

## 基于视觉注意模型和进化规划的感兴趣区检测方法

张菁 沈兰荪 高静静

(北京工业大学信号与信息处理研究室 北京 100124)

**摘要:** 根据生物注意机制, 该文提出了一种基于视觉注意模型和进化规划的感兴趣区检测方法。采用进化规划方法分割图像候选区域; 区域兴趣度由视觉注意模型产生的局部显著和进化规划计算的全局显著共同度量。在视觉注意模型中, 图像经过小波多尺度变换和计算中央周边差得到局部显著度。注意焦点在显著度增强因子的作用下, 选取候选区域得到感兴趣区。实验结果表明, 所提方法检测的感兴趣区更接近人眼的视觉注意机制, 并取得了较为满意的对象检测和兴趣度量结果。

**关键词:** 计算机视觉; 视觉注意模型; 感兴趣区; 进化规划; 显著度; 注意焦点

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2009)07-1646-07

## Region of Interest Detection Based on Visual Attention Model and Evolutionary Programming

Zhang Jing Shen Lan-sun Gao Jing-jing

(Signal & Information Processing Lab, Beijing University of Technology, Beijing 100124, China)

**Abstract:** According to biological attention mechanism, a region of interest detection based on visual attention model and evolutionary programming is proposed in this paper. Candidate image regions are segmented with evolutionary programming algorithm. The interest of a region is measured with local saliency produced by visual attention model and global saliency based on evolutionary programming. After wavelet multiscale transform, local saliency in visual attention model is computed by center-surround differences. Under the action of saliency enhancement factors, focus of attention selects region of interest from candidate regions. The experimental results show the proposed approach is closer to human attention mechanism and performs well for object detection and interest measurement.

**Key words:** Computer vision; Visual Attention Model (VAM); Region Of Interest (ROI); Evolutionary Programming (EP); Saliency; Focus of Attention (FOA)

### 1 引言

“视觉感兴趣区(Regions Of Interest, ROI)”是人们观察和理解图像时产生兴趣(关注或注意)的区域<sup>[1]</sup>。文献[2]进一步指出, 所谓感兴趣区是指图像中最能引起人们兴趣, 最能表现图像内容的区域。基于感兴趣区的图像表示方法区分出了区域的重要程度, 在一定程度上消除了图像的冗余信息, 突出了图像的主要内容, 从而在一定程度上弥补了语义鸿沟。而如何从一幅图像中自动的提取出感兴趣区, 就是感兴趣区检测技术<sup>[1]</sup>。

生物体特别是人类拥有强大的图像理解和模式识别能力, 建立由生物学启发下的计算机视觉模型一直是图像处理领域中的一个研究热点<sup>[3]</sup>。依据生物视觉理论, 视觉注意机制 (visual attention

mechanism)首先由视觉内容中最“与众不同”的部分, 即与周围其他部分相比更加显著的部分引起, 然后根据个体兴趣, 去选择注意<sup>[4]</sup>。现有研究表明, “显著”(saliency)的产生是由于视觉对象本身具有某种能够引起观察者注意的特殊属性, 并用“显著度”表示显著的程度或显著值的大小, 显著部分称为显著区。建立视觉注意力模型 (Visual Attention Model, VAM)是为了利用视觉注意机制得到图像中最容易引起注意的部分, 并用一幅灰度图像表示这种显著程度。已有研究表明, 人们往往对其感兴趣的对象付诸更多的注意。使用视觉注意力模型得到图像中的显著度较高的区域作为感兴趣区, 是一种行之有效的检测方法<sup>[5]</sup>。

基于视觉注意的 ROI 检测方法由于加入了视觉注意机制模拟人的视觉感知过程, 使得检测结果更符合人们的生理特征, 取得了明显的进展。但是在检测感兴趣区的过程中没有考虑图像对象信息, 因此检测的结果和实际对象仍有差异。

2008-06-03 收到, 2009-02-27 改回

国家自然科学基金(60772069, 60402036)和北京市自然科学基金(4062006)资助课题

人们往往对图像中包含的对象或目标感兴趣,对象特征是实现基于感兴趣区图像检索的基本前提。这些对象通常是一些具有实际物理意义的物体,对象的检测依赖于良好的图像分割技术。本文将基于视觉注意和基于对象的 ROI 检测方法相结合,提出了一种自动提取图像 ROI 的检测方法,一方面有效利用了视觉注意机制度量区域兴趣,另一方面可以获取具有实际物理意义的区域。

## 2 感兴趣区检测方法

最经典的视觉注意模型是由美国加州大学 Itti<sup>[6]</sup>于 1998 年依据 Treisman 等的特征整合理论提出的,简称 Itti 模型。Itti 模型从输入图像中提取多方面的特征,如亮度、颜色、方向等,形成各个特征维上的关注图(conspicuity map),然后分析、合并得到显著图(saliency map)。关注图根据特征的中央周边(center-surround)差计算得到,显著图则由各个特征关注图融合而成。

本文对 Itti 模型进行了改进,并用进化规划方法分割图像对象,提出了一种基于视觉注意模型和进化规划的感兴趣区检测方法。

改进的视觉注意模型(图 1)主要包括 3 部分:(1)视觉显著度量:由局部显著和全局显著共同度量,局部显著度量用改进的 Itti 模型方法,全局显著度量用进化规划方法,视觉显著度用一幅灰度图(称为显著图)表示,每个像素灰度值表示输入图像相应位置的显著度,此时显著图中的像素点称为显著点;(2)显著区的生成:由注意焦点选择候选区域作为显著区,输出区域兴趣度最大的 ROI;(3)注意焦点的选择和转移:显著图中的显著点经过赢家取全神经网络竞争产生注意焦点,然后使用视觉显著度增强因子进行修正,在返回抑制的作用下,继续寻找显著图中新的注意焦点,以得到下一个 ROI。

本文对 Itti 模型的改进主要体现在(1)视觉显著度量方法:由进化规划方法定义全局显著度,改进的 Itti 模型定义局部显著度,两者合并计算视觉显著度;(2)显著区生成选用进化规划方法确定的候选区域,这些候选区域是由一系列图像对象组成的;

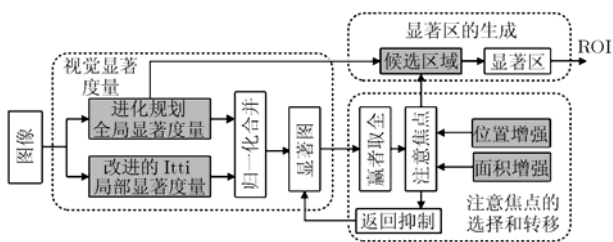


图 1 感兴趣区检测方法

(3)注意焦点使用显著度增强因子进行修正。如图 1 中灰色框部分。

## 3 视觉显著度量

视觉显著度表示的是图像特征的显著程度。在 Itti 模型中,显著度仅依据对象区域及其邻域的局部信息,没有考虑图像的全局信息,因此,本文提出对图像分别计算局部显著度和全局显著度,归一化合并得到视觉显著度,并用显著图表示输入图像的特征显著程度。

### 3.1 局部显著度量

Itti 模型从视觉特征提取、中央周边差、归一化合并等方面较好地模拟了人眼视觉系统的注意机制,其显著度量依据的是对象区域及邻域等局部信息,因此,本文采用 Itti 模型的方法计算局部显著度,并将 Itti 模型中的显著图称为局部显著图,如图 2 所示。

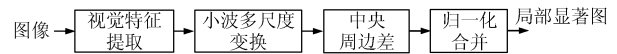


图 2 局部显著性度量

在多尺度变换部分,常用的方法有:高斯金字塔方法和小波变换的方法。由于小波变换利用人眼视觉系统在响应频带与空间方向选择上敏感性不同,将原图像在独立频带与不同空间方向上进行分解,因而更加符合人眼的视觉特性,为此,本文采用基于小波的多尺度变换替换 Itti 模型的高斯金字塔。

(1)视觉特征提取 设  $r, g, b$  分别对应输入图像的红、绿、蓝通道,则亮度特征为:  $I = (r + g + b)/3$ 。建立 4 个宽调谐的颜色通道作为颜色特征:红色  $R = r - (g + b)/2$ , 绿色  $G = g - (r + b)/2$ , 蓝色  $B = b - (r + g)/2$  和黄色  $Y = (r + g)/2 - |r - g|/2 - b$ 。方向特征为 Gabor 小波在  $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$  4 个方向分量。

(2)小波多尺度变换 人眼的视觉特性是一个多信道模型,具有多信道分解特性。图像经小波变换后,形成了一种锥形数据结构,它具有与人眼视觉特性相匹配的特点。在此提取每级小波变换后的低频子图进行中央周边差的计算。小波分解采用的低通滤波器系数为:  $\{0.602949, 0.26686411, -0.0782233, -0.016864118, 0.02674876\}$ ; 高通滤波器系数为:  $\{1.11508704, -0.59127176, -0.05754353, 0.091271762, 0.0\}$ 。

(3)中央周边差 经过中央周边差计算(式(1))以及归一化之后,得到的各个特征关注图如图 3 所

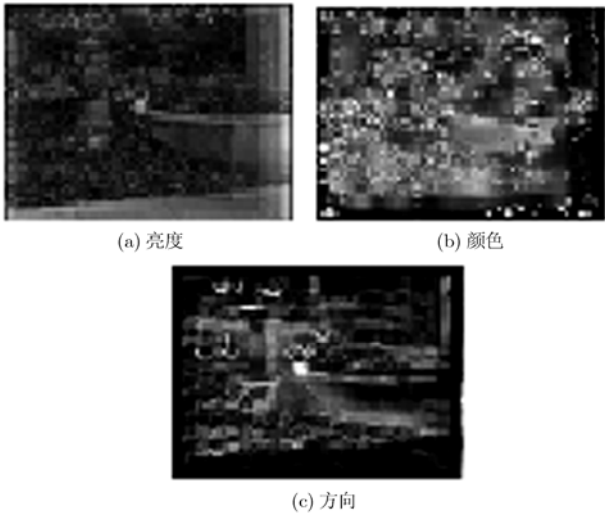


图3 特征关注图

示。

$$DOG(x, y) = \frac{1}{2\pi\sigma_c^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_c^2}\right) - \frac{1}{2\pi\sigma_s^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_s^2}\right) \quad (1)$$

(4)归一化组合 将亮度、颜色和方向3个特征上的关注图  $\tilde{I}, \tilde{C}, \tilde{O}$  归一化(归一化因子  $N(\cdot)$ )组合得到局部显著图  $S_l$ :

$$S_l = \frac{1}{3}[N(\tilde{I}) + N(\tilde{C}) + N(\tilde{O})] \quad (2)$$

这是一幅灰度图像, 其灰度值表示了局部显著度, 如图4所示。

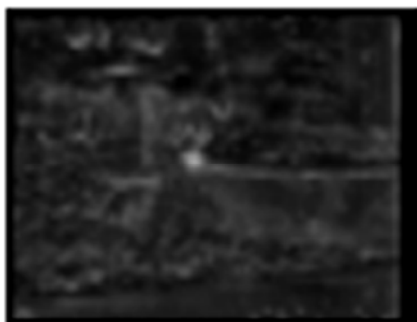


图4 局部显著图

### 3.2 全局显著度量

进化规划(Evolutionary Programming, EP)是由 Fogel 于 1962 年<sup>[7]</sup>提出的, 其基本思想是从整体的角度出发来模拟生物的进化过程, 它着眼于整个群体的进化, 强调的是“物种的进化过程”<sup>[8]</sup>。基于上述思想, 可以将整幅图像看成是一个生物圈, 其中有  $n$  个生物种群, 这些种群在一个有限的空间内为

了争夺有限的资源而生存, 在某一地区, 如果某种种群占有绝对优势, 那么其他种群在这一地区就处于劣势, 有可能被种群所排挤, 而如果各种种群在此地区内都不能占优势, 那么这一地区的各种竞争将相对残酷, 此时个体变异的可能性较大。

Stentiford<sup>[9]</sup>依据图像中邻域间的不相似性, 采用进化规划的查找方法, 提出了一种视觉注意度量方法。在 Stentiford 模型中, 通过随机对比其它  $t(t=100)$  个邻域结构计算视觉注意值, 这种随机对比使得视觉注意值具有不确定性。此外, 该模型得到的显著区不能覆盖整个对象区域。

为了去除 Stentiford 模型中随机选取产生的不确定性, 本文提出采用对比每个像素及其邻域结构, 由此产生的视觉注意值即为全局显著度。具体计算步骤如下。

(1)对图像进行下采样 由于对比全部像素会增加计算量, 因此先对原始图像进行下采样。通过双线性下采样将原始图像中的多个像素映射为图像对象中的单个像素, 以得到图像对象。

(2)计算像素的显著度 依据进化规划的原理, 图像中的每个像素点视为一个个体, 取其周围半径为 3 的  $7 \times 7$  邻域作为一个种群, 则整幅图像构成一个种群。在一个像素种群中随机挑选 3 个像素构成一种布局(也称为分布模式), 依次比较其他像素种群与这种布局或分布模式的差异, 相同为“0”, 不同则为“1”。这些差异的累积和表示了该像素与其他像素的差别, 即该像素种群在整幅图像中出现的频率。差别越大, 意味该像素的显著度越高。图像中像素  $(x, y)$  的显著度  $S'_g(x, y)$  为

$$S'_g(x, y) = \sum_{i=0}^{H \times W - 1} m_i, \quad m_i = \begin{cases} 1, & \text{不相似} \\ 0, & \text{相似} \end{cases} \quad (3)$$

其中  $H$  和  $W$  分别为采样后图像的高和宽,  $(H \times W)$  表示图像的像素个数,  $(x, y)$  表示像素点的坐标位置,  $m_i$  表示像素  $(x, y)$  种群与其他像素种群的布局差异, 相似为“0”, 否则为“1”, 此处比较的是像素的 RGB 值。用式(3)表示的图像记为  $S'_g$  图。

(3)计算全局显著度 在得到所有像素的显著度后, 归一化计算每个像素点的全局显著度  $S_g(x, y)$ :

$$S_g(x, y) = \frac{S'_g(x, y) - S'_{g\_min}}{S'_{g\_max} - S'_{g\_min}} \times 255 \quad (4)$$

其中  $S'_{g\_min}$  为  $S'_g$  图中的最小值,  $S'_{g\_max}$  为  $S'_g$  图中的最大值。用式(4)表示的全局显著图记为  $S_g$  图, 如图 5 所示。

### 3.3 归一化合并计算视觉显著度

令局部显著图记为  $S_l$ , 全局显著图记为  $S_g$ , 对



图5 全局显著图

二者进行归一化合并得到总的视觉显著图  $S$ (图6), 计算公式如下( $N(\cdot)$ 为归一化因子):

$$S = \frac{1}{2}[N(S_l) + N(S_g)] \quad (5)$$



图6 显著图

#### 4 显著区的生成

区域检测的目的是将图像中感兴趣的对象区域从背景中提取出来。常用的区域检测方法包括灰度阈值法, 边缘检测法等。这些方法要求对象和背景之间有明显的灰度变化。但是在实际情况中, 由于光照、对焦等影响, 这种明显的灰度变化很难出现。进化规划也是一种区域检测方法, 它是通过模拟自然进化过程而得到最优解的一种随机的进化方法<sup>[10]</sup>。

进化规划是通过抑制图像中具有相同模式的对象以区分显著和不显著区域的。当图像某候选区域在其他区域中出现频率越少时, 则该候选区域的兴趣度就越高, 反之, 兴趣度就越低。本文采用进化规划的方法, 从全局寻优性来提取候选区域, 为生成显著区做准备。

对于上节讨论的灰度显著图, 当候选区域出现的频率小于一定的阈值时, 即图像的灰度值大于一定值  $T$  时, 则该处的兴趣度为“1”, 反之, 为“0”, 其公式如下:

$$g(x, y) = \begin{cases} 1, & S_g(x, y) > T \\ 0, & S_g(x, y) \leq T \end{cases} \quad (6)$$

阈值处理后的图像, 会存在一些孔洞, 对于较小的孔洞可以采用形态学闭运算的方法来去除; 对于较大的孔洞, 本文用边缘为种子的区域填充方法去除。

#### 5 注意焦点的选择和转移

注意焦点(Focus Of Attention, FOA)是视觉显著度最强的点, 是场景中需要首先关注的地方。注意焦点的选择和转移决定了感兴趣区的位置及其重要程度。

注意焦点的选择和转移首先通过赢者取全(Winner-Take-All, WTA)神经网络从显著图中找到注意焦点, 选取候选区域得到显著区, 然后引入对象面积、位置等显著度增强因子, 在返回抑制机制的作用下, 进行注意焦点的转移。

##### 5.1 注意焦点的选择

(1)WTA 神经网络选择注意焦点 WTA 网络<sup>[11]</sup>是选择注意模型中经常被采用的网络结构。该结构利用神经元间的相互竞争快速找到具有最大值的神经元, 此神经元被称为获胜者。然后其输入将传入下一级进行处理, 而其他神经元的输入将被抑制。

选取注意焦点时, 显著图由一个二维的积分发放神经元阵列实现, 神经元的输入电流是显著图对应点的初始值。显著图中每个神经元的电位通过电阻转换成 WTA 网络中神经元的输入电流。WTA 网络也是一个二维的积分发放神经元阵列, 其神经元的时间常数比显著图中神经元的时间常数小, 电位上升得比显著图中的神经元快。把时间离散化为固定间隔的时间段。每个时间段内显著图的神经元的输入电流在电容上积分(充电), 电位升高。然后显著图的输出电流在 WTA 网络中的神经元上进行一次充电。这样, WTA 网络的神经元总是比显著图的神经元先产生发放, 最先产生发放的神经元就对应于显著图中显著值最大的神经元。该神经元的位置点就是注意焦点。

(2)显著区的生成 目前大多数方法是在显著图的基础上用区域增长或者  $k$ -均值聚类的方法得到显著区, 这类方法一般将注意焦点作为区域增长或  $k$ -均值聚类的起始点(也称为种子点)形成区域, 但是由于显著图的界限比较模糊, 很难准确的选取种子点。此外, 每幅图像的显著区数目不尽相同, 种子的数目也难以确定。为了解决以上问题, 本文提出由注意焦点选取候选区域的方法得到显著区, 即当注意焦点在候选区域中时, 这个区域被认为是显著区; 当注意焦点不在任何候选区域时, 则转移到下一个注意焦点进行处理; 依次重复, 直到候选区域

选取完毕。由以上过程可知，每个注意焦点均有一个与其相对应的显著区。

## 5.2 注意焦点的转移

已有研究表明，注意焦点倾向于转移到与当前注意焦点比较接近的位置，因此注意焦点的转移应遵循“邻近优先”的原则。另外，对象大小也是影响视觉注意的一个较为重要的因素。因此，本文提出引入对象面积、位置等显著度增强因子，在返回抑制机制的作用下，影响注意焦点的转移。

(1)返回抑制 根据返回抑制机制的特点<sup>[12,13]</sup>，在注意焦点转移中，已经被选择过的显著区域的显著性会受到抑制。如式(7)所示，在进行第 $k$ 次注意焦点转移时，将当前显著图 $S$ 中属于第 $(k-1)$ 次显著区 $SR$ 内的像素值 $(x,y)$ 均设为“0”：

$$S^0(x,y) = 1, S^k(x,y) = \begin{cases} 0, & (x,y) \in SR^{k-1} \\ S^{k-1}(x,y), & \text{其它} \end{cases} \quad (7)$$

(2)位置增强因子 根据邻近优先性<sup>[12,13]</sup>，在注意焦点转移中，与当前显著区接近的区域显著度会受到增强。为此，用位置增强(Position Enhancement, PE)因子(取值0-1)以调整接近区域的显著度：在进行第 $k$ 次注意焦点转移时，离第 $(k-1)$ 个注意焦点位置(Pos $X^{k-1}$ , Pos $Y^{k-1}$ )越近，其显著度就越高，用高斯函数表示为

$$PE^k(x,y) = \alpha_k G(x - \text{Pos}X^{k-1}, y - \text{Pos}Y^{k-1}, \sigma_k) \quad (8)$$

其中 $G(\cdot)$ 为高斯函数； $\sigma_k$ 为高斯函数的标准偏差，与显著区大小成一定的比例； $(x,y)$ 为坐标点； $\alpha_k$ 为加权系数，控制邻近优先性的显著程度； $PE^k(x,y)$ 表示选取第 $k$ 个注意焦点时，第 $k$ 个注意焦点到第 $k-1$ 个注意焦点的位置。

(3)面积增强因子 对象大小也是影响视觉注意的重要因素：通常大对象比小对象更容易吸引视觉注意，当对象面积超过饱和度时，其重要性又会明显降低。为此，用面积增强(Area Enhancement, AE)因子(取值1-2)来表示：对象越大，则显著度就越高；当对象面积不到图像1%时，认为其显著度一样。区域 $k$ 的面积增强因子 $AE(k)$ 为

$$AE(k) = 1 + b_k \frac{\max(\text{Area}(k)/AR, 1)}{\sum_j \max(\text{Area}(j)/AR, 1)} \quad (9)$$

其中 $\text{Area}(k)$ 为第 $k$ 个区域的面积； $AR$ 是一个常量，其值为图像总面积的1%，目的是为了实现区域面积的饱和度，防止给予面积过大的区域的权值过高； $b_k$ 为加权系数，控制尺寸的影响程度。

## 5.3 区域兴趣度量

通过以上分析可知，兴趣度量是以视觉显著度为主，受对象面积、位置增强因子的影响。因此，

本文提出区域 $k$ 的兴趣度量为

$$\text{interest}(k) = S^k \times PE^k \times AE(k) \quad (10)$$

其中 $PE^k$ 表示区域 $k$ 的位置增强因子， $AE(k)$ 为区域 $k$ 的面积增强因子， $S^k$ 表示第 $k$ 个注意焦点的视觉显著度。最后将其归一化到(0~255)之间即为最终的兴趣度。

## 6 感兴趣区检测实验结果与分析

为了验证上述算法的有效性，使用 Pentium(R) D3.00GHz CPU 以及 1GB 内存的计算机对含有一个对象和多个对象的大约 500 幅图像(植物、鸟、马、标志、飞机等几类)进行了实验。

### 6.1 算法结果

感兴趣区及注意焦点的转移顺序如图 7 所示，两个对象的白色连线表示了注意焦点的转移，起点是第 1 个感兴趣区，终点指向了下一个感兴趣区。为了方便后面的测试，本文根据算法得出的注意焦点的转移顺序依次将其区域标记为 1, 2, 3, ... 等，并且作为后面测试用的区域标记符号。从实验结果来看，本文方法检测的感兴趣区更接近对象边界，检测一个对象时间大约需要 1 s，几乎接近 Itti 模型的运算时间。此外，由于采用的进化规划方法对比的是图像颜色，当对象中存在与背景接近的颜色区域时，出现了欠分割的现象，如图 7 中第 3 行“马腿”以及“花”图像的“绿叶”。对于这类图像，本文方法仍难以做到精确分割。

表 1 是对含有多个对象图像进行兴趣度量的部分测试数据(对应图 7 中第 3~12 幅图像顺序)，其中位置增强因子是某一对象与前一对象之间的位置信息，在选择第 1 个注意焦点时，以图像中心为参照。

### 6.2 兴趣度量评价

为了验证所提检测方法在首次感兴趣区和注意转移顺序与人主观观察结果的符合率，本文采用主观的区域兴趣度评价方法对 5 位用户进行了测试，要求每位用户分别指出每幅测试图像的兴趣区及注意顺序，然后统计 Itti 模型和本文方法在检测首次感兴趣区及转移顺序的符合率，结果如表 2 所示。

由表 2 可以看出，本文方法在首次感兴趣区和注意转移顺序的符合率均比 Itti 模型有明显提高，也进一步说明所提方法的检测结果更接近人眼的视觉注意机制。这是由于：(1)Itti 模型仅利用局部的显著性控制注意焦点的转移，而本文方法加入了全局信息，并且引入对象大小和位置信息的影响；(2)Itti 模型中的“返回抑制性”不能完全抑制已选择

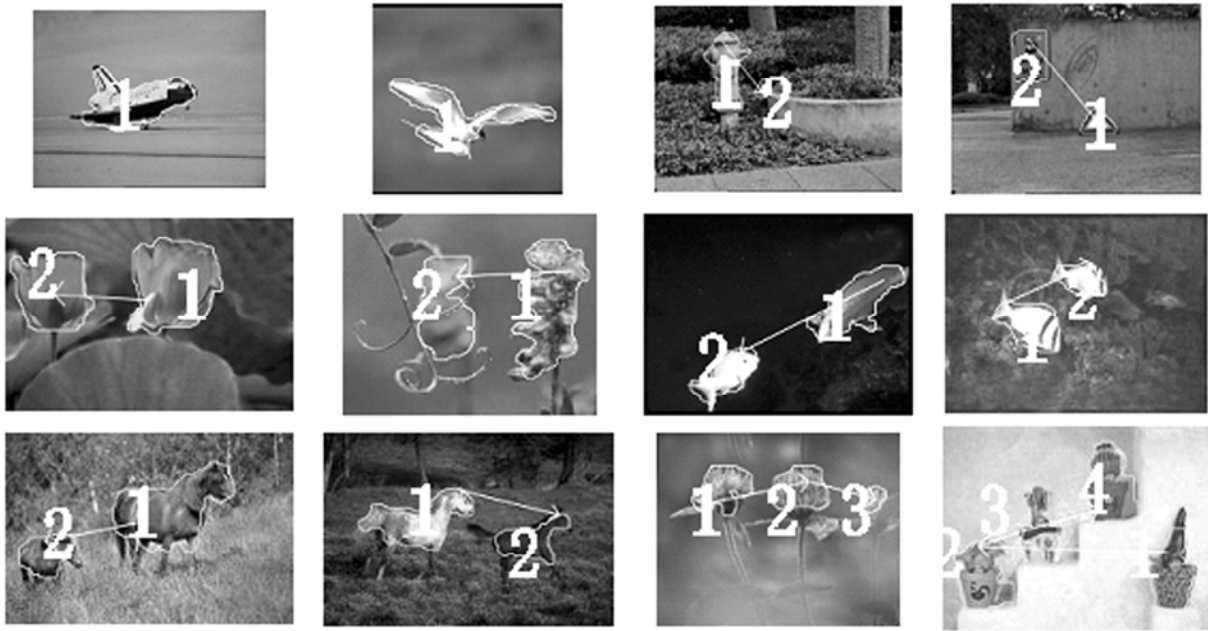


图 7 本文方法的感兴趣区检测结果

表 1 图像 ROI 兴趣度测试数据

图像	ROI	局部显著度	全局显著度	面积增强因子	位置增强因子	兴趣度
3	1	166.0000	231.9894	1.4010	0.9964	255.0000
	2	224.0000	255.0000	1.0989	0.9736	216.5097
4	1	255.0000	255.0000	1.1226	0.9910	255.0000
	2	173.0000	189.0775	1.3773	0.9696	232.9338
5	1	236.0000	255.0000	1.2878	0.9999	255.0000
	2	161.0000	224.5275	1.2121	0.9985	210.9123
6	1	255.0000	255.0000	1.2942	0.9981	255.0000
	2	168.0000	209.8995	1.2057	0.9971	195.8295
7	1	100.0000	255.0000	1.3163	0.9928	255.0000
	2	136.0000	255.0000	1.1836	0.9743	229.7931
8	1	76.0000	255.0000	1.3307	0.9947	255.0000
	2	56.0000	255.0000	1.1692	0.9820	222.6766
9	1	255.0000	255.0000	1.3955	0.9999	255.0000
	2	209.0000	235.4865	1.1044	0.9954	182.1674
10	1	88.0000	255.0000	1.3362	0.9992	255.0000
	2	70.0000	255.0000	1.1637	0.9899	218.5855
11	1	150.0000	243.5509	1.2111	0.9906	255.0000
	2	138.0000	238.6909	1.2287	0.9839	246.5941
	3	226.0000	255.0000	1.0600	0.9680	137.5546
12	1	214.0000	255.0000	1.1201	0.9914	255.0000
	2	190.0000	248.5520	1.1001	0.9763	243.8338
	3	192.0000	246.3838	1.1401	0.9975	238.9418
	4	94.0000	200.2834	1.1396	0.9889	203.1466

表 2 首次感兴趣区和注意转移顺序符合率对比(%)

方法	单个对象	多个对象		综合(多个对象 + 单个对象)	
		首次感兴趣对象	注意转移顺序	首次感兴趣对象	注意转移顺序
Itti 模型	83	56	41	66	56
本文方法	100	75	63	84	76

的注意内容, 会发生重叠转移, 而本文使用的“返回抑制性”能完全抑制已选择的注意内容。

### 7 结束语

根据生物视觉注意理论, 本文提出了一种基于视觉注意模型和进化规划的自动检测感兴趣区的方法。这种方法在处理小对象时获得了较为满意的区域检测和兴趣度量结果。这是因为本文依据的是人眼视觉注意通常集中到特殊对象上, 即出现频率小的对象, 而大对象在整幅图中出现的频率较高, 使其显著度下降, 因此需要进一步研究对大对象的处理。下一步可以考虑将大对象映射为小对象, 提取出对象区域后再还原回初始大小。此外, 还需进一步减小算法复杂度, 提高运算时间, 以及增强处理复杂图像的能力。

这是一种无需用户参与的、非监督的感兴趣区检测方法, 并且符合生物的视觉注意机制。利用视觉注意模型检测感兴趣区为进一步实现基于感兴趣区的图像检索奠定了基础。研究视觉注意模型对相关领域如图像压缩, 视频监控和人机交互等技术的

发展具有重要意义。

### 参考文献

- [1] 沈兰荪, 卓力. 小波编码与网络视频传输[M]. 北京: 科学出版社, 2005: 9-11.  
Shen L S and Zhuo L. Wavelet Coding and Network Video transmission [M]. Beijing: Science Press, 2005: 9-11.
- [2] Khanh V, Hua K A, and Tavanapong W. Image retrieval based on regions of interest [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2003, 15(4): 1045-1049.
- [3] Bulthoff H H, Lee S W, and Poggio T, *et al.* Biologically Motivated Computer Vision [M]. New York: Springer Publishing Company, 2003: 150-159.
- [4] Huang L and Pashler H. Working memory and the guidance of visual attention: Consonance-driven orienting [J]. *Psychonomic Bulletin & Review*, 2007, 14(1): 148-153.
- [5] 张菁, 沈兰荪, 等. 基于视觉感知的图像检索的研究[J]. 电子学报, 2008, 36(3): 494-499.  
Zhang J and Shen L S, *et al.* A Survey of image retrieval based on visual perception [J]. *Acta Electronica Sinica*, 2008, 36(3): 494-499.
- [6] Itti L and Koch C. Computational modeling of visual attention [J]. *Nature Reviews Neuroscience*, 2001, 2(3): 194-203.
- [7] Fogel L J. Toward inductive inference automata [C]. In Proceedings of the International Federation for Information Processing Congress. Munich, Germany, August 27-September 1, 1962: 395-399.
- [8] 周方俊, 王向军, 张民. 基于 t 分布变异的进化规划[J]. 电子学报, 2008, 36(4): 667-671.  
Zhou F J, Wang X J, and Zhang M. Evolutionary programming using mutations based on the t probability distribution [J]. *Acta Electronica Sinica*, 2008, 36(4): 667-671.
- [9] Stentiford F W M. Attention based similarity [J]. *Pattern Recognition*, 2007, 40(3): 771-783.
- [10] Lai C C and Chang C Y. A hierarchical evolutionary algorithm for automatic medical image segmentation [J]. *Expert Systems with Applications*, 2009, 36(1): 248-259.
- [11] Richards W, Seung H S, and Pickard G. Neural voting machines [J]. *Neural Networks*, 2006, 19(8): 1161-1167.
- [12] 张鹏, 王润生. 基于视点转移和视区追踪的图像显著区域检测[J]. 软件学报, 2004, 15(6): 891-898.  
Zhang P and Wang R S. Detecting salient regions based on location shift and extent trace [J]. *Journal of Software*, 2004, 15(6): 891-898.
- [13] He D J, Shao J M, and Gen N, *et al.* A model for image categorisation based on a biological visual mechanism [J]. *New Zealand Journal of Agricultural Research*, 2007, 50(5): 781-787.
- 张菁: 女, 1975年生, 讲师, 博士生, 研究方向为多媒体信息检索.
- 沈兰荪: 男, 1938年生, 教授, 博士生导师, 主要研究领域为图像/视频编码、传输、处理和中医舌象采集与分析技术等.
- 高静静: 女, 1984年生, 硕士生, 研究方向为智能化信息处理.