

文章编号: 2095-2163(2020)09-0208-06

中图分类号: TP391.4

文献标志码: A

基于样本倍增、深度神经网络与 SVM 的少样本图像识别技术

秦俊举¹, 曹选平¹, 夏校朋²

(1 成都纺织高等专科学校 机械工程学院, 成都 611731; 2 电子信息控制重点实验室 成都 610036)

摘要: 基于深度神经网络的图像识别技术在少样本情况下会出现识别准确率下降, 甚至训练失败的情况。本文针对该问题, 提出了一种将少样本倍增、深度神经网络、SVM(support vector machine)相结合的算法。通过样本倍增提升样本数量, 通过深度神经网络训练得到图像的高维特征, 通过经典特征提取算法提取图像常规特征, 构建包含高维特征和常规特征的支持向量, 使用 SVM 进行训练。实验表明, 本文所提方法对于特定样本集, 在样本数量为总样本 1/40 的情况下, 识别准确率达到 61.36%(仅仅使用 YOLO 神经网络算法识别准确率为 44.08%), 大大提高了少样本下目标识别准确率, 具有重要的应用价值。
关键词: 样本倍增; SVM; 深度神经网络

Image recognition technology for few samples based on sample multiplication, deep neural network and SVM

QIN Junju¹, CAO Xuanping¹, XIA Xiaopeng²

(1 College of Mechanical Engineering, Chengdu Textile College, Chengdu 611731, China;

2 Key Laboratory of Electronic Information Control, Chengdu 610036, China)

[Abstract] In the case of few samples, the recognition accuracy of image recognition technology based on deep neural network will decrease. To solve this problem, this paper proposes an algorithm which combines sample multiplication, deep neural network and SVM (support vector machine). Firstly, the sample multiplication is used to increase the number of samples. Secondly, deep neural network is used to obtain the high-dimensional features of the image. Thirdly, the classical feature extraction algorithms is performed to extract conventional features of the image, and then the support vector including the high-dimensional features and the conventional features is constructed, which is trained by SVM. Experiments show that the accuracy of the proposed method is 61.36% (the accuracy of using Yolo algorithm only is 44.08%) for a specific sample set, whose number of samples is 1 / 40 of the total samples, which greatly improves the accuracy of recognition under a small number of samples and has important application value.

[Key words] sample multiplication; SVM; deep neural network

0 引言

基于深度神经网络的图像识别技术已得到广泛的应用。但是其缺点在于需要大量图像样本进行训练, 才能保证较好的识别效果。但某些应用下, 如遥感图像或红外图像的目标识别中, 由于图像采集的高成本和高难度使得不易获得大量的训练图像, 会导致训练不收敛, 或识别准确率大大降低。针对该问题, 人们研究了一系列方法, 以提升少样本下图像识别的准确度^[1]。

任佳等引入支持向量机, 利用贝叶斯网络各观测节点不同时刻下的预测信息, 通过样本回归对缺失数据进行修复, 解决了小样本下的 BN 参数学习问题^[2]; 郑欣悦提出结合表征学习和注意力机制的少样本学习方法 VAE-ATTN^[3]。首先, 通过预训练变分自编码器 (variational autoencoder, VAE) 学习丰富的隐特征; 其次, 对提取出的隐特征构建注意力机

制; 最后, 使用分类器进行图像分类; 田霞针对基于元学习的少样本图像分类, 设计了多尺度关系网络^[4]。首先, 在特征提取器加入多尺度特征, 将卷积神经网络的输出在深度方向进行拼接; 其次, 设计多尺度特征的结合方式; 最后, 将关系特征通过神经网络进行度量学习, 提高了少样本学习基准集的分类准确率, 缓解了过拟合情况; 徐培针对少样本的应用场景, 提出了如何有效地组织样本和高效检测目标等系列方法^[5]; 王宁针对人脸图像容易受外界影响, 且存在少样本情况, 提出先验采样约束结合扩展遮挡字典的细化稀疏识别算法, 在光照、表情、角度变化情况下, 取得了较满意的识别结果^[6]。

除了对少样本情况下的识别算法优化改进, 还对少样本倍增算法进行研究。例如通过图片的亮度改变、加噪声、翻转、随机裁切、缩放等制造出大量样本; 使用如 GAN 的生成模型, 生成大量接近真实的

作者简介: 秦俊举(1988-), 女, 硕士, 讲师, 主要研究方向: 智能工业机器人、图像识别、机电一体化; 曹选平(1972-), 男, 学士, 副教授, 主要研究方向: 机电一体化; 夏校朋(1987-), 男, 博士, 高级工程师, 主要研究方向: 多谱综合识别、综合一体化。

收稿日期: 2020-07-23

样本;使用经典的特征提取算法,再使用 SVM 等浅层网络来训练,需要较少的数据量。

1 传统深度神经网络识别算法的局限

基于深度神经网络的图像识别算法有很多,比较著名的有 YOLO (You Only Look Once)、SSD、FasterRCNN 等算法。其中,YOLO 算法以运算速度快,可以达到实时运算,且识别准确率高而闻名。YOLO 把一整张图片应用到一个神经网络中去,网络把图片分成不同的区域,给出每个区域的边框预测和概率,并依据概率大小对所有边框分配权重。最后设置阈值,只输出得分(概率值)超过阈值的检测结果。

本文以该算法为例说明其识别准确率与样本数量的关系。首先对 YOLO 算法性能进行测试。选取的图像样本集为在网上搜集的渔船、货轮、驱逐舰、护卫舰、航母等 5 个类别的 7 200 余张图片,每个类别约 1 200 张。利用其中约 60% 的图像作为训练样本集 A1;20% 的图像作为验证样本集 A2;另外 20% 作为测试样本集 A3,进行网络训练。YOLO 算法识别过程如图 1 所示。

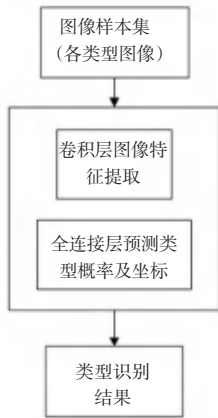


图 1 YOLO 图像识别算法

Fig. 1 The algorithm of YOLO image recognition

深度神经网络的识别准确率随样本数量而变化。样本数量越少,识别准确率下降的就越大,见表 1。而对某些特定应用,如雨雾环境监视等,高质量、足量的样本图像并不易获得,因此研究少样本下的识别有重要意义。

表 1 识别准确率随样本数量变化

Tab. 1 The recognition accuracy varies with the number of samples

少样本数量/全样本数量	识别准确率
1	Avg: 0.7700
1/20	Avg: 0.5150
1/40	Avg: 0.4408

2 本文提出方法

本文提出一种将少样本倍增、深度神经网络与 SVM 相结合的识别算法。

首先,采用样本倍增手段,对现有的少样本图像进行旋转、加噪、视角变换等操作,从而生成更多类似实际场景的样本,在一定程度上缓解样本数量偏少给深度神经网络训练带来的不利影响;

其次,采用 YOLO 深度神经网络对倍增后的样本进行训练,得到目标的高维特征,包括类型识别结果(目标类型、属于该类型的概率)及其它特征(目标像素的长宽比)等;

最后,采用 SVM 支持向量机。首先利用经典的图像特征提取算法(SIFT 特征点提取、Hough 直线检测等)提取目标特征(船舶姿态及 SIFT 特征点等),再与 YOLO 提取的高维目标特征一起构建支持向量,使用 SVM 对支持向量进行训练和识别。

但本文采用深度神经网络获取目标的高维特征信息,使用经典特征提取算法提取目标常规特征,构建包含高维特征信息、常规特征信息的特征向量,使用 SVM 支持向量机训练,达到深度学习神经网络与浅层神经网络的统一运用,以提高少样本下的分类识别能力。算法流程如图 2 所示。

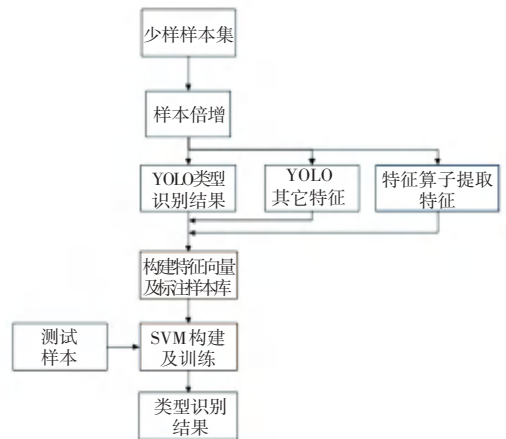


图 2 基于样本倍增、深度神经网络、SVM 的少样本图像识别算法

Fig. 2 The algorithm of few-sample image recognition based on sample multiplication, deep neural network, and SVM

2.1 样本倍增

样本倍增可以通过旋转、缩放、平移、模糊、加噪等操作来实现。缩放是最简单的操作,可以模拟物体的远近程度。旋转、平移可以通过透视变换来实现。

2.1.1 透视变换样本倍增

世界坐标系 (X, Y, Z) 到摄像机坐标系 (x, y, z) 的变换。该变换可以由世界坐标系到摄像机坐标系的平移和旋转得到,式(1):

$$\begin{pmatrix} \hat{e}^x \\ \hat{e}^y \\ \hat{e}^z \end{pmatrix} = M \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + C = [M, r] \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (1)$$

其中, M 为 3×3 的旋转矩阵; C 和 r 表示 3×1 的平移单位矢量; (X, Y, Z) 和 (x, y, z) 分别表示变换场景前和变换后的点坐标。

摄像机坐标系 (x, y, z) 到二维图像坐标系 (u, v) 的变换得到, 式(2):

$$S \begin{pmatrix} \hat{e}^u \\ \hat{e}^v \\ \hat{e}^1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{e}^x \\ \hat{e}^y \\ \hat{e}^z \end{pmatrix} + \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} \hat{e}^x \\ \hat{e}^y \\ \hat{e}^z \end{pmatrix} + \begin{pmatrix} \hat{e}^0 \\ \hat{e}^1 \\ \hat{e}^2 \end{pmatrix} \quad (2)$$

其中, $\begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$ 是相机内参数矩阵,

$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$ 是摄像头的标定参数, 指示镜头的畸变程度。由式(1)和(2)可知, 相机根据拍摄物

体的三维世界坐标 (X, Y, Z) 可以得到二维图像坐标 (u, v) , 反之是无法实现的。对同一相机来说, 其相机内参数矩阵和外参数矩阵是一定的, 对同一空间物体, 通过施加不同的透视变换矩阵 $[M, r]$, 可以得到不同视角的样本图像, 达到样本倍增的目的。图3表示的是一幅原图经透视变换后生成另一幅图。可以看出, 该透视图与原来的图像样本相比有较大的视角差异, 可以达到样本倍增目的。



图3 原图及经透视变换后生成的样本图像

Fig. 3 The original image and image generated after perspective transformation

2.1.2 模糊变换样本倍增

模糊变换可以通过图像下采样后插值放大来实现。也可以使用低通滤波的方法来实现, 如高斯滤波等, 如图4所示。这些滤波操作削弱了图像中的高频分量, 使边缘信息变得模糊。



图4 图像缩放后再插值放大后的模糊效果、高斯模糊产生的模糊效果

Fig. 4 The blur effect of image by zooming and interpolating, and the blur effect produced by Gaussian blur

2.1.3 加噪变换样本倍增

噪声可以模拟图像在实际场景中所受到的成像影响。但对图像所加的噪声影响需要在最大程度上满足实际真实情况。以海上监视应用为例, 船舶图像主要受海雾影响。根据经典的大气散射物理模型, Narasimhan 等人给出雾、霾天气条件下单色大气散射模型^[7], 即窄波段摄像机所拍摄的图像像素 $I(x)$ 可表示为式(3):

$$I(x) = A\rho(x)e^{-\beta d(x)} + A(1 - e^{-\beta d(x)}), \quad (3)$$

其中, x 为空间坐标; A 表示天空的亮度; $\rho(x)$ 为空间坐标 x 处的场景反照率; $d(x)$ 为空间坐标 x 处的场景深度; β 为大气散射系数。因此, 对原图 $I(x)$ 处理, 模拟典型条件下的场景反照率、大气散射系数及场景深度, 从而生成模拟的海上舰船图像, 如图5所示。

除此之外, 还有其它变换倍增方法, 如水平反转变换(如图6所示)等。

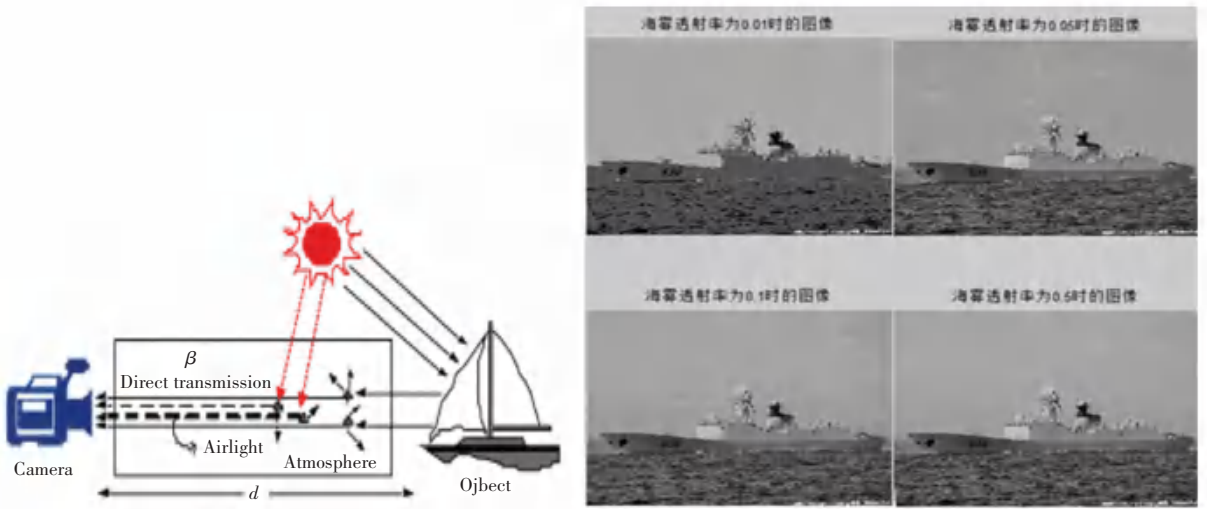


图 5 大气散射物理模型及海雾对图像质量影响模拟

Fig. 5 The physical model of atmospheric scattering and simulation of the influence of sea fog on image quality



图 6 图像水平翻转倍增

Fig. 6 The image multiplying by horizontally flipping

2.2 深度神经网络 (YOLO) 特征提取

经过倍增后的样本集按 60% 划分为新的训练集 B1, 20% 划为新的验证集 B2, 另外 20% 作为新的测试集 B3。对训练样本集 B1 和验证样本集 B2 进行 YOLO 算法的训练和验证, 并使用训练后的权重对测试样本集 B3 进行识别, 可以得到目标的类型信息。除此之外, 还能得到目标属于该类型的概率, 及目标在图像中的位置 (4 个坐标点)、目标的长宽比信息。这些信息均可作为 SVM 支持向量机的特征输入。

2.3 特征算子特征提取

对于新的测试样本集 B3, 借助传统的特征提取算法, 如边缘提取、hough 直线检测、SIFT 特征点检测等技术, 可以提取船舶相对图像水平方向的角度, 作为目标姿态的表征。如图 7 所示, 首先通过边缘提取算法得到目标的边缘信息, 然后对目标的水平纹理图像进行 hough 变换, 提取水平线条, 进而得到

水平线条与图像水平方向的夹角, 作为支持向量机的一个特征输入。除此之外, 还可以提取目标的特征点, 如 SIFT 特征点, 将提取出的特征点作为支持向量的特征输入。

2.4 SVM 支持向量构建

支持向量机 (Support Vector Machine) 是 Cortes 和 Vapnik 于 1995 年首先提出的, 它在解决小样本、非线性及高维模式识别中表现出许多特有的优势。

利用 YOLO 训练得到的权重, 通过测试可以得到测试样本集 B3 的目标类型、识别概率、目标长宽比, 以及通过特征算子可以提取目标姿态、目标 SIFT 特征点等信息。这些信息共同构成多维支持向量输入, 测试样本集 B3 的真实类型作为 SVM 支持向量机的输出, 对 SVM 开展训练。训练完成后, 可以使用训练好的 SVM 对原始的测试样本集 A3 进行识别。SVM 支持向量构成见表 2。



图7 Hough 直线检测结果及船舶的 SIFT 特征点分布图

Fig. 7 The result of Hough transforming and the distribution map of the ship's SIFT feature points

表2 SVM 支持向量构成

Tab. 2 The composition of SVM

支持向量	数据结构
少样本倍增后 YOLO 识别结果	目标类型 目标概率
目标长宽比	比值
舰船姿态	舷线水平角度均值
舰船 SIFT 特征点	SIFT 特征向量(降维)

3 仿真实验

采用各种变换方法对 1/40 总样本数的少样本图像进行倍增,共倍增约 22 倍。将上述样本的 60% 作为训练样本集 B1,20% 作为验证样本集 B2,另外 20% 作为测试样本集 B3。

采用倍增后的训练样本集 B1 及验证样本集 B2 对 YOLO 算法进行训练。进行 21 000 次迭代,选取 IOU 最优的迭代次数,最终确定的迭代次数为 18 000 次。使用其权重识别验证测试样本集 B3,得到包括目标识别类型及类型概率、目标像素长宽比

等信息,作为 SVM 的特征向量输入。

根据 YOLO 识别算法框选出的目标图像,提取目标侧舷的水平线条,计算相对于图像水平方向的角度,度量船身姿态,作为 SVM 支持向量的输入。

同时,对框选出的目标图像进行 SIFT 特征点提取。由于 SIFT 特征点为 128 维向量,维度高,数据量大。因此,为避免 SIFT 算子导致 SVM 训练时间明显增加,这里只选取 20 个 SIFT 特征点。并对 128 维的向量降维,使用 SIFT 向量各维度平方和作为 SIFT 特征的近似表征。由表 2 可知,本文的 SVM 支持向量由 24 个维度组成,对应输出则是 B3 的真实类型标签。

接着使用 SVM 训练测试样本集 B3。最优的处罚因子 c 和核函数 g 的选择对分类精度影响很大,在本文的试验中,用函数 SVMcgForClass 自动选择 c 和 g 的值。图 8 是优化过程中的参数选择结果。用得到的最优 c 和 g ,通过 svmtrain 函数训练分类模型。

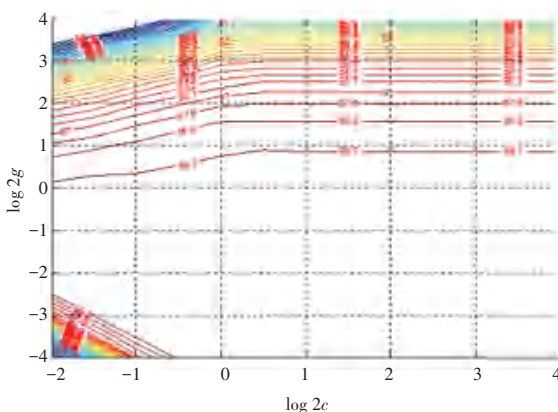
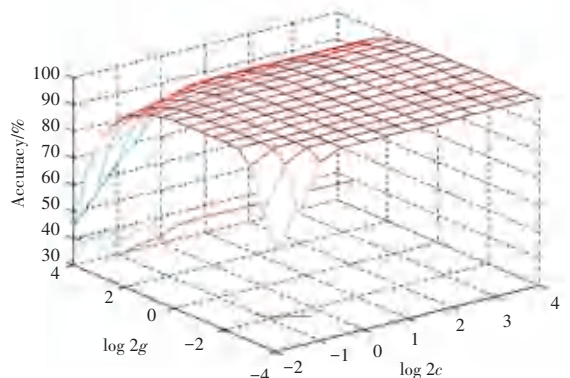
SVC 参数选择结果图(等高线图)[GridSearchMethod]
Best $c=0.25$ $g=0.25$ CV Accuracy=100%SVC 参数选择结果图(3D 视图)[GridSearchMethod]
Best $c=0.25$ $g=0.25$ CV Accuracy=100%

图8 训练过程中的参数选择

Fig. 8 The parameter selection during training

SVM 训练完成后,使用训练生成的 model 模型对待测试样本集 A3 (共计 5 类,每类 240 幅)进行

类型识别测试。使用 libsvm 中的 svmpredict 函数来测试分类精度,实验结果如图 9 所示。

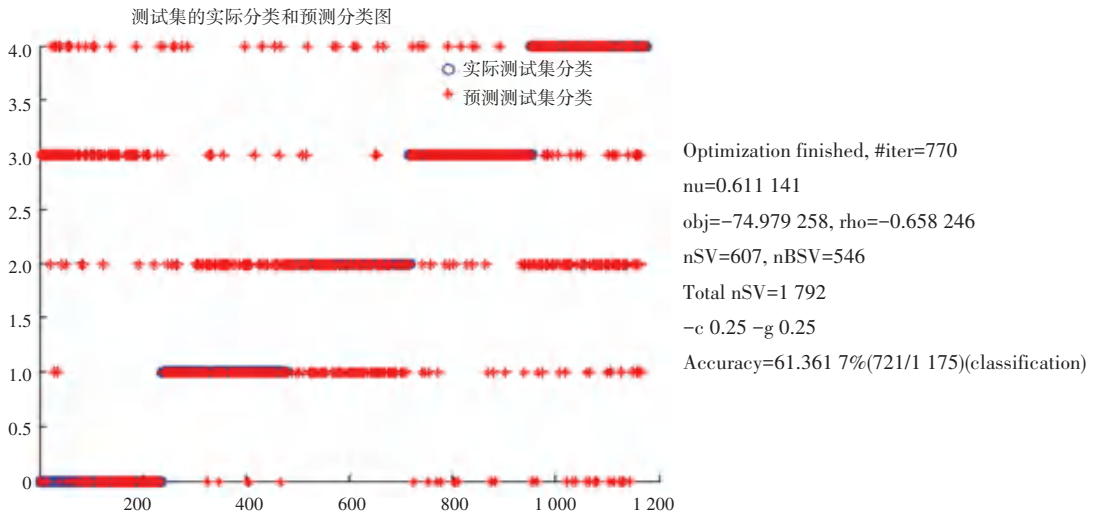


图9 SVM最终识别结果

Fig. 9 The recognition result of SVM

从图9可以看出,本文所采用的算法最终识别准确率达到61.36%,显著优于原YOLO算法在1/40样本时44.08%的识别率。

4 结束语

本文采用少样本倍增、深度神经网络与支持向量机相结合的方法,对于特定样本集,在样本数量为总样本1/40的情况下,识别准确率达到61.36%,显著优于原YOLO算法在1/40样本时44.08%的识别准确率。验证了该算法在少样本识别中的有效性。

但该算法性能受样本倍增手段是否符合实际情况、特征向量的构建是否合理等诸多因素影响,存在一定的不确定性。因此,提高识别能力最核心、有效的保证仍然首选高质量、足量的真实样本。但尽管如此,作为一种重要的补充手段,该算法可以在一定

程度上弥补由于样本数量不足引起的识别能力下降,仍然具有重要的研究价值。

参考文献

- [1] 李新叶,龙慎鹏,朱婧. 基于深度神经网络的少样本学习综述[J/OL]. 计算机应用研究:1-8[2019-11-26].<https://doi.org/10.19734/j.issn.1001-3695.2019.03.0036>.
- [2] 任佳,高晓光,茹伟. 数据缺失的小样本条件下BN参数学习[J]. 系统工程实践,2011,31(1):172-177.
- [3] 郑欣悦. 基于深度学习的少样本图像分类方法[D]. 中国科学院大学(中国科学院国家空间科学中心),2019.
- [4] 田霞. 基于元学习的少样本图像分类方法研究[D]. 电子科技大学,2019.
- [5] 徐培. 基于少量样本的快速目标检测与识别[D]. 电子科技大学,2014.
- [6] 王宁. 非充分样本环境下扩展字典稀疏表示人脸识别算法研究[D]. 燕山大学,2015.
- [7] 马忠丽,文杰,梁秀梅. 海雾下视频图像去雾算法研究[J]. 计算机应用研究,2014,31(9):2836-2840.

(上接第207页)

参考文献

- [1] ZHU H, GAO Y, HOU Y, et al. Multi-time slots real-time pricing strategy with power fluctuation caused by operating continuity of smart home appliances[J]. Engineering Applications of Artificial Intelligence, 2018, 71:166-174.
- [2] QIAN L P, ZHANG Y J A, HUANG J, et al. Demand Response Management via Real-Time Electricity Price Control in Smart Grids[J]. IEEE Journal on Selected Areas in Communications, 2013, 31(7):1268-1280.
- [3] 朱红波,高岩,后勇,等. 马尔可夫过程下多类用户智能电网实时电价[J]. 系统工程理论与实践, 2018, 38(3): 807-816.
- [4] SAMADI P, MOHSENIAN-RAD H, WONG V W S, et al. Adaptive energy consumption scheduling with load uncertainty for the smart grid[C]// Communications (ICC), 2013 IEEE International Conference on. IEEE, 2013.
- [5] John W. Labadie. Closure to "Optimal Operation of Multireservoir

Systems: State-of-the-Art Review" by John W. Labadie[J]. Journal of Water Resources Planning and Management, 2005, 131(5):407-408.

- [6] 王成山. 微电网分析与仿真实例[M]. 科学出版社,2013.
- [7] 胡海松,张保会,张嵩,等. 微电网中的储能设备及飞轮储能特性的研究[J]. 电网与清洁能源, 2010, 26(4):21-24.
- [8] 梅亚东,熊莹,陈立华. 梯级水库综合利用调度的动态规划方法研究[J]. 水力发电学报,2007,26(2):1-4.
- [9] 胡建军,周冀衡,柴家荣,等. 多指标正交试验数据的优化分析及应用[J]. 中国烟草学报, 2008, 14(2):9-14.
- [10] 翟国富,梁慧敏,王峰,等. 基于正交试验设计的极化磁系统参数优化设计方法的研究[J]. 中国电机工程学报, 2003, 23(10):158-163.
- [11] HOWSON H R, SANCHO N G F. A new algorithm for the solution of multi-state dynamic programming problems[J]. Mathematical Programming, 1975, 8(1):104-116.
- [12] 冯仲恺,廖胜利,程春田,等. 库群长期优化调度的正交逐步优化算法[J]. 水利学报, 2014(8):903-911.