows a well-defined phase transition at a temperature of 2.26 ± 0.01 (see supplemental information). The well-known temperature-dependent phase transition of the computed energies of the system is shown for comparison. **c** Relationship between spatial information and energy in the 2D Ising ferromagnet

The Basis Set Issue

A limitation of the FT approach to computing information is that the FT does not provide an optimal basis set for all images. This is effectively the same problem that plagues the Kolmogorov complexity. Thus, the FT approach provides an upper bound to the spatial information in an image. The extent to which that limits the utility of the FT approach is not easily determined, but it is a subject for further study. Also, the approach used here can be used with many other basis functions, should there be other functions that are superior.

Higher Dimensions

In this article we describe the treatment of a one-component image—that is, an image with one type of element (a pixel) that varies only in intensity (the value of the pixel), as well as the projection of that image into three dimensions. However, the approach is readily extended to deal with higher (or lower) dimensional data as well as multiple components. In the case of multiple components, additional dimensions can be used to encode additional qualities of a pixel. The higher dimensions can represent any variable, atom type, time, temperature, connectivity, etc. Such vectorization is, as earlier noted, commonly seen in bioinformatic analysis of DNA sequences (Jackson et al. 2000; Wang and Johnson 2002). For example, sequences of DNA can be represented as a string of numbers with an arbitrary scale, such as $A = 1$, $T = 2$, $G = 3$ and $C = 4$. However, a number assignment is typically arbitrary, and the results of an analysis depend on the particular assignment used in a meaningless way. On the other hand, the base at each position in the sequence by a vector that has four elements can have values of either 1 or 0 to represent the presence or absence of a base, respectively, at that position. In this case the representation is independent of the assignment, and the entire sequence is described in a histogram-independent manner. Extending that logic, we can examine complex data structures with input from multiple sources by representing the different types of data in separate dimensions. The FT can be readily applied to structures of many dimensions, permitting the kSI for almost any data set to be calculated.

Conclusion

The *k*-space-based approach to computing information presented here is a facile and general method that yields a

metric for spatial information, and that provides the basis for further developments in many fields of study. We also note that the *k*-space framework can be extended to almost any application in which information theory is currently used, such as computing joint or mutual information between images. Information is used to characterize the complexity of physical systems (Ma 1999; Feldman and Crutchfield 2003), including the biochemical and biological systems of interest to us, and the kSI approach lends itself to such applications.

Acknowledgements The authors thank Dr. Thomas Woolf and Dr. David Haviland for interesting and helpful discussions, as well as Mr. Gardner Swan for programming assistance.

References

Applebaum D (1996) Probability and information: an integrated approach. Cambridge University Press, New York

Bardera A, Feixas M, Boada I, et al (2006) Compression-based image registration. Proceedings of the IEEE international symposium on information theory, July 9–14, Seattle, WA

Bekenstein JD (1973) Black holes and entropy. Phys Rev D 7:2333–2346

Bilbao JI, Ulloa JL, Guarini M, et al (2004) *K*-space information map. Proceedings of the International Society for Magnetic Resonance in Medicine Twelfth Scientific Meeting, May 15–21, Kyoto, Japan

de Polavieja GG (2004) Reliable biological communication with realistic constraints. Phys Rev E Stat Nonlin Soft Matter Phys 70:061910

Drenth J (1999) Principles of protein X-ray crystallography, 2nd edn. Springer, New York

Feldman DP, Crutchfield JP (2003) Structural information in two-dimensional patterns: entropy convergence and excess entropy. Phys Rev E Stat Nonlin Soft Matter Phys 67:051104

Fink TMA, Ball RC (2001) How many conformations can a protein remember? Phys Rev Lett 87:198103

Gatlin LL (1966) The information content of DNA. J Theor Biol 290:61974

Geiger B, Bershadsky A, Pankov R et al (2001) Transmembrane crosstalk between the extracellular matrix–cytoskeleton crosstalk. Nat Rev Mol Cell Biol 2:793–805

Honerkamp-Smith AR, Cicuta P, Collins MD et al (2008) Line tensions, correlation lengths, and critical exponents in lipid membranes near critical points. Biophys J 95:236–246

Hummer G, Garde S, Garcia AE et al (1996) An information theory model of hydrophobic interactions. Proc Natl Acad Sci USA 93:8951–8955

Jackson JH, George R, Herring PA (2000) Vectors of shannon information from Fourier signals characterizing base periodicity in genes and genomes. Biochem Biophys Res Commun 268: 289–292

Kimmel JM, Salter RM, Thomas PJ (2007) An information theoretic framework for eukaryotic gradient sensing. Adv Neural Inf Proc Syst 19:705–712

Lenaerts T, Ferkinghoff-Borg J, Schymkowitz J et al (2009) Information theoretical quantification of cooperativity in signalling complexes. BMC Syst Biol 3:9

Li M, Vitányi PMB (2008) An introduction to Kolmogorov complexity and its applications. Springer, New York

Ma YG (1999) Application of information theory in nuclear liquid gas phase transition. Phys Rev Lett 83:3617–3620

McConnell HM (2008) Understanding membranes. ACS Chem Biol 3:265–267

Page DL, Koschan AF, Sukumar SR et al (2003) Shape analysis algorithm based on information theory. Proc IEEE Int Conf Image Proc 1:229–232

Quastler H (1953) Essays on the use of information theory in biology. University of Illinois Press, Urbana

Rueckert D, Clarkson MJ, Hill DLG et al (2000) Non-rigid image registration using higher-order mutual information. Proc SPIE 2000:438

Russ JC (2002) The image processing handbook. CRC Press, Boca Raton, FL

Russakoff DB, Tomasi C, Rohlfing T et al (2004) Image similarity using mutual information of regions. Proceedings of the 8th European Conference on Computer Vision, May 11–14, Prague, Czech Republic, pp 596–607

Scott HL (1988) Anisotropic Ising model with four-spin interactions: application to lipid bilayers. Phys Rev A 37:263–268

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423

Shannon CE (1951) Prediction and entropy of printed English. Bell Syst Tech J 30:50–64

Tkacik G, Callan CG, Bialek W (2008) Information capacity of genetic regulatory elements. Phys Rev E Stat Nonlin Soft Matter Phys 78:011910

Veatch SL, Cicuta P, Sengupta P et al (2008) Critical fluctuations in plasma membrane vesicles. ACS Chem Biol 3:287–293

Vergassola M, Villermaux E, Shraiman BI (2007) "Infotaxis" as a strategy for searching without gradients. Nature 445:406–409

Wang W, Johnson DH (2002) Computing linear transforms of symbolic signals. IEEE Trans Signal Proces 50:628–634

Werbin JL, Heinz WF, Romer LH et al (2007) Micropatterns of an extracellular matrix protein with defined information content. Langmuir 23:10883–10886

Yang CN (1952) The spontaneous magnetization of a 2-dimensional Ising model. Phys Rev 85:808–815