

基于双稀疏优化的空域错误隐藏

严静文^{a,b}, 肖晶^{a,b}, 高戈^{a,b}

(武汉大学 a. 国家多媒体软件工程技术研究中心; b. 计算机学院, 武汉 430072)

摘要: 现有空域错误隐藏算法通常利用线性插值或者常规稀疏表达恢复丢失像素, 但线性插值在恢复不平滑图像时因邻域信息不一致导致恢复图像模糊, 而常规稀疏表达因字典构建不当造成丢失像素重建效果较差。为此, 提出一种改进的空域错误隐藏算法, 采用动态阈值搜索潜在集合和模板集合提高字典构建精度, 利用典型相关分析获得双稀疏优化的初值, 通过稀疏重建恢复丢失像素。实验结果表明, 与现有主流算法相比, 该算法的峰值信噪比至少提高 1.23 dB, 具有较好的错误隐藏效果。

关键词: 空域错误隐藏; 线性插值; I 帧丢失; 字典构建; 稀疏优化

中文引用格式: 严静文, 肖晶, 高戈. 基于双稀疏优化的空域错误隐藏[J]. 计算机工程, 2019, 45(2): 245-249.

英文引用格式: YAN Jingwen, XIAO Jing, GAO Ge. Spatial error concealment based on coupled sparse optimization[J]. Computer Engineering, 2019, 45(2): 245-249.

Spatial Error Concealment Based on Coupled Sparse Optimization

YAN Jingwen^{a,b}, XIAO Jing^{a,b}, GAO Ge^{a,b}

(a. National Engineering Research Center for Multimedia Software;

b. School of Computer, Wuhan University, Wuhan 430072, China)

[Abstract] Linear interpolation algorithm or conventional sparse representation algorithm are used to recover the lost pixels currently. However, linear interpolation restores image blur due to inconsistent neighborhood information when restoring unsmooth images. For the conventional sparse representation algorithm, improper dictionary construction will result in a poor recovered image quality. To solve these problem, an improved spatial error concealment algorithm is proposed. The proposed algorithm optimizes the process of potential set and template set search by means of dynamic threshold searching, which improves the precision of the constructed dictionary. It can obtain the value of double sparse optimization using Canonical Correlation Analysis(CCA), and recover lost pixels by sparse reconstruction. Experimental results show that the proposed algorithm improves the Peak Signal to Noise Ratio(PSNR) by at least 1.23 dB compared with the current mainstream algorithm, and has a good error hiding effect.

[Key words] spatial error concealment; linear interpolation; I-frame loss; dictionary construction; sparse optimization

DOI: 10.19678/j.issn.1000-3428.0050093

0 概述

H.264/AVC 编码标准^[1]对比其他视频压缩标准在图像/视频压缩上具有更高的效率, 但是其带来高效率的同时, 容易引起压缩码流对信道误码更加敏感, 导致压缩码流在传输过程中更易发生丢包和错误^[2]。一个比特的丢失会在后续帧中引起错误扩散, 这将导致当前编码帧的失真, 严重影响接收端的图像质量。重传视频是解决这一问题的最简单方法, 但该方法在实际应用中因为带宽和实时性的限制, 实现代价高、实用性差。而错误隐藏的基本原理是在不改变编码器的情况下, 在解码端采用一系列

错误隐藏算法恢复丢失信息来达到错误掩盖的效果。这样做的优势在于能保持原有码率, 不需要额外的带宽要求, 就可以较好地保证丢失信息图像的质量。

通常错误隐藏算法分为空域错误隐藏^[3]、时域错误隐藏以及时空域错误隐藏^[4]。时域错误隐藏算法的核心思想是利用前一帧的运动矢量去预测后一帧, 然而在很多情况下, 运动矢量不易得到, 因而导致时域错误隐藏难以实现, 本文主要针对空域错误隐藏进行研究。空域错误隐藏算法通常基于邻域信息相关的特性, 利用周围信息进行线性插值来恢复丢失的像素。但空域错误隐藏算法存在的问题是:

基金项目: 国家自然科学基金(61471271)。

作者简介: 严静文(1990—), 女, 硕士研究生, 主研方向为音频信号处理; 肖晶、高戈, 副教授。

收稿日期: 2018-01-15 **修回日期:** 2018-02-27 **E-mail:** yanjingwen@whu.edu.cn

对于不平滑的图像,甚至是丢失域占比很大的图像,用于插值的邻近像素的选择会受到限制,将会导致图像许多细节恢复的效果不好。

近年来,信号稀疏表达被广泛应用于图像处理中,现有稀疏表达图像恢复问题所用技术基于离散余弦变换(Discrete Cosine Transform, DCT)和快速傅里叶变换(Fast Fourier Transform, FFT)^[5],并且用于稀疏重建的字典是线下训练的,需要大量图片作为训练前提,实际应用场景中无法使用单张图像进行训练。文献[6]提出的利用双稀疏表达(Dual Sparse Representation, DSR)恢复丢失区域的方法使用 y 和 y_j 的欧式距离小于某一阈值来确定用于构建字典的潜在集合和模板集合,但由固定阈值带来的匹配块和潜在块选取不当会引起局部线性相关模型建立不准,并且局部线性相关模型的选取不当会造成稀疏重建的初始值不准。针对上述问题,本文提出基于稀疏优化的错误隐藏算法,将固定阈值改为自适应阈值,并利用典型相关分析(Canonical Correlation Analysis, CCA)作为局部线性相关模型,以提高稀疏优化的准确性。

1 现有研究

文献[7]提出一种凸集投影(Projections onto Convex Sets, POCS)约束方法。该方法利用先验知识及丢失块周围的信息确定凸集,并通过傅里叶变换以及傅里叶逆变换不断迭代最终确定丢失像素值,但 POCS 恢复的图像具有严重的块效应。文献[8]提出一种基于马尔可夫随机场(Markov Random Field, MRF)的错误隐藏技术,通过计算丢失宏块的 Sobel 算子来确定用于错误隐藏的像素位置,采用加权方式估计当前丢失的像素。文献[9]提出一种自适应序列插值(Orientation Adaptive Interpolation, OAI)方法,将 8 个方向的预测值加权作为丢失像素的估计值,该方法固定的扫描顺序会

导致错误扩散。文献[10]提出一种线性稀疏预测(Sparse Linear Prediction, SLP)方法,通过分析用于预测的权值与周围可知像素的关系,推断两者呈一定的指数关系,利用这一关系顺序地恢复出受损图像,从而恢复纹理信息,但是图像细节恢复效果不佳。文献[11]提出一种线性自适应预测(Adaptive Linear Prediction, ALP)方法,可以根据当前环境自适应地调整扫描顺序和潜在像素集的形状,但是丢失块内部像素的恢复效果不理想。目前基于稀疏表达的错误隐藏方法选择线下训练字典,需要大量图片作为训练前提,对于当前单张图像的恢复问题不适用,因而无法应用于实际传输中错误的恢复。文献[6]提出一种双稀疏表达的方法。该方法通过主成分分析(Principal Component Analysis, PCA)进行字典构建,解决了因数据不足造成字典训练无法完成的问题。上述方法存在的问题在于寻找用于重建的潜在像素不准,因而重建图像质量不高。本文提出的稀疏优化算法,利用自适应阈值提高搜索潜在像素集合和模板像素集合的准确性,改善稀疏重建的效果,使得受损图像的恢复效果更好。

2 基于双稀疏优化的空域错误隐藏算法

2.1 错误隐藏框架

在 H. 264/AVC 编码标准中,视频/图像通常被划分为独立的宏块来进行帧内/帧间预测。经过 H. 264/编码后的码流通常分片打包传输,丢失的一个数据包至少为一个 16×16 的宏块,本文针对 16×16 宏块独立丢失的情况进行研究。本文算法的具体框架如图 1 所示,针对 2×2 的小块进行错误隐藏,当输入一张受损图像时,本文算法将利用动态阈值确定与当前待恢复像素块对应的潜在像素块和模板像素块,并建立对应潜在像素块和模板像素块的字典,同时利用典型相关分析作为两者的局部线性相关模型,通过稀疏重建恢复出丢失的像素。

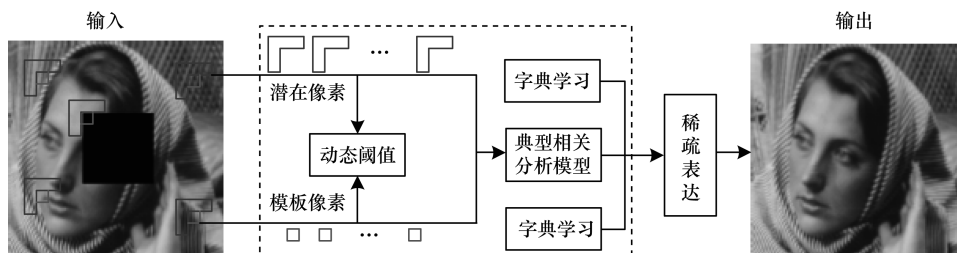


图 1 本文算法整体框架

2.2 图像生成

由于恢复 16×16 的块通常会造成错误扩散,因此本文采用与文献[6]相同的设置。具体设置如

图 2(a)所示, x 是丢失宏块中的一个 2×2 的小方块, y 表示 x 周围 6×6 区域中所有像素值已知的区域, x_j 和 y_j 所组成的区域形状与 x 和 y 组成的区域形

状完全相同, x_j 和 y_j 均为像素值已知的宏块。本文将所有像素值未知的区域定义为 L , 所有像素值已知的区域定义为 S 。由图 2(a) 可知, 丢失的宏块大小为 16×16 , 并划分为 64 个 2×2 的小方块, 小方块的恢复顺序如图 2(b) 所示, 较亮区域的像素将优先恢复, 并为后续像素的恢复提供最大的置信度, 较暗区域表示像素随之被恢复。

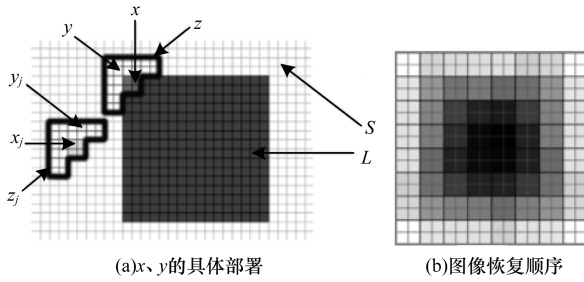


图 2 图像恢复过程

2.3 稀疏表示

研究表明, 图像块可以被表示为字典的线性稀疏形式^[12], 这样可以通过为每个丢失块建立稀疏表达模型来恢复丢失的像素。稀疏模型的一般表示为:

$$Y = D\alpha + \varepsilon \quad (1)$$

其中, Y 表示原始信号, $D = [d_1, d_2, \dots, d_p]$ 表示稀疏表达的字典矩阵, $d_i (i \in [1, p])$ 表示字典的基本向量, ε 表示残差。当 $\|\varepsilon_x\|_2 \ll \|\psi_x\|_2$ 并且 $\|\alpha\|_0 \ll k$ 时, 该模型是稀疏的。为了获得更高效的算法, 通常将 l_0 范式转换为 l_1 范式, 于是更新为如下形式:

$$\alpha^* = \min_{\alpha} \{ \|Y - D\alpha\|^2 + \lambda \|\alpha\|_1 \} \quad (2)$$

其中, λ 为正则化参数, 其目的是为了平衡重建效果和稀疏性。最终原始信号的重建值为:

$$Y^* = D\alpha^* \quad (3)$$

同理, 丢失的像素块 x 的重建值可以表示为:

$$x^* = d_x \xi_x^* \quad (4)$$

其中, d_x 表示用于重建 x 的字典, ξ_x^* 表示用于稀疏重建中字典 d_x 对应的系数矩阵。通过式(4)可知, 解决本文的关键在于字典 d_x 及字典 d_x 对应系数矩阵的求解。

2.3.1 集合匹配

为确定潜在集合与模板集合, 即找到与 x 对应的潜在像素块 x_j 以及与 y 对应的模板像素块 y_j 。通过计算 y 与 y_j 的欧式距离, 并使其小于某一阈值 τ 即可解决该问题^[13], 具体表示如下:

$$\begin{aligned} \phi_y &= \{y_j \mid \|y - y_j\|^2 < \tau, x_j, y_j \in (S \cup L)\} \\ \phi_x &= \{x_j \mid \|y - y_j\|^2 < \tau, x_j, y_j \in (S \cup L)\} \end{aligned} \quad (5)$$

其中, τ 表示一个常数。由于固定的阈值会造成某些块匹配的不准确, 这将严重影响字典构建以及线性相关模型的准确性, 从而影响重建图像质量, 因此对式(4)进行改进, 得到:

$$L_j = \|y - y_j\|^2 \quad (6)$$

其中, L_j 为 y 和 y_j 的欧式距离, 将所有的欧式距离进行升序排序, 获得排序后的欧式距离集合 $L = \{L_1, L_2, \dots, L_K, L_{K+1}, \dots, L_N\}$ 。最终的 K 个潜在像素块和模板像素块可以通过式(7)获得:

$$\begin{aligned} \phi_y &= \{y_j \mid \|y - y_j\|^2 < L_K, x_j, y_j \in (S \cup L)\} \\ \phi_x &= \{x_j \mid \|y - y_j\|^2 < L_K, x_j, y_j \in (S \cup L)\} \end{aligned} \quad (7)$$

其中, L_K 表示当前待恢复像素周围的 y 与潜在像素块 y_K 的欧式距离。由此可以得到最终的 K 个潜在像素块和模板像素块组成的集合分别为: 潜在集合 $\phi_y = \{y_1, y_2, \dots, y_K\} \in \mathbb{R}^{K \times N_y}$ 和模板集合 $\phi_x = \{x_1, x_2, \dots, x_K\} \in \mathbb{R}^{K \times N_x}$ 。

2.3.2 局部线性相关模型

CCA 作为多元统计学的一个重要部分, 是用于研究 2 组变量之间的一种统计分析方法, 能够有效地揭示 2 组变量之间的相互线性依赖关系, 通常应用于多重回归、判别分析和相应分析^[14]。给定两组变量 $U = \{u_1, u_2, \dots, u_m\}$ 和 $V = \{v_1, v_2, \dots, v_n\}$, CCA 旨在找到两组向量 w_u 和 w_v 使得 $w_u^T U$ 和 $w_v^T V$ 的相关性最大, U 和 V 的典型相关关系可以表示为:

$$\rho = \frac{\text{cov}(w_u^T U, w_v^T V)}{\sqrt{(w_u^T U U^T w_u)(w_v^T V V^T w_v)}} \quad (8)$$

其中, ρ 表示 U 和 V 的典型相关系数, $\text{cov}(w_u^T U, w_v^T V)$ 表示 $w_u^T U$ 和 $w_v^T V$ 的协方差, 将式(8)变形可得:

$$\begin{aligned} \min_{w_u, w_v} & \|w_u^T U - w_v^T V\|_F^2 \\ \text{s. t. } & w_u^T U U^T w_u = 1, w_v^T V V^T w_v = 1 \end{aligned} \quad (9)$$

受到 H. 264/AVC 中帧内预测的启发, 获知相邻像素之间存在强相关性。CCA 通常用于两组变量之间的线性关系, 鉴于典型相关分析可以找到使两组变量最相关的对应系数矩阵。本文利用典型相关分析作为局部线性相关模型, 即找到可以使 ϕ_y 和 ϕ_x 相关性最大的对应系数矩阵 W_y 和 W_x , 建立潜在集合 ϕ_y 和模板集合 ϕ_x 的相关关系, 相应的 CCA 模型为:

$$\begin{aligned} \min_{W_x, W_y} & \|W_x^T \phi_x \Lambda - W_y^T \phi_y \Lambda\|_F^2 \\ \text{s. t. } & W_x^T \Lambda \phi_x \phi_x^T \Lambda W_x = 1, W_y^T \Lambda \phi_y \phi_y^T \Lambda W_y = 1 \end{aligned} \quad (10)$$

其中, Λ 是一个对角矩阵, 反映了 y 和 y_j 的相似性, 表达式为:

$$\Lambda_{jj} = \exp\left\{-\frac{\|y - y_j\|^2}{\Sigma^2}\right\} \quad (11)$$

为了更好地变形式(10),使 $x' = x\Lambda, y' = y\Lambda$, $C_{xx}' = \Lambda x x^T \Lambda, C_{yy}' = \Lambda y y^T \Lambda$,将式(10)转换为拉格朗日函数:

$$L = \|W_x^T \phi_x' - W_y^T \phi_y'\|_F^2 + u(1 - W_x^T C_{xx} W_x) + u(1 - W_y^T C_{yy} W_y) \quad (12)$$

根据式(12),最终 x 的初值为:

$$x = W_x^{-1} W_y y \quad (13)$$

通过学习 ϕ_y 和 ϕ_x 的映射关系,可以得到 ϕ_y 和 ϕ_x 的稀疏表达关系为:

$$d_x \xi_x = W_x^{-1} W_y d_y \xi_y \quad (14)$$

式(14)可以转换为:

$$\xi_x = d_x^T W_x^{-1} W_y d_y \xi_y \quad (15)$$

为方便表示,令 $S = d_x^T W_x^{-1} W_y d_y$ 。

本文利用核主成分分析(Kernel Principal Component Analysis, KPCA)来学习潜在集合 ϕ_y 和模板集合 ϕ_x 对应的最相关字典 d_y 和 d_x 。KPCA 方法结合 KSVD 方法与 PCA 方法,可以降低构建字典的误差^[15]。

最终采用双字典优化来求解 ξ_x ,具体表达式为:

$$\min_{\xi_x, \xi_y} \left\{ \|x - d_x \xi_x\|_2^2 + \lambda \|\xi_x\|_F^2 + \|y - d_y \xi_y\|_2^2 + \lambda \|\xi_y\|_F^2 + \lambda \|\xi_x - S \xi_y\|_2^2 \right\} \quad (16)$$

其中, x 是最初的估计值。对式(16)进行求解, x 可以表示为:

$$x^* = d_x \xi_x \quad (17)$$

当恢复 x 后,剩余像素将按照图 2(b)所示的顺序依次进行恢复,先恢复较亮区域的像素,再恢复较暗区域的像素。

3 实验结果与分析

3.1 实验数据

为测试本文算法的有效性,在 25% 丢包率场景下选择错误隐藏算法常用的数据进行测试,测试图像包括 Barbara(512 像素 × 512 像素)、Cameraman(256 像素 × 256 像素)、Baboon(512 像素 × 512 像素)、Goldhill(512 像素 × 512 像素)、Butterfly(256 像素 × 256 像素)、Plants(256 像素 × 256 像素)、S-tarfish(256 像素 × 256 像素)、Leaves(256 像素 × 256 像素)、Peppers(512 像素 × 512 像素)、Lena(512 像素 × 512 像素),如图 3 所示。

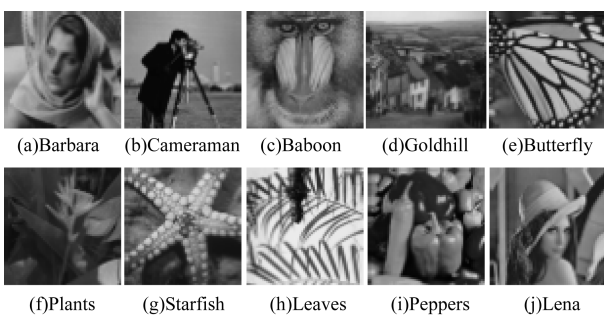


图 3 测试图像

3.2 参数设置

本文需要设置的参数如下:当利用动态阈值寻找最佳潜在像素 y_j 和模板像素 x_j 时,本文选出 $K = 40$ 个对应的潜在像素块和模板像素块。将式(11)中的 Σ^2 根据经验设置为: $\Sigma^2 = \text{length}(y)$ ^[9],即 Σ^2 的值为 y 区域中用于估计 x 的像素个数。在本文所有实验中, λ 值设置为 0.000 1。

3.3 实验结果

3.3.1 主观实验结果

为验证本文算法的有效性,选择 3 组主流算法与本文算法进行对比,包括 SLP-E^[10]、ALP^[11] 这 2 种线性插值算法及双稀疏表达 DSR 算法^[6]。图 4 给出了 Barbara 使用其他空域错误隐藏算法和本文算法对图像进行错误恢复的主观结果。从图 4(c)中可以看出,本文算法可以较好地恢复受损图像,尤其对受损图像眼睛和鼻子部分的恢复效果比其他算法好,恢复的图像更加接近原始图像。

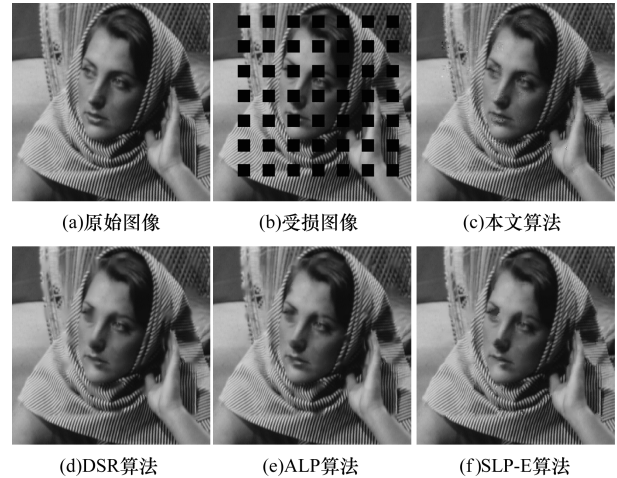


图 4 基于空域错误隐藏算法的 Barbara 恢复效果

3.3.2 客观实验结果

本文使用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和多尺度结构相似性(Multi-scale Structural Similarity Index, MS_SSIM)这 2 个客观评价指标对算法进行评价。PSNR 定义为:

$$PSNR = 10 \cdot \lg \frac{255^2}{MSE} \quad (18)$$

其中, MSE 是受损图像及原图像的均方差。

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2 \quad (19)$$

其中, m, n 分别表示图像的行列数。MS_SSIM 定义为:

$$SSIM(I, K) = \frac{(2u_I u_K + C)(2\sigma_{IK} + C)}{(u_I^2 + u_K^2 + C)(\sigma_I^2 + \sigma_K^2 + C)} \quad (20)$$

其中, u_I, u_K 分别表示图像 I 和图像 K 的均值, σ_I, σ_K 分别表示图像 I 和图像 K 的标准差, σ_{IK} 表示 2 幅图像的协方差, C 为常数。通过不同算法恢复的图像 PSNR 以及 MS_SSIM 如图 5、图 6 所示。

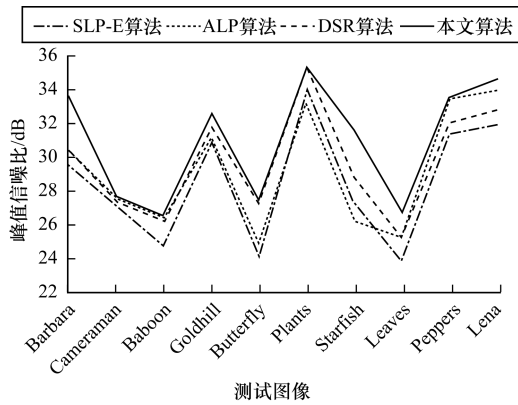


图 5 不同空域错误隐藏算法的 PSNR 对比

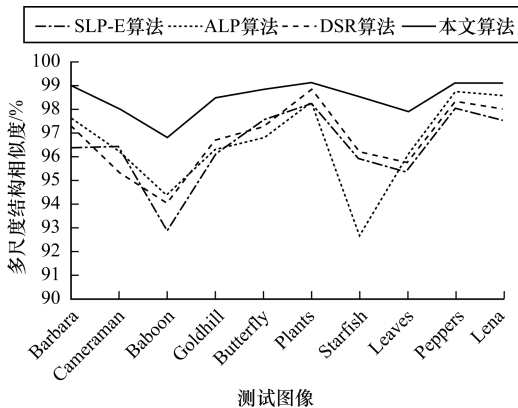


图 6 不同空域错误隐藏算法的 MS_SSIM 对比

从图 5 可以看出,本文算法的 PSNR 比 ALP 错误隐藏算法平均提高 1.9 dB,比 DSR 算法平均提高 1.23 dB。另外,本文算法对于 Barbara、Lena 的提升效果更明显,而对于 Camera、Plants 的提升幅度较小。这是由于 Cameraman 中的三脚架与地面的区别很小,在恢复过程中通常出现三脚架柱子截断的现象;而对于 Plants 来说,内部的白色区域在恢复时会发生边缘抖动的现象。从图 6 可以看出,本文算法的 MS_SSIM 对比 DSR 算法平均提高 1.72%。对于 Barbara、Plants、Peppers、Lena 这些图像,恢复图像的结构相似性高达 99%,由此可以看出本文算法可以更好地还原图像。

4 结束语

针对码流传输过程中图像部分像素丢失的情况,线性插值算法在处理非平滑图像时效果不理想,稀疏表达算法在构建算法字典时不能满足实时传输的需要。为解决上述问题,本文提出一种基于双稀疏优化的空域错误隐藏算法。该算法采用动态阈值搜索潜在像素集合与模板像素集合,利用 CCA 建立集合之间的关系,基于 KPCA 构建集合字典,通过稀疏表达重建丢失的像素。实验结果表明,本文算法恢复的图像主观效果相较于其他算法具有明显提升,错误隐藏效果较好。但该算法在搜索潜在像素集合与模板像素集合时使用多层循环,需要进行多

次迭代恢复,从而造成运算复杂度较高,因此下一步将对集合搜索算法进行优化,降低运算复杂度。

参考文献

- [1] WIEGAND T, SULLIVAN G J, BJNTGARD G. Overview of H.264/AVC video coding system [J]. IEEE Transactions on Circuits System Video Technology, 2003, 13 (7): 560-576.
- [2] 卢刘明,陆肖元. 基于网络丢包的网路视频质量评估[J]. 中国图像学报, 2009, 14(1): 52-58.
- [3] WANG Yao, WU Zhenyu, BOYCE J M. Modeling of transmission-loss-induced distortion in decoded video[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(6): 716-732.
- [4] 马宇峰,魏维,杨科利. 视频通信中的错误隐藏技术[M]. 北京:国防工业出版社,2007:122-156.
- [5] LI X. Image recovery via hybrid sparse representation: a deterministic annealing approach [J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 953-962.
- [6] LIU X, ZHAI D, ZHOU J. Sparsity-based image error concealment via adaptive dual dictionary learning and regularization [J]. IEEE Transactions on Image Processing, 2016, 26(2): 782-796.
- [7] SUN H, KWOK W. Concealment of damaged block transform coded images using projections onto convex sets [J]. IEEE Transactions on Image Processing, 1995, 4(4): 470-477.
- [8] SHIRANI S, KOSENTINI F, WARD R. An adaptive Markov random field based error concealment method for video communication in an error prone environment [C]// Proceedings of 1999 IEEE International Conference on the Acoustics, Speech, and Signal Processing. Washington D. C., USA: IEEE Press, 1999: 3117-3120.
- [9] LI X, ORCHARD M T. Novel sequential error-concealment techniques using orientation adaptive interpolation [J]. IEEE Transactions on Circuits System Video Technology, 2002, 12(10): 857-864.
- [10] KOLODA J, OSTERGAARD J, JENSEN S H, et al. Sequential error concealment for video/images by sparse linear prediction [J]. IEEE Transactions on Multimedia, 2013, 15(4): 957-969.
- [11] LIU J, ZHAI G, YANG X, et al. Spatial error concealment with an adaptive linear predictor [J]. IEEE Transactions on Circuits System Video Technology, 2015, 25(3): 353-366.
- [12] 练秋生,张伟. 基于图像块分类稀疏表示的超分辨率重构算法[J]. 电子学报, 2012, 40(5): 920-925.
- [13] ZHAI Deming, LIU Xianming, ZHOU Jiantao. Spatial error concealment via model based coupled sparse representation [C]// Proceedings of ICMEW '13. Washington D. C., USA: IEEE Press, 2013: 1-4.
- [14] HARDOON D, S ZEDMAK S, SHAWA-TAYLOR J. Canonical correlation analysis: an overview with application to learning methods [J]. Neural Computation, 2004, 16(12): 2639-2664.
- [15] 余付平,冯有前,范成礼. 基于主成分分析的字典学习[J]. 控制与决策, 2013, 28(7): 1109-1112.