

基于非局部均值和总变分最小化的单视频超分辨率算法

陈 诚¹ 常 侃^{1,2,3} 莫彩网¹ 李天亦¹ 覃团发^{1,2,3}

(广西大学计算机与电子信息学院 南宁 530004)¹

(广西大学广西多媒体通信与网络技术重点实验室 南宁 530004)²

(广西大学广西高校多媒体通信与信息处理重点实验室 南宁 530004)³

摘 要 传统的基于重建的单视频超分辨率方法能够获得较好的重建效果。然而,已有算法没有充分利用视频内的帧间、帧内相关性,重建效果仍有待提升。针对这一问题,提出了一种新的单视频超分辨率算法。为充分利用帧内相关性,采用非局部均值模型表征帧内非局部结构特性,采用总变分模型表征帧内局部结构特性;为了探索帧间相关性,采用光流法进行帧间预测。最后,为了求解所建立的优化问题,提出了基于 split-Bregman 方法的快速迭代算法。实验结果表明,与同类算法相比,所提算法在主、客观质量上均有相应的提升。

关键词 视频超分辨率,非局部均值,总变分,光流法

中图分类号 TP751 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.03.042

Single Video Super-resolution Algorithm Based on Non-local Means and Total Variation Minimization

CHEN Cheng¹ CHANG Kan^{1,2,3} MO Cai-wang¹ LI Tian-yi¹ QIN Tuan-fa^{1,2,3}

(School of Computer and Electronic Information, Guangxi University, Nanning 530004, China)¹

(Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning 530004, China)²

(Guangxi Colleges and Universities Key Laboratory of Multimedia Communications and Information Processing, Guangxi University, Nanning 530004, China)³

Abstract The traditional reconstruction-based single video super-resolution algorithms are able to solve the video super-resolution problem well. However, the existing algorithms have not fully exploited the correlation in intra-frames and inter-frames, which leaves much space for further improvement. This paper proposed a new single video super-resolution algorithm to solve this problem. When exploiting the spatial correlations, the non-local means model is used to get the non-local structural property and the total variation model is utilized to get the local structural property. In order to exploit inter-frame correlation, optical flow method is applied to perform inter-frame estimation. Finally, to solve the established optimization problem, a split-Bregman method based fast iteration algorithm was proposed. The experimental results demonstrate the effectiveness of the proposed algorithm. Compared with other algorithms, the proposed algorithm is able to achieve better subjective and objective results.

Keywords Video super-resolution, Non-local means, Total variation, Optical flow

1 引言

单视频超分辨率(Super-Resolution, SR)技术可以从单个低分辨率(Low Resolution, LR)视频中恢复出高分辨率(High Resolution, HR)视频。现有的单视频 SR 方法可以大致分为基于学习的方法和基于重建的方法。

在基于学习的方法中,通常需要从外部数据库中学习出 LR 图像与 HR 图像之间的映射关系。例如,徐煜明等人^[1]基于亚像素图像块进行配准,并依据配准结果构造 LR 字典; Hung 等人^[2-3]利用混合分辨率视频中的关键帧训练得到相

应的字典,再采用 LR 帧与字典匹配生成 HR 帧;李玲慧等人^[4]采用神经网络方法学习拟合参数并进行快速的 HR 视频重建,同时利用 HR 视频内的时空特征计算得到最终的 HR 重建结果。近年来,生成对抗网络(Generative Adversarial Network, GAN)技术逐渐被应用于超分辨率领域。Goldstein 等人^[5]通过将一系列低分辨率图像和与之对应的高分辨率图像作为训练数据,学习一个由卷积神经网络表示的映射函数,实现了单幅图像的超分辨率。虽然基于学习的 SR 方法取得了一定的成功,但是该类方法的性能依赖于目标图像与外部图像集之间的相似性,且该类方法对噪声的鲁棒性也不强。

收稿日期:2017-04-06 返修日期:2017-06-22 本文受国家自然科学基金项目(61761005, 61401108),广西自然科学基金项目(2016GXNSFAA380154)资助。

陈 诚(1993—),男,硕士生,主要研究方向为图像及视频超分辨率,E-mail:vigil1993@163.com;常 侃(1983—),男,博士,副教授,CCF 会员,主要研究方向为多媒体通信、压缩感知、稀疏表示等,E-mail:pandack0619@163.com(通信作者);莫彩网(1993—),女,硕士生,主要研究方向为图像与视频处理;李天亦(1993—),男,硕士生,主要研究方向为图像复原;覃团发(1966—),男,博士,教授,主要研究方向为多媒体通信、网络编码、图像和视频检索等。

在基于重建的方法中,通常将 HR 图像的重建过程视为反问题,采用多种先验信息对解空间进行约束。相较于基于学习的方法,该方法具有较好的鲁棒性。总变分(Total Variation, TV)模型^[6]是一种简单而有效的先验约束模型。Chen 等人^[7]提出了一种针对 TV 最小化变分模型的快速解法。Farsiu 等人^[8]对 TV 模型进行了改进,提出了双边(Bilateral) TV 模型,并采用梯度下降法进行求解。Yuan 等人^[9]提出根据图像的空域特征对 TV 模型的加权参数进行自适应调整。TV 模型属于空域约束。为了综合应用视频内的时域相关性,Belekos 等人^[10]提出了一种基于空间自适应多信道先验的视频 SR 算法,并将帧间相关性考虑其中,但是,该方法所用的帧间相关性约束不适用于包含复杂或剧烈运动的场景。文献^[11]提出了一种视频内的帧间相似性约束,采用时域参考帧上的相似像素点对插值结果进行非局部均值滤波。Zhu 等人^[12]采用了一种基于特征引导的变分光流法来探索帧间相关性,但是所提出的约束模型未考虑图像内的结构特性。Li 等人^[13]提出了一种基于自适应超像素引导的自回归(Adaptive Superpixel-guided Auto-Regressive, ASAR)模型,该模型联合使用光流法和 AR 模型探索非关键帧的时空相关性。但是由于该算法采用的 AR 模型仅表征了帧内的局部相关性,缺少对非局部相关性的表征,因此其 SR 效果存在可提升的空间。为解决文献^[13]中存在的问题,本文提出一种新的基于重建的单视频 SR 算法。该算法采用非局部均值(Non-Local Means, NLM)模型与 TV 模型替代文献^[13]中提出的 AR 模型作为约束项,实现了对帧内空间相关性的建模,从而实现了对视频序列更有效的 SR 重建,获得了更好的主观效果。

2 相关背景知识

2.1 TV 模型与 NLM 模型

TV 模型由 Rudin 等人^[6]首次提出。假定存在一个 HR 图像 x , 则该图像的 TV 模型为:

$$D(x) = \sum_i \sqrt{(\Delta_h^v x)^2 + (\Delta_v^v x)^2} \quad (1)$$

其中, $\Delta_h^v x$ 和 $\Delta_v^v x$ 分别为水平和垂直方向上的一阶差分算子, 即对像素 i , 有 $\Delta_h^v x = x_i - x_{r(i)}$ 和 $\Delta_v^v x = x_i - x_{b(i)}$, 其中 $r(i)$ 和 $b(i)$ 分别表示像素点 i 右侧和下方最接近于 i 的像素点。

在 Rudin 等人提出的各向同性 TV 模型的基础上, Esedoglu 等人^[14]提出了各向异性 TV 模型:

$$D(x) = \sum_i |\Delta_h^v x| + |\Delta_v^v x| \quad (2)$$

NLM 模型由 Buades 等人^[15]提出, 最早被应用于视频和图像的去噪。该方法的基本思想是, 当前像素的估计值由图像中与其具有相似邻域的像素取加权平均得到。假设含噪声图像为 $v = \{v(i) | i \in I\}$, 像素 i 的估计值 $u(i)$ 由图像中所有像素值的加权平均得到:

$$u(i) = \sum_{j \in I} s(i, j) \times v(j) \quad (3)$$

加权值 $s(i, j)$ 表示像素点 i 与 j 之间的相似度, 其值取决于灰度级矢量 $V(\xi_i)$ 和 $V(\xi_j)$, 其中 ξ_i 和 ξ_j 分别为图像中以 i 和 j 为中心、大小固定的像素块。加权值 $s(i, j)$ 的计算过程如下:

$$s(i, j) = \frac{1}{Z(i)} \exp\left(-\frac{\|V(\xi_i) - V(\xi_j)\|_2^2}{h^2}\right) \quad (4)$$

其中, $\|V(\xi_i) - V(\xi_j)\|_2^2$ 为两个像素块之间的欧氏距离。归一化系数为:

$$Z(i) = \sum_j \exp\left(-\frac{\|V(\xi_i) - V(\xi_j)\|_2^2}{h^2}\right) \quad (5)$$

其中, h 为平滑参数, 控制高斯函数的衰减程度, h 越大, 高斯函数变化越平缓, 去噪水平越高, 但同时图像也会越模糊。

2.2 基于 ASAR 模型的视频超分辨率算法

基于 ASAR 模型的视频超分辨率算法由 Li 等人^[13]提出, 其在 Song 等人^[16]提出的基于关键帧重建的视频 SR 算法的基础上, 实现了对帧间、帧内相关性的建模。

具体地, 首先将输入的 LR 视频序列划分为关键帧和非关键帧, 采用 Kim 等人^[17]提出的基于岭回归(Ridge Regression)的单图像 SR 方法实现关键帧的 SR, 从而得到 HR 关键帧; 然后利用距离当前非关键帧最近的关键帧来预测当前非关键帧, 其过程可以表达为:

$$\tilde{U}_{i,j}^{\text{non}} = aU_{r_1, s_1}^{\text{key}} + bU_{r_2, s_2}^{\text{key}} + cU_{r_3, s_3}^{\text{key}} \quad (6)$$

其中, U^{key} 为距离最近的 HR 关键帧; (r_1, s_1) , (r_2, s_2) 和 (r_3, s_3) 分别为距离像素点 (i, j) 最近的 3 个点, 且 $r = i + O_r^h$, $s = j + O_s^v$, O_r^h 和 O_s^v 分别为水平和垂直位移距离; a, b, c 则为加权参数。则非关键帧的 SR 求解过程等价于以下优化问题:

$$\hat{U} = \arg \min_{U^{\text{non}}} \left(\frac{1}{2} \|U^{\text{non}} - HU^{\text{non}}\|_2^2 + \frac{\alpha}{2} \|U^{\text{non}} - \tilde{U}^{\text{non}}\|_2^2 + \frac{\beta}{2} \|U^{\text{non}} - AU^{\text{non}}\|_2^2 \right) \quad (7)$$

其中, 第一项为数据保真项, H 为下采样矩阵; 第二项为帧间相关性约束项; 最后一项为非关键帧上的 AR 模型约束项, A 为 AR 模型的系数。通过求解式(7), 即可得到非关键帧的 SR 结果。

3 基于 NLM 模型和 TV 模型的视频超分辨率算法

由 2.2 节可知, 视频序列中存在帧间、帧内相关性, 文献^[13]将 LR 视频序列分为关键帧和非关键帧两部分进行处理。对关键帧使用单图像 SR 方法实现上采样, 再利用光流法进行帧间预测, 同时采用 AR 模型作为约束项实现非关键帧的 SR。但是由于 AR 模型存在局限性, 导致文献^[13]中的方法存在一定不足。因此, 本文提出一种基于 NLM 模型和 TV 模型的视频 SR 算法, 以有效解决文献^[13]中算法存在的问题。

3.1 关键帧超分辨率

由于利用单帧或者多帧 LR 图像 SR 的方法存在运算复杂度较高且很难获得高质量的 HR 图像的问题, 因此 Brandi 等人^[18]假定输入为含有少量 HR 帧的混合分辨率视频序列, 并通过提取 HR 关键帧中有效的高频信息实现非关键帧的 SR, 获得了较好的视频 SR 效果。而在现实生活中, 较难满足将混合分辨率视频序列作为输入这个条件, 因此本文自行指定关键帧, 并采用文献^[17]提出的单图像 SR 方法对关键帧实现上采样。该方法具有相对较低的运算复杂度, 且在保证良好的 SR 效果的同时, 运算时间短, 能满足视频 SR 对运算时间的要求。

3.2 非关键帧超分辨率

文献^[13]中提出的 ASAR 模型为表征帧内的结构特性, 将 AR 模型作为约束项构建最优化问题, 但是由于 AR 模型采用局部邻域的像素值通过加权来表示当前像素点的值, 仅探索了帧内的局部结构特性, 缺少对非局部结构特性的表征,

并未充分利用图像的结构特征,因此其 SR 效果仍有可提升的空间。

为解决此问题,希望采用一种新的表征帧内相关性的模型,以同时描述帧内的局部相似性及非局部相似性。本文提出了基于 NLM 模型和 TV 模型的单视频 SR 算法,其中 NLM 模型用于探索帧内的非局部结构特性,过利用像素周围固定大小的窗口内的信息来表征该像素的特征;TV 模型用于探索帧内的局部结构特性,在图像重建时依赖于图像的结构特征,能在去除图像噪声的同时保护图像边缘。将 NLM 模型和 TV 模型相结合来作为 SR 重建框架中的一个约束项,通过最小化约束项,可以在保证去噪效果的同时保留视频帧中的细节信息。

改进并引入 NLM 模型和 TV 模型后,为保证表达的简洁,用 U 替换式(7)中的 U^{non} ,并用范数形式将其表达为:

$$\hat{U} = \arg \min_U \left(\frac{1}{2} \|Y - HU\|_2^2 + \frac{\alpha}{2} \|U - E\|_2^2 + \frac{\beta}{2} \|U - SU\|_2^2 + \gamma \|DU\|_1 \right) \quad (8)$$

其中, α, β, γ 为加权参数, Y 为原始 LR 帧, H 为下采样矩阵, E 为待重建非关键帧的预测值, S 为 NLM 滤波系数矩阵, $D = [D_h^T, D_v^T]^T$, D_h 和 D_v 分别为水平方向和垂直方向的梯度算子。

在式(8)中,第一项约束是数据保真项,强调待重建的非关键帧经过下采样后应与 LR 测量值相近;第二项约束强调待重建的非关键帧应与其预测结果 E 相近;第三项约束强调待重建的非关键帧应与其 NLM 滤波结果接近;第四项约束强调待重建的非关键帧应具有较小的 TV 范数值。

越准确的 E 越有助于获取更准确的非关键帧重建结果。本文采用如图 1 所示的方法来计算单个非关键帧的预测值 E 。首先,采用传统插值方法得到初始的 HR 非关键帧;其次,采用光流法,以前后相邻的两个 HR 关键帧为参考帧,对当前非关键帧做双向运动估计及运动补偿。仅利用关键帧进行运动补偿的好处是,能够有效限制重建误差在视频序列中的扩散。

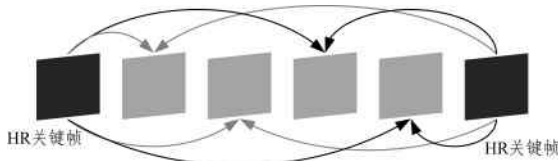


图 1 采用光流法对非关键帧进行运动估计及运动补偿

Fig. 1 Motion estimation and motion compensation of non-key frames by using optical flow method

为在降低运算复杂度的同时保证运算速度,基于 split-Bregman^[19]迭代算法求解式(8)。首先引入一个中间变量 d ,并令 $d = DU$,再将式(8)分解为 U 和 d 两个子问题,并迭代求解,具体计算公式如下:

$$\begin{cases} U^{k+1} = \arg \min_U \left(\frac{1}{2} \|Y - HU\|_2^2 + \frac{\alpha}{2} \|U - E\|_2^2 + \frac{\beta}{2} \|U - S^k U\|_2^2 + \frac{\eta}{2} \|d^k - DU - b^k\|_2^2 \right) \\ d^{k+1} = \arg \min_U \left(\frac{\eta}{2} \|d^k - DU^{k+1} - b^k\|_2^2 + \gamma \|d\|_1 \right) \\ b^{k+1} = b^k + DU^{k+1} - d^{k+1} \end{cases} \quad (9)$$

其中, η 为加权参数, b 是 Bregman 迭代中间变量, NLM 滤波系数矩阵 S 在迭代中更新。

在式(9)中,直接求解“ U ”子问题的复杂度较高,因此采用上一次的迭代结果 U^k 计算 NLM 滤波结果。由此,“ U ”子问题被更改为:

$$U^{k+1} = \arg \min_U \left(\frac{1}{2} \|Y - HU\|_2^2 + \frac{\alpha}{2} \|U - E\|_2^2 + \frac{\beta}{2} \|U - S^k U^k\|_2^2 + \frac{\eta}{2} \|d^k - DU - b^k\|_2^2 \right) \quad (10)$$

在求解“ U ”子问题时,通过对式(10)求导,并令导数等于 0,可以得到关于 U 的闭式解:

$$U^{k+1} = \frac{\alpha E + H^T Y + \beta S^k U^k + \eta D^T (d^k - b^k)}{(\alpha + \beta) I + H^T H + \eta D^T D} \quad (11)$$

但是,直接计算式(11)的复杂度仍然较高,因此采用共轭梯度法进行计算。

在求解“ d ”子问题时,可直接写出该问题的解为:

$$d^{k+1} = \text{shrink}(DU^{k+1} + b^k, \gamma/\eta) \quad (12)$$

其中, $\text{shrink}(\cdot)$ 为收缩算子,对于输入 x 与阈值 T ,其定义如下:

$$\text{shrink}(x, T) = \max(|x| - T, 0) \odot \text{sgn}(x) \quad (13)$$

至此,最优化问题(8)的求解推导完成。

为了便于阅读,将上述算法的流程总结为算法 1。将初始图像 U^0 、待重建非关键帧的预测值 E 、原始 LR 帧 Y 、下采样矩阵 H 、NLM 模型系数矩阵 S 、TV 模型梯度计算矩阵 D 和约束项参数 $\alpha, \beta, \gamma, \eta$ 作为算法的输入;初始化条件为 $d^0 = DU^0, b^0 = 0$;迭代次数的上限为 K ;根据式(11)计算 U^{k+1} ,根据式(12)计算 d^{k+1} ,根据式(9)更新 U^{k+1} ;为保证运算速度,将该算法的收敛条件设定为 $\|U^{k+1} - U^k\|_2 / \|U^{k+1}\|_2 < 10^{-4}$ 。

算法 1 基于 TV 模型和 NLM 模型的视频 SR 算法

输入: $E, Y, H, S^0, D, U^0, \alpha, \beta, \gamma, \eta$

输出: U

初始化: $d^0 = DU^0, b^0 = 0$

For $k = 0, 1, 2, \dots, K$

 根据式(11)计算 U^{k+1} ;

 根据式(12)计算 d^{k+1} ;

 计算 $b^{k+1} = b^k + DU^{k+1} - d^{k+1}$,得到 b^{k+1} ;

 更新 NLM 系数矩阵 S^k ;

 If $\|U^{k+1} - U^k\|_2 / \|U^{k+1}\|_2 < 10^{-4}$ Break

End For

4 实验结果与分析

为了验证所提出的视频 SR 算法的有效性,分别使用 5 个分辨率为 352×288 的视频序列 foreman, football, news, mobile 和 flower 进行实验,并与 Li C 的算法^[7]、Farisu 的算法^[8]、Zhu 的算法^[12]和 Li K 的算法^[13]在同等实验条件下进行比较。选取各序列的前 25 帧进行实验;所选取的 5 个视频序列中包含了不同快慢的运动以及不同程度的图像细节,因此能够较好地测试各类算法的性能。算法中关键帧之间的间隔设定为 4,正则化参数设置为 $\alpha = 0.02, \beta = 0.04, \gamma = 0.003, \eta = 0.001$ 。其中, α, β 和 γ 分别代表了帧间相似性约束、帧内非局部相似性约束以及帧内局部相似性约束的强度。

一般来说,为了取得较好的结果,数据保真项约束强度应足够大,因为在式(8)中,数据保真项的加权系数已固定为0.5,所以 α 、 β 和 γ 选择了较小的值。目前, α 、 β 和 γ 的取值是通过在不同序列上进行遍历测试得到的最佳经验值。由式(12)可知,参数 γ 与 η 的比值是收缩算子的输入阈值。根据实验经验,选取的该阈值在范围[2.0,5.0]内可取得较为理想的结果,因此本文选取的阈值为3.0,由此计算得到 η 为0.001。

在实验中,首先对原始HR视频序列中的每一帧使用Bicubic方法下采样,得到LR视频序列中的各帧(下采样比例分别为2,3,4);然后将LR视频序列作为各个算法的输入,分别使用4种算法和本文提出的算法进行处理。在本文实验中,对YUV格式的视频进行处理。对于Y分量信号,分别采用文献[7]、文献[8]、文献[12]、文献[13]以及本文算法进行SR重建,对色度分量则统一采用Bicubic方法进行SR重建。本文实验选用YUV模式而不是RGB模式的原因在于:1)文献[7]、文献[8]、文献[12]、文献[13]以及本文算法都没有探索色彩通道间的相关性,若在RGB模式下进行实验,需要轮流在R,G,B3个色彩通道上进行SR重建,会极大地增加运算复杂度;2)YUV模式将亮度分量和色度分量分离,因为人眼对亮度分量比较敏感,对色度分量不敏感,所以可以对亮度分量采用较为复杂的方法进行重建,而对色度分量采用简单的插值算法,从而在不明显损失重建质量的前提下有效提升运算速度。这样的处理方式也与相关文献中的实验设置一致。

表1和表2分别列出了文献[7]、文献[8]、文献[12]、文献[13]以及本文提出的算法在相同的实验条件下SR结果的Y分量的峰值信噪比(Peak Signal to Noise Ratio, PSNR)值与结构相似性(Structural Similarity Index, SSIM)值的比较。表中所列数值为序列中各帧数值的平均值。

表1 不同上采样率和视频序列下的PSNR比较

Table 1 Comparison of PSNR under different sampling rates and video sequences

(单位: dB)

序列	上采样率	文献 [7]	文献 [8]	文献 [12]	文献 [13]	本文 算法
foreman	×2	36.96	35.40	37.08	37.13	37.33
	×3	33.70	32.34	34.26	34.30	34.39
	×4	30.44	29.95	31.86	31.94	31.95
football	×2	31.76	31.63	33.25	33.21	33.36
	×3	27.69	27.14	27.93	28.07	28.11
	×4	25.58	25.25	25.85	25.91	25.94
news	×2	32.68	31.70	33.78	33.90	34.06
	×3	28.54	27.67	28.79	28.91	29.02
	×4	25.67	25.31	25.86	25.94	25.95
mobile	×2	24.94	24.67	25.68	26.08	26.27
	×3	21.85	21.43	22.04	22.20	22.28
	×4	20.58	20.29	20.71	20.78	20.76
flower	×2	24.92	24.92	25.27	25.31	25.50
	×3	22.83	22.78	22.97	22.99	23.11
	×4	21.71	21.77	21.91	21.92	21.94
平均	×2	30.25	29.66	31.01	31.13	31.30
	×3	26.92	26.27	27.20	27.29	27.38
	×4	24.80	24.51	25.24	25.30	25.31

表2 不同上采样率和视频序列下的SSIM比较

Table 2 Comparison of SSIM with different sampling rates and video sequences

序列	上采样率	文献 [7]	文献 [8]	文献 [12]	文献 [13]	本文 算法
foreman	×2	0.9604	0.9399	0.9585	0.9586	0.9591
	×3	0.9227	0.8981	0.9233	0.9240	0.9244
	×4	0.8799	0.8534	0.8886	0.8896	0.8896
football	×2	0.9231	0.9145	0.9340	0.9311	0.9344
	×3	0.8005	0.7780	0.8030	0.8039	0.8074
	×4	0.7016	0.6787	0.7036	0.7052	0.7078
news	×2	0.9619	0.9401	0.9676	0.9681	0.9683
	×3	0.9033	0.8707	0.9052	0.9071	0.9076
	×4	0.8335	0.8053	0.8421	0.8435	0.8439
mobile	×2	0.8491	0.8211	0.8667	0.8720	0.8769
	×3	0.6678	0.6212	0.6740	0.6808	0.6846
	×4	0.5534	0.5155	0.5623	0.5669	0.5645
flower	×2	0.8509	0.8308	0.8572	0.8574	0.8658
	×3	0.7328	0.7156	0.7373	0.7368	0.7425
	×4	0.6545	0.6429	0.6579	0.6589	0.6598
平均	×2	0.9091	0.8893	0.9168	0.9174	0.9209
	×3	0.8054	0.7767	0.8086	0.8105	0.8133
	×4	0.7246	0.6992	0.7309	0.7328	0.7331

从表1、表2中可以看出,在不同上采样率下,本文提出的算法都能有效提高视频序列的分辨率,得到较好的客观质量;与其他4种算法相比,所提算法在大部分情况下可以获得最高的PSNR值和SSIM值。例如,在上采样率为2时,对于foreman序列,本文的算法相较于文献[7]、文献[8]、文献[12]、文献[13]等的算法,PSNR值分别提高了0.37dB, 1.93dB, 0.25dB和0.20dB。但上采样率为4时,在mobile序列下,本文算法的客观效果会比文献[13]的算法稍差,主要原因在于mobile序列细节丰富,在上采样率较高时,NLM模型加权值的计算准确度降低,因此在一定程度上影响了算法性能。

从主观视觉质量上评价5种算法的优劣,图2和图3分别给出了flower序列和mobile序列在上采样率为2时,不同算法得到的HR重建图像。

从SR效果图及其放大区域可清楚地看出,相较于其他几种方法,本文算法获得了更好的主观质量且保留了更多的纹理细节。

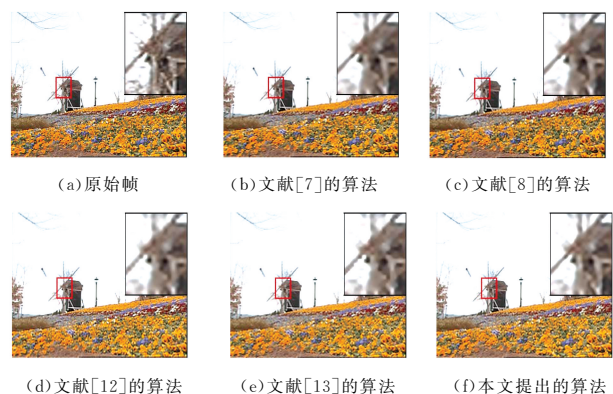


图2 上采样比例为2时flower视频序列上不同算法的SR结果的对比

Fig. 2 Comparison of SR results of different algorithms on flower sequence while up-sampling is 2



图3 上采样比例为2时 mobile 视频序列上不同算法的 SR 结果的对比

Fig. 3 Comparison of SR results of different algorithms on mobile sequence while up-sampling is 2

综上所述,可以总结出:

1) 文献[7]、文献[8]和文献[12]的算法没有有效应用视频内的时空相关性,导致其 SR 效果不佳。

2) 相较于文献[7]及文献[8]的算法,文献[13]中算法的质量获得了较明显的提升;文献[13]的算法效果也优于文献[12]的结果,原因在于 ASAR 模型对空域和时域相关性的建模较为准确。

3) 相较于文献[13],本文采用了 NLM 模型与 TV 模型替代了 AR 模型,更准确地挖掘了图像帧内的结构特征,从而得到了最佳的 SR 效果。需要注意的是,本文算法在文献[13]的基础上获得的 PSNR 及 SSIM 增益还不够明显,主要原因在于各帧图像内的空域相关性已经得到了较好的表达,从而在一定程度上限制了算法性能的提升幅度。

结束语 本文提出了一种基于 NLM 模型和 TV 模型的视频 SR 算法。一方面采用 NLM 模型和 TV 模型表征帧内结构特性;另一方面采用光流法进行帧间预测,探索帧间相关性。从实验结果可以看出,与目前比较经典的视频 SR 算法如文献[7]、文献[8]、文献[12]及文献[13]的算法相比,本文算法的 SR 结果在主观视觉质量和客观评价指标上均更优。尽管本文已在单视频 SR 问题上证明了所提模型的有效性,但对于其他应用,如视频去噪、去模糊、多视频 SR 问题等,本文模型的效果未知;另外,本文算法并没有探索色彩通道之间的相关性。在未来的工作中,将进一步对上述问题进行研究和探讨。

参考文献

- [1] XU Y M, SONG J W, XIAO X J. Super resolution algorithm based on sub-pixel block matching and dictionary learning[J]. Computer Science, 2016, 43(8): 304-308. (in Chinese)
徐煜明,宋佳伟,肖贤建.基于亚像素块匹配和字典学习的超分辨率算法[J].计算机科学,2016,43(8):304-308.
- [2] HUNG E M, GARCIA D C, DE QUEIROZ R L. Example-based enhancement of degraded video[J]. IEEE Signal Processing Letters, 2014, 21(9): 1140-1144.
- [3] HUNG E M, DE QUEIROZ R L, BRANDI F, et al. Video super-resolution using codebooks derived From key-Frames[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 9(22): 1321-1331.
- [4] LI L H, DU J P, LIANG M Y, et al. Video super-resolution algorithm based on spatial-temporal feature and neural network [J]. Journal of Beijing University of Posts and Telecommunications, 2016, 39(4): 1-6. (in Chinese)
李玲慧,杜军平,梁美玉,等.基于时空特征和神经网络的视频超分辨率算法[J].北京邮电大学学报,2016,39(4):1-6.
- [5] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [J]. arXiv preprint arXiv:1609.04802, 2016.
- [6] RUDIN L I, OSHER S, FATEMI E. Nonlinear total variation based noise removal algorithms[J]. Physica D: Nonlinear Phenomena, 1992, 60(1): 259-268.
- [7] LI C, YIN W, JIANG H, et al. An efficient augmented Lagrangian method with applications to total variation minimization [J]. Computational Optimization and Applications, 2013, 56(3): 507-530.
- [8] FARSIU S, ROBINSON M D, ELAD M, et al. Fast and robust multiframe super resolution[J]. IEEE Transactions on Image Processing, 2004, 13(10): 1327-1344.
- [9] YUAN Q, ZHANG L, SHEN H. Regional spatially adaptive total variation super-resolution with spatial information filtering and clustering [J]. IEEE Transactions on Image Processing, 2013, 22(6): 2327-2342.
- [10] BELEKOS S P, GALATSANOS N P, KATSAGGELOS A K. Maximum a posteriori video super-resolution using a new multi-channel image prior[J]. IEEE Transactions on Image Processing, 2010, 19(6): 1451-1464.
- [11] ZHANG Y L, GAN Z L, ZHU X C. Video super-resolution method based on similarity constraints[J]. Journal of Image and Graphics, 2013, 18(7): 761-767. (in Chinese)
张义轮,干宗良,朱秀昌.相似性约束的视频超分辨率重建[J].中国图象图形学报,2013,18(7):761-767.
- [12] ZHU Y, LI K, JIANG J. Video super-resolution based on automatic key-frame selection and feature-guided variational optical flow[J]. Signal Processing: Image Communication, 2014, 29(8): 875-886.
- [13] LI K, ZHU Y, YANG J, et al. Video super-resolution using an adaptive superpixel-guided auto-regressive model [J]. Pattern Recognition, 2016, 51(C): 59-71.
- [14] ESEDOĞLU S, OSHER S J. Decomposition of images by the anisotropic Rudin-Osher-Fatemi model [J]. Communications on Pure and Applied Mathematics, 2004, 57(12): 1609-1626.
- [15] BUADES A, COLL B, MOREL J M. A non-local algorithm for image denoising [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005 (CVPR 2005). IEEE, 2005: 60-65.
- [16] SONG B C, JEONG S C, CHOI Y. Video super-resolution algorithm using bi-directional overlapped block motion compensation and on-the-fly dictionary training [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(3): 274-285.
- [17] KIM K I, KWON Y. Single-image super-resolution using sparse regression and natural image prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(6): 1127-1133.
- [18] BRANDI F, QUEIROZ R, MUKHERJEE D. Super-resolution of video using key frames and motion estimation [C] // 2008 15th IEEE International Conference on Image Processing. IEEE, 2008: 321-324.
- [19] GOLDSTEIN T, OSHER S. The split Bregman method for L1-regularized problems [J]. SIAM Journal on Imaging Sciences, 2009, 2(2): 323-343.