

## 基于图像显著性的人工视觉图像处理策略

王静<sup>1,2</sup>, 张羽婷<sup>3</sup>, 张云<sup>1,2</sup>, 韩彦岭<sup>1</sup>, 洪中华<sup>1</sup>, 杨树瑚<sup>1</sup>

1. 上海海洋大学信息学院, 上海 201306; 2. 农业农村部渔业信息重点实验室, 上海 201306; 3. 中国移动通信集团上海有限公司, 上海 200082

**【摘要】目的:**在假体设备中的视觉信息处理模块引入适当的图像处理策略,优化低分辨率下的人工视觉信息,是解决假体植入者获取的视觉信息有限问题的一种可行方法。**方法:**基于全局亮度对比度特征的图像显著性检测算法(LC),结合颜色空间变换和视觉注意力仿真处理模型,提出一种面向视网膜假体人工视觉信息处理策略。通过两个标准图像测试数据集对图像处理结果进行评估,同时对流行病学调查结果选取的盲人常用物体图片进行仿真。**结果:**评估结果验证了在提取图像前景方面与原始LC算法对比的优越性,仿真结果验证了在人工视觉条件下应用的可行性。**结论:**本文提出的策略有助于视网膜假体植入者在日常生活场景中更好地完成物体识别等基本视觉任务。

**【关键词】**视网膜假体;人工视觉信息处理;图像显著性检测

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2019)11-1277-07

## Image processing strategy based on saliency detection for artificial vision

WANG Jing<sup>1,2</sup>, ZHANG Yuting<sup>3</sup>, ZHANG Yun<sup>1,2</sup>, HAN Yanling<sup>1</sup>, HONG Zhonghua<sup>1</sup>, YANG Shuhu<sup>1</sup>

1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Fisheries Information, Ministry of Agriculture and Rural Affairs, Shanghai 201306, China; 3. China Mobile Group Shanghai Co., Ltd, Shanghai 200082, China

**Abstract: Objective** To effectively solve the problem of limited visual information obtained by prosthetic implant users by introducing the appropriate image processing strategy into the visual information processing module of the prosthesis device for optimizing the low-resolution artificial vision information. **Methods** Based on the saliency detection algorithm (luminance contrast) of global luminance contrast features, a strategy of artificial visual information optimization for retinal prosthesis was proposed by combining color space conversion and visual attention simulation processing model. Two benchmark databases were used to evaluate the result of image processing. Meanwhile, the simulation of images of the frequently used objects for the blind people which were determined by epidemiological investigation was carried out. **Results** The evaluation results verified the superiority of the proposed strategy over original luminance contrast algorithm in extracting image foreground. The simulated result verified the feasibility of the proposed strategy in artificial vision. **Conclusion** The proposed strategy is helpful for retinal prosthetic implant users to better accomplish the basic visual tasks such as object recognition in daily life.

**Keywords:** retinal prosthesis; artificial vision information processing; image saliency detection

### 前言

在视网膜致盲疾病中,常见的原发性老年黄斑

变性与视网膜色素变性尚无有效的手术或药物治疗手段。通过视网膜假体这类植入式电子设备代替视网膜中的受损部位,以实现视觉功能的修复,成为目前原发性老年黄斑变性及视网膜色素变性病患复明的可行方法<sup>[1]</sup>。目前,其典型产品第二视觉公司的Argus@ II视网膜假体已经被美国食品药品监督管理局和欧洲药品管理局批准应用于临床,超过276个植入案例<sup>[2]</sup>。视网膜假体是由外置的微型摄像头作为图像采集模块采集外界图像信息,通过图像信息处理模块把图像转换为对应的电刺激模式,通过无线

**【收稿日期】**2019-06-11

**【基金项目】**国家自然科学基金(61806123, 41871325, 41506213);上海市青年科技英才扬帆计划(16YF1415700);上海青年讲师培训资助计划(ZZZZHY15001)

**【作者简介】**王静,博士研究生,讲师,主要研究方向:计算机视觉理论及应用、医学图像处理,E-mail: wangjing@shou.edu.cn

**【通信作者】**张云,博士研究生,教授,主要研究方向:通信导航遥感技术,E-mail: y-zhang@shou.edu.cn

传输传递给体内的微电流刺激器,再由微电流刺激器将多路电脉冲加载到微电极阵列,对视觉神经系统进行电刺激,在视觉中枢产生光幻视,从而实现视觉功能修复<sup>[3]</sup>。然而,受到目前存在的各种技术局限可植入电极数目的限制,以及电极阵列植入位置和复杂的刺激诱发条件的影响,在现阶段及较长的一段时期,视网膜假体所提供的视觉感受具有非常低的分辨率、有限的视野范围、较大的光幻视点间距、较小的可表达亮度动态范围、颜色信息无规则诱发及深度信息缺乏等特点,是一种由离散光幻视点构成的低分辨率的视觉感受。从而在低分辨率视觉感受的人工视觉情况条件下,面向植入者的日常视觉任务需求,通过一定的图像处理策略,提取并优化突出最为显著的场景信息,优化呈现光幻视阵列所表达的信息,已经成为视网膜视觉假体研究中的一个重要方面。

Wang等<sup>[4]</sup>在2016年针对人工视觉仿真提出一种基于显著性计算模型的物体提取与增强算法,仿真评估实验结果表明该算法可以有效将场景图像中的显著物体分割出来。针对图像信息的实时采集和优化问题,2017年Li等<sup>[5]</sup>提出基于Lab空间图像特征的全局显著度检测算法,并在此基础上结合图像边缘信息优化表达人工视觉,开展的眼手协调仿真实验评估了该模型的有效性。Guo等<sup>[6]</sup>在2018年提出一种应用于视觉假体优化视觉信息呈现的算法,开展

了物体识别仿真实验,通过被试识别物体的准确率评估该算法的可行性。针对人工视觉信息处理,目前研究基于一定显著性检测算法提出了相应的人工视觉优化表达算法,在提取和增强效果方面具有一定的可行性。考虑到人眼的视觉生理特性,人的每一只眼睛含有对亮度敏感的视杆细胞1.2亿个,含有对色彩敏感的视锥细胞500至700万个,视杆细胞的数量远大于视锥细胞,因此眼睛对亮度的敏感程度要大于对色彩的敏感程度<sup>[7]</sup>。基于此,本文在上述研究解决思路的基础上,引入图像的全局亮度对比度特征,并结合其他颜色空间下的亮度差异特征,提出基于全局亮度差异的显著性检测优化处理算法。同时,算法考虑了在目标识别中视觉注意力的仿生学特性,进一步提高图像中显著物体的检测效果。

1 方法

本策略选择RGB、YUV和HIS这3种典型的颜色空间,分别对图像进行处理。在颜色空间变换之后,本策略对原始彩色图像进行灰度化处理,利用原始LC算法生成3幅显著图。在均衡化之后,提出一种自适应增强策略,称为“视觉注意力仿真处理”,以进一步突出显著图,即显著图YUV、显著图HSI和显著图Gray。在不同权重下,对它们进行叠加,得到最终的显著图。最后进行低像素化处理,并且仿真低分辨率的视觉感受。图1为本策略的示意图。

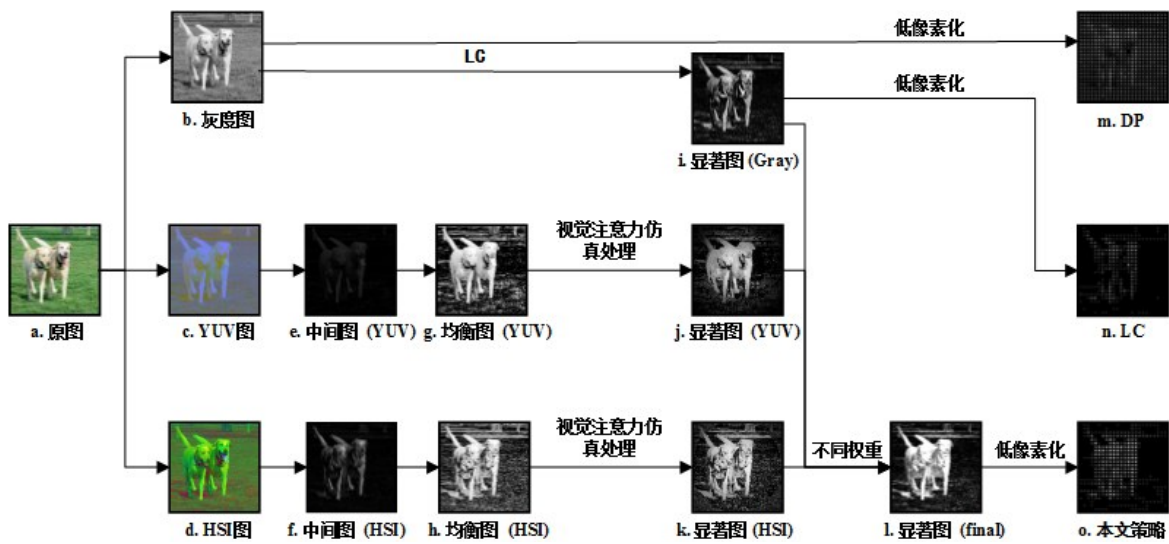


图1 基于原始LC的图像处理策略示意图

Fig.1 Schematic diagram of image processing strategy based on original luminance contrast

a是原始图像;b、c、d是颜色空间转换的结果;e、f为灰度变换和LC处理后的中间图;g和h是均衡化处理后的均衡图;m没有任何图像处理算法,即直接像素化(Direct Pixelization, DP)

1.1 颜色空间变换

YUV 颜色空间用亮度(Y)和色度(U,V)来表示

颜色信息,在YUV空间中颜色的传递可以避免迭代的颜色空间变换。

HSI模型是美国色彩学家孟塞尔Munsell于1915年提出的,以色调(H)、饱和度(S)和亮度(I)3种基本特征量来感知颜色,由这3种成分组成的理想模型是一个类似纺锤的颜色体<sup>[8]</sup>。HSI颜色空间符合人类的视觉系统特征。

### 1.2 计算图像的显著度

根据图像的亮度特征,每个像素的显著性值是其和图像中其他像素的欧式距离的总和<sup>[9]</sup>,生成显著图 $S'$ ,公式如下:

$$S'(I_k) = \sum_{\forall I_i \in I} \|I_k - I_i\| \quad (1)$$

其中, $I_i$ 为像素 $i$ 的灰度值, $I_i \in [0, 255]$ 。Gray、YUV

和 HSI 颜色空间对应的显著图分别为  $S_{\text{Gray}}$ 、 $S_{\text{YUV}}$  和  $S_{\text{HSI}}$ 。

### 1.3 视觉注意力仿真处理

基于某一种颜色空间变换下的均衡图(图1),本研究提出视觉注意力仿真处理,自适应提取最显著区域。图2为视觉注意力仿真处理的步骤图。首先,计算均衡图得到图像的重心点。然后,采用显著点检测和区域扩张的方法,得到最显著区域。基于最显著区域,我们可以得到一个自适应约束加权项。最后,约束加权项对均衡图进行加权处理,得到某一种颜色变换下的显著图。

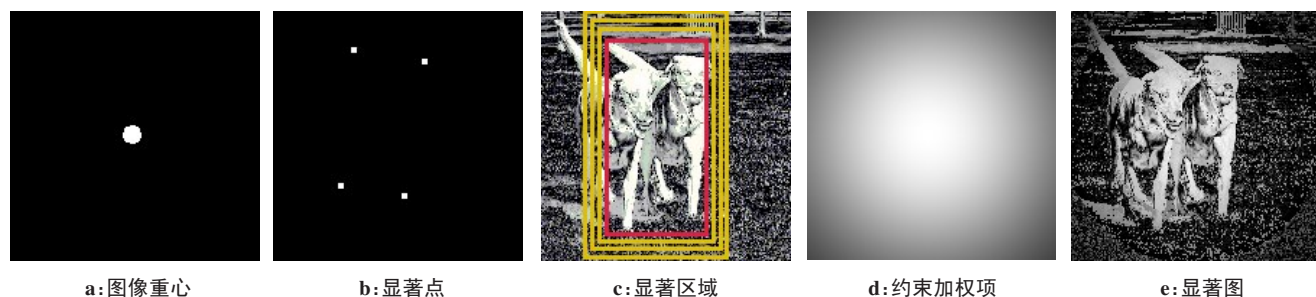


图2 视觉注意力仿真处理的步骤图(以 HSI 颜色空间为例)

Fig.2 Procedure of visual attention simulation processing (take HSI color space for example)

生物视觉感知系统中,同心圆感受野又称中心-周边感受野<sup>[10]</sup>。在二维空间中,高斯滤波器生成的曲面的等高线是从中心开始呈正态分布的同心圆。而本研究采用了同心圆,刺激源在中心 DOG 中的窄高斯来模拟人类的视觉感受野特性,其二维高斯函数可以表示为:

$$\begin{cases} G(x, y) = \sum_{y_0=0}^h \sum_{x_0=0}^w A e^{-(x^2 + y^2)/2\sigma^2} \\ x = x_0 - c_x \\ y = y_0 - c_y \end{cases} \quad (2)$$

其中, $A$ 为常数,值为255; $(c_x, c_y)$ 为高斯核的中心; $w \times h$ 分别对应图像分辨率的长和宽; $\sigma$ 是正态分布的标准偏差,决定高斯函数的高度;公式中的 $\sigma, c_x, c_y$ 是决定高斯函数的重要参数。

采用图像的重心作为高斯核的中心,而 $\sigma$ 值使用自适应算法。首先通过滑动窗口获得图像中的局部最大值点,即显著点;随后,以局部最大值点为区域扩张的种子点,区域扩张<sup>[8]</sup>的算法如下:

步骤一:边 $i$ 的外扩可能值为:

$$EP(i) = \frac{\sqrt{E(s_i)E(s'_i)}}{l_i} \quad (3)$$

其中, $i$ 为区域的边, $i \in \{1, 2, 3, 4\}$ ;  $l_i$ 为边 $i$ 的长度;

$E(s_i), E(s'_i)$ 为边内 $s_i$ 与边外 $s'_i$ 的能量值。

步骤二:如果 $EP(i) > Th$ ,边 $i$ 往外扩张一个像素, $Th$ 为扩张的停止条件。

步骤三:重复步骤一和步骤二,直到不满足扩张条件为止。

显著点扩张得到的区域,重叠率高的则为显著区域。最后, $\sigma$ 取值如下:

$$\sigma = \text{mean}(d_1 + d_2 + d_3 + d_4) \quad (4)$$

其中, $\text{mean}$ 函数求取数组 $\{d_1, d_2, d_3, d_4\}$ 的均值; $d_i$ 为显著区域边 $i$ 与图像边 $i$ 的距离。

### 1.4 显著图的生成

显著图的公式如下:

$$S = S_{\text{Gray}} + W \cdot S_{\text{YUV}} \cdot G_{\text{YUV}} + W \cdot S_{\text{HSI}} \cdot G_{\text{HSI}} \quad (5)$$

其中, $W$ 为权值0.5; $S_{\text{Gray}}$ 为灰度变换后生成的显著图; $S_{\text{YUV}}$ 为YUV颜色空间变换后生成的显著图; $S_{\text{HSI}}$ 为HSI颜色空间变换后生成的显著图; $G_{\text{YUV}}$ 和 $G_{\text{HSI}}$ 分别为 $S_{\text{YUV}}$ 和 $S_{\text{HSI}}$ 对应的高斯图。

### 1.5 人工视觉仿真

由于微电极阵列的数目有限,视网膜假体所提供的视觉感受是一种由离散光幻视点构成的低分辨率的视觉感受。在本研究中,采用局部平均法,首先将显著图分为 $24 \times 24$ 块,随后计算每块中所有像素点



的平均灰度值,最后将该灰度值作为光幻视点固定在光幻视阵列中的相应位置。一个光幻视点的公式<sup>[11]</sup>如式(6)所示:

$$p(\Delta x, \Delta y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\Delta x^2 + \Delta y^2}{2\sigma^2}\right) \quad (6)$$

其中,在一个 $n \times n$ 块中, $x$ 和 $y$ 为偏离光幻视中心坐标的大小; $\sigma$ 为块的标准差,两个光幻视点之间的间距为了满足人类的视觉光敏度, $\sigma$ 需要满足公式 $3\sigma=n/2$ 。在人工视觉仿真方面,将我们提出的策略与原始LC和DP处理后的图像进行对比。

## 2 结果

本研究提出的算法基于C++语言,在 Visual

Studio 2015 环境下实现。测试算法使用配置了 Window 7 OS 个人计算机 (Inter(R) Core(TM) i5-6500 CPU @ 3.2 GHz, 16.0 GB RAM, Lenovo, China)。算法效度评测使用MATLAB语言编程。选取 MSRA-10K<sup>[12]</sup>和 ECSSD-1K<sup>[13]</sup>两个公开基准库来验证我们提出的显著性检测策略。这两个数据库被广泛用于评测基于显著性检测的图像处理策略,其中显著区域在二值图像中被标记的图称为标注图 (Ground Truth, GT)。

选择原始的LC算法与我们的算法相比,对两个公开基准库中的图像进行测试。从图3可以看出,本文策略得到的显著图更接近GT,并且显著区域也更加精确和显著。

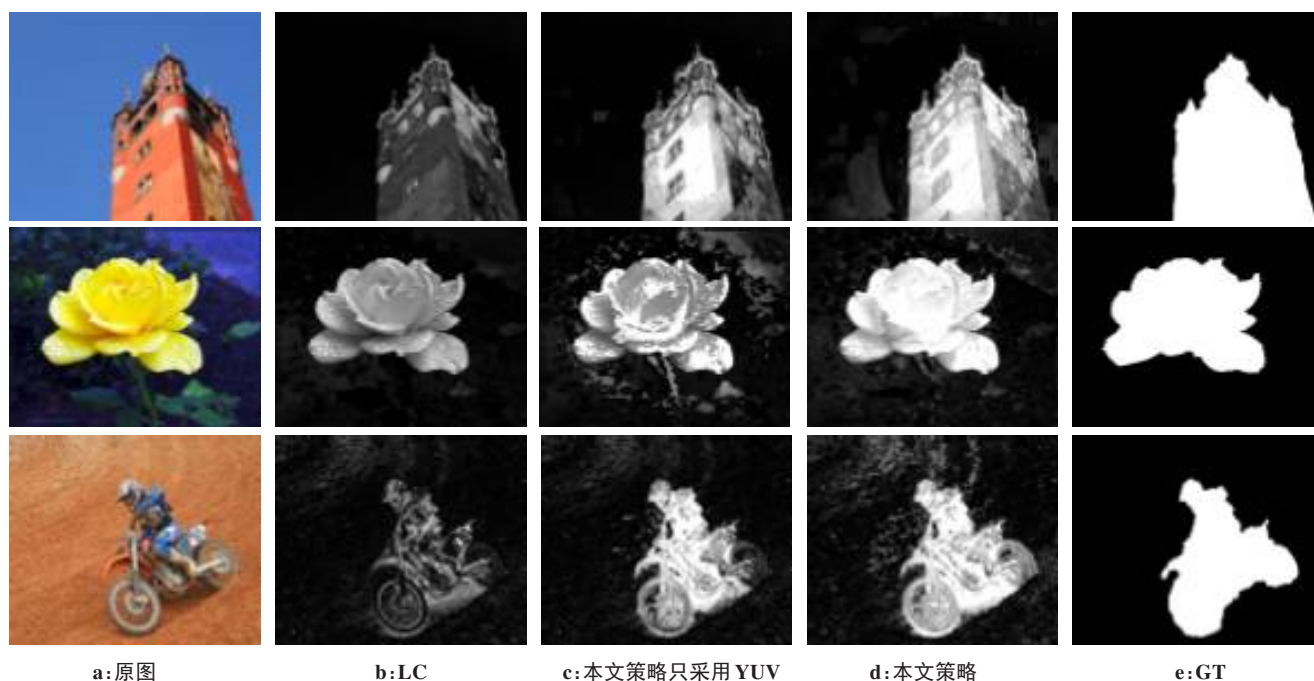


图3 显著图对比

Fig.3 Saliency map comparison

为了定量验证本研究提出的策略,选择PR<sup>[14]</sup>、F-measure<sup>[15]</sup>和ROC<sup>[16]</sup>作为评测指标。PR曲线中准确率是检测得到的区域占整张显著图的比例,召回率是检测得到的区域占标注图的比例。F-measure采用非负权重 $\beta$ ,对准确率和召回率进行加权,调和平均值,从而完善对精确度的衡量。除了PR、F-measure,我们还使用ROC评测指标进一步评测我们的改进算法。

MSRA-10K 和 ECSSD-1K 数据库集的测试结果如图4所示。图左列为各算法的PR曲线,从图中可以看出我们提出的策略优于原始的LC算法。图中间列 MSRA-10K 得到的 F-measure 结果显示我们的策略在阈值为64以上时比LC算法更高更平。从图

ECSSD-1K 可以看出,我们的结果整体比LC算法高。图右侧为ROC曲线,我们的策略远远好于LC算法。评价结果表明,我们的策略在两个基准数据库上有较好的性能,表明我们的策略对自然场景有较好的显著性检测。

为了进一步对比我们提出的策略与原始LC以及不需要任何图像处理的情况,本研究通过低像素化处理仿真视网膜假体植入后的视觉感受。根据流行病学数据,本研究选用盲人日常生活中熟悉的70个物体的数据集来测试算法。结果如图5所示。DP不能从图像中提取显著物体,也不能对图像进行增强。基于灰度变换的原始LC能够提取出显著物体,

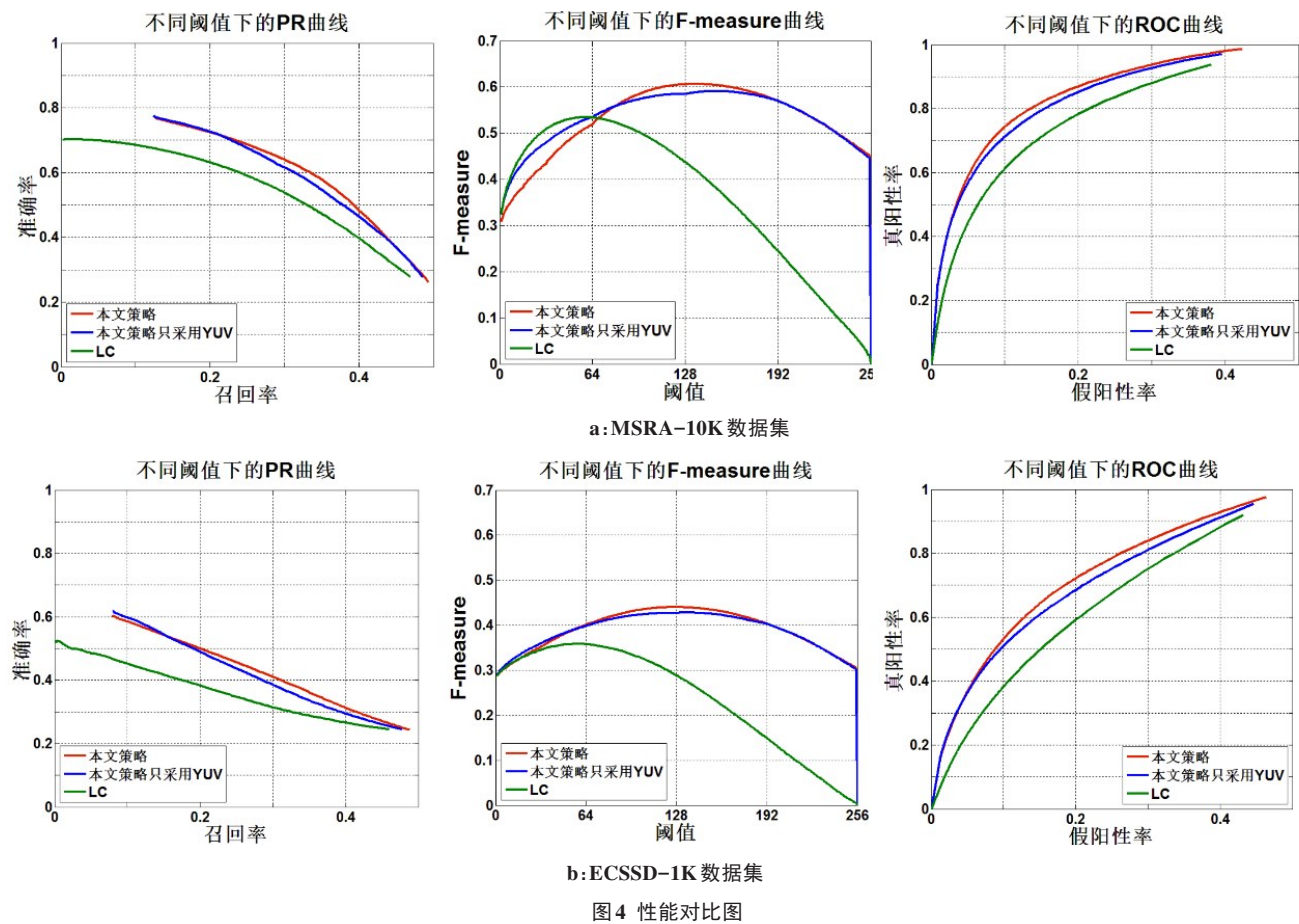


图4 性能对比图

但由于显著性的动态增强范围较低,提高效果不明显。然而,我们的策略可以更精确地获得突出的对象,使前景更加明显。

### 3 结果分析

基于全局亮度对比计算,选取LC算法作为基准,并对其进行优化。原始的LC算法可以实现更快的处理,并且可以获得比其他最新的显著检测算法更好的显著性检测结果。基于这些原因,我们提出一种基于LC的显著性检测优化策略。

本研究选择两个公共基准数据库(MSRA10K和ECSSD1K)来比较所提出的策略与原始LC的效果。MSRA10K和ECSSD1K图像数据集均具有较大的显著物体,并且靠近图像中心,由于ECSSD1K包含较多高复杂度的图像<sup>[17]</sup>。因此,ECSSD1K数据集的整体处理效果略低于MSRA10K,如图4所示。

在原始LC的基础上,除了RGB颜色空间下的强度对比外,本研究还采用YUV和HSI颜色空间变换下的亮度对比。在不同的颜色空间表示的彩色图像也具有不同的特点,颜色空间的选择是否合适对彩色图像的处理效果具有重要的影响。YUV可以避免颜色空间的迭代变换,可以实现数和指数计算的实

时处理,并且YUV颜色空间视觉上是连续的。HSI在处理彩色图像时,可仅对I分量进行处理,结果不改变原图像中的彩色种类;同时,HSI模型可以反映人感知颜色的基本属性,但H分量在不同区域的计算也会导致人类对颜色信息视觉上的不连续性。因而通过YUV和HSI两种颜色空间转换的结合,在符合人眼对颜色的感知和实时处理的同时,弥补HSI颜色空间的不连续性。

策略中提出视觉注意力仿真处理,可以自适应增强前景和过滤其他信息,主要包括区域扩张和DOG处理。原始的LC基于矩形区域进行扩张,得到显著区域。然而,当矩形选择显著区域时,往往会丢失一些前景信息。根据生物视觉感知特性,本研究通过区域扩张得到的矩形区域得到高斯差分的内径对图像进行自适应滤波,进一步增强前景和滤波背景。DOG数学模型来源于视网膜的A神经节细胞和大细胞层LGN神经元的感受野研究<sup>[18-19]</sup>,在视觉注意力的领域广泛应用于图像识别<sup>[20-21]</sup>。实验结果表明,这些处理不仅滤除了前景增强引起的背景噪声,而且与原始LC相比弥补了前景信息的缺失,实现背景抑制的同时更好地提取前景。因此,作为一种仿生注意力模型,DOG被作为获得显著区域的中间步骤。

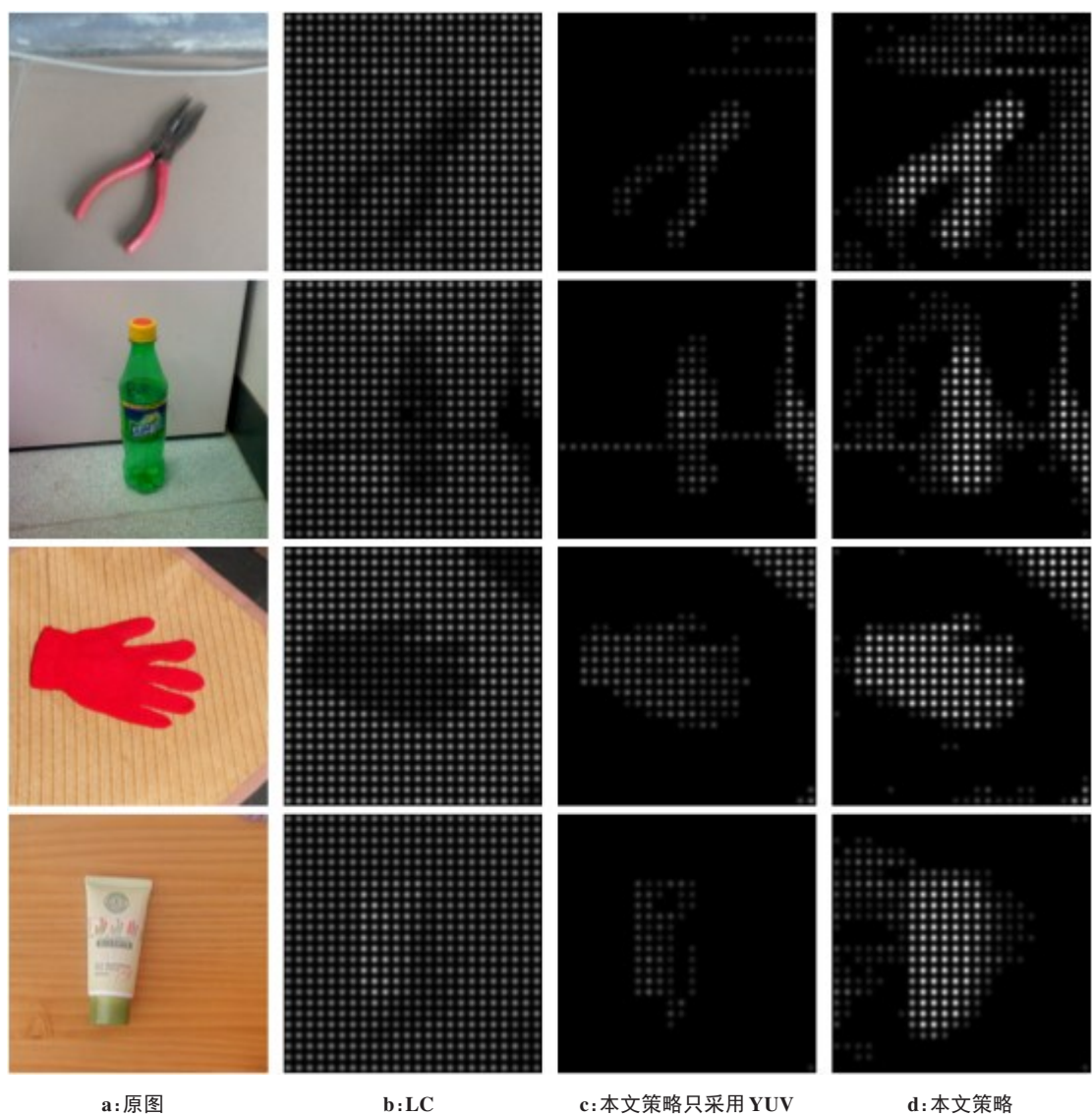


图5 仿真结果图  
Fig.5 Simulation results

#### 4 结 论

由于电极数量有限,视网膜假体只能提供低分辨率的视觉感受,这就需要图像处理策略来优化视觉信息,帮助植入者在电极数量有限的情况下完成日常生活中的目标识别和导航等任务。为了优化视觉信息,提出一种基于LC的图像显著性计算策略,将图像颜色空间转换与仿生注意机制DOG相结合。通过一些评测指标,结果表明我们的策略优于原始的LC。同时,基于盲人经常使用的物体的流行病学数据,研究表明该策略可以优化人工假体中显著物体信息的视觉表达。

#### 【参考文献】

[1] 李孟辉,任秋实.视觉假体—盲人复明的新希望[J].世界复合医学,2015,1(3): 231-237.  
LI M H, REN Q S. Visual prosthesis-a new hope for vision restoration

[J]. World Journal of Complex Medicine, 2015, 1(3): 231-237.  
[2] 徐至研. Argus II 人工视网膜系统研究进展[J]. 中华实验眼科杂志, 2019, 37(2): 157-160.  
XU Z Y. Argus II retinal prosthesis system: an update[J]. Chinese Journal of Experimental Ophthalmology, 2019, 37(2): 157-160.  
[3] 陆燕玉. 视觉假体中人工视觉最小信息需求及其优化研究[D]. 上海: 上海交通大学, 2012.  
LU Y Y. Minimal information requirement and optimization of artificial vision for visual prosthesis[D]. Shanghai: Shanghai Jiao Tong University, 2012.  
[4] WANG J, LI H, FU W, et al. Image processing strategies based on a visual saliency model for object recognition under simulated prosthetic vision[J]. Int J Artif Organs, 2016, 40(1): 94-100.  
[5] LI H, HAN T, WANG J, et al. A real-time image optimization strategy based on global saliency detection for artificial retinal prostheses[J]. Inform Sciences, 2017, 415-416: 1-18.  
[6] GUO F, YANG Y, GAO Y. Optimization of visual information presentation for visual prosthesis[J]. Int J Biomed Imaging, 2018 (3): 1-12.  
[7] 朱膺善. 关于人体解剖生理学上的一些数据[J]. 生物学教学, 1986 (2): 44-45.



- ZHU Y S. Data of human anatomy and physiology [J]. Biology Teaching, 1986(2): 44-45.
- [8] 刘伟奇, 冯睿, 周丰昆. 符合人眼视觉特性的颜色亮度模型[J]. 光学学报, 1999, 19(10): 131-134.
- LIU W Q, FENG R, ZHOU F K. A color brightness model fitting characteristics of the human vision[J]. Acta Optica Sinica, 1999, 19(10): 131-134.
- [9] ZHAI Y, SHAH M. Visual attention detection in video sequences using spatiotemporal cues[C]. ACM, 2006.
- [10] KUFFLER S W. Discharge patterns and functional organization of the mammalian retina[J]. J Neurophysiol, 1953, 16(1): 37-68.
- [11] LU Y, KAN H, LIU J, et al. Optimizing Chinese character displays improves recognition and reading performance of simulated irregular phosphene maps[J]. Invest Ophthalmol Vis Sci, 2013, 54(4): 2918.
- [12] HUANG H, FLYNN N M, KING J H, et al. Comparisons of community-associated methicillin-resistant staphylococcus aureus (MRSA) and hospital-associated MRSA infections in Sacramento, California[J]. J Clin Microbiol, 2006, 44(7): 2423-2427.
- [13] SHI J, YAN Q, XU L, et al. Hierarchical saliency detection on extended CSSD[J]. IEEE Trans Pattern Anal Mach Intell, 2014, 38(4): 717.
- [14] BUCKLAND M K, GEY F C. The relationship between recall and precision[J]. J Am Soc Inf Sci, 1994, 45(1): 12-19.
- [15] ZHANG E, ZHANG Y. F-Measure[M]. Springer, 2009.
- [16] FAWCETT T. An introduction to ROC analysis[J]. Pattern Recognit Lett, 2005, 27(8): 861-874.
- [17] CHENG M M, MITRA N J, HUANG X, et al. Global contrast based salient region detection[C]//CVPR. IEEE, 2011.
- [18] HUBEL D H, WIESEL T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. J Physiol, 1962, 160(1): 106-154.
- [19] RODIECK R W, STONE J. Analysis of receptive fields of cat retinal ganglion cells[J]. J Neurophysiol, 1965, 28(5): 832-849.
- [20] ULLMAN S. Analysis of visual motion by biological and computer systems [M]. Reading in Computer vision: issues, problems, principles, and paradigms, San Francisco: Morgan Kaufmann Publishers Inc., 1987: 132-144.
- [21] SATOH S. A visual model for object detection based on active contours and level-set method[J]. Biol Cybern, 2006, 95(3): 259-270.
- (编辑:陈丽霞)