

全局重建和位置块残差补偿的 人脸图像超分辨率算法

马祥, 齐春

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对传统基于学习的人脸图像超分辨率算法存在高频细节信息损失过多问题, 提出一种全局重建和位置块残差补偿相结合的人脸图像超分辨率新算法. 首先利用高、低分辨率训练集所有样本, 使用基于权值学习的全局重建算法得到初步的人脸图像, 再结合图像模糊和下采样过程, 产生高、低分辨率残差图像训练集, 最后使用基于位置块的残差补偿算法, 对初步的人脸图像进行高频细节补偿得到最终结果. 对比实验结果表明, 相比同类基于学习的人脸图像超分辨率算法, 在将人脸图像分辨率提高 4×4 倍的情况下, 新算法的平均峰值信噪比可提高 $0.65 \sim 3.55$ dB, 可以更好地重建出局部高频细节信息.

关键词: 人脸图像; 超分辨率; 残差补偿; 位置块

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2010)04-0009-04

A Face Image Super-Resolution Algorithm Based on Global Reconstruction and Position-Patch Residue Compensation

MA Xiang, QI Chun

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Focusing on the high loss of subtle facial details in traditional algorithm, a face image super-resolution algorithm through global reconstruction and position-patch based residue compensation is presented. The optimal coefficients of the low-resolution training images are computed and transformed into high-resolution space to reconstruct the global high-resolution image. The residue training set is obtained by a smoothing and down-sampled processing. The residue compensation based on position is performed to better recover face subtle details using the residue training set. Experimental results show that the proposed approach synthesizes high-resolution faces with more details and the average of peak signal-to-noise ratios is improved about 0.65 dB to 3.55 dB compared with some existing learning-based methods.

Keywords: face image; super-resolution; residue compensation; position-patch

图像超分辨率技术是指由一张或多张低分辨率图像重建出高分辨率图像. 目前超分辨率技术已经在很多领域得到了应用^[1], 如军事、刑侦、医学、航拍、安防等. 一般而言, 适用于各种图像的超分辨率重建算法没有针对于某一类图像设计的算法重建效果好, 主要原因在于前者没有利用该类图像自身的

一些固有特征信息. 目前针对人脸的超分辨率重建技术主要集中在基于学习的方法上^[2-4], 即研究如何由单帧输入的低分辨率人脸图像通过训练集重建出高分辨率人脸图像的方法. 这类算法通常先利用特征脸模型或流形学习模型重建出初步的人脸图像, 再进行局部高频细节补偿得到最终结果. 但是, 采用

这种模型会导致算法冗长,而且由于特征提取步骤的引入,会在重建过程中过多损失非特征信息,而这些非特征信息对图像的细节恢复是有贡献的.本文提出一种全局权值重建模型,使用所有训练集样本图像对输入的低分辨率图像进行线性组合逼近来求得重建权值,这样就相对减少了重建中信息的损失.在残差补偿阶段,本文采用一种基于位置块进行局部高频细节补偿的新方法,比起常见的基于邻域块残差补偿方法^[4],能更好地恢复出人脸高频细节信息.

1 原理

1.1 全局权值重建

将二维人脸图像用一维向量来表示,设低分辨率人脸图像训练集为 $[I_1, I_2, \dots, I_n]$,对应的高分辨率人脸图像训练集为 $[h_1, h_2, \dots, h_n]$,其中 n 为训练集样本个数, I 为输入的低分辨率人脸图像.由于人脸结构的相似性,人脸图像可以由训练集图像线性组合进行逼近,即

$$I = \sum_{i=1}^n w_i I_i + e \quad (1)$$

式中: e 为逼近误差向量; w_i 为重建权值.这种线性组合关系是由人脸的结构相似性来决定的.重建权值向量 w 中的每个元素 w_i 体现了 I_i 对重建 I 所做的贡献,所有的权值 w_i 相加之和为1,即

$$\sum_{i=1}^n w_i = 1 \quad (2)$$

使用欧式距离来度量重建误差,根据式(1),设 ϵ 为重建误差建立下式

$$\epsilon = \|I - \sum_{i=1}^n w_i I_i\|^2 \quad (3)$$

假设 ϵ 越小,重建就越成功,为求出 w_i ,结合式(2),建立一个成本函数

$$w = \arg \min_{w_i} \epsilon \quad (4)$$

这是一个在式(2)约束下的最小二乘问题.由于人脸结构的相似性,可以将 I_i 视为 I 的 n 个邻域向量,这样就将问题转化为求取 I 的 n 个邻域点权值的问题.使用 locally linear embedding (LLE)算法中求取重建权值的方法^[5-6],有

$$\epsilon = \|I - \sum_{i=1}^n w_i I_i\|^2 = \left\| \sum_{i=1}^n w_i (I - I_i) \right\|^2 = \sum_{i=1}^n \sum_{k=1}^n w_i w_k C_{ik} \quad (5)$$

构造一个局部协方差矩阵 C , C 中元素 $C_{ik} = (I - I_i)^T (I - I_k)$,使用拉格朗日乘子法,在保证所有的

w_i 相加和为1的约束下可得到

$$w_i = \left(\sum_{k=1}^n C_{ik}^{-1} \right) / \left(\sum_{l=1}^n \sum_{m=1}^n C_{lm}^{-1} \right) \quad (6)$$

式中: l, m 均为正整数; C_{ik}^{-1} 表示逆矩阵 C^{-1} 中第 i 行第 k 列元素.在实际运算中, C 可能是一个奇异矩阵,此时必须正则化 C .在实际的算法中常采用一个更快捷的方法,即求线性系统方程 $Cw=1$,最后令解得的权值满足相加和为1.

利用权值,可以得到输入图像 I 的重建图像

$$\tilde{I} = \sum_{i=1}^n w_i I_i \quad (7)$$

由式(3)可知,重建的效果依赖于

$$\epsilon = \|I - \tilde{I}\|^2 \quad (8)$$

图像分辨率降低的降质模型可以表示为^[7]

$$X_L(i, j) = \frac{1}{q^2} \sum_{k=0}^{q-1} \sum_{l=0}^{q-1} X_H(qi+k, qj+l) + n(i, j) \quad (9)$$

式中: q 为正整数,表示分辨率提升倍数; $X_L(i, j)$ 、 $X_H(i, j)$ 分别表示低分辨率和高分辨率图像;在像素灰度矩阵 (i, j) 位置的灰度值; $n(i, j)$ 表示随机噪声.为简单起见,采用矩阵向量形式表示,设 X_H 为 M 维列向量, X_L 为 N 维列向量,式(9)可简化为

$$X_L = AX_H + n \quad (10)$$

式中: A 为 $M \times N$ 维矩阵,包含一个模糊和下采样过程.

将式(7)中低分辨率训练集图像 I_i 置换为高分辨率训练集图像 h_i ,结果为

$$H = \sum_{i=1}^n w_i h_i \quad (11)$$

由式(10)和式(11)得到

$$AH + n = \sum_{i=1}^n (Ah_i + n)w_i \quad (12)$$

不考虑噪声影响,式(12)可重写为

$$AH = \sum_{i=1}^n I_i w_i = \sum_{i=1}^n I_i w_i = \tilde{I} \approx I \quad (13)$$

由式(13)可以看到, H 的降质结果近似等于输入的低分辨率人脸图像 I , H 即为初步的重建结果.由式(8)和式(13)可见, ϵ 越小, H 的降质结果就越接近 I .

1.2 位置块残差补偿

文献[4]提出一种基于邻域块的残差补偿方法,需要在残差训练集中寻找出与输入的图像块欧氏距离最小的若干个图像块来进行运算,对初步所得的结果进行局部高频细节补偿.这种算法由于包含了搜索步骤,因而运算量大,并且补偿的效果不理想.

本节提出一种基于位置图像块的残差补偿方法,不经过搜索步骤,直接使用训练集人脸图像同一位置的图像块来进行运算.用于残差补偿的低分辨率残差训练集图像 $\mathbf{L}_m = \mathbf{I}_m - \mathbf{D}(\mathbf{H}_{l,m})$,其中 $\mathbf{D}$ 为下采样函数, $\mathbf{H}_{l,m}$ 是 $\mathbf{I}_m$ 利用全局重建算法的重建结果,用于残差补偿的高分辨率残差训练集图像 $\mathbf{H}_m = \mathbf{h}_m - \mathbf{H}_{l,m}$.将每个残差训练集图像 $\mathbf{L}_m$ 、 $\mathbf{H}_m$ 分别划分为相互重叠的图像块,分别表示为图像块矩阵的形式 $\{\mathbf{L}_m(i,j)\}_p$ 、 $\{\mathbf{H}_m(i,j)\}_p$,其中 p 是每幅图像所划分的图像块的个数.以图像块 $\mathbf{L}_m(i,j)$ 为例,其周围重叠图像块的划分方式如图1所示.

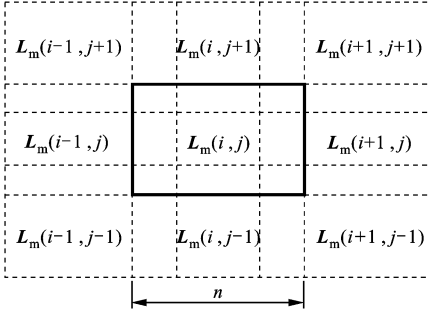


图1 图像块 $\mathbf{L}(i,j)$ 及周围重叠图像块

低分辨率人脸残差图像 $\{\mathbf{R}(i,j)\}_p$ 定义为:输入的低分辨率图像 $\mathbf{I}$ 减去 $\mathbf{H}$ 的下采样图像.

令低分辨率残差图像的每个图像块表示为

$$\mathbf{R}(i,j) = \sum_{m=1}^M \omega_m(i,j) \mathbf{L}_m(i,j) + \mathbf{e}(i,j) \quad (14)$$

式中: $\mathbf{e}(i,j)$ 为 $\mathbf{R}(i,j)$ 的重构误差; M 为残差训练集图像对个数; $\omega_m(i,j)$ 体现了残差训练集每个样本图像在 (i,j) 位置的图像块对重构高分辨率图像块所做的贡献,它满足下式

$$\sum_{m=1}^M \omega_m(i,j) = 1 \quad (15)$$

建立一个成本函数

$$\mathbf{w}(i,j) = \arg \min_{\omega_m(i,j)} \left\| \mathbf{R}(i,j) - \sum_{m=1}^M \omega_m(i,j) \mathbf{L}_m(i,j) \right\|^2 \quad (16)$$

使用同上一节相同的方法可以计算出 $\omega_m(i,j)$,从而合成相应的高分辨率残差图像块

$$\tilde{\mathbf{R}}(i,j) = \sum_{m=1}^M \mathbf{H}_m(i,j) \omega_m(i,j) + \mathbf{e}(i,j) \quad (17)$$

将求得的所有高分辨率残差图像块,按照它们原先在人脸图像的位置,拼接为整体图像,即为残差人脸图像 $\tilde{\mathbf{R}}$.将残差图像 $\tilde{\mathbf{R}}$ 叠加到初步重建结果 $\mathbf{H}$,得到最终的超分辨率结果

$$\mathbf{H}^* = \mathbf{H} + \tilde{\mathbf{R}} \quad (18)$$

2 实验结果与分析

本文采用了 CAS-PEAL-R1 大规模人脸数据库^[8]进行实验,随机选用了 290 个不同个体的平常表情正面人脸图像,手工标定并对齐了两眼中心和嘴巴中心 3 个特征点,根据需要截取人脸区域,并将图片大小统一到 128×96 像素.用大小为 7×7 像素、方差 $\sigma = 0.85$ 的高斯模糊模板,将所有高分辨率,图像采用模糊和平均下采样方法降低分辨率得到 32×24 像素的低分辨率图像.本文随机使用 150 对高低分辨率图像作为训练图像,20 个低分辨率图像作为测试图像,分别与插值算法、文献[2]和文献[4]算法进行了对比实验.文献[2]的方法为基于统计的方法,为了将其效果发挥到最佳,取 270 对样本作为训练集,并调整文献[2]和文献[4]中参数,使其算法效果发挥到最佳.在残差补偿阶段,文献[4]和本文的方法均取 3×3 像素的图像块.

采用以上 4 种算法,对低分辨率测试图像进行 4×4 倍超分辨率实验,其中一组代表性的结果见图 2.由图 2 可见,文献[2]算法对人脸中心感兴趣区域超分辨率重建的效果较好,但是其他部位尤其脸部边缘很模糊,脸部细节损失得较严重.文献[2]算法是基于统计的方法,如果将其训练集高低分辨率图像对个数取到 150,即和其他两种算法保持一致,效果还会下降很多.文献[4]算法相比文献[2]算法重建的视觉效果要好些,但是面部边缘和发际还是有

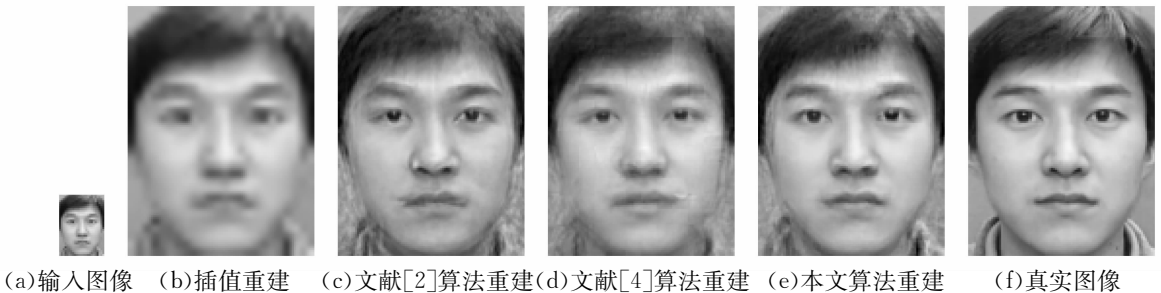


图2 采用4种算法的 4×4 倍超分辨率重建结果

些模糊,所恢复出的面部细节也比本文算法少。

采用被广泛使用的峰值信噪比(PSNR)将实验结果量化描述,4种算法的重建结果见图3。由图3可以看出,本文的所有重建结果的PSNR值几乎都高于其他方法。4种算法重建结果的平均PSNR分别为26.92 dB、26.27 dB、23.37 dB和25.57 dB。插值算法结果的PSNR值并不是最低的,这是因为PSNR对图像质量的量化评价是有局限性的。为了研究高、低分辨率训练图像对数对本文算法的影响,在训练图像对数分别为10、30、50、100、150、200的情况下,使用本文的算法对20个测试低分辨率图像进行实验,并将全部实验结果量化表示于图4。从图4可以看到,随着训练集高低分辨率图像对数的增加,利用本文算法重建出的高分辨率图像的PSNR值逐渐增大,当训练集高低分辨率图像对数超过100左右后,重建出的高分辨率图像的PSNR值逐渐趋于平稳。

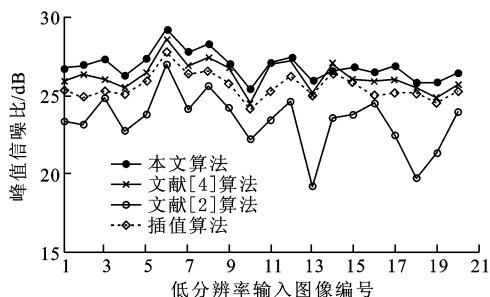


图3 4种算法的的峰值信噪比对比

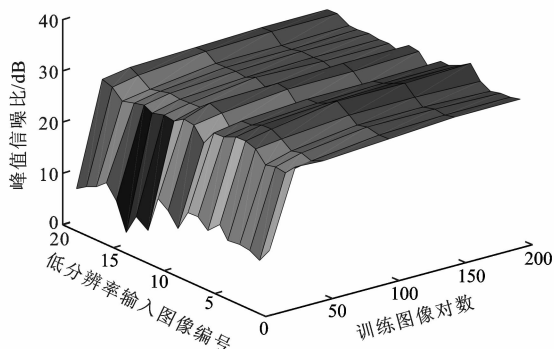


图4 训练集图像对数对本文算法的影响

3 结 论

本文提出一种基于学习的两步人脸图像超分辨率新算法,不同于使用复杂的概率模型和流形学习方法进行全局重建,本文算法使用了训练集所有人

脸样本来求取重建系数,而不是特征脸,这样就减少了一些非特征信息的损失。与传统的基于邻域块进行高频细节补偿方法不同,本文采用一种基于位置块的高频细节补偿方法,避免了搜索步骤。实验结果表明,本文算法在将人脸图像分辨率提高 4×4 倍的情况下,相对于所比较的算法,可以更好地重建出高分辨率图像。

参考文献:

- [1] 宋锐,贾媛,吴成柯,等. 自适应估计模糊参数的最大后验概率超分辨率复原算法[J]. 西安交通大学学报, 2008, 42(10):1254-1258.
SONG Rui, JIA Yuan, WU Chengke, et al. A maximum posteriori super-resolution algorithm for adaptive estimation of blur parameter [J]. Journal of Xi'an Jiaotong University, 2008, 42(10):1254-1258.
- [2] WANG Xiaogang, TANG Xiaoou. Hallucinating face by eigentransformation [J]. IEEE Transactions on Systems Man and Cybernetics; Part C, 2005, 35(3): 425-434.
- [3] PARK J S, LEE S W. An example-based face hallucination method for single-frame, low-resolution facial images [J]. IEEE Transactions on Image Processing, 2008, 17(10): 1806-1816.
- [4] ZHUANG Yueting, ZHANG Jian, WU Fei. Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation [J]. Pattern Recognition, 2007, 40(11): 3178-3194.
- [5] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 22(12): 2323-2326.
- [6] SAUL L K. An Introduction to Locally Linear Embedding [EB/OL]. (2001-01-01) [2008-04-05]. <http://www.cs.toronto.edu/~roweis/lle/publications.html>.
- [7] LIU Ce, SHUM H Y, ZHANG C S. A two-step approach to hallucinating faces: global parametric model and local nonparametric model [C]//Proceeding of the Inter Conference on Image and Graphics. Piscataway, NJ, USA: IEEE, 2001: 192-198.
- [8] GAO Wen, CAO Bo, SHAN Shiquan, et al. The CAS-PEAL large-scale Chinese face database and baseline evaluations [J]. IEEE Transactions on System Man, and Cybernetics; Part A, 2008, 38(1): 149-161.

(编辑 刘杨)