

面向社交媒体图像的显著性数据集

梁 晔<sup>1,2</sup>, 马 楠<sup>1</sup>, 郎丛妍<sup>2</sup>, 于 剑<sup>2</sup>

(1.北京联合大学 机器人学院, 北京 100101; 2. 北京交通大学 计算机与信息技术学院, 北京 100044)

**摘 要:**随着显著性研究的发展,已涌现多个显著性数据集,然而目前面向社交媒体图像的显著性数据集数量非常少.为此构建此类显著性数据集,详细论述了数据集的图像来源、图像的筛选原则、图像的标注及数据集的统计分析.为了验证新建数据集的性能,与目前流行的 7 个显著性数据集进行性能评测,新建数据集具有显著区域尺寸丰富、与图像边界连接度高、显著区域与图像的颜色差异小的优点.实验结果表明:新建数据集中显著区域与图像边界连接的比例为 17%,仅低于 ECSSD 数据集;其中显著区域和整幅图像的颜色差均值最小,且包含 10 个尺寸等级的显著区域,尺寸分布最广泛.此外,新建数据集具有标签信息,也为新的显著区域提取方法提供了实验对象.

**关键词:**数据集;显著性;社交媒体图像;标签

**中图分类号:**TP391.9      **文献标志码:**A

A saliency dataset for social images

LIANG Ye<sup>1,2</sup>, MA Nan<sup>1</sup>, LANG Congyan<sup>2</sup>, YU Jian<sup>2</sup>

(1.College of Robotics, Beijing Union University, Beijing 100101, China;  
2.School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:**With the development of saliency research, there have been a number of saliency datasets, but saliency datasets for social images are few. Aiming at the current situation, a saliency dataset for social images is constructed. The source of images, principles of image selection, annotation of images and statistical analysis of dataset are discussed in detail. In order to verify the performance of the new dataset, the new dataset and the current seven popular datasets are evaluated. The new dataset has the advantages of rich salient region sizes, high connectivity with image boundaries, small color difference of salient regions and images. In the new dataset, the ratio of salient regions connected with image boundaries is 17%, which is only lower than that in the ECSSD dataset; the mean color difference between salient regions and whole images is the smallest in the eight experimental datasets; salient regions have 10 size grades which are the widest size distribution. In particular, the new dataset has tag information and provides experimental dataset for new saliency detection methods.

**Keywords:**dataset; saliency; social images; tags

收稿日期:2017-07-16  
基金项目:国家自然科学基金(61871038,61871039);北京市自然科学基金(4182022);北京市属高校高水平教师队伍支持计划项目(IDHT20170511)  
Foundation items:National Natural Science Foundation of China (61871038,61871039); Beijing Multicipal Natural Science Foundation (4182022); Supporting Project of High-level Teachers in Beijing Municipal Universities (IDHT20170511)  
第一作者:梁晔(1978—),女,内蒙古赤峰人,讲师,博士.研究方向为图像处理. email: liangye@buu.edu.cn.  
引用格式:梁晔,马楠,郎丛妍,等.面向社交媒体图像的显著性数据集[J].北京交通大学学报,2018,42(5):135—140.  
LIANG Ye, MA Nan, LANG Congyan, et al. A saliency dataset of social images [J].Journal of Beijing Jiaotong University, 2018, 42(5):135—140.(in Chinese)

随着显著区域检测的研究,显著性的应用越来越广泛<sup>[1-3]</sup>,也涌现了数十个显著性数据集.显著性数据集来自于两个领域:一类是为了显著性研究而建立的数据集,另一类是从图像分割领域延伸过来的显著性数据集.MSRA 数据集<sup>[4]</sup>包含 20 840 幅图像,通过计算标注一致性分值进行了矩形框级别的标注.ASD 数据集<sup>[5]</sup>包含 1 000 幅图像,标注了像素级精度的显著区域. DUT-OMRON 数据集<sup>[6]</sup>包含了 5 168 幅的图像,图像中包含 1 个或多个显著区域.每幅图像由 5 个观察者进行标注,标注结果有 3 种形式:像素级的显著区域、矩形框级的显著区域和眼动视点.MSRA10000 数据集<sup>[7]</sup>包含 10 000 图像,采用一致性分值计算方法<sup>[4]</sup>,精确标注了像素级精度的显著区域. MSRA5000<sup>[8]</sup>挑选了 MSRA-B1 数据库中的 5 000 幅图像,手工标注像素级精度的显著区域.ImgSal 数据集<sup>[9]</sup>包含 235 幅图像,图像大小为 480×640,分为 6 类.观察者手工标注像素级精度的显著区域.ECSSD 数据集<sup>[10]</sup>包含 1 000 幅图像,图像来自于 BSD300 数据集、VOC 数据集和网络,由 5 位观察者标注像素级精度的显著区域. HKU-IS 数据集<sup>[11]</sup>包含 4 447 幅图像,图像符合 3 个条件:存在多个不连接的显著区域;至少一个显著区域连接到图像的边界;所有显著区域和它相邻区域的颜色差异小于 0.7,并由 3 个观察者进行像素级精度的显著区域标注.SOD 数据集<sup>[12]</sup>包含了 300 幅图像,来自于伯克利大学的 BSD 分割数据集,这是第一次将分割领域的数据集用于显著性评测,并由 7 名观察者进行像素级精度的显著区域标注.从显著性数据集的相关文献来看,数据集从矩形框标注更多地转向于像素级别的标注.数据集的图像从简单的图像结构、中心偏差以及前景和背景的明显差异变得越来越复杂,越来越具有挑战性,但数据集目前仍存在问题,例如,前景和背景差异明显、中心偏差明显、图像挑选带有很强的人为主观因素等.目前对显著性数据集的性能评测研究非常少,没有成熟的评测方法来评价数据集的性能.

随着时代的发展,人们更多的选择线上的娱乐方式,社交平台和应用变得越来越广泛.据统计,每月上传到 Facebook 网站上的视频数量超过 2 000 万个,图片数量超过 30 亿张,每年上传到 Facebook 上的图片数量超过 360 亿张;平均每分钟上传到 YouTube 网站上的视频时长为 35 h;每分钟上传到 Flickr 网站上的图片数量超过 3 000 张,每月上传到 Flickr 网站上的图片数量为 1.3 亿张,Flickr 图片总量超过到了 60 亿张.图片社交市场的迅速发展

使社交媒体图像已成为一种重要的图像类型.用户不仅可以创建自己的多媒体内容,也可以使用文字描述媒体内容.文本也被称为标签,这是社交媒体的重要特点.社交平台的飞速发展带来了海量的社交媒体图像,然而目前面向社交媒体图像的显著性数据集很少,亟需构建此类的显著性数据集为新的显著区域提取方法提供实验对象.

基于上面的分析,本文作者以社交媒体图像为研究对象,针对构建显著性数据集时没有明确的图像选取原则或者选取原则不够科学的研究现状,提出了图像的筛选原则,并构建了面向社交媒体图像的数据集.此外,对新建的数据集和目前流行的数据集进行了性能统计分析对比,验证数据集的性能.

1 数据集的图像筛选原则

目前存在大量的显著区域提取方法,这些提取方法分别从不同的角度反映出显著区域提取过程中的难点.可以认为如果数据集具有提取中的难点问题,那么数据集对于显著区域提取具有一定的难度.另外,目前文献在构建数据集时仍存在问题.以上两点是本文确定图像筛选原则的出发点.

1.1 原则的确定

ImgSal 数据集<sup>[10]</sup>包含 235 幅图像,图像被分为 6 类:包含大尺寸显著区域的图像、包含中等尺寸显著区域的图像、包含小尺寸显著区域的图像、具有大小不同的多个显著区域的图像、具有杂乱背景的图像、背景区域和显著区域非常相似的图像.然而,ImgSal 数据集相当主观,没有明确的方法衡量显著区域的大小,选择图像的方法不科学.针对此问题,本文采用定量的衡量方法,更加科学合理.

背景先验已经多次用在显著性计算中.测地显著性<sup>[13]</sup>是代表工作之一,主要依据是图像的边界区域更可能为图像的背景.然而,当显著区域与图像边界相连时算法将不再有效.逆向思考,可以认为显著区域连接边界增加了显著区域的提取难度.

众多的显著区域提取文献都把对比度看做计算显著性的关键.当前景区域和整幅图像差异较小时,自然增大了显著区域提取的难度.所以在筛选图像时,可以通过设置显著区域和整幅图像的对比度阈值来提高数据集的难度.

1.2 筛选原则

基于上面的上分析,确定了 3 个筛选图像原则.

1) 显著区域占图像的比例.将显著区域占整幅图像的比例划分为 10 个等级:[0,0.1),[0.1,0.2),[0.2,0.3),[0.3,0.4),[0.4,0.5),[0.5,0.6),[0.6,

0.7),[0.7,0.8),[0.8,0.9),[0.9,1.0].显著区域占整幅图像的比例等级覆盖越多,则认为显著区域的尺寸越丰富.

2) 显著区域与边界的连接程度.在筛选时,设置数据集与图像边界的最小连接比例,保证数据集与边界的连接比例.

3) 显著区域和图像的对比度.在筛选时,设置数据集中显著区域与图像对比度阈值,保证数据集的提取难度.数据集中的图像筛选是个反复迭代的过程,目前还不存在完美和最优的结果,只能人为控制数据集图像筛选结束的条件.

## 2 数据集的性能统计

本文制定了 3 种数据集的性能统计方法.

图像  $I$  对应的二值标注图为  $G$ . $G$  中彼此不连通的显著区域个数为  $M$ . $x_i$  代表图像  $I$  中第  $i$  块显著区域,  $1 \leq i \leq M$ .

1) 统计显著区域占整幅图像比例等级的百分比.显著区域的面积占整幅图像的比例划分为 10 个比例等级.若  $x_i$  在等级  $j$  中,则等级  $j$  内显著区域个数加 1, $n_j = n_j + 1$ ,  $1 \leq j \leq 10$ .对数据集中的每一幅图像进行上面的操作,最后计算 10 个比例等级内显著区域个数占有显著区域个数的百分比.

2) 统计与图像边界相连的显著区域占有显著区域的比例. $x_i$  代表图像  $I$  中第  $i$  块显著区域,判断  $x_i$  是否和图像边界连接.如果连接,则与边界连接的显著区域个数加 1.对数据集中的每一幅图像进行上面的操作,最后计算与图像边界相连的显著区域个数以及占数据集中所有显著区域的比例.

3) 统计显著区域和整幅图像的颜色差均值.提取图像  $I$  的颜色特征. $x_i$  代表图像  $I$  中第  $i$  块显著区域,提取  $x_i$  的颜色特征,并计算  $x_i$  的颜色特征与图像  $I$  的颜色特征之间的差异.对数据集中的每一幅图像进行上面的操作,最后计算显著区域和整幅图像的颜色特征差的均值.

## 3 数据集的构建

### 3.1 图像来源

新加坡国立大学建立了 NUS-WIDE 图像数据集<sup>[14]</sup>;包含 269 648 幅图像,选自于 Flickr 网站;提供了 81 个基准标签集.本文随机选取了 10 000 幅图像,图像来自于 NUS-WIDE 数据集的 38 个文件夹,包括 carvings、castle、cat、cell\_phones、chairs、chrysanthemums、classroom、cliff、computers、cooling\_tower、coral、cordless cougar、courthouse、

cow、coyote、dance、dancing、deer、den、desert、detail、diver、dock、close-up、cloverleaf、cubs、dall、dog、dogs、fish、flag、eagle、elephant、elk、f-16、facade 和 fawn.

### 3.2 图像的显著区域二值标注

5 个观察者对 10 000 幅图像进行像素级别的二值标注.为减少标注的不一致性,计算每个像素标注的一致性分值,像素是否为显著的阈值,计算方法为

$$s_x = \frac{\sum_{p=1}^N a_x^p}{N} \tag{1}$$

式中:  $a_x^p$  为第  $p$  个观察者对像素  $x$  的标注,若像素标注为显著,则  $a_x^p = 1$ ,否则为 0; $N$  为观察者的个数.

根据文献[5],如果一个像素有 50% 的观察者标注为显著,则认为其是显著的.使用 Adobe Photoshop 手动的从图像中分割出显著区域.

### 3.3 图像的筛选

对于 10 000 幅图像进行筛选,筛选过程如下.

- 1) 显著区域和整幅图像的颜色对比度小于 0.7;
- 2) 显著区域占整幅图像的比例要求 10 个比例等级全部覆盖;
- 3) 至少有 10% 的图像的显著区域和图像边界相连.

经过 5 次筛选之后,最终选定了 5 429 幅图像构成显著性数据集.

### 3.4 数据集的性能统计

为验证新建数据集的性能,对新建数据集和当前流行的 7 个显著性数据集进行统计分析.新建的数据集简称为 TBD(Tags based Dataset).7 个流行的显著性数据集为 ECSSD<sup>[10]</sup>、ASD<sup>[5]</sup>、MSRA5k<sup>[8]</sup>、MIT<sup>[15]</sup>、ImgSal<sup>[9]</sup>、MSRA10k<sup>[7]</sup>、DUT\_OMRON<sup>[6]</sup>.

统计与图像边界相连的显著区域占有显著区域的比例结果如表 1 所示.可以看出,TBD 数据集中显著区域与图像边界连接的比例较高,仅低于数据集 ECSSD<sup>[10]</sup>.

统计显著区域和整幅图像的颜色差均值结果如表 2 所示.众所周知,差异值越小,显著区域检测越困难.TBD 数据集的差异是最小的,对于显著区域的提取工作难度是最大的.

统计数据集中显著区域占整幅图像的百分比结果如表 3 所示.TBD 数据集和数据集 MIT<sup>[15]</sup> 是非常相似的,包含了各种尺寸的显著区域,尺寸分布最广泛,而有些数据集<sup>[5,7-9]</sup> 的显著区域在某个等级范围内就不存在了.

表 1 与图像边界连接的显著区域的比例  
Tab. 1 Proportions of salient regions which are connected with image boundaries

数据集	显著区域的总数	与图像边界连接的显著区域数量	比例/%
TBD	7 104	1 193	17.0
ECSSD	1 171	225	19.0
ASD	1 209	18	1.0
MSRA5k	5 594	191	3.0
MIT	1 015	107	11.0
ImgSal	480	1	0.2
MSRA10k	6 841	556	8.0
DUT_OMRON	10 915	590	5.0

表 2 显著区域和整幅图像的颜色差均值  
Tab. 2 Average color differences between salient regions and images

数据集	平均 RGB 颜色特征差(归一化后)
TBD	0.000 1
ImgSal	0.301 4
ECSSD	0.045 5
ASD	1.000 0
MSRA5k	0.559 6
MIT	0.183 1
DUT_OMRON	0.457 4
MSRA10k	0.689 8

表 3 10 个等级范围内显著区域的比例  
Tab.3 Proportions of salient regions in each size grade %

比例等级	TBD	ImgSal	ECSSD	ASD	MSRA5k	MIT	DUT_OMRON	MSRA10k
[0,0.1)	51.09	87.50	25.62	31.35	23.74	55.07	59.05	18.05
[0.1,0.2)	16.67	9.37	26.47	31.76	33.88	16.95	20.14	33.31
[0.2,0.3)	11.95	2.71	19.21	24.23	25.35	10.44	11.68	27.92
[0.3,0.4)	8.26	0.42	10.16	10.92	13.25	6.70	5.82	16.01
[0.4,0.5)	5.38	0	6.75	1.57	3.22	5.22	2.51	4.26
[0.5,0.6)	3.43	0	4.61	0.08	0.46	2.96	0.67	0.42
[0.6,0.7)	1.93	0	1.96	0	0.05	1.38	0.10	0.02
[0.7,0.8)	0.91	0	1.96	0.08	0.05	0.89	0	0.01
[0.8,0.9)	0.32	0	2.22	0	0	0.30	0.01	0
[0.9,1.0]	0.04	0	1.02	0	0	0.10	0	0

3.5 TBD 数据集的标签统计

此节对 TBD 数据集内的标签信息进行统计. NUS-WIDE 数据集提供了一个包含 81 个标签的集合,TBD 显著性数据集的标签集合来自于这 81 个标签的集合.TBD 数据集中出现的标签总数为 78. 每幅图像包含 1~9 个标签.标签大体可以分为两类:场景标签和对象标签.场景标签表示场景,也就是环境;对象标签描述的是实体对象,例如狗、猫、人和鸟等.其中,对象和显著区域有着密切的关系.部分标签出现的频率如图 1 所示.

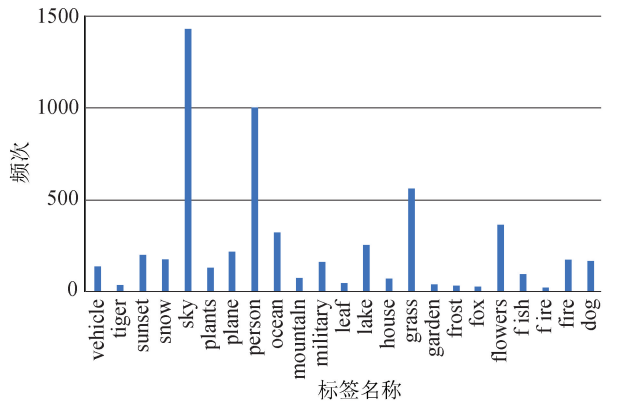


图 1 部分标签及出现的频次统计图  
Fig.1 Each tag and its frequency

3.6 数据集的典型图像

TBD 数据集中,图像可以包含多个显著区域,显著区域不一定位于图像的中心,显著区域的尺寸非常丰富,并可能与图像的边界相连.典型图像及其二值标注和带有的标签信息如图 2~图 4 所示.

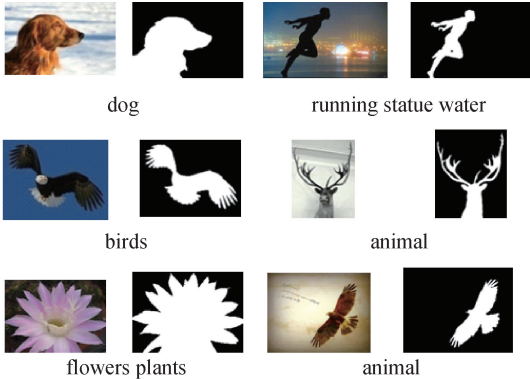


图 2 包含一个显著区域的图像示例  
Fig.2 Example images with one salient region

图 2 列举的图像只包含了单个的显著区域;图 3 列举的图像中包含多个显著区域;图 4 列举的图像中覆盖了显著区域的所有比例等级,显著区域的大小尺寸非常丰富.在图 2~图 4 中,有的图像的显著区域和边界相连,如图 2 中标签为 dog 的图像,图 3 中标签为 animal cat 的图像;有的图像的显著区域和边界不相连,如图 3 中标签为 animal birds

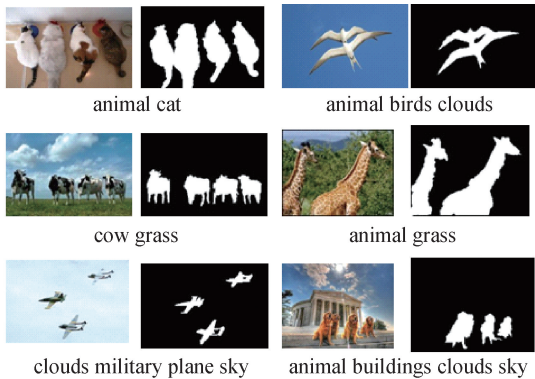


图 3 包含多个显著区域的图像示例

Fig.3 Example images with multi salient regions

显著区域 比例等级	图像显著区域二值标注和标签					
[0,0.1)			flowers			flowers plants
[0.1,0.2)			animal birds clouds			animal clouds sky
[0.2,0.3)			animal cat			animal coral fish water
[0.3,0.4)			flowers leaf			animal birds
[0.4,0.5)			flowers			animal tiger
[0.5,0.6)			animal birds			person
[0.6,0.7)			flowers plants			animal tiger
[0.7,0.8)			flowers			animal tiger
[0.8,0.9)			flowers			animal
[0.9,1.0]			animal cat			flowers

图 4 包含各种尺寸显著区域的图像示例

Fig.4 Typical images with salient regions in all size levels

clouds 的图像,图 4 中的标签为 animal coral fish water 的图像。

## 4 结论

本文围绕显著性数据集展开研究,提出了建立数据集时图像的筛选原则以及数据集的性能统计方法,构建了带有标签信息的显著性数据集.此数据集选自 NUS-WIDE 数据集中的 38 个类别,共 5 429 幅图像,包含 78 个不同的标签.数据集中的图像显著区域不限于一个;显著区域的尺寸涵盖范围非常丰富.此外,为了衡量数据集的性能,对新建数据集和 7 个流行数据集进行了统计分析.与目前流行显著性数据集的最大区别是 TBD 数据集带有标签信

息,为新的显著区域提取方法提供了实验对象。

本文下一步的工作包括:1)目前数据集在构建的时候并没有根据标签进行筛选,需要进一步设计标签的筛选原则;2)目前的筛选原则并不完善,有待补充,比如增加衡量中心先验的筛选原则;3)显著性数据集的性能如何评价,设计显著性数据集的性能评价方法有助于指导显著性数据集的科学构建。

## 参考文献 (References):

[1] ZHANG G X, CHENG M M, HU S M, et al. A shape-preserving approach to image resizing[C]// Proceedings of Computer Graphics Forum, 2009, 28 (7): 1897—1906.

[2] ZHANG W, BORJI A, WANG Z, et al. The application of visual saliency models in objective image quality assessment: a statistical evaluation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27 (6):1266—1278.

[3] XIAO C M, SHI Z L, XIA R B, et al. Edge-detection algorithm based on visual saliency[J]. Information and Control, 2014, 43 (1) : 9—13.

[4] LIU T, SUN J, ZHENG N N, et al. Learning to detect a salient object[C]// Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2007: 1—8.

[5] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection[C]// Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2009: 1597—1604.

[6] YANG C, ZHANG L H, LU H C, et al. Saliency detection via graph-based manifold ranking[C]// Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2013: 3166—3173.

[7] CHENG M M, MITRA N J, HUANG X, et al. Global contrast based salient region detection[C]// Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2011: 409—416.

[8] WANG J D, JIANG H A, YUAN Z J, et al. Salient object detection: a discriminative regional feature integration approach[J]. International Journal of Computer Vision, 2017, 123(2): 251—268.

[9] LI J, LEVINE M D, AN X J, et al. Visual saliency based on scale-space analysis in the frequency domain [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(4): 996—1010.

[10] YAN Q, XU L, SHI J P, et al. Hierarchical saliency detection[C]// Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition,

[11] LI G B, YU Y Z. Visual saliency based on multiscale deep features[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015:5455–5463.

[12] 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 IEEE International Conference on Computer Vision, 2001, 2: 416–423.

[13] WEI Y C, WEN F, ZHU W J, et al. Geodesic saliency

[14] CHUA T S, TANG J H, HONG R C, et al. NUS-WIDE: a real-world web image database from National University of Singapore[C]// Proceedings of the International Conference on Image and Video Retrieval, 2009: 48–56.

[15] JUDD T, EHINGER K, DURAND F, et al. Learning to predict where humans look[C]// Proceedings of IEEE International Conference on Computer Vision, 2010: 2106–2113.



(上接第 98 页)

[11] VAPNIK V. SVM method of estimating density, conditional probability, and conditional density[C]// IEEE International Symposium on Circuits and Systems, 2000:749-752.

[12] 王天健, 吴振升, 王晖, 等. 基于最小二乘支持向量机的改进型 GIS 局部放电识别方法[J]. 电网技术, 2011, 35(11): 178-182.

WANG Tianjian, WU Zhensheng, WANG Hui, et al. An improved approach to recognize partial discharge in GIS based on minimum least square-support vector machine[J]. Power System Technology, 2011, 35(11): 178-182. (in Chinese)

[13] 荣海娜, 张葛祥, 金炜东. 系统辨识中支持向量机核函数及其参数的研究[J]. 系统仿真学报, 2006, 18(11): 3204-3208.

RONG Haina, ZHANG Gexiang, JIN Weidong. Selection of kernel functions and parameters for support vector machines in system identification[J]. Journal of System Simulation, 2006, 18(11): 3204-3208. (in Chinese)

[14] RAKOTOMAMONJY A, BACH F, CANU S, et al. SimpleMKL [J]. Journal of Machine Learning Research, 2008, 9(3): 2491-2521.

[15] RAKOTOMAMONJY A, BACH F, GRANDVALET

Y. More efficiency in multiple kernel learning[C]//  
Proceedings of the 24th International Conference on  
Machine Learning, 2007: 775–782.

[16] 韩磊,王立威,郑艳清.GIS 典型缺陷的局部放电超高频检测及模式识别[J].内蒙古电力技术,2015,33(1):7-12.

HAN Lei, WANG Liwei, ZHENG Yanqing. UHF detection and pattern recognition of partial discharge on GIS typical defects[J]. Inner Mongolia Electric Power, 2015, 33(1): 7-12. (in Chinese)

[17] 张波.基于超声波法的 GIS 局部放电模式识别的研究 [D].保定:华北电力大学,2015.

ZHANG Bo. Research on partial discharge pattern recognition of GIS based on ultrasonic detection [D]. Baoding: North China Electric Power University, 2015. (in Chinese)

[18] 唐炬,杨东,曾福平,等.基于分解组分分析的 SF<sub>6</sub> 设备绝缘故障诊断方法与技术的研究现状[J].电工技术学报,2016,31(20):41—54.

TANG Ju, YANG Dong, ZENG Fuping, et al. Research status of SF<sub>6</sub> insulation equipment fault diagnosis method and technology based on decomposed components analysis [J]. Transactions of China Electrotechnical Society, 2016, 31(20): 41–54. (in Chinese)