图文信息技术

基于拉普拉斯金字塔的图像压缩与重构研究

常敏[a,b]，陈果[a,b]，韩帅[a]
（上海理工大学 a.光电信息与计算机工程学院 b.上海市现代光学系统重点实验室，上海 200093）

摘要：目的 研究利用深度学习辅以拉普拉斯金字塔来完成图像压缩与重构。**方法** 利用卷积神经网络提取图像的主要特征，利用双三线性插值法来减少特征尺寸，使用拉普拉斯金字塔来构建分层体系，从而逐步地减少图像大小以达到压缩的目的。在重构端上，对此系统则进行卷积操作，并采用上采样过程，进行图像的恢复重构过程，得到重构图。**结果** 采用来自法国贝尔实验室的 set 5 与 set 14 数据集进行验证，使用 2 层金字塔即在 16 倍的高倍率压缩下进行实验结果验证，结果表明在主观评价上使用深度学习的方法在清晰度和还原度上要优于 PCA，DCT 和 SVD，同时在客观评价上文中方法取得了标准差（52.73）与信息熵（7.44）的最好结果，高于 PCA 的 49.70 与 7.38。SVD 变换法与 DCT 变换法，在标准差上只有 48.69 和 49.02，远不如文中方法，同时图片的信息熵只有 7.34 与 7.35，低于文中的 7.44。**结论** 利用拉普拉斯金字塔结构来设计卷积神经网络结构来完成图像压缩与重构取得了不错的效果。

关键词：深度学习；图像压缩；图像重构；卷积神经网络；拉普拉斯金字塔
中图分类号：TN911 **文献标识码：**A **文章编号：**1001-3563(2020)15-0239-06
DOI：10.19554/j.cnki.1001-3563.2020.15.036

Image Compression and Reconstruction Based on Laplacian Pyramid

CHANG Min[a,b], *CHEN Guo*[a,b], *HAN Shuai*[a]
(a.School of Optoelectronic Information and Computer Engineering b.Shanghai Key Laboratory of Modern Optical Systems, University of Shanghai for Science and Technology, Shanghai 200093, China)

ABSTRACT: The paper aims to complete image compression and reconstruction through deep learning supplemented by Laplacian pyramid. Main features of the image were extracted with the convolutional neural network. The feature size was reduced by the bicubic linear interpolation. The hierarchy system was constructed by Laplacian pyramid to gradually reduce the image size and achieve image compression. On the reconstruction end, the corresponding convolution and up-sampling process was performed on the system; and the image reconstruction and reconstruction process was performed to obtain a reconstructed graph. Set 5 and set 14 from Bell Laboratories in France were used for verification. The experimental results were verified by the two-layer pyramid, which meant that the experimental results were verified at the 16 times of high-rate compression. The results showed that the method of deep learning was superior to PCA, DCT and SVD in terms of clarity and reduction in subjective evaluation, and the best results of standard deviation (52.73) and information entropy (7.44) were obtained in objective evaluation, which were higher than 49.70 and 7.38 of PCA. The standard deviations of SVD transform and DCT transform were only 48.69 and 49.02, which were far worse than the methods in this paper. Meanwhile, the information entropy of images was only 7.34 and 7.35, which was lower than 7.44 in this paper. Design convolutional neural network structure by Laplacian pyramid to complete image compression and re-

收稿日期：2019-12-07
基金项目：国家重大仪器专项（2016YFF0101400）
作者简介：常敏（1978—），女，博士，上海理工大学副教授，主要研究方向为光电检测仪器、图像处理。

construction achieves good results.

KEY WORDS: deep learning; image compression; image reconstruction; convolutional neural network

随着科学技术的进步，图像的分辨率得到了巨大的提升，从 720 P 到 1080 P 再到现在的 2 K 与 4 K。在图像的分辨率得到提升的同时，也为图像的存储与传输带来了前所未有的挑战，同时也对于图像压缩提出了不小的要求。深度学习在近几年的发展中体现出了十分优秀的效果与强大的学习能力，在图像处理领域同样也拥有十分瞩目的成果，同时 ResNet 的出现与 Relu 函数的优化效果，使得深度学习的功能与效果得到了进一步加强与拓展，其效果远远超出了人们的预期[1—2]。

在图像压缩和重构过程中，需要构建出由低分辨率-高分辨率映射对（LR-HR）所组成的图像特征匹配对，从而借由 LR-HR 的图像特征匹配对来实现图像的变化过程，但 LR-HR 图像特征匹配对的数量并不足以覆盖图像中的较大的纹理变化过程。Ghiasiand 等[3]通过拉普拉斯金字塔，利用自然图像所具有的自相似性构造出低分辨率输入图像的尺度-空间对的 LR-HR 的图像特征匹配对，从而证实了在图像中应用拉普拉斯金字塔的可行性。Singh 等[4]利用将补丁对分解为定向频率子带，并以此来独立地确定每个子带特征，进而在拉普拉斯金字塔中获得了更好匹配特性，对子带金字塔的优化操作取得了不错的成果。Huang 等[5]扩展了补丁对的搜索空间大小，以便对变换和透视变形来进行适当的调整，使得搜索过程更加的快捷。向静波等[6]提出了基于提升方案的拉普拉斯金字塔算法，可明显减少混淆现象，降低了图像重要细节的失真，提高了压缩性能和效率。这些方法的主要缺点在于尺度空间金字塔中的补丁搜索过程需要大量的计算成本，所以通常速度很慢，需要大量的时间。即使利用稀疏字典[7—8]或随机森林[9—10]等方法，也依然需要对图像的数据库进行相对应的分区操作，而并不是直接对复杂的分片空间进行建模，与此同时也需要为每个集群学习局部线性回归量，但是即使计算出了集群学习局部线性回归量也依旧会存在训练速度缓慢、效率低下的问题。

文中基于卷积神经网络与深度学习，提出采用拉普拉斯金字塔的图像压缩与重构方法，利用深度学习强大的学习性能与 Res 网络进行组合，以拉普拉斯金字塔[11—13]为主要的切入点，从而实现对图像的压缩与重构，为后续图像压缩与重构提供一种新的方法与思路。

1 基本原理

1.1 拉普拉斯金字塔网络结构

拉普拉斯金字塔卷积神经网络结构分为图像压缩和图像重构 2 个部分，见图 1。图像压缩部分：高分辨率图像作为输入数据，首先利用双三线性插值法（Bicubic）对输入图像进行一次缩小，将大小变为原来的 1/4，接着进行特征提取与图像重建，将得到的重建图输入第 2 层中，逐层降低图像大小，达到图像压缩的目的。特别注意在压缩过程中，每层都可对图像缩小 1/4，在通过 2 层缩小后即可实现 16 倍的压缩率。图像重构部分：对实验的结果图像进行反向操作，将前一过程获得的低分辨率图作为输入，带入重构还原体中，利用上采样操作逐层增加图像的大小，完成图像的逐层放大还原过程，进而实现图像的重构操作。

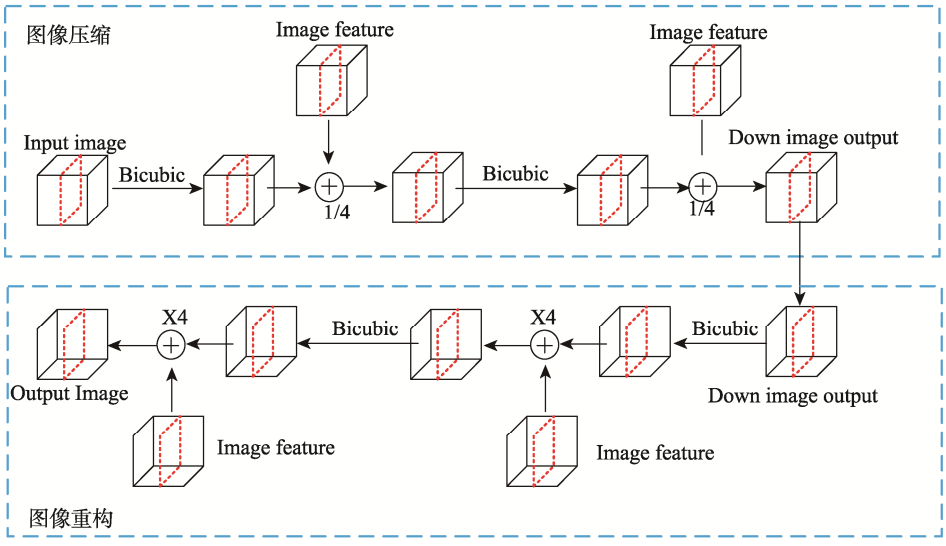

图 1 拉普拉斯金字塔网络结构

Fig.1 Structure of Laplacian pyramid network

1.2　图像压缩

1.2.1　特征提取

特征提取分支结构由 1 个残差网络层和 1 个转置卷积层组合而成，见图 2。再进行提取特征前，需要对原输入的图像进行双三线性插值处理，将原始输入图像缩小 4 倍。提取特征的过程中，需要对原始输入图像进行一次卷积来提取原始特征，然后将提取到的原始特征输入残差网络进行训练，再与原来未经过残差网络所得到的原始特征进行结合，再通过激活函数后就可以得到残差网络的结构特征图，在将结构特征图转置输出后，就会连接到 2 个不同的层上，用以完成后续的重建工作：将得到的图像特征进行卷积后将得到图像的梯度特征图，此特征图将用于与后续的特征图像进行结合；将得到的图像特征用于后续的缩小压缩过程，用于继续下一步中图像的特征压缩过程。

1.2.2　图像重建

在进行重建之前需要对原输入的图像进行插值处理，完成对原始输入图像进行压缩的过程，将其缩小 1/4 以减少后续图像训练过程中所耗的时间。这样就可以将得到的图像与特征提取分支所得到的残差预测图像特征图相组合，并通过逐个元素进行求和的方式来产生所需要的低分辨率图像。为了能获得更高比率的压缩率，需要将所输出的低分辨率图像输送到下一层，用以进行第 2 层图像的分支重建过程，这样就可以利用拉普拉斯金字塔卷积神经网络结构来对图像进行一步步的逐层压缩。在这整个网络中所包含的是一系列的 CNN 网络，在网络中的每个级别都包含有与此类似的结构，其网络结构见图 3。

1.3　图像重构

1.3.1　特征提取

实验中的特征提取分支结构和压缩过程中的分支提取结构一样，都是由残差网络和 1 个转置卷积层组合而成，但是在将所得到残差网络结构特征图通过对应的激活函数后，需要按比例 2 对提取的特征进行上采样操作，以此来获取扩大图像 4 倍的效果，并在下一层再次执行此类操作，以再次获得扩大 4 倍的图像，达到逐层放大的效果。其相应的网络结构见图 4。

1.3.2　图像重建

在重构过程的图像重建分支中，需要将在上一步的图像通过转置卷积层后，同样也要以 2 倍比例来进行上采样操作，从而增大其分辨率。这里同样采用双线性内核来对这个层进行初始化，并允许该上采样层与其余所有层进行共同优化操作。然后将上采样得到的图像与特征提取分支所得到的残差预测图像特征图相组合，这样就可得到扩大 4 倍的高分辨率的一级重建图像。然后将所输出的高分辨率重建图像输送到下一层，用以进行第 2 层图像的分支重建过程，从而得到所需要的结果。其网络结构见图 5。

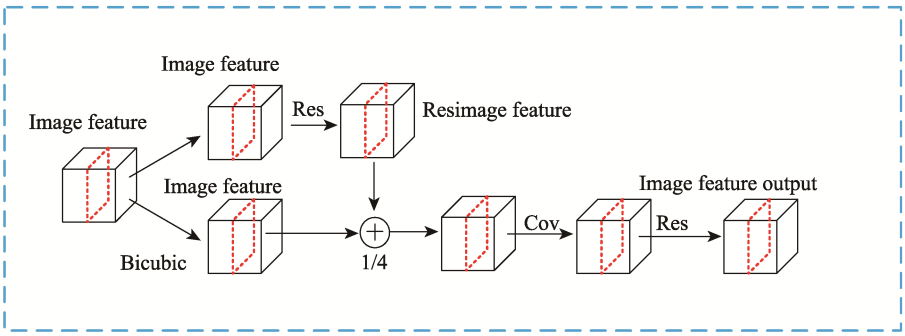

图 2　图像压缩过程中的特征提取分支
Fig.2 Feature extraction branching during image compression

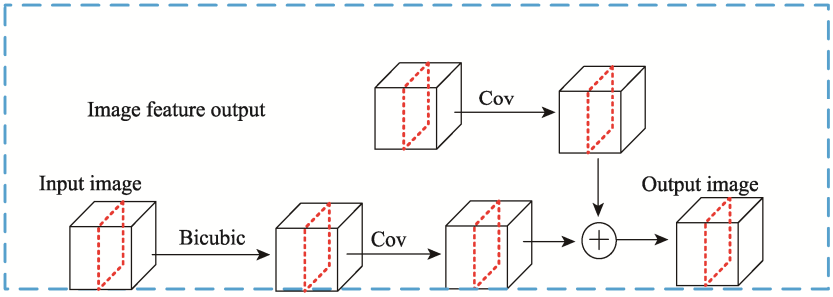

图 3　图像压缩过程中的图像重建分支
Fig.3 Image reconstruction branch during image compression

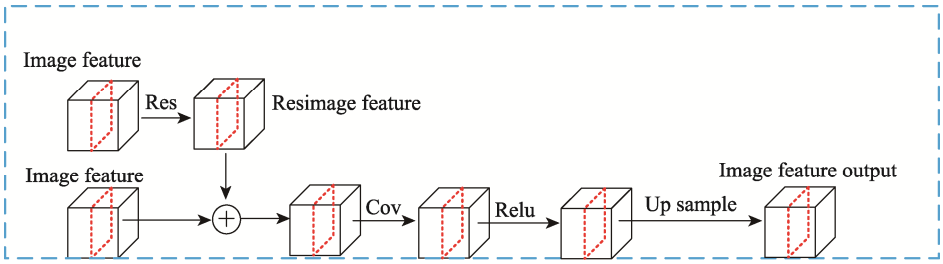

图 4 图像重构过程中的特征提取分支
Fig.4 Feature extraction branch during image reconstruction

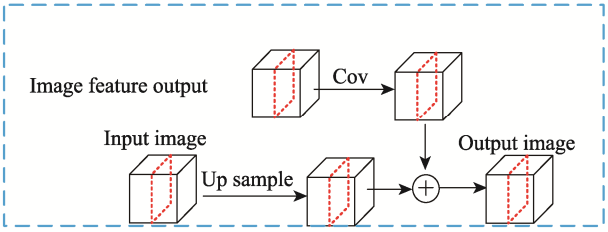

图 5 图像重构过程中的图像重建分支
Fig.5 Image reconstruction branch during image reconstruction

1.4 损失函数

在文中的实验过程中使用 Mean Square Error（MSE）作为损失函数，其计算公式为：

$$\min_{\theta}\frac{1}{n}\sum_{i=1}^{n}\left\|F(Y_s^i;\theta)-X^i\right\|_2^2$$

式中：n 为指训练样本的数量；Y_s^i 和 X^i 为训练数据中的第 i 个 LR 和 HR 子图像对；$F(Y_s^i;\theta)$ 为具有参数 θ 的 Y_s^i 的网络输出。在训练的网络中使用标准反向传播的随机梯度下降算法来使损失最小化。

2 实验结果与分析

2.1 实验设置

在所设计出的图像压缩与重构网络结构中，每个卷积层都是由 64 个滤波器组合而成，其中卷积核的大小设为 3×3[14]，而将转置卷积滤波器的卷积核大小设置为 4×4。同时在所有的卷积层和转置卷积层（重建层除外）之后都加上 1 个泄漏线性整流单元（LReLU）[15]，将其负斜率定为 0.2，残差网络层中设置其深度为 10。在训练中将动量参数初始值设置为 0.9，重量衰减初始值设置为 1×10^{-4}。与此同时将所有层的学习率初始化为 1×10^{-4}，训练周期定位 300 个周期，设定每经过 50 个时期就会将学习率降低 2 倍。在训练的过程中，使用来自 Yang 等[16]的 91 幅图像，来自 Berkeley Segmentation Dataset 训练集[17]的 200 张图像来作为文中的训练数据集。在每个培训批次中，文中随机抽取 64 个大小为 128×128 的补丁来进行相应操作。在 1 个周期中设置有 1000 次迭代的反向传播过程。由于数据量较小，未达到训练所需要的数据数量，为了增加训练的数据量，并提高训练的质量，选用白话、旋转和噪声扰动这 3 种方式来增加所需的训练数据。图像压缩的实验流程见图 6。图像重构的实验流程见图 7。

2.2 实验结果分析

为了检验文中算法的有效性，分别对 7 幅有代表性的图像进行压缩试验。所有图像都是 512×512 的灰度图，在文中将此模型与主成分分析算法、奇异值分解（SVD）算法、DCT 变换算法进行对比，选择在 16 倍的压缩率下对图片 man 的结果进行比较，实验的结果见图 8。其局部放大图见图 9。

在主观评价上，使用文中方法所得到的结果，在图像的清晰度上和还原度上要优于其余 3 种方法，与原图相比更加接近，所具有的清晰度更高，拥有较好的可视性能。能够完整地保留图像中的特征，在信息

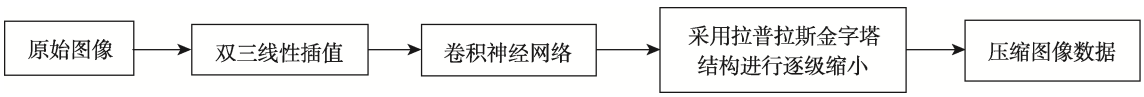

图 6 图像压缩的实验流程
Fig.6 Experimental flowchart of image compression

图 7 图像重构的实验流程
Fig.7 Experimental flowchart of image reconstruction

a 原图　b 文中方法　c PCA　d DCT　e SVD

图 8　图片 Man 的各方法结果

Fig.8 Various method results of figure Man

a 原图　b 文中方法　c PCA　d DCT　e SVD

图 9　图片 Man 的局部放大图

Fig.9 Partial enlarged view of figure Man

的还原度上实现了对图像信息的高度还原，具有很高的图像还原质量水平。利用 PCA 方法进行图像的压缩与重构值后，在还原的图像所具有的细节上模糊度增大，但不具有强烈的模糊块状感，在整体上保持了图像的大部分特征，并没有大量图像信息丢失的情况发生，且重构图像整体上完整，无较明显的差异，但是在图像质量上要劣于文中方法，在细节上也不够平滑。DCT 与 SVD 在重构之后，则有着较大的模糊度，块状感在相比之下十分明显，图像信息丢失较大，从而导致图像的模糊程度加大，图像平滑性下降，且其模糊程度与原图像相比，图像的模糊状况加大，所取得的重构结果图与原图差异度大，在整体上图像的特征损失增加。

在图像的客观评价上，表 1—3 分别表示压缩比为 4∶1，16∶1 和 64∶1 时的图像标准差和信息熵值。以压缩 16 倍为主要切入点，文中采用标准差和信息熵来检测图像的质量。据此文中在数据集之中选取 7 幅图片来表示其相应的实验结果，并且通过对这 7 幅图像进行取平均值处理，利用平均值来表示实验最终的对比结果。如表 2 所示，在所得到的结果数据中，文中方法得到了标准差为 52.73 与信息熵为 7.44 的最好结果，高于主成分分析法的 49.70 与 7.38。SVD 变换法和 DCT 变换法在标准差上只有 48.69 和 49.02，远不如文中方法，同时在信息熵上只有 7.34 与 7.35，

表 1　压缩比为 4∶1 的图像标准差和信息熵值比较

Tab.1 Comparison of 4∶1 image standard deviation and information entropy

方法	Baboon	Lena	peppers	Baby	Bridge	Liftingbody	Man	AVE
PCA	43.38/7.40	47.48/7.46	53.55/7.61	72.017.67	52.91/7.68	31.53/6.47	56.33/7.54	51.03/7.40
SVD	41.11/7.31	46.70/7.46	52.89/7.59	71.47/7.65	52.75/7.68	30.93/6.46	56.20/7.55	50.29/7.39
DCT	40.49/7.30	46.96/7.45	51.79/7.59	71.00/7.62	49.73/7.59	30.24/6.43	55.63/7.55	49.41/7.36
文中方法	45.23/7.60	54.89/7.66	55.13/7.65	72.11/7.69	53.08/7.70	31.76/6.51	60.42/7.59	53.23/7.49

注：“/”前为标准差；“/”后为信息熵

表 2　压缩比为 16∶1 的图像标准差和信息熵值比较

Tab.2 Comparison of 16∶1 image standard deviation and information entropy

方法	Baboon	Lena	peppers	Baby	Bridge	Liftingbody	Man	AVE
PCA	40.23/7.31	46.57/7.46	52.66/7.62	71.55/7.66	51.21/7.62	30.8/6.46	54.85/7.54	49.70/7.38
SVD	39.23/7.26	45.34/7.37	51.10/7.55	70.89/7.61	49.91/7.59	30.22/6.44	54.15/7.54	48.69/7.34
DCT	39.21/7.26	46.59/7.44	51.54/7.57	70.94/7.61	49.21/7.57	30.19/6.42	55.45/7.55	49.02/7.35
文中方法	47.14/7.51	54.75/7.64	54.49/7.64	71.85/7.63	50.88/7.62	30.67/6.46	59.32/7.59	52.73/7.44

注：“/”前为标准差；“/”后为信息熵

表 3 压缩比为 64∶1 的图像标准差和信息熵值比较

Tab.3 Comparison of 64∶1 image standard deviation and information entropy

方法	Baboo	Lena	peppers	Baby	Bridge	Liftingbody	Man	AVE
PCA	37.25/7.21	44.25/7.41	49.81/7.56	70.27/7.64	48.60/7.56	28.94/6.41	52.03/7.53	47.31/7.33
SVD	37.52/7.16	44.65/7.31	50.49/7.50	70.23/7.53	48.67/7.51	29.29/6.35	52.71/7.48	47.65/7.26
DCT	37.66/7.17	44.84/7.31	50.69/7.51	70.50/7.54	48.89/7.52	29.40/6.36	52.92/7.48	47.84/7.27
文中方法	39.13/7.41	46.35/7.59	51.38/7.58	71.05/7.63	49.65/7.58	30.03/6.42	56.89/7.56	49.21/7.40

注：“/”前为标准差；“/”后为信息熵

低于文中的 7.44，由此可以看出文中方法的实验结果要优于 PCA，DCT 和 SVD，所得到的图像压缩与重构结果图的质量得到了较大的提升。再看压缩 4 倍与 64 倍。在压缩 4 倍中，文中方法获得了标准差（53.23）与信息熵（7.49）的最好结果；同样在 64 倍中，文中方法以标准差 49.21 与信息熵 7.40，依旧取得了 4 种方法中最好的结果。

3 结语

提出了一种新的图像压缩与重构的方法，利用拉普拉斯金字塔结构来设计卷积神经网络结构，降低了计算量与复杂度，并通过反向结构上采样来恢复图像的大小，从而完成了图像压缩与重构的过程。与经典的算法相比在图像标准差和信息熵值上取得了不错的提升，具有一定的应用前景。

参考文献：

[1] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]// Computer Vision and Pattern Recognition, 2016: 770—778.

[2] KRIZHEVSKY, ALEX, LLYA S, et al. HINTON. ImageNet Classification with Deep Convolutional Neural Networks[C]// Association for Computing Machinery, 2012 (60): 84—90.

[3] GHIASIAND G, FOWLKES C C. Laplacian Pyramid Reconstruction and Refinement for Semantic Segmentation[C]// European Conference on Computer Vision, 2016.

[4] SINGH A, AHUJA N. Super-resolution Using Sub-band Self-similarity[C]// Asian Conference on Computer Vision in, 2014.

[5] HUANG J B, SINGH A, AHUJA N. Single Image Super Resolution from Transformed Self-exemplars[C]// Computer Vision and Pattern Recognition, 2015: 5—8.

[6] 向静波，苏秀琴. 基于提升拉普拉斯金字塔变换的图像压缩[J]. 电视技术, 2007(5): 13—15.
XIANG Jing-bo, SU Xiu-qin. Image Compression Based on Lifting Laplace Pyramid Transform[J]. TV Technology, 2007(5): 13—15.

[7] TIMOFTE R, SMET V D, GOOL L V, et al. A+: Adjusted Anchored Neighborhood Regression for Fast Super-resolution[C]// Asian Conference on Computer Vision, 2014: 5—7.

[8] 邱康，易本顺，向勉，等. 协作稀疏字典学习实现单幅图像超分辨率重建[J]. 光学学报，2018，38(9): 130—136.
QIU Kang, YI Ben-shun, XIANG Mian, et al. Collaborative Sparse Dictionary Learning to Achieve Single-image Super-resolution Reconstruction[J]. Acta Optica Sinica, 2018, 38(9): 130—136.

[9] SCHULTER S, LEISTNER C, BISCHOF H, et al. Fast and Accurate Image Upscaling with Super-resolution Forests[C]// Computer Vision and Pattern Recognition, 2015: 4—7.

[10] 蔡加欣，冯国灿，汤鑫，等. 基于局部轮廓和随机森林的人体行为识别[J]. 光学学报, 2014, 34(10): 212—221.
CAI Jia-xin, FENG Guo-can, TANG Xin, et al. Human Behavior Recognition Based on Local Contour and Random Forest[J]. Acta Optica Sinica, 2014, 34(10): 212—221.

[11] BURT P J, ADELSON E H. The Laplacian Pyramid as a Compact Image Code[J]. IEEE Transactions on Communications, 31(4): 532—540.

[12] DENTON E L, CHINTALA S, FERGUS R. Deep Generative Image Models Using a Laplacian Pyramid of Adversarial Networks[C]// Conference and Workshop on Neural Information Processing Systems, 2015.

[13] PARIS S, HASINOFF S W, KAUTZ J. Local Laplacian Filters: Edge-aware Image Processing with a Laplacian Pyramid[J]. Association for Computing Machinery Transaction on Graphics (Proc of SIGGRAPH), 2011, 30(4): 68.

[14] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[C]// International Conference on Learning Representations, 2015.

[15] MAAS A L, HANNUM A Y, NG A Y. Rectifier Nonlinearities Improve Neural Network Acoustic Models[C]// Internation Conference on Machine Learning, 2013: 723—729.

[16] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour Detection and Hierarchical Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898—916.

[17] YANG J, WRIGHT J, HUANG T S, et al. Image Super Resolution Via Sparse Representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861—2873.