



## 自然图像统计建模的进展

A. SRIVASTAVA\*

A.B. LEE†

E.P. SIMONCELLI‡

S.-C. ZHU§

### 摘要

图像的统计分析揭示了两个有趣的特性：(i) 图像统计对图像缩放的不变性，以及 (ii) 图像统计的非高斯行为，即高峰度、重尾和尖锐中心尖点。在本文中，我们回顾了自然图像统计建模的一些最新结果，这些结果试图解释这些模式。考虑了两类结果：(i) 图像或图像分解（如傅里叶或小波分解）的概率模型研究，以及 (ii) 在限制于自然图像的情况下发现底层图像流形。还考虑了这些模型在纹理分析、图像分类、压缩和去噪等领域的应用。

*Keywords:* 自然图像统计, 非高斯模型, 尺度不变性, 统计图像分析, 图像流形, 广义拉普拉斯算子, 贝塞尔 K 形式.

## 1. 简介

处理静态或动态图像的应用程序（电影）近年来变得越来越重要和流行。图像分析、压缩、去噪、传输和理解工具已广泛应用于许多科学和商业活动。例如，图像压缩在传输或存储中很重要，而从相机图像中自动识别人物对于安全目的也很重要。成像应用工具的开发始于图像的数学表示。许多此类应用程序需要“学习”一些图像模式或趋势，然后才能分析特定数据。自然图像统计领域源于观察、分离和解释自然图像所呈现的模式而努力。最近，人们更加重视图像的显式概率模型。其中一个原因可能是人们越来越欣赏图像所呈现的可变性，并意识到精确的数学/物理模型可能不切实际，需要采用统计方法。

任何统计方法都需要概率模型来捕捉“基本”图像变化，并且易于计算。研究图像统计的另一个动机是了解感官编码作为动物视觉中信息存储/处理的一种策略。对图像统计的理解可能会为动物视觉系统的架构提供线索 [79]。

有几种途径可以理解图像统计。尽管图像被表示为大向量空间（例如，正数的矩形阵列空间）的元素，这里称为图像空间，但有趣图像的子集相当小且受限。因此，一种途径是隔离这个子集（这里称为图像流形），并在其上施加一个简单的概率模型，例如均匀或高斯模型。给定图像流形及其上的概率分布，即可得到用于成像应用的统计工具。另一条路径是推导在更大的图像空间上定义的概率模型，但将任何显著的概率仅放在图像流形上。我们将本文分为这两类。

在寻找图像的统计描述时，数学和物理思想并没有被抛弃，而是密切相关。例如，图像强度的谐波分析用于将图像分解为单个组件，这些组件比原始图像更适合于模型构建。由于图像空间的维度相当高，而常见的密度估计技术主要适用于小尺寸，因此通常首先将图像分解为其组件，然后分别估计这些组件的低阶统计数据。此外，成像是物理过程，物理考虑通常有助于以某种形式寻找统计描述符。物理模型是许多研究的动机，这些研究导致了有趣的图像统计特征。

本文的布局如下：第 2 节首先介绍图像光谱分析的历史视角以及随后的一些发现。一项重要成就是发现图像统计中的非高斯性。在接下来的两节中，我们将介绍一些近期结果，分为两类：第 3 节研究分解为光谱成分的图像的概率模型，而第 4 节研究发现/近似自然图像的图像流形的方法。第 5 节概述了这些方法的一些应用，第 6 节讨论了一些未解决的问题。

## 2. 图像分解及其统计特性

为了进行统计建模，图像被视为  $\mathbb{R}^2$  中某个域上定义的空间随机过程的实现。假设域是（连续的）矩形区域或有限的均匀网格。图像建模中的一个常见假设是底层图像处理是静止的，即图像概率对图像平面中的平移不变。

### 2.1. 经典图像分析

如果将图像处理建模为二阶空间过程，则谱分析将成为一种自然工具。将两个任意像素值的协方差函数定义为  $C(x)$ ，其中  $x$  是两个像素位置之间的差值，则可以将功率谱定义为  $P(w) = \int_{\mathbb{R}^2} C(x)e^{-jwx}dx$ ，其中  $w$  表示二维空间频率。由于傅里叶基也是循环矩阵的特征基，在平稳性假设下  $C(x)$  就是这种情况，因此傅里叶分析也与流行的主成分分析相一致（稍后在第 4 节中讨论）。此外，傅里叶表示保证系数不相关。然后，图像由其傅里叶系数表示，图像统计数据通过系数统计进行研究，并且可以通过在傅里叶系数上施加随机结构来继承图像上的随机模型。系数均值和协方差的指定完全指定了二阶图像过程。空间功率谱的早期研究表明，功率  $P(w)$  衰减为  $\frac{A}{|w|^{2-\eta}}$ ，其中  $|w|$  是空间频率的幅度。这一特性称为图像的幂律谱，最早由电视工程师在 50 年代 [25, 49] 观察到，80 年代后期由 Field [30] 和 Burton 和 Moorhead [17] 在自然图像中发现。如 [61] 中所述， $\eta$  的值随图像类型而变化，但通常是一个较小的数字。

虽然傅里叶分析是经典方法的核心，但出于各种原因，其他基函数也广受欢迎。例如，为了捕捉图像中物体的局部性，使用小波基分解图像已成为一种有吸引力的工具。特别是，通常使用 Gabor 小波 [31] 用于同时在空间和频率上分解观察到的图像。此外，Marr [56] 建议使用高斯拉普拉斯滤波器来模拟早期视觉。如果将图像视为  $\mathbb{R}^2$  中有限均匀网格上的实现，图像空间变为有限维，并且可以将图像线性投影到在不同标准下最优的低维子空间中。第 4 节介绍了其中一些线性投影。再次通过投影系数的统计来研究图像统计。

## 2.2. 图像统计的尺度不变性

与幂律密切相关的一项发现是，当图像按比例放大或缩小时，图像统计数据的不变性 [16, 30]。换句话说，如果对图像进行缩放，自然图像的统计量边际分布保持不变。幂律是图像分形或尺度不变性的一种表现。通过研究多种尺度下像素对比度 ( $\log(I(x)/I_0)$ ) 的直方图，Ruderman 和 Bialek [71] 证明了其对缩放的不变性。Zhu 和 Mumford [98] 独立地通过研究图像小波分解的直方图证明了更广泛的不变性。Ruderman [69, 70] 还提供了自然图像尺度不变性的证据，并提出了一个解释它们的物理模型。Turiel 和 Parga [84] 研究了自然图像的多重分形结构，并将其与尺度不变性联系起来。Huang [43] 研究了不同类型场景的缩放。在 [85] 中，Turiel 等人表明，高光谱图像（彩色图像）也表现出多尺度特性，并且统计数据对尺度不变。Geman 和 Koloydenko [32] 研究了小窗口中像素值的顺序统计的缩放。

除了像素统计数据之外，Alvarez 等人 [1] 还展示了通过图像的形态学操作获得的拓扑统计数据的缩放。

必须强调的是，只有大型图像集合的统计数据才是尺度不变的；单个图像的统计数据会随尺度而变化。寻求图像集合概率描述的理论模型旨在实现尺度不变性，而处理单个图像的应用驱动模型旨在捕捉单个图像的变化性。

## 2.3. 边际统计的非高斯性

经典方法假设图像是二阶过程，但观察结果并不支持这一假设。研究人员发现，自然图像的高阶统计量表现出有趣的模式，接下来他们关注的是这些高阶矩。其中一个含义是，图像统计量不遵循高斯分布，需要高阶统计量。例如，一种流行的局部分解图像的机制——在空间和频率上使用小波变换会导致系数非常非高斯，即小波系数的直方图显示出重尾、中位数的尖锐尖角和不同尺度之间的大相关性。据我们所知，Field [30] 是最早强调小波滤波器响应高度峰态形状的人。Mallat [55] 指出，图像的多尺度、正交小波分解的系数可以用广义拉普拉斯密度来描述（稍后在第 3.2 节中给出）。Ruderman [69]、Simoncelli 和 Adelson [76] 也研究和建模了图像的这种非高斯行为，Moulin 和 Liu [60]，以及 Wainwright 和 Simoncelli [88]。Thomson [82] 最近利用相位二阶谱（四阶统计数据）研究了自然图像的统计数据，并证明了该统计数据的幂律和尺度不变性。

事实上，投影到任何局部零均值线性核上似乎都会产生峰度响应 [95]。

Huang [43] 表明，当用  $8 \times 8$  随机均值为 0 的滤波器对图像进行滤波时，图像的分布具有高峰度、零点处的尖锐尖角和长指数尾部。这表明线性分解可以最大化峰度或非高斯性的其他度量。这些努力 [6, 62, 86] 产生了空间定向的基，其（空间）频率带宽大约为一个倍频程，类似于许多多尺度分解。在线性分解下，通过最小化系数的独立性获得了类似的结果，从而产生了独立分量分析 [18, 23, 44]。这些观察结果证明了正交小波在一般图像分析应用中的广泛使用。

使用 Gabor 小波的另一个动机是，人们发现动物视觉皮层中简单细胞的接受域与 Gabor 函数相似 [58]。但请注意，大多数图像的小波分解都是基于一维滤波器的可分离应用，这会导致无方向（混合对角线）子带。提供更好的方向分解和更高峰度响应的替代表示包括 [24, 27, 78, 90]。

除了边际统计的非高斯行为外，许多作者还研究了滤波器响应的联合统计。特别是，大多数图像中存在的局部结构导致局部线性算子的响应具有依赖性。对小波系数联合直方图的研究表明，它们在

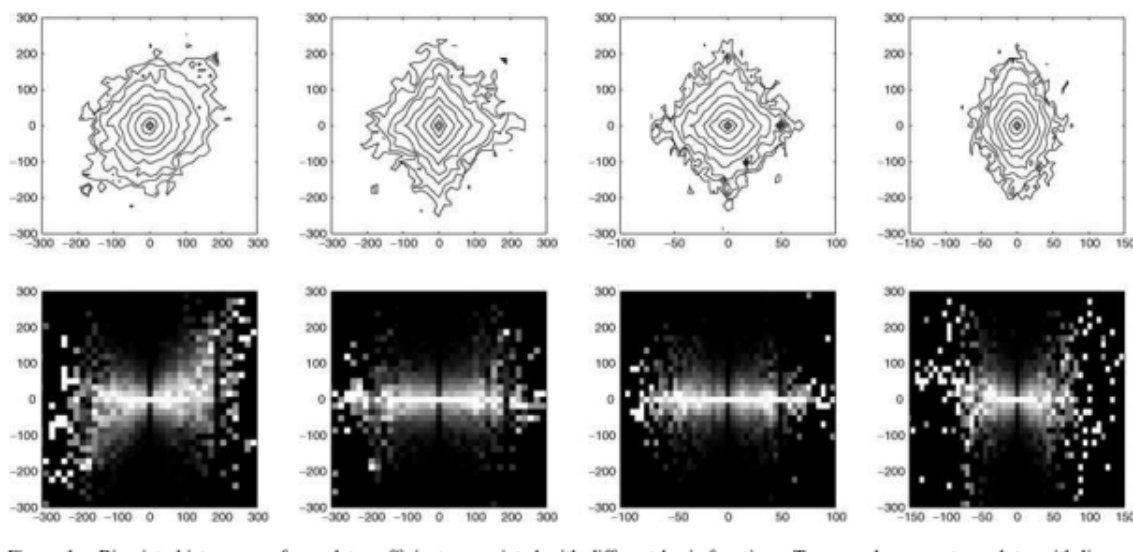


图 1: 与不同基函数相关的小波系数的二维直方图。最上面一行显示的是等值线图，线条与对数概率间隔相等。左侧两种情况是不同的空间偏移（相同的比例和方向），第三种是不同的方向（相同的比例和几乎相同的位置），最右侧对应的是相邻比例的一对（相同的方向和几乎相同的位置）。最下面一行显示了一些条件分布：亮度越高，频率越高。

尺度、方向和位置上具有依赖性。Zetzsche 及其同事指出，投影到偶对称和奇对称 Gabor 滤波器上的联合密度具有圆对称性。如果密度是高斯的，这将对应用于独立（不相关）的边缘，但高度峰态的边缘意味着响应具有强烈的依赖性 [91]。Shapiro 开发了一种启发式方法，利用小波系数之间的联合依赖性，彻底改变了图像压缩领域 [73]。

Simoncelli 及其同事研究并模拟了带通滤波器对响应之间的依赖关系，发现即使有符号响应不相关，振幅也具有很强的相关性 [14, 74]。图 1 对此进行了说明，其中显示了多对系数的条件直方图。请注意，与二阶相关性不同，这些依赖关系不能通过线性变换消除。

图像块的观测统计数据可以用多个滤波器下的系数联合直方图来描述。对于联合高斯系数，联合直方图中的等概率轮廓都是椭圆形的，但在自然图像的情况下，二维和三维轮廓表面呈现出明显的多面体形状。黄 [43] 表明，轮廓中的峰值和尖点对应于具有部分恒定强度区域和尖锐不连续性的图像中出现的简单几何图形。这些结果表明自然图像中的“类似物体”结构，并强调基于物体的模型的重要性。Grenander 和 Srivastava [36] 也将非高斯性归因于图像中物体的存在，并将该想法用于模型构建。Lee 等人已经证明，从遮挡模型合成的图像（第 3.1 节）在直方图的轮廓表面中显示出不规则的多面体形状。

### 3. 图像空间上的新兴统计模型

最早且至今仍广泛使用的图像概率模型是基于马尔可夫随机场模型（MRF）[93]。图像场是一组随机变量，表示图像平面中均匀分布网格上的像素值。在 MRF 中，给定剩余图像的像素值的条件概率

降低为给定该像素邻域的条件概率。如果像素的邻域很小，则会产生效率，此外，如果所有像素都使用相同的条件密度，则平稳性成立。Ising 和 Potts 模型是该家族中最简单的例子。Besag [8, 9] 将图像像素的联合密度表示为条件密度的乘积，并忽略了规范化器以获得伪似然公式。Clifford-Hammersely 定理，例如 [93]，指出完整条件完全指定联合密度函数（在正性假设下），并允许使用 Gibbs 采样器分析图像。Geman 和 Geman [33] 利用 MRF 和 Gibbs 分布的等效性从这些分布中进行采样。Kersten [47] 致力于在 MRF 框架中计算像素值的条件熵，给定相邻像素。Zhu 和 Mumford [98] 在图像上使用 Gibbs 模型，并使用极小极大熵标准估计模型参数。设  $H(I)$  为使用多个小波基获得的图像系数直方图的串联。则最大熵概率采用以下形式： $P(I | \lambda) \propto e^{-(\lambda, H(I))}$ 。通过将平均直方图设置为等于观察到的直方图来估计向量  $\lambda$ ，即  $\int H(I)P(I | \lambda)dI = H_{obs}$ ，并简化为  $P(I_{obs} | \lambda)$  下  $\lambda$  的最大似然估计。

### 3.1. 受物理学启发的模型

许多研究人员从物理角度研究了图像统计，试图捕捉最初生成图像的现象。所有这些模型的共同主题是根据泊松过程在图像中随机放置平面形状（线条、模板、物体等）。不同的模型在形状选择（例如原始与高级）和相互作用（例如一些模型倾向于遮挡而另一些模型倾向于叠加）方面有所不同。我们在这里总结了其中一些模型：

**3.1.1. 叠加模型。**为了捕捉图像的尺度不变性和非高斯性，Mumford 和 Gidas [61] 使用了一系列无限可分分布。他们表明，当图像被建模为物体随机放置的叠加时，就会出现这种分布。为了实现自相似性，有人提出，图像中存在的物体大小应根据密度函数  $Zr^{-3}$  在  $\mathbb{R}_+$  子集上的分布（ $Z$  是归一化常数）。使用此模型，Chi [19] 描述了物体的泊松放置，其大小根据  $1/r^3$  定律进行采样。此外，他假设了一种表面过程，该过程模拟了物体二维表面上的纹理变化。Bitouk 等人 [11] 和 Grenander 和 Srivastava [36] 也假设图像由物体的二维外观组成，这些物体位于均匀的泊松点  $z_i$  处。这些对象是从对象空间（任意大小、形状或纹理）中随机选择的。使用加权叠加（权重（或对比度）由独立标准法线  $a_i$  给出）来形成图像；它允许将线性滤波图像表示为滤波对象的叠加。此公式得出单个图像的分析（参数）概率，最终形式在第 3.2 节中给出。

**3.1.2. 遮挡模型。**在更复杂的模型中，物体像以前一样随机放置，但现在前面的物体遮挡了后面的物体。一个例子是枯叶（也称为随机拼贴）模型，该模型假设图像是相互遮挡的大致独立物体的拼贴。通过将一组有序的基本 2D 形状  $g_i$  分层放置，可以从模型生成图像。

位置  $z_i$  的大小为  $r_i$ ，这些位置是从泊松过程中采样的。对于灰度图像，随机集或“叶子” $T_i = g_i(z - z_i)$  还与强度  $a_i$  相关。叶子通常按照以下方式前后放置 1

$$I^{(i)}(z) = \begin{cases} a_i & \text{if } z \notin T_j, z \in T_i, \forall j < i \\ I^{(i-1)}(z) & \text{否则} \end{cases} \quad \text{where } i = 1, 2, \dots$$

图 2 左侧两个面板分别显示了具有椭圆形和正方形随机形状的两个枯叶模型样本。最终模型在极限  $\lim_{i \rightarrow \infty} I^{(i)}(z)$  中定义，或者等效地，当有限图像域  $\Omega \subset \mathbb{R}^2$  被完全覆盖。枯叶模型可以追溯到数学形态学中的 Matheron [57] 和 Serra [72]。他们证明了任何紧集  $K \subset \dot{\Omega}$  属于同一片叶子且不被其他叶子遮挡的概率等于比率  $\frac{E[v(X_0 \ominus \dot{K})]}{E[v(X_0 \oplus \dot{K})]}$ 。这里， $X_0$  是模型中使用的随机形状， $E[v(\cdot)]$  是预期的 Lebesgue

测度,  $X_0 \ominus \tilde{K} = \{x : x + K \in X_0\}$  和  $X_0 \oplus \tilde{K} = \{x : X_0 \cap (x + K) \neq \emptyset\}$  分别表示  $X_0$  被  $K$  侵蚀和膨胀。该模型的早期应用包括研究非均质材料和粉末中颗粒的尺寸分布 [45]。最近, 研究人员将不同版本的枯叶模型应用于自然

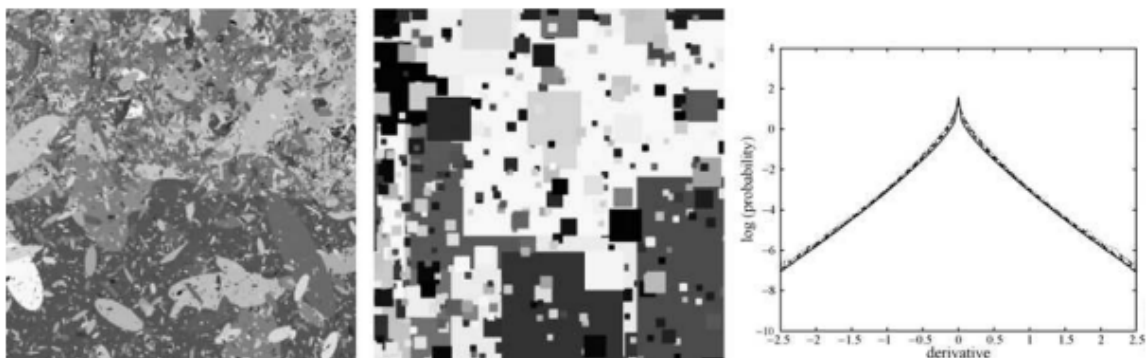


图 2: 左侧两张图片: 两种不同版本的枯叶模型, 随机形状分别为椭圆形 (左图) 和方形 (中图)。右图: 枯叶模型的衍射统计量和缩放, 圆盘分布符合大小立方定律。曲线对应于四个不同尺度 ( $N = 1, 2, 4, 8$ ) 对比度归一化后相邻像素间差异的边际分布。

场景。在 [70] 中, Ruderman 提出, 具有统计上独立的“对象”的随机拼贴描述可以解释图像中相关性的缩放。使用遮挡对象模型和手工分割图像的差异函数, 他推断出两个点属于同一对象的概率的幂律。该结果与 [1] 使用水平集和 Matheron 和 Serra 的原始枯叶形式一致。在这里, Alvarez 等人分析了自然图像的形态结构——特别是遮挡对均匀连通图像区域的面积、周长和截距长度的影响。

Lee 和 Mumford [50] 使用具有统计独立对象的枯叶模型和显著泊松过程的测度, 证明了当考虑遮挡时, 对象大小的  $1/r^3$  幂律也会导致近似全尺度不变性。假设合成图像通过例如块平均进行平滑处理, 其统计数据将与在校准的自然图像的大型数据集中观察到的统计数据相匹配。本文研究了滤波器响应直方图的尺度不变性、两个像素的完全共现统计和 Haar 小波系数的联合统计。图 2 (右图) 显示了带平滑处理的枯叶模型中在缩放下的导数统计直方图。

此外, 一些其他科学领域的类比理论也被用来解释自然图像的自相似性。例如, Ruderman 等人 [69, 71] 利用了湍流的统计数据。在 [83] 中, Turiel 等人, 将湍流的统计建模与图像的统计建模正式联系起来。在湍流中, 如果能量耗散的二阶矩与某个  $\tau$  的  $r^\tau$  成比例, 其中  $r$  是盒子大小, 则自相似性成立。作者使用基于边缘的度量来计算图像中的能量耗散, 展示了边缘统计的自相似性。

### 3.2. 分析密度: 单变量情况

统计模型的一个重要前提是它们在实时或近实时应用中具有计算效率, 而实现这一点的一个明显方法是通过参数密度。我们希望通过参数密度来表示图像的统计性质, 只使用有限的参数。在本节中, 我们将讨论参数密度系列, 它们似乎可以捕捉图像低维表示中的可变性。这里的目标不是解释大量图像的尺度不变性, 而是捕捉单个图像的可变性。对于图像  $I$  和线性滤波器  $F$ , 本节将描述  $I * F$  中像素值的边际密度。

1. 广义拉普拉斯模型：图像系数的边缘密度可以通过广义拉普拉斯密度（也称为广义高斯或拉伸指数） $f_1(x; c, p) = \frac{e^{-1x/dp}}{Z_1(p, c)}$  很好地建模，[55, 60, 76]，其中规范化器为  $Z_1(p, c) = 2 \frac{c}{\rho} \Gamma(\frac{1}{\rho})$ 。可以使用最大似然法或矩量为特定图像的子带估计参数  $p, c$ 。

另一种估算方法是通过（线性） $\log(\log(h(x) + h(-x)) - 2\log(h(0)))$  与  $\log(|x|)$  的回归，其中  $h(x)$  是  $x$  处的直方图值， $x$  是箱中心的变量。指数  $p$  的值通常在  $[0.5, 0.8]$  范围内，宽度参数  $c$  随着基函数的大小单调变化，对于较粗尺度分量，产生更高的方差 [76]。

根据 Grenander 的叠加模型公式（第 3.1 节第 1 项），如果假设随机变量  $u(z) \equiv \sum_{i=1}^n ((g_i * F)(z - z_i))^2$  具有缩放 Gamma 密度，则滤波像素的单变量密度在 [36] 中显示为：对于  $p > 0, c > 0, f_2(x; p, c) = \frac{1}{Z_2(p, c)} |x|^{p-0.5} K_{(p-0.5)}(\sqrt{\frac{2}{c}}|x|)$ ，其中  $K$  是第三类修正贝塞尔函数， $Z_2$  是归一化常数。（其特征函数形式为  $1/(1 + 0.5c\omega^2)^p$ ，当  $p = 1$  时，该函数成为拉普拉斯密度的特征函数。）这种参数密度族被称为 *BesselK* 形式，其中  $(p, c)$  称为 *Bessel* 参数。此前，Wainwright 等人 [89] 还研究了高斯尺度混合族在图像建模中的应用，并将  $f_2$  称为对称 Gamma 密度（未达到其解析参数形式）。如 [36] 所述， $p$  和  $c$  很容易从观测数据中估算出来，即  $\hat{p} = \frac{3}{SK(I^{(j)})-3}$  和  $\hat{c} = \frac{SV(I^{(j)})}{\hat{p}}$ ，其中  $SK$  是样本峰度， $SV$  是滤波图像的样本方差。该模型的一个明显优势是，参数可以在数学上与图像中存在的对象的物理特性相关联。如果将滤波器  $F$  应用于图像  $I$  以提取某些特定特征（例如垂直边缘），那么 [80] 中已显示得到的  $p$  取决于两个因素：(i) 独特性和 (ii) 该特征在  $I$  中出现的频率。边缘更清晰、独特的物体的  $p$  值较低，而具有许多物体的场景的  $p$  值较大。

2. 高斯尺度混合：滤波图像的观测统计数据表明，存在大量高度峰态和重尾的单变量密度。在统计学文献中，变量被广泛定义为正态方差-均值混合；如果给定  $u$  的  $X$  的条件密度函数为正态函数，均值为  $\mu + u\beta$ ，方差为  $u\Delta$ ，则  $X$  称为正态方差-均值混合，并且  $u$  称为混合变量。Barndorff-Nielsen [3] 引入了广义双曲分布，作为特定的正态方差-均值混合，当混合变量  $u$  属于某个类别时产生。对于  $\mu = \beta = 0$  和  $\Delta = 1$ ，所得的族也称为高斯尺度混合 [2]，并已应用于金融数学 [12] 和语音处理 [13]。此外，如果  $u$  是缩放的 Gamma 密度，则会产生 Bessel K 密度。

图 3 显示了估算该密度函数的一些示例：使用任意 Gabor 滤波器（未显示）对来自 van Hateren 数据库的自然图像进行过滤，并使用得到的像素值形成直方图  $h(x)$ 。该图的底部面板显示了  $\log(h(x))$ 、 $\log(f_1(x))$  和  $\log(f_2(x))$  的图，其参数是根据相应图像估算的。

有几种方法可以判断任何提出的概率模型的性能。最简单的方法是使用概率分布空间上的任何度量，将观察到的频率与模型预测的频率进行比较。两个常用的度量是 Kolmogorov-Smirnov 距离： $\max_{x \in \mathbb{R}} |\int_{-\infty}^x (f_1(y) - f_2(y)) dy|$  和 KullbackLeibler 散度： $\int_{\mathbb{R}} f_1(x) \log(\frac{f_1(x)}{f_2(x)}) dx$ 。在这些度量下，广义拉普拉斯算子和贝塞尔 K 形式在捕获单变量图像统计数据方面都表现良好。一般来说，对于较小的  $p$  值（特征更清晰，图像中的对象数量更少），广义拉普拉斯算子表现更好，而对于较大的  $p$ ，贝塞尔 K 密度与观察到的密度更匹配。

### 3.3. 双变量概率模型

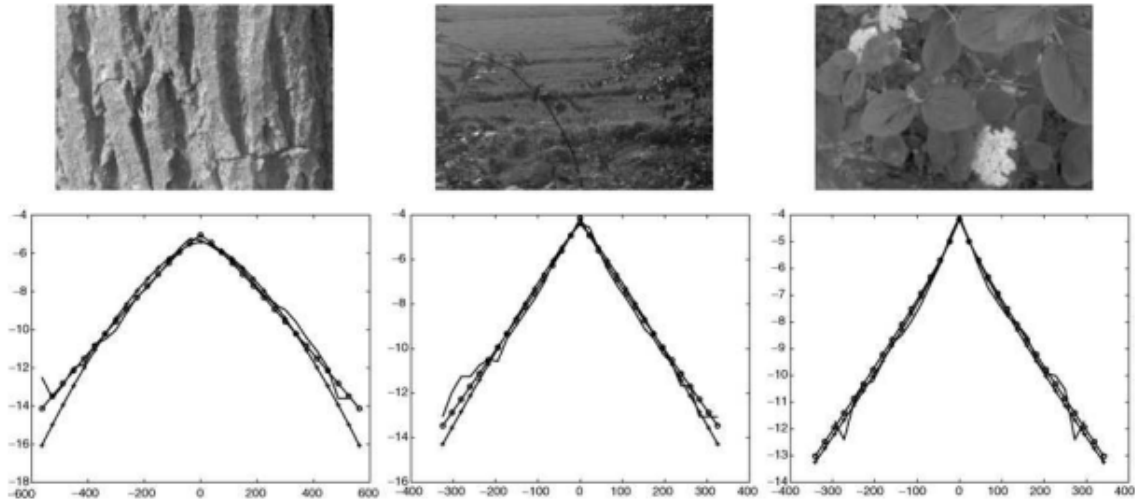


Figure 3. Estimated Bessel  $K$  and generalized Laplace densities compared to the observed histograms for the image in top. Plain lines are

图 3: 上图中贝塞尔  $K$  和广义拉普拉斯密度的估计值与观察到的直方图的比较。普通线条为直方图, 大珠子线条为贝塞尔  $K$ , 小珠子线条为广义拉普拉斯。这些密度以对数标度绘制。

到目前为止, 我们只讨论了单变量模型, 但观察到的图像的复杂性表明统计相互作用具有高阶性。作为第一个扩展, 我们研究用于捕获图像表示之间的成对相互作用的模型, 即双变量概率密度。例如, 这些模型可能涉及不同尺度、位置或方向的一对小波系数。

**3.3.1. 方差依赖性和高斯尺度混合。**如第 2.4 节所述, 即使响应 (二阶) 不相关, 与邻近位置、方向或尺度的基函数相对应的小波系数对也不是独立的。具体而言, 任何给定系数的条件方差都很大程度上取决于周围系数的值 [74]。当考虑更远的系数 (无论是在空间位置、方向还是尺度上) 时, 依赖性会变得更弱。

这种依赖性似乎无处不在, 并且可以在各种各样的图像类型中观察到。

哪种概率模型可以解释方差依赖性的观察结果? 一种候选模型是前面针对单变量情况描述的高斯尺度混合模型。在该方案中, 小波系数被建模为高斯随机变量和控制方差的隐藏“乘数”随机变量 (与前面的混合变量  $u$  相同) 的乘积。为了解释图 1 所示的成对统计数据, 可以规定相邻系数的隐藏乘数变量之间的关系。也就是说, 隐藏变量现在相互依赖, 产生数据中看到的方差缩放。这种依赖性的建模仍然是一个持续的研究课题。一种可能性是将这些变量链接到马尔可夫树中, 其中每个乘数在以其父项和子项为条件时都独立于所有其他乘数 [89]。

**3.3.2. 贝塞尔  $K$  形式的双变量扩展。**Grenander 扩展了单变量贝塞尔  $K$  形式, 以指定滤波图像的双变量密度。基本思想是通过贝塞尔  $K$  形式对滤波像素的所有线性组合进行建模, 然后调用 Cramer-Wold 装置, 请参阅 [10] 中的定义, 该定义指出, 指定多个随机变量的所有线性组合的密度唯一地指定了它们的联合密度。这里使用 Cramer-Wold 装置假设存在一个二维分布, 其 Radon 类边界 (对于半空间, 而不是在线上) 的行为符合贝塞尔  $K$  密度。在边界上施加贝塞尔  $K$  形式似乎在定性上与数据非常吻合, 但其性能仍有待在大型数据集上量化。

对于两个小波滤波器  $F^{(1)}$  和  $F^{(2)}$ , 设  $I_1 = I * F^{(1)}$  和  $I_2 = I * F^{(2)}$ , 对于  $a_1, a_2 \in \mathbb{R}$ , 设



$J(a_1, a_2, z) \equiv a_1 I_1(z) + a_2 I_2(z)$ 。Cramer-Wold 的思路是使用贝塞尔参数估计，计算所有对  $(a_1, a_2)$  的  $J(a_1, a_2, z)$  的特征函数，然后进行逆傅里叶变换，以获得联合密度  $f(I_1(z), I_2(z))$  的估计值。

该估计由八个联合矩参数化： $\mu_{0,2}, \mu_{1,1}, \mu_{2,0}, \mu_{4,0}, \mu_{3,1}, \mu_{2,2}, \mu_{1,3}, \mu_{0,4}$ ，其中  $\mu_{i,j} = \int \int I_1^i I_2^j f(I_1, I_2) dI_1 dI_2$ 。图 4 显示了这种双变量密度估计的一个示例。对于左图所示的图像，考虑其在两个 Gabor 滤波器下以相同比例但方向相差 20 度的两个滤波版本的双变量密度。

上图和下图显示了估计密度的网格和轮廓图，下图显示了观察到的双变量直方图。密度均以对数刻度绘制。

### 4. 发现图像流形

众所周知，在正数的矩形阵列空间中，只有一小分子集包含自然场景的图像。人们试图分离和表征这个子集以用于图像分析应用。主要思想是将此集合识别为低维可微流形，并使用其几何形状来表征图像。定义此流形后，简单的概率模型可以帮助捕捉图像变化。我们现在给出一个总结

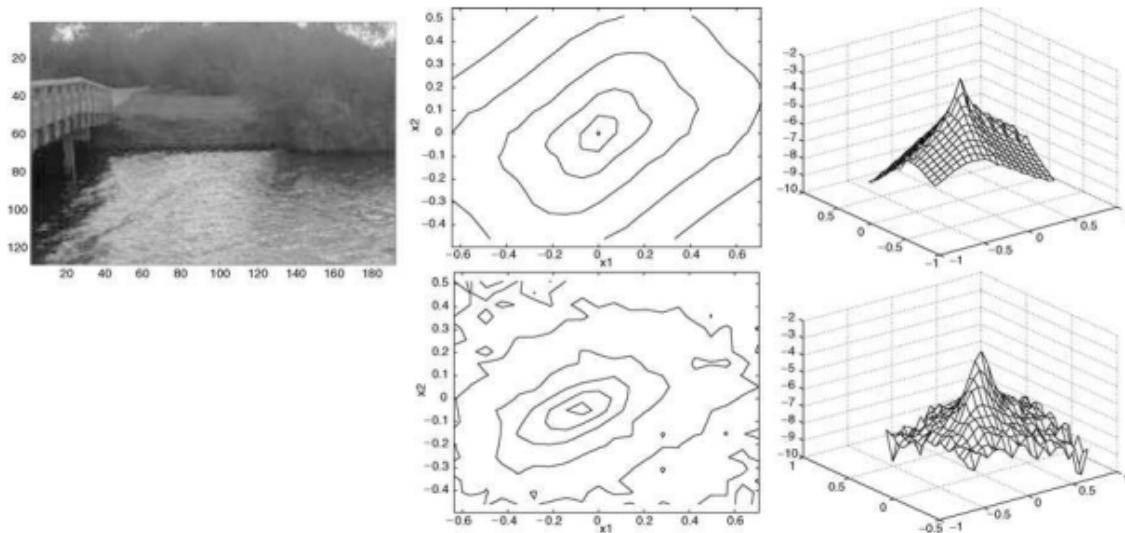


图 4: 贝塞尔 K 形式的二维扩展：对于左图所示图像，估计的（上图）和观测到的（下图）双变量密度以网格和等值线的形式绘制，比例为对数。

一些常用的估计图像流形的方法。

#### 4.1. 通过线性/局部线性子空间近似图像流形

也许最简单的近似图像流形的技术是通过观测值拟合一个线性子空间。拟合标准将指定所选的精确子空间。例如，如果想要最小化累积残差误差（观测点与子空间之间的欧几里得距离），则最佳子空间是数据矩阵的主导（或主）子空间，并且可以使用特征分解或奇异值分解轻松计算。PCA 的常见应用通过面部图像识别别人 [48] 或自然图像研究 [39]。相反，如果目标是 minimized 投影分量之间的统计相关

性，或使它们尽可能独立，则会产生独立分量基础 [6, 23, 86]。其他标准导致了子空间基础的类似公式化，例如稀疏性 [63]、Fisher 判别 [5] 和非负分解 [52]。

稀疏性的使用通常是受自然图像的尺度不变性的启发。

在一般情况下，用平坦子空间近似图像流形显然是有局限性的。在 [96] 中，作者认为，图像的线性处理会在组件之间留下相当大的依赖关系，因此需要一种非线性技术。一种扩展方法是通过用低维子空间拟合相邻图像来寻求图像流形的“局部线性嵌入”近似 [68, 81]。邻域的定义是通过欧几里得度量，但这值得进一步研究。另一种想法是将局部基元素（如小波基）组合成更高级别的结构，以更好地表示图像流形。例如，Zhu 等人 [97] 将变换后的基础元素的位置组合起来，形成称为 *texton* 的结构，以便更好地表征图像及其流形。图 5 左侧面板显示了构建“星形”texton 以匹配给定图像的示例。

## 4.2. 图像流形的可变形模板表征

Grenander 的图像理解模式理论框架受物理模型启发，并产生场景的概率代数表示。

场景中出现的物体以 3D 形式呈现形状、纹理和反射率的模型，以及它们在场景中的出现，通过对典型出现或模板进行组变换来捕获。这种方法的一个强大特点是实际的物理场景和结果图像之间的逻辑分离。场景的变化性最好用

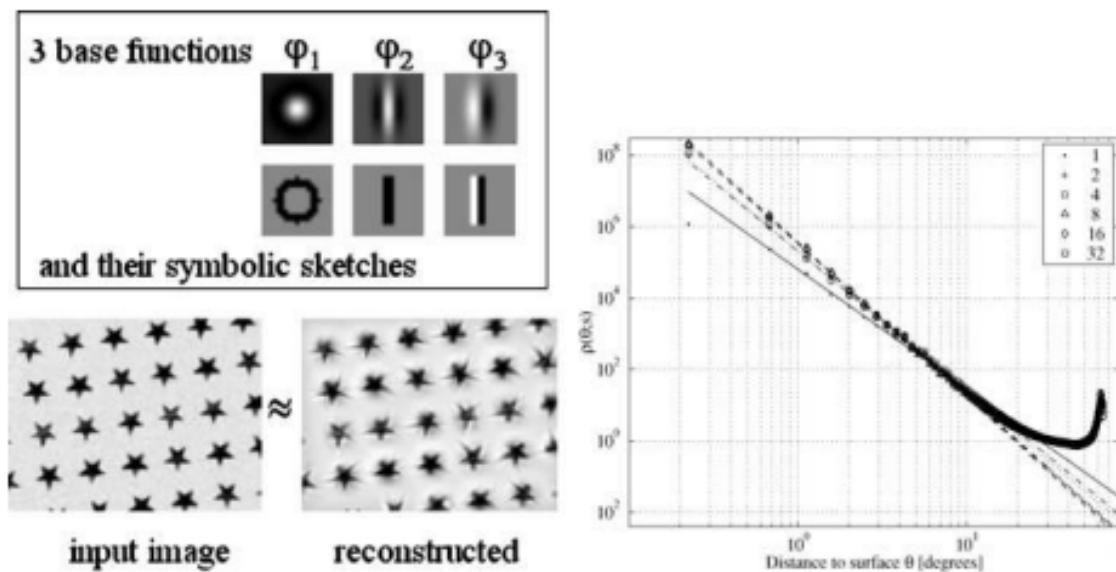


图 5: 左图为了对星空图像进行建模，一些小波基被组合成星空文本。右图自然图像在状态空间的流形处具有无限的概率密度，该流形与模糊边缘相对应。图中显示了经验概率密度与射流空间中高斯尺度空间边缘流形距离的函数关系，曲线对应于自然图像在尺度  $s = 1\ 2\ 4\ 8\ 16\ 32$  时的射流表示。

三维坐标系遵循物理原理，而不是欧几里得表示法无法很好地发挥作用的图像空间。由于三维场景和二维图像通过投影相互关联，(正交或透视)图，通过投影使用 3D 表示形成的流形，可以生成由图像形成的流形。如 [34, 59] 所述，3D 系统中物理对象的出现被建模为作用于场景中对象配置的群的

轨道。这些轨道被投影到图像空间中以形成图像所在的流形。令  $C^\alpha$  为标记为  $\alpha$  的 3D 对象的 CAD 模型，令  $S$  为改变场景中对象出现的变换群（3D 旋转、平移、变形等）。令对于  $s \in S$ ， $sC^\alpha$  表示  $s$  对模板的作用。

那么， $\{sC^\alpha : s \in S\}$  就是与物体  $\alpha$  相关联的三维轨道。进一步，如果  $T$  是从三维场景到像平面的成像投影，那么  $\{T(sC^\alpha) : s \in S\}$  就是该物体生成的像流形，该像流形的维数等于  $\dim(S)$ 。该流形的几何形状，例如切空间或指数映射，也可以通过将它们对应的项投影到更大的空间中来获得。

### 4.3. 图像基元流形的经验证据

与可变形模板理论所倡导的物理规范相反，人们也可以直接使用观察到的图像来研究图像流形。基于图像的搜索 [51, 64] 也受到了 Marr 思想 [56] 的启发，他将一系列原始强度转换成一种符号表示——所谓的“原始草图”，以边、条、斑点和终止点等图元作为基本元素，从而表示早期视觉。

计算机和人类视觉中的一个重要问题是：Marr 基元在图像数据的状态空间中如何进行几何和统计表示？

来自 Marr 假设的几何模型是一组连续流形，其一般形式为  $M(s) = [F^{(1)}(\cdot) * sC^\alpha, F^{(2)}(\cdot) * sC^\alpha, \dots]$ ，其中  $F^{(j)}(j = 1, \dots, n)$  形成一组过滤器，而  $sC^\alpha$  则是原始  $\alpha$  的图像，其参数为  $s \in S$ 。

生成的图像流形  $M(s)$  的维度由基元的固有维数决定：它通常是一个较小的数字（对比度归一化边缘为 2，条形和圆形曲线等为 3）。此外，由平面基元生成的流形形成层次结构：例如，直边的二维流形是条形的三维流形和圆形边的三维流形的子集。

在 [51] 中，Lee 等人发现  $3 \times 3$  自然图像块的状态空间极其稀疏，块密集地集中在模糊边缘的非线性流形周围。对于自然图像中对比度最高的前 20% 块，一半的数据位于边缘流形的邻域内，而该邻域仅占总状态空间的 9%。估计的概率密度是到流形的距离的函数，形式为  $f(\text{dist}) \sim 1/\text{dist}^\beta$ ，其中  $\beta = 2.5$ ；它具有

在像流形上密度无限大，其中  $\text{dist} = 0$ 。

论文 [64] 通过考虑滤波补丁来扩展这些结果，其中滤波器为（最高三阶）高斯核的导数。图 5 右侧面板显示了自然图像多尺度表示中估计的概率密度作为到边缘流形的距离的函数。密度近似尺度不变，似乎收敛到函数形式  $p(\text{dist}) \sim 1/\text{dist}^{0.7}$ 。

## 5. 统计模型的应用

开发正式统计模型的主要原因是将其应用于图像处理/分析应用。大量应用不断受益于图像统计建模的进步。这里我们选择了一个重要的子集。

### 5.1. 纹理合成

较新的统计模型的使用彻底改变了纹理分析领域。1980 年，Faugeras 等人 [29] 建议使用滤波图像的边缘来表示纹理。Bergen 和 Landy [7]、Chubb 等人 [21] 以及 Heeger 和 Bergen [41] 也提倡使用直方图。Zhu 等人 [99] 表明，使用一组滤波器获得的滤波图像的边缘分布足以表征均匀纹理。选择直方图意味着只保留出现的频率，而丢弃位置。利用这些纹理的周期性，人们可以使用几个尺度和方

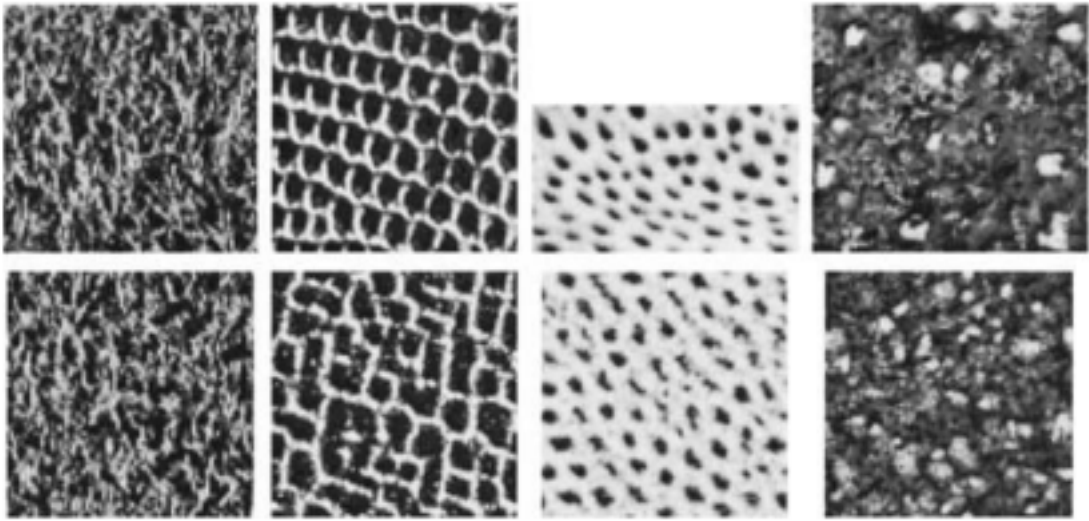


图 6: 使用 Julesz 集合模型（前两列）和 Gibbs 模型（后两列）进行纹理合成。上图：观察到的真实图像，下图：合成图像。

向的小波分解提取特征，并用它们来表示图像。已经提出了许多方案，其中两种是：(i) Julesz 集合，它对所有产生相同直方图的图像施加等效性，以及 (ii) 前面在第 3 节中提到的 Gibbs 模型。图 6 显示了在这些模型下合成的一些纹理示例：顶行显示真实图像，底行显示相应的合成纹理。必须注意的是，即使原始直方图被参数解析形式取代，该合成框架仍然有效，如 [80] 所示。

除了使用边缘之外，Portilla 和 Simoncelli [65] 开发了一个基于滤波器响应联合统计特征的模型。具体来说，他们测量了原始系数的相关性以及它们幅度的相关性，并通过迭代更新图像和强制约束，开发了一种有效的算法，用于合成受这些约束的随机图像。

## 5.2. 图像压缩

压缩似乎是新兴统计图像模型的自然应用。在典型的实现中，图像使用线性基础进行分解，该表示的系数被量化为离散值，然后利用这些值的相对发生概率对这些值进行有效编码。在这种情况下，这里考虑的统计模型提供了这些概率的估计值。最广泛使用的图像压缩方案是 JPEG 标准，它基于逐块频率分解。在 80 年代初期，人们认识到多尺度小波风格表示提供了更大的灵活性和更好的压缩性能 [15, 87]。早期的编码器基于系数的边际模型，但随着 Shapiro [73] 开发出第一个上下文编码器，这种情况发生了突然变化。这个编码器以及随后的许多编码器 [20, 67] 都利用了小波的联合统计特性的启发式优势。一些后续编码器更明确地基于此类模型 [14, 54]。

## 5.3. 图像分类

一个有趣的应用是将图像分类为一些预定义类别。如果发现某些图像的低阶统计量足以满足此目的，则可以实现一定程度的效率。我们从经典的分类方法开始讨论。

**5.3.1. 多重假设检验。**给定与类别相关的概率模型,经典的分类方法是假设选择 [37]。给定一个观察结果,目标是选择最有可能产生该观察结果的假设。令  $H_1, H_2, \dots, H_n$  对应  $n$  个图像类别,令  $P(I | H_i)$  为图像  $I$  属于  $H_i$  类的可能性,则假设选择可以通过一系列二元假设检验来完成:  $\frac{P(I|H_i)}{P(I|H_j)} \geq_{H_i} v_{ij}$ , 其中  $v_{ij}$  是阈值,通常取为 1。奈曼-皮尔逊引理在一定意义上为假设检验提供了最优性。在贝叶斯选择的情况下,阈值  $v_{ij}$  由先验比率  $v_{i,j} = P(H_i)/P(H_j)$  给出。对于所有这些测试,都需要似然函数  $P(I | H_i)$  这取决于统计模型的选择。

例如,如果选择参数形式来建模图像概率,则类别可以直接与参数值相关。参数的典型值(“平均值”)代表一个类别,并且似然值  $P(I | H_i)$  以参数形式写出,其中包含每个类别的典型参数值。可变形模板(第 4 节)已成功应用于对象识别。在存在干扰变量  $s \in S$  的情况下,通过对  $I$  进行假设检验获得推断。这里,似然值是通过干扰积分  $P(I | H_i) = \int_S P(I, s | H_i) \gamma(ds)$  计算的。组  $S$  上的干扰变量估计和估计误差界限在 [35] 中推导出来,而  $I$  的假设检验和识别误差界限在 [37] 中推导出来。

**5.3.2. 图像比较的度量。**图像分析的一个更广泛目标是量化两个给定图像之间的差异。给定这样的度量,可以执行图像聚类、图像检索、图像分类,甚至识别给定图像中的对象。如果概率模型是参数化的,则可以在图像空间上得出采用参数形式的距离度量。如果  $f(x | p_1, c_1)$  和  $f(x | p_2, c_2)$  是两个分别由  $(p_1, c_1)$  和  $(p_2, c_2)$  参数化的单变量密度函数,则距离度量采用以下形式:  $d(p_1, c_1, p_2, c_2) = \tilde{d}(f(x | p_1, c_1), f(x | p_2, c_2))$ , 其中  $\tilde{d}$  是单变量密度空间上的度量。对于  $\tilde{d}$ , 已经提出了几种形式,包括测地线长度(黎曼度量)、Earth Mover 距离 [22]、Kullback-Leibler 散度、Renyi 散度 [42]、Jensen-Renyi 散度 [38, 40]、 $x^2$ -距离和  $p = 1, 2, \dots$ , 的  $L^p$  范数。

度量的选择取决于应用和所需的计算简单性。Srivastava 等人 [80] 在两个 Bessel K 形式之间推导出了  $L^2$  度量的参数形式。可以使用估计的参数直接比较两个密度,而无需计算完整的直方图。

## 5.4. 图像去噪

一种常见的方法是将图像分解为空间频率带,并在进行一些非线性变换后对系数进行阈值处理,这种方法首先由 Bayer 和 Powell [4] 提出,后来由 Donoho [26]。非线性变换主要用于将所有小波系数收缩至零。这种收缩基于阈值处理,可以将其实现为硬阈值或软阈值。要应用统计方法(例如贝叶斯或 MAP 技术),需要对图像像素建立明确的先验模型。Simoncelli 和 Adelson [76] 研究了贝叶斯框架中的图像去噪,而 Moulin 和 Liu [60] 报告了在统计去噪方法中使用广义拉普拉斯模型。

分解图像的统计数据比原始图像的统计数据更容易表征,这一事实导致了許多基于金字塔的图像去噪方法的出现。图像被分解为多尺度表示,系数的统计数据用于在各个频带中进行去噪 [53, 66, 75]。

此外,联合统计的高斯尺度混合模型可以在贝叶斯框架中使用 [66],其结果比边际模型要好得多。

## 5.5. 其他应用

许多其他图像分析/合成应用也受益于统计思想。

Yendrikhovskij[94] 使用颜色统计聚类来模拟感知的颜色环境并计算颜色类别。物体的双向辐射分布函数 (BRDF) 完全指定了其在任意照明条件下的外观。Dror 等人 [28] 已证明这些照明图的统计数据与自然图像的统计数据相似,因此所提出的图像模型也适用于其中 [92]。

## 6. 讨论

本文讨论了自然图像统计建模的一些最新进展。这些模型不仅比传统模型提供了更好的匹配，而且还在许多成像应用中带来了显著的改进。

尽管过去二十年来，我们在理解自然图像的复杂统计特性方面取得了实质性进展，但我们距离完全概率模型还很远。例如，从现有模型中抽取的样本无法捕捉现实世界图像的多样性和复杂性，除非在同质纹理的有限情况下。

除了图像统计的单变量和双变量密度之外，计算复杂度呈指数级增长。一个重要的问题是：在特定成像应用的背景下，比如从视频图像中进行人脸识别，需要什么顺序密度才能确保合理的成功？对于同质纹理，单变量模型已经取得了成功，但对于更一般的应用，还需要多少。

充分统计的概念需要针对不同的应用环境进行精确化。即使在提出的模型中，仍有几个问题尚未解决。例如，枯叶模型中的一个问题是如何结合对象的纹理和依赖关系。目前，对于真实的枯叶模型的分析结果很少。

本文描述的许多模型都是对一组图像的统计数据建模；它们在单个图像分析中的应用需要澄清。

除了图像合成和压缩等应用之外，开发统计模型的一个重要原因是图像理解，这是一个存在许多悬而未决问题的领域。其中一个问题是：给定一个矩形外的图像，找到矩形内最可能的插值。缺乏马尔可夫性使得经典谐波分析无法进行，并指出了图像集合背后更强大的模式理论结构。

## 致谢

我们感谢 van Hateren 数据库的创建者将其数据库向公众开放。我们向我们的合作者 Ulf Grenander、Kim S. 表示感谢。

Pedersen 和 David Mumford 对本研究的帮助。非常感谢 Hamid Krim 的邀请和对撰写本文的支持。AS 部分得到了 NSF DMS-0101429、ARO DAAD19-99-1-0267 和 NMA 201-01-2010 资助。

EPS 部分得到了 NSF CAREER 拨款 MIP-9796040、Alfred P. Sloan 基金会和霍华德休斯医学研究所的支持。

## 笔记

1. 有关“前向”和“后向”枯叶算法的讨论，请参阅 [46]。

## 参考文献

1. L. Alvarez, Y. Gousseau, and J.-M. Morel, “The size of objects in natural and artificial images,” in *Advances in Imaging and Electron Physics*, P.W.H. et al. (Eds.), Academic Press: New York, 1999.

2. D. Andrews and C. Mallows, "Scale mixtures of normal distributions," *Journal of the Royal Statistical Society*, Vol. 36, pp.99–102, 1974.
3. O. Barndorff-Nielsen, "Exponentially decreasing distribution for the logarithm of a particle size," *Journal of Royal Statistical Society A*, Vol. 353, pp. 401–419, 1977.
4. B.E. Bayer and P.G. Powell, "A method for the digital enhancement of unsharp, grainy photographic images," *Advances in Computer Vision and Image Processing*, Vol. 2, pp. 31–88, 1986.
5. P.N. Belhumeur, J.P. Hefanaha, and D.J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711–720, 1997.
6. A.J. Bell and T.J. Sejnowski, "The independent components of natural scenes are edge filters," *Vision Research*, Vol. 37, No. 23, pp. 3327–3338, 1997.
7. J.R. Bergen and M.S. Landy, *Computational Modeling of Visual Texture Segregation*, MIT Press: Cambridge, MA, 1991, pp. 253–271.
8. J. Besag, "Spatial interaction and statistical analysis of lattice systems (with discussion)," *J. Royal Statist. Soc. B*, Vol. 36, pp. 192–326, 1974.
9. J. Besag, "On the statistical analysis of dirty pictures," *J. Royal Statistical Society B*, Vol. 48, No. 3, pp. 259–302, 1986.
10. P. Billingsley, *Probability and Measure*, Wiley Series in Probability and Mathematical Statistics, 1995.
11. D. Bitouk, U. Grenander, M.I. Miller, and P. Tyagi, "Fisher information measures for ATR in clutter," in *Proceedings of SPIE, Automatic Target Recognition XI*, 2001, Vol. 4379, pp. 560–572.
12. T. Bollerslev, K. Engle, and D. Nelson, *ARCH Models*, Vol. IV, North-Holland: Amsterdam, 1994, pp. 2959–3038.
13. H. Brehm and W. Stammerl, "Description and generation of spherically invariant speech-model signals," *Signal Processing*, Vol. 12, pp. 119–141, 1987.
14. R.W. Buccigrossi and E.P. Simoncelli, "Image compression via joint statistical characterization in the wavelet domain," *IEEE Transactions on Image Processing*, Vol. 8, No. 12, pp. 1688– 1701, 1999.
15. P.J. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, Vol.31, No. 4, pp. 532–540, 1983.

16. G.J. Burton, N.D. Haig, and I.R. Moorhead, "A self-similar stack model for human and machine vision," *Biological Cybernetics*, Vol. 53, No. 6, pp. 397–403, 1986.
17. G.J. Burton and I.R. Moorhead, "Color and spatial structures in natural scenes," *Applied Optics*, Vol. 26, No. 1, pp. 157–170, 1987.
18. J.-F. Cardoso, "Source separation using higher order moments," in *Proceedings of ICASSP*, 1989, pp. 2109–2112.
19. Z. Chi, "Probability models for complex systems," Ph.D. Thesis, Division of Applied Mathematics, Brown University, 1998.
20. C. Chrysafis and A. Ortega, "Efficient context-based entropy coding for lossy wavelet image coding," in *Data Compression Conf.*, 1997, Snowbird, Utah.
21. C. Chubb, J. Econopouly, and M.S. Landy, "Histogram contrast analysis and the visual segregation of IID textures," *J. Opt. Soc. Am. A*, Vol. 11, pp. 2350–2374, 1994.
22. S. Cohen and L. Guibas, "The earth mover's distance under transformation sets," in *Proceedings of Seventh IEEE International Conference on Computer Vision*, Vol. 2, pp. 1076–1083, 1999.
23. P. Comon, "Independent component analysis, a new concept?" *Signal Processing, Special Issue on Higher-Order Statistics*, Vol. 36, No. 3, 1994.
24. J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of the Optical Society of America A*, Vol. 2, No. 7, pp. 23–26, 1985.
25. N.G. Deriugin, "The power spectrum and the correlation function of the television signal," *Telecommunications*, Vol. 1, pp.1–12, 1956.
26. D. Donoho, "Denoising by soft-thresholding," *IEEE Trans. Info. Theory*, Vol. 43, pp. 613–627, 1995.
27. D.L. Donoho and A.G. Flesia, "Can recent innovations in harmonic analysis 'Explain' key findings in natural image statistics," *Network: Computation in Neural Systems*, Vol. 12, No. 3, pp. 371–393, 2001.
28. R.O. Dror, T.K. Leung, E.H. Adelson, and A.S. Willsky, "Statistics of real-world illumination," in *Proceedings of 2001 IEEE Conference on CVPR*, 2001, Vol. 2. pp. 164–171.
29. O.D. Faugeras and W.K. Pratt, "Decorrelation methods of texture feature extraction," *IEEE Pat. Anal. Mach. Intell.* Vol. 2, No. 4, pp. 323–332, 1980.



30. D.J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Optical Society of America*, Vol. 4, No. 12, pp. 2379–2394, 1987.
31. D. Gabor, "Theory of Communications," *Journal of IEE (London)*, Vol. 93, pp. 429–457, 1946.
32. D. Geman and A. Koloydenko, "Invariant statistics and coding of natural microimages," in *Proc. of the IEEE Workshop on Statistical and Computational Theories of Vision*, 1999.
33. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, pp. 721–741, 1984.
34. U. Grenander, *General Pattern Theory*, Oxford University Press, 1993.
35. U. Grenander, M.I. Miller, and A. Srivastava, "Hilbert-Schmidt lower bounds for estimators on matrix lie groups for ATR," *IEEE Transactions on PAMI*, Vol. 20, No. 8, pp. 790–802, 1998.
36. U. Grenander and A. Srivastava, "Probability models for clutter in natural images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 4, pp. 424–429, 2001.
37. U. Grenander, A. Srivastava, and M.I. Miller, "Asymptotic performance analysis of Bayesian object recognition," *IEEE Transactions of Information Theory*, Vol. 46, No. 4, pp. 1658–1666, 2000.
38. A.B. Hamza, Y. He, and H. Krim, "An information divergence measure for ISAR image registration," in *Proc. of IEEE Workshop on Statistical Signal Processing*, 2001.
39. P.J.B. Hancock, R.J. Baddeley, and L.S. Smith, "The principal components of natural images," *Network*, Vol. 3, pp. 61–70, 1992.
40. Y. He, A.B. Hamza, and H. Krim, "A generalized divergence measure for robust image registration," *IEEE Transactions on Signal Processing*, 2002, to appear.
41. D.J. Heeger and J.R. Bergen, "Pyramid-based texture analysis/synthesis," in *Proceedings of SIGGRAPH*, 1995, pp. 229–238.
42. A.O. Hero, B. Ma, O. Michel, and J. Gorman, "Alpha-divergence for classification, indexing and retrieval," Communication and Signal Processing Laboratory, Technical Report CSPL-328, U. Mich, 2001.
43. J. Huang, "Statistics of natural images and models," Ph.D. Thesis, Division of Applied Mathematics, Brown University, RI,

- 44.
45. A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley and Sons; New York, 2001.
46. D. Jeulin, I. Terol-Villalobos, and A. Dubus, "Morphological analysis of UO<sub>2</sub> powder using a dead leaves model," *Microscopy, Microanalysis, Microstructure*, Vol. 6, pp. 371–384, 1995.
47. W.S. Kendall and E. Thonnes, "Perfect simulation in stochastic geometry," Preprint 323, Department of Statistics, University of Warwick, UK, 1998.
48. D. Kersten, "Predictability and redundancy of natural images," *Journal of Optical Society of America*, Vol. 4, No. 12, pp. 2395–2400, 1987.
49. M. Kirby and L. Sirovich, "Application of the KarhunenLoeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 103–108, 1990.
50. E.R. Kretzmer, "Statistics of television signals," *Bell Syst. Tech. Journal*, Vol. 31, pp. 751–763, 1952.
51. A.B. Lee and D. Mumford, "Occlusion models for natural images: A statistical study of scale-invariant dead leaves model," *International Journal of Computer Vision*, Vol. 41, No. 1/2, 2001.
52. A.B. Lee, K.S. Pedersen, and D. Mumford, "The nonlinear statistics of high-contrast patches in natural images," *Int. J. Computer Vision*, 2002, in press.
53. D.D. Lee and H.S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, Vol. 401, pp. 788–791, 1999.
54. D. Leporini, J.-C. Pesquet, and H. Krim, "Best basis representations with prior statistical models," in *Lecture Notes in Statistics: Bayesian Inferences in Wavelet Based Models*, P. Muller and B. Vidakovic (Eds.), Springer Verlag: Berlin, 1999.
55. S.M. LoPresto, K. Ramchandran, and M.T. Orchard, "Wavelet image coding based on a new generalized gaussian mixture model," in *Data Compression Conf.*, 1997, Snowbird, Utah.
56. S.G. Mallat, "A Theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, pp. 674–693, 1989.

57. D. Marr, *Vision: A computational Investigation into the Human Representation and Processing of Visual Information*, W.H.Freeman and Company: New York, 1982.
58. G. Matheron, *Random Sets and Integral Geometry*, John Wiley and Sons: New York, 1975.
59. K.D. Miller, "A Model for the development of simple-cell receptive fields and the arrangement of orientation columns through activity dependent competition between on- and off-center inputs," *Journal of NeuroScience*, Vol. 14, pp. 409–441, 1994.
60. M.I. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *International Journal of Computer Vision*, Vol. 41, No. 1/2, pp. 61–84, 2002.
61. P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using a generalized Gaussian and complexity priors," *IEEE Trans. Info. Theory*, Vol. 45, pp. 909–919, 1999.
62. D. Mumford and B. Gidas, "Stochastic models for generic images," *Quarterly of Applied Mathematics*, Vol. 59, No. 1, pp.85–111, 2001.
63. B.A. Olshausen and D.J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, Vol. 381, pp. 607–609, 1996.
64. B.A. Olshausen and D.J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by V1?" *Vision Research*, Vol. 37, No. 23, pp. 3311–3325, 1997.
65. K.S. Pedersen and A.B. Lee, "Toward a full probability model of edges in natural images," in *Proc. of ECCV'02*, 2002.
66. J. Portilla and E.P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, Vol. 40, No. 1, pp. 49–70, 2000.
67. J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Adaptive Wiener denoising using a Gaussian scale mixture model in the wavelet domain," in *Proc 8th IEEE Int'l Conf on Image Proc.* Thessaloniki, Greece, IEEE Computer Society, 2001, pp. 37– 40.
68. R. Rinaldo and G. Calvagno, "Image coding by block prediction of multiresolution subimages," *IEEE Trans Im Proc.*, 1995.
69. S.T. Roweis and L.K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, Vol. 290, pp. 2323–2326, 2000.
70. D.L. Ruderman, "The statistics of natural images," *Network*, Vol. 5, pp. 517–548, 1994.
71. D.L. Ruderman, "Origins of scaling in natural images," *Vision Research*, Vol. 37, No. 23, pp. 3385–3398, 1997.

72. D.L. Ruderman and W. Bialek, "Scaling of natural images: Scaling in the woods," *Physical Review Letters*, Vol. 73, No. 6, pp.814–817, 1994.
73. J.P. Serra, *Image Analysis and Mathematical Morphology*, Academic Press: London, 1982.
74. J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Sig. Proc.*, Vol. 41, No. 12, pp. 3445– 3462, 1993.
75. E.P. Simoncelli, "Statistical models for images: Compression, restoration and synthesis," in *Proc. 31st Asilomar Conf on Signals, Systems and Computers*, Pacific Grove, CA. IEEE Computer Society, 1997, pp. 673–678.
76. E.P. Simoncelli, "Bayesian denoising of visual images in the wavelet domain," in *Bayesian Inference in Wavelet Based Models*, Lecture Notes in Statistics, Vol. 41, Springer Verlag: Berlin, 1999.
77. E.P. Simoncelli and E.H. Adelson, "Noise removal via Bayesian wavelet coring," in *Third Int'l. Conf on Image Proc.*, Lausanne, IEEE Sig Proc Society, 1996, Vol. I, pp. 379–382.
78. E.P. Simoncelli and R.T. Buccigrossi, "Embedded wavelet image compression based on a joint probability model," in *Proceedings of ICIP (1)*, 1997, pp. 640–643.
79. E.P. Simoncelli, W.T. Freeman, E.H. Adelson, and D.J. Heeger, "Shiftable multi-scale transforms," *IEEE Trans Information Theory*, Vol. 38, No. 2, pp. 587–607, 1992.
80. E. Simoncelli and B. Olshausen, "Natural image statistics and neural representation," *Annual Review of Neuroscience*, Vol. 24, pp. 1193–1216, 2001.
81. A. Srivastava, X. Liu, and U. Grenander, "Universal analytical forms for modeling image probability," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 9, 2002.
82. J.B. Tenenbaum, V. Silva, and J.C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, Vol. 290, pp. 2319–2323, 2000.
83. M.G.A. Thomson, "Beats, kurtosis, and visual coding," *Network: Computation in Neural Systems*, Vol. 12, No. 3, pp. 271–287, 2001.
84. A. Turiel, G. Mato, and N. Parga, "The self-similarity properties of natural images resemble those of turbulent flows," *Physical Review Letters*, Vol. 80, No. 5, pp. 1098–1101, 1998.
85. A. Turiel and N. Parga, "The multi-fractal structure of contrast changes in natural images: From sharp edges to textures," *Neural Computation*, Vol. 12, pp. 763–793, 2000.
86. A. Turiel, N. Parga, D.L. Ruderman, and T.W. Cronin, "Multiscaling and information content of natural color images," *Physical Review E*, Vol. 62, No. 1, pp. 1138–1148, 2000.

87. J.A. van Hateren, "Independent component filters of natural images compared with simple cells in primary visual cortex," Proc. Royal Statistical Society of London B, Vol. 265, pp. 359–366, 1998.
87. M. Vetterli, "Multidimensional subband coding: Some theory and algorithms," *Signal Processing*, Vol. 6, No. 2, pp. 97–112, 1984.
88. M.J. Wainwright and E.P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," Advances in Neural Information Processing Systems, S.A. Solla, T.K. Leen, and K.-R. Muller (Eds.), 2000, pp. 855–861.
89. M.J. Wainwright, E.P. Simoncelli, and A.S. Willsky, "Random cascades on wavelet trees and their use in modeling and analyzing natural imagery," *Applied and Computational Harmonic Analysis*, Vol. 11, No. 1, pp. 89–123, 2001.
90. A.B. Watson, "The cortex transform: Rapid computation of simulated neural images," *Comp. Vis. Graphics Image Proc.*, Vol.39, pp. 311–327, 1987.
91. B. Wegmann and C. Zetsche, "Statistical dependence between orientation filter outputs used in a human vision based image code," in *Proceedings of Visual Communication and Image Processing*, Society of Photo-Optical Instrumentation Engineers, Vol. 1360, pp. 909–922, 1990.
92. Y. Weiss, "Deriving intrinsic images from image sequences," in Proceedings Eight IEEE International Conference on Computer Vision, 2001, Vol. 2, pp. 68–75.
93. G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, Springer: Berlin, 1995.
94. S.N. Yendrikhovskij, "Computing color categories from statistics of natural images," *Journal of Imaging Science and Technology*, Vol. 45, No. 5, pp. 409–417, 2001.
95. C. Zetsche, "Polyspectra of natural images," in Presented at Natural Scene Statistics Meeting, 1997.
96. C. Zetsche and F. Rohrbein, "Nonlinear and extra-classical receptive field properties and the statistics of natural images," *Network: Computation in Neural Systems*, Vol. 12, No. 3, pp. 331–350, 2001.
97. S.C. Zhu, C.E. Guo, Y.N. Wu, and Y.Z. Wang, "What are textons?," in *Proc. of European Conf. on Computer Vision*, 2002.
98. S.C. Zhu and D. Mumford, "Prior learning and Gibbs reaction-diffusion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 19, No. 11, pp. 1236–1250, 1997.

99. S.C. Zhu, Y.N. Wu, and D. Mumford, “Minimax entropy principles and its application to texture modeling,” *Neural Computation*, Vol. 9, No. 8, pp. 1627–1660, 1997.



图 7: **Anuj Srivastava** 是佛罗里达州立大学统计系助理教授。他于 1996 年获得华盛顿大学电气工程博士学位，1996-1997 年期间曾在布朗大学应用数学系担任客座研究员。他在弗吉尼亚州立大学领导计算视觉实验室，利用多种成像模式从事图像理解研究。他的研究方向是统计信号处理和图像分析、流形统计和计算统计。他的研究得到了 ARO、NSF、NIMA 和 FSU 研究基金的支持。



图 8: **Ann Lee** 在瑞典查尔姆斯理工大学获得工程物理学硕士学位，在布朗大学获得物理学博士学位。她目前是布朗大学应用数学系的客座助理研究员和耶鲁大学数学系的研究导师。她的研究领域包括机器和生物视觉、统计建模和模式理论。



图 9: **Eero Simoncelli** 是纽约大学神经科学和数学副教授。他获得了哈佛大学物理学学士学位，在剑桥大学学习了一年半的数学，然后进入麻省理工学院攻读电子工程和计算机科学博士学位。1993 年 1 月，他加入宾夕法尼亚大学计算机与信息科学系任教。1996 年 9 月，他来到纽约大学，成为斯隆理论视觉神经科学中心的一员。1996 年，他因“视觉信息处理”的研究和教学获得了美国国家科学基金会（NSF）教师早期职业发展（CAREER）基金；1998 年，他又获得了斯隆研究奖学金。2000 年 8 月，他成为霍华德-休斯医学研究所（Howard Hughes Medical Institute）计算生物学新项目的副研究员。



图 10: **Song Chun Zhu** 于 1991 年获得中国科学技术大学计算机科学学士学位。他分别于 1994 年和 1996 年获得哈佛大学（哈佛机器人实验室）计算机科学硕士和博士学位。1996-1997 年，他在布朗大学应用数学系（模式理论组）任助理研究员；1997-1998 年，在斯坦福大学计算机科学系任讲师；1998-2002 年，在俄亥俄州立大学计算机与信息科学系任助理教授。目前，他是加州大学洛杉矶分校统计系和计算机科学系的副教授。他的研究重点是计算机视觉与学习、统计建模和随机计算。他发表过 50 多篇文章，并获得过各种荣誉，包括大卫-马尔奖荣誉奖、国家自然科学基金事业奖、斯隆研究员和海军青年研究员奖。

**Affiliation:**A. SRIVASTAVA<sup>1</sup>E-mail: [anuj@stat.fsu.edu](mailto:anuj@stat.fsu.edu)A.B. LEE<sup>2</sup>E.P. SIMONCELLI<sup>3</sup>S.-C. ZHU<sup>4</sup>

---

Silkman Statistical Journal

published by the Funny Project of Silkman Press

September 2024, Volume 99, Issue 2

[doi:10.18637/ssj.2024.test](https://doi.org/10.18637/ssj.2024.test)

<http://cookwhy.com/>

<http://cookwhy.com>

*Submitted:* 2024-04-12

*Accepted:* 2024-05-12

---

<sup>1</sup>Department of Statistics, Florida State University, Tallahassee, FL 32306, USA [anuj@stat.fsu.edu](mailto:anuj@stat.fsu.edu)

<sup>2</sup>Division of Applied Mathematics, Brown University, Providence, RI 02912, USA

<sup>3</sup>Courant Institute for Mathematical Sciences, New York University, New York, NY 10003, USA

<sup>4</sup>Department of Computer Science, Ohio State University, Columbus, OH 43210, USA Abstract.