arXiv:2303.12717v1 [physics.data-an] 22 Mar 2023

自然图像的多尺度相关性

Samy Lakhal, [1,2,3] Alexandre Darmon, [4] Iacopo Mastromatteo, [1,5] Matteo Marsili, [6] 以及 Michael Benzaquen[1,2,5,*]

[1] *Chair of Econophysics & Complex Systems, Ecole Polytechnique, 91128 Palaiseau Cedex, France*
[2] *LadHyX, UMR CNRS 7646, Ecole Polytechnique, 91128 Palaiseau Cedex, France*
[3] *Institut Jean Le Rond d'Alembert, UMR CNRS 7190, Sorbonne Université, 75005 Paris, France*
[4] *Art in Research, 33 rue Censier, 75005 Paris, France*
[5] *Capital Fund Management, 23 rue de l'Université, 75007 Paris, France*
[6] *Quantitative Life Sciences Section The Abdus Salam International Centre for Theoretical Physics, 34151 Trieste, Italy*

（日期：2023年3月23日）

我们采用一种非预设的信息论方法，研究自然图像的统计特性。我们提出多尺度相关性（MSR）度量 [1]，用于评估图像在所有尺度下抵抗压缩的稳健性。首先，在受控环境中，我们将合成随机纹理的 MSR 表征为图像粗糙度 H 及其他相关参数的函数。随后，我们将分析扩展到自然图像，并发现其与临界（$H \approx 0$）随机纹理存在显著相似性。我们表明，与功率谱分析等经典方法相比，MSR 更稳健，也能更充分地反映图像内容。最后，我们将 MSR 与经典度量进行比较，以校准色彩映射和去噪等常见处理流程。总体而言，MSR 方法有望成为高级图像分析和图像处理的良好选择，同时提供较高水平的物理可解释性。

目录

引言 1

I. 分辨率/相关性框架 2
A. 精度与可解释性之间的权衡 2
B. 分辨率与相关性 2
C. 在图像中的应用 3

II. 随机纹理的相关性 4
A. 关于 $1/f^\alpha$ 高斯场 4
B. 随机纹理的多尺度相关性 4
C. 最相关的分割 5

III. 自然图像的相关性 6
A. 关于灰度场 $h(\boldsymbol{r})$ 6
B. 关于梯度幅值 $|\nabla h|$ 7

IV. 在图像处理中的应用 7
A. 色彩映射 8
B. 使用 Rudin-Osher-Fatemi 算法进行去噪 9

V. 结论 9

VI. 致谢 10

A. $1/f^\alpha$ 高斯场的生成 10

参考文献 10

* michael.benzaquen@polytechnique.edu

引言

图像处理的最新进展得益于强大学习框架的兴起，这些框架将高效的架构 [2–4] 与大型高质量数据库 [5, 6] 相结合。尤其是，神经网络通过层叠简单的线性与非线性算子（如卷积矩阵或激活函数），已证明能够非常高效地对高维数据进行分类或生成。如今，它们已能以前所未有的效果捕捉图像之间的相似性。然而，尽管架构深度增加会提升其性能，但通常也会以牺牲物理可解释性为代价。理解学习动力学以及生成图像的统计特征，仍是学界面临的一项挑战 [7, 8]。

在机器学习算法出现之前，压缩、去噪 [9, 10], 或边缘检测等任务，通常（在某些情况下至今仍然）使用信号处理方法来完成。在经典方法中，第一类基于特定度量，例如广泛使用的峰值信噪比（PSNR） [12], , 这类度量建立在常见的信号处理度量指标（欧几里得距离、功率谱等）之上。第二类则通过基于视觉的实验构建半经验相似性度量，例如结构相似性指数（SSI） [13]。在这两种情况下，方法都是完全确定性的，这意味着粗糙度、平稳性或局部相关性等随机性质均被忽略。

在统计物理的背景下，近年来，人们采用一种新颖且完全不依赖先验假设的方法来解决高维数据推断问题。该方法旨在测量有限规模样本的特定属性 [14], , 通过评估预设压缩过程对简单熵度量的影响来实现

。这种方法已应用于生物学推断 [1]、金融 [15]、语言模型 [14] 和最优机器学习 [16, 17]，并取得了令人振奋的成果。本文将这一形式体系改用于图像分析和图像处理，重点关注自然图像。自然场景或景观因呈现出可区分的统计特征而长期受到研究，这些特征包括尺度不变性 [18–20]、非高斯性 [21] 或图像块临界性 [22]。

本文的结构如下。第一节通过一个示例介绍分辨率／相关性形式体系，并将其调整为适用于图像分析。第二节分析一类可参数化图像，即随机 $1/f^{\alpha}$ 高斯场，并介绍多尺度相关性（MSR）。第三节将分析扩展到自然图像及其梯度幅值，并讨论它们与合成高斯场之间有意义的统计相似性。第四节展示如何将 MSR 方法用于常见的图像处理任务。

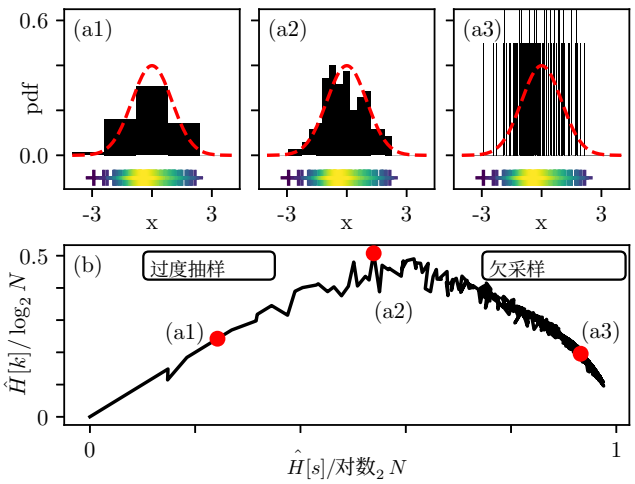

图 1. 高斯分布样本（$N = 100$）的相关性分析。(a) 箱数 n 对归一化直方图（黑色柱）的影响，其中 (a1) $n = 5$、(a2) $n = 23$ 和 (a3) $n = 400$。红色曲线表示潜在分布。底部标记（+）表示初始样本数据点，颜色表示局部数据密度。(b) 分辨率/相关性曲线。

I. 分辨率/相关性框架

本文介绍近期由我们中的一位构建的信息论框架，该框架用于对高维数据样本及其在压缩过程中的行为进行无先验分析。相关度量指标由压缩样本的简单统计量推导而来。

A. 精度与可解释性之间的权衡

我们来考虑分箱问题，即将随机变量 X 的样本聚类为若干组，每组中的 X值相近。如果采样得到的数据点 $\mathcal{S} = \{x_1, \ldots, x_N\}$ 都处于不同状态（例如，当 X 的分布为连续分布时），经验分布就是一个狄拉克梳状函数。为了深入了解采样变量，可以使用经过精心选择的箱/区间绘制直方图来可视化数据。实际上，这一过程通过压缩降低数据分辨率，促使结构显现，从而提高可解释性。随后可以对底层过程作出假设，并寻找最优参数，以最佳地描述数据。

我们通过对高斯变量 $X \sim \mathcal{N}$（0，1）的 $N = 100$ 次实现进行采样来说明这一直觉，如图 1 所示。数据被分入 n 个相同的箱，分别对应 $n = 5$、23 和 400 这三个不同取值。我们还定义了箱宽 ℓ 作为压缩参数，将原始样本 $\mathcal{S}$ 转换为压缩样本 $\mathcal{S}^{\ell}$。压缩步骤是将每个数据点替换为相应的直方图柱索引。图 1(a1)（较大ℓ）展示了过度抽样的情况。仅使用五个箱，便会损失大量数据分辨率。在相反，图 1(a3)（较小 ℓ）对应欠抽样 状态，箱非常窄（大多仅包含一个数据点），因此所得分布接近狄拉克梳状函数。图 1(a2)（中间 ℓ）则似乎是一个合理的折中：直方图在视觉上接近生成器密度，这表明我们可能接近数据压缩的最佳级别。基于这一观察，人们可能会倾向于采用高斯模型，并为均值和方差选取合适的估计量。然而，这一决定完全取决于特定的压缩级别，因此未能充分利用当前样本。

我们在下一节中引入的形式体系提供了一种有原则的方法，将压缩级别的选择与一个最优性准则联系起来；该准则不受生成数据所依据的生成模型性质影响。

B. 分辨率与相关性

Marsili等人 [15] 先前的研究通过引入可观测量，解决了<覆采样/欠采样 $\mathcal{S}^{\ell} = \{s_1^{\ell}, \ldots, s_N^{\ell}\}$ 转变的问题；这些可观测量可用于监测由参数 ℓ压缩 $\mathcal{S}$后得到的缩减样本 $\mathcal{S}^{\ell} = \{s_1^{\ell}, \ldots, s_N^{\ell}\}$ 中的变化。首先，考虑相同状态的数据点数量 s 以及在 $\mathcal{S}^{\ell}$中出现 k次的状态数量 m_k^{ℓ}。由此可得$\sum_s k_s^{\ell} = \sum_k k m_k^{\ell} = N$。例如，在图 1(a1)所示的压缩样本中，$k_s^{\ell}$ 的取值为$\{2, 26, 49, 23, 0\}$；由于直方图中的每个柱条高度各不相同，因此有 $m_2 = m_{26} = m_{49} = m_{23} = m_0 = 1$，否则为 $m_k = 0$。

定义分辨率$\hat{}H^{\ell}[s]$ 和相关性$\hat{}H^{\ell}[k]$如下：

$$\begin{cases} \hat{H}^{\ell}[s] &= -\sum_s \frac{k_s^{\ell}}{N} \log_2 \frac{k_s^{\ell}}{N}, \\ \hat{H}^{\ell}[k] &= -\sum_k \frac{km_k^{\ell}}{N} \log_2 \frac{km_k^{\ell}}{N}. \end{cases} \tag{1}$$

《style id='1'》分辨率《/style》是经验分布 $\{p_s^{\ell} = k_s^{\ell}/N\}_s$ 的熵，描述了对 $\mathcal{S}^{\ell}$中的一个状态概率进行编码所需的平均比特数。压缩会将数据点聚集在一起，从而降低平均编码成本。对于原始数据，分辨率 达到最大值，并随着 ℓ单调递减，直到达到熵最小的完全压缩样本。相关性 是分布$\{q_k^{\ell} = km_k^{\ell}/N\}_k$的熵，也就是从 $\mathcal{S}^{\ell}$ 采样的数据点在样本中出现 k 次的概率。这是对 p_s^{ℓ}的压缩版本，其中相同的频数状态被聚集在一起，并在此过程中舍弃其标签 s 。因此，只需知道 q_k^{ℓ} 即可构建不带标签的直方图，这等同于假设采样次数相同的状态不可区分。按频数值降序排列这些状态，就会得到著名的齐普夫图。最终，相关性 编码了每个柱子的高度；当 $\{km_k^{\ell}/N\}_k$均匀分布时，其值最大，从而得到 $m_k \propto k^{-1}$。我们在表 I 中报告了典型的采样情形及其对应的分辨率/相关性值。

回到高斯抽样的例子，图 1(b) 展示了随 ℓ 变化而得到的 $\hat{}H^{\ell}[k]$ 与 $\hat{}H^{\ell}[s]$ 的函数关系。 $n = 5$ (a1)、 $n = 23$ (a2) 和 $n = 400$ (a3) 的对应值也特别标出。请注意，(a2) 使相关性最大，而 (a1) 和 (a3) 分别对应过度抽样和欠采样。此处需要强调的是，尽管在这个特定例子中视觉上看起来如此，样本 (a2) 并不一定会使底层分布与经验分布之间的距离最小。有趣的是，分辨率/相关性属性仅取决于原始样本 $\mathcal{S}$ 和压缩参数 ℓ，因此整个方法与生成过程无关。因此，最值得关注的是样本在压缩过程中，从欠采样逐渐过渡到过度抽样时的变化方式。由此，必须选择一种能够实现这两种状态相互转换的压缩过程。

表 I. 典型采样情形。

情形	采样状态	$\{m_k\}_k$	$H[s]$	$H[k]$
Full 过度抽样	相同	$m_N = 1$ $m_k = 0$	0	0
Full 欠采样	不同	$m_1 = N$ $m_k = 0$	对数 N	0
中间采样	中间	$m_k \propto k^{-1}$	H_0	最大值 $H[k]$

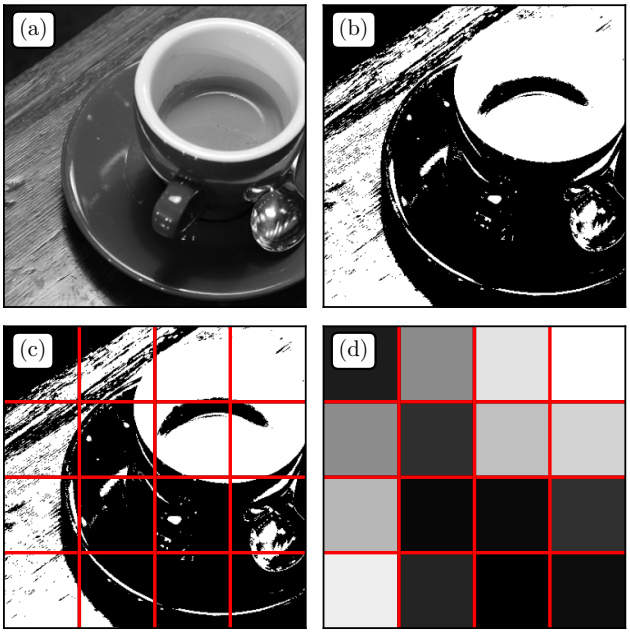

图2. 经典基准图像上的分割/压缩过程示意图。(a) 原始图像。(b) 在给定分位数值 a 下的阈值化图像。(c) 网格缩减后的阈值化图像。(d) 缩减样本，其中每个网格单元都由平均像素颜色替代。

C. 在图像中的应用

图像通常被描述为场 $h(\boldsymbol{r})$，其中 $\boldsymbol{r} \in \{1, \ldots, N_X\} \times \{1, \ldots, N_Y\}$。这等价于一个由 $\mathcal{S} = \{(\boldsymbol{r},\ h(\boldsymbol{r})\}$ 组成的样本，其大小为 $N = N_X N_Y$，用于描述每个像素的位置和颜色。自然地， $\mathcal{S}$ 处于完全欠采样状态，因为每个数据点都是唯一的。

因此，为了压缩灰度图像，我们提出一种包含两个步骤的简单流程：(i) 分割；(ii) 空间压缩，如图2所示。分割是指使用阈值 a （黑色像素的比例）将灰度级转换为黑白像素，从而得到二值图像 $h_a(\boldsymbol{r})$（图2(b)）。这降低了样本中可能的颜色状态数量，是达到完全过采样状态的必要条件。注意，可以通过对所有分割结果取平均值来重建原始图像。此步骤也可推广至颜色图像，例如在 RGB 色彩空间中使用三元组 $(a_{\mathrm{R}},\ a_{\mathrm{G}},\ a_{\mathrm{B}})$。第二步是压缩像素位置（图2(c)）。将每个坐标 $\boldsymbol{r}$ 替换为其在步长为 ℓ 的网格上的位置索引 $\boldsymbol{r}^{\ell}$ 。最终得到一个压缩样本：

$$\mathcal{S}_a^{\ell} = \{(\boldsymbol{r}^{\ell}, h_a(\boldsymbol{r}))\}. \tag{2}$$

随后，将每个像素值替换为缩小后的网格中对应的平均值（图2(d)）。最后，$k^{\ell}_{(\boldsymbol{r}_{\ell},0)}$ 和$k^{\ell}_{(\boldsymbol{r}_{\ell},255)}$ 分别定义为单元格中的黑白像素数量

$\boldsymbol{r}^\ell$，而 m_k^ℓ 定义为在尺度 ℓ下包含 k 黑色或白色像素的像素单元数量。利用公式（1），可以计算出后续将使用的$\hat{}H^\ell[s]$ 和$\hat{}H^\ell[k]$的值。

我们可以将这一压缩过程与卷积神经网络（CNN）等图像处理架构直接类比 [2]。首先，它们的组成层通常将空间压缩过程（即第一次线性卷积）与可训练层或预设层结合起来。随后，通过对像素值进行称为激活函数的非线性变换来执行分割。类似地，我们的过程由一个单层网络构成，以 $\mathcal{S}$ 作为输入，并输出 $\mathcal{S}_a^\ell$。有趣的是，我们无需将特定卷积矩阵指定为算法的输入，只需提供一个尺寸参数，从而使我们的方法更具通用性。最后请注意，任何能够实现欠采样/过采样转换的压缩过程都可以采用。例如，可以使用离散傅里叶系数或小波系数，这些系数常用于 JPEG 压缩算法 [9, 10]。另一种方法是使用经过训练或未经训练、具有二值激活函数且层大小可调的网络（类似感知机）的中间表示，如深度神经网络架构中的分辨率/相关性权衡所示 [16]。

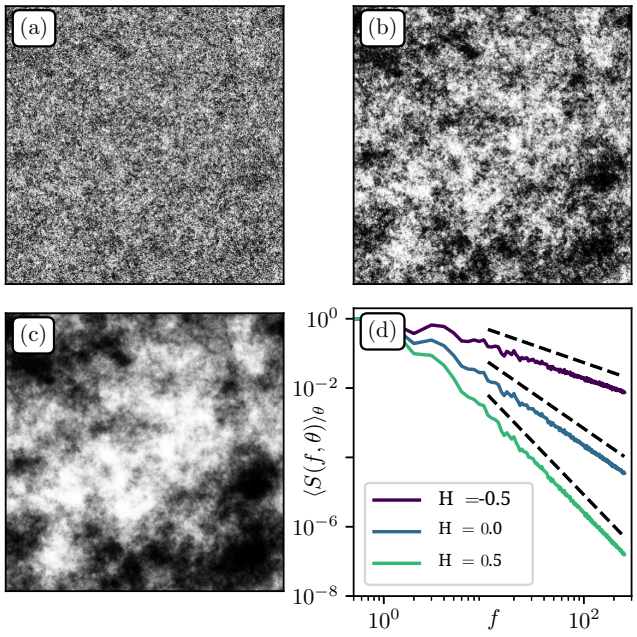

图3. $1/f^\alpha$ 由相同白高斯噪声种子生成的纹理。(a)、(b)、(c) 分别为粗糙度 H $= -0.5$，0，0。5、空间分辨率为 512×512 的 $1/f^\alpha$ 随机场表征。(d) 方位角平均功率谱 $\langle S(f,\ \theta)\rangle_\theta$。黑色虚线表示理论功率谱随 H $\alpha = 1 +$的衰减 $1/f^{2\alpha}$。

II. 随机纹理的相关性

本节将说明如何在一类简单却广泛存在的过程中使用度量指标（$\hat{}H^\ell[s]$，$\hat{}H^\ell[k]$）：二维 $1/f^\alpha$ 随机高斯场。我们首先回顾这类场的性质，然后研究 α 对分辨率与相关性的影响。

A. 关于 $1/f^\alpha$ 高斯场

$1/f^\alpha$ 高斯场由初始互不相关的二维白噪声经过线性滤波构成（见附录 A）。后者具有平坦的傅里叶谱，随后乘以$1/f^\alpha$，因此得到按 $1/f^{2\alpha}$缩放的幂谱。这会在空间域中形成空间相关性。这种幂律滤波器引入了缩放特性，通常用粗糙度赫斯特指数 H 描述：$= \alpha - d/2$ 其中，d 为场维度（此处为 $d = 2$）。根据 H 的正负，可以得到两类过程。当 H < 0 时，随机场是平稳的，即均值固定，且相关性 $C(\delta r) \propto \delta r^{2\mathrm{H}}$ 在滞后距离 δr处。当 H $= -d/2$ 时，对应未修改的谱（白噪声）。当 H

> 0，该过程不再是平稳的，但具有平稳增量，其标度为 $\langle[h(\boldsymbol{r}+\delta\boldsymbol{r}) - h(\boldsymbol{r})]^2\rangle \propto \delta r^{2\mathrm{H}}$。我们生成了三个粗糙度值 H $\in \{-1/2, 0, 1/2\}$ 各不相同的样本，分别如图3(a)、(b)和(c)所示。赫斯特指数会影响粗糙度的视觉表现，随着 H 增大，图像会变得更加平滑。图3(d)显示了方位角平均功率谱 $S(f) = \langle|\tilde{h}(f,\ \theta)|^2\rangle_\theta$，由此可验证生成方法具有稳健性，因为得到了预期的标度行为和指数。

B. 随机纹理的多尺度相关性

现在，我们对图3所示的场执行上述分割。阈值 $a = 0.5$ 对应的纹理如图4(a)-(c)所示，相应的分辨率/相关性曲线 $(\hat{}H^\ell[s],\ \hat{}H^\ell[k])_{\ell\in\{1,\ldots,N\}}$ 绘于图4(d)。

可以看出，对于 H $= -0.5$ （图4(a)）和 H $= 0$ （图4(b)），图案基本保持不变；但对于 H $= 0.5$ （图4(c)），情况并非如此，分割过程产生了大片色调均匀的区域。这是由于空间相关性更强，使图案更具持续性，并减少了围绕平均值的涨落。此外，正如视觉质量评估实验 [23], 所报告的，H $= 0$ 纹理在所有尺度上都呈现出有趣的视觉特征，而 H $= -0.5$的这些特征似乎仅限于小尺度。要将这些观察结果与图4(d)中的相关性曲线联系起来并不容易，因为相对相关性会随分辨率而变化。因此，考察所有压缩级别下的相关性似乎更为自然。

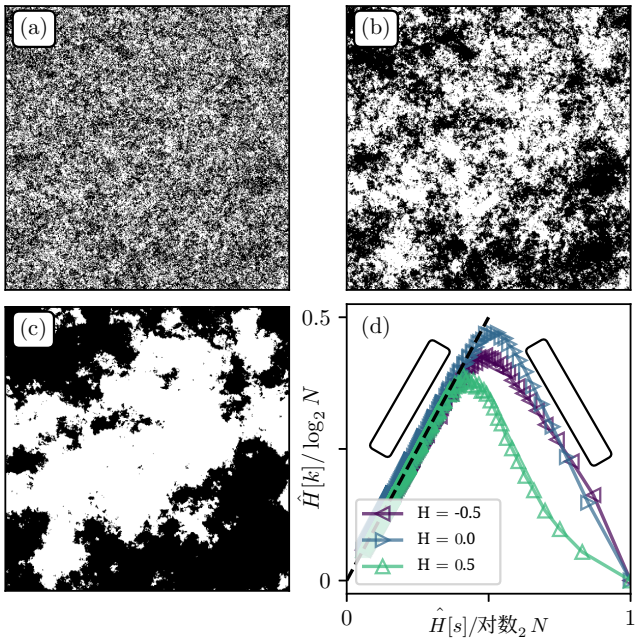

图4。(a)、(b)、(c) 分别为图3中 H $=-0.5$，0，0。5 时纹理的分割版本，阈值为 $a=0.5$。(d) 以最大熵 $\log_2 N$ 归一化的分辨率/相关性曲线。

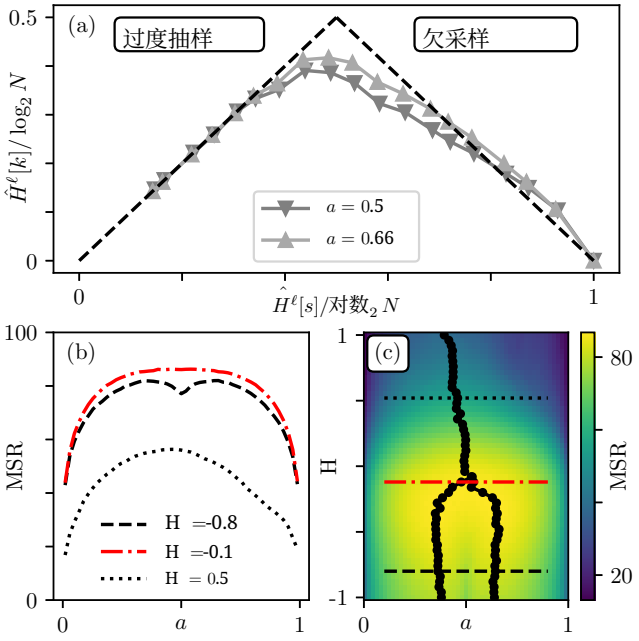

图5. 分割值 a的影响。(a) 两个 a值下 H $=-0.8$ 的相关性曲线。(b) H $=-0.8$ （黑色虚线）、H $=-0.1$ （红色点划线）和 H $=0.5$ （黑色点线）的 MSR 随 a 的变化关系。(c) MSR(H a) 密度图。极大值以黑色标记表示。

为此，我们引入一种衡量样本对压缩整体稳健性的指标，称为多尺度相关性（MSR），其定义如下：

$$\mathrm{MSR} := \int \hat{H}^\ell[k]\, d\hat{H}^\ell[s], \tag{3}$$

这实际上就是分辨率/相关性曲线下方的面积。这一度量最初在[1] 中被提出，用作表征神经元活动时间序列的序参量；正如复杂度度量所预期的那样，它成功地区分了有用信息与环境噪声 [24]。需要注意的是，尽管文献中已提出若干基于多尺度熵贡献的复杂度度量[25, 26], ，多尺度相关性（MSR）的不同之处在于，每个尺度的贡献都由分辨率自然加权。其他度量通常对每个压缩级别赋予相同权重。对于图4中的图像，可得 MSR(H $=0.5$) $<$ MSR(H $=-0.5$) $<$ MSR(H $=0$)。这与我们先前的视觉印象一致：图3(b)中的纹理似乎在不同尺度上包含更多信息。

C. 最相关的分割

人们自然会认为，分割阈值 a 会影响相关性。在给定的 H <0 下，相关性最高的表征似乎并不对应于 $a=0.5$。图5(a)证实了这一点： $H=-0.8$ 的相关性曲线在 $a=0.66$ 时高于 $a=0.5$。图5(b)展示了三个 H 值下，多尺度相关性随 a 变化的函数曲线。当 H $=-0.8$ 时（虚线），可以观察到两个对称的极大值，分别位于 $a_c=0.5\pm.13$，这与图5(a)一致。有趣的是，通过选择“背景画布”打破像素分布的对称性，可以获得在分辨率/相关性方面更有趣的样本。如图5(c)所示，当 H ≈ 0 时会出现分岔，低于该值时，多尺度相关性的两个极大值并存。H $<-1/2$ 时得到的 a_c 值接近二维方格晶格 [27] 上的经典渗流阈值 $a^*=0.59$。事实上，我们分割后的图像等价于相关渗流位点问题的样本。特别是，Prakash 等人 [28] 观察到，正如我们在此所见，当 H $\to 0$ 从下方趋近时，两个极大值会在 $a_c=0.5$ 处连续汇合，同时多尺度相关性(a)曲线在该值附近变平（见图5(b)）。在这一临界点，图像的信息含量对分割过程的敏感度降低。

当 H $\gtrsim 0$ 时，MSR(a) 在 $a_c=0.5$ 处呈现唯一的最大值。然而，随着 H 进一步增大，相关性的范围也随之扩大，从而产生有限尺寸效应。由此得到的 a_c 会非常依赖噪声，因为不同样本会导致不同的临界阈值。有趣的是，二维分数布朗运动渗流 [29]中也曾报告过这种行为。

III. 自然图像的相关性

现在，我们将关注自然图像，即自然场景和景观的图片。由于这类图像具有稳健的统计特征，如尺度不变性和临界性，文献中长期以来一直对其进行研究 [18–21, 30], 。

A. 关于灰度场 $h(\boldsymbol{r})$

图 6(a) 展示了 Tkacik等人 [31] 在博茨瓦纳奥卡万戈三角洲拍摄的照片，照片被描述为“[...] 热带稀树草原栖息地，与人眼被认为进化形成的环境相似”。图像被划分为十五个大小为 512×512 像素的图像块。可以观察到各种各样的图案，从天空中均匀的浅灰色调，到树枝处明显的不连续，以及嘈杂的植被纹理。

所有图像块的功率谱分析如图 6(b) 所示。高频极限下的形状是由相机标定、光学模糊或后处理程序造成的，这些因素与图像块内容无关。在低频处，我们观察到按指数 -2.0 ± 0.1 衰减的幂律。请注意，尽管存在一些可能与图像块特征有关的微小涨落 [30]，但功率谱分析似乎无法捕捉——

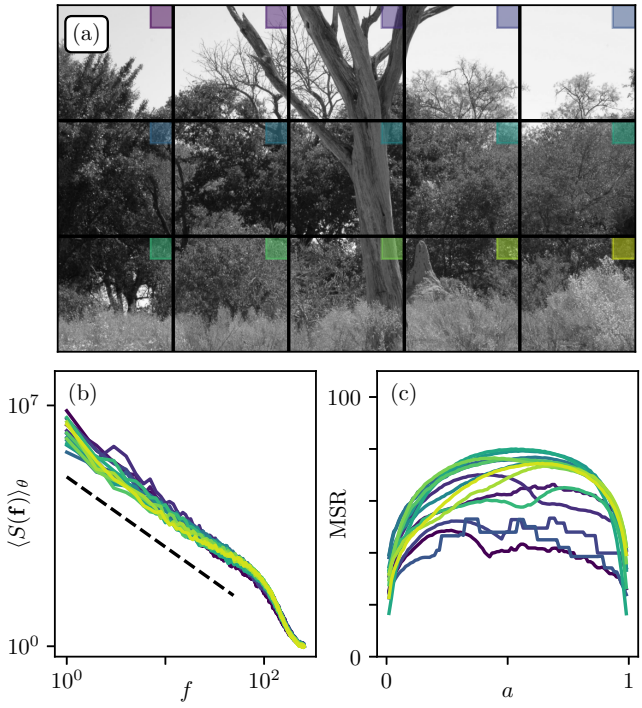

图 6. (a) 来自 [31], 的自然灰度图像，按大小为 512×512 的图像块进行分割。(b) 每个图像块的功率谱。虚线表示指数为 -2 的衰减幂律。(c) 每个图像块的 MSR 随 a 的变化。

上文提到的不同图像块之间的视觉异质性。

也就是说，$S(f) \sim 1/f^2$ 对应于粗糙度指数 $\mathrm{H} = 0.0 \pm 0.1$，这恰好是第二节中 MSR 对随机纹理表现出临界且非平凡行为的范围。因此，我们预计 MSR 方法或许能够更精细地刻画每个图像块。经典谱分析的另一个问题是，图像的功率谱预计会对其颜色直方图的非线性变换极其敏感，即使是保持视觉效果不变的单调变换也是如此。使用多尺度分辨率方法则不存在这一问题，因为分割参数 a 决定了黑白像素的比例，而与颜色直方图的形状无关。

图 6(c) 显示了所有图像块的多尺度分辨率曲线。首先可以观察到，多尺度相关性的取值范围与第 II 节中 $\mathrm{H} \approx 0$ 纹理的取值范围大致相当。然后，

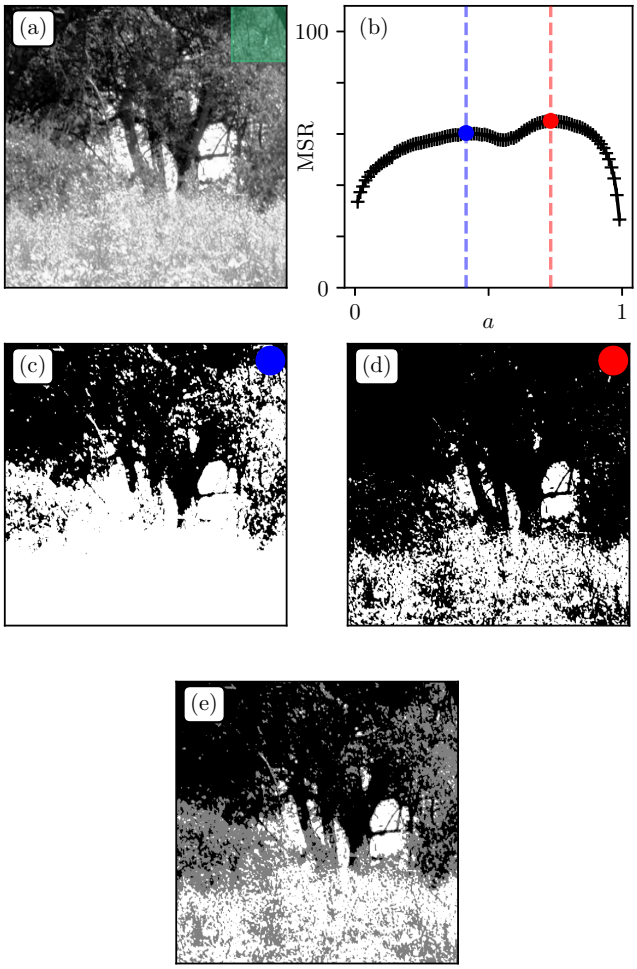

图 7。(a) 图 6(a) 的左下图像块。(b) 多尺度相关性随 a 变化的函数图，并突出显示了临界阈值 $(a_1,\ a_2) = (0.42,\ 0.73)$。(c)(d) 对应的分割图块。(e) 将 (c) 和 (d) 相加得到的图像，具有三个颜色层级 $\{0,127,255\}$。

可以清楚地看到，各图像块的多尺度分辨率存在显著差异。主要包含茂密纹理、图案没有突变的图像块，其多尺度分辨率曲线有一个独特的极大值。请注意，某些情况下出现的奇异点是由于直方图中某些颜色所占比例过高（均匀的天空）造成的。包含异质色调，或包含大小不一的实体物体（如树干、树枝和灌木）的图像块（例如图 6(a) 左下方）往往有两个极大值，这与 $\mathrm{H} < 0$ 的情况类似（见第 II 节）。

图 7 聚焦于图 6(a) 左下方的图像块。这个子图像似乎呈现出两个明显占主导地位的颜色层级。实际上，这些层级对应于图 7(b) 中多尺度分辨率曲线的极大值。图 7(c) 和 7(d) 的分割结果直观地证实了这一点，它们分别最能捕捉图像顶部和底部的涨落。我们强调，后两种表征共同构成了对 (a) 最具信息量的分割结果。将它们叠加（图 7(e)）确实能以仅含三个颜色层级的结果良好近似原始图像 {0,127,255}。因此，多尺度分辨率方法似乎能够很好地反映自然图像内容的多样性，而这种多样性无法通过经典功率谱分析获取。

B. 关于梯度幅值 $|\nabla h|$

为了进一步理解自然图像的结构，我们现在聚焦于梯度幅值场，它旨在捕捉轮廓或边界等显著的空间不规则性。此外，求取梯度的优点在于可以使初始场平稳化。梯度分析是各种图像处理流程中的基础环节，应用范围从经典边缘检测 [32], 到机器学习中的有监督 [33] 或无监督 [2] 分类架构。从更侧重感知的心理物理学视角来看，研究表明，可以直接通过对梯度场的视觉评估推断方向、几何形状和位置等关键信息 [34–36]。我们通过小波卷积计算梯度 $|\nabla h|$ 。这种方法如今已得到广泛应用，因为它在信号处理任务中表现出出色的稳健性 [37–41]。在有：

$$|\nabla h| = |h * \boldsymbol{\psi}_j(\boldsymbol{r})|^2, \tag{4}$$

其中，$\boldsymbol{\psi}_j := (\psi_{j,\,x},\ \psi_{j,\,y})$ 是一个特征二进尺度为 2^j 的小波梯度滤波器。该小波由梯度与高斯窗混合而成，后者的标准差为 $\sigma_j = 2^j$ 像素。

使用 $j = 0$ 得到图 8(a) 所示的图像。正如预期，对于植被的起伏纹理或树枝等锐利轮廓，得到的信号较强（亮色调）；而对于天空等平滑且均匀的区域，信号值较低（暗色调）。

接着，我们对这些新的图像块进行多尺度相关性分析（图8(b)），并发现大多数图像块的多尺度分辨率曲线都很平坦。这相当于第 II ≈ 0

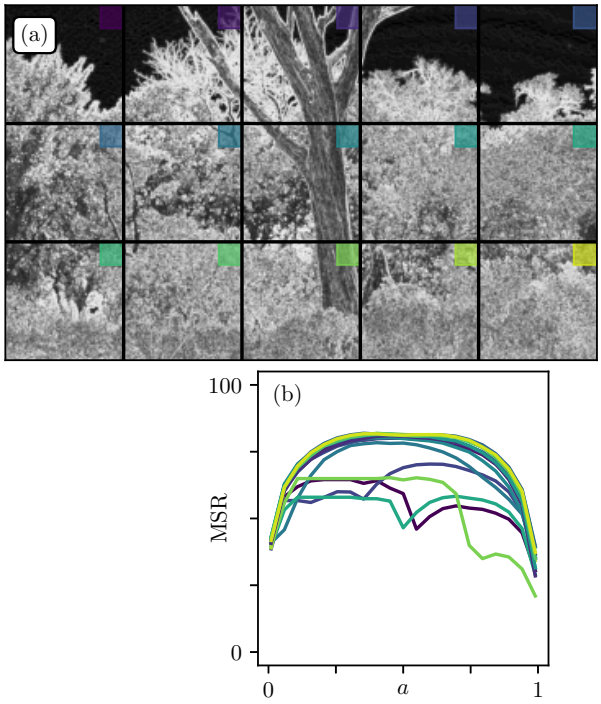

图 8。（a）图 6(a) 的梯度幅值场，其中 $j = 0$，划分为 512×512 个图像块。（b）不同图像块的多尺度相关性随 a 的变化。

节中所述具有对数相关性的临界情形（见图5）。确实可以将自然图像视为由各种尺寸物体拼贴而成；这种图案的叠加让人联想到加性级联过程 [42] ，它们也呈现出对数相关性。

现在我们来探究改变小波尺寸的影响（见图 9）。我们选取图 8(a) 中上方中间的图像块，因为其中包含较大的物体和小尺度细节。如图 9(c) 和 (d) 所示，增大 j 会对较小的涨落进行粗粒化，只留下较大的涨落。这会使低压缩率下的相关性降低，进而降低整体多尺度相关性（图 9(b) ）。最后，在临界阈值下得到的分割后梯度场（图 9(e) 和 (f)）在视觉上仍与初始场（图 9(c) 和 (d)）相近。这是意料之中的，因为梯度幅值在实体物体的轮廓处已经呈现出很大比例的黑白像素。

IV. 图像处理中的应用

这里我们展示 MSR 在常见数字图像处理任务中的潜力，即色彩映射和去噪。

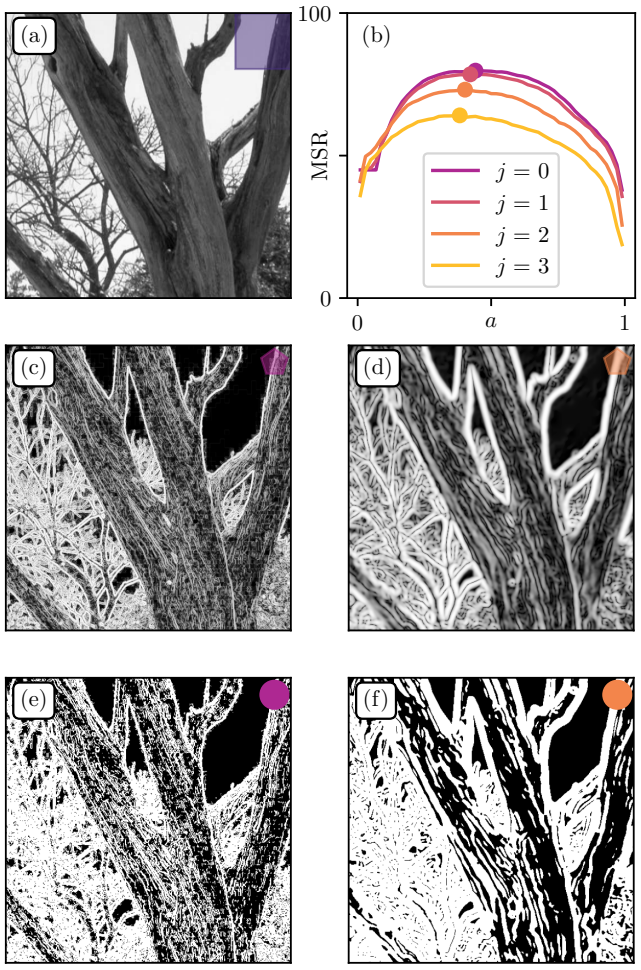

图 9. 梯度小波尺寸的影响。(a) 来自图 6(a) 的原始图像块。(b) 二进尺度梯度小波（2^j）的多尺度相关性随 a 变化的函数关系 $j \in \{0,1,2,3\}$。(c)(d) 分别对应 $j=0$ 和 $j=2$ 的梯度幅值。(e)(f) 分别对应 $j=0$ 和 $j=2$ 在临界阈值 a_c 处的分段梯度幅值。

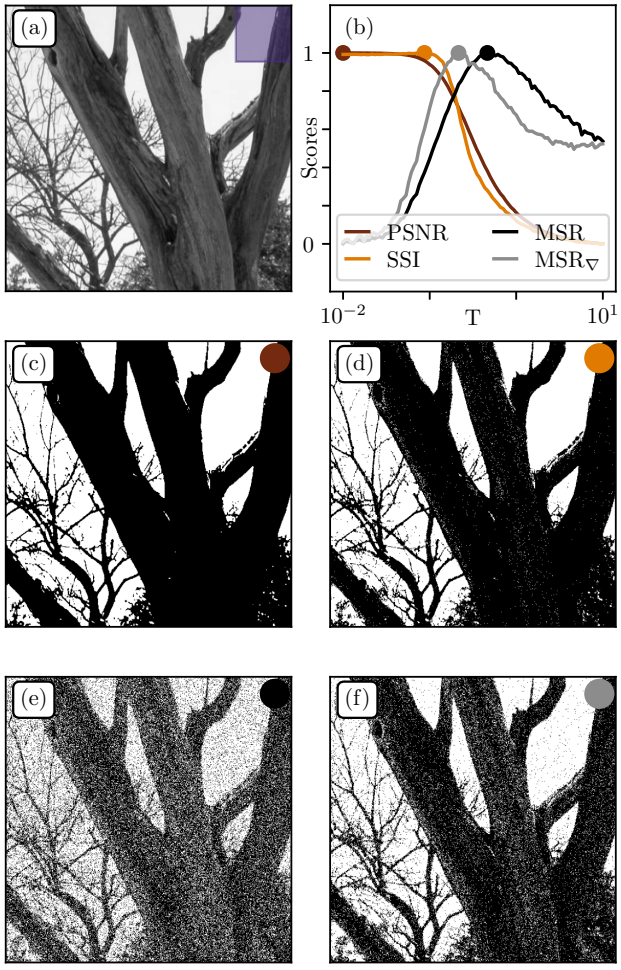

图 10. 色彩映射。(a) 来自图 6(a) 的原始图像块。(b) 不同性能度量的重缩放得分随温度变化的函数关系：峰值信噪比（PSNR）、结构相似性指数（SSI）、直接多尺度相关性（MSR），以及梯度场上的 MSR（MSR_∇）。(c)、(d)、(e)、(f) 分别为 PSNR、SSI、MSR 和 MSR_∇ 在最优温度 T^* 下的色彩映射。

A. 色彩映射

考虑这样一个色彩映射问题：将像素值映射到缩减调色板上。为简单起见，我们考虑将初始灰度调色板映射为二值 $\{0,255\}$（黑白）的情形。我们使用玻尔兹曼分布 $\mathbb{P}(c|h_{ij}) \propto e^{-(h_{ij}-c)^2/T}$ 实现一个随机映射过程，其中 h_{ij} 是像素的原始颜色，其坐标为 ij，$c \in \{0,255\}$ 是缩减调色板中的颜色，[43]，而 T 是温度参数，参见 [44]。$T=0$ 对应于选择缩减调色板中最接近的颜色，而 $T \to \infty$ 则会产生均匀噪声。

优化这一流程需要校准 T，以最大化原始图像与缩减后图像之间的某种高级相似性度量，希望它能捕捉到比简单的像素间欧氏距离最小化。这里我们提出一种替代方法，即最大化一种信息度量——多尺度相关性（MSR），并将其与经典指标进行比较，即峰值信噪比（PSNR）[12] 和结构相似性指数（SSI）[13]。PSNR 与原始图像和映射图像之间的均方误差（MSE）直接相关：$\text{PSNR} = 10 \log_{10}(\Delta^2/\text{MSE})$，其中 Δ 为信号范围；对于典型的灰度编码，其值为 255。SSI 基于对两幅图像中图像块的比较，并考虑亮度和对比度等属性。这两种指标在数字图像处理领域都得到了广泛应用。

图 10(a)展示了从图 6(a)中提取的原始图像块。图 10(b)展示了各项度量指标随温度 T 的变化。可以看到，原始图像与映射图像之间的峰值信噪比达到最大值

在 $T = 0$ 时。这并不令人意外，因为根据定义，峰值信噪比与均方误差呈单调关系。图 10(c) 中相应的映射结果显得过于锐利、对比度过高，将植被与天空清晰地区分开来，同时引入了阈值化伪影。对结构相似性指数的优化得到一个非零但较小的温度 $T = 0.1$，几乎没有改善最终图像（见图 10(d)）。接着，我们分别计算直接场和梯度场的多尺度相关性。将多尺度相关性（T）最大化后，得到图 10(e) 所示的图像。该图保留了更忠实的视觉特征，在大尺度上与原始图像具有较好的相似性，但代价是产生了人工的小尺度特征。最后，图 10(f) 所示的梯度幅值多尺度相关性 [45]，最大化结果，似乎在 (c)、(d) 和 (e) 之间取得了不错的折中：它呈现了中尺度特征（树干细节），同时没有模糊更精细的细节（小树枝）。

因此，在强烈颜色缩减的情况下，多尺度相关性方法比结构相似性指数等传统度量指标能带来更好的视觉效果；此外，后者还需要对原始图像具备先验语义知识。请注意，分析还可以扩展到更复杂的色彩映射过程，例如误差扩散 [46, 47]或基于蒙特卡洛的算法 [48]。

B. 使用Rudin-Osher-Fatemi 算法去噪

现在我们关注一种去噪过程，其目的在于校正由信号处理或相机伪影引起的不需要的噪声。解决这一问题的经典算法是 Rudin-Osher-Fatemi (ROF) [11]，它通过最小化以下泛函来实现：

$$\mathcal{L}[f] = \lambda \|\nabla f\|_2^2 + \|h - f\|_2^2 , \tag{5}$$

其中，h 是原始含噪图像，f 是目标去噪图像，λ 是正则化/惩罚项，用于防止梯度爆炸并获得平滑解。自由参数 λ 通常由操作者通过视觉评估来选取。

在此，我们再次使用 PSNR、SSI 和 MSR_∇ 度量指标来校准此类模型。我们考虑图 11(a) 中的图像，该图像是将高斯白噪声添加到图 6(a) 的图像块后得到的。我们特意选择较高的噪声值，使去噪过程更加困难，因此原始图像中的某些细节可能永远无法恢复。我们的目标是寻找能够带来最佳视觉效果的最优λ^*。各方法的得分随 λ 变化的情况如图 11(b) 所示。分别使用 PSNR、SSI 和 MSR_∇ 进行最优去噪后得到的图像见图 11(c)、(d) 和 (e)。随着

峰值信噪比图像仍然有较高水平的噪声，但树干表面或树枝上的细节得以保留。相比之下，结构相似性指数去除了大量噪声，但代价是小尺度细节变得模糊。虽然不像色彩映射处理那样明显，采用多尺度相关性进行最佳去噪$_\nabla$ 似乎是一个不错的

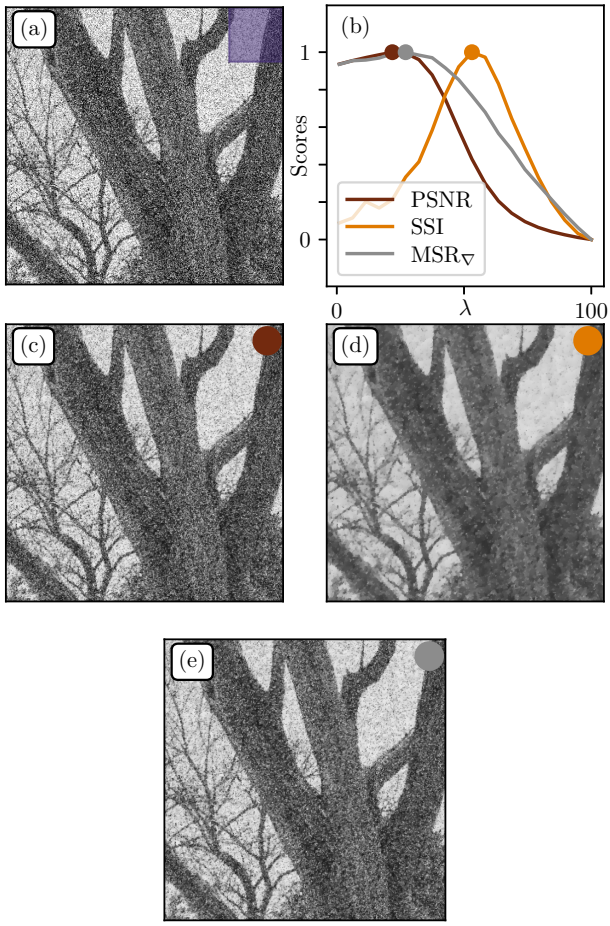

图 11. 去噪。(a) 向图 10(a) 中的同一图像块添加高斯噪声 ($\sigma = 100$) 后得到的含噪图像块。不同性能度量指标的重标度得分随 λ 变化的情况：峰值信噪比 (PSNR)、结构相似性指数 (SSI) 以及基于梯度场的 MSR，即 MSR_∇。(c)、(d)、(e) 分别为 PSNR、SSI 和 MSR_∇ 在最优正则化参数 λ^* 下的去噪结果。

折中方案，既避免了峰值信噪比图像噪声过多，也避免了结构相似性指数图像过度平滑。

V. 结论

让我们总结一下已取得的成果。我们首先通过一个简单的示例介绍了分辨率/相关性框架，并展示了这种形式体系如何应用于图像分析。为了在可控环境中研究该框架，我们从随机纹理入手。随后，我们定义了多尺度相关性（MSR），用于衡量所有压缩尺度上的熵贡献，并获得了让人联想到相关渗流问题的统计特征。具体而言，我们指出存在一个临界粗糙度参数 $\mathrm{H}_c \approx 0$，它对

应于对数相关性，并讨论了最优分割。接着，我们将分析扩展到自然图像，并将其与随机纹理成功地进行了比较；我们观察到它们与临界随机高斯场有很强的相似性。对梯度幅值场的考察则揭示了与粗糙度临界性更强的相似性。最后，在色彩映射和去噪这两项简单图像处理任务中，我们将 MSR 流程与经典信号处理度量进行了比较。所得结果颇具意义，展现了非特定的 MSR 方法在图像处理中的潜力。

本节若能扩展至更复杂的图像处理技术，将会更有裨益，但这超出了本文的范围。未来研究还应着重于在简单情形下，对相关性和分辨率进行解析上易于处理的拓展，例如采用精心选择的级联过程生成高斯白噪声。还应注意，我们考虑的是在直接空间中进行简单的压缩过程，但也可以使用等效表征（例如离散余弦 [9]或小波谐波 [10],）来定义缩减样本 $\mathcal{S}$。最后，我们发现，多尺度相关性能够捕捉最相关的分割值，这些值可用作学习框架的预处理方法。

VI. 致谢

我们衷心感谢 Jean-Philippe Bouchaud 富有成效的讨论。本研究是在经济物理与复杂系统研究讲席的框架下开展的，该讲席由 Fondation du Risque 和 Fondation de l'Ecole polytechnique、巴黎综合理工学院和 Capital Fund Management 共同支持。S.L. 感谢定量生命科学组接待他，使他得以在意大利的里雅斯特国际理论物理中心（ICTP）访问。

附录 A： $1/f^\alpha$ 高斯场的生成

A 节中使用的 $1/f^\alpha$ 纹理通过滤波程序生成。考虑一个具有给定自相关函数 $C(\boldsymbol{r})$ 的场 $\phi(\boldsymbol{r})$。根据卷积定理，在傅里叶空间中有 $\tilde{C}(\boldsymbol{q}) = |\tilde{\phi}(\boldsymbol{q})|^2$，其中 $|\boldsymbol{q}| = f$。生成具有指定相关性的随机高斯场 ϕ 的一种自然方法是：

$$\phi = \mathcal{F}^{-1}\left(\sqrt{|\tilde{C}(\boldsymbol{q})|}\tilde{\eta}(\boldsymbol{q})\right), \tag{A1}$$

其中 $\eta(\boldsymbol{r})$是高斯白噪声。因此，纹理呈现周期边界条件。还要注意，$\mathbf{H} = 0$ （等价地， $\alpha = 1$）情形对应于如下形式的对数空间相关性（例如参见 [49]）：

$$C(\boldsymbol{r}) = -\lambda\left(\log\frac{r}{\xi} + \log 2 - \gamma\right), \tag{A2}$$

其中 $r = |\boldsymbol{r}|$, ξ^{-1} 是正则化常数，可将其解释为低频截止；γ 是欧拉常数。

[1] R. J. Cubero、M. Marsili 和 Y. Roudi，《计算神经科学杂志》**48**，85（2020）。[2] A. Krizhevsky、I. Sutskever 和 G. E. Hinton，《ACM 通讯》**60**，84（2017）。[3] Y. LeCun、L. Bottou、Y. Bengio 和 P. Haffner，《IEEE 会议录》**86**，2278（1998）。[4] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、 L. Kaiser 和 I. Polosukhin，《神经信息处理系统进展》**30**（2017）。[5] A. Krizhevsky 和 G. Hinton，多伦多大学计算机科学系硕士论文（2009）。[6] J. Deng、W. Dong、R. Socher、L.-J. Li、K. Li 和 L. Fei-Fei，《IEEE 计算机视觉与模式识别会议》，248（2009）。[7] A. M. Saxe、J. L. McClelland 和 S. Ganguli，arXiv 预印本 arXiv:1312.6120（2013）。[8] S. Goldt、M. Advani、A. M. Saxe、F. Krzakala 和 L. Zdeborov´a，《神经信息处理系统进展》**32**（2019）。[9] G. K. Wallace，《IEEE 消费电子学报》**38**，xviii（1992）。[10] A. Skodras、C. Christopoulos 和 T. Ebrahimi，IEEE 信号处理杂志 **18**，36（2001）。[11] L. I. Rudin、S. Osher 和 E. Fatemi，《物理学 D：非线性现象》**60**，259（1992）。[12] J. Korhonen 和 J. You，第四届多媒体体验质量国际研讨会，37（2012）。[13]Z. Wang、A. C. Bovik、H. R. Sheikh 和 E. P. Simoncelli，IEEE 图像处理汇刊 **13**，600（2004）。[14] M. Marsili 和 Y. Roudi，《物理学报告》**963**，1（2022）。[15] A. Haimovici 和 M. Marsili，《统计力学杂志：理论与实验》**2015**，P10013（2015）。[16] J. Song、M. Marsili 和 J. Jo，《统计力学杂志：理论与实验》**2018**，123406（2018）。[17] O. Duranthon、M. Marsili 和 R. Xie，《统计力学杂志：理论与实验》**2021**，033409（2021）。[18] R. M. Balboa 和 N. M. Grzywacz，《视觉研究》**43**，2527（2003）。[19] D. L. Ruderman，《神经系统计算》**5**，517（1994）。

[20] D. L. Ruderman 和 W. Bialek，《物理评论快报》**73**，814（1994）。[21] D. Zoran 和 Y. Weiss，IEEE 第十二届计算机视觉国际会议，2209（2009）。[22] G. J. Stephens、T. Mora、G. Tkaˇcik 和 W. Bialek，《物理评论快报》**110**，018701（2013）。[23] S. Lakhal、A. Darmon、J.-P. Bouchaud 和 M. Benzaguen，《物理评论研究》**2**，022058（2020）。[24] S. Aaronson、S. M. Carroll 和 L. Ouellette，arXiv 预印本 arXiv:1405.6903（2014）。[25] Y.-C. Zhang，《物理学杂志 I》**1**，971（1991）。[26] A. Humeau-Heurtier，《熵》**22**（2020）。[27]M. B. Isichenko，《现代物理评论》**64**，961（1992）。[28] S. Prakash、S. Havlin、M. Schwartz 和 H. E. Stanley，《物理评论 A》**46**，R1724（1992）。[29]C. Du、C. Satik 和 Y. Yortsos，《美国化学工程师学会会刊》**42**（1996）。[30] A. Van der Schaaf 和 J. van Hateren，《视觉研究》**36**，2759（1996）。[31] G. Tkaˇcik、P. Garrigan、C. Ratliff、G. Milˇcinski、J. M. Klein、L. H. Seyfarth、P. Sterling、D. H. Brainard 和 V. Balasubramanian，《公共科学图书馆·综合》**6**，e20409（2011）。[32] T. Peli 和 D. Malah，《计算机图形与图像处理》**20**，1（1982）。[33] M. Andreux、T. Angles 等，《机器学习研究杂志》**21**，2256（2020）。[34] M. S. Keil、G. Cristobal 和 H. Neumann，《视觉研究》**46**，2659（2006）。[35] M. S. Keil，——《视觉研究》**47**，3360（2007）。[36] M. Kilpel¨ainen 和 M. A. Georgeson，《科学报告》**8**，1593（2018）。[37] R. Morel、G. Rochette、R. Leonarduzzi、J.-P. Bouchaud 和 S. Mallat，arXiv 预印本 arXiv:2204.10177（2022）。[38] S. G. Mallat，《IEEE声学、语音与信号处理汇刊》**37**，2091（1989）。[39] J.-P. Antoine、P. Carrette、R. Murenzi 和 B. Piette，《信号处理》**31**，241（1993）。[40] P. Abry、S. Jaffard、S. Roux、B. Vedel 和 H. Wendt，未爆弹药探测与处置，1（2009）。[41] H. Wendt、P. Abry、S. G. Roux、S. Jaffard 和 B. Vedel，《信号处理》**26**，47（2009）。[42] B. Duplantier、R. Rhodes、S. Sheffield 和 V. Vargas，《几何、分析与概率：献给 Jean-Michel Bismut 的纪念文集》，191（2017）。[43]该概率密度由最大熵分布得到，该分布与最小化原始图像和映射图像之间的均方误差（均方误差）有关。[44] S. Lakhal、A. Darmon 和 M. Benzaquen，《统计力学杂志：理论与实验》**2023**，033401（2023）。[45] 请注意，要计算梯度幅值的多尺度相关性（MSR），必须对梯度处理过程得到的灰度图像进行分割，并对 a求平均。[46] J. F. Jarvis、C. N. Judice 和 W. Ninke，《计算机图形学与图像处理》**5**，13（1976）。[47] R. W. Floyd 和 L. Steinberg，信息显示学会会议录 **17**（1976）。[48] J. Puzicha、M. Held、J. Ketterer、J. M. Buhmann 和 D. W. Fellner，IEEE 图像处理汇刊 **9**，666（2000）。[49] M. Peskin，《量子场论导论》（CRC出版社，2018）。