

基于改进的非下采样 Contourlet 变换的 超分辨率复原算法

吴 炜 杨晓敏 陈 默 何小海*

(四川大学电子信息学院图像信息研究所, 四川 成都 610064)

摘要 Contourlet 变换应用于图像复原时容易引入伪吉布斯现象。非下采样 Contourlet 变换(NSCT)具有平移不变性,能够克服伪吉布斯现象,但是由于基于学习的超分辨率复原需要建立不同分辨率的关系,而 NSCT 变换的结果是每一层图像大小都一样,不能像拉普拉斯金字塔那样建立高低分辨率图像的对应关系及运算量较大。针对这些问题,提出了基于改进的非下采样 Contourlet 变换(INSCT)的超分辨率复原算法。为了表示人脸特征,算法首先建立了 INSCT 金字塔。然后针对人脸的特殊性,在匹配过程中,采用对应点进行匹配的方法。实验表明该算法具有较好的性能,复原出的超分辨率人脸图像无论在主观视觉效果上还是在客观评价指标上都取得较好的结果,复原的图像具有更好的视觉效果,更逼真,更接近于原始高分辨率图像。

关键词 图像处理; 基于学习的超分辨率; 幻觉脸; 非下采样 Contourlet 变换; INSCT 金字塔

中图分类号 TP391.4 **文献标识码** A **doi:** 10.3788/AOS20092906.1493

Super-Resolution Restoration Algorithms Based on Improved Nonsubsampled Contourlet Transform

Wu Wei Yang Xiaomin Chen Mo He Xiaohai

(*Institute of Image Information, College of Electronics and Information Engineering,*
Sichuan University, Chengdu Sichuan 610064, China)

Abstract Contourlet transform in image restoration is apt to bring pseudo-Gibbs phenomenon. Nonsubsampled Contourlet (NSCT) has better frequency selectivity and regularity compared with the contourlet transform, and it can overcome pseudo-Gibbs phenomenon. However, learning-based super-resolution need establish the relation of different resolutions. Different from Laplacian pyramid, result of NSCT at each layer is with the same size, but cannot establish multi-resolution pyramid, and it needs large amount of computation. According to the problem of NSCT in learning-based super-resolution, a face image super-resolution restoration algorithm based on INSCT is proposed. In order to represent face images features, the INSCT pyramid is established. For particularity of the face images, corresponding points in matching process are used. Experimental results show that this method has good performance, and achieves better results on subjective visual effects and on objective peak signal-to-noise ratio. And the results are closer to the real images.

Key words image processing; learning-based super-resolution; face hallucination; nonsubsampled contourlet transform; INSCT pyramid

1 引 言

超分辨率技术^[1~4]是提高图像分辨率的主要方法之一,具有广阔前景。其主要优点是可利用现有

的成像系统,在不改变原有系统硬件设备的前提下,提高图像的分辨率。超分辨率技术在医学、遥感、视频转换和安全监控等领域都有十分重要的应用。

收稿日期: 2008-05-06; **收到修改稿日期:** 2008-10-15

基金项目: 教育部视频超分辨率重建关键技术研究重点项目(107094)和四川大学电子信息学院青年教师基金资助课题。

作者简介: 吴 炜(1975-),男,讲师,博士,主要从事图像处理、图像识别和图像通信方面的研究。

E-mail: wuwei@scu.edu.cn

* 通信联系人。E-mail: cias@263.net

图像超分辨率复原的目的就是要对退化图像通过一定的技术处理,使图像获得更多的细节和信息,从而复原出没有退化前的理想图像。目前超分辨率技术主要分为两大类:1) 基于重建的超分辨率方法^[5~7];2) 基于学习的超分辨率方法^[8~15]。

基于重建的超分辨率技术通常需要输入多幅(多帧)图像。而基于学习的超分辨率技术通常只需要输入单幅图像,通过一个事先建立的图像库进行超分辨率复原。本文的研究重点集中于基于学习的超分辨率研究。国内外学者已经提出一些基于学习的超分辨率算法^[8~16]。

由于 Contourlet 变换中的结构具有非常有限的冗余性,在图像复原时奇异点周围容易引入伪吉布斯现象。而非下采样的 Contourlet 变换(NSCT)采用非抽样的金字塔结构和非抽样方向滤波器组,使得 NSCT 具有平移不变性,能够克服伪吉布斯(Pseudo-Gibbs)现象。在基于学习的超分辨率算法中需要建立高低分辨率的关系,但是 NSCT 变换结果每一层图像大小都一样,不能像拉普拉斯金字塔那样建立高低分辨率图像的对应关系以及 NSCT 运算量比较大。针对这些缺陷,本文提出了基于改进的非下采样 Concourlet (INSCT)变换的人脸图像超分辨率复原算法。为了表示人脸特征,算法首先建立了 INSCT 金字塔。然后针对人脸的特殊性,在匹配过程中,采用对应点进行匹配的方法。实验表明基于 INSCT 的超分辨率复原算法具有较好的性能,复原的图像无论是主观的视觉效果还是客观的评价指标上都取得较好的结果。

2 基于学习的超分辨率的理论基础

设 I_H 为高分辨率(HR)图像, I_L 为低分辨率(LR)图像,超分辨率复原需要解决的问题是在已知 I_L 的条件下,求出最优的 I_H 。最常用方法是最大后验概率法(MAP)。根据贝叶斯估计理论,后验概率为

$$P(I_H/I_L) = \frac{P(I_L/I_H)P(I_H)}{P(I_L)}, \quad (1)$$

其中 $P(I_H)$ 和 $P(I_L)$ 分别为 HR 和 LR 的先验概率; $P(I_L/I_H)$ 为给定 HR 图像时,观测 LR 图像的条件概率。为了求得最优的高分辨率图像 I_H , 必须找到使后验概率 $P(I_H/I_L)$ 取得最大值的 I_H 。因为 I_L 已知,所以 $P(I_L)$ 是常数。

$$I_H = \arg \max_{I_H} P(I_H/I_L) =$$

$$\arg \max_{I_H} P(I_L/I_H)P(I_H). \quad (2)$$

由于对数函数是单调递增函数,因此可以利用对数运算将(2)式中的乘法转换为加法,即

$$I_H = \arg \max_{I_H} \ln P(I_L/I_H) + \ln P(I_H). \quad (3)$$

可见预测 I_H 的三个步骤:(1) 建立先验模型,求取 $P(I_H)$;(2) 建立观测模型,求取 $P(I_L/I_H)$;(3) 将先验模型和观测模型集成到 MAP 框架中,求得最优的高分辨率图像 I_H 。其中第一步建立先验模型,求取先验模型 $P(I_H)$ 是最复杂、最重要也是最关键的一步。

1) 求取先验模型

假设低分辨率图像 I_L 通过对高分辨率训练样本图像的学习,复原出的高分辨率图像是 \hat{I}_H ;真实的高分辨率图像为 I_H 。 \hat{I}_H 和 I_H 的特征分别为 $K(\hat{I}_H)$ 和 $K(I_H)$ (K 为特征提取算子,通常提取的是高频信息)。在理想情况下, \hat{I}_H 和 I_H 应该是相等的,这样 $K(\hat{I}_H)$ 和 $K(I_H)$ 也应该相等的。但是实际情况中, $K(\hat{I}_H)$ 和 $K(I_H)$ 之间存在一定的误差,假设 $K(\hat{I}_H)$ 和 $K(I_H)$ 之间的误差满足高斯分布,并且方差为 σ_p 。可以建立如下求取先验模型 $P(I_H)$,该式的概率分布是方差为 σ_p 的高斯分布:

$$P(I_H) = P(\eta_p) |_{\eta_p = K(\hat{I}_H) - K(I_H)} = \exp\left(-\frac{\eta_p^2}{\sigma_p}\right). \quad (4)$$

对先验模型 $P(I_H)$ 的求取, \hat{I}_H 起着关键性的作用。从这点意义上来说,在基于学习的超分辨率中先验模型 $P(I_H)$ 充分考虑了图像的特殊性。

2) 求取观测模型

观测模型就是求取条件概率 $P(I_L/I_H)$ 的值,与先验模型相比,求取观测模型相对来说较为简单。待求高分辨率人脸图像 I_H ,所对应的低分辨率人脸图像 $D(I_H)$ 可以用平滑和下采样过程得到,输入的低分辨率人脸图像 I_L 应该与 $D(I_H)$ 是相同的。但是通常来说, I_L 和 $D(I_H)$ 之间存在一定的误差,假设 I_L 和 $D(I_H)$ 之间的误差满足高斯分布,并且方差为 σ_m 。可以建立方差为 σ_m 的高斯分布的概率分布公式

$$P(I_L/I_H) = P(\eta_m) |_{\eta_m = D(I_H) - I_L} = \exp\left(-\frac{\eta_m^2}{\sigma_m}\right). \quad (5)$$

这样把观测模型 $P(I_L/I_H)$ 转化为高斯分布函数。将(4)式,(5)式代入(3)式,并且进行简化得到

$$I_H = \arg \min_{I_H} (\|D(I_H) - I_L\|^2 + \lambda \|K(I_H) - K(\hat{I}_H)\|^2), \quad (6)$$

式中 D 是降采样算子, K 表示高频特征提取算子。

在(6)式中, $\|D(I_H) - I_L\|$ 主要是保证复原的高分辨率图像降质后与低分辨率图像相似; $\|K(I_H) - K(\hat{I}_H)\|$ 主要是保证 I_H 的高频信息尽可能的与 \hat{I}_H 的高频信息相似。也就是 I_H 应尽可能地与 \hat{I}_H 相似, 同时要使得 I_H 的降质图像应与 I_L 相似。先验(信息)知识的获取, 全体现在 \hat{I}_H 中。 \hat{I}_H 的预测对求解 I_H 是关键因素, 复原图像质量的高低, 对 \hat{I}_H 的依赖程度很高, 如何准确的预测 \hat{I}_H 是基于学习的超分辨率研究的难点和重点。

λ 是系数, 用于调节(6)式第一项和第二项权重的参数。如果 $\lambda \rightarrow 0$, (6)式退化为

$$I_H = \arg \min_{I_H} (\|D(I_H) - I_L\|^2). \quad (7)$$

(7)式是基于重建的超分辨率算法中的最大似然(ML)算法, 它是一个病态问题, 求解它需要加入正则化条件。它完全没有考虑 \hat{I}_H , 也就是忽略了图像(库)的先验知识, 退化为基于重建的超分辨率算法。

如果 $\lambda \rightarrow \infty$, (6)式退化为

$$I_H = \arg \min_{I_H} (\|K(I_H) - K(\hat{I}_H)\|^2). \quad (8)$$

这样求解(8)式, 得到 $I_H = \hat{I}_H$ 。也就是说在 $\lambda \rightarrow \infty$ 时, 通过训练样本学习复原出的高分辨率图像是 \hat{I}_H 就是最终的复原图像 I_H 。

当 $\lambda \neq 0$ 以及 $\lambda \neq \infty$ 时对(6)式求解, 通常采用最速下降法求解^[8]。由于误差函数是一个二次型, 因此算法能够收敛于一个全局最小值, 最后求解得到 I_H 。

由于提出的基于 INSCT 的人脸图像超分辨率复原算法复原得到的 \hat{I}_H 效果较好, 图像的高频信息相当丰富, 而且不会出现 Baker 方法^[8]中的方块效应以及基于小波和基于 Contourlet 变换的超分辨率^[16]复原结果中会出现的 Pseudo-Gibbs 现象, 并且降质后的 $D(\hat{I}_H)$ 与低分辨率图像 I_L 非常相似; 及采用最速下降法求解相当费时。因此本文算法取 $\lambda \rightarrow \infty$ 。

3 非下采样 Contourlet 变换

信号处理中一种重要的思想方法是采用一种可逆的线性变换, 变换的效率体现在能够用尽量少的基函数反映所处理信号的本质特征。但是在允许一定冗余度的情况下, 扩大基函数集合可以使变换更加灵活、完善的表示图像信息。冗余性基函数通常能更有效获取信号的一些特征, 在图像复原、去噪、增强以及边缘检测应用中, 冗余性变换处理效果要

优于非冗余性变换。

Contourlet 变换中的塔型滤波带结构具有非常有限的冗余性, 在拉普拉斯金字塔分解和方向滤波带分解中都有上下采样, 该变换不是平移不变的。由此在使用 Contourlet 变换用于图像复原时奇异点周围容易引入 Pseudo-Gibbs 现象。

NSCT 变换是在 Contourlet 变换的基础上由 A. L. Cunha 等^[17]提出的。非下采样 Contourlet 变换(Nonsampled Contourlet Transform, NSCT)去掉了 Contourlet 变换中的采样操作, 是一种具有平移不变性的表示方法, 可克服 Pseudo-Gibbs 现象。目前 NSCT 已经成功地用于图像融合等方面^[18~20]。NSCT 是一种超完备的变换。NSCT 与 Contourlet 变换^[21,22]在结构上很相似, 不同的是, NSCT 去掉了抽样环节, 它是由一个非下采样的金字塔滤波器(Nonsampled pyramid, NSP)和一个非下采样的方向滤波器组(Nonsampled directional filter bank, NSDFB)构成。NSCT 不仅具有多尺度、良好的空域和频域局部特性和方向特性外, 还具有平移不变特性以及更高的冗余度。

NSCT 中的金字塔滤波器和 DFB 都是非下采样的, 这使得其具有平移不变性。并由此衍生出两个好处: 1) 更高的冗余度保证了图像经 NSCT 后在所得到的各个子带中的视觉特征及信息更加的完整; 2) 根据多抽样率理论, NSCT 的低频子带不会有频率混淆现象产生, 因而具有更强的方向选择性。

NSCT 首先采用 NSP 获得图像的多尺度分解, 然后采用 NSDFB 对得到的各尺度子带图像进行方向分解, 从而得到不同尺度、不同方向的子带图像(系数)。图 1 所示为 NSCT 变换流程图。

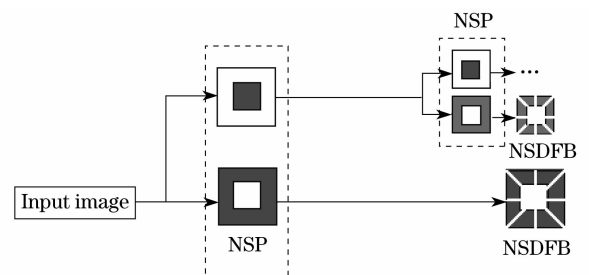


图 1 NSCT 的流程图

Fig. 1 Nonsampled contourlet (NSCT) algorithm

3.1 非下采样金字塔(NSP)

非下采样金字塔分级结构是通过多级迭代的方式实现的。首先提供满足下列完全重建条件一组基本的低通、高通滤波器组

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 1, \quad (9)$$

其中 $H_0(z)$ 为低通分解滤波; $H_1(z)$ 为高通分解滤波器; $G_0(z)$ 为低通重建滤波器; $G_1(z)$ 为高通重建滤波器。图 2 为非下采样金字塔滤波器组一级分解重建的结构示意图。

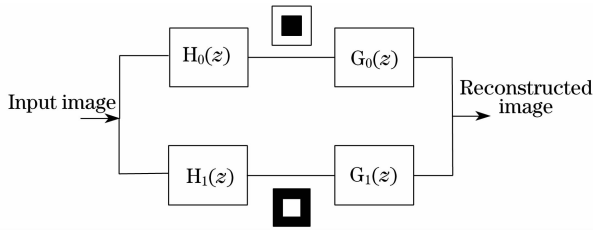


图 2 非下采样金字塔滤波器示意图

Fig. 2 Nonsubsampled pyramid filter

为了实现对图像的多尺度分解,采用“a trous”小波算法相似的思想,对低通滤波器和高通滤波器分别进行上采样(对于第 J 尺度上的分解,在相邻的两个滤波器系数间插入 $2^j - 1$ 个零),然后对上一尺度的低频图像用上采样后的低通滤波器进行低通滤波,得到低频图像,对上一尺度低频图像用上采样后的高通滤波器进行高通滤波,得到分解后的高频图像。该过程可以反复进行分解,类似于拉普拉斯金字塔变换。

3.2 非下采样方向滤波器组

非下采样方向滤波器组(NSDFB)是 Contourlet 变换中临界采样 DFB 的平移不变版本。它也是采用“a trous”算法的思想,即对级联的每级滤波器进行上采样,来获得平移不变性。非下采样 DFB 的基本模块是一个双通道的非下采样滤波器组。

NSCT 在基本扇形滤波器组或钻石滤波器组的基础上,通过对滤波器的操作来得到需要的象限滤波器组、平行滤波器组,避免了对图像的采样操作。非下采样方向滤波器组的结构为:1) 先经过扇形滤波器组和象限滤波器组将图像分为 4 个方向子带;2) 经过平行滤波器组的迭代分为各个不同的方向子带。

为了实现更好的、更加精细的方向分解,对非下采样的方向滤波器组进行迭代。在第一层方向分解之后,都用一个梅花形矩阵 Q 对方向滤波器组中的所有滤波器进行上采样,作为下一层方向分解的方向滤波器组。定义

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (10)$$

图 3 显示了非下采样 DFB 分解中双通道滤波器的迭代结构。

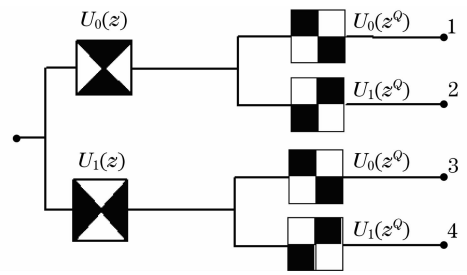


图 3 迭代 NSDFB 的分解结构图

Fig. 3 Iterative nonsubsampled directional filter bank

如果对某尺度下子带图像进行 k 级方向分解,可得到 2^k 个与原始输入图像尺寸大小相同的方向子带图像。图像经 J 级 NSCT 分解后可得到 1 个低频子带图像和 $\sum_{j=1}^J 2^k$ 个带通方向子带图像,其中 k_j 为尺度 j 下的方向分解级数。

NSCT 继承了 Contourlet 变换的多尺度、多方向以及良好的空域和频域局部特性,变换后能量更加集中,能够更好地捕捉和跟踪图像中重要的几何特征。由于 NSCT 分解和重构中不存在下采样和上采样环节,因此图像的分解和重构过程中不具有频率混叠项,因而滤波器的设计自由度更大,也使得 NSCT 具有平移不变性以及得到各级子带图像与源图像具有尺寸大小相同等特性,在图像复原或者融合过程中能够有效减少配准误差对融合结果的影响以及较容易找到各个子带图像之间的对应关系,从而有利于运算的实现。

4 改进的 NSCT

由于 NSCT 结果每一层图像大小都一样,不能像拉普拉斯金字塔^[23]那样建立高低分辨率图像的对等关系。虽然在文献[24]中将低分辨率图像插值放大后的结果图像与高分辨率图像的 NSCT 的低通部分作为对等图像,但是实际上两者存在差异。

考虑到拉普拉斯金字塔可建立多分辨率金字塔结构,并且建立拉普拉斯金字塔需要的运算量比建立 NSCT 的非抽样金字塔(需要采用“a trous”算法,而“a trous”算法计算比较耗时)运算量少很多。虽然拉普拉斯金字塔理论上会产生 Pseudo-Gibbs 现象,但是在实际复原过程中,产生的 Pseudo-Gibbs 现象相对来说不是很明显。因此对 NSCT 进行改进。NSCT 采用非抽样的金字塔结构和非抽样方向滤波器组。提出的 INSCT 是将 NSCT 的非抽样金字塔换为拉普拉斯金字塔,而非

抽样方向滤波器组还是保持不变,在处理速度和冗余度上进行了折中,并且可建立多分辨率金字塔结构。

对人脸图像进行 INSCT 结果图如图 4 所示,人脸图像进行了 4 层拉普拉斯金字塔,拉普拉斯金字塔的每一层进行 4 个方向的非下采样方向滤波。

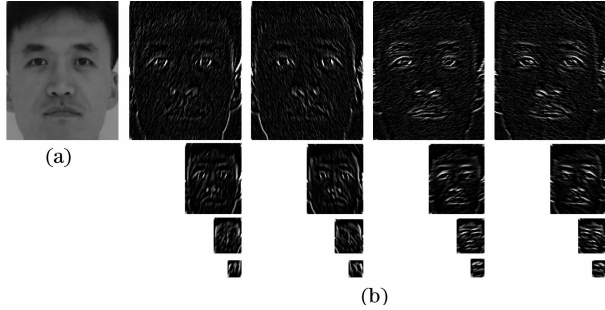


图 4 人脸图像及其 INSCT 变换结果图(a)原始人脸图像; (b)4 层拉普拉斯金字塔进行 4 个方向的非下采样方向滤波后的结果(INSCT 变换结果)

Fig. 4 Face image and its 4 layers INSCT pyramids. (a) Original face image; (b) results of INSCT transformation

4.1 基于改进的 NSCT(INSCT)的人脸图像超分辨率

采用 INSCT 的方法提取特征。INSCT 不仅能够准确地捕获到不同尺度,不同方向,不同频率的子带中的图像边缘信息,而且与 Contourlet 变换相比在图像复原时不容易产生明显的 Pseudo-Gibbs 现象,同时相对于 NSCT 减少了计算量。

为了表示人脸特征,建立 INSCT 的金字塔结构。分辨率越高,信息也相对越多;而分辨率越低,信息也相对越少。因此对分辨率高的层,分解的方向可以增加,而分辨率低的层,对其分解的方向可减少。如图 5(a)所示,对待复原的低分辨率图像,使用 INSCT 将其分解为 2 层,上面的一层使用非下采样 DFB 分解为 8 个方向,下面的一层使用非下采样 DFB 分解为 4 个方向。如图 5(b)所示,对训练库中的高分辨率图像进行 4 层分解,对最上面的两层分解为 16 个方向。

基于 INSCT 的超分辨率复原需要获得待复原的低分辨率图像细节部分的系数(即图 5(a)虚线部分的系数)。这就需要待复原的低分辨率图像的中频特征(即第三、四层特征)在高分辨率图像库中进行特征匹配,并将最匹配的特征对应的高频系数(即第一、二层系数)拷贝到待复原的低分辨率图像的细节部分。

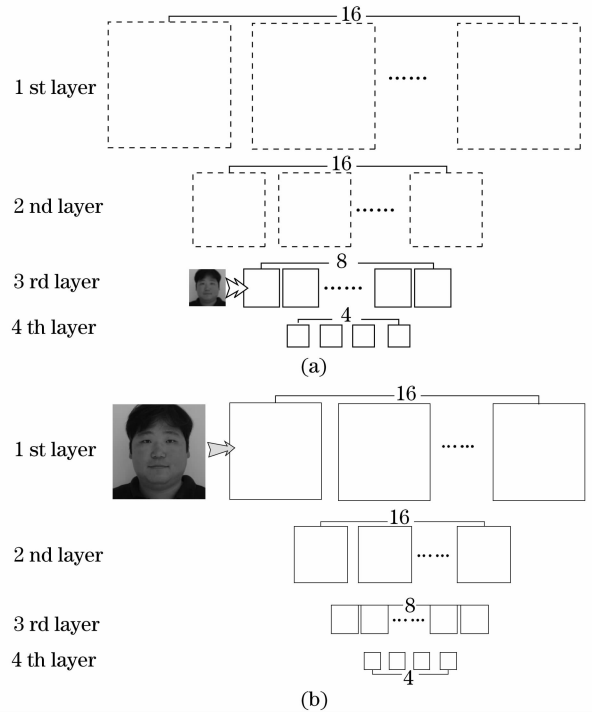


图 5 高低分辨率图像分解示意图。(a)低分辨率图像分解示意图,虚线部分表示需要通过学习获取的部分,(b) 高分辨率图像分解示意图

Fig. 5 Illustration of decomposition of low resolution image (a) and high resolution image (b)

由于 INSCT 有较大的冗余度,以及得到各级子带图像与同级的拉普拉斯金字塔图像具有尺寸大小相同的特性,在超分辨率复原过程中能够有效减少配准误差对融合结果的影响及较容易找到各个子带图像之间的对应关系。因此只是使用 INSCT 金字塔系数作为特征。

设输入的低分辨率人脸图像 I , 它的 INSCT 金字塔从第三层开始构建,定义 I 中的任一像素点 p 的父结构为从第三层和第四层中与 p 对应的一组像素的特征向量,如图 6 所示。

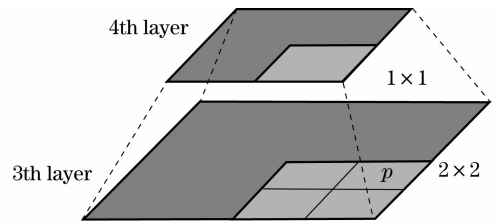


图 6 父结构示意图

Fig. 6 Parents structure

像素的特征向量由该层和该层以下的非下采样的方向滤波器提取的特征信息组成。像素点 p 的父结构特征就是一个 $8+4$ 维的向量(其中,第三层分解为 8 个方向,第四层分解为 4 个方向特征)定义为

$$\mathbf{S}_3(I)(m,n) = \begin{bmatrix} L_3 D_1(I)(m,n), & L_3 D_2(I)(m,n), & L_3 D_3(I)(m,n), & L_3 D_4(I)(m,n) \\ L_3 D_5(I)(m,n), & L_3 D_6(I)(m,n), & L_3 D_7(I)(m,n), & L_3 D_8(I)(m,n) \\ L_4 D_1(I)\left(\frac{m}{2}, \frac{n}{2}\right), & L_4 D_2(I)\left(\frac{m}{2}, \frac{n}{2}\right), & L_4 D_3(I)\left(\frac{m}{2}, \frac{n}{2}\right), & L_4 D_4(I)\left(\frac{m}{2}, \frac{n}{2}\right) \end{bmatrix}, \quad (11)$$

其中 $L_i D_j(I)(m,n)$ 表示图像位置为 (m,n) 的第 i 层的第 j 个方向的特征。采用同样的方法,对训练库中所有高分辨率人脸图像训练样本 T_i ,分别在第三层构建父结构 $\mathbf{S}_3(T_i)(m,n)$ 。

由于对人脸图像进行复原,正面人脸图像相对于其它图像具有一些特殊性,即人脸图像具有全局约束,即不同人的鼻子,嘴,眼睛的位置相对于正面人脸来说,基本上是固定的。因此利用这种性质,进行寻找最优匹配时,只是在对应的位置寻找最优匹配。也就是说,低分辨率图像的位置 (m,n) 点进行复原时,只需要在训练样本库的低分辨率图像中对于的位置 (m,n) 寻找最优匹配的。使用这样的方法,可以大量减少计算的时间,提高运算效率。

4.2 算法实现

算法分为两个部分,即训练部分和学习部分。具体的步骤如下:

(1) 训练过程

对输入的每一个训练样本 T_i 建立 INSCT 金字塔并构建父结构特征 $\mathbf{S}(T_i)$ 。在实验中父结构特征从第三层开始构建。

(2) 学习过程

1) 对待复原的图像 I 建立父结构特征 $\mathbf{S}_3(I)$

2) 将待复原的图像位置为 (m,n) 的特征 $\mathbf{S}_3(I)(m,n)$ 与训练样本相应位置的特征 $\mathbf{S}_3(T_i)(m,n)$ 进行比较,计算欧式距离,

$$d_i = \min \|\mathbf{S}_3(T_i)(m,n) - \mathbf{S}_3(I)(m,n)\|. \quad (12)$$

3) 将最小距离 d_i 对应的 $\mathbf{S}_3(T_i)(m,n)$ 的低层(第一、二层)系数(高频信息)拷贝到到复原图像相应的低层位置(即图 5(a)虚线部分)。

4) 待复原图像和获取的系数一起进行 INSCT 反变换得到高分辨率复原图像。

5 实验结果与分析

使用亚洲人脸标准图像数据库 IMDB^[25] 中的人脸图像,提取人脸的面部图像,并进行归一化,归一化成大小为 192 pixel × 160 pixel 的人脸图像。把 192 pixel × 160 pixel 人脸图像作为高分辨率人脸

图像,对其进行降质处理,降质为大小为 48 pixel × 40 pixel 的图像,将其作为低分辨率人脸图像(图像复原时需要放大 16 倍)。实验选取了所有不带眼镜的 75 人进行实验,每人选正面人脸,总共 75 幅。随机选择其中的 8 人(8 幅人脸图像)作为测试数据。

实验结果如图 7 所示,图 7(a)真实的高分辨率图像;(b)为基于 INSCT 变换的结果;(c)为最近邻插值算法的结果;(d)为 Cubic B-spline 插值算法的结果;(e)为基于 Contourlet 变换的结果。

分别从复原图像的主观视觉效果和客观评价指标峰值信噪比和熵对算法的性能进行分析。从主观视觉效果来看基于 INSCT 变换的结果明显比最近邻插值算法和 Cubic B-spline 插值算法效果好得多。本文算法复原结果接近于原始高分辨率图像,与文献[16]的基于 Contourlet 变换的方法相比,在细节上的结果相对较好,并且基本没有明显的 Pseudo-Gibbs 现象,而 Contourlet 变换的复原结果在轮廓部有明显的 Pseudo-Gibbs 现象。图 8 为基于 INSCT 变换的复原结果与基于 Contourlet 变换的复原结果及其局部放大比较图。可以看出基于 Contourlet 变换的复原结果在轮廓部有明显的 Pseudo-Gibbs 现象。而基于 INSCT 的复原结果基本上不存在此现象。

由于测试中采用了 8 幅人脸图像,因此最终客观评价采用平均峰值信噪比及平均熵。(平均峰值信噪比是指 8 幅测试样本峰值信噪比的平均;平均熵类似)。

峰值信噪比(PSNR)定义为

$$R_{\text{PSNR}} = 10 \times \lg \left(\frac{255^2}{\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [\text{HR}(x,y) - \text{SR}(x,y)]^2} \right), \quad (13)$$

其中原始的高分辨率图像为 HR,由低分辨率图像复原得到的超分辨率图像为 SR,它们在像素点 (x,y) 上的灰度值分别为 $\text{HR}(x,y)$ 和 $\text{SR}(x,y)$ 。图像的长度和宽度以像素点数表征分别为 M 和 N 。

图像的熵定义为



图 7 各种不同方法的实验结果(a) 真实的高分辨率图像；(b) 基于 INSCT 变换的结果；(c) 最近邻插值算法的结果；(d) Cubic B-Spline 插值算法的结果；(e) 基于 Contourlet 变换的结果
 Fig. 7 Comparison of the results by different methods. (a)Original high-resolution image; (b) INSCT's results; (c) nearest neighbor; (d) cubic B-spline; (e) contourlet's result

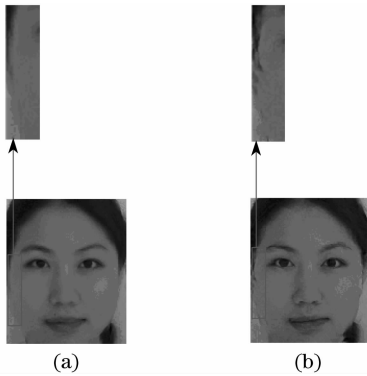


图 8 实验结果及其局部放大比较图。(a) 本文方法的结果；(b) contourlet 系数方法
 Fig. 8 Comparison of the results and its local enlarged images. (a) Proposed method's results; (b) contourlet's results

$$H = - \sum_{i=0}^{255} p(i) \log_2 p(i) \quad (14)$$

其中复原图像的灰度值从 1 到 255 的分布为 $p = \{p(1), p(2), \dots, p(i), \dots, p(255)\}$ 。其中 $p(i)$ 表示灰度值 i 在两幅图像中出现的概率。

图 9 为各种不同方法的平均峰值信噪比, 本文方法具有最高的平均峰值信噪比。最近邻插值算法与 Cubic B-Spline 插值算法的结果的峰值信噪比较低, 只有 29 dB 左右。Contourlet 方法的结果只有 31.1 dB, 而本文算法的结果能够到达 32.6 dB。表 1 为各种不同方法复原结果的平均熵, 本文算法和 Contourlet 方法结果的熵都较插值算法结果的熵有一定的提高。虽然本文方法复原结果的熵较 Contourlet 方法复原结果的熵略低, 但是基本不存在 Contourlet 方法复原结果的 Pseudo-Gibbs 现象,

并且峰值信噪比也较 Contourlet 方法提高 1.5 dB。

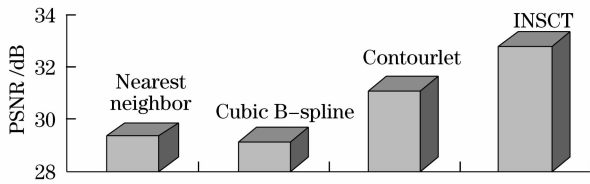


图 9 各种不同方法的平均峰值信噪比

Fig. 9 Average peak signal-to-noiseratio (PSNR) comparison by different methods

表 1 各种不同方法复原图像的平均熵

Table 1 Aaverage entropy comparison of different methods

Algorithm	Entropy
Nearest neighbor	6.89
Cubic B-Spline	6.91
Contourlet transform	7.00
INSCT	6.97

本文算法的效果与非下采样方向滤波器的方向数有一定的关系,方向数的选取在算法中起到一定的作用。在下面的实验中,还是以前面随机选择的 8 人作为测试样本,以剩下的 67 人(67 幅人脸图像)作为训练数据。

第四、三层的方向系数用于特征的匹配,而第二、一层是需要复原的高频系数。因此在实验中分别对二、一层的方向和四、三层的方向进行实验分析。首先固定第四、三层的方向数分别为 2^2 、 2^3 , 然后对二、一层的方向数进行变化。

图 10 为第二、一层方向数对复原图像的平均峰值信噪比的影响,其中横坐标表示第二、一层的方向数,例如(1,2)表示第二、一层的方向数分别为 2^1 、 2^2 。从图 10 可以看出随着第二、一层的方向的增加,复原图像的峰值信噪比也随之增加,但是随着方向数的继续增加,峰值信噪比增加的幅度越来越小。随着第二、一层方向的增加复原效果也越来越好,这主要是由于 INSCT 有较大的冗余度,并且随着方向的增加,冗余度越大。并且得到的各级子带图像与同级的拉普拉斯金字塔图像具有尺寸大小相同的特性,在超分辨率复原过程中能够有效减少配准误差对复原结果的影响。

第四、三层的方向数对平均峰值信噪比的影响如图 11 所示。固定第二、一层的方向数分别为 2^4 、 2^4 , 然后对第四、三层的方向数进行变化。如果方向数取值太大,会引入噪声,可能使得匹配出现偏差;反之,不能较好地表示方向信息,同样使得匹配出现偏差。由于本实验中预先对人脸图像进行了对齐处

理,并匹配时是相应位置点进行匹配,这导致实验的结果对第四、三层方向的变化不是很敏感。

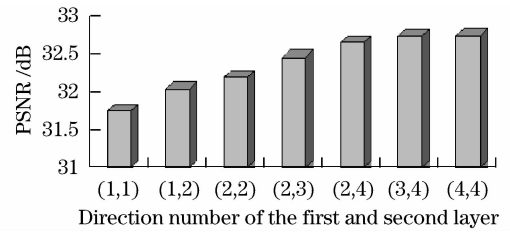


图 10 第二、一层方向数对复原图像的平均峰值信噪比的影响

Fig. 10 PSNR versus the direction number of the first and second layer

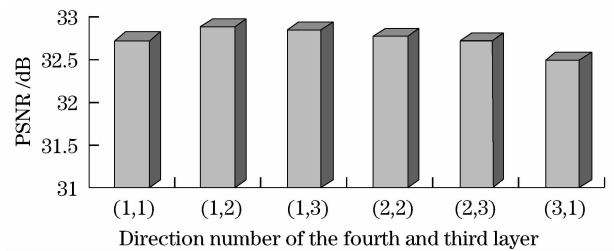


图 11 第四、三层方向数对复原图像的平均峰值信噪比的影响

Fig. 11 PSNR versus the direction number of the fourth and third layer

6 总 结

基于学习的超分辨率技术是超分辨率领域重要的研究热点之一。它在在医学、遥感、视频转换和安全监控等领域都有十分重要的应用。针对 Contourlet 变换在图像复原时奇异点周围容易引入 Pseudo-Gibbs 现象,提出了基于改进的 INSCT 的超分辨率复原算法。对 IMDB 人脸库的实验表明该算法具有较好的性能,复原出的超分辨率人脸图像无论在主观视觉上还是在客观评价指标上都取得较好的结果,复原的图像具有更好的视觉效果。

参 考 文 献

- 1 J. D. van Ouwerkerk. Image super-resolution survey[J]. *Image and Vision Computing*, 2006, **24**(10): 1039~1052
 - 2 Sung Cheol Park, Min Kyn Park, Moon Gi Kang. Super-resolution image reconstruction: a technical overview[J]. *IEEE Signal Processing Magazine*, 2003, **20**(3): 21~36
 - 3 David Capel, Andrew Zisserman. Computer vision applied to super resolution[J]. *IEEE Signal Processing Magazine*, 2003, **20**(3): 75~86
 - 4 Zheng Lixian, He Xiaohai, Wu Wei *et al.*. Learning-based super-resolution technique[J]. *Computer Engineering*, 2008, **34**(5): 193~195
- 郑丽贤,何小海,吴 炜 等. 基于学习的超分辨率技术[J]. *计算机工程*, 2008, **34**(5): 193~195

- 5 Zhuo Ning, Sun Huayan, Zhang Haijiang. A new approach for improvement of CCD imaging resolution[J]. *Acta Optica Sinica*, 2005, **25**(6):777~780
卓 宁, 孙华燕, 张海江. 一种新的提高 CCD 成像分辨率的方法[J]. *光学学报*, 2005, **25**(6):777~780
- 6 Qin Fengqing, He Xiaohai, Wu Wei. Image super-resolution reconstruction based on sub-pixel registration and iterative back projection[C]. *Proc. the International Conference, 2007 on Information Computing and Automation ICICA (2007)*, 2007, (1):71~74
- 7 Schultz R, Stevenson R. Extraction of high resolution frames from video sequences [J]. *IEEE Transactions on Image Processing*, 1996, **5**(6):996~1011
- 8 Baker S, Kanade T. Limits on super-resolution and how to break them [J]. *IEEE Conf. Computer Vision and Pattern Recognition*, 2000, **9**(2):372~379
- 9 Baker S, Kanade T. Hallucinating faces[C]. *Proc. the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (ICAFGR), Grenoble, France*, 2000; 83~88
- 10 Jian Sun, Nanning Zheng. Image Hallucination with Primal Sketch Priors[C]. *Computer Vision and Pattern Recognition, Proc.*, 2003 *IEEE Computer Society Conference*, 2003, **2**(7): 729~736
- 11 Ayan Chakrabarti, A. N. Rajagopalan. Super-resolution of face images using kernel PCA-based prior[J]. *IEEE Transactions on Multimedia*, 2007, **9**(4):888~892
- 12 Ni. K. S., Kumar, S. Single Image Superresolution Based on Support Vector Regression[C]. *Acoustics, Speech and Signal Processing, ICASSP 2006 Proceedings*, 2006, (2):601~604
- 13 Ni. K. S. Ni, Truong Q. Nguyen. Image superresolution using support vector regression [J]. *IEEE Transactions on Image Processing*, 2007, **16**(6):1596~1610
- 14 Wang X., Tang X. Hallucinating face by eigentransformation [J]. *IEEE Transactions on Systems, Man and Cybernetics, Part C*, 2005, **35**(3): 425~434
- 15 C. Su, Y. Zhuang, L. Huang *et al.*. Steerable pyramid based face hallucination [J]. *Pattern Recognition*, 2005, **38** (6): 813~824
- 16 Jiji. C. V, Chaudhuri. S. Single-frame images super-resolution through contourlet learning[J]. *EURASIP Journal on Applied Signal Processing*, 2006, **2006**: 1~11
- 17 Arthur L. da Cunha, Jianping Zhou, Minh N. Do. The nonsubsampling contourlet transform; theory, design, and applications[J]. *IEEE Transactions on Image Processing*, 2006, **15**(10):3089~3101
- 18 Bin Yang, Shutao Li, Fengmei Sun. Image Fusion Using Nonsampled Contourlet Transform[C]. *Fourth International Conference on Image and Graphics*, 2007: 719~724
- 19 Jianping Zhou, Arthur L. Cunha, Minh N. Do. Nonsampled contourlet transform; construction and application in enhancement[C]. *ICIP 2005. IEEE International Conference on Image Processing*, 2005, (1):469~472
- 20 Zhang Qiang, Guo Baolong. Remote sensing image fusion based on the nonsubsampling contourlet transform [J]. *Acta Optica Sinica*, 2008, **28**(1):74~80
张 强, 郭宝龙. 基于非采样 Contourlet 变换的遥感图像融合算法[J]. *光学学报*, 2008, **28**(1):74~80
- 21 Liu Kun, Guo Lei, Chang Weiwei. Regional feature self-adaptive image fusion algorithm based on contourlet transform[J]. *Acta Optica Sinica*, 2008, **28**(4):681~686
刘 坤, 郭 雷, 常威威. 基于 Contourlet 变换的区域特征自适应图像融合算法[J]. *光学学报*, 2008, **28**(4):681~686
- 22 Zhang Jingjing, Fang Yonghua. Novel denoising method for remote sensing image based on contourlet transform[J]. *Acta Optica Sinica*, 2008, **28**(3):462~466
张晶晶, 方勇华. 基于 Contourlet 变换的遥感图像去噪新算法 [J]. *光学学报*, 2008, **28**(3):462~466
- 23 Burt P J, Adelson E H. The Laplacian pyramid as a compact image code[J]. *IEEE Transactions on Communications*, 1983, **31** (4): 532~540
- 24 Wei Ni, Baolong Guo, Liu Yang. Example based super-resolution algorithm of video in contourlet domain[C]. *Fourth International Conference on Image and Graphics*, 2007:13~19
- 25 Dong H, Gu N. Asian face image database PF01[R]. Pohang University of Science and Technology, 2001.