

基于对比敏感度和马尔可夫链的注意信息提取算法

张孝临^{1,2},赵宏伟^{1,2},王 慧^{1,2},戴金波^{1,2,3}

(1. 吉林大学计算机科学与技术学院,吉林长春 130012;

2. 吉林大学符号计算与知识工程教育部重点实验室,吉林长春 130012;

3. 长春师范学院计算机科学与技术学院,吉林长春 130032)

摘 要: 借鉴生理学的研究成果,提出了一种新的基于对比敏感度和马尔可夫链的视觉注意信息提取算法.在注意特征向量提取之前,先用与离心率有关的对比敏感度函数对输入的图像进行加权,用以模拟视网膜神经节的反应机制;在特征向量上定义马尔可夫链,用它的平稳分布做为活动图上的显著度.算法的平均计算时间和以神经生物学家的研究成果为标准计算的接受者操作特性曲线下面积证明了算法的有效性.

关键词: 视觉注意;视觉显著性;对比敏感度;马尔可夫链

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2010) 2A-213-05

Extracting Attention Information Algorithm Based on Contrast Sensitivity and Markov Chain

ZHANG Xiao-lin^{1,2}, ZHAO Hong-wei^{1,2}, WANG Hui^{1,2}, DAI Jin-bo^{1,2,3}

(1. College of Computer Science and Technology, Jilin University, Changchun 130012, China;

2. Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China;

3. College of Computer Science and Technology, Changchun Normal University, Changchun 130032, China)

Abstract: Inspired by the research in physiology, a novel algorithm for extracting bottom-up attention information (integration of contrast sensitivity and Markov chain, ACSMC) is proposed in this paper. In our algorithm, the original image is weighted with a contrast sensitivity formula which is a function retinal eccentricity to simulate the mechanism of retinal ganglion. A Markov chain is defined on feature maps. The equilibrium distribution of this chain is taken as saliency values. The average of algorithm cost time and area under receiver operating characteristic curve (AUROC) based on the research of neurobiologist demonstrate its effectiveness.

Key words: visual attention; visual saliency; contrast sensitivity; Markov chain

1 引言

视觉注意可以把系统中有限的处理资源优先分配给少数几个显著的视觉区域上.视觉注意的分配算法对智能模拟、图像压缩和目标识别等多个领域都具有重要意义,近几年来,已成为一个热点课题.

目前视觉注意模型在提取自底向上的视觉信息时大体上可以分为特征向量提取(extraction)、活动图形成(activation)和整合(combination)三个步骤.目前大多数视觉注意计算模型在提取特征向量时均采用基于生物学的研究成果的多通道多尺度滤波器组对图像进行滤波

的方法^[1-6].在特征提取之后,一些学者用中央-外围算子^[1-5]来模拟细胞的感受野特性从而获得局部显著性,也有一部分学者在特征向量上计算附加函数来形成活动图^[6,7],还有学者将协同识别理论引入到这一过程^[8].最后,根据计算得到的活动图用局部最大值^[1-3]、高斯差^[4]和多尺度熵^[9]等算法来确定注意固定点等信息.以上模型无论具体采用哪种算法,其中心思想均是在视觉空间中具有较高的显著性的区域在获得有限的注意资源时具有优先性.因此,图像显著性的计算在注意焦点转移的确定中起决定作用.

本文提出了一种新的视觉注意信息提取算法

ACSMC(Algorithm Based on Contrast Sensitivity and Markov Chain),用对比敏感度和马尔可夫链来提高注意固定点确定的可靠性和计算效率.在构建金字塔之前用与离心率有关的对比敏感度函数对输入的图像进行加权,可以更好的模拟视网膜神经节细胞在视网膜不同位置上对相同的明暗刺激产生不同响应的情况.令质点在得到的特征向量上做具有反射边界的二维随机游动,以相邻两点间的差异和距离做为相互游动的可能性度量,以马尔可夫链的平稳分布做为该点的活动显著值.具体算法和实验数据如下.

2 基于对比敏感度的特征向量提取

经典的注意计算模型在特征向量提取时通常将原始视觉图像分为颜色通道、亮度通道、方向通道和纹理通道等通道分别传递,利用图像金字塔降采样得到不同等级的特征向量^[1-6].图像金字塔降采样可以有效的模拟感受野的等级特性,然而大多数模型在金字塔的底层上采用固定尺寸的滤波器进行滤波,它们没有考虑到在金字塔底层上因视网膜上光感受器、神经节细胞的分布不均而造成的视网膜神经节细胞对相同的明暗刺激随着离心率的增加而产生急剧减小的刺激的情况.事实上,大多数视觉注意的模型在提取视觉显著度的时候是依靠亮-暗、红-绿和黄-蓝的视觉对比得到的^[1-5],即使有的模型在程序上没有直接做这种视觉对比,它们在最终获取注意固定点时也是根据全局的显著度情况来确定的^[6,7],全局显著度计算落到神经节的层次上也应随着离心率的不同而不同.这里,我们用对比敏感度函数做为视网膜变换的权系数.

对比敏感度为视觉系统能察觉的对比度阈值的倒数,它描述了人眼区分细微明暗和它们边界的能力.对比度阈值越低,则对比敏感度越高,视觉功能越好.视网膜的生理结构决定了,在中央凹区域的对比敏感度是最高的,当从中央凹往外围测试时,对比敏感度急剧下降.

文献[10]给出了一个符合心理学关于对比敏感度的实验数据的公式,如下所示

$$CT(f, e) = CT_0 \exp(\alpha f \frac{e + e_2}{e_2}) \quad (1)$$

其中, CT 是可视对比阈值, f 是空间频率, e 是视网膜离心率, CT_0 为最大对比阈值, α 是空间频率衰减常数, e_2 是半分辨离心率常数.最佳的参数设置是 $\alpha = 0.106$, $e_2 = 2.3$, $CT_0 = 1/64$.

对比敏感度是对比阈值的倒数,

$$CS(f, e) = 1/CT(f, e) \quad (2)$$

f 通常被设置为图像的奈奎斯特频率,这样对比敏感度函数 $CS(f, e)$ 实际上是关于视网膜离心率的函数.

设输入图像为 $O: [n]^2 \rightarrow R$ (实际上,文中用到的图像

不是严格正方的,而是矩形的,这时我们取矩形的短边为 n),该幅图像的整个视角为 θ ,则视距 v 是使得图像边缘的离心率为 $\theta/2$ 的从眼睛到图像的观察距离.可得,

$$v(\theta) = \frac{n}{2} / \tan(\frac{\theta}{2}) \quad (3)$$

设图像中心点坐标为 (x_0, y_0) ,则图像上任一点 $(x, y) \in [n]^2$ 的离心率 e 是该点到中心点的距离与视距之比的反正切.

$$e(x, y, \theta) = \tan^{-1} \left(\frac{((x - x_0)^2 + (y - y_0)^2)^{1/2}}{v(\theta)} \right) \quad (4)$$

则用对比敏感度函数变换后的图像为

$$O'(x, y) = O(x, y) * CS(e(x, y, \theta)) \quad (5)$$

当 $\theta = 55^\circ$, $f = 5.4 \text{ cpd}$ 时,视网膜变换权系数如图1所示.在图像正中央视觉敏感度最大,归一化权系数为1.可以看出,权系数随着该点与中心点距离的增加(离心率的增加)而减小.

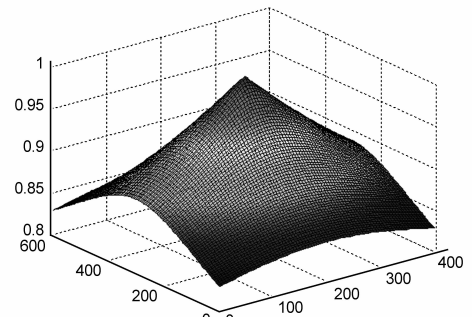


图1 视网膜变换权系数

用敏感度对图像进行加权变换可以有效的抑制图像外围信息在注意显著图形成过程中的活跃性.

本文选择视觉上较为敏感的亮度、颜色和方向这三类特征参与检测.用文献[1]中特征通道提取方法对变换后的图像 $O'(x, y)$ 分别计算,得到亮度图 I ,颜色通道中红绿特征图 RG 、蓝黄特征图 BY ,方向通道特征图 $O(\theta)$,其中 $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$.分别对三个通道的7幅特征图用9层的高斯金字塔滤波,得到 $I(\sigma)$, $RG(\sigma)$, $BY(\sigma)$ 和 $O(\theta, \sigma)$,其中 $\sigma = 0, \dots, 8$.选取金字塔的第二层和第三层并把它们归一化 N (局部平均^[11])到统一尺寸做为特征向量.

3 基于马尔可夫链的活动图形成

在特征向量上定义马尔可夫链,用它的平稳分布做为活动图上的显著度.在我们之前,文献[6]把特征图上的所有点做为节点,把两点之间的差异与它们距离的乘积做为边的连接权值,以此构成完全有向图,并在完全图上定义马尔可夫链,用马尔可夫链的平稳分布来估计活动图上的显著度,取得了较好的效果.但文献[6]中计算在完全有向图上定义的马尔可夫链的平

稳分布具有较高的时间复杂度,约为 $O(n^4)$. 我们对其定义马尔可夫链的方法进行了改进,在得到注意固定点效果损失不那么明显的情况下,大幅降低计算所耗费的时间,改进后的算法时间复杂度约为 $O(n^2)$,具体算法如下:

设特征向量 F 的维数为 $M \times N$, 则在该特征向量上做具有反射边界的二维随机游动,如果游动点 Q 现在位于任一非边界上的点 (x, y) 上,则下一时刻 Q 只能位于与当前点 Q 相邻的周边八个点上(如果当前位于边界点,则下一时刻 Q 只能位于周边五个点上;若位于角点,下一时刻只能位于周边三个点上), Q 向周边点 (x', y') 移动的概率与两点的向量值差异成正比,与两点的距离指数成反比. 令 $p'((x, y), (x', y'))$ 为未归一化之前 Q 从 (x, y) 点向 (x', y') 点游动的可能性,则

$$p'((x, y), (x', y')) = \frac{|F(x, y) - F(x', y')|}{\exp\left(\frac{(x - x')^2 + (y - y')^2}{2\sigma^2}\right)}$$

其中, $|x - x'| \leq 1, |y - y'| \leq 1, \sigma$ 为常数. (6)

由于 Q 点向周围八个方向的游动概率之和应为 1, 我们归一化 p' , 得到

$$P\{X_{n+1} = (x', y') | X_n = (x, y)\} = \frac{p'((x, y), (x', y'))}{\sum_{i=-1}^1 \sum_{j=-1}^1 p'((x, y), (x + i, y + j))}$$

其中, $|x - x'| \leq 1, |y - y'| \leq 1$ (7)

Q 当前位于边界点或角点的情况以此类推.

设置一步转移概率矩阵 P 的维数为 $(MN) \times (MN)$ 维, 其矩阵元素

$$P_{ij} = P\{X_{n+1} = (\lceil j/N \rceil, (j \bmod N) + N(\lfloor j/N \rfloor - \lceil j/N \rceil + 1)) | X_n = (\lceil i/N \rceil, (i \bmod N) + N(\lfloor i/N \rfloor - \lceil i/N \rceil + 1))\} \quad (8)$$

其中, $i, j = 1, 2, \dots, MN$

此时, 其平稳分布 $\pi = (\pi_1, \pi_2, \dots, \pi_{MN})$ 为满足方程组(9)的解.

$$\begin{cases} \pi = \pi P \\ \sum_{i=1}^{MN} \pi_i = 1 \end{cases} \quad (9)$$

这个分布给出了经过长时间游动以后 Q 点位于某点的可能概率. 用本文方法定义的马氏链, 其平稳分布反映了特征图上该点在竞争过程中获胜的可能性, 用它做为活动图的显著度具有良好的物理意义.

本文整合算法即显著图的生成方法, 与活动图生成方法类似, 游动点 Q 在活动图上做具有反射边界的二维随机游动. 不同之处是, 未归一化之前 Q 从 (x, y) 点向 (x', y') 点游动的可能性

$$p'((x, y), (x', y')) = \frac{|F(x', y')|}{\exp\left(\frac{(x - x')^2 + (y - y')^2}{2\sigma^2}\right)}$$

其中, $|x - x'| \leq 1, |y - y'| \leq 1, \sigma$ 为常数. (10)

显著图的生成方法其它部分与活动图生成方法相同, 由于篇幅限制, 不再累述.

4 实验结果及分析

实验原始数据采用 Einhäuser 等人^[11]的数据集. Einhäuser 等人测试并记录了 7 个人在自由注视的情况下对 108 幅自然场景图像和它们的修改版本图像(共 972 幅图片)进行观察所获得的注意固定点位置. 他们将 972 幅图片分为 18 组(每组 54 幅), 7 个人分为四组(前三组二个人, 第四组一个人), 人物前三组分别对 18 组图片中的 6 组图片进行测试, 人物第四组对所有 18 组图片进行测试. 这样, 每幅图片均有三个人对其进行测试. 我们在测试本文算法性能时只选取 Einhäuser 等人的数据集中未被修改的 108 幅自然场景图片和它们的三个人的注意固定点位置坐标做为标准数据集, 其中 108 幅图片只用到其中心位置 600×400 像素部分, 对应的视角范围是 $76^\circ \times 55^\circ$. 由于标准数据集中的测试图片是灰度图, 因此实际在计算算法特征图时只选择亮度通道和方向通道参与检测. 把本文算法 ACSMC 与目前被普遍认可的 Gbvs^[6]和 Saliency^[2]算法的部分与整体分别做比较, 比较结果如下.

4.1 活动图形成和整合算法性能比较

为实现比较的公平性, 进行这一部分的比较时特征提取部分采用的方法相同. 对亮度通道和方向通道 5 幅特征图进行金字塔滤波, 选用金字塔的第二层和第三层共 10 幅特征图, 并把它们分别归一化(局部平均^[1])到 25×37 尺寸做为特征向量. 其中 Center-surround(c-s)^[2]活动图生成算法需要做多尺度特征图的相减操作, 这里我们选取的 c 为金字塔的 2 层和 3 层, δ 为 4 层. 我们把 108 幅自然图片分为 11 组进行测试(前 10 组分别含有 10 幅图片, 第 11 组含有 8 幅图片), 分别用文中基于马尔可夫

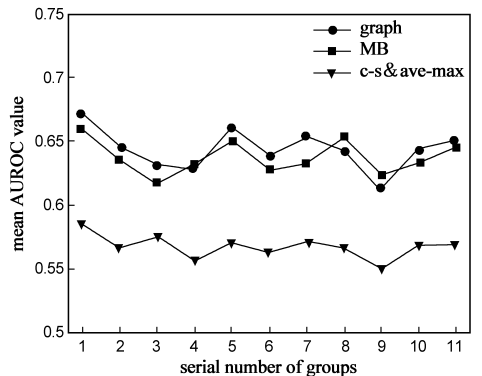


图2 活动图形成和整合算法的比较图

链的活动图形成和整合算法 MB (Markovian-based), graph^[6]和 c-s&ave-max^[2]对它们进行计算得到显著图, 根据 Einhäuser 数据集中的注意固定点位置计算 ROC 曲线下面积(area under receiver operating characteristic curve, AUROC). 以图片分组的序号为横坐标, 以每组图片的 AUROC 平均值为纵坐标的实验结果图如图 2 所示.

表 1 给出了这三种算法对 108 幅图片进行计算得到的平均 AUROC 值和在主频为 1.5GHz, 512MB 内存的机器上算法的平均计算时间.

虽然 MB 算法的平均 AUROC 值略低于 graph 算法的平均 AUROC 值, 但 MB 算法的平均计算时间大幅小于 graph 算法的平均计算时间. 以上三种算法的平均计算时间以 c-s&ave-max 算法为最小, 仅为 0.097s, 但同时

它的平均 AUROC 值也是最小的.

表 1 算法性能测试结果

算法性能	算法	graph	c-s&ave-max	MB
平均 AUROC		0.6431	0.5670	0.6364
平均计算时间(s)		11.278	0.097	6.271

4.2 注意信息提取算法比较

为符合标准数据集中图片对应的视角范围, 文中 ACSMC 算法视网膜变换部分设置 $\theta = 55^\circ$. 用 ACSMC, Gbvs 和 Saliency 算法分别计算测试集中图片的显著图, 并计算它们的 ROC 曲线下面积. 部分测试结果如图 3 所示. 用 Einhäuser 数据集给出的注意固定位置做为评判性能的标准. 为观察方便, 在图 3 中的测试原图上用黄色“■”标记了相关注意固定点位置.

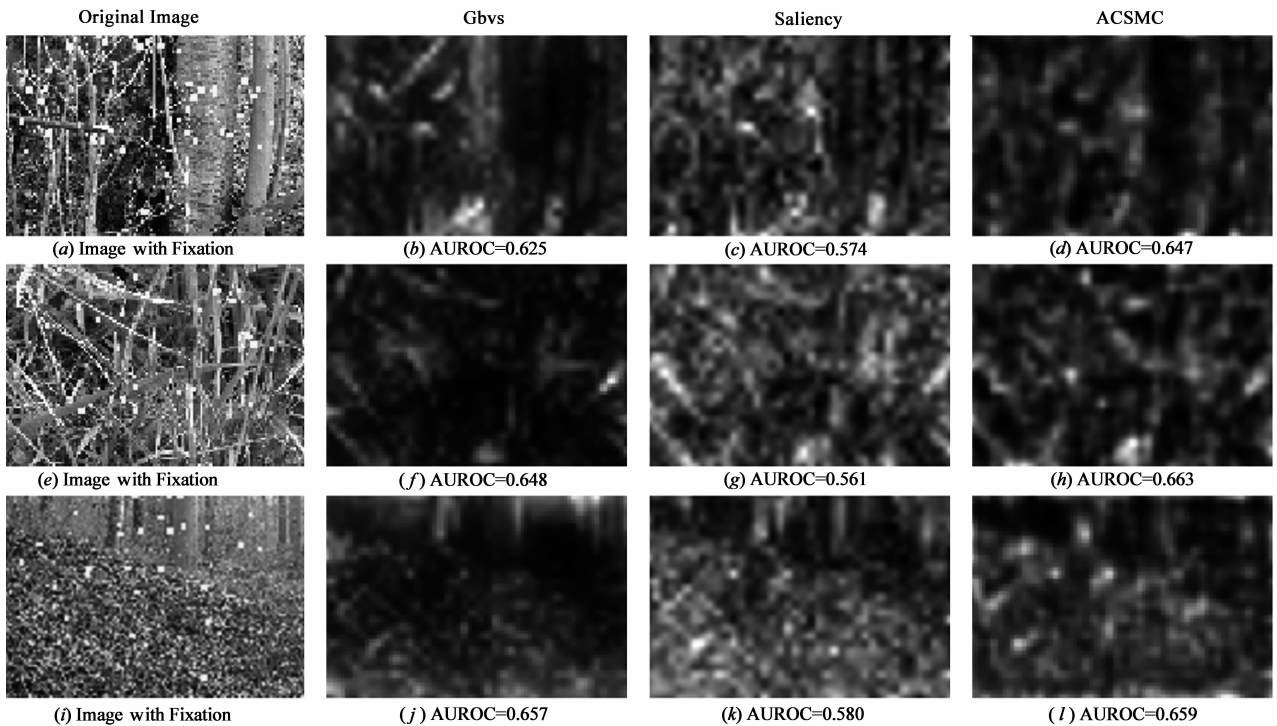


图3 部分测试结果. (a),(e),(i)为三幅用黄色“■”标记了注意固定位置的数据集中的图片;(b),(c),(d)分别为对图3(a)用Gbvs, Saliency, ACSMC算法计算得到的显著图;(f),(g),(h)分别为对图3(e)用Gbvs, Saliency, ACSMC算法计算得到的显著图;(j),(k),(l)分别为对图3(i)用Gbvs, Saliency, ACSMC算法计算得到的显著图;在标注中分别给出了它们的ROC曲线下面积(AUROC).

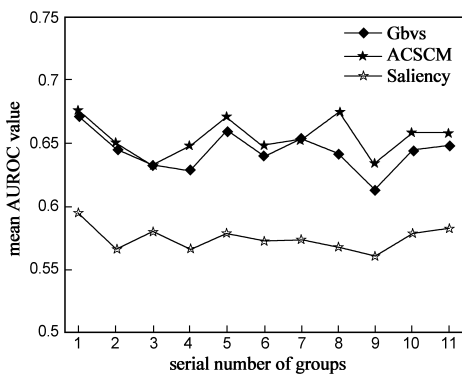


图4 注意信息提取算法的比较图

仍旧把 108 幅自然图片分为 11 组进行测试, 计算平均 AUROC 值, 注意信息提取算法 ACSMC, Gbvs 和 Saliency 算法的分组测试结果如图 4 所示.

由图 4 可知, ACSMC 结果曲线略高于 Gbvs 曲线, 且两者的曲线值均大于 Saliency 曲线值. 图 4 中 Saliency 的检测结果要好于图 2 中 c-s&ave-max 的检测结果, 这可能与 Saliency 根据原文^[2]选择的相减尺度比图 2 中 c-s&ave-max 选择的尺度多有关系. 这三种算法对数据集所有图片进行注意信息提取的平均 AUROC 值和平均计算时间如表 2 所示.

由表 2 可知, ACSMC 的平均 AUROC 值略大于 Gbvs

且平均计算时间小于 Gbvs,虽然 ACSMC 的平均计算时间大于 Saliency,但 ACSMC 的计算准确度相对于 Saliency 要高很多.综合考虑,ACSMC 算法的整体性能要好于 Gbvs 算法和 Saliency 算法.

表 2 算法性能测试结果

算法性能	Grbvs	Saliency	ACSMC
平均 AUROC	0.6431	0.5737	0.6541
平均计算时间(s)	15.374	4.728	11.107

5 结论

本文提出了一种新的基于对比敏感度和马尔可夫链的视觉注意信息提取算法.通过计算以神经生物学家的研究成果为标准的接受者操作特性曲线下面积和算法平均计算时间证实,该算法是一种行之有效的自底向上的注意信息提取方法.

同时,本文算法还存在有待改进之处.算法只估计了注意固定点的位置,没有对显著区域的尺寸进行估计.通常上注意模型在得到注意固定点以后以注意固定点为圆心以一定尺寸范围内的区域做为显著区域,然而根据心理学的研究发现显著区域的尺寸在注意过程中会不断发生变化.因此设计有效的算法来估计显著区域尺寸将是下一步工作的方向.

参考文献:

- [1] L Itti, C Koch E Niebur. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254 – 1259.
- [2] L Itti. Models of bottom-up attention and saliency[A]. IL. Itti ed. Neurobiology of attention[C]. San Diego, CA: Elsevier, 2005. 576 – 582.
- [3] C Siagian, L Itti. Rapid biologically-inspired scene classification using features shared with visual attention[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(2): 300 – 312.
- [4] L Itti, C Koch. A saliency-based search mechanism for overt and covert shifts of visual attention[J]. Vision Research, 2000, 40(10 – 12): 1489 – 1506.
- [5] D Parkhurst, K Law, E Niebur. Modeling the role of Saliency in the allocation of overt visual attention[J]. Vision Research, 2002, 42(1): 107 – 123.
- [6] J Harel, C Koch, P Perona. Graph-based visual saliency[J]. Advances in Neural Information Processing Systems, 2007, 19: 545 – 552.
- [7] N Bruce, J Tsotsos. Saliency based on information maximization [J]. Advances in Neural Information Processing Systems,

2006, 18: 155 – 162.

- [8] 邵静,高隽.基于协同感知的视觉选择注意计算模型[J].中国图象图形学报,2008,13(1):129 – 136.
Shao Jing, Gao Jun. Visual selective attention computational model based on synergetic perception[J]. Journal of Image and Graphics, 2008, 13(1): 129 – 136.
- [9] 田媚,罗四维,黄雅平等.基于局部复杂度和初级视觉特征的自底向上注意信息提取算法[J].计算机研究与发展,2008,45(10):1739 – 1746.
Tian Mei, Luo Si-wei, Huang Ya-ping et al. Extracting bottom-up attention information based on local complexity and early visual features[J]. Journal of computer research and development, 2008, 45(10): 1739 – 1746.
- [10] W S Geisler, J S Perry. A real-time foveated multiresolution system for low-bandwidth video communication[J]. Human Vision and Electronic Imaging, Proc. SPIE, 1998, 3299: 294 – 305.
- [11] Einhäuser W, Kruse W, Hoffmann KP, König P. Differences of monkey and human overt attention under natural conditions [J]. Vision Research, 2006, 46: 1194 – 1209.

作者简介:



张孝临 男,1979年12月出生于辽宁鞍山.吉林大学计算机科学与技术学院博士研究生.主要研究方向:智能信息系统与嵌入式技术. E-mail: zxl2257cn@yahoo.com.cn



赵宏伟 男,1962年11月出生于辽宁沈阳.吉林大学计算机科学与技术学院教授,博士生导师.主要研究方向:智能信息系统与嵌入式技术、计算机图像处理与可视化. E-mail: zhaohw@jlu.edu.cn



王慧(通信作者) 女,1982年12月出生于黑龙江密山.吉林大学计算机科学与技术学院博士研究生.主要研究方向:智能信息系统与嵌入式技术. E-mail: email_wanghui@126.com

戴金波 女,1971年10月出生于吉林蛟河.吉林大学计算机科学与技术学院博士研究生,长春师范学院计算机硬件教研室副教授.主要研究方向:智能信息系统,思维计算. E-mail: email_wanghui@126.com