

基于背景原型对比度的显著性物体检测

罗辰辉¹, 张 伟¹, 沈琼霞², 叶 波¹

(1. 湖北大学 计算机与信息工程学院, 武汉 430062; 2. 烽火通信科技股份有限公司, 武汉 430073)

摘要: 针对传统显著性模型在自然图像的显著性物体检测中存在的缺陷, 提出了一种利用背景原型 (background prototypes) 进行对比的视觉关注模型, 以实现显著性物体的检测与提取; 传统显著性模型主要通过计算区域中心与四周区域差异性实现显著性检测, 而自然场景中显著性区域和背景区域往往都存在较大差异, 导致在复杂图像中难以获得理想检测效果; 基于背景原型对比度的显著性物体检测方法在图像分割生成的超像素图基础上, 选择距离图像中心较远的图像区域作为背景原型区域, 通过计算图像中任意区域与这些背景原型区域的颜色对比度准确检测和提取图像中的显著性物体; 实验结果表明, 基于背景原型对比度的显著性模型可以更好地滤除杂乱背景, 产生更稳定、准确的显著图, 在准确率、召回率和 F-measure 等关键性能和直观视觉效果上均优于目前最先进的显著性模型, 计算复杂度低, 利于应用推广。

关键词: 视觉关注; 背景原型; 视觉显著度

Saliency Detection via Background Prototypes Contrast

Luo Chenhui¹, Zhang Wei¹, Shen Qiongxia², Ye Bo¹

(1. School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China;

2. Service and CPE Business Unit, Fiberhome Telecommunication Technologies Co. Ltd., Wuhan 430073, China)

Abstract: To overcome the disadvantages of existing saliency models in saliency detection, a novel object-based attention model is presented to predict visual saliency using the contrast against the background prototypes. Traditional saliency models mainly detect salient regions by comparing the differences between center and surround regions, which makes hard to get desired results in complex scenes for significant differences often appear both in salient and background regions in real images. Saliency detection via background prototypes contrast firstly over-segment the input image into perceptually homogeneous superpixels, and automatically identifies a series of regions far away from image center as background prototypes. The visual saliency is then accurately calculated using the color contrast with respect to the selected background prototypes. Promising experimental results demonstrate that the proposed model, which outperforms the compared state-of-the-art saliency models in average precision, recall, F-measure and visual effect, can better exclude the cluttered backgrounds, and thus produces more robust and accurate saliency maps. Moreover, due to its computational efficiency, our model is easy to be widely applied.

Keywords: visual attention; background prototype; visual saliency

0 引言

人类注意力系统 (human attention system, HAS) 具备从一个既定场景中迅速定位兴趣目标的能力, 这也是人类的生物视觉注意力相对普通光学设备的主要优势之一。因此, 如何在现有光学设备控制系统中模拟实现 HAS, 使光学镜头能够自动聚焦并捕获目标对象, 成为一个重要的研究课题。由于这些光学设备主要为基于嵌入式系统设计的便携式设备, 系统集成的计算资源有限 (一般所用的处理器为嵌入式处理器, 性能不高, 并且内存容量也有限), 难以进行大量高速运算, 这就要求所设计的显著性模型必须在保证准确度和可靠性的同时, 也要具有较低的复杂度, 易于实现。性能需求与计算复杂度的矛盾, 是图像显著区域检测的主要困难, 近年来众多研究人员在

此问题上付出了巨大的努力, 取得了一些有效成果^[1-7]。

目前显著性模型主要分为两类: 自顶向下模型^[8-9]和自底向上模型^[10-14]。自顶向下的显著性模型一般通过学习得到的高层语义信息 (如人脸检测) 来辅助进行显著性物体检测。由于这类方法在学习过程中需要耗费大量的计算资源, 因此很难应用到光学设备中。而自底向上的显著性模型主要利用简单的图像特征信息 (例如亮度、颜色、方向等) 来计算前景显著度。这类方法复杂度较低, 更适于移植到光学设备中。Itti 等^[10]率先进行了相关研究, 提出基于中心-四周对比度的显著性模型。文献 [11-12] 基于马尔可夫随机理论提出了基于图模型的显著性度量算法。还有一些研究人员通过检测频域的不规则性来判断视觉显著性^[1, 13-14]。这些方法在图像显著性物体检测问题上获得了一定成效。然而, 自然场景中显著性区域和背景区域往往都具有很大的差异性, 传统的显著性模型很难在复杂的自然图像中获得较好的检测效果。图 1 对传统显著性模型 (包括 CA^[9]、FT^[14]、GB^[12]、IT^[11]、MS^[13] 模型等) 检测效果进行了对比。可以看出, 现有方法都受到了背景杂乱或对象形状信息丢失的影响, 在显著性物体检测中存在明显缺陷: CA 模型可以较好地检测到显著性区域的边界, 但对于显著性物体的形状信息则损失严重; FT 模型除了检测到图像的显著区域, 还将部分复杂背景也误检为显著物体; GB 模型和

收稿日期: 2017-06-19; 修回日期: 2017-07-07。

基金项目: 国家自然科学基金 (61301144, 61601175)。

作者简介: 罗辰辉 (1991-), 男, 湖北黄冈人, 研究生, 主要从事图像处理、物联网、无线通信方向的研究。

通讯作者: 张 伟 (1979-), 男, 湖北武汉人, 博士, 讲师, 硕士研究生导师, 主要从事图像处理、无线通信方向的研究。

IT 模型未能将显著物体与图像背景进行准确区分, 导致检测结果出现明显失真; MS 模型在复杂背景下, 也将部分背景图案检测为显著性物体, 降低了检测准确性。可见, 在较低的计算复杂度下实现图像显著性物体的精准提取, 具有重要的理论研究价值和前景。

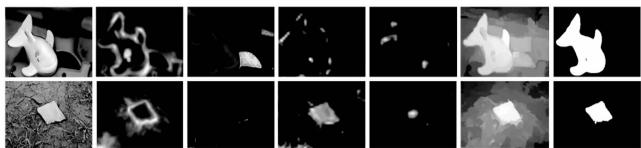


图 1 现有显著性模型缺点图示说明。从左到右分别为: 原始图像, 通过 CA、FT、GB、IT、MS 和本文模型获得的显著图。

传统的显著性模型一般基于一个生物学研究结果, 即人类视觉注意力会集中于图像的中心区域^[15]。因此, 以往的模型往往通过计算图像区域中心与四周区域的差异性来进行显著性物体检测^[10-15]。因此, 当实际图像中显著性物体偏离图像中心时, 就可能造成与背景图案的混淆, 难以得到理想的检测结果。与传统的显著性检测模型不同, 本文提出了一种基于背景原型对比度的自底向上显著性模型。这里, “背景原型”指的是通过预分割算法^[16]获得的超像素图像区域。显著性建模的另一个重要问题是如何定量测量视觉显著度。在本文的模型中, 由于显著区域与背景存在较大的外观差异, 因此可以通过计算与背景原型的颜色对比差异判断显著区域。图 1 同时给出了本文模型获得的显著图结果。如图 1 所示, 相对传统显著性模型的检测结果, 本文提出的模型可以更好地滤除杂乱背景, 产生更稳定、准确的显著图。此外, 该模型计算复杂度低, 适合应用于光学器件中。

1 原理与方法

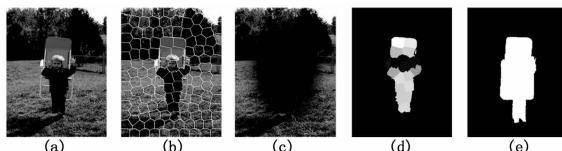
显著性物体检测的关键问题是如何准确区分背景区域和显著性目标。传统的显著性模型基于区域中心与四周的差异度计算, 在显著性物体偏离中心区域或背景图案复杂时区分效果不佳, 造成检测精度降低。为了更精确地检测目标与背景, 本文基于背景原型对比度建立显著性模型: 先分析图像计算产生背景原型, 再通过比较原始图像与背景原型的特征信息比对确定显著性目标物体。由于邻近区域的像素存在较大冗余性, 基于像素分析背景原型计算量大并且欠缺必要性, 因此本文基于超像素进行背景原型图的分析计算。同时, 由于背景原型与显著目标存在较大颜色差异, 选择颜色作为比对的特征信息。

基于背景原型对比度的显著性物体检测方法实施可以分为 3 个步骤:

- 1) 图像分割, 通过分割算法被分解成为一些同质的超像素;
- 2) 选择背景原型, 参考在整幅图像中的相对位置, 远离图像中心的区域被选择为背景原型;
- 3) 显著性度量, 通过与背景模型的颜色对比来判断出视觉显著性区域。

图 2 是对本文方法实施过程的例子。首先, 对原始图像 (a) 进行超像素分割, 得到超像素分割图 (b); 再计算超像素位置值确定背景原型图 (c); 最后, 通过与背景模型的颜色对比得到显著图 (d); (e) 为对显著物体人工标注结果。可以看

出, 基于背景原型对比度的显著性物体检测方法得到的显著图与人工标注结果接近度高, 效果理想。下面, 对显著性对象检测模型进行详细分析说明。



(a) 原始图像; (b) 超像素分割图 (超像素以白色边界线区分); (c) 视为背景的图像边缘区域; (d) 显著图; (e) 人工标注结果
图 2 方法过程概述

1.1 图像分割

基于背景原型对比度的显著性物体检测方法第一步是对原始图像进行超像素分割。超像素分割将图像分析过程由基于像素转换为基于区域, 可以显著降低图像分析复杂度, 提升算法速度。本文使用文献 [16] 中的方法把输入图像分割成超像素。这个分割算法有两个重要参数, 即 $rgnSize$ 和 $regularizer$, 分别表示分割图像的超像素数和空间规律的平衡表。使用超像素分割技术可以将相似区域融合起来的同时避免破坏目标物体的边界特性。由于颜色直方图是一个鲁棒的全局外观特征描述子, 因此本文选用颜色直方图作为超像素的表示特征。为了保证感知精度, 直方图在 CIELab 颜色空间进行计算。在此基础上, 第 i 个超像素 R_i 则可由一个 K 维颜色直方图 h_i 来描述, 其中第 k 维分量记作 $h_i(k)$ 。

1.2 选择背景原型

在超像素分割图基础上, 选择准确的背景原型。根据文献 [15], 在像素 x 到图像中心 c 的欧式距离 $d(x, c)$ 基础上利用逆高斯分布计算生成一个预处理图 (即图 3 的中间图):

$$p(x) = 1 - \exp\{-d(x, c)/\sigma^2\} \quad (1)$$

这里 σ^2 一般设为经验值 1。如图 3 所示, 测试图像经过预处理图 $p(x)$ 加权处理, 并且定义超像素 R_i 的位置值 s_i 为所含像素的平均位置值。当超像素位置值高于预定义的阈值 θ , 即被确认为背景原型, 这些超像素的集合记作 $R^B = \{R_1^B, R_2^B, \dots, R_M^B\}$, 其中:

$$R_i^B = R_i, \text{ if } s_i \geq \theta \quad (2)$$

θ 一般取经验值 0.8。靠近图像边界的区域往往会被优先选择, 这样处理有利于囊括不同特点的背景区域。

1.3 基于背景原型的显著性度量

最后, 基于背景原型, 将各个超像素区域通过全局颜色对比算法^[17]来获得显著图。在这里, 本文通过计算超像素到所选背景原型的 χ^2 距离来判断显著区域, 而不是对整个图像场景进行对比计算, 算法复杂度低, 易于实现。 R_i 与背景原型 $R_j^B \in R^B$ 之间的差异度可通过以下公式计算:

$$C_{ij}(R_i, R_j^B) = \frac{1}{2} \sum_{k=1}^K \frac{(h_i(k) - h_j^B(k))^2}{h_i(k) + h_j^B(k)} \quad (3)$$

其中: $h_i(k)$ 是区域 R_i 的颜色直方图, $h_j^B(k)$ 是区域 R_j^B 的颜色直方图, K 表示直方图 bin 的个数。

尽管显著性区域有时也会具有多样性 (如图 2、图 3、图 4 所示), 但这些区域的颜色依然和背景原型有着较大区别。此外, 背景区域常常与背景原型的有着类似的颜色分布。因此, R_i 的显著图应为所有背景原型的最小值:

$$S(R_i) = \min_j C_{ij} (R_i, R_j^B) \quad (4)$$

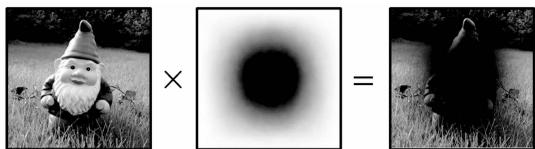


图 3 背景原型的产生图示

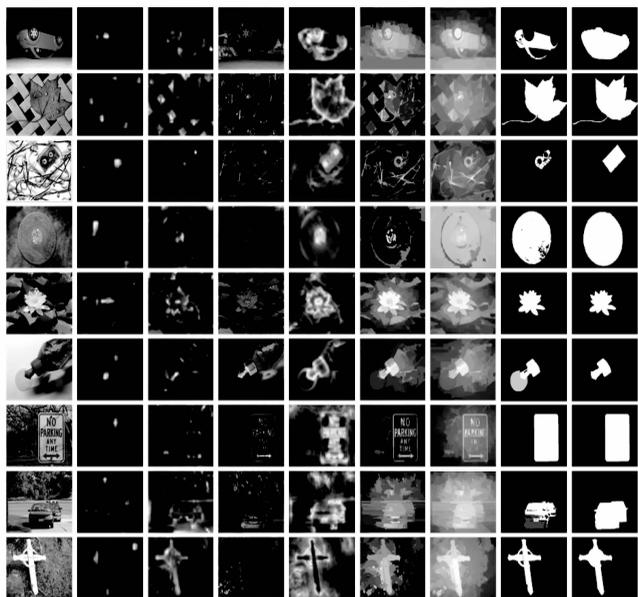


图 4 本文方法与已有方法的视觉效果对比。从左到右分别为: 原始图像、IT、GB、FT、CA、RC、MS、本文方法和人工标注的显著图

2 实验及性能比较

本节通过对比实验证明基于背景原型对比度的显著性物体检测方法的有效性。为了增强实验结果的说服力, 本节实验均在文献 [13, 17] 给出的数据集上进行, 同时选择以下六个具有代表性的显著性模型作为比较的对象, 即视觉关注测量模型 (visual attention measure, IT^[10])、基于图的显著性模型 (graph-based saliency, GB^[11])、频谱调谐显著性模型 (frequency-tuned saliency, FT^[13])、上下文感知模型 (context-aware saliency, CA^[8])、整体对比显著性模型 (global-contrast saliency, RC^[17]) 和中层视觉显著性模型 (midlevel saliency, MS^[12])。实验通过在相同数据集上运行本方法及各对比方法获得显著图, 采集数据进行定量性能对比分析及定性效果对比分析。

实验中本文方法进行超像素分割的参数 (rgnSize, regularizer) 设为 (30, 1.0), 每个颜色直方图中的超像素为 $K = 60$ 维的向量。对于 IT^[10], GB^[11], FT^[13] 和 RC^[17] 模型, 本文通过实现其算法运行获得最终的显著性图, 从而得到相关比较数据, 对于 CA^[8] 和 MS^[12] 模型, 本文直接利用作者公布的显著性图做性能比较。

2.1 性能定量分析

实验对各个显著性模型复杂度、准确率、返回率和 F-measure 等性能参数进行测量比较。

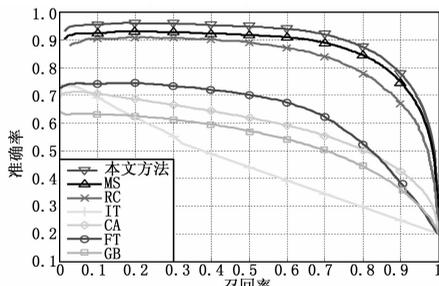
通过对在相同硬件平台上, 对相同图像的算法运行时间的比较, 可以直观反映出计算复杂度高低。表 1 是这几种方法平均运行时间的对比结果。

表 1 平均运行时间的对比
(图像的分辨率为 400×300)

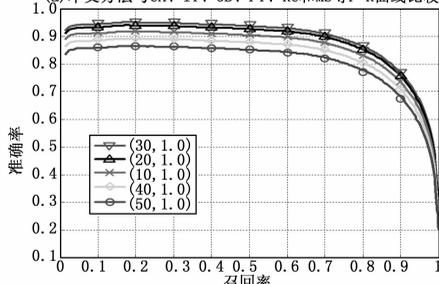
算法	IT	GB	FT	RC	本文方法
时间/s	0.611	1.614	0.016	0.253	0.174
代码	Matlab	Matlab	C++	C++	Matlab

通过表 1 可以看出, FT 模型平均运行时间最短, 复杂度最低; 除了 FT 模型, 本文方法平均运行时间均少于其它方法, 说明其为计算复杂度较低的模型之一。结合后面的效果对比实验结果, 本文方法获取的显著图明显优于 FT 模型, 综合而言, 本文方法具备较低的复杂度, 较高的计算效率。

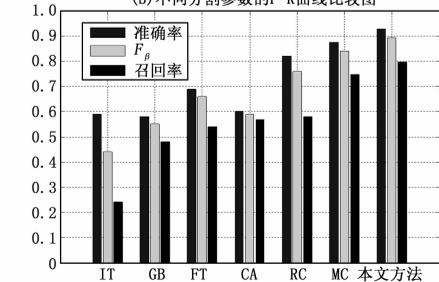
本文采用文献 [13, 17] 中的方法对以上算法准确率、返回率和 F-measure 等性能参数进行量化比较, 结果如图 5 所示。实验中, 归一化显著图分割阈值变化范围为 $0 \sim 255$ 。图 5 (a) 为各显著性模型精度-召回曲线。可以看出, 本文提出的模型性能明显优于其它模型。本文方法中第一步对图像进行超像素分割是非常关键的, 而分割参数对于分割结果乃至最终显著性物体检测结果有较大影响。图 5 (b) 为采用不同分割参数获得的结果, 表明本文采用的参数设置是性能最佳的。在最佳分割参数下, 本文对通过自适应阈值得到的显著图进行二元分割



(a) 本文方法与 CA、IT、GB、FT、RC 和 MS 等 P-R 曲线比较图



(b) 不同分割参数的 P-R 曲线比较图



(c) 最佳分割阈值下的平均准确率、返回率和 F-measure

图 5 MSRA 1000 数据集的定量比较

计算 F-measure, 公式中 P 表示准确率, R 表示召回率, $\beta^2 = 0.3^{[13-17]}$ 。各模型方法计算的平均准确率、返回率和 F-measure 等实验结果如图 5 (c) 所示。可以看到, 本文提出的方法无论在准确率、返回率还是 F-measure, 均达到最好的性能。相比于其他方法, 本文方法对返回率性能的提升效果最为显著, 这表明本文方法可以准确地检测出更多的显著区域。

2.2 效果定性对比

通过对各种方法产生的显著图结果进行二元分割后, 可以对其视觉效果进行直接对比, 判断方法的准确度。图 5 展示了用不同方法获得的视觉显著图, 以便进行定性的效果比较。由图中结果可以看出, 本文的模型能够产生高分辨率的显著图, 是各个方法中最接近于人工标注显著图的。相比之下, 其他方法产生的显著图均存在明显不足之处: 有的方法提取的显著对象形状出现失真, 如 IT 和 GB; 有的过分强调显著对象的边缘轮廓, 如 CA; 有的在不适于从杂乱复杂的背景中难以达到满意的效果, 如 FT、RC 和 MS。

定量性能对比及定性效果对比实验结果说明基于背景原型对比度的显著性物体检测方法在较低的计算复杂度下获得了高效、准确的显著图, 具有较强的应用价值。

3 结论

本文在背景原型对比的基础上提出了一个新的显著性模型。实验结果表明, 它在准确率、召回率和 F-measure 等关键性能和直观视觉效果上均优于目前最先进的显著性模型。此外, 本方法计算复杂度低, 对硬件资源要求不高, 非常适于移植到软硬件资源有限的光学设备 (如数码相机) 上。

参考文献:

[1] Xu Y, Zhao Y, Jin C, et al. Salient target detection based on pseudo-wigner-ville distribution and renyi entropy [J]. 2010, 35: 475-477.

[2] Gao C, Sang N, Huang R. Instance-based attention: where could humans look first when searching for an object instance [J]. 2012, 37: 76-78.

[3] Vazquez E, Gevers T, Lucassen M, et al. Saliency of color image derivatives: a comparison between computational models and human perception [J]. J. Opt. Soc. Am., 2010; 613-621.

[4] 张永梅, 吴攀, 李炜. 一种图像敏感对象的识别方法 [J]. 计算机测量与控制, 2014, 22 (7): 2167-2169.

[5] Ko B, Nam J. Object-of-interest image segmentation based on human attention and semantic region clustering [J]. J. Opt. Soc. Am., 2006, 23: 2462-2470 (2006).

[6] Kim W, Kim C. Saliency detection via textural contrast [J]. Opt. Lett., 2012, 37: 1550-1552.

[7] Itti L, Gold C, Koch C. Visual attention and target detection in cluttered natural scenes [J]. Opt. Eng. 2001, 40: 1784-1793.

[8] Goferman S, Zelnik-Manor L, Tal A. Context aware saliency detection [J]. IEEE Trans. Pattern Anal. Machine Intell., 2012, 34: 1915-1926.

[9] Torralba A. Modeling global scene factors in attention [J]. J. Opt. Soc. Am., 20, 2003, 20: 1407-1418.

[10] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Trans. Pattern Anal. Machine Intell., 1998, 20: 1254-1259.

[11] Harel, J, Koch C, Perona P. Graph-based visual saliency [D]. Proceedings of the Advances in Neural Information Processing Systems, MIT, 2006.

[12] Yu J, Tian J. Saliency detection using midlevel visual cues [J]. Opt. Lett., 2012, 37: 4994-4996.

[13] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned saliency region detection [A]. Proceedings of the IEEE conference on Computer Vision and Pattern Recognition [C]. IEEE, 2009: 1597-1604.

[14] Li C, Xue J, Tian Z, et al. Saliency detection based on biological plausibility of hyper-complex fourier spectrum contrast [J]. Opt. Lett., 2012, 37: 3609-3611.

[15] Tatler B. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions [J]. J. of Vis., 2007, 7: 1-17.

[16] Achanta R, Shaji A, Smith K, et al. Slic superpixels [R]. EPEL, Tech. Rep, 149300, 2010.

[17] Cheng M, Zhang G, Mitra N, et al. Global contrast based salient region detection [A]. Proceedings of the IEEE conference on Computer Vision and Pattern Recognition [C]. IEEE, 2011: 409-416.

[18] 汪永伟, 赵荣彩, 常德显, 等. 一种基于改进证据理论的推理决策方法 [J]. 计算机测量与控制, 2014, 22 (7): 2167-2169.

[19] 朱庆, 徐胜华, 韩李. 基于 D-S 证据理论的彩色航空影像阴影提取方法 [J]. 自动化学报, 2007, 33 (6): 588-595.

[20] Zadeh L A. Review of Shafer's a mathematical theory of evidence [J]. AI Magazine, 1984, 5: 81-83.

[21] 韩德强, 杨艺, 韩崇昭. DS 证据理论研究进展及相关问题探讨 [J]. 控制与决策, 2014, 29 (1): 1-11.

[22] 杨艺, 韩德强, 韩崇昭. 基于多准则排序融合的证据组合方法 [J], 自动化学报, 2012, 38 (5): 823-831.

[23] 崔家玮, 李弼程. 一种新的 Dempster 组合规则适用条件 [J]. 计算机科学, 2014, 41 (10): 286-290.

[24] 汪永伟, 赵荣彩, 常德显, 等. 一种基于改进证据理论的推理决策方法 [J]. 计算机科学, 2014, 41 (12): 24-29.

(上接第 250 页)

参考文献:

[1] Dempster A P. Upper and lower probabilities induced by a multivalued mapping [J]. Annals of Mathematical Statistics, 1967, 38 (2): 325-339.

[2] Dempster A P. A Generalization of Bayesian Inference [J]. Journal of the Royal Statistical Society, 2008, 30 (30): 73-104.

[3] Shafer G. A Mathematical Theory of Evidence [M]. Princeton: Princeton University Press, 1976.

[4] 吴亮, 彭圆, 闫祎. 基于 Dempster-Shafer 证据推理的水中目标多特征融合识别技术 [J]. 计算机测量与控制, 2013, 21 (9): 2503-2505.

[5] 邵俊宇, 向超, 曹玉华, 等. 无人机远程故障测试与诊断技术研究 [J]. 计算机测量与控制, 2015, 23 (9): 2950-2953.

[6] Chen N, Yu D, Xia B. Evidence theory based analysis for the pre-