

基于层次和动态阈值的图像显著区域检测方法

葛 涛,冯松鹤

(北京交通大学 计算机与信息技术学院,北京 100044)

(urfrdgt@163.com)

摘 要:发展完善了基于层次的显著区域检测模型。把原始图像分层,对每层分别提取底层特征,计算显著性特征,包括亮度显著性、饱和度显著性和色彩显著性。考虑到图像的概率特性,通过引入二维正态分布函数,准确合理地表示图像像素间的相关性。把各层计算所得的结果叠加,通过引入动态阈值,得到最后的显著区域。动态阈值方法不受原始图像的影响,有较好的适应性。将此方法应用于真实图像,得到了令人满意的检测结果。

关键词:显著区域;显著性特征;视觉注意;动态阈值

中图分类号: TP391.41 **文献标识码:** A

A method for image salient regions detection based on layers and dynamic threshold

GE Tao, FENG Song-he

(College of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

Abstract: The model for detecting salient regions based on layers was developed and consummated. Primary images were divided into layers. And then for every layer, basic characteristics were extracted, and saliency characteristics including intensity saliency, saturation saliency and color saliency were calculated. Because of the possibility characteristics of images, the planar normal distribution function was adopted to present the pertinence of pixels accurately and reasonably. The results of layers were added and the dynamic threshold was used to reach the final saliency regions. Dynamic threshold method can be influenced hardly by the primary images, so it can adapt to different situations. When the method was applied to real images, satisfying results were got.

Key words: salient region; saliency characteristic; visual attention; dynamic threshold

显著区域检测^[1]主要是通过图像固有的影响人类视觉的信息,根据有关心理学的规律,检测出图像的最明显的区域,也就是对人类视觉刺激最强烈的区域。目前已有的比较有效的方法有基于对比的模糊增长方法^[2]、基于 SVM 的显著区域提取方法^[3]等。在本文中,主要使用几种颜色显著性特征实现对显著区域的检测。

1 检测模型

如图 1 所示,首先把原图像进行分层,分层的方法多种多样,我们这里采用了隔点采样的方法,对原始图像每隔一个点采样一次,得到长、宽都为原始图像二分之一的图像。第一层的图像为原始图像,第二层的图像是第一层图像面积的四分之一,而第三层图像为第二层图像面积的四分之一。

针对每一层,提取各种底层特征,然后根据底层特征分别计算亮度显著性特征、饱和度显著性特征和彩色显著性特征。由于人的视觉对对比强烈的区域较为敏感,因此这三种特征实际上包含了某两个像素点之间的差值,差值越大越显著。

由于二维正态分布函数有很好的概率特性,并且与实际情况非常吻合,为了综合考虑各个像素点与周围其他像素点的关联情况,我们引入二维正态分布函数。采取二维正态分布函数对各像素点的显著性值加权平均,得到的加权平均值

最能反映实际情况。进一步,把每层得到的加权平均值进行叠加:第三层的结果经放大后与第二层结果叠加,然后再经放大后与第一层结果叠加。这是因为:每一层计算出来的加权平均值,实际上只代表了显著区域的轮廓,而我们想要得到的是整个显著区域,第三层的结果正好可以“填充”到第二层的“轮廓”中去,第二层的结果正好可以“填充”到第一层的“轮廓”中去,依此类推,反复叠加,最终获得显著区域的形状。

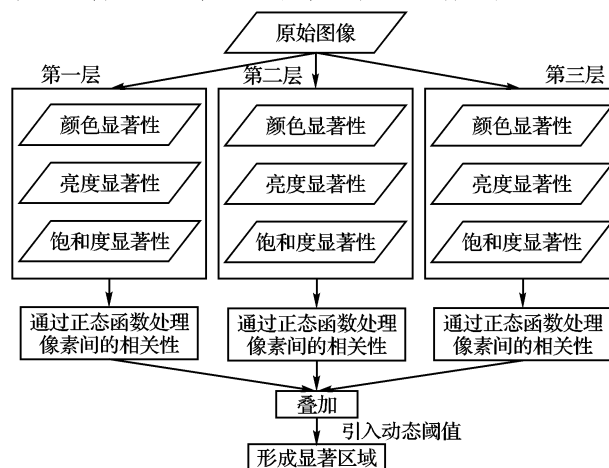


图 1 显著区域检测模型

收稿日期:2006-05-25;修订日期:2006-06-30

作者简介:葛涛(1981-),男,北京人,硕士研究生,主要研究方向:视频数据库、图像处理;冯松鹤(1981-),男,江苏盐城人,博士研究生,主要研究方向:视频数据库。

如果想进一步获得明显的显著区域(显著区域用白色表示,其余区域用黑色表示),就需要根据上一步叠加的结果计算出一个动态阈值,大于此阈值的像素点用白色表示,其余像素点用黑色表示,这样,就得到了最终的显著区域。

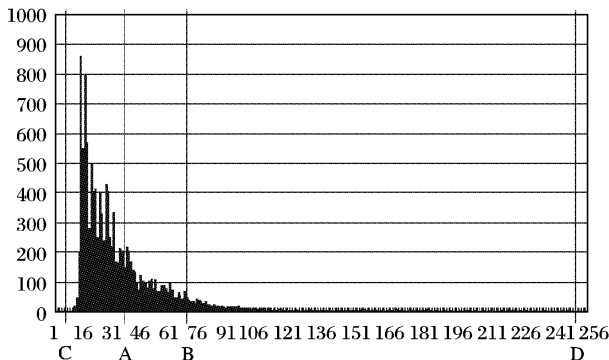


图2 叠加后图像的灰度直方图

2 显著性计算

要想获得更高层次的显著性特征,就需要首先从原始图像中提取出底层特征。这些底层特征是图像直接反映出来的,包括:亮度特征、饱和度特征和四个颜色通道; R (red), G (green), B (blue)和 Y (yellow)。通过这些底层特征,计算出高一层的显著性特征:亮度显著性特征、饱和度显著性特征以及彩色显著性特征。这些特征能够较好地反映出像素间的对比关系,是显著区域检测的关键。文献[4]中还提到方向等其他一些特征。但是在实际检测过程中,这些特征并没有明显提升检测效果,反而增加了计算强度,增加了检测时间。因此,我们只选用了以上几种显著性特征进行显著性检测。为了解决像素间的相关性表示的问题,我们引入了二维正态分布函数。各种显著性特征值与二维正态分布函数结合,完成图像显著性的计算。

2.1 基本底层特征

2.1.1 亮度特征

亮度与物体的反射率成正比,对于彩色来说,颜色中掺入白色越多就越明亮,掺入黑色越多则亮度越小。这说明,光照越强,亮度值就越大。而光照越强的区域,对视觉的刺激就越明显,因此,把亮度特征作为显著性检测的底层特征。亮度特征的计算方法^[5]为: $I(P_{ij}) = [r(P_{ij}) + g(P_{ij}) + b(P_{ij})]/3$ 。其中, P_{ij} 为坐标为 (i, j) 的像素点; $r(P_{ij}), g(P_{ij}), b(P_{ij})$ 为像素点 P_{ij} 的 r, g, b 三个分量值; $I(P_{ij})$ 为像素点 P_{ij} 的亮度值。

2.1.2 饱和度特征

饱和度是指颜色的纯洁性。可见光谱的各种单色光是最饱和的彩色。当光谱色加入白光成分时,就变得不饱和。物体的饱和度取决于该物体表面选择性反射光谱辐射能力。物体对光谱某一较窄波段的反射率高,而对其他波长的反射率很低或没有反射,则表明它有很高的选择性反射的能力,这一颜色的饱和度就高。这说明,饱和度特征有助于检测图像的显著区域。其计算公式^[6]为: $S(P_{ij}) = \max(r(P_{ij}), g(P_{ij}), b(P_{ij})) - \min(r(P_{ij}), g(P_{ij}), b(P_{ij}))$ 。其中, $S(P_{ij})$ 为像素点 P_{ij} 的饱和度值。

2.1.3 彩色通道

为了方便获得更高层次的显著性特征,我们采用了下面四个颜色通道^[4](负数设为0),它们是高层次显著性特征的

基础。

$$R(P_{ij}) = r(P_{ij}) - [g(P_{ij}) - b(P_{ij})]/2;$$

$$G(P_{ij}) = g(P_{ij}) - [r(P_{ij}) + b(P_{ij})]/2;$$

$$B(P_{ij}) = b(P_{ij}) - [r(P_{ij}) + g(P_{ij})]/2;$$

$$Y(P_{ij}) = [r(P_{ij}) + g(P_{ij})]/2 -$$

$$|r(P_{ij}) - g(P_{ij})|/2 - b(P_{ij})$$

其中, $R(P_{ij}), G(P_{ij}), B(P_{ij}), Y(P_{ij})$ 分别为颜色的四个通道: Red, Green, Blue 和 Yellow。

2.2 显著性特征

本文综合了亮度显著性、饱和度显著性和色彩显著性,它们都影响显著区域的检测。

2.2.1 亮度显著性和饱和度显著性

本文采用基于对比的方法,像素 x 与像素 y 之间的亮度差和饱和度差越大,越有利于对显著区域的区分。定义^[4]为: $\Delta I(x, y) = |I(x) - I(y)|$; $\Delta S(x, y) = |S(x) - S(y)|$ 。其中, $I(x)$ 和 $I(y)$ 分别表示像素 x 和 y 的亮度值, $S(x)$ 和 $S(y)$ 分别表示像素 x 和 y 的饱和度值。

2.2.2 色彩显著性^[7]

通过前面提取的底层特征,可以得到: $RG(x, y) = |(R_x - G_x) - (R_y - G_y)|/2$; $BY(x, y) = |(B_x - Y_x) - (B_y - Y_y)|/2$; $\Delta C(x, y) = \sqrt{\eta_{RG}^2 RG^2(x, y) + \eta_{BY}^2 BY^2(x, y)}$ 。其中,

$$\eta_{RG} = \frac{R_x + R_y + G_x + G_y}{R_x + R_y + G_x + G_y + B_x + B_y + Y_x + Y_y}, \quad \eta_{BY} = \frac{B_x + B_y + Y_x + Y_y}{R_x + R_y + G_x + G_y + B_x + B_y + Y_x + Y_y},$$

$$\frac{2\sqrt{B_x^2 + B_y^2 + Y_x^2 + Y_y^2}}{3 \times 255}, x, y \text{ 表示两个像素点。} \Delta C(x, y) \text{ 为色彩显著性值。}$$

2.3 引入二维正态分布函数,综合考虑各显著性特征

综合考虑亮度显著性、色彩显著性和饱和度显著性,我们有如下公式:

$$S_{cls}(x, y) = \sqrt{\alpha \Delta I(x, y)^2 + \beta \Delta C(x, y)^2 + \chi \Delta S(x, y)^2}$$

在实际中,我们把 α, β, χ 都设为1。

由概率论可知,二维正态随机变量 (X, Y) 的概率密度为:

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

把上面的 x, y 对应于图像的横纵坐标,这样就可以把图像上各个像素的关系和正态分布的概率密度函数联系起来了。通过设置上面公式中的 $\rho, \mu_1, \mu_2, \sigma_1, \sigma_2$, 可以得到非常切合实际的对应关系。在很多情况下, x 与 y 的相关性不大,因此,设 $\rho = 0$ 。进一步,可以得到图像上任何两个像素点的关系:

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}\left[\frac{(x_1-x_2)^2}{\sigma_1^2} + \frac{(y_1-y_2)^2}{\sigma_2^2}\right]\right\}$$

其中, $(x_1, y_1), (x_2, y_2)$ 分别为图像中的两个像素点。 $f(x, y)$ 表示两点间的相关性。

如果两个像素点距离很近,则这两个点的相关性就越强, $f(x, y)$ 的值就越大;反之,相关性就越小, $f(x, y)$ 的值也就越小。我们利用正态分布函数对显著性特征加权,就得到了各个像素点最终的显著性值:

$$S_{CIS}(x) = \frac{\sum_{i=1}^{n \times m-1} S_{CIS}(x, y_i) \cdot f(x, y_i)}{\sum_{i=1}^{n \times m-1} f(x, y_i)}$$

其中, m, n 分别为每层图像的长度和宽度。

3 基于阈值的显著区域

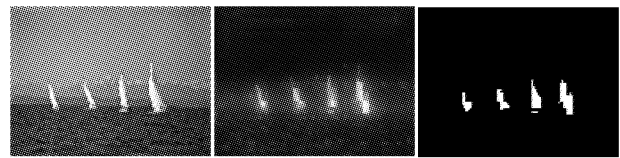
通过对每层计算显著性,可以得到显著区域的轮廓。要想得到清晰准确的显著区域,需要把各个层的结果叠加,使得各层的轮廓互相填充。如果想进一步获得明显的显著区域(显著区域用白色表示,其余区域用黑色表示),就需要根据上一步叠加的结果计算出一个阈值。这里,我们提出了一个设置动态阈值的方法。

图 2 所示的是各层叠加后的灰度图像(如图 3(b)所示)的直方图。其中, A 为直方图的数学期望, B 为数学期望与标准差的和, C 为数学期望与标准差的差, D 为 256 个直方图数据中最接近第 256 个分量的一个不为零的点。根据概率论和数理统计可知,直方图中大部分数据都集中在 C 和 B 之间。根据这一特点,我们可以把阈值设置为 B 和 D 之间的某一个值,一般设为 B 和 D 的中点或者直接把 B 点设为阈值。这样,无论原始图像如何,都可以找到一个适合它的阈值。这个阈值可以有效地消除叠加后图像中的噪声点,达到检测显著区域的目的。

4 实验结果

对于原始图像(如图 3(a)所示),把原始图像分为 3 层,对其中的每一层,应用上述方法计算其显著性,然后把各层的显著性值叠加,结果如图 3(b)所示。对于叠加后的值,按照上述方法计算阈值。大于此阈值的像素点用白色表示,其余

像素点用黑色表示,形成最终的显著区域(如图 3(c)所示)。可以看出,采用正态分布函数,很好地解决了像素间的关联问题,并产生了预期的结果。由于对每一层计算所得的值只反映了物体的轮廓,而叠加后的图像恰恰弥补了每一层的不足。此外,由于采用了动态阈值的方法,很好地消除了叠加后的图像(如图 3(b)所示)中的噪音点。



(a) 原始图像 (b) 各层显著性值叠加后的图像 (c) 显著区域
图 3 实验图像

参考文献:

- [1] 沈庭之, 方子文. 数字图像处理及模式识别[M]. 北京: 北京理工大学出版社, 1998.
- [2] MA Y-F, ZHANG H-J. Contrast-based Image Attention Analysis by Using Fuzzy Growing[A]. Proceedings of the 11th ACM International Conference on Multimedia[C], 2003. 374-381.
- [3] KO B, KWAK SY, BYUN H. SVM-based Salient Region(s) Extraction Method for Image Retrieval[A]. Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004[C], 2004. 977-980.
- [4] LTTI 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.
- [5] 章毓晋. 图像处理和基础[M]. 北京: 高等教育出版社, 2002.
- [6] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003.
- [7] SUN Y. Hierarchical Object-Based Visual Attention for Machine Vision[J]. Artificial Intelligence, 2003, 146(1): 77-123.

(上接第 2697 页)

仿真实验中,指定的平移周期数为 16,在特定的信号采样频率条件下,均进行 50 次实验,然后分别计算两种算法的平均时延实际值和平均所耗时间。表 1 是不同采样频率条件下的仿真结果。

表 1 不同采样频率条件下的仿真结果

采样频率 /KHz	时延值/ms			算法耗时/ms	
	理论	改进前	改进后	改进前	改进后
1	15.6250	6.6437	13.8049	1.131	0.362
2	7.8125	4.4563	6.6125	1.141	0.371
4	3.9062	2.8049	3.7975	1.184	0.423
8	1.9531	1.4234	1.8523	1.521	0.635
16	0.9766	0.6542	0.7542	2.036	0.752
32	0.4883	0.2546	0.3542	2.852	0.954
64	0.2441	0.0825	0.1923	3.215	1.235

从仿真结果可以看出,在同信噪比不同采样频率的情况下,由于改进算法能够有效地消除回响的干扰,所以其时延估计精度相对于改进前算法的时延估计精度,平均提高了 20%~30%。另外,从表 1 可知改进算法所耗时间平均仅为改进前算法的 30% 左右,计算复杂度大大降低,便于在工程中实时应用。

4 结语

本文针对基于最大似然权函数的广义互相关时延估计算法的计算量大以及无法消除回响干扰的不足,提出了一种改

进的时延估计算法,并利用实测噪声数据进行了仿真实验。仿真结果表明,改进算法能够有效消除噪声和回响的干扰,同时具有计算复杂度低、对硬件要求不高、实时性好等优点,具有潜在的应用前景。

参考文献:

- [1] SROUR N, ROBERTSON J. ARL-TR-706, Remote Netted Acoustic Detection System: Final Report[R]. US Army Research Laboratory, 1995.
- [2] BRANDSTEIN M. Microphone Arrays[M]. New York: Springer-Verlag, 2001.
- [3] RUI Y, FLORENCIO D. Time delay estimation in the presence of correlated noise and reverberation[J]. IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004, 2(2): 133-136.
- [4] KNAPP CH, CARTER GC. The generalized correlation method for estimation of time delay[J]. IEEE Transaction on Acoustics, Speech and Signal Processing, 1976, 24(4): 320-327.
- [5] CARTER GC. Coherence and time delay estimation[J]. Proceedings of the IEEE, 1987, 75(2): 236-255.
- [6] 陈华伟, 赵俊渭, 郭业才. 二次加权频域自适应时延估计算法与应用[J]. 声学学报, 2003, 28(1): 61-65.
- [7] 邱天爽, 王宏禹. 加窗 LMS-ML 自适应时间延迟估计[J]. 大连理工大学学报, 1994, 34(4): 470-476.
- [8] GU Y, TANG K, CUI H. Convergence analysis of a deficient-length LMS filter and optimal-length sequence to model exponential decay impulse response[J]. Signal Processing Letters, 2003, 10(1): 4-7.