

三维稀疏变换域协同滤波图像去噪 (BM3D TIP2007)

摘要

我们提出了一种基于变换域中增强稀疏表示的新型图像去噪策略。通过将类似的二维图像片段（例如块）分组成为我们称之为“组”的三维数据阵列来实现稀疏度的增强。协同滤波是为处理这些三维组而开发的一种特殊过程。我们使用三个连续步骤来实现它：组的三维变换，变换谱的收缩和三维逆变换。结果是由联合滤波的分组图像块组成的三维估计。通过衰减噪声，协同滤波甚至可以揭示分组块共享的嵌套细节，同时保留每个块本质特征。随后，已滤波的块将返回其原始位置。因为这些块是重叠的，所以对于每个像素，我们获得了许多需要合并起来的不同估计。聚合是一种特殊的平均过程，它利用了这种冗余。通过使用专门开发的协同维纳滤波，我们的方法获得了显著的改进。本文详细介绍了一种基于该方法的去噪算法及其高效实现，更进一步地还开发了相应的彩色图像去噪算法。实验结果表明，该算法在峰值信噪比和主观视觉质量两方面均取得了当前最先进的去噪效果。

关键词：图像去噪，稀疏性，自适应分组，块匹配，三维变换收缩。

1. 介绍

目前已存在大量的去噪方法，它们源自各种学科，如概率论，统计学，偏微分方程，线性和非线性滤波，光谱和多分辨率分析。所有这些方法都依赖于关于真实（无噪声）信号的一些明确或隐含的假设，以便将其与随机噪声正确地分离。

特别的，变换域去噪方法通常假设真实信号可以通过几个基本元素的线性组合很好地近似。也就是说，信号在变换域中稀疏地表示。因此，通过保留少量传递主要是真实信号能量的高幅变换系数，并丢弃主要是由于噪声引起的其余部分，就可以有效地估计真实信号。表示的稀疏性取决于变换和真实信号的属性。

多分辨率变换可以实现空间局部细节的良好稀疏性，例如边缘和奇点。因为这些细节通常在自然图像中很丰富并且传达了嵌入其中的信息的重要部分，所以这些变换已经在图像去噪领域取得了显著的应用效果。最近，许多基于多分辨率变换的高级去噪方法被开发出来，这些方法依赖于典型过完备变换（例如平移不变变换和面向多重变换）系数之间的详细统计相关性。在[1],[2],[3],[4]中可以看到这种图像去噪方法的例子。

不仅限于小波技术，过完备的表示在提高方案恢复能力方面一贯起着重要的作用，即使是对于最基本的基于变换的方法来说也是如此。这一点在代表“滑动窗口变换域图像去噪方法”的文献[5],[6]中就有所体现，其基本思想是在局部（窗口化）变换域中应用收缩变换。其中，连续窗口之间的重叠解释了过完备性，而变换本身通常是正交的，例如2D DCT。

然而，如果所采用的变换不能获得某些图像细节的稀疏表示，则过完备性本身不足以补偿无效收缩。例如，2D DCT在表示锐利过渡和奇点方面无效，而小波变换通常对纹理和平滑过渡表现不佳。自然图像的多样性使得任何固定的2D变换都不可能在所有情况下实现良好的稀疏性。因此，常用的

正交变换只能在特定图像模式上实现稀疏表示。

Muresan和Parks [7]提出了局部图像块的自适应主成分作为克服标准正交变换的上述缺点的工具。这种方法为高度结构化的图像模式产生了良好的结果。然而，正确的PCA基实际上由于噪声的存在而劣化。出于类似的意图，Elad和Aharon的K-SVD算法[8]利用通过初步训练程序获得的高度完备的词典。这些技术的缺点是，PCA和学习词典都会产生非常高的计算负担。

另一种方法[9]是应用形状自适应变换于那些形状适配于图像的显著细节，因此主要包含同质信号的邻域上。形状自适应变换可以在这些自适应邻域中实现真实信号的非常稀疏的表示。

最近，一种精细的自适应空间估计策略，即NL-means方法[10]被提出来了。这种方法与变换域方法不同。其基本思想是建立图像的逐点估计，其中获得每个像素，作为以与估计像素为中心的区域相似的区域为中心的像素的加权平均。估计是非局部的，因为原则上可以在图像的所有像素上计算平均值。这种方法的一个重要扩展是基于样本的估计[11]，它利用成对假设检验来定义自适应非局部估计邻域，并获得了与基于变换的技术的最佳成果相比更具竞争力的结果。

在本文中，我们提出了一种基于增强变换域中稀疏表示的新型图像去噪策略。通过将图像的类似2D片段分组为3D数据阵列（我们称之为“组”）来实现稀疏性的增强。协同过滤是为处理这些3D组而开发的一种特殊程序。它包括三个连续的步骤：组的3D变换，变换谱的收缩和逆3D变换。借此，我们获得由一组联合滤波的2D片段组成的组的3D估计。由于分组块之间的相似性，变换可以实现真实信号的高度稀疏表示，从而可以通过收缩很好地分离噪声。通过这种方式，协同过滤甚至可以显示分组片段共享的嵌套细节，同时保留每个片段的基本独特特征。

基于该新策略的图像去噪算法已被开发并详细描述。它概括和改进了我们在[12]中介绍的初步算法。它已被有效实现，提供了有效的复杂性/性能权衡。实验结果表明，它在峰值信噪比和主观视觉质量方面都达到了出色的去噪性能，优于目前的最新技术水平。本文也给出了基于[13]的彩色图像去噪扩展。

本文的结构如下。我们在第二部分介绍了分组和协同过滤概念。开发的图像去噪算法在第III节中描述。该算法的有效且可扩展的实现可以在第IV节中找到，其对彩色图像去噪的扩展在第V节中给出。实验结果在第VI节中给出。第七节对已提出的方法进行了全面讨论，第八节包含了相关结论。

II. 分组和协同过滤

我们将术语“组”的概念定义为：收集与给定信号的类似的 d 维片段，形成 $d + 1$ 维的数据结构。例如，在图像的情况下，信号片段可以是任意2D邻域（例如图像块或块）。在那里，组是通过将类似的图像邻域堆叠在一起而形成的3D阵列。如果邻域具有相同的形状和大小，则形成的3D阵列是广义圆柱体。分组的重要性在于能够使用每个组的更高维度的过滤，这利用了潜在的相似性（分组，片段等）在分组片段之间，以便估计每个片段中的真实信号。这种方法我们称之为协同过滤。

A. 分组

分组可以通过各种技术实现；例如，K-means聚类[14]，自组织映射[15]，模糊聚类[16]，矢量量化[17]等。关于这个主题存在大量文献；我们建议读者参考文献[18]来获取详细、系统地概述这些方法

的信息。

通常将信号片段之间的相似度计算为某个距离测量的倒数。因此，较小的距离意味着较高的相似性。可以采用各种距离测量，例如两个信号片段之间差异的 p 范数。其他例子是在非局部均值估计器[10]中使用的加权欧几里德距离（ $p=2$ ），以及在基于范例的估计器[11]中使用的归一化距离。当处理复杂或不确定（例如噪声）数据时，可能需要首先从信号中提取一些特征，然后仅测量这些特征的距离[18]。

B. 按匹配分组

诸如矢量量化或K均值聚类之类的分组技术基本上基于分区的思想。这意味着它们构建不相交的组或集群（类），使得每个片段属于一个且仅属于一个组。构造其元素具有高度相互相似性的不相交群通常需要递归过程并且可能在计算上要求很高[18]。此外，分割导致不同片段的不均等处理，因为接近该组中心的那些比远离它的那些更好地表示。这种情况总是发生，即使在信号的所有片段等距分布的特殊情况下也是如此。

对于相互类似的信号片段，通过匹配可以实现更简单更有效的分组，其与上述分割方法相反，所形成的组不一定是相邻的。匹配是一种用于查找与给定参考信号片段类似的信号片段的方法。这是通过成对地测试参考片段和位于不同空间位置的候选片段之间的相似性来实现的。与参考片段的距离（即不相似性）小于给定阈值的片段，将被认为是彼此相似的并且随后被分组。相似性起到所关注组的隶属函数的作用，并且参考片段可以被认为是该组的某种中心。任何信号片段都可以用作参考片段，因此可以为其构建组。

我们注意到，对于大多数距离测量，建立参考片段与所有匹配片段之间距离的界限意味着，该组中任何两个片段之间的距离也是有界的。粗略地说，这个界限是该组的直径。虽然对于任意距离测量，这样的陈述可能不能精确地保持，比如对于度量（例如， l_p -范数）的情况，它只是三角不等式的直接结果。

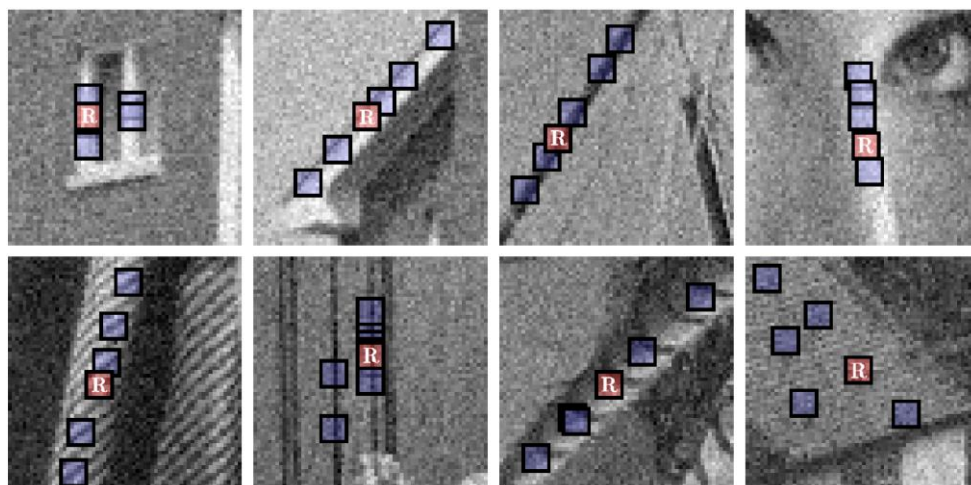


图 1 来自带噪的自然图像的块分组图示，其中噪声为标准偏差 15 的零均值白噪声。每个片段中标有“R”的为参考块，其余块为与其匹配的一些块。

块匹配（BM）是一种特殊的匹配方法，已广泛用于视频压缩（MPEG 1,2和4以及H.26x）中的运动估计。作为分组的特定方式，它用于找到类似的块，然后将它们堆叠在一起形成3D阵列（即组）。

在图1中给出了通过图像的块匹配进行分组的说明性示例，其中我们示出了一些参考块以及与它们相似的匹配的参考块。

C. 协同过滤

给定一组 n 个片段，该组的协同过滤产生 n 次估计，每个分好组的片段一次。通常，这些估计可能不同。术语“协同”是从字面上理解的，即每个分好组的片段协同过滤所有其他片段，反之亦然。

让我们考虑一个用协同滤波估计图2中的由加性零均值独立噪声破坏图片的（未示出）的说明性示例。特别的，让我们关注同一图中所示的已经分组的块。这些块表现出完美的相互相似性，这使得元素平均（即，在相同相对位置的像素之间的平均）成为合适的估计器。因此，对于每个组，通过协同平均的方法就可以产生所有分组块的估计。因为假设相应的无噪声块是相同的，所以估计是无偏的。因此，最终估计误差仅归因于与组中的块数成反比的残差方差。只要这些组包含大量片段，无论信号片段有多复杂，我们都可以获得非常好的估计。

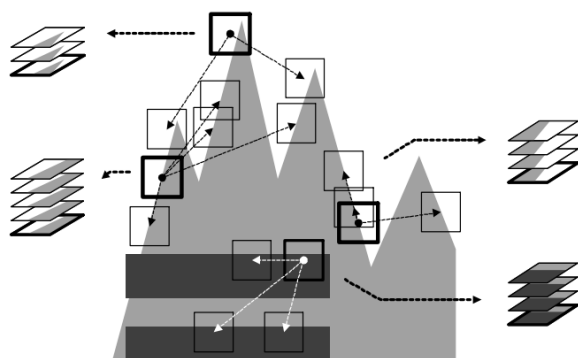


图 2 在人工绘制的图像中分组的简单示例，其中对于每个参考块（带有粗边框），存在完全相似的参考块。

但是，在自然图像中不太可能完全相同的块。如果在同一组内允许不相同的片段，则通过元素平均获得的估计变得有偏差。除非使用允许对每个分组片段产生不同估计的估计器，否则偏差误差有可能占估计中总体最终误差的最大份额。因此，应采用比平均更有效的协同过滤策略。

D. 变换域中收缩的协同过滤

有效的协同过滤可以通过在变换域中的收缩来实现。假设已经形成了类似信号片段的 $d + 1$ 维组，则协同收缩包括以下步骤。

- 将 $d + 1$ 维线性变换应用于该组。
- 收缩（例如通过软阈值和硬阈值或维纳滤波）变换系数以衰减噪声。
- 反转线性变换以生成所有分组片段的估计值

当应用于自然图像片段组时，这种协作的变换域收缩可以特别有效，例如图1中的组。这些组的特征在于：

- 片段内相关性出现在每个分组片段的像素之间——自然图像的独特性
- 片段间相关性出现在不同片段的相应像素之间——分组片段之间的相似性

3D变换可以利用两种相关性，从而在组中生成真实信号的稀疏表示。这种稀疏性使收缩非常有效地衰减了噪声，同时保留了信号的特征。

让我们通过考虑图1所示的分组图像块，简单地说明这种协同收缩的好处。让我们首先考虑以下情况：不执行协作过滤，而是将2D变换分别应用于给定一组 n 个片段。由于这些分组的块非常相似，因此对于它们中的任何一个，我们都应获得大约相同数量的有效变换系数，我们记之为 α 。这意味着整个 n 个片段组由 $n\alpha$ 个系数表示。相反，在协同过滤的情况下，除了2D变换外，我们还对分组后的块应用了1D变换（等效于对整个组应用可分离的3D变换）。如果此一维变换具有基于DC的元素，那么由于块之间的相似度很高，大约只有代表整个组的 α 个有效系数，而不是 $n\alpha$ 个。因此，分组增强了稀疏性，稀疏性随分组块的数量而增加。

如图1所示，在不同空间位置的小图像块之间的强烈相似性在自然图像中确实非常普遍。它是属于均匀区域，边缘，纹理，平滑强度梯度等的块的特征。因此，在对自然图像进行建模时，可以将相互相似的块的存在视为非常现实的假设，这使我们有强地动机去将分组和协同过滤应用于图像去噪算法。

III. 算法

在提出的算法中，分组是通过块匹配实现的，而协同过滤是通过在3D变换域中收缩来实现的。使用的图像片段是固定大小的正方形块。该算法中执行的一般过程如下。通过从输入噪声图像中依次提取参考块并对每个这样的块进行处理：

- 查找与参考块相似的块（块匹配），并将它们堆叠在一起以形成3D阵列（组）
- 对群组执行协同过滤，并将获得的所有分组块的2D估计值返回到其原始位置。

在处理所有参考块之后，获得的块估计可能重叠，因此每个像素有多个估计。我们汇总这些估算值以形成整个图像的估算值。

此通用过程以两种不同的形式实现，以组成一个两步算法。此算法在图3中进行了说明，其过程如下：

步骤1.初步估计。

a) 块估算。对于嘈杂图像中的每个块，请执行以下操作。

- i) 分组。查找与当前处理的块相似的块，然后将它们堆叠在一起形成3D数组（组）。
- ii) 协同性硬阈值。将3D变换应用于所形成的组，通过对变换系数进行硬阈值化来衰减噪声，将3D变换反转以生成所有分组块的估计，并将这些块的估计返回到其原始位置。

b) 聚集。通过对所有获得的重叠的块估计值进行加权平均，计算出真实图像的初步估计值。

步骤2.最终估计：使用基本估算，进行改进的分组和协同维纳滤波。

a) 块级估计。对于每个块，请执行以下操作。

i) 分组。在基本估算范围内使用BM查找与当前处理的块相似的块的位置。使用这些位置，形成两组（3D阵列），一组来自噪声图像，另一组来自基本估计。

ii) 协同维纳过滤。在两个组上应用3D变换。使用关于真实（先导）能谱的基本估计值的能谱对有噪声的组进行维纳滤波。通过对滤波后的系数应用逆3D变换来生成所有分组块的估计，并将块的估计返回到其原始位置。

b) 聚集。通过使用加权平均值汇总所有获得的局部估计，计算出真实图像的最终估计。

上面算法的第二步有两个重要动机：

- 使用基本估计而不是嘈杂的图像可以通过块匹配来改善分组
- 使用基本估计作为经验维纳滤波的导频信号比对噪声数据的3D频谱进行简单的硬阈值处理要有效和准确得多。

观察模型及符号

我们考虑噪声图像 $z: X \rightarrow \mathbb{R}$ 由下述形式构成：

$$z(x) = y(x) + \eta(x), x \in X$$

其中， x 是属于图像域 $X \subset \mathbb{Z}^2$ 的2D空间坐标，并且 η 是以 $\sigma^2, \eta(\cdot) \sim \mathcal{N}(0, \sigma^2)$ 为方差的零均值高斯噪声。

我们用取自 z 的一个固定大小为 $N_1 \times N_1$ 的块来表示 Z_x ，其中 x 是块左上角的坐标。换言之，我们称 Z_x 位于 z 的 x 处。一组归纳好的2D块用带有下标的粗体大写字母表示，该下标是其分组的块的坐标的集合，例如， \mathbf{Z}_S 是一个由位于 $x \in S \subseteq X$ 的块 Z_x 构成的3D数组。为了区分第一步和第二步中使用的参数，我们分别使用上标“ht”（硬阈值）和“wie”（维纳滤波）做标注。例如， N_1^{ht} 是第一步中使用的块尺寸，而 N_1^{wie} 是第二步中使用的块尺寸。类似的，我们使用 \hat{y}^{basic} 表示基本估计，用 \hat{y}^{final} 表示最终估计。

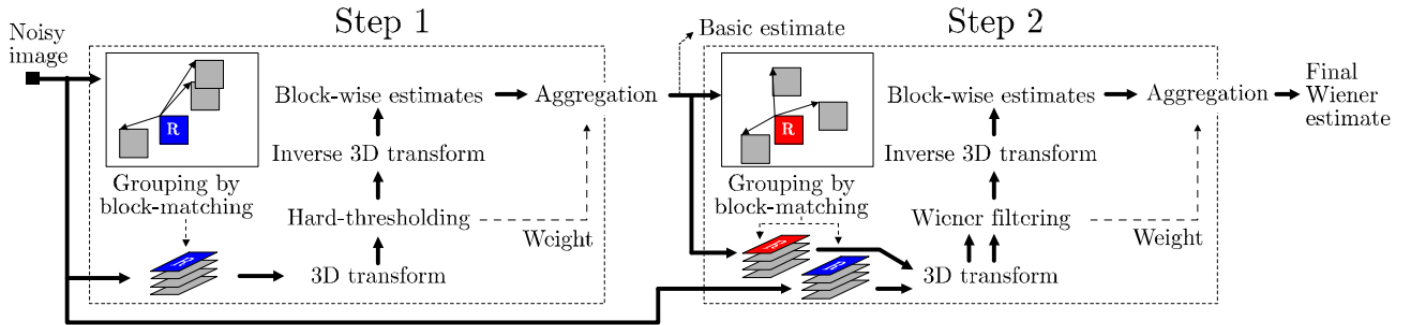


图 3 本文提出的图像去噪算法的流程图。对于每个处理的块（标记为“R”）重复用虚线包围的操作。

接下来的小节中详细介绍了我们所提出的降噪方法的具体步骤。

A. 步骤 1a 和 2a：块估计

在此步骤中，我们以滑动窗口方式处理参考图像块。在此，“过程”代表通过以下方式对所有分组块的真实信号进行分组和估计：

- 步骤1（a）ii中的协同性硬阈值
- 步骤2（a）ii中的协同维纳滤波

所得的估计结果称为“块估计值”。

由于步骤1a和2a具有相同的结构，因此我们分别在以下两个小节中介绍它们。其中，我们将当前处理的图像块固定为 Z_{x_R} （位于当前坐标 $x_R \in X$ ），并将其命名为“参考块”。

1) 步骤1(a)i和步骤1(a)ii：分组和协同性硬阈值：

如II-B节中所述，我们通过在噪声图像 z 内进行块匹配来实现分组，仅有那些相对于参考距离的距离（即不相似度）小于固定阈值的块会被认为是相似的并且被分组。具体的，我们将“L2距离”作为不相似度的量度指标。

理想情况下，如果真实图像 y 可用，则块距离可以由下述公式（1）计算。

$$d^{ideal}(Z_{x_R}, Z_x) = \frac{\|Y_{x_R} - Y_x\|_2^2}{(N_1^{ht})^2} \quad (1)$$

其中 $\|\cdot\|_2$ 表示L2范数，块 Y_{x_R} 和 Y_x 分别位于 y 中的 x_R 和 $x \in X$ 处。但是，只有噪声图像 z 可用，并且距离只能由噪声块 Z_{x_R} 和 Z_x 通过下述公式计算。

$$d^{noisy}(Z_{x_R}, Z_x) = \frac{\|Z_{x_R} - Z_x\|_2^2}{(N_1^{ht})^2} \quad (2)$$

如果块 Z_{x_R} 和 Z_x 不重叠，该距离是非中心卡方随机变量，均值和方差分别为

$$\begin{aligned} E\{d^{noisy}(Z_{x_R}, Z_x)\} &= d^{ideal}(Z_{x_R}, Z_x) + 2\sigma^2 \\ \text{var}\{d^{noisy}(Z_{x_R}, Z_x)\} &= \frac{8\sigma^4}{(N_1^{ht})^2} + \frac{8\sigma^2 d^{ideal}(Z_{x_R}, Z_x)}{(N_1^{ht})^2} \end{aligned} \quad (3)$$

方差随 $\mathcal{O}(\sigma^4)$ 的复杂度逐渐增长。因此，对于相对较大的 σ 或较小的 N_1^{ht} ，不同 $d^{noisy}(Z_{x_R}, Z_x)$ 的概率密度可能会严重重叠，这将导致错误的分组。也就是说，具有比阈值大的理想距离的块被相似地匹配，而具有这样的距离的小块会被忽略。

为了避免上述问题，我们建议使用粗略的预滤波来测量块距离。我们可以通过以下方式来实现这种预滤波：在两个块上应用归一化的2D线性变换，然后对所获得的系数进行硬阈值化，即可得到

$$d(Z_{x_R}, Z_x) = \frac{\|Y'(\mathcal{T}_{2D}^{ht}(Z_{x_R})) - Y'(\mathcal{T}_{2D}^{ht}(Z_x))\|_2^2}{(N_1^{ht})^2} \quad (4)$$

其中， Y' 是阈值 $\lambda_{2D}\sigma$ 的硬阈值运算符， \mathcal{T}_{2D}^{ht} 表示归一化的2D线性变换。

使用 d 距离，块匹配（BM）的结果是一个集合，其中包含类似于 Z_{x_R} 的块的坐标，

$$S_{x_R}^{ht} = \{x \in X: d(Z_{x_R}, Z_x) \leq \tau_{match}^{ht}\} \quad (5)$$

其中 τ_{match}^{ht} 匹配是两个块被视为相似的最大 d 距离。参数 τ_{match}^{ht} 由理想差的可接受值的确定性推测所选择，主要用于忽略信号的噪声成分。

显然， $d(Z_{x_R}, Z_x) = 0$ 意味着 $|S_{x_R}^{ht}| \geq 1$ ，其中 $|S_{x_R}^{ht}|$ 贡献了 $S_{x_R}^{ht}$ 的基数。在获得了 $S_{x_R}^{ht}$ 后，通过堆叠匹配的噪声块 $Z_{x \in S_{x_R}^{ht}}$ 以形成大小为 $N_1^{ht} \times N_1^{ht} \times |S_{x_R}^{ht}|$ 的3D阵列来形成一个组，我们称之为 $\mathbf{Z}_{S_{x_R}^{ht}}$ 。匹配的块通常可以重叠。我们不限制块的排列顺序，这将在IV-B节中讨论。

$Z_{x \in S_{x_R}^{ht}}$ 的协同滤波是通过3D变换域中的硬阈值实现的。本文选用的归一化3D线性变换记作 \mathcal{T}_{3D}^{ht} ，它将兼具在II-D节中讨论的两种相关类型的优点，使真实信号组 $Y_{S_{x_R}^{ht}}$ 获得良好的稀疏性。这样可以通过硬阈值进行有效的噪声衰减，然后进行逆变换，以产生块估计值的3D数组

$$\hat{Y}_{S_{x_R}^{ht}} = \mathcal{T}_{3D}^{ht^{-1}} \left(\gamma \left(\mathcal{T}_{3D}^{ht} \left(Z_{S_{x_R}^{ht}} \right) \right) \right) \quad (6)$$

其中， γ 是具有阈值 $\lambda_{3D}\sigma$ 的硬阈值操作符。数组 $\hat{Y}_{S_{x_R}^{ht}}$ 包括 $|S_{x_R}^{ht}|$ 个堆叠的块估计值 $\hat{Y}_x^{ht, x_R}, \forall x \in S_{x_R}^{ht}$ 。在 \hat{Y}_x^{ht, x_R} 中，下标 x 表示此块估计的位置，上标 x_R 表示参考块。

2) 步骤2(a)i和2(a)ii: 分组和协同维纳滤波:

给定在步骤1b中获得的真实图像的初步估计值 \hat{y}^{basic} ，可以通过在该初步估计值内进行分组并进行协同经验维纳滤波来改善降噪效果。

因为按照假设， \hat{y}^{basic} 中的噪声已经被显著地衰减掉了，因此我们将基于阈值的d距离（4）替换为在基本估算值内计算出的归一化的平方距离（L2距离）。这是理想距离（1）的近似值。因此，匹配块的坐标是集合 $S_{x_R}^{wie}$ 的元素。

$$S_{x_R}^{wie} = \left\{ x \in X: \frac{\| \hat{Y}_{x_R}^{basic} - \hat{Y}_x^{basic} \|_2^2}{(N_1^{wie})^2} < \tau_{match}^{wie} \right\} \quad (7)$$

我们使用集合 $S_{x_R}^{wie}$ 去规范两个组，一个来自初步估计而另一个来自观察噪声：

- $\hat{Y}_{S_{x_R}^{wie}}^{basic}$ 由初步估计块 $\hat{Y}_{x \in S_{x_R}^{wie}}^{basic}$ 堆叠构成
- $Z_{S_{x_R}^{wie}}$ 由噪声块 $Z_{x \in S_{x_R}^{wie}}$ 堆叠构成

我们从基础组估计的3D变换系数的能量中定义经验维纳收缩系数为

$$W_{S_{x_R}^{wie}} = \frac{\left| \mathcal{T}_{3D}^{wie} \left(\hat{Y}_{x \in S_{x_R}^{wie}}^{basic} \right) \right|^2}{\left| \mathcal{T}_{3D}^{wie} \left(\hat{Y}_{x \in S_{x_R}^{wie}}^{basic} \right) \right|^2 + \sigma^2} \quad (8)$$

然后，将 $Z_{S_{x_R}^{wie}}$ 的协同维纳滤波实现为噪声数据的3D变换系数 $\mathcal{T}_{3D}^{wie} \left(Z_{S_{x_R}^{wie}} \right)$ 与维纳收缩系数 $W_{S_{x_R}^{wie}}$ 的逐元素相乘。随后，逆变换 $\mathcal{T}_{3D}^{wie^{-1}}$ 产生一组估计

$$\hat{Y}_{S_{x_R}^{wie}}^{basic} = \mathcal{T}_{3D}^{wie^{-1}} \left(W_{S_{x_R}^{wie}} \mathcal{T}_{3D}^{wie} \left(Z_{S_{x_R}^{wie}} \right) \right) \quad (9)$$

该组包括位于匹配位置 $x \in S_{x_R}^{wie}$ 处的块估计值 \hat{Y}_x^{wie, x_R} 。

B. 步骤 1b 和 2b：通过聚合实现的全局估算

在步骤1a和2a中分别获得了块估计值 $\hat{Y}_{S_{x_R}^{ht}}^{ht,x_R}$ 和 $\hat{Y}_{S_{x_R}^{wie}}^{wie,x_R}$, $\forall x_R \in S_{x_R}^{ht}$, 然而两个集合都是真实图像的不完整表示, 因为通常块估计区域是可以重叠的。另外, 可以在完全相同的坐标上定位多个块估计, 例如, $\hat{Y}_{x_b}^{ht,x_a}$ 和 $\hat{Y}_{x_b}^{ht,x_b}$ 都位于 x_b , 但分别在处理参考块 x_a 和处理参考块 x_b 时获得。可以预期, 在有大量重叠的块估计的区域, 即某个块附近有很多其他相似块的区域中, 信号的表示大概率上是不完整的。因此, 该方法的冗余度取决于分组操作, 也取决于具体图像。

为了分别在步骤1b和2b中计算真实图像的基本估计值和最终估计值, 我们聚合了相应的块估计值 $\hat{Y}_{x \in S_{x_R}^{ht}}^{ht,x_R}$ 和 $\hat{Y}_{x \in S_{x_R}^{wie}}^{wie,x_R}$, $\forall x_R \in X$ 。该聚合操作通过在块估计区域重叠的那些像素位置处进行加权平均来执行。权重的选择在以下小节中讨论。

1) 聚合权重:

通常, 逐块估计在统计上是相关的, 有偏差的, 并且每个像素的方差都不同。但是, 非常有必要考虑上述因素影响。与工作[6]和[9]相似, 我们发现, 一个令人满意的加权聚合选择(权重分配方案), 是与相应的块估计的总样本方差成反比的选择(权重分配方案)。也就是说, 噪声较大的块估计值应被赋予较小的权重。如果 $\mathbf{Z}_{S_{x_R}^{ht}}$ 和 $\mathbf{Z}_{S_{x_R}^{wie}}$ 组中的加性噪声是独立的, 则相应的组估计值(6)和(9)中的总样本方差分别等于 $\sigma^2 N_{har}^{x_R}$ 和 $\sigma^2 \|\mathbf{W}_{S_{x_R}^{wie}}\|_2^2$, 其中 $N_{har}^{x_R}$ 是经过硬阈值处理后保留的(非零)系数的数量, $\mathbf{W}_{S_{x_R}^{wie}}$ 是维纳滤波器系数(8)。基于此, 在每个 $x_R \in X$ 的步骤1b中, 我们为组估计 $\hat{Y}_{x \in S_{x_R}^{ht}}^{ht,x_R}$ 分配权重

$$w_{x_R}^{ht} = \begin{cases} \frac{1}{\sigma^2 N_{har}^{x_R}}, & N_{har}^{x_R} \geq 1 \\ 1, & otherwise \end{cases} \quad (10)$$

相似的, 在每个 $x_R \in X$ 的步骤2b中, 我们为组估计 $\hat{Y}_{x \in S_{x_R}^{wie}}^{wie,x_R}$ 分配权重

$$w_{x_R}^{wie} = \sigma^{-2} \|\mathbf{W}_{S_{x_R}^{wie}}\|_2^{-2} \quad (11)$$

我们指出, 只有在构成该组的噪声块彼此不重叠时, 才能实现组中噪声的独立性。因此, 一方面确保独立性的高昂成本将对BM造成严重的限制, 即仅允许在非重叠的块之间进行匹配。另一方面, 如果考虑到可能的重叠, 则以 $\mathcal{T}_{3D}^{ht}(\mathbf{Z}_{S_{x_R}^{ht}})$ 或 $\mathcal{T}_{3D}^{wie}(\mathbf{Z}_{S_{x_R}^{wie}})$ 为单位的每个变换系数的独立方差计算工作就会变得令人难以置信的复杂, 因为需要考虑相应变换系数中的协方差项。考虑到上述因素, 在我们的算法中, 我们使用重叠块, 不考虑协方差。因此, (10)和(11)中提出的权重仅与相应的总样本方差大致成反比。

1) 加权平均聚合:

全局初步估计值 \hat{y}^{basic} 由步骤1a中获得的块估计值 $\hat{Y}_{x \in S_{x_R}^{ht}}^{ht,x_R}$ 的加权平均得到。其中的权重 $w_{x_R}^{ht}$ 定义于(10), 即

$$\hat{y}^{basic}(x) = \frac{\sum_{x_R \in X} \sum_{x_m \in S_{x_R}^{ht}} w_{x_R}^{ht} \hat{y}_{x_m}^{ht, x_R}(x)}{\sum_{x_R \in X} \sum_{x_m \in S_{x_R}^{ht}} w_{x_R}^{ht} \chi_{x_m}(x)}, \forall x \in X \quad (12)$$

其中 $\chi_{x_m}: X \rightarrow \{0,1\}$ 是位于 $x_m \in X$ 处的块的方形支撑的特征函数，而块估计 $\hat{y}_{x_m}^{ht, x_R}$ 在其支撑区域外进行了零填充。

全局最终估计 \hat{y}^{final} 由公式(12)计算得到，其中 \hat{y}^{basic} ， $\hat{y}_{x_m}^{ht, x_R}(x)$ ， $S_{x_R}^{ht}$ 和 $w_{x_R}^{ht}$ 分别由 \hat{y}^{final} ， $\hat{y}_{x_m}^{wie, x_R}(x)$ ， $S_{x_R}^{wie}$ 和 $w_{x_R}^{wie}$ 替代。

IV. 快速且高效的实现方案

上一节中介绍的方法的直接实现对计算的要求很高。为了实现一种实用高效的算法，我们施加了约束并利用了以下列表中列出的某些权宜之计。

减少处理块的数量。

- 与其在水平和垂直方向上使用 $N_{step} \in \mathbb{N}$ 个像素的步距，不如将其滑动到每个下一个参考块。

因此，参考块的数量从大约 $|X|$ 减少到 $\frac{|X|}{N_{step}^2}$ 步

减少分组的复杂性。

- 通过在分组的块数上设置上限 $N_2 \in \mathbb{N}$ 来限制组的最大大小；即确保 $|S_{x_R \in X}| \leq N_2$ 。
- 在以当前处理的坐标 $x_R \in X$ 为中心的受限制的 $N_S \times N_S$ 的局部邻域中搜索候选匹配块。
- 为了进一步加快BM的速度，我们使用预测搜索，即搜索邻域是非矩形的并取决于先前匹配的块。我们形成这样一个邻域，即以先前匹配的坐标为中心的 $N_{PR} \times N_{PR}$ （其中 $N_{PR} \ll N_S$ ）的并集在处理图像的方向上相应地偏移了 N_{step} 步，例如在水平方向上进行逐格扫描。但是，对于每个第 N_{FS} 个已处理块，我们仍在较大的 $N_S \times N_S$ 邻域中进行穷举搜索BM。特别地， $N_{FS} = 1$ 意味着仅在 $N_S \times N_S$ 的区域中使用穷举搜索。

降低应用变换的复杂度。

- 将变换 \mathcal{T}_{3D}^{ht} 和 \mathcal{T}_{3D}^{wie} 限制为可分离变换的类别，并沿匹配的块分别使用 \mathcal{T}_{2D}^{ht} 和 \mathcal{T}_{2D}^{wie} ，沿着组的第三维进行一维变换 \mathcal{T}_{1D} （沿着块堆叠方向）。
- 光谱 $\mathcal{T}_{2D}^{ht}(\mathbf{Z}_x)$ 、 $\mathcal{T}_{2D}^{wie}(\mathbf{Z}_x)$ 以及 $\mathcal{T}_{2D}^{wie}(\hat{y}_x^{basic})$ 是针对当前处理的坐标周围的 $N_S \times N_S$ 附近的每个块预先计算的。接下来，它们被重新用于随后的参考块，这些参考块的 $N_S \times N_S$ 邻域与当前参考块重叠。因此，这些变换对于每个处理的坐标仅计算一次。例如，它们不会在(4)中每次都重新计算。此外，在(6)，(8)和(9)中，我们分别通过在预先计算的 \mathcal{T}_{3D}^{ht} 和 \mathcal{T}_{3D}^{wie} 变换块上分别应用 \mathcal{T}_{1D} 来计算正向 \mathcal{T}_{2D}^{ht} 和 \mathcal{T}_{2D}^{wie} 变换。

高效地实现聚合

- 首先，如步骤1(a)ii和2(a)ii所述，对获得的块估计值进行加权并在缓冲区中累加（保持图像的尺寸）。同时，将相应的权重累积在另一个缓冲区的相同位置。然后，如步骤1b和2b中所述，聚合操作(12)最终通过两个缓冲器之间的简单的元素除法得以实现。

消除边缘效应

- 使用 $N_1 \times N_1$ 的Kaiser窗口（带有参数）作为(12)中权重的一部分，以减少使用某些2D变换（例如2D DCT、2D DFT以及周期小波）时可能出现的边框效果。

A. 复杂度

该算法的时间复杂度为 $\mathcal{O}(|X|)$ ，因为所有参数都是固定的，所以时间复杂度线性地取决于输入图像的大小。

鉴于前一小节中介绍的限制，在不利用预测搜索BM的情况下，每个像素的操作次数约为

$$3C_{T_{2D}} + \frac{2(N_1^2 + N_2)N_S^2}{N_{\text{step}}^2} + \frac{3(N_2C_{T_{2D}} + N_1^2C_{T_{1D}})}{N_{\text{step}}^2}$$

其中为了简单起见，我们在其中省略了参数/运算符的上标“ht”和“wie”，以及：

- 第一项加数来自每个滑块（在 $N_S \times N_S$ 的邻域内）预先计算的 T_{2D}
- 第二项来自 $N_S \times N_S$ 邻域中通过穷举搜索BM得到分组，并且
- 第三项加数来自变换 T_{3D} ，它是由可分离的 T_{1D} 和 T_{2D} 组成的

以上， C_T 表示变换 T 所需的算术运算数，它取决于快速算法的可用性、可分离性等属性。例如可以通过快速傅立叶变换算法有效地计算DFT，并且可以使用迭代滤波器组有效地实现二进小波分解。

通过利用快速可分离变换和预测搜索BM，我们可以显着降低算法的复杂性。

B. 参数选择

我们研究了所提出算法的各种变换选择和参数选择。根据实验结果，我们建议将参数配置方案分为两种，即“常规”和“快速”，如表1所示。这两种配置方案的主要特征如下。

- 常规配置方案。此配置方案在计算复杂度和降噪性能之间提供了一种合理的折衷方案。根据噪声级别分为两种情况：
 $\sigma \leq 40$ 噪声不太严重，不会影响分组的准确性，因此通过设置 $\lambda_{2D} = 0$ 禁用d距离(4)中的阈值设置，并且使用相对较小的块大小，即 $N_1^{ht}, N_1^{wie} = 8$ 。
 $\sigma > 40$ 对应于高噪声水平，因此使用 $\lambda_{2D} = 2$ 来提高分组的准确性，并使用更大的块大小，即 $N_1^{ht} = 12, N_1^{wie} = 11$ 。
- 快速配置方案。以降低降噪性能为代价提供较低的计算复杂度。它利用了建议的快速预测搜索BM（不同于仅使用穷举搜索BM的常规配置方案）。

图9说明了当 $\sigma > 40$ 时对d距离使用阈值化（ $\lambda_{2D} = 2$ ）和更大的块大小的好处。这部分将在第VI节中进行讨论。

表 1 快速配置与常规配置的参数对照表

		Fast Profile	Normal Profile	
			$\sigma \leq 40$	$\sigma > 40$
Approx. exec. time for a 256×256 grayscale image on 1.5 GHz Celeron M		0.7 sec	4.1 sec	5.8 sec
Parameters for Step 1 (ht)	T_{2D}^{ht}	2D-Bior1.5	2D-Bior1.5	2D-DCT
	N_1^{ht}	8	8	12
	N_2^{ht}	16	16	16
	N_{step}^{ht}	6	3	4
	N_S^{ht}	25	39	39
	N_{ES}^{ht}	6	1	1
	N_{PR}^{ht}	3	-	-
	β^{ht}	2.0	2.0	2.0
	λ_{2D}	0	0	2.0
	λ_{3D}	2.7	2.7	2.8
	τ_{match}^{ht}	2500	2500	5000
Parameters for Step 2 (wie)	T_{2D}^{wie}	2D-DCT	2D-DCT	2D-DCT
	N_1^{wie}	8	8	11
	N_2^{wie}	16	32	32
	N_{step}^{wie}	5	3	6
	N_S^{wie}	25	39	39
	N_{ES}^{wie}	5	1	1
	N_{PR}^{wie}	2	-	-
	τ_{match}^{wie}	400	400	3500
	β^{wie}	2.0	2.0	2.0
Common	T_{1D}	1D-Haar	1D-Haar	1D-Haar

表 2 输出 PSNR (dB)对使用的转换的依赖关系。与 T_{2D}^{ht} 对应的列包含基础估计 \hat{y}^{basic} 的 PSNR 结果，所有其他列包含最终估计 \hat{y}^{final} 的结果。观测噪声的 $\sigma=25$ 。

Transform	Boats			Lena		
	T_{2D}^{ht}	T_{2D}^{wie}	T_{1D}	T_{2D}^{ht}	T_{2D}^{wie}	T_{1D}
Haar	29.31	29.84	29.91	31.24	31.93	32.08
Db2	29.22	29.83	29.90	31.19	31.97	32.06
Db4	29.34	29.88	29.89	31.31	32.01	32.06
Db6	29.30	29.86	29.89	31.28	31.98	32.06
Bior1.3	29.42	29.87	29.90	31.35	31.96	32.06
Bior1.5	29.43	29.88	29.90	31.37	31.97	32.06
WHT	29.22	29.84	29.88	31.24	32.00	32.07
DCT	29.35	29.91	29.88	31.42	32.08	32.07
DST	29.33	29.91	29.79	31.36	31.97	31.92
DC+rand	29.07	29.75	29.88	31.06	31.88	32.06
DC-only	-	-	28.03	-	-	30.65

为了说明变换 T_{2D}^{ht} 、 T_{2D}^{wie} 和 T_{1D} 的选择对去噪性能的影响，我们在表2中列出了一些实验结果。如前文所述，在我们方法的步骤1和2中使用的3D变换 T_{3D}^{ht} 和 T_{3D}^{wie} 是分别由 T_{2D}^{ht} 、 T_{2D}^{wie} 与 T_{1D} 计算得到的。并且， T_{2D}^{ht} 和 T_{2D}^{wie} 都是一维变换的可分离成分，例如表中指定的那些变换。我们的实验中使用了以下规范化变换：

- DST, DCT: 离散正弦和余弦变换，
- WHT: 沃尔什-阿达玛变换，以及
- 一些全二进小波分解，使用：
 - Dbp: 带 p 个消失矩的Daubechies小波，其中 $p = 1,2,4,6$ ；当 $p = 1$ 时，它与Haar小波重合
 - Bior1. N_r : 双正交样条小波，其中分解和重构小波函数的消失矩分别为1和 N_r
- DC+rand:一种标准正交变换，它有一个DC基元，它的其余基元具有随机性，即是通过一个高

斯白过程的实现的标准规范化得到的。

此外，我们对 \mathcal{T}_{1D} 额外实验了元素平均法，即只在第三维中保留DC(丢弃所有其他变换系数)，因此它的名字是“DC-only”。在这种情况下，所有分组的块都是通过元素平均来估计的，就像第II-C节中的示例一样。

在表2中，我们给出了各种 \mathcal{T}_{2D}^{ht} 、 \mathcal{T}_{2D}^{wie} 和 \mathcal{T}_{1D} 对应的结果。表中的实验在所有情况下都使用了常规配置方案的参数，这里只更改了表中各列所对应的具体转换函数。粗体显示的结果对应于性能最好的转换。我们观察到， \mathcal{T}_{2D}^{ht} 、 \mathcal{T}_{2D}^{wie} 的选择对去噪性能没有显著影响。即使是“DC+rand”变换，其除DC以外的基本元素都是随机的，其PSNR也只在0.1~0.4dB范围内有小幅下降。这可以解释为协同过滤主要依赖于 \mathcal{T}_{1D} 来挖掘分组块之间的片段间相关性。估计能力不显著依赖于 \mathcal{T}_{2D}^{ht} 、 \mathcal{T}_{2D}^{wie} 的能量压缩能力。从这个意义上说，片段间的相关性比片段内的相关性更为重要。

现在让我们关注表II中各种 \mathcal{T}_{1D} 变换对应的结果。与其他标准变换和“DC+rand”相比，DST的性能要差一些。我们认为，出现这种差异的原因是DST中缺少DC基元——与所有其他具有此元素的变换相反。为什么 \mathcal{T}_{1D} 的DC很重要?简单地说，这是因为DC基元捕获了一个组的第3维元素之间的相似性。由于分组的块相似，因此它们对应的2D光谱也是如此，DC项反映了这种相似性。

但是，正如第二节所讨论的，不可能存在完全匹配的块。为了避免生成仅包含参考块的无效组，我们在式(5)和式(7)中使用了严格的正阈值。由式(3)可以看出，噪声对块距离的精度有影响，在实践中，这意味着在一个组中可能存在这样的块，对于这些块，底层的真实信号 Y_x 比 τ_{match}^{ht} 离 Y_{x_R} 远得多。因此，单独的DC元素无法捕获分组块之间的潜在差异。这一点可以从 \mathcal{T}_{1D} 的“DC-only”的糟糕结果中得到证实。在任何其他变换中，甚至是“DC+rand”中的随机变换中，额外的基元素能都带来比“DC-only”更多的性能改进。

实验发现，组块的排序对算法的估计能力没有显著影响。 \mathcal{T}_{1D} 的“DC+rand”结果证实了这一点，该结果与任何其他正交变换(结构化的、非随机的)的结果相同。对于此变换，顺序与DC无关，只与随机生成的其他基元素相关。因此，我们可以得出结论，组中块的顺序不影响最终结果。考虑到这一点，再加之我们的实现中BM已经生成了按块距离排序的块集合，所以我们使用的正是这种排序。当然，在一个组中总是有参考块，因为到它到自己的距离几乎等于零。

请注意，即使一个组是基于与给定参考块的相似性来构造的，这并不意味着这个参考块比其他任何块都更适合由这个组来表示。例如，可能发生的情况是，所有匹配的块(除了参考块)都与参考块非常不相似，但彼此之间非常相似。这样一个群体可以被称为“不平衡的”。

对于 \mathcal{T}_{1D} ，我们选择Haar全二进分解，因为它可以使用2抽头分析/合成滤波器通过迭代滤波器组有效地实现。要应用这种标准正交全二进分解，变换大小必须是2的幂。我们通过将 S_x^{ht} (5)和 S_x^{wie} (7)的元素数限制为小于或等于 S_x^{ht} 和 S_x^{wie} 中原始元素数的2的最大乘方来达到这一要求。

V. 拓展到彩色图像去噪

我们考虑一个自然的RGB图像，在它的每个通道中都有加性的独立同分布零均值高斯噪声。让一个亮度-色度变换应用在这样一个噪声图像中，其中亮度通道用Y表示，色度通道用U和V表示。此类变换的典型示例是YCbCr和对应的色彩空间反变换，其变换矩阵分别是

$$A_{YCbCr} = \begin{bmatrix} 0.30 & 0.59 & 0.11 \\ -0.17 & -0.33 & 0.50 \\ 0.50 & -0.42 & -0.08 \end{bmatrix}, A_{opp} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} & \frac{-\sqrt{2}}{3} & \frac{1}{3\sqrt{2}} \end{bmatrix}. \quad (13)$$

由于底层自然彩色图像的特性，例如其R、G和B通道之间的高度相关性，可以观察到以下情况：

- Y的信噪比高于U和V (R、G、B通道的去相关)；
- Y包含了大部分有价值的信息(边缘、阴影、物体、纹理图案等)；
- U和V主要包含低频信息(这些信道通常来自未充分采样的数据)；
- 在图像等亮度区域上，U和V通道的对应区域不会发生剧烈变化；

用于彩色图像去噪的直接扩展是将已提出的灰度去噪方法分别应用于Y、U和V通道中的每个通道。然而，由于分组操作对噪声水平很敏感，这种朴素的方法会因为色度中的低信噪比而受到影响。因为适当的分组对于我们的方法的有效性而言是至关重要的，因此我们建议仅对亮度Y执行一次分组，并且在色度U和V上应用协作过滤时，重复使用完全相同的分组。使用分别在步骤1(a)和2(a)中的公式(5)和(7)找到Y所对应的分组块坐标集合，再在两个U和V通道上重用；使用这些集合，分别在三个通道上执行协同过滤（步骤1(a)ii和步骤2(a)ii）和聚合（步骤1b和2b）。对色度的分组约束是基于这样的假设：如果两个色块的亮度是彼此相似的，那么它们的色度也是彼此相似的。此外，考虑到通过块匹配进行分组需要BM3D总执行时间大约一半的时间，与在三个通道上分别应用灰度BM3D相比，分组约束将使计算减少大约三分之一。

VI. 实验结果

在本节中，我们介绍和讨论了用所开发的算法得到的实验结果；灰度版本以块匹配和3D筛选(BM3D)命名，相应的彩色版本缩写为C-BM3D。对于所有的实验，我们使用的Matlab代码可以在<http://www.cs.tut.fi/~foi/gcf.bm3d>找到。在这个网站上，我们还提供了进一步的结果和我们在实验中使用的原始和去噪的测试图像。除非另有说明，否则我们对BM3D和C-BM3D都使用表I中的“常规配置”的参数。

A. 灰度图像去噪

表 3 灰度图像去噪：本文提出的 BM3D 算法输出图像的 PSNR(dB)。

σ / PSNR	<i>C.man</i> 256 ²	<i>House</i> 256 ²	<i>Peppers</i> 256 ²	<i>Montage</i> 256 ²	<i>Lena</i> 512 ²	<i>Barbara</i> 512 ²	<i>Boats</i> 512 ²	<i>Fprint</i> 512 ²	<i>Man</i> 512 ²	<i>Couple</i> 512 ²	<i>Hill</i> 512 ²	<i>Lake</i> 512 ²
2 / 42.11	43.96	44.63	43.48	46.47	43.59	43.66	43.18	42.90	43.61	43.17	43.04	43.02
5 / 34.16	38.29	39.83	38.12	41.14	38.72	38.31	37.28	36.51	37.82	37.52	37.14	36.58
10 / 28.14	34.18	36.71	34.68	37.35	35.93	34.98	33.92	32.46	33.98	34.04	33.62	32.85
15 / 24.61	31.91	34.94	32.70	35.15	34.27	33.11	32.14	30.28	31.93	32.11	31.86	31.08
20 / 22.11	30.48	33.77	31.29	33.61	33.05	31.78	30.88	28.81	30.59	30.76	30.72	29.87
25 / 20.18	29.45	32.86	30.16	32.37	32.08	30.72	29.91	27.70	29.62	29.72	29.85	28.94
30 / 18.59	28.64	32.09	29.28	31.37	31.26	29.81	29.12	26.83	28.86	28.87	29.16	28.18
35 / 17.25	27.93	31.38	28.52	30.46	30.56	28.98	28.43	26.09	28.22	28.15	28.56	27.50
50 / 14.16	25.84	29.37	26.41	27.35	28.86	27.17	26.64	24.36	26.59	26.38	27.08	25.78
75 / 10.63	24.05	27.20	24.48	25.04	27.02	25.10	24.96	22.68	25.10	24.63	25.58	24.11
100 / 8.14	22.81	25.50	22.91	23.38	25.57	23.49	23.74	21.33	23.97	23.37	24.45	22.91

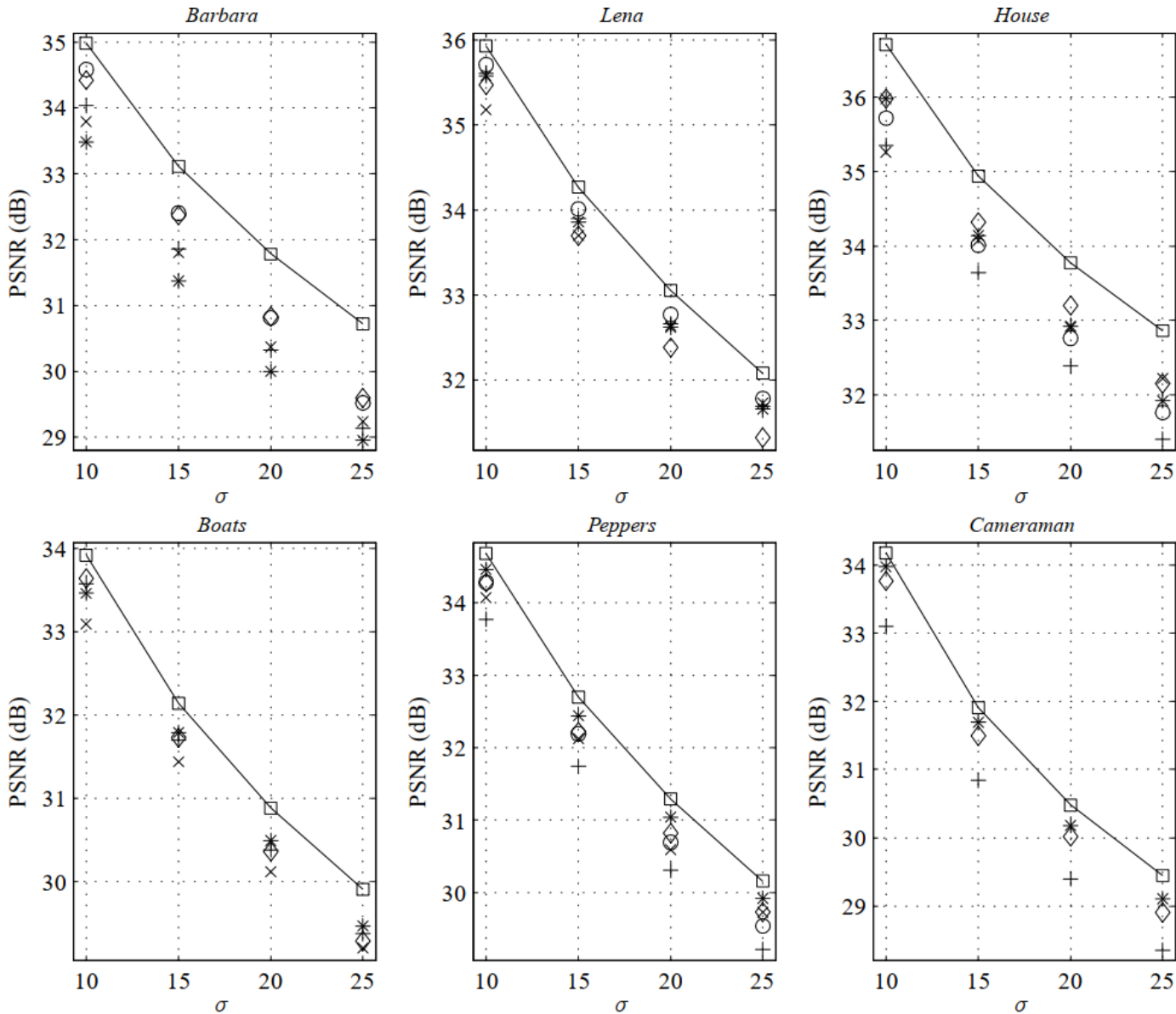


图 4 灰度图像去噪效果对比。对于以下方法，输出 PSNR 作为 σ 的函数：'——': 本文所提出的 BM3D, 'o': FSP+TUP BLS-GSM [4], '+': BLS-GSM [3], 'x': exemplar-based [11], '◇': K-SVD [8], '*': 逐点 SA-DCT [9]。(注意，船图像上的[4]方法和摄影师图像上的[4]、[11]方法的结果都不存在，因为它们既没有在相应的文章中展示，也没有公开这些方法的实现。)

表3给出了BM3D算法对一组标准灰度图像的输出结果及其PSNR。估计图像 \hat{y} 与原图 y 的PSNR标准计算公式为

$$\text{PSNR}(\hat{y}) = 10\log_{10} \left(\frac{255^2}{|X|^{-1} \sum_{x \in X} (y(x) - \hat{y}(x))^2} \right)$$

在图4中,我们将本文所提出的BM3D的输出结果的PSNR与最先进的BLS-GSM[3]、FSP+TUP BLS-GSM[4]、exemplar-based[11]、K-SVD[8]、逐点SA-DCT [9]技术的输出结果的PSNR进行了比较;对于K-SVD方法[8],我们展示了它的最佳结果,这些结果是使用在有噪图像上训练的自适应词典获得的。从图中可以看出,本文所提出的BM3D表现出了最好的性能,并且始终优于其他所有技术。特别地,由于这些图像包含结构化的对象(House中的边缘和Barbara中的纹理),可以进行非常有效地执行分组和协同过滤,因此BM3D在House和Barbara上有显著的改进。



图 5 带噪的房屋灰度图像($\sigma=25$)及其 BM3D 去噪估计(PSNR 32.86 dB)

在图5中,我们显示了嘈杂的($\sigma=25$)房屋图像和其相应的BM3D估计。在该测试图像中,相邻块之间的相似性很容易在均匀区域和沿着规则形状的结构中感知到,其中一些如图1所示。因此,这样的细节在估计中得到了很好的保留。

BM3D算法的降噪性能在图6中得到了进一步说明,其中显示了一些噪声较大的($\sigma=25$)测试图像的样例以及相应的降噪图像的样例。去噪后的图像保持良好:

- 均匀的区域和平滑的强度过渡 (Lena的脸颊, 以及其他图像的背景),
- 纹理和重复图案 (Barbara中的围巾), 以及
- 锐利的边缘和奇异点 (摄影师和船上物体的边界)。

图7中显示了一个极端噪声水平(例如 $\sigma=100$)的去噪示例。鉴于原始图像几乎完全被掩盖在噪声中,因此得出的估计图像显示出的结果算得上保留了合理的细节。尤为值得关注的是,估计图忠实地重建了重复的图案模式,例如衣服上的条纹。

在主观视觉质量方面,我们发现各种图像细节被很好地保存下来,同时引入的伪像很少。我们可以在图6、7和8中观察到这一点。Vansteenkiste等人进行的心理-视觉实验结果证实了我们的算法具有最先进的主观视觉质量[19]。在实验结果中,有35个评估人员将BM3D算法的初步版本[12]归类为8种评估的最新技术中最好的。评估的标准是感知噪声水平、感知模糊水平和总体视觉质量。此外,我们认为目前的BM3D算法的主观视觉质量(在细节保存方面)明显优于[19]中评估的初步版本。在图8中,

我们展示了由当前版本和BM3D算法的初始版本表示的图像。仔细观察就会发现，当前的BM3D所显示的图像具有更少的振铃伪影和更高的细节保留能力。

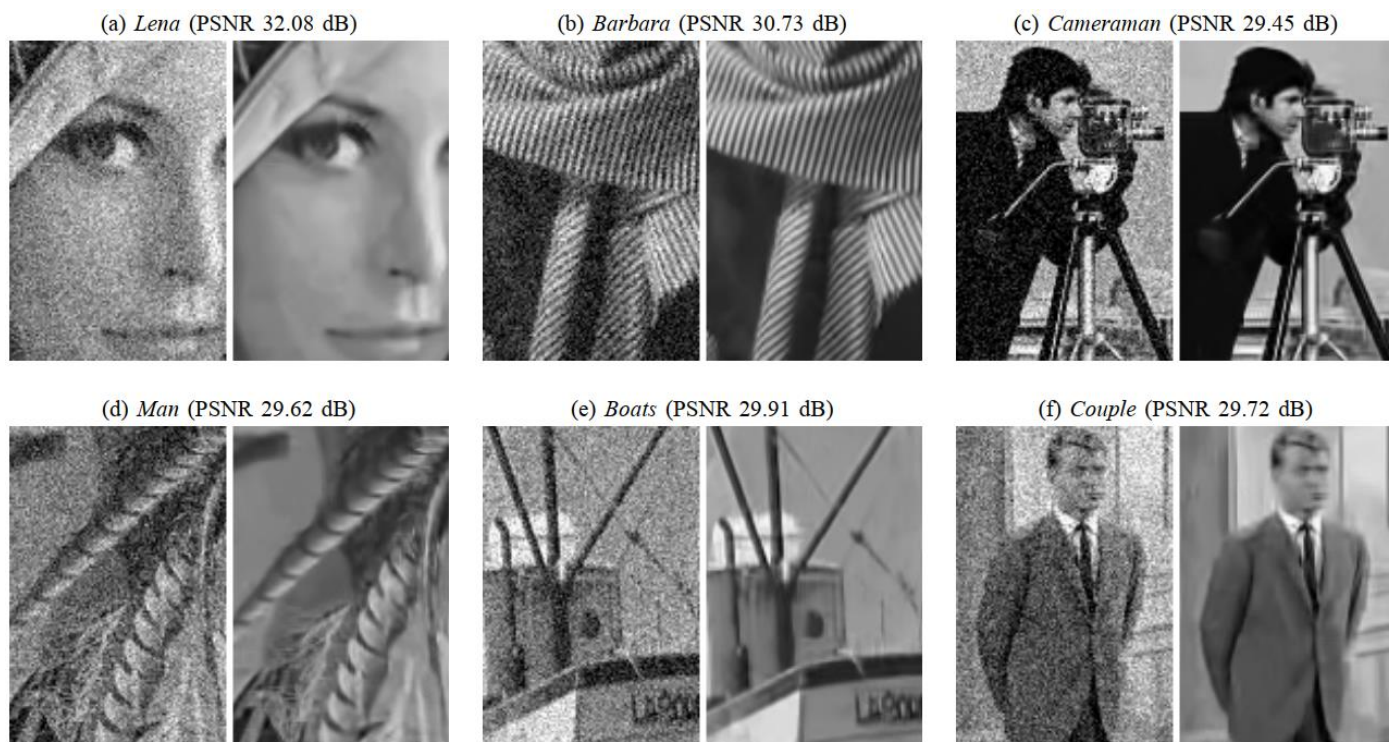


图 6 带噪的灰度图像的片段($\sigma=25$, PSNR 20.18 dB)和相应的 BM3D 估计。

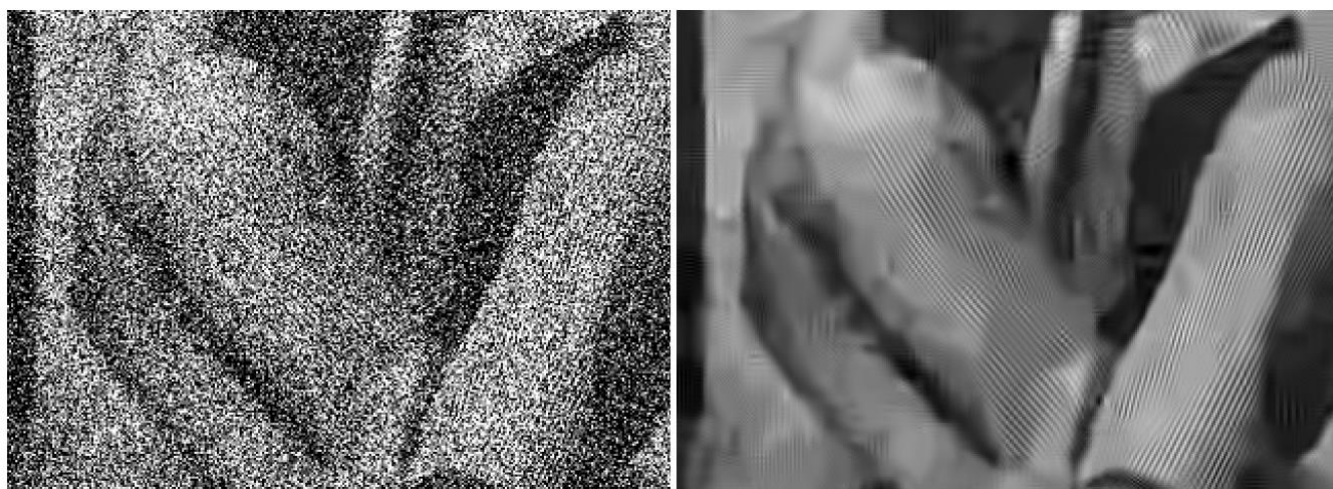


图 7 左图为的带噪的灰度图像 *Barbara* 的片段($\sigma=100$, PSNR 8.14dB); 右图为 BM3D 去噪估计中的对应片段 (PSNR 23.49dB)

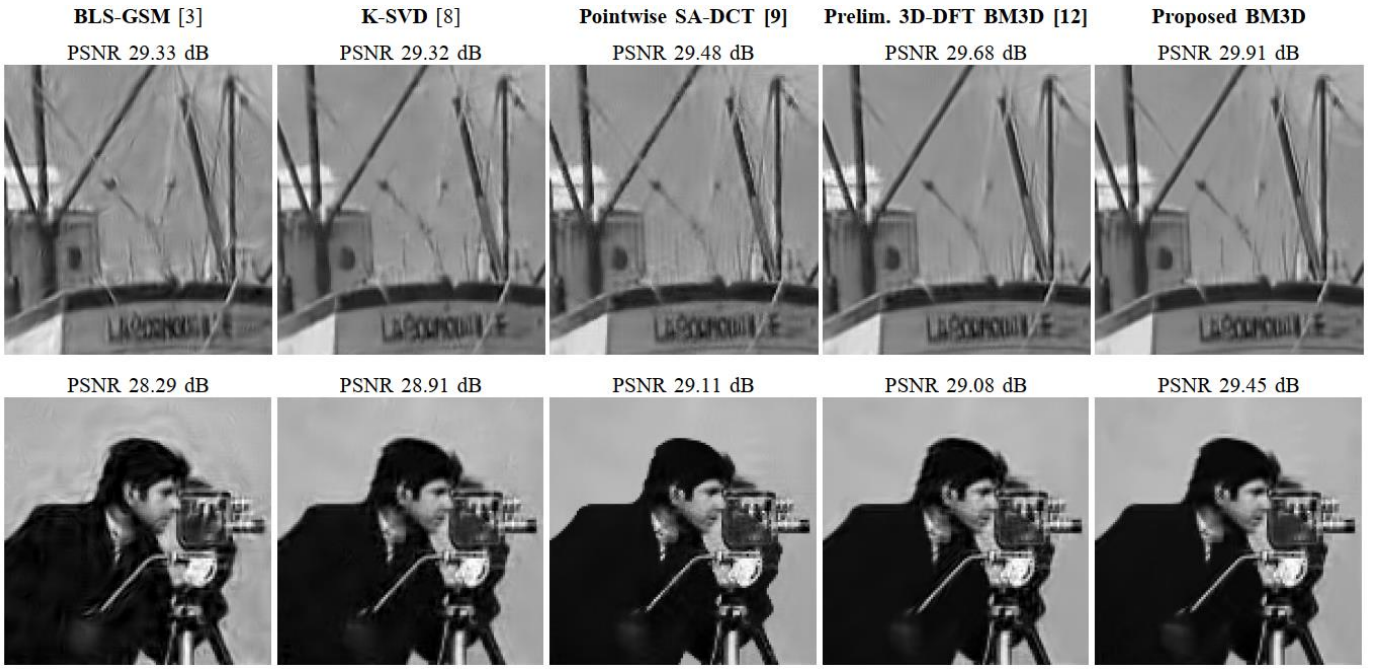


图 8 分别使用算法[3],[8],[9],[12]和本文提出的 BM3D 对灰度图像船（第一行）和摄影师（第二行）的进行去噪的效果片段，噪声图像的噪声水平为 $\sigma=25$ （噪声图像片段已在图 6 中列出）

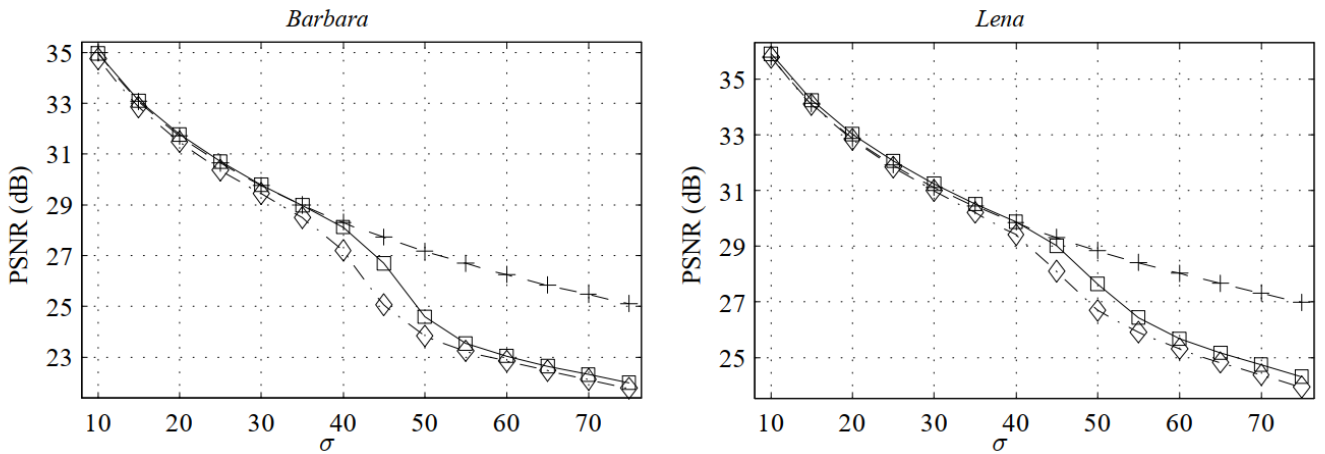


图 9 表 1 中不同配置方案下的 BM3D 算法 PSNR 输出对比图。请注意：'◇'代表“ $\sigma \leq 40$ ”分支下的常规配置方案，'+'代表快速配置方案，'+'代表“ $\sigma > 40$ ”分支下的常规配置方案。常规配置方案下的两个分支均在 σ 完整值域 [10,75]上进行了测试。

我们在图9中展示了快速配置方案和常规配置方案下的BM3D的PSNR表现。表1中的常规配置方案中的两种情况在 $\sigma \in [10, 75]$ 的区间上的表现被分别标记了出来，可以看到，对于“ $\sigma \leq 40$ ”的常规配置方案，由于错误分组的原因在 $\sigma = 40$ 附近相较于另一种方案出现了明显的PSNR下降。另一方面，可以观察到对于图中“ $\sigma > 40$ ”的方案，基于阈值的d距离(4)与相对较大的块尺寸 N_1 一起使用时，算法结果没有出现明显的PSNR下降。值得注意的是，对于高达中等水平的噪声，例如 $\sigma < 35$ ，快速配置方案和常规配置方案之间的PSNR差在0.05~0.2dB的范围内。相较于执行时间减少6倍(如表1所示)，这是一个可以接受的代价；更确切地说，对于在1.5 GHz赛扬M处理器上计算256×256的图像，其近似执行时间(去噪)从常规配置方案的4.1秒降至快速配置方案的0.7秒。BM3D算法允许通过改变 N_{step} 来实现进一步的复杂性/性能权衡。本文粗略地比较了一下，图4中考虑的其他方法的执行时间（在1.5GHz

赛扬M处理器上对256×256的图像进行去噪)分别为: BLS-GSM为22.1秒, SA-DCT滤波器为6.2秒, 在输入的噪声图像上训练自适应K-SVD需要9~30分钟(取决于 σ), 而使用训练好的字典进行滤波则需要25 – 120秒。在2Ghz奔腾IV处理器上测试时, 在[11]中报告的基于样本的方法的执行时间约为1分钟。我们没有测试FSP+TUP BLSGSM的执行时间, 然而考虑到它是一个两步的BLS-GSM的扩展, 其执行速度不会比BLS-GSM更快。

B. 彩色图像去噪

我们使用颜色空间反变换(13)和常规配置方案的参数对C-BM3D进行了实验。在所有的实验中, 我们都考虑了在R、G和B三个信道中均有方差为 σ^2 的独立同分布零均值高斯噪声的噪声图像。估计图像 \hat{y}_c 与原图 y_c 的PSNR标准计算公式为

$$\text{PSNR}(\hat{y}_c) = 10\log_{10} \left(\frac{255^2}{(3|X|)^{-1} \sum_{c=R,G,B} \sum_{x \in X} (y_c(x) - \hat{y}_c(x))^2} \right)$$

其中下标 $c \in \{R, G, B\}$ 表示颜色通道。

表 4 彩色图像去噪：本文提出的 C-BM3D 算法输出图像的 PSNR

σ / PSNR	<i>Lena</i>	<i>Peppers</i>	<i>Baboon</i>	<i>F16</i>	<i>House</i>
5 / 34.15	37.82	36.82	35.25	39.68	38.97
10 / 28.13	35.22	33.78	30.64	36.69	36.23
15 / 24.61	33.94	32.60	28.39	35.00	34.85
20 / 22.11	33.02	31.83	26.97	33.77	33.84
25 / 20.17	32.27	31.20	25.95	32.78	33.03
30 / 18.59	31.59	30.62	25.14	31.94	32.33
35 / 17.25	30.91	30.00	24.46	31.13	31.58
50 / 14.15	29.72	28.68	23.14	29.41	30.22
75 / 10.63	28.19	27.12	21.71	27.60	28.33

表 5 彩色图像去噪：本文所提出的 BM3D 和最近的两种最先进算法[9]和[20]的 PSNR 对比

Image	Method	Standard deviation σ			
		10	15	20	25
<i>Lena</i> 512 ² RGB	Proposed C-BM3D	35.22	33.94	33.02	32.27
	P.wise SA-DCT [9]	34.95	33.58	32.61	31.85
	ProbShrink-MB [20]	34.60	33.03	31.92	31.04
<i>Peppers</i> 512 ² RGB	Proposed C-BM3D	33.78	32.60	31.83	31.20
	P.wise SA-DCT [9]	33.70	32.42	31.57	30.90
	ProbShrink-MB [20]	33.44	32.05	31.12	30.35
<i>Baboon</i> 512 ² RGB	Proposed C-BM3D	30.64	28.39	26.97	25.95
	P.wise SA-DCT [9]	30.62	28.33	26.89	25.86
	ProbShrink-MB [20]	30.17	27.83	26.38	25.27

表4给出了所提出的C-BM3D算法对一些标准测试图像的输出的PSNR。表5给出了[20]和最近两种最先进的方法[9]的比较, 我们可以看到, 我们所提出的算法在三张测试图像上的表现都优于[9]和[20]。BM3D的高视觉质量可以从图10中得到体现, 图中显示了在高噪声($\sigma = 50$)的Lena图像下C-BM3D的估计结果。可以观察到帽子、尖锐边缘和光滑区域上的细节都被完好地保留了。对于256×256的RGB图像, C-BM3D在1.5 GHz赛扬M处理器上的执行时间大约为7.6秒。

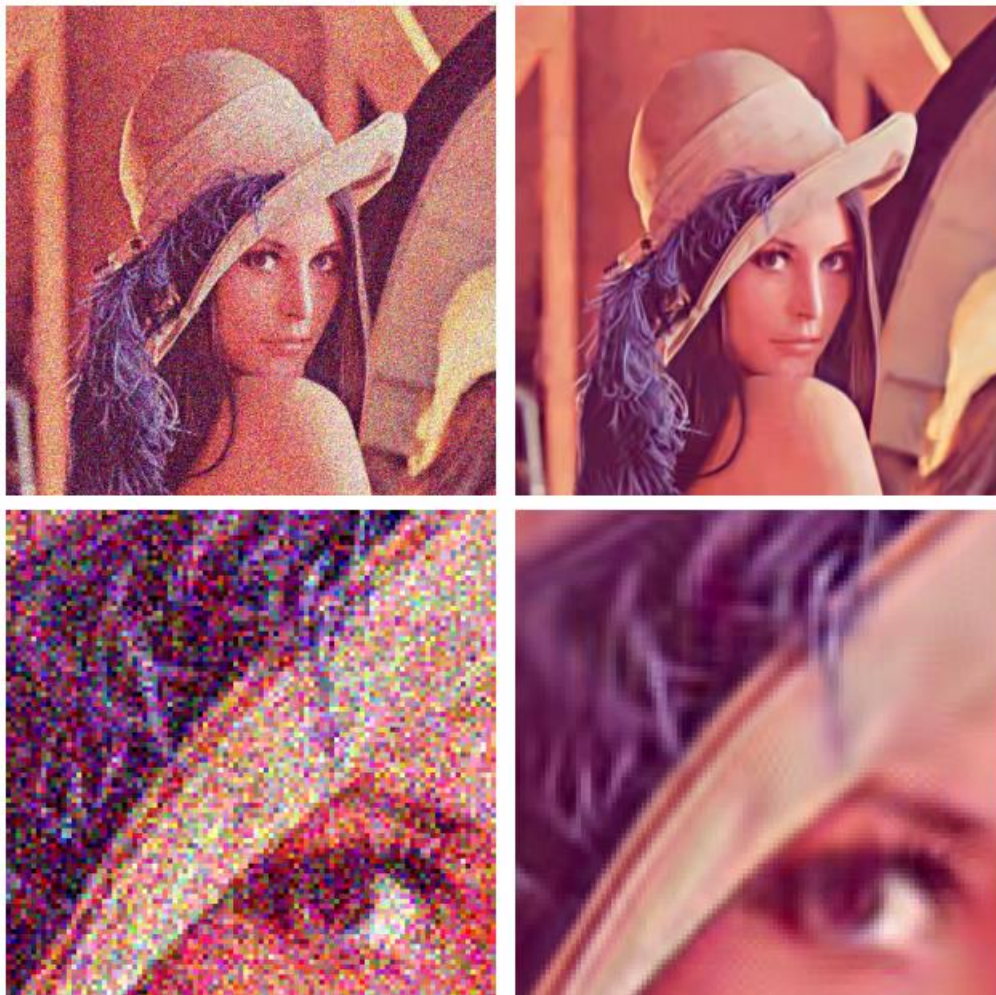


图 10 彩色图像去噪：左侧为带噪图像 Lena($\sigma=50$, PSNR 14.15dB)和其片段。右侧为 C-BM3D 算法的估计图 (PSNR 29.72dB)和其对应的片段。（图像的彩色版本可以在 IEEE Xplore 的文章中找到在线版本）

VII. 讨论

本文提出的方法是在局部逼近技术方面工作的发展。它从具有简单对称邻域的经典局部多项式逼近出发。自适应点大小变化的邻域是实现高效算法的第一步。下一步是实现基于自适应星型邻域的各向异性估计，该方法允许存在非对称的估计区域。对于自然图像相关的估计，这些估计方案的不对称性是设计的关键。最近的图书[21]中总结了这一发展。

突破性进展得益于引入了自适应阶的局部逼近。首先，它是根据具有不同窗口大小的标准正交变换完成的[22]。这些变换的谱的硬阈值的存在意味着逼近序列中的一些项会被自适应地丢弃，因此模型的阶数变得与数据相关[23]。点状自适应DCT滤波器[9]是在局部邻域中进行自适应阶数估计的最有效方法，该方法在具体统计规则定义的自适应形状邻域中计算正交变换。

本文给出了局部逼近方法发展的下一个关键点。空间自适应是通过选择与给定参考块相似的块组来实现的。因此，局部估计变成了非局部估计。所选的块被分组在三维阵列中，联合滤波，并在它们的采样处聚合。三维阵列中块的联合滤波是通过频谱项的收缩来实现的，因此再次采用阶数自适应估计的思想，但是以一种比较具体的方式。该方法的主要优点是非局部性和协同滤波。后者可以有效地

保留图像块中的局部特征，并且去噪效果非常好。

我们谨提及其他几位作者的工作，以澄清我们的贡献的背景，并说明本文的方法与其他类似方法的不同之处。

由于我们的方法与非局部估计器[10]和[11]在信号上基于同样的假设，因此有必要将这类技术与我们的方法进行比较。在非局部估计中使用的加权平均值对应于零阶多项式近似。它的有效性取决于对自适应权重的精致计算，而这具体取决于以估计像素为中心的图像块与平均过程中使用的图像块之间的相似性。我们的方法与此不同。我们通过使用一组更灵活的基函数（嵌入到变换中），使模型具有阶自适应的能力，并能够更有效地利用分组块之间的相似性。这是通过协同滤波实现的，它允许计算所有分组块的高阶估计值（而不仅仅是加权平均值）。

在[8]中提出的算法是从一个全局优化公式中推导出来的。将图像分割成一组重叠的块，并通过在每个块拟合最小复杂度模型来实现滤波。通过融合这些模型得到最终的图像估计。该算法十分出色的性能主要来自于利用训练得到的一组基函数(字典)。相比之下，我们的协同滤波本质上是不同的，因为硬阈值所产生的模型只在整个组上具有较低的复杂性。对于块估计和整体的图像而言，模型可以是高度复杂和冗余的，因为每个块可以进入许多组，因此可以参与许多协同估计。这种冗余提供了非常好的噪声衰减，并允许我们避免标准阈值方案中常见的伪影。因此，我们可以说，与[8]中的一些低复杂度建模不同，我们利用了具体的超完备的表示。

第二步中使用的协同维纳滤波和使用自适应权值的块估计的聚合是我们方法的主要特征。维纳滤波利用基本估计的功率谱对形成的组进行滤波。结果，估计明显优于第一步中使用的硬阈值。从表II中可以看出PSNR的提高（通过将“ \mathcal{T}_{2D}^{ht} ”列中的数字与其他两列“ \mathcal{T}_{2D}^{wie} ”或“ \mathcal{T}_{1D} ”中的数字进行比较）。可以观察到，改进是十分显著的，通常大于0.5dB。

我们的算法使用的基函数是标准函数，计算效率高且与图像无关。们认为，可以通过使用更复杂的基础（例如自适应PCA [7]或学习知识字典[8]不完整）来改善所提出的降噪方法。但是，由于这些变换通常是不可分离的并且没有快速算法，因此计算复杂度也将大大增加。如上一节所示，即使使用当前使用的标准转换，我们的算法也已经比[8]和[11]表现出了更好的效果。

表 6 三种不同的彩色图像去噪算法的 PSNR 对比。实验图像在 RGB 域上被 $\sigma=25$ 的加性噪声污染，所有的 PSNR(dB)都是在 RGB 域上计算得到的。

Approach to color-image denoising	<i>Lena</i>	<i>House</i>	<i>Peppers</i>
BM3D independently on the <i>R</i> , <i>G</i> , and <i>B</i> color channels	31.44	32.18	30.93
BM3D independently on each opponent color channel	32.01	32.64	31.01
C-BM3D, with grouping constraint in opponent color space	32.27	33.03	31.20

本文提出的应用于彩色图像的扩展方法是值得一提的，因为我们没有三个亮度-色度通道上单独应用灰度版BM3D，而是对两个色度通道都施加了分组约束。分组约束意味着分组仅在亮度通道(通常比色度通道具有更高的信噪比)中进行一次，并且在两个色度中重复使用完全相同的分组进行协同滤波。将所提出的C-BM3D的性能与灰度BM3D在单个颜色通道上的独立应用进行比较是值得的。从表6可以看出，C-BM3D的PSNR比BM3D在各色彩通道上的独立应用好0.2~0.4 dB，比BM3D在RGB通

道上的独立应用好0.3~0.8 dB。这一改进显示了在C-BM3D中对色度通道使用分组约束的重要意义。我们注意到,在逐点SA-DCT去噪方法[9]中已经采用了类似的想法,利用亮度通道的信息来过滤色度通道。在那里,仅针对Y通道确定自适应形状估计邻域,然后将其重用于U和V。与[9]相比,本文方法的PSNR改进(0.1~0.4 dB)与这两种方法的灰度版本之间的改进是一致的。

VIII. 结论

本文提出的图像建模和估计算法可以解释为一种新的非局部自适应非参数滤波方法。该算法表现出了目前最先进的性能。据我们所知,表3和表4所示的PSNR结果在去除灰度图像和彩色图像中的加性高斯白噪声方面是最高的。此外,如表I所示,该算法以合理的计算成本实现了这些结果,并允许有效的复杂性/性能权衡。

该方法通过修改算法基本部分和维纳部分的系数方差计算,可以适应各种噪声模型,如加性有色噪声、非高斯噪声等。此外,本文所提出的方法只需要稍作调整,就可以用于对一维信号和视频进行降噪、图像恢复,同时也可以用于能够从高度稀疏的信号表示中受益的其他问题。

参考文献

- [1]. L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, November 2002.
- [2]. A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising," *IEEE Trans. Image Process.*, vol. 11, no. 5, pp. 545–557, May 2002.
- [3]. J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, "Image denoising using a scale mixture of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, November 2003.
- [4]. J. Guerrero-Colon and J. Portilla, "Two-level adaptive denoising using Gaussian scale mixtures in overcomplete oriented pyramids," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, Genova, Italy, September 2005.
- [5]. L. Yaroslavsky, K. Egiazarian, and J. Astola, "Transform domain image restoration methods: review, comparison and interpretation," in *Proc. Nonlinear Image Process. and Pattern Analysis XII*, vol. 4304, 2001, pp. 155–169.
- [6]. O. Guleryuz, "Weighted overcomplete denoising," in *Proc. Asilomar Conf. on Signals, Systems and Computers*, vol. 2, Pacific Grove, CA, USA, November 2003, pp. 1992–1996.
- [7]. D. Muresan and T. Parks, "Adaptive principal components and image denoising," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, September 2003.
- [8]. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned

- dictionaries,” *IEEE Trans. on Image Process.*, vol. 15, no. 12, pp. 3736–3745, December 2006.
- [9]. A. Foi, V. Katkovnik, and K. Egiazarian, “Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images,” *IEEE Trans. Image Process.*, vol. 16, no. 5, May 2007.
 - [10]. A. Buades, B. Coll, and J. M. Morel, “A review of image denoising algorithms, with a new one,” *Multisc. Model. Simulat.*, vol. 4, no. 2, pp. 490–530, 2005.
 - [11]. C. Kervrann and J. Boulanger, “Optimal spatial adaptation for patch based image denoising,” *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2866–2878, October 2006.
 - [12]. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising with block-matching and 3D filtering,” in *Proc. SPIE Electronic Imaging: Algorithms and Systems V*, vol. 6064A-30, San Jose, CA, USA, January 2006.
 - [13]. —, “Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space,” in *IEEE Int. Conf. Image Process.*, San Antonio, Texas, September 2007, accepted.
 - [14]. J. B. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proc. Berkeley Symp. Math. Stat. Probab.*, Berkeley, 1967, pp. 281–297.
 - [15]. T. Kohonen, *Self-Organizing Maps*, 2nd ed., ser. Information Sciences. Springer, Heidelberg, 1997, vol. 30.
 - [16]. F. Höppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis*. Wiley, Chichester, 1999.
 - [17]. A. Gersho, “On the structure of vector quantizers,” *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 157–166, 1982.
 - [18]. A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: A review,” *ACM Computing Surveys*, vol. 31, no. 3, pp. 264–323, 1999.
 - [19]. E. Vansteenkiste, D. Van der Weken, W. Philips, and E. E. Kerre, “Perceived image quality measurement of state-of-the-art noise reduction schemes,” in *Lecture Notes in Computer Science ACIVS*, vol. 4179, Antwerp, Belgium, September 2006, pp. 114–124.
 - [20]. A. Pizurica and W. Philips, “Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising,” *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 654–665, March 2006.
 - [21]. V. Katkovnik, K. Egiazarian, and J. Astola, *Local Approximation Techniques in Signal and Image Process.* SPIE Press, 2006, vol. PM157.
 - [22]. —, “Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule,” *Math. Imaging and Vision*, vol. 16, no. 3, pp. 223–235, May 2002.
 - [23]. A. Foi and V. Katkovnik, “From local polynomial approximation to pointwise shape-adaptive transforms: an evolutionary nonparametric regression perspective,” in *Proc. Int. TICSP Workshop Spectral Methods Multirate Signal Process.*, Florence, Italy, September 2006.