

基于学习的超分辨率技术

郑丽贤, 何小海, 吴 炜, 杨晓敏, 陈 默

(四川大学电子信息学院图像信息研究所, 成都 610064)

摘 要: 基于学习的超分辨率算法使用一个图像训练集来产生一个学习模型, 运用该模型为输入的低分辨率图像创建更多的高频信息, 获得比基于重建算法更好的结果。该文介绍了基于学习的超分辨率技术的相关工作、理论基础和主要算法, 提出基于学习的超分辨率算法中仍需解决的关键问题, 展望其在未来的研究发展方向。

关键词: 超分辨率; 马尔可夫随机场; 图像金字塔

Learning-based Super-resolution Technique

ZHENG Li-xian, HE Xiao-hai, WU Wei, YANG Xiao-min, CHEN Mo

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

【Abstract】 Learning-based super-resolution technique predicts the high-resolution images from the input low-resolution ones, through learning from a training set which consists of a large number of other high-resolution images. And the results are better than the reconstruction based super-resolution algorithms. The related work, theory and algorithms of learning-based super-resolution are illustrated. The crucial problems which need to be resolved in further work are proposed. Directions of future research are pointed.

【Key words】 super-resolution; Markov random field; image pyramid

目前的图像超分辨率技术主要分为两类: 基于重建的超分辨率技术和基于学习的超分辨率技术。基于重建的方法是指所有的可用信息都从输入数据中得到, 没有任何附加的背景知识, 整个解决过程相当于一个信息提取和信息融合的问题, 最典型的方法是凸集投影法^[1]和最大后验概率估计法^[2]。由于图像含义未知, 因此所有信息只能全部从输入的图像序列中获得。随着分辨率放大系数的增加, 需要提供的输入图像样本数量急剧增加, 直到达到放大系数的上限后, 无论增加多少输入图像样本, 都无法再改善重建效果。针对重建算法的局限性, 基于学习的超分辨率技术作为一个前沿的研究领域应运而生。它已知图像含义, 从大量的训练样本集中获取先验知识作为超分辨率的依据。训练样本都是与输入图像包含同类信息的图像, 以输入图像为依据, 用学习过程中获得的知识对输入图像中的信息进行补充。基于学习的方法充分利用了图像本身的先验知识, 在不增加输入图像样本数量的情况下, 仍能产生新的高频细节, 获得比基于重建的算法更好的复原结果, 并能够较好地运用于人脸和文字等图像的复原。

1 相关工作

基于学习的超分辨率技术最早是由卡耐基—梅隆实验室的Baker S^[3]在2000年提出的。他们提出一种基于识别先验知识的方法, 通过算法去学习训练指定类别, 将得到的先验知识用于超分辨率。随后, 多伦多大学的Hertzmann A^[4]等提出了基于多尺度自动回归的图像类比算法。麻省理工学院的Freeman W T^[5]等提出了一个基于例子的方法, 利用马尔可夫网络来学习训练库中与低分辨率图像不同区域相对应的高分辨率图像的细节, 再用学习得到的关系来预测输入低分辨率图像的细节信息。Christopher M^[6]等利用一个图像块数据库来获取自然图像的空间中频段和高频段之间的关系, 并用已

被增强的超分辨率图像的一部分作为当前图像增强的训练集, 从而添加了附加的高频信息。

如上几种算法面向一般图像, 针对人脸图像, Baker S^[7]提出“人脸幻想”(face hallucination)思想, 选择人脸图像高斯金字塔的水平与垂直方向的导数以及拉普拉斯金字塔, 作为人脸图像的特征空间。通过离线学习得到映射, 该映射反映了原图在不同分辨率下的图像灰度对应关系, 以此作为基于识别的先验信息, 但得到的人脸图像在有些部位仍存在较大的噪声。清华大学的Liu^[8]等给出一个全局参数模型与局部非参数模型相结合的方法, 对“幻觉检测方法”做了改进。香港中文大学的Wang^[9]等对一系列低分辨率图像和对应高分辨率图像进行主成分分析获得人脸形状和纹理模型, 求得待重建图像对应的模型系数, 从而利用这些系数及模型重建人脸图像。浙江大学的Su^[10]等提出一种基于多尺度和多方向特征的人脸超分辨率算法, 采用可操纵金字塔学习人脸图像中的低层次局部特征的空间分布, 并结合塔状父结构和局部最优匹配算法来预测最佳先验模型。

2 基于学习的超分辨率理论基础

设 I_H 为高分辨率图像(HR), I_L 为低分辨率图像(LR), 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)。据贝叶斯估计理论, I_H 的概率, 即后验概率由式(1)生成:

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

基金项目: 国家自然科学基金资助项目(60372079)

作者简介: 郑丽贤(1982-), 女, 硕士研究生, 主研方向: 图像处理, 图像通信; 何小海, 教授、博士生导师; 吴 炜, 讲师、在职博士研究生; 杨晓敏, 博士研究生; 陈 默, 硕士研究生

收稿日期: 2007-03-30 **E-mail:** zlx.2235@163.com

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 \bar{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。

$$\bar{I}_H = \arg \max_{I_H} P(I_H/I_L) = \arg \max_{I_H} P(I_L/I_H)P(I_H) \quad (2)$$

预测 \bar{I}_H 的 3 个步骤如下:

- (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 \bar{I}_H 。

3 基于学习的超分辨率主要研究方法

3.1 基于学习的超分辨率复原框架

基于学习的超分辨率复原是使用一个图像训练集来产生学习模型, 并运用这个模型创建图像的高频细节。复原框架如图 1 所示。具体步骤为: (1) 将高分辨率图像按照退化模型进行恶化, 产生一个图像训练集。(2) 根据高低分辨率块的对对应关系, 通过一定算法进行训练学习, 获取先验知识, 建立学习模型。(3) 以输入的低分辨率块为依据, 在建立好的学习模型中搜索最匹配的高分辨率块, 最终复原出最优的超分辨率图像。

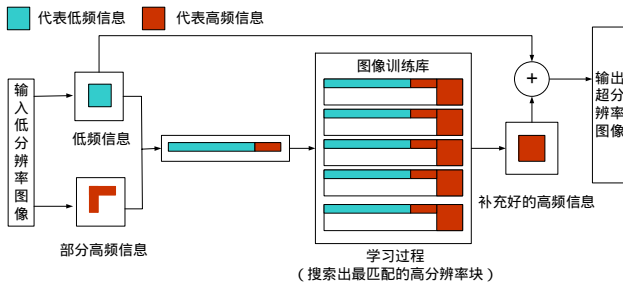


图 1 基于学习的超分辨率复原框架

3.2 学习模型的设计

如何获取先验知识, 即学习模型的设计是基于学习的超分辨率技术的核心问题, 下面分别对目前主要采用的学习模型进行介绍。

3.2.1 马尔可夫随机场模型

马尔可夫随机场模型提供了关于图像的一种统计描述, 着眼考虑每个像素关于它的一组邻近像素的条件分布, 能够有效地描述图像的局部统计特征, 因此, 可以使用马尔可夫随机场来建立图像的先验模型。马尔可夫网络模型如图 2 所示^[5], 提供了一个贝叶斯框架, 允许对图像进行最大后验概率(MAP)估计。

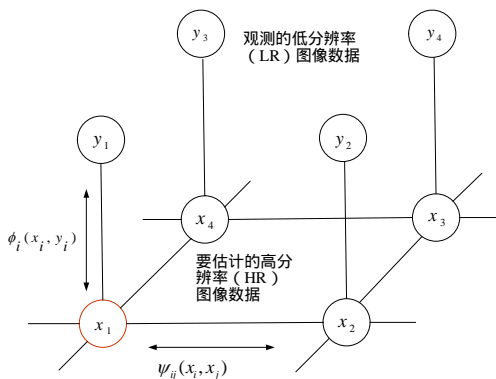


图 2 马尔可夫网络模型

在图 2 中, y 代表观测的低分辨率图像数据; x 为待估计的高分辨率图像数据; $\phi_i(x_i, y_j)$ 是对变量 x_i 的观测函数; $\psi_{ij}(x_i, x_j)$ 是 x_i 和相邻的变量 x_j 之间的相关函数。把观测图像和训练库的图像都分成小的色块, 每一个色块对应马尔可夫网络上的一个节点。

文献[5]利用参数化马尔可夫网络对高低分辨率图像建模, 从大量的例子中学习并获得网络参数, 然后利用贝叶斯信任传播机制, 为所要处理的图像找到一个后验概率的局部最大值, 即为从观测图像中获得相应的高分辨率图像而进行 MAP 估计, 从而得到合成的超分辨率图像。文献[11]将马尔可夫随机场的学习模型应用到了彩色图像的超分辨率复原中, 并取得了良好的效果。

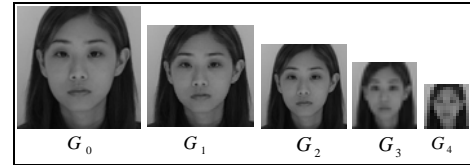
3.2.2 图像金字塔模型

使用图像金字塔来表示图像的空间尺度信息, 常用的有 高斯金字塔和拉普拉斯金字塔^[12]。设原图像为 $G_0(m, n)$, M, N 为图像的行、列数, 原图 G_0 为高斯金字塔的最底层, 则第 l 层高斯金字塔由式(3)生成:

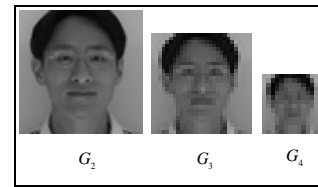
$$G_l(m, n) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_{l-1}(2i+m, 2j+n) \quad (3)$$

$$i \ll \frac{m}{2}, j \ll \frac{n}{2}, 0 \leq l \leq N$$

其中, $w(m, n)$ 是一个窗口函数, 近似为高斯低通滤波器。可见一幅图像的高斯金字塔是对下一层进行低通滤波然后作隔行和隔列的降采样而生成的。这一系列上一级比下一级缩小 4 倍的图像从低到高排列就形成了所谓图像的高斯金字塔, 如图 3 所示。超分辨率复原的目标就是要估计出金字塔的最底层 G_0 。



(a) 高分辨率图像的高斯金字塔



(b) 低分辨率图像的高斯金字塔

图 3 高斯金字塔

拉普拉斯金字塔是高斯金字塔与其上一层通过插值扩大的差值图像, 反映的是高斯金字塔两级间的信息差, 实际上就是图像的细节部分。完整的拉普拉斯金字塔定义由式(4)生成, 图 4 显示了高斯金字塔和拉普拉斯金字塔的区别。

$$L_l(I) = \begin{cases} G_l(I) - \text{EXPAND}(G_{l+1}(I)) & 0 \leq l \leq N-1 \\ G_l(I) & l = N-1 \end{cases} \quad (4)$$

使用图像金字塔表示人脸图像的空间尺度信息, 建立一个涵盖较多的人脸图像特征的人脸训练库, 然后通过学习训练库中大量高分辨人脸图像的深层次局部特征得到人脸图像的先验概率。对输入图像的任一像素, 找出训练库中每一张图像在金字塔上与其对应的像素, 及其相邻局部区域。如果训练库中包含有与输入图像的局部特征比较相似的高分辨率图像的局部特征, 那么超分辨率结果图像的局部特征就会学习训练库中这个高分辨率图像的对应特征。图像金字塔算法如图 5 所示。

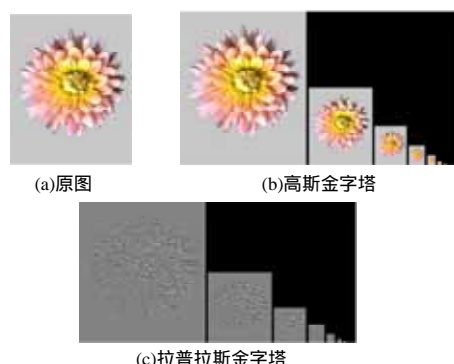


图4 高斯金字塔和拉普拉斯金字塔的区别

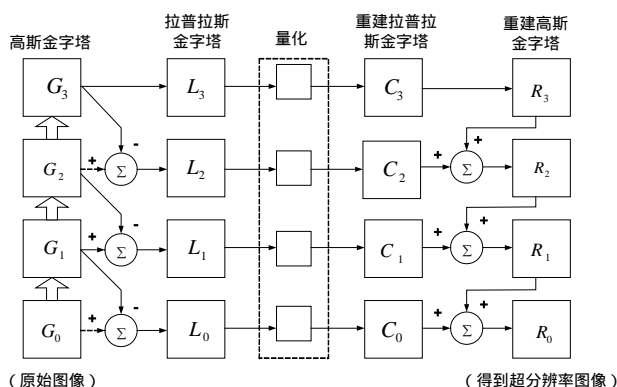


图5 图像金字塔算法

3.2.3 其他模型

除了以上两种典型学习模型外，还可以采用主成分分析(PCA)模型和神经网络模型等。

采用 PCA 模型进行学习复原的框图如图 6 所示。文献[8]中提出了一种 2 步人脸图像分辨率增强算法：第 1 步利用 PCA 模型和最大后验概率方法计算整体图像；第 2 步利用马尔可夫网络对真实高分辨率图像与整体图像之间的残差图像建模，并利用图像块采样的方法进行匹配计算。得到的结果是第 1 步得到的整体图像与第 2 步得到的残差图像之和。

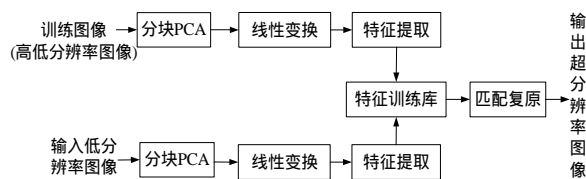


图6 采用 PCA 模型进行学习复原的框图

神经网络模型：高低分辨率图像之间有一定的对应关系，因此先验知识可以通过神经网络按照一定的学习规则训练得到。首先对高分辨率图像集做退化处理产生训练集，接着利用高低分辨率图像间的对应关系，通过神经网络进行训练学习，然后根据输入的低分辨率块搜索最匹配的高分辨率块。相比于概率统计方法，利用神经网络模型的学习速度要快得多，但训练后网络性能的好坏却受到多方面因素影响。

4 结束语

目前有关基于学习的图像超分辨率技术的研究尚处于理论探索阶段，远远没有达到实用阶段。未来研究方向主要集中在以下几方面：(1)建立更有效的学习模型，以提供更完备的先验知识；(2)完善学习算法和搜索算法，提高学习训练速

度和匹配速度，减小运算量；(3)考虑在彩色图像和三维图像的超分辨率复原研究。

基于学习的超分辨率技术虽起步较晚，但能弥补重建算法的很多不足，在许多领域都有着广泛的应用前景，例如：(1)证件识别，如身份证、驾照、护照等证件上的人脸复原等；(2)公安刑事侦察，提高事发现场、涉案嫌疑人员的照片质量；(3)海关、机场等部门的监控系统；(4)生物医学领域中生物组织连续切片图像的三维重建和三维显示工作等。随着各应用领域的发展，相关理论和技术的不断提高，必将推动基于学习的超分辨率技术的进一步发展。

参考文献

- [1] Patti A, Sezan M, Tekalp A. Super-resolution Video Reconstruction with Arbitrary Sampling Lattices and Nonzero Aperture Time[J]. IEEE Transactions on Image Processing, 1997, 6(8): 1064-1076.
- [2] Schultz R, Stevenson R. Extraction of High-resolution Frames from Video Sequences[J]. IEEE Transactions on Image Processing, 1996, 5(6): 996-1011.
- [3] Baker S, Kanade T. Limits on Super-resolution and How to Break Them[J]. Computer Vision and Pattern Recognition, 2000, 9(2): 372-379.
- [4] Hertzmann A, Jacobs C E, Oliver N, et al. Image analogies[C]// Proceedings of Computer Graphics, Annual Conferences Series, ACM SIGGRAPH. Los Angeles, California: [s. n.], 2001: 327-340.
- [5] Freeman W T, Jones T R, Pasztor E C. Example-based Superresolution[J]. IEEE Computer Graphics and Applications, 2002, 22(2): 56-65.
- [6] Christopher M. Super-resolution Enhancement of Video[C]// Proceedings of the 9th International Conference on Artificial Intelligence and Statistics. Key West, Florida: [s. n.], 2003: 410-414.
- [7] Baker S, Kanade T. Hallucinating Faces[C]// Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition. Grenoble, France: [s. n.], 2000: 83-88.
- [8] Liu Ce, Shum H Y, Zhang Changshun. A Two-step Approach to Hallucinating Faces: Global Parametric Model and Local Nonparametric Model[C]// Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai Marriott, Hawaii: [s. n.], 2001: 192-198.
- [9] Wang Xiaogang, Tang Xxiaoou. Hallucinating Face by Eigentransformation[J]. IEEE Transactions on Systems, Man and Cybernetics, Part C, 2005, 35(3): 425-434.
- [10] Su Congyong, Zhuang Yueting, Huang Li, et al. Steerable Pyramid Based Face Hallucination[J]. Pattern Recognition, 2005, 38(6): 813-824.
- [11] McAuley J J, Caetano T S, Smola A J, et al. Learning High-order MRF Priors of Color Images[C]// Proceedings of the 23rd International Conference on Machine learning. Pittsburgh, Pennsylvania: [s. n.], 2006: 617-624.
- [12] Burt P J, Adelson E H. The Laplacian Pyramid as a Compact Image Code[J]. IEEE Transactions on Communications, 1983, 31(4): 532-540.