

基于改进视觉注意模型的显著区域检测

宋 侃

(安徽大学 计算机科学与技术学院, 安徽 合肥 230601)

摘 要:文中在专家学者提出的视觉注意模型的基础上,提出一种新的视觉注意模型改进算法。首先,采用改进后的 Itti 模型来计算图像视觉显著性,从底层特征,包括亮度、方向、颜色,然后建立小波金字塔,利用中央周围算子作差运算,得到颜色、亮度和方向每个特征的显著图,然后通过归一化算子,融合成 Itti 模型下的显著图;接着,借鉴了 MA 和 Zhang 等提出的显著图模型算法的一些方法,从全局的角度计算显著图;最后将以上得到的显著图进行归一化融合。

关键词:视觉注意模型;Itti 模型;显著图;多尺度变换

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2015)07-0234-04

doi:10.3969/j.issn.1673-629X.2015.07.053

Region of Interest Detection Based on Improved Visual Attention Model

SONG Kan

(College of Computer Science and Technology, Anhui University, Hefei 230601, China)

Abstract: In this paper, propose a new visual attention model improvement algorithm based on the visual attention model proposed by experts. First, the improved Itti model is used to compute the visual saliency of image, from low-level features including color, brightness and direction, then construct a wavelet Pyramid, using the central peripheral operator for difference operation to attain the concern graph of each feature of color, brightness and direction, and then through the normalized operator, merging into the saliency map of Itti model. Next, drawing on the experience of algorithm proposed by MA and Zhang etc, calculate the saliency map from a global perspective. Finally the saliency map above are normalized fusion.

Key words: visual attention model; Itti model; saliency; multi-scale transform

0 引 言

在图像中最能吸引观察者的兴趣,也最能表达图像内容的区域称之为感兴趣区域。观察者观察图像时,自觉地使用了自底向上和自上而下两种不同形式方法的融合。自底向上,这是一种原始快速的机制,是人眼不带有主观目的的,仅仅从图像的亮度颜色等特征去观察;自上而下,是一种意识指导认知机制,是人从先验知识的角度看待图像。在图像分析过程中,基于图像视觉注意的显著区域检测有着非常重要的作用。注意在人类信息加工过程中处于重要的调节环节,它可以对有限资源进行分配,使感知具有一定的选择能力。如果可以将这些机制引入到图像分析领域,将仅有的计算资源率先分配给那些容易吸引观察者注

意的区域,这样势必会极大提高图像处理分析等方法的效率。正是在这个基础上才有了显著性区域检测的提出和发展。

1 ROI 检测方法概述

正如人们观察图像一样,对图像的显著性区域检测一般分为两大类:自下而上基于数据驱动的显著性区域突现和自上而下任务驱动的目标突现。其中自下而上的显著性检测方法研究方兴未艾,代表性算法是文献[1]中提出的模拟生物视觉注意机制的选择性注意算法(Itti 算法);还有没有基于任何生物视觉原理的纯数学计算方法,如 Achanta 等^[2]提出的全分辨率算法(AC 算法)和 Hou 等^[3]提出的基于空间频域分

收稿日期:2014-08-18

修回日期:2014-11-24

网络出版时间:2015-06-23

基金项目:安徽省科技攻关计划科技强警专项资金资助项目(1301b042020);安徽大学青年骨干教师培养对象经费资助项目;高等学校博士学科点专项科研基金联合资助课题(20133401110009)

作者简介:宋侃(1986-),男,硕士研究生,研究方向为图像处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150623.1046.037.html>

析的剩余谱算法(Spectral Residual, SR)。

1998年Itti等在Koch和Ullman的工作基础上提出了模拟生物体视觉注意机制的选择性注意模型。Itti模型的方法简单方便,运算速度快,它是目前影响范围最广的可计算模型。它利用神经生理学的感受野、侧抑制神经网络等研究成果,模拟了人的视觉注意力机制。认为像素与背景的对比是吸引人注意的重要因素,并把这种对比定义为该点的显著值^[4]。它的具体方法是:

(1)先把输入的图像通过线性滤波,分成颜色、亮度、方向等多种特征图,然后利用高斯滤波和向下采样对输入的图像建立9层的Gauss金字塔,分别形成亮度、方向和颜色的金字塔。

(2)然后使用中央周边算子对每个独立的金字塔进行处理,形成各个特征的显著图。

(3)最后将形成的特征显著图归一化,线性融合为一幅显著图。

2 改进后的算法

Ma等^[5]提出了用局部区域对比的方法来提取显著图,这个方法是将图像划分为若干像素块,通过计算像素块之间的特征距离来得到图像的显著图。文中在此基础上借鉴了Itti多尺度分析方法的思想,也提出了一种多尺度分析方法。这种算法所定义的显著值是输入图像的一个局域区域 R_1 的平均特征向量和另外一个区域 R_2 平均特征向量的差值。一般情况下, R_2 的区域包括了 R_1 的区域。因此适当缩放 R_2 区域,即可得到不同尺度下的特征向量差值。文中提出的模型创新点主要体现在:提出了单个像素显著度的概念。Itti模型对图像处理的显著度是局部的显著图,记为 S_1 ,Ma和Zhang算法对图像处理的显著度看成是一个全局的显著度,记为 S_2 ,因此处理后像素的总显著度 $S = S_1 + S_2$ 。

2.1 改进Itti模型

从上文可知,Itti模型较好地模拟了人眼视觉系统的视觉感受野机制,原模型使用了高斯金字塔方法,但是笔者认为小波变换可能更加符合人眼的视觉特性,这是因为小波在空间方向和相应频带选择上的敏感性不同,导致了原图在这两个方面进行了解。为此,用小波替换原模型中的高斯金字塔。具体改动后的方法如下所示:

(1)视觉特征:设 (I, R, G, B) 分别对应输入图像的亮度、红、绿、蓝特征。其中第0层是输入图像,1到8层分别是小波对图像进行变换和向下采样所得,第 n 层的图像为 $n-1$ 层图像的 $1/2$ 。然后对金字塔的每一层分别提取亮度、红色、绿色、蓝色、黄色、方向,组成

不同的金字塔。方向特征为Gabor小波在 0° 、 45° 、 90° 、 135° 四个方向的分量。亮度和各种颜色的特征由如下公式得到:

$$I = \frac{(r + g + b)}{3} \quad (1)$$

$$R = r - \frac{(g + b)}{2} \quad (2)$$

$$G = g - \frac{(r + b)}{2} \quad (3)$$

$$B = b - \frac{(r + g)}{2} \quad (4)$$

$$Y = \frac{(r + g)}{2} - \frac{(r - g)}{2} - b \quad (5)$$

(2)在需要建立类似于人类视觉系统的模型中,Gabor小波表现出了良好的特性,这是由它本身的特点决定的。具体的表现在提取频率域信息和局部空间等方面。在下面实验里,笔者采用小波分解低通滤波器系数是 $\{0.602\ 949, 0.266\ 864\ 11, -0.078\ 223\ 3, -0.016\ 864\ 118, 0.026\ 748\ 76\}$ ^[6];高通滤波器系数是 $\{1.115\ 087\ 04, -0.591\ 271\ 76, -0.057\ 543\ 53, 0.091\ 271\ 762, 0.0\}$ ^[6]。

(3)中央周围运算的具体方式是:将金字塔的2,3,4层作为中心 c ,分别跟 $c+3$, $c+4$ 层的对应像素点相减,而由于不同的特征图分辨率不同,需要通过插值的方式使得图像大小相同。

(4)最后进行归一化,将所有特征映射图归一化到规定的区间 $[0 \cdots M]$,这样可以消除各个特征通道之前的幅度存在的差异;找到每个特征映射图的全局最大值 M 及其位置,并计算其他所有局部最大值的均值 m ;每个特征映射图全局乘以 $(M - m)^2$ 。这样可以保证具有高对比度的、极值点且突出的特征映射图具有较大的权重。归一化之后,就可以将方向、亮度、颜色3个特征图进行线性相加:

$$S_1 = \frac{[N(I) + N(C) + N(O)]}{3} \quad (6)$$

2.2 另一种算法改进后的具体实现

具体实现方法如图1所示^[7]。区域 A 为一个固定大小的像素。区域 B 为包含区域 A 的一个可变大小的像素区域。通过自左向右、自上向下遍历整个图像,计算 A 和 B 之间的特征距离,并把特征距离可视化地转为灰度特征图。由于 B 是可变的,在 B 的某一个特定的尺度下,将 A 区和 B 区之间的特征距离作为显著图,就可得到若干幅不同尺度下的灰度显著图。对这几幅不同的显著图,可以在空间上全部映射到一幅图像上,映射后生成的图的像素点的灰度值等于原若干幅图像的平均值。

区域 A 的显著图在区域 B 某个固定尺度下的计算方式:

$$C_{i,j} = D \left[\left(\frac{1}{N_1} \sum_{p=1}^{N_1} V_p \right) \left(\frac{1}{N_2} \sum_{q=1}^{N_2} V_q \right) \right] \quad (7)$$

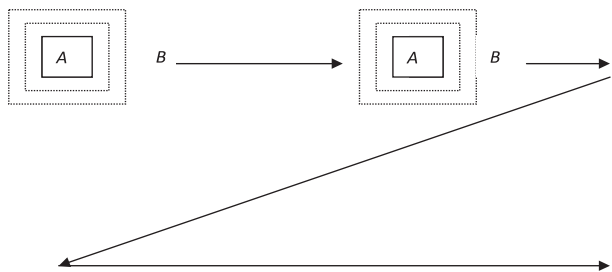


图 1 改进后算法的实现

其中, N 代表的是区域中的像素个数; V 为其中各具体像素的特征向量; D 为欧氏距离。由于在人眼的视觉系统中^[8-9], 人对颜色差距的判定略等于欧氏距离, 在基于 Lab 空间的情况下提取颜色特征时, 将 D 近似表示为:

$$D = \| V_1 - V_2 \| \quad (8)$$

其中, $V_1 = [L_1, a_1, b_1]^T$ 和 $V_2 = [L_2, a_2, b_2]^T$ 分别表示区域内的平均特征向量。对于宽为 x , 高为 y 像素的图像, 区域 B 的大小变动阈值 w_B 为:

$$\frac{x}{8} \ll w_B \ll \frac{x}{2} \quad (9)$$

这里假定 x 小于 y , 否则式(3)中的 x 更换为 y 。若 B 的变化范围过大时, 会将许多非显著区域判为显著区域; 而若 B 的变化范围又过小时, 则只能得到物体的边缘部分。之所以通过对 B 进行缩放变动, 而不是对原图像进行, 是因为这样可以方便地得到跟原图大小一样的显著图。

文中, w_B 的变化范围为 3 个尺度, 这样在不同尺度大小下显著度的和就是最终的显著度图, 其中 $m_{i,j}$ 为图像中各对应点的显著值, 如式(10):

$$m_{i,j} = \sum_s C_{i,j} \quad (10)$$

可以按照下面的公式归一化像素点的全局显著度。

$$S_2(x, y) = \frac{S'_2(x, y) - S'_{2_{\min}}}{S'_{2_{\max}} - S'_{2_{\min}}} \times 255 \quad (11)$$

其中, $S'_{2_{\min}}$ 为 S'_2 图中的最小值; $S'_{2_{\max}}$ 为图中的最大值。令局部显著图记为 S_1 , 全局显著图记为 S_2 , 对二者进行合并后的最终显著图记为 S , 计算公式如下:

$$S = \frac{1}{2} [N(S_1) + N(S_2)] \quad (12)$$

3 显著区的生成

对于最后形成的灰度显著图, 可以选取一个阈值来判定, 即图像的灰度值大于阈值 T 时, 则该范围区域的兴趣度则记为“1”, 反之, 记为“0”。公式如下:

$$g(x, y) = \begin{cases} 1, & S_g(x, y) > T \\ 0, & S_g(x, y) \ll T \end{cases} \quad (13)$$

当前从显著图中提取显著区域的方法大多数是用区域增长或者 K -means 等^[10-11], 这些方法几乎都将注意焦点当作 K -means 聚类 and 区域增长的起始来逐步形成区域, 但因为上节方法形成的显著图的界限非常难以界定, 所以在如何准确选取起始点上将带来难题。此外, 由于不同图像的显著区数量不尽相同, 种子的具体个数也很难把握。

为了解决这个问题, 文中将注意焦点和候选区联系起来解决显著区域的问题, 即只有当注意焦点落在了这篇区域中, 才能把当前区域判定为显著区, 若没有则移动到下一个注意焦点。然后循环直到结束。由此可知, 每一个注意焦点均会得到与其相对应的显著区。

4 实验结果与分析

为了验证上述算法的有效性, 笔者用个人家用计算机对含有多个不同对象的 200 幅图片 (包括植物、鸟、马、狮子等几类) 进行了实验, 如图 2 所示。从实验结果来看, 文中的检测方法对感兴趣区计算的结果更加接近对象边缘, 检测的时间大概需要 1 s。



图 2 显著区域提取的实验结果

为了验证所给出的方法在第一次 ROI 和之后的注意焦点的转移顺序和实验者的主观观测结果的相似程度, 文中对 5 个用户进行了测试, 要求他们分别给出对观测图像的 ROI 及观察顺序, 然后逐一统计了文中以及 Itti 方法模型在观测第一次 ROI 及观察顺序的相似度, 结果如表 1 所示。

5 结束语

根据生物视觉的相关理论^[12-13], 文中给出了一种

表 1 注意转移顺序和首次感兴趣区和的对比					
方法	单个对象	多个对象		综合对象	
		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	83	56	41	66	56
文中方法	100	75	63	84	76

新的基于改进后的两种不同视觉注意模型合并的非人为控制检测感兴趣区的方法。此方法在处理对象较少的图片时获得了令人满意的度量结果。出现这种结果的原因是由于人类视觉模型比较轻易地会关注于特殊对象上,即出现频率略小的对象,而对那些在区域内逐渐大量出现的对象,其显著度则会逐步下降,因此需要进一步研究大量出现对象的情况。但也有一些明显的不足:比如 ROI 将被作为非显著区域被去掉当它与周围区域的对比不是非常的明显时。这是由于在判断过程中阈值的选取是由先验知识所得,不是根据算法来自适应选取的;若图像中 ROI 比较大时,提取出来以后的区域中间会有一些空洞点。这是由于变动范围 B 只给出了三个不同尺度来计算区域 A 的显著度,若再增加 B 的次数,将会显著减少空洞点的产生,但是这样又会增加计算的时间复杂度。因此,还需更进一步减少时间复杂度,增强计算效率,以及提高处理复杂图像的能力。

参考文献:

[1] Itti L,Koch C. Computational modeling of visual attention[J].

Nature Reviews Neuroscience,2001,2(3):194-203.

[2] Achanta R,Hemami S,Estrada F,et al. Frequency-tuned salient region detection[C]//Proc of IEEE international conference on computer vision and pattern recognition. [s.l.]: IEEE,2009:1597-1604.

[3] Hou Xiaodi,Zhang Liqing. Saliency detection;a spectral residual approach[C]//Proc of IEEE conference on computer vision and pattern recognition. Minneapolis: IEEE, 2007: 11-19.

[4] 张 菁,沈兰荪,David Dagan Feng. 基于视觉感知的图像检索的研究[J]. 电子学报,2008,36(3):494-499.

[5] Ma Y F,Zhang H J. Contrast-based image attention analysis by using fuzzy growing[C]//Proceedings of the eleventh ACM international conference on multimedia. [s.l.]: ACM,2003.

[6] 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.

[7] 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.

[8] 罗四维. 视觉感知系统信息处理理论[M]. 北京:电子工业出版社,2006.

[9] 耿永政,陈 坚. 结合图论的 JSEG 彩色图像分割算法[J]. 计算机技术与发展,2014,24(5):15-19.

[10] 夏召强,冯晓毅,彭进业. 基于边缘与深度特征的感兴趣区域检测技术[J]. 计算机仿真,2009,26(7):248-251.

[11] 孔祥宇. 基于显著区域的图像检索方法研究[D]. 北京:北京交通大学,2008.

[12] 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.

[13] 李树艳. 基于人类视觉系统的选择性注意模型研究[D]. 成都:电子科技大学,2013.

2015 中国计算机大会 (CNCC2015)

China National Computer Congress

主题:互联网催生新经济

2015 中国计算机大会 (CNCC2015) 将于 2015 年 10 月 22~24 日在合肥安徽世纪金源大饭店举行。会议由中国计算机学会 (CCF) 主办,中国科学技术大学和合肥市人民政府承办。大会主题为“互联网催生新经济”,探讨互联网+新经济中面临的技术挑战和问题。会议邀请到包括 2014 年图灵奖获得者 Michael Stonebraker 教授在内的 10 余名国内外计算领域顶级知名专家作大会特邀报告,专题论坛 20 余场,讨论类脑计算、智能语音、互联网新经济学、自主可控基础软硬件等前沿热点问题。大会还为知名 IT 企业和创业公司搭建了自由开放的互动平台,组织了丰富多彩活动,满足与会者和企业的不同需求。

诚挚邀请您参加 2015 中国计算机大会!

报名方式:即日起至 2015 年 10 月 15 日,登陆 CNCC2015 官网报名:<http://cncc.ccf.org.cn/>。

注册链接:http://cnccreg.ccf.org.cn/user_cn/user_login.asp?hid=。报满为止。

联系人:曾 菲 E-mail: cncc_pr@ccf.org.cn 电 话:010-6260 0336

基于改进视觉注意模型的显著区域检测

作者：[宋侃](#)，[SONG Kan](#)

作者单位：[安徽大学 计算机科学与技术学院, 安徽 合肥, 230601](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：2015(7)

引用本文格式：[宋侃](#), [SONG Kan](#) [基于改进视觉注意模型的显著区域检测](#)[期刊论文]-[计算机技术与发展](#) 2015(7)