

## 基于残差字典及协作表达的单图像超分辨率算法

田旭<sup>1</sup> 常侃<sup>1,2,3</sup> 黄升<sup>1</sup> 覃团发<sup>1,2,3</sup>

1 广西大学计算机与电子信息学院 南宁 530004

2 广西大学广西多媒体通信与网络技术重点实验室 南宁 530004

3 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

(tianxu775825@163.com)

**摘要** 通过传统的单图像超分辨率(Super Resolution, SR)算法重建的高分辨率图像往往存在高频信息不足、边缘模糊的问题。为了提升重建图像的质量,提出了一种基于残差字典及协作表达的单图像 SR 算法(Residual Dictionary and Collaborative Representation, RDCR)。在训练环节,该算法结合字典学习及协作表达的思想,首先训练一个主字典及主投影矩阵,其次利用重建的样本图像训练多层残差字典及多层残差投影矩阵;在测试环节,通过逐层重建残差信息,得到不断精细化的高频信息,以提升重建的高分辨率图像的质量。通过实验证明,相比传统算法 A+,所提算法在 4 倍上采样下的 Set5 及 Set14 图像集上可以分别获得 0.20 dB 及 0.18 dB 的峰值信噪比增益,在运算时间上所提算法与 A+ 接近。

**关键词** 超分辨率;字典学习;协作表达;稀疏表示

中图法分类号 TP751

## Single Image Super-resolution Algorithm Using Residual Dictionary and Collaborative Representation

TIAN Xu<sup>1</sup>, CHANG Kan<sup>1,2,3</sup>, HUANG Sheng<sup>1</sup> and QIN Tuan-fa<sup>1,2,3</sup>

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

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

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

**Abstract** Usually, the traditional single image super resolution (SR) algorithms generate the high resolution (HR) images with insufficient high-frequency information and blurred edges. To improve the quality of the reconstructed HR images, this paper proposes a single image SR algorithm by using residual dictionary and collaborative representation (Residual Dictionary and Collaborative Representation, RDCR). In the training phase, firstly, based on the ideas of dictionary learning and collaborative representation, a main dictionary and the corresponding main projection matrices are learned. After that, the reconstructed image samples are utilized to train multiple layers of residual dictionaries and residual projection matrices. In the testing phase, high-frequency information is gradually refined by reconstructing the residual information layer by layer. Extensive experimental results show that, at a scale factor of 4, the average peak signal-to-noise ratio (PSNR) values obtained by the proposed method on Set5 and Set14 are 0.20 dB and 0.18 dB higher than the traditional method A+, respectively. And the running time of the proposed method is close to that of A+.

**Keywords** Super resolution, Dictionary learning, Collaborative representation, Sparse representation

## 1 引言

为了解决图像在获取和传输过程中受损降质的问题,研究者们提出了图像 SR 重建的概念。单幅图像 SR 重建是利用一幅低分辨率(Low Resolution, LR)图像来估计相应的高分辨率(High Resolution, HR)图像的过程。SR 算法可大致

分为基于插值的算法、基于重建的算法、基于学习的算法 3 类。

常用的插值运算包括最邻近插值、双三次插值(Bicubic)等。插值法没有引入额外的高频信息,因此重建后的图像锯齿痕迹和模糊现象比较严重。

基于重建的方法以正则化的形式来表达 HR 图像的先验知识,把重建问题转化为带约束的最优化问题,例如,文献[1]

到稿日期:2019-06-26 返修日期:2019-11-12 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金项目(61761005,61761007);广西自然科学基金项目(2016GXNSFAA380154)

The work was supported by the National Natural Science Foundation of China (61761005,61761007) and Natural Science Foundation of Guangxi Zhuang Autonomous Region(2016GXNSFAA380154).

通信作者:常侃(pandack0619@163.com)

提出了自适应稀疏域选择与自适应正则化算法,文献[2]提出了非局部集中稀疏表示算法,文献[3]提出了基于非局部全变分模型的算法,文献[4]提出了数据自适应低阶建模与外梯度先验的算法。基于重建的方法减少了重建图像边缘和纹理的锯齿现象,但存在复杂度高、收敛速度慢、图像重建时间长等缺点。

基于学习的方法通过样本学习得到映射模型,不仅在图像的重建速度上明显优于基于重建的方法,而且能较好地引入高频信息。例如,文献[5-10]分别利用样例学习、稀疏表示、协作表达进行 SR 重建。上述算法只通过单一字典对进行重建,导致重建细节不够丰富,重建结果具有不稳定性。

因此,有学者在字典训练阶段引入残差、聚类字典学习方法来解决单一字典学习的不足<sup>[11-14]</sup>。其中较有代表性的是文献[11]提出的算法,其虽然在残差信息的利用上有所改善,但重建的核心思想仍是稀疏编码,因此图像重建的速度相对较慢。此外,一些学者还通过自相似性<sup>[15-16]</sup>、卷积神经网络<sup>[17-21]</sup>等方法处理 SR 问题,取得了不错的重建效果。基于卷积神经网络的方法的重建效果虽好,但存在训练时间长、对计算机硬件要求高等缺点。

综合上述问题,本文提出一种新的基于多层残差字典学习的 SR 算法。该算法在训练阶段使用协作表达的方法重建训练样本,训练多层残差字典,构建相应的投影矩阵;在测试阶段通过计算好的投影矩阵恢复出残差信息,实现更有效的 SR 重建,从而获得更好的主客观效果。

## 2 相关背景知识

### 2.1 基于字典学习的单图像超分辨率算法

基于字典学习的单图像超分辨率算法(Single Image Super Resolution, SISR)<sup>[7]</sup>分为训练阶段和测试阶段两部分。其在训练阶段通过 K-SVD(K-Singular Value Decomposition)算法进行字典训练,构建 LR 字典  $\mathbf{D}_l$ 。其中,LR 字典  $\mathbf{D}_l$  应满足式(1):

$$\begin{aligned} \mathbf{D}_l = \arg \min_{\mathbf{D}_l} \sum_k \|\mathbf{p}_l^k - \mathbf{D}_l \cdot \mathbf{q}^k\|_2^2 \\ \text{s. t. } \|\mathbf{q}^k\|_0 \leq L, \forall k \end{aligned} \quad (1)$$

其中,  $\mathbf{q}^k$  为训练样本中 LR 图像块的稀疏系数,  $\mathbf{p}_l^k$  为按列排成向量的第  $k$  个 LR 图像块,  $L$  为充分小的常数。

SISR<sup>[7]</sup>算法认为 HR 图像块与 LR 图像块之间具有映射关系,即二者在相应的冗余字典中具有相同的稀疏表示系数,因此 HR 字典  $\mathbf{D}_h$  可由式(2)计算得到:

$$\mathbf{D}_h = \mathbf{C}_h \mathbf{Q}^+ = \mathbf{C}_h \mathbf{Q}^T (\mathbf{Q} \mathbf{Q}^T)^{-1} \quad (2)$$

其中,矩阵  $\mathbf{C}_h$  的每列为一个 HR 图像块,矩阵  $\mathbf{Q}$  的每列与 HR 图像块对应的稀疏系数  $\mathbf{q}^k$ 。

在测试阶段,利用 OMP(Orthogonal Matching Pursuit)算法计算出测试图像的 LR 图像块对应的稀疏系数  $\boldsymbol{\alpha}^*$ ,计算式为:

$$\boldsymbol{\alpha}^* = \arg \min_{\boldsymbol{\alpha}} \|\mathbf{F} \mathbf{D}_l \boldsymbol{\alpha} - \mathbf{F} \mathbf{y}_l\|_2^2 + \gamma \|\boldsymbol{\alpha}\|_1 \quad (3)$$

其中,  $\mathbf{y}_l$  表示 LR 测试图像块,  $\mathbf{F}$  为特征提取算子,  $\gamma$  为权衡系数。

最后,利用式(4)计算得到 HR 图像块:

$$\mathbf{x}_h = \mathbf{D}_h \boldsymbol{\alpha}^* \quad (4)$$

### 2.2 基于双字典及稀疏表示的超分辨率算法

基于双字典的稀疏表示超分辨率算法(Dual Dictionary and Sparse Representation, DDSR)<sup>[11]</sup> 将待重建的图像的高频信息(High Frequency, HF)分为主要高频信息(Main High Frequency, MHF)与残差信息(Residual Frequency, RF)两部分。针对 MHF 训练 HR 主字典  $\mathbf{D}_{MH}$  和 LR 主字典  $\mathbf{D}_{ML}$ , 针对 RF 训练 HR 残差字典  $\mathbf{D}_{RH}$  和 LR 残差字典  $\mathbf{D}_{RL}$ 。

在训练阶段,根据式(1)和式(2)训练主字典  $\mathbf{D}_{MH}$  和  $\mathbf{D}_{ML}$ , 并通过式(3)与式(4)重建出主要的高频  $\mathbf{I}_{MHF}$ , 将其与低频(Low Frequency, LF)图像  $\mathbf{I}_{LF}$  相加后得到重建的训练图像  $\mathbf{I}_{TMP}$ , 然后由式(5)计算残差图像  $\mathbf{I}_{RF}$ :

$$\mathbf{I}_{RF} = \mathbf{I}_{HF} - \mathbf{I}_{MHF} \quad (5)$$

将  $\mathbf{I}_{TMP}$  当作新的训练样本,用残差图像  $\mathbf{I}_{RF}$  作为新训练样本的 LF, 由式(1)、式(2)训练残差字典对  $\mathbf{D}_{RL}$  和  $\mathbf{D}_{RH}$ 。

在测试阶段,首先通过主字典  $\mathbf{D}_{MH}$  重建出测试图像的主要 HF 图像  $\mathbf{I}_{TMP}$ , 其次利用  $\mathbf{I}_{TMP}$  和残差字典  $\mathbf{D}_{RL}$ ,  $\mathbf{D}_{RH}$  重建出残差信息  $\mathbf{I}_{RF}$ , 最后由式(6)得到最终估计的图像  $\mathbf{I}_{EST}$ :

$$\mathbf{I}_{EST} = \mathbf{I}_{TMP} + \mathbf{I}_{RF} \quad (6)$$

### 2.3 锚点邻域回归算法

锚点邻域回归算法(Anchored Neighborhood Regression, ANR)<sup>[8-9]</sup>用  $l_2$  范数取代  $l_1$  范数,将求解稀疏系数的问题转换为协作表达问题:

$$\alpha = \arg \min_{\alpha} \|\mathbf{N}_l^j \alpha - \mathbf{p}_l\|_2^2 + \gamma \|\alpha\|_2^2 \quad (7)$$

上述问题可通过式(8)直接求解:

$$\alpha = (\mathbf{N}_l^{jT} \mathbf{N}_l^j + \gamma \mathbf{E})^{-1} \mathbf{N}_l^{jT} \mathbf{p}_l \quad (8)$$

其中,  $\mathbf{E}$  表示单位阵,  $\mathbf{N}_l^j$  为 LR 字典  $\mathbf{D}_l$  中第  $j$  个原子对应的邻域,  $\mathbf{D}_l$  为 LR 训练图像块,则重建的 HR 图像块可表示为:

$$\mathbf{x}_h^j = \mathbf{N}_h^j \alpha = \mathbf{N}_h^j (\mathbf{N}_l^{jT} \mathbf{N}_l^j + \gamma \mathbf{E})^{-1} \mathbf{N}_l^{jT} \mathbf{p}_l = \mathbf{P}_j \mathbf{p}_l \quad (9)$$

其中,  $\mathbf{N}_h^j$  是与 HR 字典  $\mathbf{D}_h$  对应的 HR 邻域,  $\mathbf{P}_j$  表示与第  $j$  个 HR 字典原子对应的投影矩阵。

上述方法是 ANR 算法的核心思想,用  $l_2$  范数取代  $l_1$  范数来约束目标函数  $\alpha$ , 将 HR 图像重建的计算方式从式(4)转换为式(9)。在测试阶段无需像 SISR<sup>[7]</sup>算法一样,利用 OMP 算法,通过式(3)计算 LR 图像块的稀疏系数,而是直接通过投影矩阵,利用式(9)计算 HR 图像,避免了迭代计算,缩短了计算时间,因此重建图像的耗时会大大降低。调整的锚定邻域回归算法(Adjusted Anchored Neighborhood Regression, A+)与 ANR 算法类似,只是将上述计算原子的最邻近域的范围扩大到整个训练集。

## 3 基于残差字典及协作表达的 SR 算法

DDSR 算法使用了双字典进行重建,但是其存在 3 个主要问题:1)重建时需要求解稀疏系数,因此重建速度较慢;2)DDSR 算法仅通过训练 HR 字典引入高频信息,因此能够提取的残差信息有限;3)DDSR 算法仅设计了一层残差结构,算法性能还有提升空间。另一方面,ANR 及 A+算法只使用单一字典,没有考虑残差信息的利用,导致重建图像的细节不够丰富。

针对这些问题,本文提出了一种基于残差字典及协作表

达的SR算法。一方面,该算法通过引入协作表达的思想,在训练阶段计算并存储投影矩阵,在测试阶段读取对应的投影矩阵,由相应的投影矩阵与LR图像块相乘来获得重建图像,因此不必通过OMP算法迭代计算稀疏系数 $\alpha^*$ ,减小了计算的复杂度,从而缩短了图像重建的时间。另一方面,为了充分利用残差信息,该算法训练多层残差字典。

### 3.1 训练阶段

权衡算法复杂度与重建效果后,本文算法选择训练三层残差字典(关于残差字典层数的实验见4.1节)。但为了方便叙述,本节以两层残差字典算法为例,三层残差字典算法可根据本节所述流程简单推导获得。训练流程如图1所示,具体分为3步。

#### (1)主字典与主投影矩阵的训练

HR训练图像 $I_{ORG}$ 经下采样后得到LR训练图像 $I_{DS}$ ,通过Bicubic插值放大到原始大小,得到LF图像 $I_{LF}$ , $I_{ORG}$ 减去 $I_{LF}$ 得到高频图像 $I_{HF}$ 。

通过滤波器提取 $I_{LF}$ 的特征后,将其特征图像进行分块,得到LR训练图像块集 $\{p_{ML}\}$ , $I_{HF}$ 直接分块得到HR训练图像块 $\{p_{MH}\}$ ,根据2.1节所述算法训练主字典对 $D_{MH}$ 及 $D_{ML}$ 。对于 $D_{ML}$ 的每个原子,在训练集合 $I_{LF}$ 中寻找 $K$ 个最邻近样本,将最邻近样本集组合成LR邻域矩阵集合 $\{S_{ML}^j\}_{j=1}^n$ ,其中 $n$ 为字典的原子数。同理,对 $D_{MH}$ 的每个原子在 $I_{HF}$ 中寻找 $K$ 个最邻近样本,形成HR邻域矩阵集合 $\{S_{MH}^j\}_{j=1}^n$ 。根据 $D_{MH}$ 的第 $j$ 个字典原子 $d_M^j$ ,由式(10)计算得到对应的主字典投影矩阵 $P_M^j$ :

$$P_M^j = S_{MH}^j (S_{ML}^j S_{ML}^j + \gamma E)^{-1} S_{ML}^j (10)$$

根据 $I_{LF}$ 中的每一个图像块 $p_{ML}$ 和它的最邻近原子 $d_M^j$ 及其对应的主字典投影矩阵 $P_M^j$ ,根据式(11)计算出MHF图像块 $x_{MHF}^j$ 。

$$x_{MHF}^j = P_M^j p_{ML} (11)$$

#### (2)第一层残差字典与第一层残差邻域的训练

通过合并所有计算出的 $x_{MHF}^j$ 及在重叠块之间取平均,求出主高频图像 $I_{MHF}$ 。由式(5)计算第一层残差图像 $I_{RF1}$ ,主字典重建图像 $I_{TMP1}$ 由 $I_{MHF}$ 与 $I_{LF}$ 相加获得。

将 $I_{TMP1}$ 当作新的训练图像的低频信号, $I_{RF1}$ 当作新的训练图像的高频信号,将 $I_{TMP1}$ 与 $I_{RF1}$ 采用与 $I_{LF}$ 和 $I_{HF}$ 相同的处理方法,得到第一层残差投影矩阵的训练集 $\{p_{RL1}, p_{RH1}\}$ ,针对LF残差字典 $D_{RL1}$ 及HF残差字典 $D_{RH1}$ 中的原子,分别在 $\{p_{RL1}\}$ 和 $\{p_{RH1}\}$ 中找到最近邻域 $\{S_{RL1}^j\}_{j=1}^n$ 与 $\{S_{RH1}^j\}_{j=1}^n$ 。

根据第一层残差字典的第 $j$ 个字典原子 $d_{R1}^j$ ,可由式(12)得到对应的第一层残差字典的投影矩阵 $P_{R1}^j$ :

$$P_{R1}^j = S_{RH1}^j (S_{RL1}^j S_{RL1}^j + \gamma E)^{-1} S_{RL1}^j (12)$$

#### (3)第二层残差字典与第二层残差投影矩阵的计算

将第一层残差投影矩阵集 $\{P_{R1}^j\}_{j=1}^n$ 与 $p_{RL1}$ 代入式(10)后计算出第一层残差高频图像 $I_{RHF1}$ ,将其与 $I_{TMP1}$ 相加得到第一层残差字典重建图像 $I_{TMP2}$ 。

重复之前的步骤,根据 $I_{TMP2}$ 与 $I_{RF2}$ 训练出第二层残差字典 $D_{RL2}, D_{RH2}$ 和第二层残差投影矩阵集 $\{P_{R2}^j\}_{j=1}^n$ ,存储各层残差字典以及对应的投影矩阵,以备测试阶段使用。

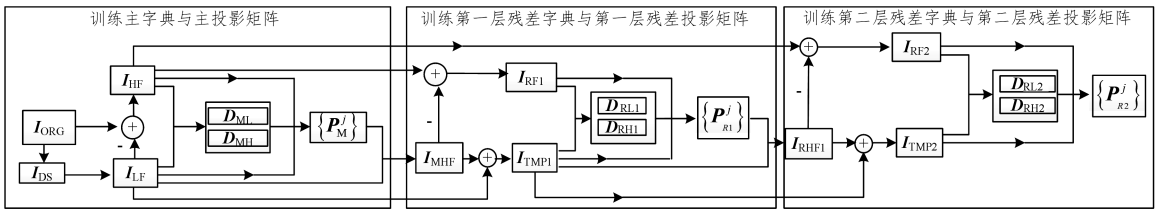


图1 训练阶段的流程图

Fig. 1 Flowchart of training phase

### 3.2 测试阶段

测试阶段的计算过程如图2所示,分为3步。其通过读

取之前训练好的字典与对应的投影矩阵对测试图像进行重建。

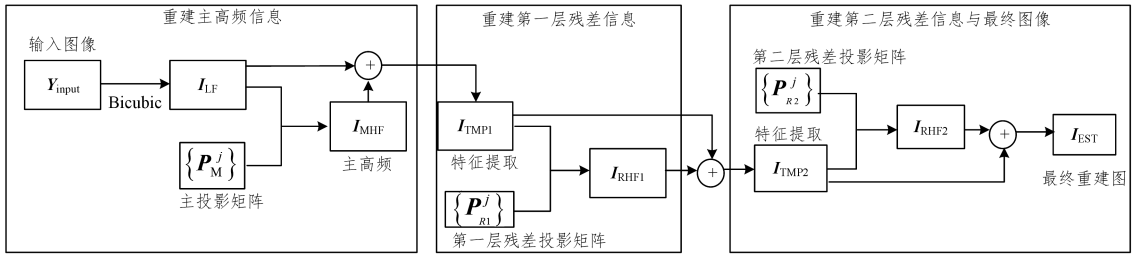


图2 测试阶段的流程图

Fig. 2 Flowchart of testing phase

(1)重建主要高频信息。将测试图像 $Y_{input}$ 通过Bicubic放大到原始大小,经特征提取算子提取低频特征 $I_{LF}$ 后,利用主投影矩阵集合 $\{P_M^j\}_{j=1}^n$ 以及 $I_{LF}$ 的每一个图像块,直接由式(11)生成 $I_{MHF}$ 。主字典重建图像 $I_{TMP1}$ 由式(13)获得:

$$I_{TMP1} = I_{LF} + I_{MHF} (13)$$

(2)重建第一层残差高频。将 $I_{TMP1}$ 当作新的输入图像,经特征提取滤波器及分块处理后,将每个图像块与对应的第一层残差投影矩阵 $P_{R1}^j$ 一起代入式(10),计算出第一层残差高频特征 $I_{RHF1}$ ,并由式(14)得到第一层残差字典重建图像

$$I_{TMP2} :$$

$$\mathbf{I}_{\text{TMP2}} = \mathbf{I}_{\text{TMP1}} + \mathbf{I}_{\text{RHF1}} \quad (14)$$

(3)重建第二层残差高频。将  $\mathbf{I}_{\text{TMP2}}$  看作新的输入图像,重复以上步骤,根据  $\mathbf{I}_{\text{TMP2}}$  与  $\mathbf{P}_{\text{R2}}$  重建出第二层残差高频图像  $\mathbf{I}_{\text{RHF2}}$ ,将  $\mathbf{I}_{\text{TMP2}}$  和  $\mathbf{I}_{\text{RHF2}}$  回代式(14)得到最终的 HR 图像  $\mathbf{I}_{\text{EST}}$ 。为了方便阅读,本文算法的流程如算法 1 所示。

#### 算法 1 RDCR

训练阶段:

- 1.由式(1)、式(2)、式(10)训练出主字典对  $\mathbf{D}_{\text{MH}}, \mathbf{D}_{\text{ML}}$  以及主字典投影矩阵集  $\{\mathbf{P}_{\text{M}}^j\}_{j=1}^n$ ;
- 2.根据式(11),利用  $\{\mathbf{P}_{\text{M}}^j\}_{j=1}^n$  重建出训练图像  $\mathbf{I}_{\text{TMP1}}$ ,将  $\mathbf{I}_{\text{TMP1}}$  作为新的训练集,重复步骤 1,得到第一层残差字典对  $\mathbf{D}_{\text{RL1}}, \mathbf{D}_{\text{RH1}}$ ,根据式(12)计算第一层残差投影矩阵集  $\{\mathbf{P}_{\text{R1}}^j\}_{j=1}^n$ ;
- 3.重复步骤 2,训练出第二层残差字典对  $\mathbf{D}_{\text{RL2}}, \mathbf{D}_{\text{RH2}}$  及计算第二层残差投影矩阵  $\{\mathbf{P}_{\text{R2}}^j\}_{j=1}^n$ 。

测试阶段:

- 1.由测试图像  $\mathbf{Y}_{\text{input}}$ 、主投影矩阵  $\{\mathbf{P}_{\text{M}}^j\}_{j=1}^n$ ,根据式(11)重建出主高频信息  $\mathbf{I}_{\text{MHF}}$ ,根据式(13)主字典重建图像  $\mathbf{I}_{\text{TMP1}}$ ;
- 2.将  $\mathbf{I}_{\text{TMP1}}$  作为新的输入图像,使用第一层残差投影矩阵集  $\{\mathbf{P}_{\text{R1}}^j\}_{j=1}^n$ ,由式(11)重建出一层残差信息  $\mathbf{I}_{\text{RHF1}}$ ,并由式(14)重建出第一层残差字典重建图  $\mathbf{I}_{\text{TMP2}}$ ;
- 3.将  $\mathbf{I}_{\text{TMP2}}$  作为新的输入图像,将第二层残差投影矩阵  $\{\mathbf{P}_{\text{R2}}^j\}_{j=1}^n$ ,代入式(11)重建出二层残差信息  $\mathbf{I}_{\text{RHF2}}$ ,将  $\mathbf{I}_{\text{RHF2}}$  与  $\mathbf{I}_{\text{TMP2}}$  代入式(14)重建出最终的 HR 图像  $\mathbf{I}_{\text{EST}}$ 。

## 4 实验结果与分析

为了验证提出的 SR 算法的有效性,实验分析将分为算法模块性能分析和算法性能比较分析两部分。本文采用 Matlab 2016A 仿真软件进行实验,CPU 为 I7-6700,内存 32 GB。

除了第 3 节提出的算法外,本文算法改进了特征提取滤波器,本实验使用 Roberts 算子与  $[0 \ 1 \ -2 \ 2 \ -1 \ 0], [-1 \ 0 \ 2 \ 0 \ -1]$  两组梯度滤波器代替 SCSR<sup>[5]</sup> 等算法中使用的一阶、二阶梯度滤波器进行特征提取。实验均采用经  $90^\circ, 180^\circ, 270^\circ$  旋转后扩展的文献[5]中的图像训练集。其他参数设置:字典大小  $n$  为 1024,邻域大小  $K=2 \ 048$ ,图像块大小为  $9 \times 9$ ,正则化参数  $\gamma$  为 0.01, A+ 算法<sup>[8]</sup> 与本文算法均提取 10 000 000 个训练向量。

为了较好地评估本文算法的性能,在测试集 Set5 和 Set14 上进行实验,将其与 SCSR<sup>[6]</sup>, SISR<sup>[7]</sup>, DDSR<sup>[11]</sup>, ANR<sup>[8]</sup>, A+<sup>[9]</sup>, SRCNN<sup>[17]</sup> 这 6 种较经典的 SR 算法进行比较。对测试图像采用 Bicubic 下采样,以生成 LR 图像。客观质量评价标准采用结构相似性 (Structural Similarity Index, SSIM) 和峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR)。

### 4.1 算法性能分析

本节在测试集 Set5 上测试不同层数的残差字典和残差投影矩阵,以及新特征提取滤波器对图像重建效果的贡献。表 1 列出了本文算法的几种变型算法的重建结果的客观质量。其中, L1RD, L2RD, L3RD, L4RD 分别表示使用了一层、两层、三层、四层残差字典及残差投影矩阵进行重建的算法。其中 L1RD, L2RD, L3RD, L4RD 均采用与 SCSR 相同的特征

提取滤波器,而 RDCR 即算法 1,采用三层残差字典,并采用了实验设置部分所述的新特征提取滤波器。

表 1 3 倍上采样下本文算法的 5 种变型算法的 PSNR 与 SSIM

Table 1 PSNR and SSIM results obtained by 5 versions of proposed algorithm ( $\times 3$ )

图像	指标	L1RD	L2RD	L3RD	L4RD	RDCR
Baby	PSNR/dB	35.21	35.22	35.22	35.23	35.24
	SSIM	0.9224	0.9228	0.9231	0.9233	0.9241
Bird	PSNR/dB	35.56	35.65	35.70	35.73	35.96
	SSIM	0.9566	0.9571	0.9581	0.9585	0.9629
Butterfly	PSNR/dB	27.49	27.54	27.55	27.55	27.60
	SSIM	0.9141	0.9160	0.9172	0.9175	0.9202
Head	PSNR/dB	33.81	33.82	33.83	33.83	33.84
	SSIM	0.8277	0.8278	0.8285	0.8287	0.8303
Woman	PSNR/dB	31.23	31.29	31.29	31.29	31.32
	SSIM	0.9306	0.9310	0.9317	0.9319	0.9331
平均	PSNR/dB	32.67	32.70	32.72	32.72	32.79
	SSIM	0.9102	0.9109	0.9117	0.9120	0.9141

由表 1 可知,多层残差字典算法的重建效果优于单层残差字典算法,这是因为使用一层残差字典重建图像后,残差图像中依然有可以利用的高频信息。对比 L3RD, L4RD 可知,随着残差字典层数的增加,重建的提升程度减小, PSNR 维持不变, SSIM 仅提升 0.0002。这是因为图像可供挖掘的剩余高频信息越来越少。由表 2 可知,每增加一层残差字典,在训练与测试时间上均要多花费大约 1/3 的时间。因此,权衡复杂度与重建效果后,本文最终使用三层残差字典。在改进特征提取算子后,相比未改进特征提取的 L3RD 算法, RDCR 算法的重建效果提升明显,这是因为低频信息提取效果越好,稀疏字典与投影矩阵的表示就越准确,重建图像的高频信息恢复得就越好。

表 2 本文算法的 5 种变型算法的运行时间

Table 2 Running time of 5 versions of proposed algorithm

时间	L1RD	L2RD	L3RD	L4RD	RDCR
训练时间/h	3.1	6.3	9.4	12.7	9.4
测试时间/s	0.68	0.88	1.34	1.75	1.33

### 4.2 与主流算法的性能比较

各类算法的对比结果如表 3—表 6 所列,本文算法的重建结果相比文中提到的几种经典算法有较大的提升,例如,在表 3 中,相比 DDSR<sup>[11]</sup>, A+<sup>[9]</sup> 和 SRCNN<sup>[17]</sup>, 本文算法的 PSNR 增益分别为 0.56 dB, 0.18 dB, 0.27 dB。当放大倍数为 4 时,本文算法的 PSNR 增益与 DDSR<sup>[11]</sup>, A+<sup>[9]</sup>, SRCNN<sup>[17]</sup> 相比分别为 0.40 dB, 0.20 dB, 0.41 dB。与放大倍数为 3 时相比, DDSR<sup>[11]</sup> 的 PSNR 提升量稍微降低, A+<sup>[9]</sup> 和 SRCNN<sup>[17]</sup> 的 PSNR 提升量稍微提升。这是因为 DDSR<sup>[11]</sup> 与本文算法均引入了残差字典的思想,说明在重建放大倍数较大的图像时,残差字典能更好地利用图像的剩余高频信息,提高重建效果。在测试集 Set14 上,当上采样倍数为 3 时,本文算法的 PSNR 增益与 DDSR<sup>[11]</sup>, A+<sup>[9]</sup>, SRCNN<sup>[17]</sup> 相比分别为 0.35 dB, 0.17 dB, 0.18 dB。与表 3 对比,提升量下降。这是因为测试集 Set14 中包含了许多高复杂度的图像,重建时图像的高频信息更难恢复,因此提升量相对于测试集 Set5 稍降。

表 3 不同算法在 Set5 上 3 倍上采样的 PSNR 与 SSIM

Table 3 PSNR and SSIM results obtained by different methods on Set5 ( $\times 3$ )

图像	指标	SCSR	SISR	ANR	DDSR	A+	SRCNN	RDCR
Baby	PSNR/dB	34.94	35.09	35.13	35.19	35.20	35.02	<b>35.24</b>
	SSIM	0.9043	0.9215	0.9221	0.9230	0.9225	0.9176	<b>0.9241</b>
Bird	PSNR/dB	34.29	34.61	34.64	35.08	35.60	35.05	<b>35.96</b>
	SSIM	0.9391	0.9465	0.9498	0.9530	0.9561	0.9518	<b>0.9629</b>
Butterfly	PSNR/dB	25.67	25.93	25.94	26.54	27.29	<b>27.92</b>	27.60
	SSIM	0.8611	0.8713	0.8796	0.8886	0.9092	<b>0.9245</b>	0.9202
Head	PSNR/dB	33.26	33.68	33.64	33.72	33.78	33.56	<b>33.84</b>
	SSIM	0.8024	0.8187	0.8236	0.8262	0.8273	0.8192	<b>0.8303</b>
Woman	PSNR/dB	30.05	30.26	30.35	30.65	31.18	31.05	<b>31.31</b>
	SSIM	0.9037	0.9160	0.9185	0.9215	0.9288	0.9224	<b>0.9331</b>
平均	PSNR/dB	31.65	31.92	31.94	32.23	32.61	32.52	<b>32.79</b>
	SSIM	0.8821	0.8948	0.8978	0.9021	0.9089	0.9065	<b>0.9141</b>

表 4 不同算法在 Set14 上 3 倍上采样的 PSNR 与 SSIM

Table 4 PSNR and SSIM results obtained by different methods on Set14 ( $\times 3$ )

图像	指标	SCSR	SISR	ANR	DDSR	A+	SRCNN	RDCR
Baboon	PSNR/dB	23.47	23.52	23.56	23.61	23.63	23.62	<b>23.65</b>
	SSIM	0.5878	0.5895	0.5990	0.5994	0.6064	0.6069	<b>0.6092</b>
Barbara	PSNR/dB	26.41	26.67	<b>26.80</b>	26.75	26.51	26.62	26.64
	SSIM	0.7633	0.7818	<b>0.7836</b>	0.7816	0.7795	0.7803	0.7812
Bridge	PSNR/dB	24.82	25.02	25.01	25.12	25.17	25.15	<b>25.26</b>
	SSIM	0.6920	0.7014	0.7012	0.7065	0.7091	0.7082	<b>0.7139</b>
Coastguard	PSNR/dB	27.01	27.10	27.07	27.13	27.27	<b>27.32</b>	27.31
	SSIM	0.6393	0.6539	0.6577	0.6609	0.6625	<b>0.6645</b>	0.6629
Comic	PSNR/dB	23.91	23.97	24.04	24.20	24.38	24.45	<b>24.50</b>
	SSIM	0.7557	0.7578	0.7606	0.7706	0.7794	0.7853	<b>0.7899</b>
Face	PSNR/dB	33.46	33.55	33.62	33.70	33.76	33.56	<b>33.83</b>
	SSIM	0.8041	0.8195	0.8232	0.8237	0.8274	0.8120	<b>0.8281</b>
Flowers	PSNR/dB	28.38	28.42	28.49	28.95	29.15	29.09	<b>29.29</b>
	SSIM	0.8297	0.8385	0.8403	0.8446	0.8527	0.8492	<b>0.8551</b>
Foreman	PSNR/dB	32.08	33.22	33.21	33.89	34.23	33.63	<b>34.34</b>
	SSIM	0.9129	0.9292	0.9301	0.9355	0.9420	0.9321	<b>0.9423</b>
Lenna	PSNR/dB	32.76	33.01	33.08	33.28	33.52	33.48	<b>33.73</b>
	SSIM	0.8648	0.8782	0.8804	0.8812	0.8856	0.8847	<b>0.8867</b>
Man	PSNR/dB	27.82	27.91	27.92	28.06	28.28	28.26	<b>28.42</b>
	SSIM	0.7829	0.7862	0.7890	0.7950	0.7996	0.7992	<b>0.8053</b>
Monarch	PSNR/dB	30.84	31.11	31.09	31.85	32.15	32.68	<b>32.72</b>
	SSIM	0.9290	0.9382	0.9377	0.9440	0.9458	0.9502	<b>0.9508</b>
Pepper	PSNR/dB	33.41	34.03	33.91	34.38	34.74	34.41	<b>34.89</b>
	SSIM	0.8778	0.8863	0.8851	0.8892	0.8930	0.8894	<b>0.8940</b>
PPT3	PSNR/dB	24.97	25.21	25.24	25.82	26.20	<b>26.73</b>	26.53
	SSIM	0.8948	0.9085	0.9117	0.9276	0.9346	<b>0.9374</b>	0.9366
Zebra	PSNR/dB	28.12	28.51	28.56	28.79	28.98	28.83	<b>29.32</b>
	SSIM	0.8359	0.8429	0.8424	0.8487	0.8498	0.8491	<b>0.8536</b>
平均	PSNR/dB	28.39	28.63	28.69	28.96	29.14	29.13	<b>29.31</b>
	SSIM	0.7978	0.8078	0.8096	0.8146	0.8204	0.8177	<b>0.8222</b>

表 5 不同算法在 Set5 上 4 倍上采样的 PSNR 与 SSIM

Table 5 PSNR and SSIM results obtained by different methods on Set5 ( $\times 4$ )

图像	指标	SCSR	SISR	ANR	DDSR	A+	SRCNN	RDCR
Baby	PSNR/dB	32.63	32.96	33.06	33.15	33.28	32.98	<b>33.48</b>
	SSIM	0.8756	0.8767	0.8824	0.8829	0.8843	0.8564	<b>0.8859</b>
Bird	PSNR/dB	31.45	31.70	31.77	32.41	32.53	31.97	<b>32.72</b>
	SSIM	0.8991	0.9005	0.9027	0.9052	0.9131	0.9045	<b>0.9155</b>
Butterfly	PSNR/dB	23.36	23.52	23.53	24.16	24.49	<b>25.06</b>	24.69
	SSIM	0.8085	0.8168	0.8175	0.8324	0.8425	<b>0.8658</b>	0.8447
Head	PSNR/dB	32.07	32.26	32.28	32.43	32.56	32.19	<b>32.72</b>
	SSIM	0.7594	0.7636	0.7675	0.7796	0.7828	0.7630	<b>0.7848</b>
Woman	PSNR/dB	27.54	27.83	27.85	28.24	28.64	28.20	<b>28.86</b>
	SSIM	0.8472	0.8518	0.8526	0.8695	0.8820	0.8687	<b>0.8840</b>
平均	PSNR/dB	29.41	29.67	29.70	30.09	30.29	30.08	<b>30.49</b>
	SSIM	0.8379	0.8418	0.8445	0.8541	0.8609	0.8520	<b>0.8630</b>

表 6 不同算法在 Set14 上 4 倍上采样的 PSNR 与 SSIM

Table 6 PSNR and SSIM results obtained by different methods on Set14 ( $\times 4$ )

图像	指标	SCSR	SISR	ANR	DDSR	A+	SRCNN	RDCR
Baboon	PSNR/dB	22.47	22.66	22.69	22.68	22.74	22.70	22.82
	SSIM	0.4931	0.4952	0.4960	0.4958	0.5033	0.5002	0.5075
Barbara	PSNR/dB	25.38	25.59	25.60	25.67	25.75	25.70	25.77
	SSIM	0.7174	0.7189	0.7192	0.7241	0.7259	0.7251	0.7266
Bridge	PSNR/dB	23.50	23.65	23.63	23.66	23.71	23.68	23.84
	SSIM	0.5982	0.5993	0.5989	0.5996	0.6005	0.5998	0.6014
Coastguard	PSNR/dB	25.58	25.74	25.79	25.82	25.99	25.93	26.12
	SSIM	0.5475	0.5483	0.5489	0.5507	0.5581	0.5562	0.5583
Comic	PSNR/dB	22.07	22.31	22.33	22.45	22.60	22.53	22.70
	SSIM	0.6366	0.6385	0.6394	0.6455	0.6597	0.6513	0.6633
Face	PSNR/dB	32.01	32.18	32.24	32.31	32.41	32.12	32.50
	SSIM	0.7710	0.7729	0.7742	0.7765	0.7808	0.7708	0.7822
Flowers	PSNR/dB	26.33	26.44	26.50	26.73	26.90	26.84	27.12
	SSIM	0.7581	0.7602	0.7610	0.7674	0.7746	0.7725	0.7785
Foreman	PSNR/dB	30.77	31.00	30.93	31.36	32.23	31.46	32.51
	SSIM	0.8845	0.8882	0.8879	0.8964	0.9089	0.9116	0.9126
Lenna	PSNR/dB	30.64	30.92	31.10	31.24	31.41	31.20	31.52
	SSIM	0.8344	0.8365	0.8373	0.8421	0.8454	0.8414	0.8462
Man	PSNR/dB	26.23	26.45	26.44	26.62	26.79	26.65	27.04
	SSIM	0.7118	0.7135	0.7131	0.7244	0.7276	0.7251	0.7319
Monarch	PSNR/dB	28.52	28.72	28.74	29.06	29.41	29.89	29.83
	SSIM	0.8933	0.8986	0.8993	0.9045	0.9122	0.9244	0.9140
Pepper	PSNR/dB	31.81	32.12	31.94	32.24	32.88	32.34	33.17
	SSIM	0.8530	0.8558	0.8539	0.8587	0.8631	0.8610	0.8644
PPT3	PSNR/dB	22.75	23.06	22.86	22.99	23.64	23.84	24.08
	SSIM	0.8281	0.8367	0.8365	0.8392	0.8764	0.8844	0.8860
Zebra	PSNR/dB	25.25	25.47	25.49	25.58	25.94	25.97	26.14
	SSIM	0.7318	0.7342	0.7345	0.7396	0.7495	0.7504	0.7517
平均	PSNR/dB	26.65	26.88	26.89	27.03	27.33	27.20	27.51
	SSIM	0.7327	0.7355	0.7357	0.7403	0.7494	0.7481	0.7518

表 7 列出了几种算法在 Set5 和 Set14 两个图像集上的平均重建时间。可知,由于采用协作表达取代了传统的稀疏表达,本文算法在提高重建效果同时,运算速度明显快于 DDSR<sup>[11]</sup>,SCSR<sup>[6]</sup>等算法,与 SISR<sup>[7]</sup>算法的速度接近。

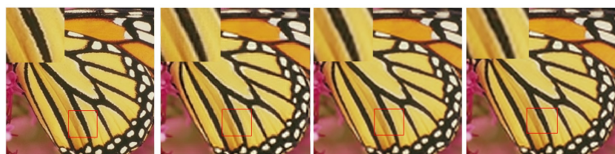
表 7 上采样倍数为 3 时不同算法的重建时间

Table 7 Reconstruction time of different methods ( $\times 3$ )

(单位:s)

图像库	SCSR	SISR	ANR	DDSR	A+	SRCNN	RDCR
Set5	10.48	0.75	0.36	3.94	0.33	1.75	1.33
Set14	73.45	1.6	0.55	10.23	0.51	3.42	1.96

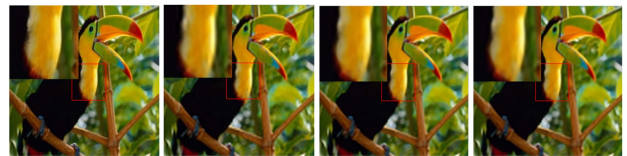
为了从主观视觉质量上评价上述算法和本文算法的优劣,图 3 和图 4 分别展示了各种算法在上采样倍数为 3 时重建的图像。



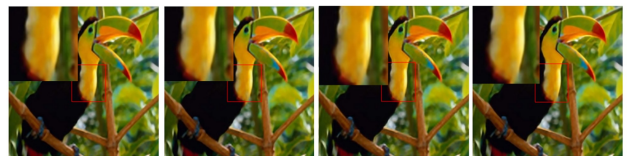
(a) Original Image (b) SCSR (25.67 dB, 0.8611) (c) SISR (25.93 dB, 0.8713) (d) ANR (25.94 dB, 0.8794)



(e) DDSR (26.54 dB, 0.8886) (f) A+ (27.29 dB, 0.9092) (g) SRCNN (27.92 dB, 0.9245) (h) RDCR (27.60 dB, 0.9202)

图 3 不同算法的 Butterfly 图像 SR 重建结果( $\times 3$ )Fig. 3 SR reconstruction results of Butterfly image by different methods( $\times 3$ )

(a) Original Image (b) SCSR (34.29 dB, 0.9391) (c) SISR (34.61 dB, 0.9465) (d) ANR (34.64 dB, 0.9498)



(e) DDSR (35.08 dB, 0.9530) (f) A+ (35.60 dB, 0.9561) (g) SRCNN (35.05 dB, 0.9518) (h) RDCR (35.96 dB, 0.9629)

图 4 不同算法的 Bird 图像 SR 重建结果( $\times 3$ )Fig. 4 SR reconstruction results of Bird image by different methods ( $\times 3$ )

从图中可看到,SCSR<sup>[6]</sup>的重建效果最差,边缘锯齿现象较明显,纹理不清晰。SISR<sup>[7]</sup>的重建效果稍好,不过图像边缘仍然模糊,且细节不够清晰。与 ANR<sup>[8]</sup>,SISR<sup>[7]</sup>相比,DDSR<sup>[11]</sup>重建的图像细节更完整,但图像纹理丰富的区域仍然存在一些伪影现象。相比 ANR<sup>[8]</sup>和 DDSR<sup>[11]</sup>,A+<sup>[9]</sup>和 SRCNN<sup>[17]</sup>很大程度地丰富了重建图像的细节,改善了重建图像锯齿和模糊的现象。而本文算法考虑了残差图像的剩余信息,重建的图像恢复了更多的高频信息,细节更完善,更好地抑制了锯齿现象和伪影现象,边缘更锐利,具有更好的视觉效果。

**结束语** 传统的基于稀疏表达的单图像 SR 方法获得了较大的成功,但存在着对高频信息的表达能力偏弱的问题。

为了解决上述问题,本文提出了一种基于残差字典及协作表达的单图像 SR 算法,通过训练主字典及多层残差字典,并在此基础上引入协作表达,来提升算法对高频信息的表达能力。本文算法与目前较为经典 SCSR<sup>[6]</sup>算法、SISR<sup>[7]</sup>算法、ANR<sup>[8]</sup>算法、DDSR<sup>[11]</sup>算法、A+<sup>[9]</sup>算法、SRCNN<sup>[17]</sup>算法相比,图像的重建结果无论在主观视觉上或是客观评价指标上,表现均为最佳。由于引入了协作表达,本文算法具备较快的运算速度,运算复杂度相比稀疏表示类算法更低。需要注意的是,包括本文算法在内的绝大多数基于学习的单图像 SR 算法并未考虑图像中噪声的影响。因此,下一步将重点研究如何提升本文算法在含噪情况下的鲁棒性。

### 参 考 文 献

- [1] DONG W S,ZHANG L,SHI G M,et al. Image Deblurring and super resolution by adaptive sparse domain selection and adaptive regularization [J]. IEEE Transactions on Image Processing, 2011,20(7):1838-1857.
- [2] DONG W S,ZHANG L,SHI G M,et al. Nonlocally centralized sparse representation for image restoration [J]. IEEE Transactions on Image Processing,2013,22(4):1620-1630.
- [3] REN C,HE X,PU Y,et al. Enhanced non-local total variation model and multi-directional feature prediction prior for single image super resolution [J]. IEEE Transactions on Image Processing,2019,28(8):3778-3793.
- [4] CHANG K,ZHANG X,DING P L K,et al. Data-adaptive low-rank modeling and external gradient prior for single image super-resolution [J]. Signal Processing,2019,161:36-49.
- [5] FREEMAN W T,JONES T R,PASZTOR E C. Example based super-resolution [J]. IEEE Computer Graphics and Applications,2002,22(2):56-65.
- [6] YANG J C,WRIGHT J,HUANG T,et al. Image super resolution via sparse representation [J]. IEEE Transactions on Image Processing,2010,19(11):2861-2873.
- [7] ZEYDE R,ELAD M,PROTTER M. On single image scale-up using sparse-representations[C]// International Conference on Curves and Surfaces. 2010:711-730.
- [8] TIMOFTE R,SMET V D,GOOL L V. Anchored neighborhood regression for fast example based super resolution[C]// International Conference on Computer Vision(ICCV). IEEE,2013:241-246.
- [9] TIMOFTE R,SMET V D,GOOL L V. A+: adjusted anchored neighborhood regression for fast super resolution[C]// Asian Conference on Computer Vision (ACCV). IEEE,2014:111-126.
- [10] TIMOFTE R,RASMUS R,GOOL L V. Seven ways to improve example based single image super resolution[C]// Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016:375-379.
- [11] ZHANG J,ZHAO C,XIONG R,et al. Image Super-Resolution via Dual-Dictionary Learning and Sparse Representation [C]//

IEEE International Symposium on Circuits and Systems (ISCAS). 2012:1688-1691.

- [12] PAN Z X,YU J,XIAO C B,et al. Single Image Super Resolution Based on Adaptive Multi-Dictionary Learning [J]. Acta Electronica Sinica,2015,43(2):209-216.
- [13] WANG R G,LIU L L,YANG J,et al. Image super-resolution based on clustering and collaborative representation [J]. Opto-Electronic Engineering,2018,45(4):9-18.
- [14] QIU K,YI B S,XIANG M,et al. Collaborative Sparse Dictionary Learning for Reconstruction of Single Image Super Resolution [J]. Acta Optica Sinica,2018,38(9):130-136.
- [15] QING X J,SHAN Y Y,XIAO J J,et al. Self-Learning Single Image Super-resolution Reconstruction Based on Compressive Sensing and SVR [J]. Computer Science,2017,44(11):169-174.
- [16] HUANG J B,SINGH A,AHUJA N. Single image super-resolution from transformed self-exemplars[C]// Conference on Computer Vision and Pattern Recognition (CVPR). IEEE,2015:5197-5206.
- [17] DONG C,LOY C C,HE K,et al. Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2016,38(2):295-307.
- [18] DONG C,LOY C C. Accelerating the super-resolution convolutional neural network[C]// European Conference on Computer Vision (ECCV). IEEE,2016:391-407.
- [19] BEE L,SANGHYUM S,HEEWON K,et al. Enhanced deep residual networks for single image super-resolution[C]// Computer Vision and Pattern Recognition Workshop (CVPRW). IEEE,2017:136-144.
- [20] LI J,FANG F,MEI K,et al. Multi-scale residual network for image super-resolution[C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018:517-532.
- [21] LIAN Y Y,WU X J. Research on Image Super-Resolution Reconstruction of Super Deep Convolutional Neural Network[J]. Computer Engineering,2019,45(1):217-220.



**TIAN Xu**, born in 1993, postgraduate. His main research interests include image super-resolution and image denoising.



**CHANG Kan**, born in 1983, Ph.D, associate professor, master's supervisor, is a member of China Computer Federation. His main research interests include image and video processing and compressive sensing, and video coding, etc.