论 文

卡通 — 纹理字典学习的图像分解方法

李亚峰^{①②}, 冯象初^{①*}^① 西安电子科技大学应用数学系, 西安 710071^② 宝鸡文理学院计算机科学系, 宝鸡 721016

* 通信作者. E-mail: xcfeng@mail.xidian.edu.cn

收稿日期: 2012-02-15; 接受日期: 2012-05-27

国家自然科学基金 (批准号: 60872138, 61001156)、教育部科学技术研究重点项目基金 (批准号: 211155) 和宝鸡文理学院 2012 年院级科研重点项目基金 (批准号: ZK12055) 资助项目

摘要 结合变分方法与字典学习, 提出一个新的图像分解变分模型和由扩散流引导的字典学习算法. 实验结果表明, 和已有图像分解方法相比, 新算法自适应性强, 不仅可以更好地将图像的卡通和纹理分开, 而且能有效地将纹理中的噪声去除.

关键词 字典学习 卡通 — 纹理分解 图像去噪 全变分 优化算法

1 引言

图像处理和计算机视觉中一个重要的问题就是区分图像的不同特征. 一幅已知的图像可以表示为 $f = u + v$, 其中 u 是 f 的卡通部分, 包含了图像主要的几何特征信息, 由分片光滑的区域及清晰的边缘构成, v 是振荡纹理或噪声. Meyer^[1] 从理论上对图像的卡通 — 纹理分解进行了阐述, 并使用计算调和与分析中的 G 空间、 F 空间和 E 空间刻画纹理或噪声. G 空间、 F 空间或 E 空间的范数记为 $\|\cdot\|_*$ (以下统称为 $*$ 范数). Meyer 的理论指出纹理部分在 $*$ 范数度量下较小, 而卡通部分在 BV(bounded variation) 半范的度量下较小. 因此在 ROF^[2] 模型的基础上, Meyer 给出了图像卡通 — 纹理分解模型:

$$\inf_{u,v} \{E(u,v) = |\nabla u|_1 + \lambda \|v\|_*, u+v=f\}, \quad (1)$$

这里 f 为观察图像, λ 是平衡参数, u 是卡通部分, v 是纹理部分. 由于 Meyer 模型中刻画纹理部分的 G 、 F 、 E 三个空间的特殊形式, 该模型没有标准的 Euler-Lagrange 方程, 从而难以数值求解. 自 Meyer 的分解模型提出后有大量的工作围绕着这一模型展开. Vese^[3] 等提出用 L^p 的散度空间近似 G 空间 (简称 VO 模型). Osher 等^[4] 继续拓展 VO 模型的观点, 提出了一种基于全变差极小化和用 H^{-1} 范数代替 $*$ 范数的图像分解模型 (简称 OSV 模型):

$$\inf_{u \in BV} \left\{ E(u) = |\nabla u|_1 + \lambda \|f - u\|_{H^{-1}}^2 \right\}, \quad (2)$$

其中 $\|f - u\|_{H^{-1}}^2 = \int_{\Omega} |\nabla(\Delta^{-1}(f - u))|^2$. 模型 (2) 有明确的 Euler-Lagrange 方程, 容易求解. Yin 等^[5,6] 提出 TV-L1 图像分解模型和求解该模型的二次锥规划内点算法, 该模型基于 L1 数据项, 可以

引用格式: 李亚峰, 冯象初. 卡通 — 纹理字典学习的图像分解方法. 中国科学: 信息科学, 2013, 43: 644-656

将图像分解为不同尺度的特征. 他们在文献 [7] 中给出了 TV-L1、VO、Meyer 三种变分图像分解模型的比较. 随后, Duval 等 [8] 给出 TV-L1 模型的一个几何观点. 最近, Buades 等 [9] 提出一对非线性滤波器, 本质上是通过引入全变差的高通和低通滤波器, 他们建议的分解方法仍然保持 Meyer 变分模型 (1) 的本质特征. 在文献 [9] 中, 对比于目前其他变分图像分解方法, Buades 等提出的方法速度快、分解效果好.

目前, 基于字典的稀疏表示方法在图像处理中得到越来越多的应用. 在字典设计方面, 构造出了许多有效的解析字典, 如小波包字典、局部余弦字典、曲线波字典等. 基于稀疏表示方法, Starck 等 [10] 提出了图像分解的形态学成分分析方法 (MCA, morphological component analysis). MCA 的基本思想是用两个预先设定的字典, 一个用来描述纹理部分, 另一个用来描述卡通部分, 通过交替最小化目标泛函, 以达到图像分解的目的. 随着对稀疏表示方法进一步深入和细致的研究, MCA 图像分解方法在图像处理和模式识别的各个领域得到不断的发展和应用 (包括图像放大、图像超分辨、图像修补、CS (compressed sensor)、盲源分离、人脸识别等等). MCA 是基于如下假设: 自然图像的卡通部分和纹理部分在特定域 (例如曲线波域和小波域等) 中是稀疏的. 因此 MCA 的关键问题是稀疏域的选择或选取合适的字典. 图像分解的质量依赖于这些稀疏域的选择或字典的构造, 预先选择的字典对图像的卡通与纹理分解结果有很大的影响. 然而, 在理论上很难证明如何选取两个合适的字典分别描述图像的卡通与纹理部分. 不同性质的图像有不同的结构, 所以它们的最优稀疏域是不同的. 对于不同性质的图像使用预先设定的字典并不能很好的描述图像的不同成分. 我们从稀疏域的近似、解析字典的选择、算法的改进以及参数选取等方面做了大量的工作 [11~14], 虽然 MCA 的图像分解方法已较为成熟, 但仍然有很多难以解决的问题. 选取合适的字典, 至今仍然是一个公开问题 [10].

不同于预先设定的解析字典, 字典学习 (dictionary learning) 是通过优化某个目标泛函, 由样本训练出更适应于当前图像表示的自适应字典. 字典学习算法已用于图像去噪、修补、识别等, 例如最优方向算法 (MOD, method of optimal directions) [15]、KSVD (K singular value decomposition) 算法 [16]、多尺度字典学习算法 [17]、彩色图像字典学习算法 [18] 等. 针对图像分解问题, 本文建立一个新的模型, 通过模型给出卡通-纹理字典学习算法. 我们学习的两个字典可以较好地分别刻画图像的卡通和纹理部分, 用学习得到的字典和稀疏系数不仅可以重构出图像的卡通部分和纹理部分, 而且能有效地将纹理中的噪声去除.

2 新模型

2.1 新模型

假定图像 f 含有高斯白噪声 n , 且可以分解为 $f = u + v + n$. 利用变分方法和字典学习的优势, 我们将图像的卡通部分 u 和纹理部分 v 限制在两个合适的空间里, 通过学习两个字典分别描述图像的卡通部分和纹理部分. 建议的模型如下:

$$\arg \min_{\{u, v, D_1, D_2, \alpha_i, \beta_i\}} \left\{ \begin{aligned} & \lambda_1 \|f - (u + v)\|^2 + \lambda_2 |\nabla u|_1 + \lambda_3 \|v\|_{H^{-1}}^2 \\ & + \left(\sum_i \|R_i u - D_1 \alpha_i\|^2 + \sum_i \mu_i |\alpha_i|_0 \right. \\ & \left. + \sum_i \|P_i v - D_2 \beta_i\|^2 + \sum_i \eta_i |\beta_i|_0 \right) \end{aligned} \right\}, \quad (3)$$

$$\text{s.t. } d_l^T d_l = 1, s_m^T s_m = 1, l = 1, \dots, k_1, m = 1, \dots, k_2,$$

其中 $\lambda_1, \lambda_2, \lambda_3$ 是平衡参数, R_i, P_i 分别表示在卡通部分和纹理部分第 i 个位置的抽取像素块算子 (注意 R_i, P_i 抽取图像块的维数可以不同). 对卡通部分 u , 使用 BV 半范数刻画, 同时学习一个卡通字典 $D_1 = [d_1, d_2, \dots, d_{k_1}]$, 要求卡通部分能被 D_1 较好的稀疏表示; 对纹理部分 v , 使用 H^{-1} 范数刻画, 同时学习一个纹理字典 $D_2 = [s_1, s_2, \dots, s_{k_2}]$, 要求纹理部分能被 D_2 较好的稀疏表示. 模型中字典 D_1, D_2 的元素满足单位化约束, $d_l^T d_l = 1, l = 1, \dots, k_1, s_m^T s_m = 1, m = 1, \dots, k_2$. 稀疏表示系数的参数 μ_i 和 η_i 须匹配于噪声水平.

2.2 模型分析

模型 (3) 是一个联合优化问题, 本文采用交替最小化方法求解, 从而模型 (3) 分解为 4 个子问题. 下面分析每一个子问题, 并解释子问题在整个优化算法中的作用. 用 $u^{(k)}, v^{(k)}$ 分别表示第 k 次更新的卡通与纹理图像.

子问题 1 固定纹理部分 $v = v^{(k)}, D_2, \beta_i$, 求解 u, D_1, α_i . 由模型 (3), 考虑下面的子优化问题

$$\begin{aligned} (u^{(k+1)}, D_1, \alpha_i) = \arg \min_{u, D_1, \alpha_i} & \lambda_1 \| (f - v^{(k)}) - u \|^2 + \sum_i \| R_i u - D_1 \alpha_i \|^2 + \sum_i \mu_i |\alpha_i|_0, \\ \text{s.t. } & d_l^T d_l = 1, l = 1, \dots, k_1. \end{aligned} \quad (4)$$

求解模型 (4) 分为卡通字典学习与卡通图像重构, 这一过程学习到一个表示卡通部分的字典 D_1 , 并对初始卡通图像 $f - v^{(k)}$ 做去噪处理, 因此得到一个去噪版本的卡通图像 $u^{(k+1)}$, 下面通过扩散流再次更新卡通图像.

子问题 2(扩散流引导) 固定 $D_1, \alpha_i, D_2, \beta_i$, 由模型 (3), 考虑求解最小化问题

$$u^{(k+2)} = \arg \min_u \left(E(u) = \lambda_2 |\nabla u|_1 + \lambda_3 \|v\|_{H^{-1}}^2 \right), \quad \text{s.t. } v = u^{(k+1)} - u. \quad (5)$$

可以看到 (5) 式的形式与 OSV 模型一致, 它可以通过扩散流来求解能量泛函的最小值点, 因此求解子问题 2 的作用是将卡通图像中剩余的纹理成分分离出去, 从而引导下面纹理字典的学习. 由于扩散流的初始值为去噪版本的卡通图像 $u^{(k+1)}$, 所以此时的分解受到噪声的影响较小. 由子问题 2 的解, 得到分离出纹理成分的卡通图像 $u^{(k+2)}$.

子问题 3 固定卡通部分 $u = u^{(k+2)}, D_1, \alpha_i$, 求解 v, D_2, β_i . 由模型 (3), 考虑求解子优化问题

$$\begin{aligned} (v^{(k+1)}, D_2, \beta_i) = \arg \min_{v, D_2, \beta_i} & \lambda_1 \| (f - u^{(k+2)}) - v \|^2 + \sum_i \| P_i v - D_2 \beta_i \|^2 + \sum_i \eta_i |\beta_i|_0, \\ \text{s.t. } & s_m^T s_m = 1, m = 1, \dots, k_2. \end{aligned} \quad (6)$$

这个子问题与子问题 1 一致, 求解时分为纹理字典学习与纹理图像重构, 与问题 1 不同的是纹理图像的初始值为 $f - u^{(k+2)}$. 通过 (6) 式学习得到一个纹理字典 D_2 , 并去除 $f - u^{(k+2)}$ 的噪声, 得到一个干净版本的纹理图像 $v^{(k+1)}$.

子问题 4(扩散流引导) 固定 $D_1, \alpha_i, D_2, \beta_i$, 由模型 (3), 考虑求解最小化问题

$$v^{(k+2)} = \arg \min_v \left(E(v) = \lambda_2 |\nabla v|_1 + \lambda_3 \|u\|_{H^{-1}}^2 \right), \quad \text{s.t. } u = v^{(k+1)} - v. \quad (7)$$

子问题 4 通过扩散流来求解能量泛函的最小值点, 其作用是将纹理图像中剩余的卡通成分分离出去, 从而引导下一次卡通字典的学习. (7) 式中的初始值为去噪版本的纹理图像 $v^{(k+1)}$, 所以此时的分解受

到噪声的影响较小. 由子问题 4 的解得到去除卡通成分的纹理图像 $v^{(k+2)}$, 继而返回到子问题 1 进行迭代.

以上对 4 个子问题的分析表明: 子问题 1, 3 通过学习不同性质的字典, 用其表示图像的不同成分, 实现卡通部分与纹理部分的分解, 并保证了卡通部分与纹理部分的独立处理和噪声的去除; 子问题 2, 4 利用不同性质的空间范数, 约束图像的不同成分, 分别修正图像中结构与纹理部分 (去除结构部分中的纹理, 去除纹理部分中的结构), 实现卡通部分与纹理部分的进一步分解; 上述过程交替迭代, 完成字典的学习以及图像的分解和去噪.

算法具体描述如下:

第一步: 输入平衡参数 $\lambda_1, \lambda_2, \lambda_3$, 卡通字典与纹理字典 D_1, D_2 (单位化的随机字典), $u^{(0)} = f, v^{(0)} = 0, k = 0, \varepsilon$.

第二步: 迭代.

固定纹理部分 $v = v^{(k)}, D_2, \beta_i$, 求解子问题 1, 得到 $u^{(k+1)}, D_1, \alpha_i$.

扩散流引导, 固定 $D_1, \alpha_i, D_2, \beta_i$, 求解子问题 2, 得到 $u^{(k+2)}$.

固定卡通部分 $u = u^{(k+2)}, D_1, \alpha_i$, 求解子问题 3, 得到 $v^{(k+1)}, D_2, \beta_i$.

扩散流引导, 固定 $D_1, \alpha_i, D_2, \beta_i$, 求解子问题 4, 得到 $v^{(k+2)}$.

$k = k + 2$.

第三步: 输出 u, v .

3 算法

这一节给出每个子问题的详细解法.

子问题 1 固定纹理部分 $v = v^{(k)}, D_2, \beta_i$, 求解 u, D_1, α_i . 考虑求解子优化问题

$$\begin{aligned} \arg \min_{u, D_1, \alpha_i} \lambda_1 \| (f - v^{(k)}) - u \|^2 + \sum_i \| R_i u - D_1 \alpha_i \|^2 + \sum_i \mu_i |\alpha_i|_0, \\ \text{s.t. } d_l^T d_l = 1, l = 1, \dots, k_1. \end{aligned} \quad (8)$$

求解模型 (8) 分为卡通字典学习与卡通图像重构. 初始的卡通图像为 $u^{(k)} = f - v^{(k)}$, 假设有 M 个训练样本, 则初始训练样本集 $U = [R_1 u^{(k)}, R_2 u^{(k)}, \dots, R_M u^{(k)}]$.

卡通字典学习: 求解

$$\arg \min_{D_1, \alpha_i} \sum_i \| R_i u - D_1 \alpha_i \|^2 + \sum_i \mu_i |\alpha_i|_0 \quad \text{s.t. } d_l^T d_l = 1, l = 1, \dots, k_1. \quad (9)$$

模型 (9) 的求解同样采用交替优化的方法, 首先给定字典 D_1 , 求解所有训练样本在 D_1 上的稀疏表示系数 $\alpha_i (i = 1, \dots, M)$, 从而得到稀疏表示矩阵 $\mathbf{A} = [\alpha_1, \alpha_2, \dots, \alpha_M]$; 然后根据 $\mathbf{A}$ 更新字典 D_1 , 即求解 (9) 式每次迭代包含如下两个步骤:

第一步, 固定字典 D_1 , 求解 $\mathbf{A}$.

在 (9) 式中, 参数 μ_i 须匹配于噪声水平, 通过合适的选择^[19], 模型 (9) 等价于

$$\arg \min_{\alpha_i} |\alpha_i|_0 \quad \text{s.t. } \| R_i u - D_1 \alpha_i \|^2 \leq \varepsilon, \quad (10)$$

这里 ε 是固定的逼近误差常数. 利用正交匹配追踪算法^[20], 可求解 (10) 式. 当求得所有训练样本 $R_i u^{(k)}$ 的稀疏系数 α_i , 则获得稀疏表示矩阵 $\mathbf{A}$.

第二步, 固定 $\mathbf{A}$, 更新字典 D_1 .

在这一阶段, 稀疏表示矩阵 $\mathbf{A}$ 是已知的, (9) 式退化为

$$\arg \min_{D_1} \|U - D_1 \mathbf{A}\|_F^2 \quad \text{s.t.} \quad d_l^T d_l = 1, l = 1, \dots, k_1. \quad (11)$$

为了表达方便, 记 $\mathbf{A} = [\mathbf{a}_1; \mathbf{a}_2; \dots; \mathbf{a}_{k_1}]$, 这里 $\mathbf{a}_i (i = 1, \dots, k_1)$ 是矩阵 $\mathbf{A}$ 的行向量. 下面对字典 D_1 进行逐列更新, 即更新 $d_j (j = 1, \dots, k_1)$ 时, 字典 D_1 的其他元素 ($d_i (i \neq j)$) 固定, (11) 式变为

$$\arg \min_{d_j} \left\| U - \sum_{i \neq j} d_i \mathbf{a}_i - d_j \mathbf{a}_j \right\|_F^2 \quad \text{s.t.} \quad d_j^T d_j = 1. \quad (12)$$

令 $V = U - \sum_{i \neq j} d_i \mathbf{a}_i$, (12) 式写为

$$\arg \min_{d_j} \|V - d_j \mathbf{a}_j\|_F^2 \quad \text{s.t.} \quad d_j^T d_j = 1. \quad (13)$$

使用 Lagrange 乘子法, (13) 式变为

$$\arg \min_{d_j} [\text{tr}(-V \mathbf{a}_j^T d_j^T - d_j \cdot \mathbf{a}_j V^T + d_j \cdot (\mathbf{a}_j \mathbf{a}_j^T - \gamma) d_j^T + \gamma)]. \quad (14)$$

这里 γ 是一个纯量, 令 (14) 式的泛函关于 d_j 的导数为零, 得到 $d_j = V \mathbf{a}_j^T (\mathbf{a}_j \mathbf{a}_j^T - \gamma)^{-1}$, 由单位化约束得到 (11) 式的解为 $d_j = V \mathbf{a}_j^T / \|V \mathbf{a}_j^T\|, j = 1, 2, \dots, k_1$, 因此得到更新的字典 D_1 . 注意, 这里给出的字典更新方式与已有的不同 (如 MOD, KSVD 等), 其更新方式简单, 并且也能达到较好的收敛速度.

通过以上两步进行迭代, 得到 (9) 式的解 D_1, α_i .

卡通图像重构: 固定 $v = v^{(k)}$, 利用 D_1, α_i 重构 $u^{(k+1)}$. 由 (8) 式得

$$u^{(k+1)} = \arg \min_u [\lambda_1 \|(f - v^{(k)}) - u\|^2 + \sum_i \|D_1 \alpha_i - R_i u\|^2]. \quad (15)$$

(15) 式为可微的优化问题, 通过求导, 容易得到

$$u^{(k+1)} = \left(\lambda_1 I + \sum_i R_i^T R_i \right)^{-1} \left(\lambda_1 (f - v^{(k)}) + \sum_i R_i^T D_1 \alpha_i \right). \quad (16)$$

至此, 我们得到子优化问题 1 的解 $u^{(k+1)}, D_1, \alpha_i$.

子问题 2 (扩散流引导) 固定 $D_1, \alpha_i, D_2, \beta_i$, 求解最小化问题

$$u^{(k+2)} = \arg \min_u (E(u) = \lambda_2 |\nabla u|_1 + \lambda_3 \|v\|_{H^{-1}}^2), \quad \text{s.t.} \quad v = u^{(k+1)} - u. \quad (17)$$

令 $E'(u) = 0$, Euler-Lagrange 方程为

$$2\lambda_3 \Delta^{-1} (u^{(k+1)} - u) = \lambda_2 \text{div} \left(\frac{\nabla u}{|\nabla u|} \right), \quad \frac{\partial u}{\partial n} \Big|_{\partial \Omega} = 0, \quad \frac{\partial \text{div}(\frac{\nabla u}{|\nabla u|})}{\partial n} \Big|_{\partial \Omega} = 0. \quad (18)$$

在 (18) 式两边同乘以 Laplacian 算子 Δ 有

$$2\lambda_3 (u^{(k+1)} - u) = \lambda_2 \Delta \text{div} \left[\frac{\nabla u}{|\nabla u|} \right], \quad \frac{\partial u}{\partial n} \Big|_{\partial \Omega} = 0, \quad \frac{\partial \text{div}(\frac{\nabla u}{|\nabla u|})}{\partial n} \Big|_{\partial \Omega} = 0. \quad (19)$$

(19) 式相应的扩散流为

$$\frac{du}{dt} = -\frac{\lambda_2}{2\lambda_3} \Delta \operatorname{div} \left[\frac{\nabla u}{|\nabla u|} \right] + (u^{(k+1)} - u). \quad (20)$$

由 (20) 式得到 (17) 式的解 $u^{(k+2)}$.

子问题 3 固定卡通部分 $u = u^{(k+2)}$, D_1, α_i , 求解 v, D_2, β_i , 考虑求解子优化问题

$$\begin{aligned} \arg \min_{v, D_2, \beta_i} & \lambda_1 \|(f - u^{(k+2)}) - v\|^2 + \sum_i \|P_i v - D_2 \beta_i\|^2 + \sum_i \eta_i |\beta_i|_0, \\ \text{s.t. } & \mathbf{s}_m^T \mathbf{s}_m = 1, m = 1, \dots, k_2. \end{aligned} \quad (21)$$

采用与子问题 1 同样的解法, 分为纹理字典学习和纹理图像重构, 即求解下面的两个优化问题.

纹理字典学习: 给定初始训练样本 $P_i(f - u^{(k+2)})$, 求解

$$\arg \min_{D_2, \beta_i} \sum_i \|P_i v - D_2 \beta_i\|^2 + \sum_i \eta_i |\beta_i|_0 \quad \text{s.t. } \mathbf{s}_m^T \mathbf{s}_m = 1, m = 1, \dots, k_2. \quad (22)$$

纹理图像重构: 用 $u^{(k+2)}, D_2, \beta_i$ 重构纹理部分

$$v^{(k+1)} = \arg \min_v \lambda_1 \|f - (u^{(k+2)} + v)\|^2 + \sum_i \|P_i v - D_2 \beta_i\|^2. \quad (23)$$

(22), (23) 式的优化方法与求解子问题 1 一致, 但这里是纹理字典的学习与纹理图像重构, 体现在纹理图像的初始值为 $f - u^{(k+2)}$.

子问题 4(扩散流引导) 固定 $D_1, \alpha_i, D_2, \beta_i$, 由模型 (3), 考虑求解最小化问题

$$v^{(k+2)} = \arg \min_v (\lambda_2 |\nabla u|_1 + \lambda_3 \|v\|_{H^{-1}}^2), \quad \text{s.t. } u = v^{(k+1)} - v. \quad (24)$$

子问题 4 的求解方法与子问题 2 一致. (24) 式的解是从 $v^{(k+1)}$ 中去除卡通成分后的纹理图像 $v^{(k+2)}$. 使用 $v^{(k+2)}$ 返回子问题 1 进行迭代.

从算法的细节可以看到, 建议的算法有如下几个方面的特点: (1) 与传统字典学习算法不同, 新算法是由图像卡通和纹理部分的空间限制来引导的, 每次迭代通过扩散流的校正, 从而学习得到与限制空间相对应的两个不同性质的字典. (2) 新算法与预先去除图像噪声后再进行图像分解本质上是不同的. 新算法是一个交替最小化过程, 将字典学习、图像表示耦合于扩散流中. 交替迭代中, 不断地修正残差图像中的结构和纹理成分, 使得不同结构的字典更好地稀疏表示相应成分, 以保证图像不同成分的独立. 在实验中可以清楚的看到本文算法优于预处理再分解方法. (3) 字典学习 (子问题 1, 3) 过程需要进行迭代更新字典与稀疏系数. 在实际计算中, 每个子优化问题不必求出确切的最小值点, 字典的更新只进行一次, 并且扩散流也只需数次的迭代即可. (4) 尽管没有从理论上证明迭代的收敛性, 但实际算例表明交替迭代过程较快地收敛到稳定点. (5) 算法中两个学习字典的维数可以不同, 因此可以根据不同图像的结构灵活选取字典维数. (6) 众所周知, TV 约束是图像分解模型的一个成功的正则化约束, 然而 TV 约束造成了边界模糊及图像的阶梯效应 (块效应), 新算法通过引入两个学习字典来更好地刻画图像的不同成分, 从实验结果可以看到新算法有效地减少了 TV 约束带来的边界模糊与图像的阶梯效应.

图 1 测试图像

Figure 1 The test images. (a) Barbara; (b) Baboon; (c) Fingerprint; (d) Noisy square; (e) Textured square

4 数值实验与结果分析

这一节使用本文算法、OSV、MCA 和文献 [9] 中提出算法对给定的测试图像做卡通 — 纹理分解实验. MCA 方法中两个字典分别为曲线波和 LDCT (代码下载于 <http://jstarck.free.fr/mca.html>), 文献 [9] 的算法中, 尺度参数选择为默认值 3 (代码下载于 [http:// mw. cmla. ens-cachan.fr/ megawave/ algo/ cartoon_ texture/](http://mw.cmla.ens-cachan.fr/megawave/ algo/ cartoon_ texture/)). 测试图像为 Barbara 图像 (256×256)、Baboon 图像 (256×256)、以纹理成分为主体的指纹 Fingerprint 图像 (121×117)、以结构为主体的 Noisy square 图像 (309×304) 和包含两种纹理成分的 Textured square 图像 (256×253), 如图 1 所示.

4.1 参数的选取

由于本文算法中刻画卡通图像和纹理图像的字典维数可以不同, 为了比较在不同字典维数选择下, 图像分解的质量和计算时间, 实验中给出了不同测试图像上, 分别选择相同和不同字典维数的实验结果. 两个字典维数不同时, 卡通字典 D_1 的大小选择为 100×200 , 冗余度为 2, 子块的大小为 10×10 , 纹理字典 D_2 的大小选择为 64×256 , 冗余度为 4, 子块的大小为 8×8 ; 两个字典维数相同时, 卡通字典和纹理字典的大小均选择为 64×256 , 冗余度为 4, 子块的大小为 8×8 ; 所有实验中, 从图像中有重叠的提取所有子块用于训练字典. 稀疏表示逼近误差经验的选取为 $\varepsilon = (1.14\sigma)^2 \times n$, σ 是噪声标准方差, n 表示图像块的维数. 在提出的模型 (3) 中, 参数 λ_1 控制着图像分解的两个成分之和 $u + v$ 与观测图像 f 的接近程度, 所有实验中, 取 $\lambda_1 = 30$. 参数 λ_2, λ_3 控制着图像的分解尺度, λ_2/λ_3 可以解释为图像分解的尺度参数, 如果 λ_2/λ_3 较大, 则全变差能量在总模型能量中所占比例较大, 因此最小化能量泛函 (3) 的过程迫使卡通图像全变差能量相对较小, 而震荡纹理图像的 H^{-1} 范数相对较大, 因此得到较大尺度的纹理成分; 同理, 如果 λ_2/λ_3 较小, 得到较小尺度的纹理成分. 实验中, 选择 $\lambda_2 = 1 \sim 1000, \lambda_3 = 1$. 因此, 使用本文算法, 图像分解的尺度仅由尺度参数 λ_2 来控制. 扩散流 (20) 中的 Laplacian 算子 Δ 采用中心差分格式, 演化时间 $dt = 0.1$.

4.2 实验结果与分析

实验 1 对带噪 Barbara 图像做卡通 — 纹理分解实验, 噪声标准方差 $\sigma = 20$. 图 2(a) 和 (b) 给出本文算法在两种字典尺寸下的实验结果. 实验结果表明: 本文算法在两种字典维数的选择下, 图像分解和去噪结果相当. 在计算速度上, 选择维数较低的字典运算速度更快. 图 2(c) 和 (d) 给出 OSV、MCA 的实验结果. 本文算法与 OSV 方法相比, 在结构部分减少了阶梯效应 (如脸部、眼睛、额头等的块状结构), 在纹理部分得到了较干净的纹理成分 (图 2(c)); 与 MCA 方法相比, 新算法更好地分离出卡通、

图 2 Barbara 图像分解和去噪实验结果比较

Figure 2 Cartoon-texture decomposition and denoising results by the three methods. (a) The proposed method, the size of D_1 is set to 100×200 , the size of D_2 is set to 64×256 , Cartoon+Texture PSNR=30.1267 dB, the runtime is 1392 s; (b) the proposed method, the size of D_1 and D_2 is set to 64×256 , Cartoon+Texture PSNR=30.1582 dB, the runtime is 1067 s; (c) OSV, the runtime is 201 s; (d) MCA, Cartoon+Texture PSNR=27.7235 dB, the runtime is 284 s

图 3 Barbara 图像的卡通, 纹理字典

Figure 3 The learned dictionaries by the proposed algorithm. (a) Cartoon dictionary, the size of D_1 is 100×200 , the size of atoms is 10×10 ; (b) texture dictionary, the size of D_2 is 64×256 , the size of atoms is 8×8

纹理和噪声 (图 2(d)). 实际上, 学习字典的原子图 3 显示出, 卡通字典不含有纹理信息, 而纹理字典也不含有卡通信息, 从而保证了由相应字典重构的卡通与纹理部分较好的分离, 并且可以看到新算法在峰值信噪比上具有明显优势.

实验 2 对带噪 Baboon 图像的卡通 — 纹理分解实验, 噪声标准方差 $\sigma = 20$. 由于 Baboon 图像中的纹理和噪声很类似, 将噪声与纹理分开的难度更大. 我们首先实验了本文算法中不同字典尺寸的影响 (图 4(a) 和图 4(b)). 实验结果表明: 与具有相同较低维数字典的分解结果相比, 对尺度较大的结构部分选择维数较高的结构字典其分解和去噪结果略好. 图 4(d) 和图 4(e) 给出 OSV、MCA 方法的实验结果. 本文算法与 OSV 方法相比, 在分解、去噪的同时能够保持更好的对比度 (如眼睛), 并且看到 OSV 方法 (图 4(d)) 出现明显的块效应; 本文算法与 MCA 方法相比 (图 4(e)), 具有更好的分解结果 (例如胡须) 和更高的信噪比.

实验 3 与实验 2 不同, 实验 3 使用在 Barbara 图像学习的卡通 — 纹理字典 (如图 3 所示), 对加噪 Baboon 图像做卡通 — 纹理分解. 从图 4(b) 可以看到利用样例图像 (example image) Barbara 预学习的字典, 对于 Baboon 图像也能得到较好的分解结果, 信噪比依然高于 MCA 方法. 这个实验表明: 图 3 所示的卡通 — 纹理字典具有较宽的适用范围. 由于实验中字典是固定的, 因此节省了运算量.

实验 4 这个实验将验证本文算法与预先去除图像噪声后再进行图像分解之间的区别. 预先去除图像噪声后再进行图像分解算法由如下两步构成: (1) 使用本文的字典学习去噪算法对带噪 Barbara

图 4 Baboon 图像分解和去噪实验结果比较

Figure 4 Cartoon-texture decomposition and denoising results by the three methods. (a) The proposed method, the size of D_1 is set to 100×200 , the size of D_2 is set to 64×256 , Cartoon+Texture PSNR=26.5188 dB, the runtime is 1536 s; (b) the proposed method, the size of D_1 and D_2 is set to 64×256 , Cartoon+Texture PSNR=26.5005 dB, the runtime is 1128 s; (c) the proposed method with pre-learned dictionaries, Cartoon+Texture PSNR=25.8503 dB, the runtime is 713 s; (d) OSV, the runtime is 202 s; (e) MCA, Cartoon+Texture PSNR=24.6187 dB, the runtime is 301 s

图 5 Barbara 图像分解实验

Figure 5 Decomposition result of “Barbara” image by the preprocessed method. (a) The result by applying the denoising algorithm to the noisy “Barbara” image; (b) cartoon part with OSV; (c) texture part with OSV

图像进行去噪处理。(2) 使用 OSV 方法对去噪后的 Barbara 图像做分解。图 5 给出了上述两步方法的结果。由图 5 可以清楚地看到纹理图像中的结构成分。实验结果表明: 本文算法的分解结果 (图 2(a), 图 2(b)) 比直接预处理再分解图像的结果 (图 5) 更优。本文的迭代算法通过引入两个学习字典来刻画图像的不同成分, 并且考虑去除图像的噪声, 空间引导过程的作用在于修正图像的结构和纹理部分, 保证了卡通部分与纹理部分的独立处理。在每次迭代中, 从残差图像中提取图像的结构和纹理成分, 使得不同性质的学习字典更好的表示相应成分, 从而由两个学习字典重构的图像具有不同的特征。

实验 5 用本文方法、OSV 和 MCA 对不同强度噪声下的测试图像进行分解实验。表 1 给出三种方法去噪后峰值信噪比 (PSNR, peak signal-to-noise ratio) 的比较。这些实验结果说明新方法可以

表 1 测试图像的分解去噪实验比较

Table 1 PSNR results of the three methods

Test images	σ	Original image	OSV	MCA	Our method
"Barbara"	10	28.1317	30.7830	29.4862	33.8924
	20	22.1111	26.9101	27.7235	30.1582
	30	18.5892	24.1136	24.7960	27.4303
"Baboon"	10	28.0971	29.5166	28.7130	30.4217
	20	22.0765	25.0445	24.6187	26.5005
	30	18.5547	23.5977	23.0852	24.6014

图 6 Fingerprint 图像分解实验结果比较

Figure 6 Decomposition results by the proposed method and the method in [9]. (a) Noisy "Fingerprint" image ($\sigma = 20$); (b) the proposed method applied to Noisy "Fingerprint" image, Cartoon+Texture PSNR=24.3543 dB; (c) the proposed method applied to cleaned "Fingerprint" image; (d) the method in [9] applied to cleaned "Fingerprint" image

分解图像的同时有效抑制噪声.

实验 6 使用本文算法和文献 [9] 中算法做对比实验. 测试图像是以纹理成分为主体的 Fingerprint 图像 (图 1(c)). 由于文献 [9] 中的算法没有考虑图像含有噪声的情形, 在本实验中只将其用于无噪声的图像上. 本文算法分为带噪和不带噪两种情形进行图像分解实验. 从实验结果 (图 6(b), 图 6(c)) 可以看到本文算法将指纹图像成功的提取到纹理部分, 结构部分较好地保留了指纹图像的背景成分, 且其边界较清晰, 没有太多的模糊效应.

实验 7 在这个实验中, 测试图像是以结构为主体的 Noisy square 图像 (图 1(d)). 从实验结果 (图 7) 可以看到, 对比文献 [9] 的算法结果, 本文算法将 Noisy square 图像中的结构部分和 "噪声" 纹理成功分离开, 纹理部分没有出现明显的边界结构.

实验 8 测试图像为包含两种纹理成分的 Textured square 图像 (图 1(e)), 图 8 给出使用本文算法在参数 λ_2 取不同值时的实验结果. 从实验结果可以看到, 本文算法的分解尺度参数 λ_2 控制着提取

图 7 Noisy square 图像分解实验结果比较

Figure 7 Decomposition results by the proposed method and the method in [9]. (a) The proposed method; (b) the method in [9]

图 8 本文算法在不同参数下对 Textured square 图像的分解结果

Figure 8 Decomposition results by the proposed method with the different scale parameter λ_2 . The first row results are cartoon parts. The second row results are texture parts. Each column is the resulting image by the proposed method with different values of parameter λ_2 . (a) $\lambda_2 = 1$; (b) $\lambda_2 = 100$; (c) $\lambda_2 = 1000$

纹理图像的尺度, 不同的参数值得到不同尺度下的图像分解结果. 随着尺度参数 λ_2 的增大, 提取纹理图像的尺度逐渐增大, 对应的卡通图像更加平滑.

5 结束语与讨论

本文提出卡通 — 纹理字典学习算法. 该算法结合变分法, 由扩散流引导字典学习, 用学习的两个字典分别描述卡通部分和纹理部分. 实验结果表明, 新算法成功地将卡通、纹理与噪声分开, 更好地实现了图像的分解和去噪. 本文算法迭代进行稀疏编码、卡通和纹理字典更新、图像重构, 故运算量较大. 一个解决办法如文中所述, 通过样例图像预学习字典来做图像分解. 目前如何更快、更精确的进

行字典学习和稀疏编码是众多学者研究的前沿问题. 在模型中, 我们使用 BV 和 H^{-1} 空间引导字典学习, 是否有更合理的空间限制 (例如 Besov 空间和 $\text{div}(\text{BMO})$ 空间)? 字典维数如何选择? 这些问题有待于进一步研究.

参考文献

- 1 Meyer Y. Oscillating patterns in image processing and nonlinear evolution equations. Boston: University Lecture Series, American Mathematical Society, 2001
- 2 Rudin L, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D*, 1992, 60: 259–268
- 3 Vese L, Osher S. Modeling textures with total variation minimization and oscillating patterns in image processing. *J Sci Comput*, 2003, 19: 553–572
- 4 Osher S, Sole A, Vese L. Image decomposition and restoration using total variation minimization and the H^{-1} Norm. *SIAM Multiscale Model Simul*, 2003, 1: 349–370
- 5 Yin W T, Goldfarb D, Osher S. Image Cartoon-Texture Decomposition and Feature Selection Using the Total Variation Regularized L_1 Functional. Technical Report 05-47, UCLA CAM Reports. 2005
- 6 Yin W T, Goldfarb D, Osher S. The total variation regularized L_1 model for multiscale decomposition. *SIAM Multiscale Model Simul*, 2007, 6: 190–211
- 7 Yin W T, Goldfarb D, Osher S. A comparison of three total variation based texture extraction models. *J Vis Commun Image Represent*, 2007, 18: 240–252
- 8 Duval V, Aujol J F, Gousseau Y. The TVL1 model: a geometric point of view. *SIAM Multiscale Model Simul*, 2009, 8: 154–189
- 9 Buades A, Le T M, Morel J M, et al. Fast cartoon + texture image filters. *IEEE Trans Image Process*, 2010, 19: 1978–1986
- 10 Starck J L, Elad M, Donoho D L. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Trans Image Process*, 2005, 14: 1570–1582
- 11 Jiang L L, Feng X C, Yin H Q. Variational image restoration and decomposition with curvelet shrinkage. *J Math Imag Vision*, 2008, 30: 125–132
- 12 Li Y F, Feng X C. The split Bregman method for L_1 projection problems. *Acta Electron Sin*, 2010, 38: 2471–2475 [李亚峰, 冯象初. L_1 投影问题的分裂 Bregman 方法. *电子学报*, 2010, 38: 2471–2475]
- 13 Li M, Feng X C. A variational model for image decomposition based on wavelet method. *Acta Electron Sin*, 2008, 36: 184–187 [李敏, 冯象初. 基于小波空间的图像分解变分模型. *电子学报*, 2008, 36: 184–187]
- 14 Zhang L, Paul B, Wu X L. Multiscale LMMSE-based image denoising with optimal wavelet selection. *IEEE Trans Circuits Syst Video Technol*, 2005, 15: 469–481
- 15 Engan K, Aase S O, Hakon H J. Method of optimal directions for frame design. In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*. Phoenix, 1999. 2443–2446
- 16 Aharon M, Elad M, Bruckstein A M. The K-SVD: an algorithm for designing of overcomplete dictionaries for sparse representations. *IEEE Trans Image Process*, 2006, 15: 4311–4322
- 17 Mairal J, Sapiro G, Elad M. Learning multiscale sparse representations for image and video restoration. *SIAM Multiscale Model Simul*, 2008, 7: 214–241
- 18 Mairal J, Elad M, Sapiro G. Sparse representation for color image restoration. *IEEE Trans Image Process*, 2008, 17: 53–69
- 19 Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans Image Process*, 2006, 15: 3736–3745
- 20 Pati Y C, Rezaiifar R, Krishnaprasad P S. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, 1993. 40–44

Coupled dictionary learning method for image decomposition

LI YaFeng^{1,2} & FENG XiangChu^{1*}

1 Department of Applied Mathematics, Xidian University, Xi'an 710071, China;

2 Department of Computer Science, Baoji University of Arts and Science, Baoji 721016, China

*E-mail: xcfeng@mail.xidian.edu.cn

Abstract A novel variational model for image decomposition is proposed. Meanwhile a new cartoon-texture dictionary learning algorithm, which is guided by diffusion flow, is presented. Numerical experiments show that the proposed method has better performance than the existing algorithms in image decomposition and denoising.

Keywords dictionary learning, cartoon-texture decomposition, image denoising, total variation, optimization algorithm

LI YaFeng was born in 1977. He received the Ph. D. degree in 2011 from Xidian University. Currently, he is a Lecturer in Department of Computer Science, Baoji University of Arts and Science. His research interests include color image processing, texture analysis, and optimization algorithm.

FENG XiangChu was born in 1962. He received the Ph. D. degree in 1999 from Xidian University. Currently, he is a Professor in the Department of Applied Mathematics, School of Science, Xidian University. His research interests include numerical analysis, wavelets, and partial differential equations for image processing.