



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2016
05
VOL.21

ISSN1006-8961
CN11-3758/TB



图像自动聚类 P574



第21卷第5期（总第241期）
2016年5月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文
编辑出版《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮 编 100101
电子信箱 jig@radi.ac.cn
电 话 010-64807995
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

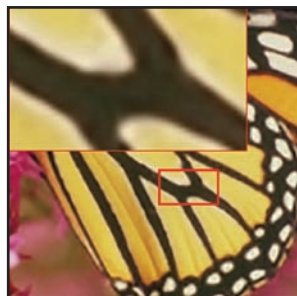
Title inscription: Song Jian Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

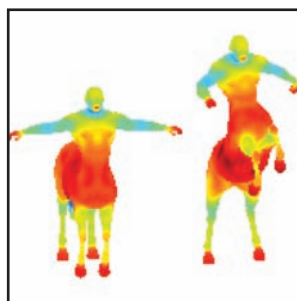
Editor-in-Chief LI Xiaowen
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

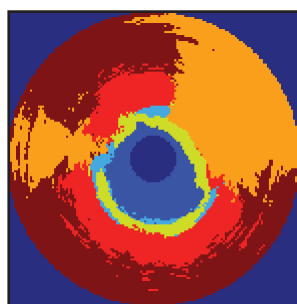
CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



利用双通道卷积神经网络的
图像超分辨率算法(第556页)



融合特征描述约束的3维
等距模型对应关系计算(第
628页)



结合先验形状信息和序贯学
习的心血管内超声外弹力膜
检测(第646页)

综述

中国图像工程: 2015

章毓晋 533

图像处理和编码

基于代价敏感Adaboost目标跟踪

薛一哲, 王拓 544

利用双通道卷积神经网络的图像超分辨率算法

徐冉, 张俊格, 黄凯奇 556

图像分析和识别

局部均匀模式描述和双加权融合的人脸识别

任福继, 李艳秋, 许良凤, 胡敏, 王晓华 565

K步稳定的鞋印花纹图像自动聚类

王新年, 舒莹莹 574

空间约束混合高斯运动目标检测

董俊宁, 杨词慧 588

图像理解和计算机视觉

分层信息融合的物体级显著性检测

李波, 金连宝, 曹俊杰, 冷成财, 卢春园, 苏志勋 595

结合区域协方差分析的图像显著性检测

张旭东, 吕言言, 缪永伟, 郝鹏翼, 陈佳舟 605

前景判别的局部模型匹配目标跟踪

刘大千, 刘万军, 费博雯 616

计算机图形学

融合特征描述约束的3维等距模型对应关系计算

杨军, 闫寒, 王茂正 628

一类加权有理插值样条曲面及局部约束控制

刘植, 肖凯, 陈晓彦, 江平, 谢进 636

医学图像处理

结合先验形状信息和序贯学习的心血管内超声外弹力膜检测

林慕丹, 杨丰, 梁淑君, 赵海升, 黄铮, 崔凯 646

遥感图像处理

局部显著特征下的光学遥感图像舷靠舰船检测

李轩, 刘云清, 卞春江, 毛博年 657

长时序高光谱图像清晰度影响因素分析

尚任, 黑保琴, 李盛阳, 覃帮勇, 张九星 665

基于变差函数的中高分辨率SAR影像农村建筑区提取

林晨曦, 周艺, 王世新, 刘文亮, 田野, 张燕楠 674

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS

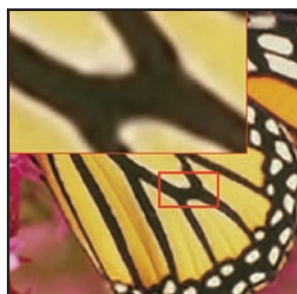
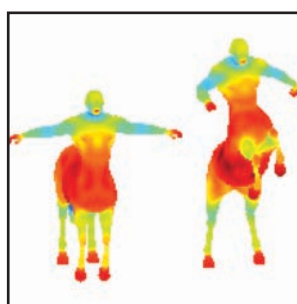
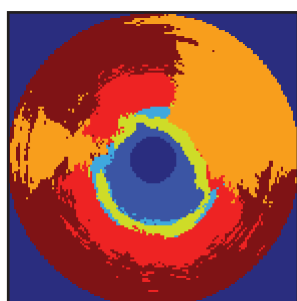


Image super-resolution using two-channel convolutional neural networks(P556)



Calculation of correspondences between three-dimensional isometric shapes with the use of a fused feature descriptor (P628)



External elastic membrane border detection based on for intravascular ultrasound images(P646)

Review

Image engineering in China: 2015	
Zhang Yujin	533

Image Processing and Coding

Object-tracking method based on improved cost-sensitive Adaboost	
Xue Yizhe, Wang Tuo	544
Image super-resolution using two-channel convolutional neural networks	
Xu Ran, Zhang Junge, Huang Kaiqi	556

Image Analysis and Recognition

Face recognition method based on local mean pattern description and double weighted decision fusion for classification	
Ren Fuji, Li Yanqiu, Xu Liangfeng, Hu Min, Wang Xiaohua	565
K-steps stabilization based on automatic clustering of shoeprint images	
Wang Xinnian, Shu Yingying	574
Moving object detection using improved Gaussian mixture models based on spatial constraint	
Dong Junning, Yang Cihui	588

Image Understanding and Computer Vision

Object level saliency detection by hierarchical information fusion	
Li Bo, Jin Lianbao, Cao Junjie, Leng Chengcai, Lu Chunyuan, Su Zhixun	595
Image saliency detection using regional covariance analysis	
Zhang Xudong, Lyu Yanyan, Miao Yongwei, Hao Pengyi, Chen Jiazhou	605
Foreground discrimination in local model-matching tracking	
Liu Daqian, Liu Wanjun, Fei Bowen	616

Computer Graphics

Calculation of correspondences between three-dimensional isometric shapes with the use of a fused feature descriptor	
Yang Jun, Yan Han, Wang Maozheng	628
Weighted rational interpolation spline surfaces and local constraint control	
Liu Zhi, Xiao Kai, Chen Xiaoyan, Jiang Ping, Xie Jin	636

Medical Image Processing

External elastic membrane border detection based on sequential learning and prior shape information for intravascular ultrasound images	
Lin Mudan, Yang Feng, Liang Shujun, Zhao Haisheng, Huang Zheng, Cui Kai	646

Remote Sensing Image Processing

Inshore ship detection method in optical remote sensing images using local salient characteristics	
Li Xuan, Liu Yunqing, Bian Chunjiang, Mao Bonian	657
Analysis of factors influencing clarity of hyper spectral images in long time series	
Shang Ren, Hei Baoqin, Li Shengyang, Qin Bangyong, Zhang Jiuxing	665
Variogram-based rural build-up area extraction from middle and high resolution SAR images	
Lin Chenxi, Zhou Yi, Wang Shixin, Liu Wenliang, Tian Ye, Zhang Yannan	674

中图分类号: TP391 文献标识码: A 文章编号: 1006-8961(2016)05-556-09

论文引用格式: Xu R, Zhang J G, Huang K Q. Image super-resolution using two-channel convolutional neural networks[J]. Journal of Image and Graphics, 2016, 21(5): 556-564. [徐冉, 张俊格, 黄凯奇. 利用双通道卷积神经网络的图像超分辨率算法[J]. 中国图象图形学报, 2016, 21(5): 556-564.] [DOI:10. 11834/jig. 20160503]

利用双通道卷积神经网络的图像超分辨率算法

徐冉, 张俊格, 黄凯奇

中国科学院自动化研究所智能感知与计算研究中心, 北京 100190

摘要: **目的** 图像超分辨率算法在实际应用中有着较为广泛的需求和研究。然而传统基于样本的超分辨率算法均使用简单的图像梯度特征表征低分辨率图像块, 这些特征难以有效地区分不同的低分辨率图像块。针对此问题, 在传统基于样本超分辨率算法的基础上, 提出双通道卷积神经网络学习低分辨率与高分辨率图像块相似度进行图像超分辨率的算法。**方法** 首先利用深度卷积神经网络学习得到有效的低分辨率与高分辨率图像块之间相似性度量, 然后根据输入低分辨率图像块与高分辨率图像块字典基元的相似度重构出对应的高分辨率图像块。**结果** 本文算法在 Set5 和 Set14 数据集上放大 3 倍情况下分别取得了平均峰值信噪比 (PSNR) 为 32.53 dB 与 29.17 dB 的效果。**结论** 本文算法从低分辨率与高分辨率图像块相似度学习角度解决图像超分辨率问题, 可以更好地保持结果图像中的边缘信息, 减弱结果中的振铃现象。本文算法可以很好地适用于自然场景图像的超分辨率增强任务。**关键词:** 图像超分辨率; Pair-wise 卷积神经网络; 双通道卷积神经网络; 图像块相似度学习

Image super-resolution using two-channel convolutional neural networks

Xu Ran, Zhang Junge, Huang Kaiqi

Center for Research on Intelligent Perception and Computing (CRIPAC), Institute of Automation,
Chinese Academy of Sciences, Beijing 100190, China

Abstract: **Objective** All traditional example-based super-resolution methods adopt image-gradient features for low-resolution images and thus, these methods are unable to characterize the low-resolution space satisfactorily. To address this issue, this paper proposes a novel unified framework for image super-resolution that effectively combines example-based method with deep learning models. **Method** The proposed method consists of three main stages: low- and high-resolution similarity-learning, high-resolution patch-dictionary-learning, and high-resolution patch-generating stages. At the first stage, two different convolutional neural networks are proposed for learning a novel similarity metric between high- and low-resolution image patches. At the second stage, the high-resolution patch dictionaries are learned from training sets. At the last stage, the high-resolution patches are generated based on learned similarities between the input low-resolution patch and the atoms in the high-resolution patch dictionary. **Result** Experimental results on several commonly adopted datasets show that the proposed two-channel model quantitatively and qualitatively achieves improved performance compared with other methods. **Conclusion** The proposed two-channel model can preserve more detailed information and reduce ringing artifacts in the resulting images. **Key words:** image super resolution; Pair-wise convolutional neural networks; two-channel convolutional neural networks; patch similarity learning

收稿日期: 2015-11-26; 修回日期: 2016-01-11

基金项目: 国家重点基础研究发展计划 (973) 基金项目 (2012CB316302); 国家自然科学基金项目 (61322209, 61175007)

第一作者简介: 徐冉 (1992—), 男, 2013 年于中国科学院自动化研究所模式识别实验室攻读硕士学位, 主要研究方向为图像处理、图像超分辨率。E-mail: ran.xu@nlpr.ia.ac.cn

Supported by: National Basic Research Program of China (2012CB316302); National Natural Science Foundation of China (61322209, 61175007)

0 引言

图像超分辨率增强是指根据输入的一幅或多幅低分辨率图像重构出一幅具有更高分辨率图像的过程。由于高分辨率图像中包含有更多的细节信息, 图像超分辨率增强技术在视频监控、医学成像、遥感卫星成像等各个领域有着广泛的应用。然而由于根据有限的输入信息重构出具有更多信息图像的过程是一个典型的病态反过程, 图像超分辨率增强始终是一个极具挑战的任务。

现有的超分辨率算法可以分为三大类: 基于插值^[1-4]、基于重建^[5-8]以及基于样本^[9-11]的方法。其中基于样本的方法因其有着较好的实现效果是近些年来研究的热点。这类方法的主要思想是从一个额外的样本集中根据输入低分辨率图像的特征搜索产生对应的高分辨率图像。低分辨率图像与高分辨率图像本质上处于两个不同的数据空间, 而超分辨率增强任务实质上是在提供低分辨率空间数据的前提下尽可能准确地产生出其在高分辨率空间中的对应图像。因此如何准确有效地对这两个空间之间存在的复杂关系进行建模是超分辨率增强任务的关键所在。传统的基于样本学习的方法主要从两种思路解决该问题: 一是直接利用样本学习到两个空间之间的映射关系, 二是将两个空间中的数据通过不同的映射函数映射到同一个特征空间。Hong 等人^[9]提出邻域嵌入 (neighbor embedding) 算法利用局部线性嵌入理论来解决超分辨率问题。该算法假设低分辨率图像块与高分辨率图像块形成的流形空间具有相似的局部几何形状, 对于输入的低分辨率图像块, 根据其在 K 个最近邻低分辨率图像块上的重构系数产生出对应的高分辨率图像块。为了解决邻域嵌入算法中存在的字典尺寸过大问题, Yang 等人^[10]提出了基于稀疏表示的超分辨率增强算法, 该算法首先通过训练样本学习得到一对低分辨率与高分辨率图像块字典, 接着对于输入的低分辨率图像块, 首先求得到其在低分辨率图像块字典下的稀疏表示系数, 然后利用该系数直接与高分辨率图像块字典相乘产生对应的高分辨率图像块。该算法中学习到的低分辨率与高分辨率图像块字典所起到的作用即是各自空间投影到同一特征空间。为了解决图像超分辨率算法时间效率低的问题, Radu 等人^[11]提出

了 ANR 方法。该方法首先通过对字典中的基元进行聚类, 根据不同邻域样本计算每个基元对应的低分辨率图像块与高分辨率图像块之间的映射函数。测试时对于输入的低分辨率图像块, 选取与其最为相似的基元对应的映射函数产生高分辨率图像块。该算法由于事先计算并存储了每个基元的映射矩阵, 在测试时可以大大提升算法的时间效率并取得了较好的处理结果。

然而上述传统的基于样本超分辨率算法针对低分辨率图像块使用的都是简单的图像特征, 如图像的一阶、二阶梯度等。这些简单的特征不足以充分地表征低分辨率图像所处空间, 因此传统方法难以重构出准确的高分辨率图像。在图像分类识别等相关领域, 如何利用图像特征充分表示图像的数据结构是关键所在^[12]。最近深度学习工具已经在计算机视觉领域中的许多任务中取得了巨大的成功, 深度学习网络学习出的特征在大量的任务中被证实比传统的手工设计的特征具有更强的表征能力^[13]。因此本文提出了一种结合深度学习与传统基于样本方法的超分辨率算法。该算法首先通过深度卷积神经网络学习得到低分辨率与高分辨率图像块之间的相似性度量, 接着计算输入的低分辨率图像块与高分辨率图像块字典中每个基元之间相似度, 最后根据计算出的相似度重构出对应的高分辨率图像块。近期 Chao 等人^[15]已经提出了使用卷积神经网络进行超分辨率增强的算法 (SRCNN) 并且取得了较好的效果。该算法使用卷积神经网络直接学习低分辨率图像到高分辨率图像的映射关系。同时 Cui 等人^[16]提出了 DNC 算法, 该算法首先采用非局部自相似搜索方法增加低分图像块中的高频信息, 接着利用自动编码器产生高分辨率图像块。上述两种算法都是通过深度学习工具学习低分辨率与高分辨率两个空间之间的映射关系, 与之不同的是本文提出的算法中, 深度学习工具仅是被用来判断输入的低分辨率与高分辨率图像块是否对应。与直接学习映射关系的任务相比, 对输入图像块对进行相似度计算的任务因其输入的信息更多, 因此相对而言较为简单。实验结果表明, 本文提出的双通道模型超分辨率算法在结果图像中能够保留更多高频细节信息, 同时也可以很好地减弱振铃现象的发生。实验结果 PSNR 指标也证明了本文双通道模型算法的有效性。

1 本文提出的算法

本文算法从低分辨率图像块与高分率图像块匹配角度出发解决图像超分辨率任务。图1显示了本文算法的整体框架。该算法包含有3个主要阶段,第1阶段为低分辨率图像块与高分率图像块相似度计算过程,在这个阶段中本文尝试了两种不同结构的深度卷积神经网络模型来学习低分辨率图像块与高分率图像块之间的相似度。第2阶段为高分率图像块字典学习过程,在这个阶段,分别使用了K均值算法与K-SVD(K奇异值分解)算法学习产生高分率图像块字典。第3阶段为高分率图像块生成阶段,在这个阶段中本文提出了K近邻相似度加权高分率图像块生成算法。

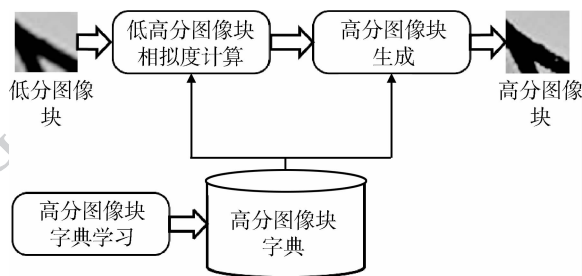


图1 本文算法整体框架

Fig. 1 The framework of the proposed method

1.1 低分、高分图像块相似度计算

传统计算两个图像块之间的相似度的主要解决思路为首先针对每个图像块分别提取特征,然后根

据特征之间的欧氏距离判定两个图像块是否对应。传统思路中特征表达的好坏严重影响着最终相似度计算的准确性。在图像超分辨率任务中,由于低分辨率图像块与高分率图像块所处特征空间不同,因此不适合对其采用直接提取手工特征计算相似度。然而利用深度学习工具计算图像块的相似度有着多种选择。本文从两种不同的思路出发分别提出了Pair-wise模型与双通道模型用以学习高分率图像块与低分辨率图像块之间的相似度。

1.1.1 Pair-wise 模型

Pair-wise思路首先应用在图像检索领域,通过训练阶段使得相同类图像特征之间的欧氏距离尽可能小而不同类图像特征之间的欧氏距离尽可能的大,最终使得学习出的图像特征表示具有较好的类别判别能力^[17]。图2显示了该模型的网络结构。模型具有两个独立的分支,每个分支具有相同的网络结构,均由卷积层、ReLU层、Pooling层以及全连接层构成。训练与测试时每个分支针对输入中的一个图像块进行处理并且每个分支的最后全连接层64维的输出即是对应输入图像块的特征。训练时每个分支的输出连同训练样本的标签一同输入到对比损失函数层(contrastive-loss)进行损失计算与反传。对比损失函数的具体计算为

$$E = \frac{1}{2N} \sum_{n=1}^N y_n d_n^2 + (1-y_n) \max(m - d_n, 0)^2 \quad (1)$$

$$d_n = \|a_n - b_n\|_2$$

式中, y_n 为第 n 个输入训练样本是否对应的标签, y_n 为1表示输入高分率与低分辨率图像块相互对应,

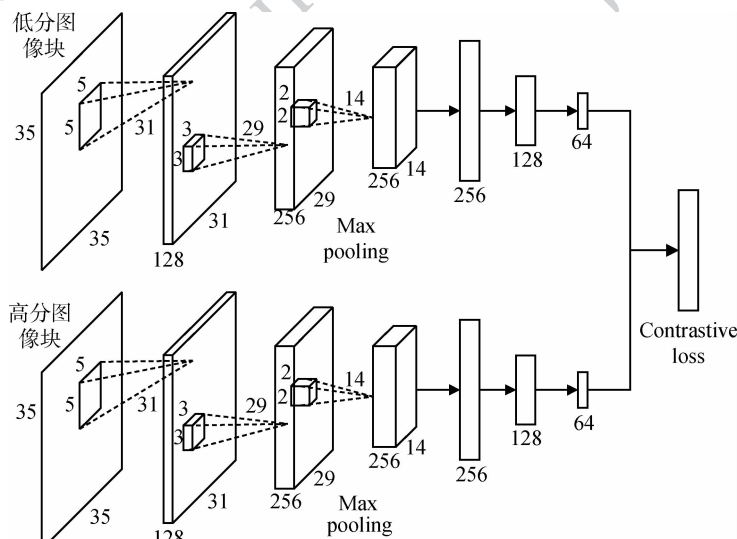


图2 Pair-wise模型网络结构

Fig. 2 The architecture of the Pair-wise model

0 则相反。 m 为对比损失函数的间隔参数, 在对比损失函数中, 间隔参数的意义为控制同类样本特征之间的欧氏距离范围。同类样本特征之间的欧氏距离应该小于间隔参数, 同时不同类样本特征之间的欧氏距离应该大于间隔参数。式(1)中 E 为所要优化的目标损失。 d_n 为第 n 个训练样本中高低分辨率图像块经过对应分支生成的特征之间的欧氏距离。通过训练最小化该损失函数可以使得对应的高分辨率与低分辨率图像块特征之间的欧氏距离尽可能小而不对应图像块特征之间的欧氏距离尽可能大。训练好的模型中每个分支均可独立地接受一个图像块作为输入并产生其对应的特征。

根据该模型中两个分支的参数是否共享, 该模型可以划分为 Siamese 模型与 Pseudo-siamese 模型^[18]。Siamese 模型中两个分支的参数是共享的, 意味着针对高低分辨率图像块使用相同的参数提取特征。而在 Pseudo-siamese 模型中两个分支的参数是不共享的, 网络训练时各自独立更新, 意味着针对高低分辨率图像块使用不同的参数提取特征。同时也意味着 Pseudo-siamese 模型中的参数数量是 Siamese 模型的两倍, 更多的参数可以为模型提供更多的复杂性和灵活性。

1.1.2 双通道模型

不同于上述 Pair-wise 模型根据传统思路出发解决高低分辨率图像块相似度学习问题, 双通道模型则是直接将图像块相似度计算问题转换成二分类问题。图3显示了双通道模型的网络结构, 该模型与通常的分类网络类似包含有一系列的卷积层、ReLU 层、Pooling 层以及全连接层。输入数据中的两个通道分别存储了低分辨率图像块与高分辨率图像块。两个通道的数据一同输入到网络中进行处理。不同于 Pair-wise 模型, 该模型中没有针对每个图像块的特征之分, 网络从开始时就一同利用两个图像块提供的信息进行判别, 模型最终的全连接层 2 维的输出经过 Soft-max 函数处理则为输入高低分辨率图像块之间的相似度。同 Pair-wise 模型一样, 训练阶段中, 训练数据的标签为 1 则表明该训练数据中的低分辨率图像块与高分辨率图像块相互对应, 标签为 0 则相反。

1.2 高分辨率图像块字典学习

通过上阶段中的模型训练, 对于输入的高低分

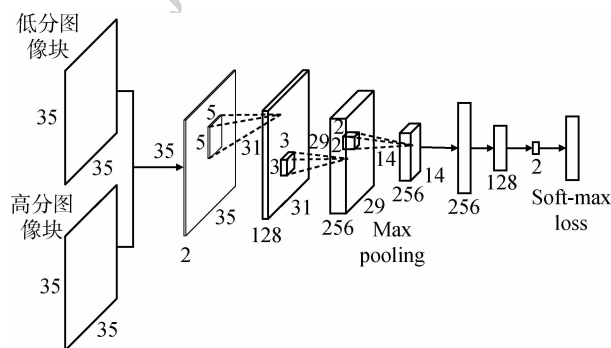


图3 双通道模型网络结构

Fig. 3 The architecture of the two-channel model

辨率图像块已经可以利用训练好的模型得到他们之间较为准确的相似度。本阶段主要解决候选高分辨率图像块产生问题。针对上阶段中两种不同的网络结构, 本阶段分别使用 K 均值聚类与 K-SVD 算法产生高分辨率图像块字典。

针对 Pair-wise 模型, 由于模型中的每个分支可以独立的接受图像块作为输入并生成其对应的特征, 本文采用 K 均值聚类算法对训练样本中的高分辨率图像块进行聚类, 每个聚类中心即被保存为高分辨率图像块字典中 D_{hp} 的基元 d_i 。接着将每个基元 d_i 输入到 Pair-wise 模型中对应的分支产生其对应的特征 d'_i 并将其保存为高分辨率特征字典 DF_h 。需要指出的是在 Pseudo-siamese 模型中, 由于两个分支的参数不共享, 使用训练时接受高分辨率图像块作为输入的分支对高分辨率图像块字典中的基元进行特征提取工作。

针对双通道模型, 使用稀疏表示的算法产生字典。对于训练样本中的高分辨率图像块 Y_h , 首先对其进行均值归一化。接着使用 K-SVD 字典学习算法^[19]产生高分辨率图像块字典, 即

$$\begin{aligned} (D_h, \partial) = \arg \min_{D_h, \partial} \|Y_h - D_h \partial\|^2 \\ \text{s. t. } \|\partial\|_0 \leq L \end{aligned} \quad (2)$$

式中, D_h 为学习到的高分辨率图像块字典, ∂ 为训练样本中的高分辨率图像块在该字典下的稀疏表示。 L 为表示系数的最大稀疏性约束。传统的基于样本超分辨率算法中, 高分辨率图像块字典与低分辨率图像块字典的学习过程是整个算法的关键, 而本文算法由于不再需要低分辨率图像块字典, 因此降低了字典学习算法的好坏对超分辨率算法最终结果的影响。

1.3 高分辨率图像块生成

在获得了低分辨率与高分辨率图像块相似度计算模型以及高分辨率图像块字典之后,本阶段主要根据输入低分辨率图像块与高分辨率图像块字典基元的相似度产生其对应的高分辨率图像块,提出了K近邻相似度加权高分辨率图像块生成算法。

针对 Pair-wise 模型,对于输入的低分辨率图像块 X_j ,首先将其输入到 Pair-wise 模型中对应的分支提取出其对应的特征 f_j ,同理,在 Pseudo-siamese 模型的情况下使用训练时接受低分辨率图像块作为输入的分支进行处理。接着计算其特征 f_j 与高分辨率特征字典 DF_h 中每个基元 df_i 的欧氏距离,选取与其最为接近的 K 个基元重构出其对应的高分辨率图像块

$$Y_j = \frac{1}{Z} \sum_{i=1}^K \exp\left(\frac{1}{\|f_j - df_i\|_2^2}\right) d_i \quad (3)$$

式中, Y_j 为生成的高分辨率图像块, f_j 为输入低分辨率图像块经过 Pair-wise 模型产生的特征, df_i 为高分辨率特征字典 DF_h 中的第 i 个基元, d_i 则为对应的高分辨率图像块字典 D_{hp} 中的第 i 个基元, Z 为归一化常数。

针对双通道模型,由于该模型的输入要求为高低分辨率图像块对,因此对于输入的低分辨率图像块 X_j ,首先将其均值补偿到高分辨率图像块字典 D_m 中的每个基元 d_i ,并将其与每个基元一一组合输入到训练好的双通道模型中。模型最后的全连接层输出经过 Soft-max 函数处理即可作为该低分辨率图像块 X_j 与对应字典基元 d_i 的相似度 $\text{sim}(X_j, d_i)$ 。将其与每个基元的相似度作为权重选取与其最为相似的 K 个基元进行加权平均即可产生对应的高分辨率图像块

$$Y_j = \frac{1}{Z} \sum_{i=1}^K \exp(\text{sim}(X_j, d_i)) d_i \quad (4)$$

$$\text{sim}(X_j, d_i) = \frac{\exp(f_1)}{\exp(f_1) + \exp(f_0)}$$

式中, d_i 为高分辨率图像块字典 D_m 中第 i 个与输入低分辨率图像块 X_j 最为相似的基元, $\text{sim}(X_j, d_i)$ 为输入低分辨率图像块 X_j 与基元 d_i 的相似度, f_0 与 f_1 分别表示将 X_j 与 d_i 组合输入到双通道模型中时网络最后一层全连接层输出的 2 维特征中的第 0 维与第 1 维数据。

2 实验与分析

2.1 实验设置

出于实验公平性的考虑,使用超分辨率增强领域通用的包含有 91 幅自然图像的数据集^[10,15]作为模型和高分辨率图像块字典学习的训练数据。同传统方法一样,数据集中的真实高分辨率图像经过人为模拟的降采样、高斯模糊过程产生低分辨率图像。本文人工模拟产生低分辨率图像的过程中使用与对比算法中一样的处理参数。与传统的基于样本的方法一样,将这些人工模拟的低分辨率图像经过双三次插值缩放到原有尺寸。接着从这些低分辨率图像与原始高分辨率图像中以滑动窗口的方式提取图像块。对应位置处的高低分辨率图像块组合构成正样本,不对应位置处的高低分辨率图像块则构成负样本。

考虑到传统的基于图像块的超分辨率算法中对于图像块尺寸的要求不能太大,然而太小尺寸的图像块难以提供足够的信息供深度卷积神经网络进行学习。为了平衡两者之间的矛盾,本文算法采用了对小尺寸的图像块进行重复堆叠生成较大尺寸图像块的方法。模型训练时采用尺寸为 35×35 像素的图像块,高分辨率图像块字典学习时则采用尺寸为 7×7 像素的图像块。测试时从输入的低分辨率图像中提取出尺寸为 7×7 像素图像块,并将其重复堆叠生成 35×35 像素的图像块输入到模型中。与 SRCNN^[15]方法相同,针对 2 倍与 3 倍分别训练了不同的高低分辨率图像块相似度计算模型。

深度学习模型的效果很大程度上依赖于训练数据的有效性,尤其对需要输入组合类型数据的深度学习模型而言,负样本的选择至关重要。本文首先针对每一个高分辨率图像块从训练数据集中随机选取出一系列与之不对应的低分辨率图像块组合成为负样本。在使用这些初选的样本进行训练之后,针对每个高分辨率图像块从训练数据集中选取出其最为相似但并不对应的低分辨率图像块组合形成难负(hard-negative)样本。接着用新产生的难负样本对已经训练好的模型进行调优(fine-tune)。

Pair-wise 模型与双通道模型的具体参数根据实践经验设置如图 2、图 3 所示。Pair-wise 模型中对比损失函数的间隔参数 m 设置为 1。本文使用 Caffe 深度学习工具箱^[20]进行模型的搭建和训练。根

据实践经验模型训练的初始学习率设为 0.001, 动量参数 (momentum) 设为 0.9。字典学习中表示系数的最大稀疏约束 L 设为 10, 同时式 (3) (4) 中的参数 K 也设为 10。高分辨率图像块字典的维数为 1 024。整体训练集包含有 800 000 对样本, 正负样本比例为 1:2。对于三通道的彩色图像, 实验中遵循传统算法的处理方法, 先将其转换成 YUV 图像, 只对 Y 通道图像进行超分处理, 剩余 U、V 通道图像使用双三次插值方法缩放到目标尺寸。需要指出的是本文算法可以轻易地扩展成对多通道图像进行处理。

2.2 实验结果与分析

为了验证本文算法的有效性, 实验中使用超分辨率增强领域通用的测试数据集 Set5^[15]、Set14^[15]、BSD200^[21] 进行测试。实验中本文算法与双三次插值 (Bicubic)、Yang^[10]、ANR^[11]、SRCNN^[15] 4 种方法进行了对比。本文算法按照不同的网络模型分别命名为 Siamese 方法、Pseudo-Siamese 方法以及双通道方法。图 4 显示了不同算法在 butterfly 图像上放大 3 倍的处理结果。从图 4 放大区域可以看出, 其他算法产生的结果图像中纹理区域的边缘部分有较明显的振铃现象, 本文提出的双通道方法的结果图像中振铃现象有明显的减弱。同时双通道方法的结果

图像中的黑色区域也更为纯净噪声较少。图 5 显示的是不同算法在 bird 图像上放大 3 倍的处理结果。放大区域中的黑色条状纹理显示出双通道方法可以产生更为锐利的边缘与真实高分辨率图像最为接近。表 1 为 Set5 数据集上本文算法与不同对比算法的实验结果 PSNR 值指标。表 2 则列出了不同算法在 Set4 和 BSD200 数据集上的平均 PSNR 指标值。由上述两表可知, 双通道方法取得了比传统基于样本方法与 SRCNN^[15] 算法更好的增强效果。

在本文提出的两种不同结构的网络模型中, 双通道模型的方法显示出了较 Pair-wise 模型方法更好的实现效果。表明从二分类问题角度解决高低分辨率图像块相似度学习问题更为有效。这是因为 Pair-wise 模型中两个分支之间相互独立, 没有信息的交流和互动, 每个分支只能利用单个输入图像块单独提供的信息。然而在双通道模型中, 由于网络从一开始就利用了两个图像块的所有信息, 可以更好的利用两个图像块之间的差异信息来判别他们是否对应。同时在 Pair-wise 模型中, Pseudo-Siamese 方法取得了较 Siamese 方法更好的效果, 这表明模型中两个分支的参数独立更新可以更好地解决高低分辨率图像块相似度学习问题。这是因为如前所述, 高低分辨率图像块处在两个不同的特征空间, 而

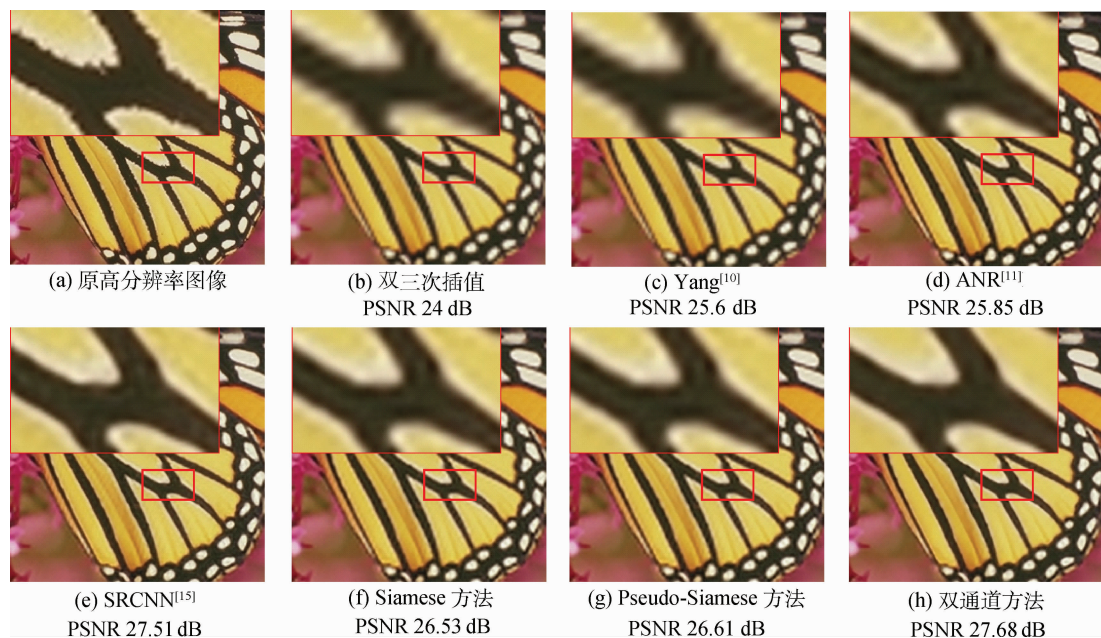


图4 butterfly 图像超分辨率增强放大 3 倍结果

Fig. 4 Result of super resolution on butterfly image with 3 upscaling factor ((a) original high resolution image; (b) Bicubic result; (c) Yang^[10] result; (d) ANR^[11] result; (e) SRCNN^[15] result; (f) Siamese result; (g) Pseudo-Siamese result; (h) two-channel result)

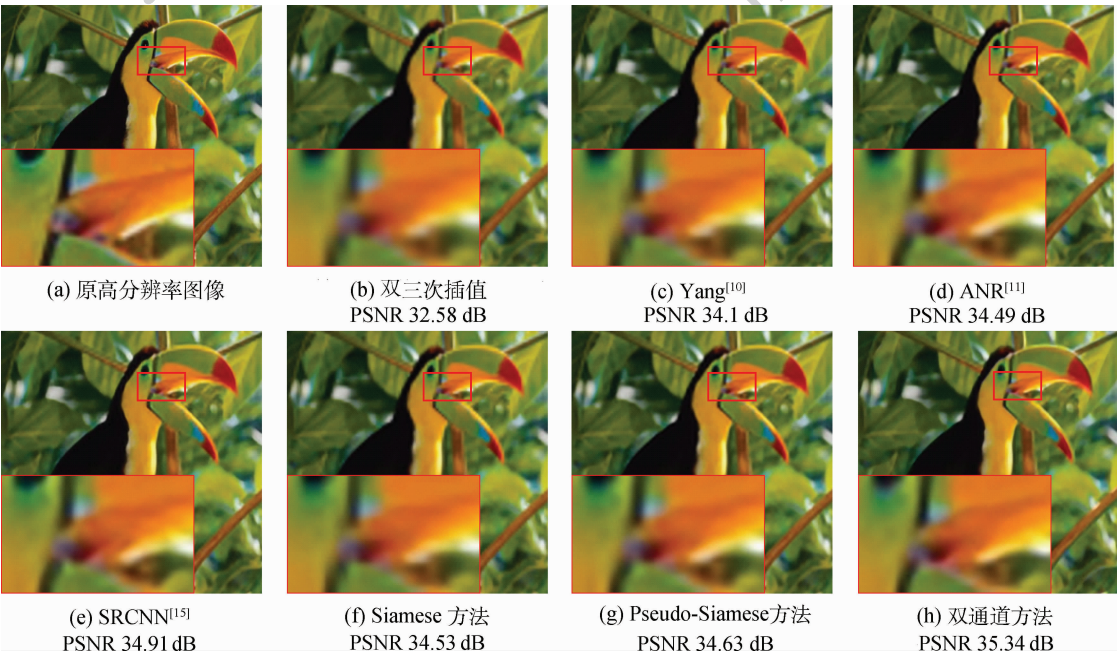


图 5 bird 图像超分辨率增强放大 3 倍结果

Fig. 5 Result of super resolution on bird image with 3 upscaling factor ((a) original high resolution image; (b) Bicubic result; (c) Yang^[10] result; (d) ANR^[11] result; (e) SRCNN^[15] result; (f) Siamese result; (g) Pseudo-Siamese result; (h) two-channel result)

表 1 Set5 数据集上不同算法实验结果 PSNR
Table 1 PSNR results of different methods on Set5 dataset

/dB								
图像	倍数	Bicubic	Yang ^[10]	ANR ^[11]	SRCNN ^[15]	Siamese	Pseudo-Siamese	双通道方法
baby	2	37.05	37.34	38.29	38.28	37.96	38.01	38.41
bird	2	36.68	37.26	39.96	40.64	39.68	39.79	40.71
butterfly	2	27.45	27.93	30.25	32.2	29.72	29.86	32.42
head	2	34.85	35.17	35.6	35.64	35.36	35.51	35.71
woman	2	32.14	33.03	34.48	34.94	34.57	34.69	35.09
baby	3	33.91	34.3	35.06	35.01	34.48	34.74	35.15
bird	3	32.58	34.1	34.49	34.91	34.53	34.63	35.34
butterfly	3	24	25.6	25.85	27.51	26.53	26.61	27.68
head	3	32.89	33.2	33.63	33.51	33.48	33.53	33.67
woman	3	28.14	29.9	30.33	30.92	30.2	30.33	30.83
平均	2	33.63	34.15	35.72	36.34	35.46	35.57	36.47
	3	30.30	31.42	31.87	32.37	31.84	31.97	32.53

表 2 Set14 和 BSD200 数据集上不同算法实验结果平均 PSNR
Table 2 Average PSNR results of different methods on Set14 and BSD200 datasets

/dB								
数据集	倍数	Bicubic	Yang ^[10]	ANR ^[11]	SRCNN ^[15]	Siamese	Pseudo-Siamese	双通道方法
Set14	2	30.23	30.68	32.28	32.31	31.89	32.05	32.41
	3	27.53	28.11	29.15	29.01	28.59	28.77	29.17
BSD200	2	28.39	29.03	30.14	30.16	29.78	29.89	30.23
	3	25.89	26.51	27.05	26.93	26.69	26.80	27.08

Pair-wise 模型中两条分支的主要作用是将高低分辨率图像块映射到同一特征空间中, 因此两个分支若使用相同的参数则很难将不同空间的数据映射到同一个特征空间。

与传统基于样本的方法 ANR^[11] 与 Yang^[10] 相比, 双通道方法的效果提升明显, 表明深度卷积神经网络可以学到有效的特征用以计算高低分辨率图像块相似度。同时与 SRCNN^[15] 算法相比, 双通道方法也取得了效果上的提升, 表明通过双通道模型学习低分辨率与高分辨率图像块之间相似度来解决超分辨率增强任务较直接利用卷积神经网络学习两者之间映射关系更为有效。时间效率上通过 GPU 并行, 双通道方法可以达到与 Yang^[10] 方法相近的速度。

3 结 论

本文在传统基于样本的超分辨率算法的基础上, 引入了深度学习工具进行高低分辨率图像块相似度学习。提出了新的利用深度学习工具进行图像超分辨率的框架。大量的实验结果表明本文提出的双通道模型算法处理后的结果图像从主客观评价上均好于已有的算法。将来的研究重点将包括如何改善现有高分辨率图像块字典学习方法以及提高算法的时间效率。

参考文献 (References)

- [1] Tsai R, Huang T. Multiple frame image restoration and registration[C]//Advances in Computer Vision and Image Processing. Greenwich, CT: JAI Press Inc., 1984: 317-339.
- [2] Komatsu T, Aizawa K, Igarashi T, et al. Signal-processing based method for acquiring very high resolution images with multiple cameras and its theoretical analysis[J]. IEEE Proceedings of Communications, Speech and Vision, 1993, 140(1): 19-24. [DOI: 10.1049/ip-i-2.1993.0005]
- [3] Ur H, Gross D. Improved resolution from subpixel shifted pictures[J]. CVGIP: Graphical Models and Image Processing, 1992, 54(2): 181-186. [DOI: 10.1016/1049-9652(92)90065-6]
- [4] Komatsu T, Igarashi T, Aizawa K, et al. Very high resolution imaging scheme with multiple different-aperture cameras[J]. Signal Processing: Image Communication, 1993, 5(5-6): 511-526. [DOI: 10.1016/0923-5965(95)90014-K]
- [5] Tappen M F, Russell B C, Freeman W T. Exploiting the sparse derivative prior for super-resolution and image demosaicing[C]//Proceedings of the IEEE Workshop on Statistical and Computational Theories of Vision. Washington, DC: IEEE, 2003.
- [6] 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. [DOI: 10.1109/TPAMI.2010.25]
- [7] Dai S Y, Han M, Xu W, et al. Soft edge smoothness prior for alpha channel super resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN: IEEE, 2007: 1-8. [DOI: 10.1109/CVPR.2007.383028]
- [8] Dai S Y, Han M, Xu W, et al. Softcuts: a soft edge smoothness prior for color image super-resolution[J]. IEEE Transactions on Image Processing, 2009, 18(5): 969-981. [DOI: 10.1109/TIP.2009.2012908]
- [9] Chang H, Yeung D Y, Xiong Y M. Super-resolution through neighbor embedding[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE, 2004: 1. [DOI: 10.1109/CVPR.2004.1315043]
- [10] Yang J C, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873. [DOI: 10.1109/TIP.2010.2050625]
- [11] Timofte R, De V, Van Gool L. Anchored neighborhood regression for fast example-based super-resolution[C]//Proceedings of the IEEE International Conference on Computer Vision. Sydney, NSW: IEEE, 2013: 1920-1927. [DOI: 10.1109/ICCV.2013.241]
- [12] Li Z C, Liu J, Tang J H, et al. Robust structured subspace learning for data representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(10): 2085-2098. [DOI: 10.1109/TPAMI.2015.2400461]
- [13] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the Advances in Neural Information Processing Systems. South Lake Tahoe, Nevada: Curran Associates Inc., 2012: 1097-1105.
- [14] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA: IEEE, 2015: 1-9. [DOI: 10.1109/CVPR.2015.7298594]
- [15] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution[C]//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014: 184-199. [DOI: 10.1007/978-3-319-10593-2]
- [16] Cui Z, Chang H, Shan S G, et al. Deep network cascade for image super-resolution[C]//Proceedings of the 13th European Con-

- ference on Computer Vision. Zurich, Switzerland: Springer, 2014: 49-64. [DOI: 10.1007/978-3-319-10602-1_4]
- [17] Li Z C, Tang J H. Weakly supervised deep metric learning for community-contributed image retrieval[J]. IEEE Transactions on Multimedia, 2015, 17 (11): 1989-1999. [DOI: 10.1109/TMM.2015.2477035]
- [18] Zagoruyko S, Komodakis N. Learning to compare image patches via convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA: IEEE, 2015: 4353-4361. [DOI: 10.1109/CVPR.2015.7299064]
- [19] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[C]//Proceedings of the 7th International Conference on Curves and Surfaces. Berlin Heidelberg: Springer, 2012: 711-730. [DOI: 10.1007/978-3-642-27413-8_47]
- [20] Jia Y Q, Shelhamer E, Donahue J, et al. Caffe: convolutional architecture for fast feature embedding[C]//Proceedings of the 22nd ACM International Conference on Multimedia. Orlando, FL: ACM, 2014: 675-678. [DOI: 10.1145/2647868.2654889]
- [21] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, BC: IEEE, 2001, 2: 416-423. [DOI: 10.1109/ICCV.2001.937655]