

# 基于协同感知的视觉选择注意计算模型

邵 静 高 隽

(合肥工业大学计算机与信息学院图像信息处理研究室, 合肥 230009)  
(中国科学院合肥智能机械研究所仿生感知与控制研究中心, 合肥 230031)

**摘 要** 鉴于在任务相关的视觉注意中, 需要建立基于任务的视觉注意显著图来引导视觉注意, 为此利用与人认知过程相接近的协同感知理论来研究基于任务的视觉注意计算模型, 即首先利用协同识别理论研究二义及多义模式的视觉感知, 得到协同视觉感知理论; 然后将协同视觉感知中的模式与从视觉注意模型中提取的底层视觉特征相对应, 利用偏置矩阵的性质计算底层视觉特征间受任务影响而产生的偏置, 再由此偏置和底层视觉特征生成基于任务的视觉注意显著图; 最后提出了基于协同感知理论的视觉选择注意计算模型。该算法用于基于任务的视觉搜索的实验结果表明, 该算法是有效的, 在认知上是合理的。

**关键词** 视觉注意 任务相关 协同感知 计算模型

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2008)01-0129-08

## Visual Selective Attention Computational Model Based on Synergetic Perception

SHAO Jing GAO Jun

(Department of Computer and Information, Laboratory of Image Information Processing Hefei University of Technology, Hefei 230009)  
(Centre for Biomimetic Sensing and Control Research, Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031)

**Abstract** In the task-relevant visual attention, the saliency map based on task is required to be built to direct the visual attention. The synergetic perception theory which is similar to the cognition of the human is utilized to research the computational model of task-relevant visual attention. Firstly, the perception of ambiguous model is researched through the synergetic recognition theory. The synergetic visual perception theory is available. The patterns in the synergetic visual perception are corresponding to the visual features in the visual attention model. Then, the bias between the visual features influenced by the task is computed by the property of the bias matrix. The task-relevant visual attention saliency map is built from the bias and the visual features. Finally, a computational model of visual selective attention based on synergetic perception is presented. The algorithm is applied to the visual search task. The validity and the rationality in cognition of the algorithm are demonstrated through the experiments.

**Keywords** visual attention, task relevant, synergetic perception, computational model

## 1 引 言

众所周知, 视觉选择注意 (visual selective attention) 机制是灵长类动物处理视觉信息的本质特征。

另外, 人类视网膜对图像是非均匀采样的, 这是视觉选择注意机制的生物基础; 同时, 由于高层视觉处理只是对初始传感器信息的一个子集进行处理, 因此需要对初始得到的视觉信息进行选择。众所周知, 人不是被动地接受来自视网膜上的视觉信息, 而是

基金项目: 国家自然科学基金项目 (60375011, 60575028); 安徽省优秀青年科技基金项目 (04042044); 教育部“新世纪优秀人才支持计划”项目 (NCET-04-0560)

收稿日期: 2006-04-28 改回日期: 2006-09-05

第一作者简介: 邵静 (1981~), 男, 2003年毕业于合肥工业大学电子信息工程专业, 现为合肥工业大学信号与信息处理专业博士研究生。主要研究方向为视觉选择注意机制、协同信息处理方法等。E-mail: shaw81@gmail.com

主动地感知它,是通过眼动、头动和移动身体等行为去选择有用的信息,而且这种选择是通过基于图像数据的底层视觉特征和任务驱动的高层知识,以“注意集中”的方式同时实现。

视觉注意计算模型中,是用显著图(saliency map)<sup>[11]</sup>来表示视觉区域的显著性。显著图就是一幅和原始图像大小相同的 2 维“图像”,其中的每个像素值表示原图像对应点的显著性大小。显著图不仅表达视觉区域每个位置的显著性,并且基于显著性的空间分布来引导注意区域的选择。

显著图的生成是视觉注意计算模型的核心,目前主要有以下 4 种研究思路:(1)基于空间和特征整合模型思路,例如 Itti 先采用中央/周边的计算策略先在每种特征图内部进行竞争,然后通过线性组合得到总的显著图<sup>[12]</sup>;(2)基于物体和整合竞争假设的思路,例如, Mahmood 先将视觉注意分为吸引注意模块和付出注意模块,再通过颜色、纹理和平行线特征来生成显著图<sup>[3]</sup>,而 Sun 则是通过扩展整合竞争假设,使用一种基于分组的竞争策略,首先假设输入的图像已经完成感知分类,然后利用灰度、颜色和方向信息来形成显著图<sup>[4]</sup>;(3)基于尺度空间表示和信息理论思路,例如 Jägersand 通过计算相邻尺度图像间的 Kullback 距离函数,该函数在空间维上集中,就可以用来找当前图像中物体的最佳尺度,而在尺度维上集中,就可以用来建立一个基于信息量度量的显著图<sup>[5]</sup>,Ferraro 和 Boccignone 等人通过建立图像的尺度空间表示,利用尺度维上的熵来建立显著图<sup>[6-7]</sup>,Kadir 在统一框架下讨论了尺度、显著性和场景描述的问题,先用香农信息熵作为特征空间的显著性度量函数,然后用相邻尺度间的差分作为尺度空间的显著性度量,选择局部熵最大时的尺度为最佳尺度,从而完成对图像的显著性度量<sup>[8]</sup>;(4)基于图像布局和 Bayes 学习与推理的思路,例如 Oliva 和 Torralba 等人将整幅图像看作一个整体,通过建立其描述,对图像所表示的场景进行分类,先得到场景的全局和先验知识,然后通过样本图像的训练和 Bayes 推理来进行图像中显著物体提取<sup>[9-12]</sup>。

现有模型的共同特点是:(1)显著图的生成都是基于各种底层的视觉特征(如,强度、颜色、方向、深度等);(2)各种底层视觉特征生成显著图时,由于没有考虑任务对底层视觉特征之间偏置的影响,因此选择出的区域缺乏与任务的相关性。

在任务相关的视觉注意中,如何将给定的任务与底层视觉特征结合来形成基于任务的视觉显著图,以便更迅速地搜索到与任务相关的区域一直是研究的难点。这一方面与视觉生理学中,缺乏自上而下的神经通道模型有关;另一方面在实际应用中,由于任务的多样性,也增加了研究的难度。本文利用与人认知过程非常相近的协同识别理论来研究二义及多义模式的视觉感知,首先得到协同视觉感知理论,并由此研究任务引导情况下,视觉注意计算模型中底层视觉特征相互间产生的偏置,然后提出了基于协同感知理论的视觉注意计算模型。

## 2 基于图像的底层视觉特征

文中采用由 Koch 和 Ullman 最先提出,由 Itti 完善的基于图像底层视觉特征的计算方法<sup>[1,2]</sup>,该方法对输入图像,首先利用 2 维高斯金字塔模型得到图像的尺度空间表示<sup>[13]</sup>;然后由图像的不同尺度空间表示,通过一个线性的“中央-周边差”算子来计算底层视觉特征。这种“中央-周边差”算子是通过高分辨尺度和低分辨尺度之间的差建立的,可用  $\ominus$  来标记两个不同尺度图像之间的差。

早期视觉特征提取方法如下:

记  $r$ 、 $g$  和  $b$  分别为原图像的红、绿和蓝色通道分量,则强度为

$$I = (r + g + b) / 3 \quad (1)$$

$I$  用来产生高斯金字塔尺度图像  $I(\sigma)$ ,其中  $\sigma$  为尺度因子。这样就可以得到以下广义上的红、绿、蓝和黄 4 个通道值:

$$\begin{aligned} R &= r - (g + b) / 2 \\ G &= g - (r + b) / 2 \\ B &= b - (r + g) / 2 \\ Y &= (r + g) - |r - g| / 2 - b \end{aligned} \quad (2)$$

其中,负值设为 0。红、绿、蓝、黄高斯金字塔尺度图像  $R(\sigma)$ 、 $G(\sigma)$ 、 $B(\sigma)$  和  $Y(\sigma)$  可以从这 4 个颜色通道得到。

(1) 强度特征图

强度特征图由强度对比度产生。而通过计算  $I(c, s)$  则可以表示强度特征图

$$I(c, s) = |I(c) \ominus I(s)| \quad (3)$$

其中,  $c$  为高分辨率尺度因子,  $s$  为低分辨率尺度因子。

(2) 颜色特征图

在视觉皮层中, 有一个叫作“颜色双对立”(color double-opponent)系统, 这种空间和色彩对抗存在于红/绿、绿/红、蓝/黄、黄/蓝对中。如式(4)所示,  $RG(c, s)$ 用来计算红/绿、绿/红双对立, 式(5)表示蓝/黄、黄/蓝双对立。

$$RG(c, s) = |(R(c) - G(c)) \ominus (G(s) - R(s))| \quad (4)$$

$$BY(c, s) = |(B(c) - Y(c)) \ominus (Y(s) - B(s))| \quad (5)$$

### (3) 方向特征图

V1区是视觉信息处理的前端, 其数学模型通常用 Gabor滤波器表示。若利用 Gabor金字塔  $O(\sigma, \theta)$ 表示, 则可以从强度  $I$ 得到局部方向信息, 其中  $\sigma$ 为尺度因子,  $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ 。而通过局部方向对比度的计算, 则可将方向特征图  $O(c, s, \theta)$ 编码成一组, 即

$$O(c, s, \theta) = |O(c, \theta) \ominus O(s, \theta)| \quad (6)$$

各种早期视觉特征计算出来后, 要考虑的问题是有任务引导的情况下, 如何生成基于任务的视觉注意显著图, 用来引导视觉注意? 有任务引导的情况下, 需要先计算各底层视觉特征之间受任务影响产生的偏置, 再由计算出的偏置和底层视觉特征来生成基于任务的视觉注意显著图。下面利用协同识别理论研究二义及多义模式的视觉感知, 首先得到协同感知理论, 并利用协同感知理论计算底层视觉特征相互间产生的偏置, 进而提出基于协同感知的视觉选择注意计算模型。

## 3 基于协同识别理论的视觉感知

20世纪 80年代末, Haken将协同学(synergetics)原理运用于模式识别<sup>[14]</sup>, 提出了一个重要观点: 模式识别的过程即为模式形成的过程。

### 3.1 协同识别理论

根据协同学的思想<sup>[14, 15]</sup>, 对待识别模式  $q$ 可以构造一个动力学过程, 使  $q$ 经过一个中间状态  $q(t)$ 进入到诸原型模式中的一个模式  $v_k$ 。该过程可以描述为  $q(0) \rightarrow q(t) \rightarrow v_k$ , 满足模式识别的方程为

$$\dot{q} = \sum_k \lambda_k v_k (v_k^+ q) - B \sum_{k \neq l} (v_k^+ q)^2 (v_l^+ q) v_k - C (q^+ q) q + F(t) \quad (7)$$

其中,  $q$ 是以输入模式  $q(0)$ 为初始值的状态向量;  $\lambda_k$ 为注意参数;  $q$ 为待识别模式;  $v_k$ 为原型模式;  $v_k^+$ 为  $v_k$ 的正交伴随向量;  $F(t)$ 为涨落力;  $B$ 与  $C$ 为指定系数。

协同识别方法可以描述为一个有势的动力学过程。若忽略式(7)中  $F(t)$ 和暂态量, 则其势函数表达式为

$$V = -\frac{1}{2} \sum_{k=1}^M \lambda_k (v_k^+ q)^2 + \frac{B}{4} \sum_{k \neq l} (v_k^+ q)^2 (v_l^+ q)^2 + \frac{C}{4} \left| \sum_{k=1}^M (v_k^+ q)^2 \right|^2 \quad (8)$$

势能函数  $V$ 由以下 3项构成: 第 1项定义了在意面上对应原型模式的极小值, 每个极小值的深度与注意参数  $\lambda_k$ 有关; 第 2项定义了各原型模式之间的竞争, 并决定了各原型模式吸引域之间的边界, 这些靠常数  $B$ 来决定; 第 3项是限制项, 由常数  $C$ 决定。

序参量的引入可使协同动力学的表达简化。定义序参量  $\xi_k = v_k^+ q$  相应的序参量动力学方程为

$$\dot{\xi}_k = \lambda_k \xi_k - B \sum_{k \neq l} \xi_k \xi_l - C \left| \sum_{k=1}^M \xi_k \right|^2 \xi_k \quad (9)$$

此时, 势函数方程为

$$V = -\frac{1}{2} \sum_{k=1}^M \lambda_k \xi_k^2 + \frac{B}{4} \sum_{k \neq l} \xi_k \xi_l^2 + \frac{C}{4} \left| \sum_{k=1}^M \xi_k \right|^2 \quad (10)$$

### 3.2 二义模式的协同感知研究

不同观察者对二义模式的每一种解释会产生不同的概率。如, 模式  $v_1$  的概率为  $P_1$ ,  $v_2$  的概率为  $P_2$ , 由此重构试验向量为

$$q = P_1 v_1 + P_2 v_2 \quad (11)$$

下面用协同势函数理论分析重构的试验向量。图 1(a)为无偏向的情况,  $P_1 = P_2 = 50\%$ , 两个模式的吸引域大小相等; 图 1(b)为有偏向情况, 两个模式的吸引域大小变得不等。

如图 2所示, 无偏向情况下, 势函数  $V$ “地形图”上沿着对角线方向的隆脊位置(即吸引域之间的分界)与平面  $\xi_1, \xi_2$ 上的对角线相对应; 有偏向的情况下, 势函数的隆脊移动, 可通过改变与平面  $\xi_1, \xi_2$ 上对角线之间的偏向角  $\alpha_0$ 来表示势函数“地形图”中的隆脊位置的移动。

无偏情况下, 二义模式的两种可能解释则分别对应两个模式, 由式(9)得

$$\dot{\xi}_1 = \xi_1 (\lambda_1 - A_1 \xi_1 - A_2 \xi_2) \quad (12)$$

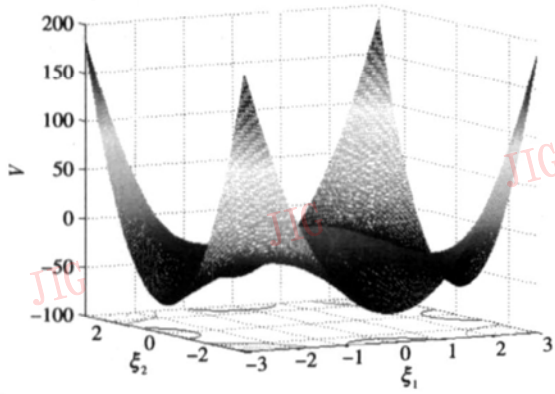
$$\dot{\xi}_2 = \xi_2 (\lambda_2 - A_2 \xi_1 - A_1 \xi_2) \quad (13)$$

其中,  $A_1 = C, A_2 = (B + C)$ 。同时导出  $\lambda_1, \lambda_2$  的微分方程如下<sup>[14, 16, 17]</sup>:

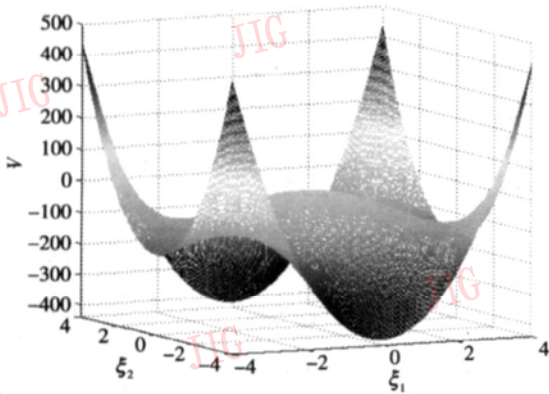
$$\dot{\lambda}_1 = \gamma (1 - \lambda_1 - \xi_1) \quad (14)$$

$$\dot{\lambda}_2 = \gamma (1 - \lambda_2 - \xi_2) \quad (15)$$

其中,  $\gamma > 0$ 表示振荡程度。



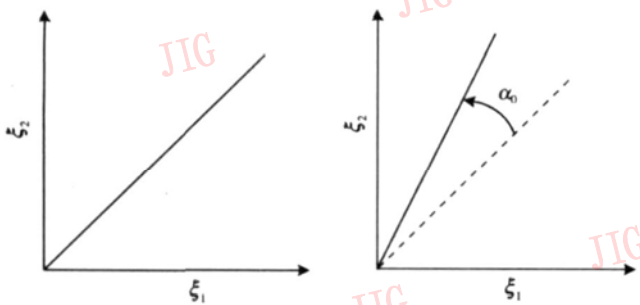
(a) 无偏向情况



(b) 有偏向情况

图 1 协同势函数地形图

Fig 1 Reliefmap of synergetic potential function



(a) 无偏向情况

(b) 有偏向情况

图 2  $\xi_1, \xi_2$  平面上对角线的位置

Fig 2 Location of diagonal line in  $\xi_1, \xi_2$  plane

通过对协同识别理论的研究表明<sup>[14-18]</sup>：动力学方程中的参数  $B$  定义了各原型模式吸引域之间的边界。对二义模式来说，通常有两个吸引子，而决定吸引域边界的势函数隆脊的位置则可由模式  $v_1$  决定，即势函数  $V$  的隆脊是由势函数  $V$  中的  $V_1$  决定的，而且，可以通过改写势函数方程 (式 (10)) 中含有  $B$  的第 2 项来表示感知的偏向。

本文用  $\alpha_0$  表示模式 2 对模式 1 偏向的角位移，

并引入以下新势函数  $\tilde{V}_1$  代替原势函数  $V_1$ ：

$$\tilde{V}_1 = \frac{B}{4} \xi_1 \xi_2 \left| 1 - 4\alpha_0 \frac{\xi_1 - \xi_2}{\xi_1 + \xi_2} \right| \quad (16)$$

若用  $\tilde{V}_1$  改写原势函数方程 (式 (10)) 中的第 2 项，则势函数变为

$$V = -\frac{1}{2} \sum_{k=1}^M \lambda_k \xi_k + \frac{B}{4} \sum_{k \neq l} \xi_k \xi_l \left| 1 - 4\alpha_k \frac{\xi_k - \xi_l}{\xi_k + \xi_l} \right| + \frac{C}{4} \left| \sum_{k=1}^M \xi_k \right|^2 \quad (17)$$

在式 (14) 式 (15) 不变的情况下，如果有偏向，则序参量的动力学方程为

$$\dot{\xi}_1 = \xi_1 \left[ \lambda_1 - A_1 \xi_1 - A_2 \xi_2 + 4(A_2 - A_1) \alpha_0 \xi_2 \left| 1 - \frac{2\xi_1}{(\xi_1 + \xi_2)} \right| \right] \quad (18)$$

$$\dot{\xi}_2 = \xi_2 \left[ \lambda_2 - A_2 \xi_2 - A_1 \xi_1 - 4(A_2 - A_1) \alpha_0 \xi_1 \left| 1 - \frac{2\xi_2}{(\xi_1 + \xi_2)} \right| \right] \quad (19)$$

当  $\nu = 0.1, A_1 = 1, A_2 = 1.8$  及  $\alpha_0$  分别为 0.02, 0.04 时，图 3(a) 图 3(b) 显示了序参量  $\xi_1$  和  $\xi_2$  及其对应的注意参数  $\lambda_1$  和  $\lambda_2$  的动力学振荡过程，图中横坐标 Iterat 表示迭代次数。在有偏感知情况下，当  $\alpha_0$  增大时，每个模式注意参数饱和 (注意的饱和定义为：当模式  $v_k$  的序参量  $\xi_k$  的幅值增加时，则对模式  $v_k$  的注意就饱和) 的时间变得不同，对偏向大的模式的注意饱和时间将变长。当  $\alpha_0$  超过临界值  $\alpha_0^{crit}$  时，感知振荡停止，即如果偏向太大，观察者只能看出一种模式。临界值  $\alpha_0^{crit}$  可以由  $A_1, A_2$  来确定，即

$$\alpha_0^{crit} = \frac{1}{4} \left| \frac{1}{A_2 - A_1} - 1 \right| = \frac{1}{4} \left| \frac{1}{B} - 1 \right| \quad (20)$$

可见式 (20) 定义了各模式吸引域边界的参数  $B$  与偏向临界值  $\alpha_0^{crit}$  之间的关系。

### 3.3 多义模式的协同感知

二义模式感知研究的结论可以推广到多义模式的情况。记多义模式的序参量为  $\xi_k$ ，注意参数为  $\lambda_k$  ( $k = 1, 2, \dots, M$ )，则可以写出以下与各个模式对应的序参量和注意参数方程：

$$\dot{\xi}_k(t) = \lambda_k(t) \xi_k(t) - 4B \xi_k(t) \sum_{k \neq l}^M \xi_l(t) - 4C \xi_k(t) \sum_k^M \xi_k(t) \quad (21)$$

$$\dot{\lambda}_k(t) = \nu [1 - \lambda_k(t) - \xi_k(t)] \quad (22)$$

$M = 2$  时，描述偏向在用角度  $\alpha_0$  来描述多义模

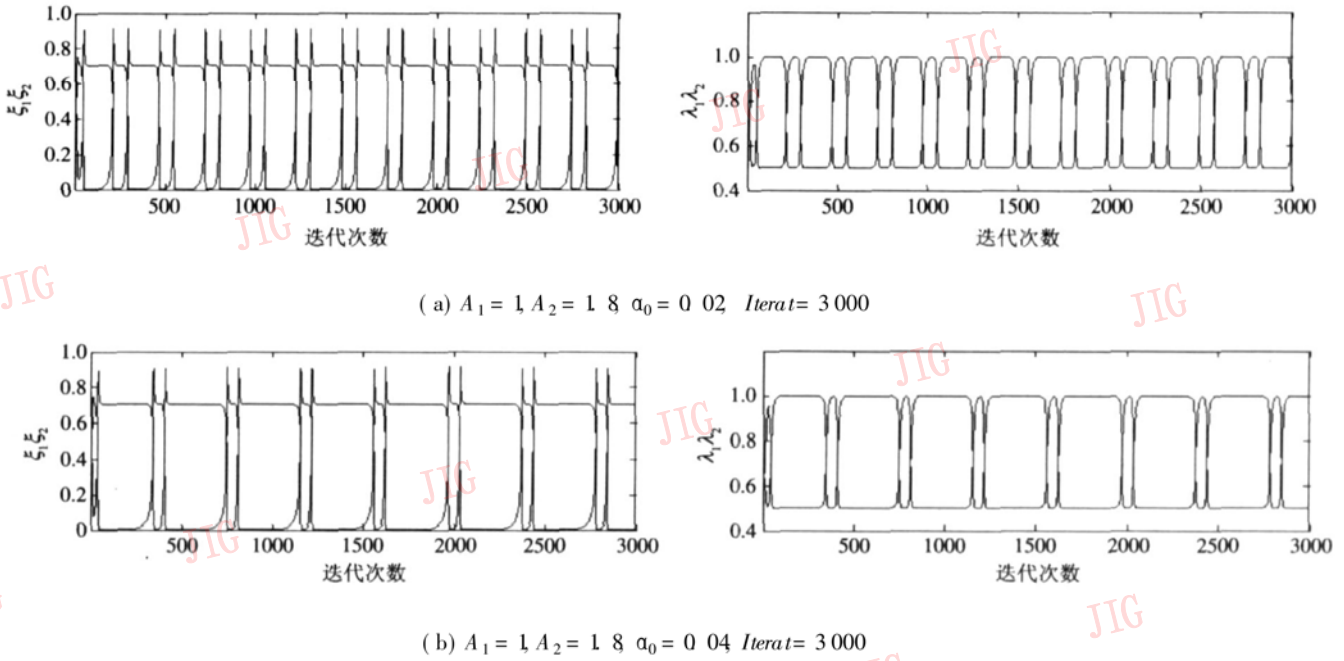


图 3 有偏向情况下, 二义模式感知的序参量及其注意参数的振荡过程

Fig 3 In bias condition, the oscillation of the order parameters and attention parameters of the ambiguous model

式时, 势函数  $V_1$  的偏向需要用矩阵  $\alpha$  来代替, 即可将势函数  $V_1$  推广至多维:

$$V_1 = \frac{B}{8} \sum_{k \neq \hat{k}} \xi_k \xi_{\hat{k}} \left| 1 - 4\alpha_{k, \hat{k}} \frac{\xi_k - \xi_{\hat{k}}}{\xi_k + \xi_{\hat{k}}} \right| \quad (23)$$

将式 (23) 带入到式 (21) 序参量  $\xi_k$  的动力学方程, 就可得到有偏向情况下的序参量  $\xi_k$  的动力学方程, 即

$$\begin{aligned} \dot{\xi}_k(t) = & \lambda_k(t) \xi_k(t) - \\ & 4B \xi_k(t) \sum_{\hat{k} \neq k} \xi_{\hat{k}}(t) \left| 1 - 4\alpha_{k, \hat{k}} \frac{\xi_k(t) - \xi_{\hat{k}}(t)}{\xi_k(t) + \xi_{\hat{k}}(t)} \right| + \\ & 32B \xi_k(t) \sum_{\hat{k} \neq k} \alpha_{k, \hat{k}} \frac{\xi_{\hat{k}}(t)}{\xi_k(t) + \xi_{\hat{k}}(t)} - \\ & 4C \xi_k(t) \sum_{\hat{k}} \xi_{\hat{k}}(t) \end{aligned} \quad (24)$$

### 3.4 偏置矩阵的性质

将图 2 中二义模式的偏置扩展至多义模式, 即得到

$$\tan \left| \alpha_{k, \hat{k}} + \frac{\pi}{4} \right| = \frac{\xi_k}{\xi_{\hat{k}}} \quad (25)$$

进而可以得到

$$\begin{aligned} \alpha_{k, \hat{k}} = & \arccot \left| \frac{\xi_k}{\xi_{\hat{k}}} \right| - \frac{\pi}{4} \\ = & \frac{\pi}{2} - \arctan \left| \frac{\xi_k}{\xi_{\hat{k}}} \right| - \frac{\pi}{4} \\ = & - \left| \frac{\pi}{4} - \arctan \left| \frac{\xi_k}{\xi_{\hat{k}}} \right| \right| \\ = & - \alpha_{k, \hat{k}} \end{aligned} \quad (26)$$

可以证明:  $\alpha_{k, \hat{k}} = \alpha_{\hat{k}, k} - \alpha_{k, k}$ 。偏置矩阵  $\alpha$  是一个对角分量为零的反对称矩阵, 为了能确定偏置矩阵的值, 需要知道  $\alpha$  中  $M-1$  个分量。设分量  $\alpha_{k, \hat{k}}$  ( $\hat{k} = 1, 2, \dots, M$ ) 已知, 其中,  $\alpha_{1, 1} = 0$  并可由式 (25) 可以计算  $k=1$  时的序参量

$$\xi_k = \xi_{\hat{k}} \tan \left| \alpha_{1, \hat{k}} + \frac{\pi}{4} \right|, \quad x = k, \hat{k} \quad (27)$$

对于其他分量  $\alpha_{k, \hat{k}}$  ( $k \neq 1$ ), 计算如下

$$\alpha_{k, \hat{k}} = \arctan \left| \frac{\tan(\alpha_{1, \hat{k}} + \pi/4)}{\tan(\alpha_{1, \hat{k}} + \pi/4)} - \frac{\pi}{4} \right| \quad (28)$$

然后就可由已知的  $\alpha_{k, \hat{k}}$  ( $\hat{k} = 1, 2, \dots, M$ ) 计算偏置矩阵中余下的其他分量。

## 4 基于协同感知的视觉注意计算模型

在任务相关的视觉注意中, 需要建立基于任务的视觉注意显著图, 本文将这种基于任务的视觉注意显著图称为任务显著图。建立任务显著图的一个问题就是: 如何以符合人感知过程的标准来度量各底层视觉特征间受任务影响而产生的偏置, 并由此偏置和底层视觉特征来生成任务显著图。

由对协同感知理论的研究可以看出, 协同模式感知与人的感知过程有着深刻的相似之处, 其体现了人在认知过程中如何通过各模式之间的竞争发展和协同合作来得到需要的模式。可以利用已研究得到的协同感知理论来计算各底层视觉特征间因受任

务影响而产生的偏置。若将协同视觉感知中的模式与提取的底层视觉特征相对应,即若提取了 7 种底层视觉特征——强度、颜色 (RG, BY)、方向 (0°, 45°, 90°和 135°),则可以看作具有 7 种解释的多义模式感知问题,可先采用计算多义模式感知偏置矩阵的方法来计算这 3 种底层视觉特征间受任务影响而产生的偏置;然后就可由得到的偏置和底层视觉特征来共同生成任务显著图。

生成任务图时,首先讨论视觉任务。视觉任务一般由语义给出,实际应用中,任务是多样性的,比如在一幅图像中寻找一张桌子、一辆汽车、一个穿红色衣服的人等。为了和所提取的底层视觉特征相对应,可将给定的任务分解为颜色、方向或者亮度方面的语义信息。本文提取的底层视觉特征有亮度特征图、颜色特征图和方向特征图,其中颜色特征图分为红绿通道和蓝黄通道,方向特征图有 0°, 45°, 90°和 135° 4 个方向,本文是采用  $M = 7$  个多义模式的协同感知来计算 7 个特征图受任务影响产生的偏置。

由协同感知偏置矩阵的性质可知,需要确定偏置矩阵中  $M - 1 = 6$  个偏置值,每个偏置的临界值都以振荡停止为前提,即先由任务给定的某一显著特

征对其他特征的偏置值满足感知振荡停止的条件:  $\alpha > \alpha_0^{crit}$ ,然后用式 (27)、式 (28) 计算偏置矩阵中各元素的值。

本文提出的基于协同感知的视觉选择注意计算模型的算法框图如图 4 所示,算法具体描述如下:

- (1) 对输入图像,首先建立图像的高斯金字塔尺度表示模型,文中是采用 5 个空间尺度表示,即  $\sigma \in \{0, 1, 2, 3, 4\}$ ,其中,  $\sigma = 0$  即为原始输入图像,尺度每增加一个,表示对低一级尺度图像进行一次高斯滤波和降采样;
- (2) 选择两组“中央-周边”尺度因子  $c = 1, s = 3$  和  $c = 2, s = 4$  同时利用“中央-周边差”算子计算特征图,并归一化到尺度  $\sigma = 0$  即得到 1 个强度, 2 个颜色, 4 个方向特征图;
- (3) 分析给定的任务,将其与提取的 7 种底层视觉特征中某一个对应,再利用协同模式感知理论来计算与任务相关的特征和其他特征之间的偏置;
- (4) 由计算得到的偏置矩阵的值和各底层视觉特征图来计算任务显著图;
- (5) 利用 Winner-Take-All 机制选择显著图中的显著区域,并用圆圈来标注显著区域,圆圈中的数字表示该显著区域被注意的次序;

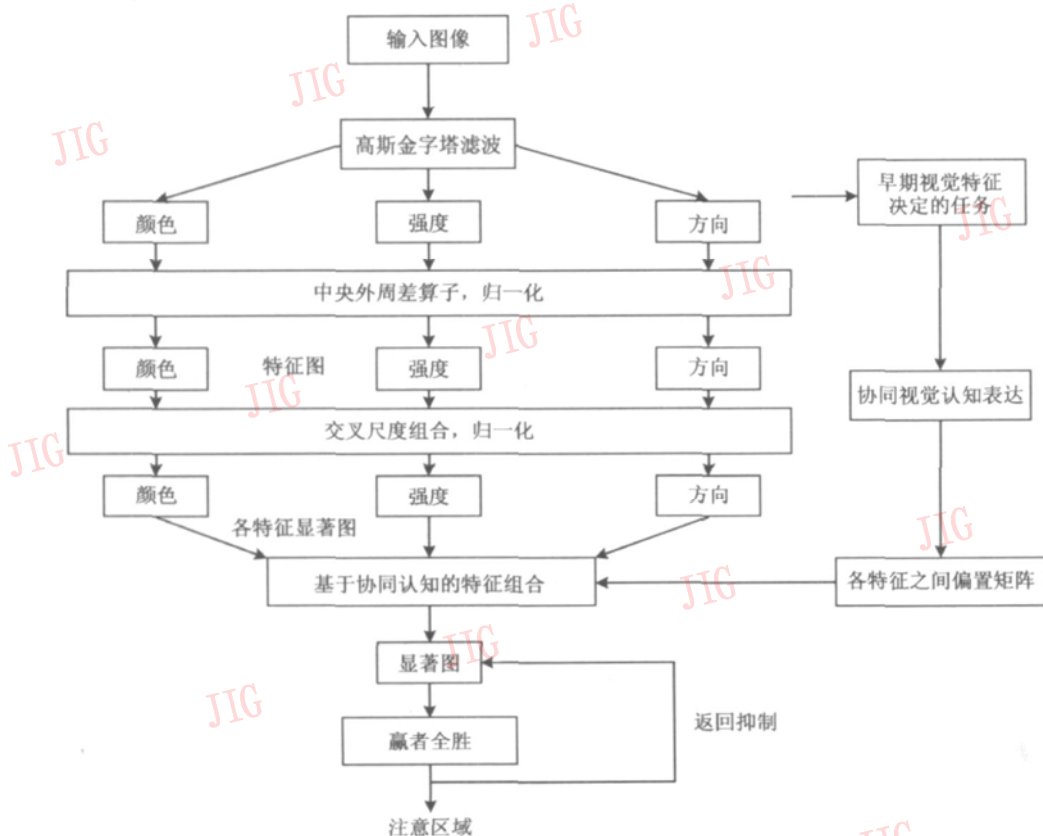


图 4 基于协同感知的视觉选择注意计算模型

Fig 4 Visual selective attention computational model based on synergistic perception

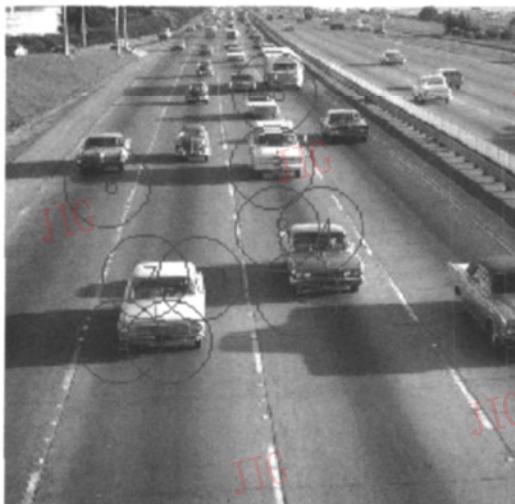
(6) 由返回抑制机制 (inhibition of return, IR) 控制视点转移, 直至找到所有的显著区域。

将文中的计算模型应用于基于任务的视觉搜索实验。实验中, 原始输入图像是含有汽车的道路场景图像, 给定的任务是搜寻图像中的汽车。搜寻时, 首先分析给定的语义信息, 对于汽车, 根据经验, 一般都具有水平方向的显著性, 因此可将给定的任务

搜索汽车分解为提取的底层视觉特征中的水平方向特征图, 并将感知偏向于水平方向特征图; 然后利用协同模式感知理论来计算水平方向特征图对其他特征图的偏置值, 最后由计算出的偏置值和底层视觉特征图生成任务显著图。图 5 为实验结果, 图中的显著区域用圆圈标注, 圆圈内的数字为该显著区域被注意的次序, 控制视点转移 10 次。



(a) 文献 [13]方法计算的显著区域



(b) 本文算法选择出的显著区域

图 5 文献 [13]方法和本文方法选择显著区域对比

Fig 5 The comparison of selected salient regions based on the proposed algorithm in reference [13] and in this paper

从实验结果可以看出, 基于协同感知的视觉注意计算模型所选择出的显著区域, 基本上与任务相关, 而对其他一些显著区域则没有注意。另外, 从各个显著区域被注意的次序可以看出, 由于任务的引导, 与任务相关的显著区域可以更快地被注意到。从实验结果可得出本文算法具有以下优点:

(1) 与基于图像底层视觉特征的模型相比, 本

文算法可以更快而有效地选择出与任务相关的显著区域:

(2) 将给定的视觉任务与底层视觉特征相对应, 就避免了对给定任务复杂的表达过程;

(3) 运用协同感知理论来计算底层视觉特征因受任务影响而产生的偏置, 并可由得到的偏置和底层视觉特征生成任务显著图, 算法与人感知过程有

深刻的相似之处。

## 5 结论与讨论

本文利用协同感知理论研究了基于任务的视觉注意显著图的生成,并提出了基于协同感知的视觉选择注意计算模型,即首先研究了协同识别理论在二义及多义模式感知时的特性,然后利用协同感知理论计算底层视觉特征间受任务影响而产生的偏置,生成了基于任务的显著图。该模型从视觉选择注意机制出发,利用与人感知过程有深刻相似之处的协同识别理论来研究视觉感知,并建立了基于协同感知的视觉选择注意计算模型。实验结果表明,该算法是有效的,且与人的感知具有相似性。

在基于任务的视觉注意计算模型中,如何更好地定义知觉物体,如何利用与人类视觉生理特性和认知过程更符合的多尺度来选择显著区域,以及更准确地度量高层知识对各底层视觉特征之间造成的偏置等,这些问题还有待进一步研究。

### 参考文献 (References)

- 1 Koch C, Ullman S. Shifts in selective visual attention: Towards the underlying neural circuitry [J]. *Human Neurobiology*, 1985, **4**: 219~ 227.
- 2 Itti L. Models of Bottom-Up and Top-Down Visual Attention [D]. Los Angeles, California, USA: California Institute of Technology, 2000.
- 3 Mahmood F. Attentional Selection in Object Recognition [D]. Boston, Massachusetts, USA: Massachusetts Institute of Technology, 1993.
- 4 Sun Yao-nu. Hierarchical Object-Based Visual Attention for Machine Vision [D]. Edinburgh, UK: University of Edinburgh, 2003.
- 5 Jagersand M. Saliency maps and attention selection in scale and spatial coordinates: An information theoretic approach [A]. In Proceedings of Fifth International Conference on Computer Vision [C], Boston, MA, USA, 1995. 195~ 202.
- 6 Ferraro M, Boccignone G, Caelli T. Entropy-based representation of image information [J]. *Pattern Recognition Letters*, 2002, **23**(12): 1391~ 1398.
- 7 Ferraro M, Boccignone G, Caelli T. Generalized spatio-chromatic diffusion [J]. *Pattern Recognition Letters*, 2002, **24**(10): 1298~ 1309.
- 8 Kadir T. Scale, Saliency and Scene Description [D]. Oxford, UK: University of Oxford, 2002.
- 9 Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope [J]. *International Journal of Computer Vision*, 2001, **42**(3): 145~ 175.
- 10 Oliva A, Torralba A, Castelhanom, *et al*. Top-down control of visual attention in object detection [A]. In Proceedings of the IEEE International Conference on Image Processing 2003 (ICIP'03) [C], Barcelona, Spain, 2003. 1: 253~ 256.
- 11 Torralba A. Contextual priming for object detection [J]. *International Journal of Computer Vision*, 2003, **53**(2): 169~ 191.
- 12 Torralba A. Modeling global scene factors in attention [J]. *Journal of Optical Society of America*, 2003, **20**(7): 1407~ 1418.
- 13 Itti L, Koch C, Niebur E. 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.
- 14 Haken H. *Synergetic Computers and Recognition—A Top-down Approach to Neural Nets* [M]. Berlin: Springer-Verlag, 1991.
- 15 Gao Jun. *The Theory of Artificial Neural Networks and Simulation* [M]. Beijing: China Machine Press, 2003. [高隽. 人工神经网络原理及仿真实例 [M]. 北京: 机械工业出版社, 2003.]
- 16 Haken H. A brain model for vision in terms of synergetics [J]. *Journal of Theoretical Biology*, 1994, **171**: 75~ 85.
- 17 Ditzinger T, Tuller B, Haken H, *et al*. A synergetic model for the verbal transformation effect [J]. *Biological Cybernetics*, 1997, **77**(1): 31~ 40.
- 18 Gao Jun, Dong Houming, Shao Jing, *et al*. Parameters optimization of synergetic recognition approach [J]. *Chinese Journal of Electronics (English)*, 2005, **14**(2): 192~ 197.