

自组织映射与流形学习的图像显著度检测^{*}

陈加忠¹, 曹华¹, 苏曙光¹, 伊斯刚²

¹(华中科技大学 计算机科学与技术学院, 湖北 武汉 430074)

²(重庆师范大学 物理与电子工程学院, 重庆 408331)

通讯作者: 陈加忠, E-mail: chenjz70@163.com

摘要: 提取反映图像内容的结点以及为这些结点分配初始标签, 是半监督学习用于显著度检测的关键问题. 通过自组织映射把图像分成多个结点, 这些结点不但反映图像内容的颜色特征, 还能够反映图像内容的轮廓特征. 然后通过把二维结点图嵌入到高维的空间构造带权无向图. 由于无向边的对称性, 进一步采用流形学习的方法, 把无向图和半监督学习结合起来, 通过预设边界结点预期的显著度, 最终计算出所有结点的显著度. 实验结果表明, 与近年提出的几种经典的显著度检测算法相比, 所提出的方法取得了较好的 Precision-Recall 性能和较舒服的视觉效果.

关键词: 显著度检测; 自组织映射; 流形学习; 半监督学习; 无向图

中文引用格式: 陈加忠, 曹华, 苏曙光, 伊斯刚. 自组织映射与流形学习的图像显著度检测. 软件学报, 2015, 26(Suppl. (2)): 137-144. <http://www.jos.org.cn/1000-9825/15024.htm>

英文引用格式: Chen JZ, Cao H, Su SG, Yi SG. Saliency detection using self organizing map and manifold learning. Ruan Jian Xue Bao/Journal of Software, 2015, 26(Suppl. (2)): 137-144 (in Chinese). <http://www.jos.org.cn/1000-9825/15024.htm>

Saliency Detection Using Self Organizing Map and Manifold Learning

CHEN Jia-Zhong¹, CAO Hua¹, SU Shu-Guang¹, YI Si-Gang²

¹(College of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

²(College of Physics and Electronic Engineering, Chongqing Normal University, Chongqing 408331, China)

Abstract: Extracting nodes that reflect image content and assigning initial labels for these nodes are two critical technologies for saliency detection. A novel method of saliency detection is proposed by this work. It consists of two main parts, one is self organizing map (SOM), and the other is manifold learning (ML). Hundreds of nodes are obtained by the SOM. These nodes can capture not only the color, but also the contour of image content. By means of embedding a two dimension map into higher Euclid space, a weighted undirected graph is constructed. In consideration of edge symmetry in undirected graph, a manifold learning method, which combines undirected graph and semi-supervision, is further proposed. With supplied initial saliency values for nodes along image borders, the saliency values are computed for all nodes. Experimental results demonstrate the proposed model not only achieves high performance on precision and recall, but also presents a pleasing visual effect.

Key words: saliency detection; self organizing map; manifold learning; semi-supervised learning; undirected graph

在大数据时代, 非常有必要帮助人脑滤去不重要的信息从而使得人能关注于特别感兴趣的信息. 图像的显著度检测可以帮助人们把有限的视觉资源用于感知最有用的视觉信息. 根据特征集成理论(feature integration theory)^[1], Itti 等人提出了一种显著度检测模型, 通过在不同尺度上计算强度、颜色、方向等特征得到图像的显著度图^[2]. 在 Itti 的模型启发下, 后续又有多种新的利用图像颜色、结构、方向及运动等信息的显著度检测方法. 这些方法通过场景的底层特征刺激来驱动人的关注, 因此被称作 Bottom-up 框架. 该框架又分成两个流派, 即生物激励的和基于计算的方法. 前一种主要用于视觉关注点的检测^[2,3]. 后一种的检测目的为点亮整个关注的对

* 基金项目: 国家自然科学基金(61300140); 现代信息科学与网络技术北京市重点实验室开放课题(XDXX1307)

收稿时间: 2014-06-20; 定稿时间: 2014-08-20

象^[4-8],比如,Cheng 等人通过对图像进一步的量化使得颜色接近的像素集中到不同的聚类中,通过计算不同聚类之间的对比度来计算显著度^[5].Goferman 等人先对图像分块,再采用最近邻算法求取不同类型块之间的对比度^[6].Perazzi 等人把图像分成数百个类,根据这些类的位置和颜色分布来检测图像的显著度^[7].Achanta 等人基于结构是颜色的重要补充的假设,通过移除图像的低频信息来检测显著度^[4].它们的实质都是利用图像颜色或者结构的对比度.为了让对比度更好地体现对象和背景的区别,Wei 等人通过背景优先的方法,对边界和中心部分采用不同的权重,取得了很好的检测效果^[8].

然而这些方法都忽视了一个重要的线索,即图像的对象和背景具有明显不同的组织结构,因此本文提出一种图像的自组织映射(self organizing map,简称 SOM)方法.与已有用于图像分割的 SOM 方法相比^[9],提出的方法充分考虑到拓扑邻域的收缩,因而可以通过输出结点更好地刻画体现图像颜色和轮廓的组织结构.从图 1(c)的结点图中可以看出,对象结点很少触及图像的边界,而触及边界的大多属于背景结点.如图 1(d)所示,虽然对象和边界之间的结点从位置上难以区别,但它们的轮廓和颜色具有明显的特征:属于背景则轮廓和颜色非常接近边界结点,反之亦然.然而如何让算法去区分它们,仍然具有很大的难度.因此,我们进一步提出一种流形学习(manifold learning,简称 ML)的方法,根据 SOM 得到的结点图构造带权无向图,采用拉普拉斯正则化最小二乘(Laplacian regularized least square,简称 LRLS)检测背景结点和对象结点的显著度.LRLS 算法在 SAR 目标识别上得到了很好的应用^[10].这类半监督学习方法的关键在于数据的预分类,而提出的 SOM 具有良好的预分类性能,因此通过设置边界结点和对象结点的显著度预期值,我们可以采用 LRLS 算法来实现图像的显著度检测.

据我们所知,还未见其他文献把 SOM 和 ML 用于图像显著度检测.实验结果表明,本文方法和近几年提出的几种经典方法相比,无论是在 Precision-Recall 性能还是主观视觉效果方面,都得到了明显的提高.

本文第 1 节描述采用 SOM 得到反映图像固有拓扑的结点图.第 2 节描述基于带权无向图和拉普拉斯正则化的流形学习与显著度检测.第 3 节是详细的实验结果与分析.第 4 节是结束语.

1 图像的自组织映射

SOM 是一种特殊的人工神经网络,其输出神经元之间相互竞争以求被激活.在 SOM 中,神经元被放置在二维网格结点上.在神经系统的信息处理基本结构中,一个计算映射由神经元阵列定义,它们并行处理携带信息的传感信号.神经元将输入信号转变为空间位置编码的概率分布,通过映射中最大相关激活的位置标示参数的计算值.网络在突触权值初始化之后,SOM 算法可以分成 3 个主要过程,即竞争、合作以及突触调节.

1.1 竞争过程

首先把一幅输入图像划分为 N 个区域,记作向量 $\bar{x} = [x_1, x_2, \dots, x_N]^T$.网络中每个神经元的突触权值向量和输入空间的维数相同,神经元 l 的突触权值向量记为 $\bar{w}_l = [w_{l1}, w_{l2}, \dots, w_{lN}]^T$, $l = 1, 2, \dots, L$,其中 L 是网络中神经元的总数.为了找到输入矩阵 \bar{x} 与突触权值向量 \bar{w}_l 的最佳匹配,求取 \bar{x} 和 \bar{w}_l 的最小欧氏距离.用索引 i 表示输入矩阵 \bar{x} 的最优匹配神经元编号,它可以通过(1)式的条件决定.

$$i = \arg \min_l \|\bar{x} - \bar{w}_l\| \quad (1)$$

如果把 L 看成是 k -means 算法的迭代总次数,那么神经元 \bar{w}_i 就是第 l 次分类得到的 N 个聚类中心.实际上,广义 Lloyd 算法就是一种 SOM,而 k -means 算法则是一种不受侧向距离限制的 SOM.

1.2 合作过程

获胜神经元位于合作神经元的拓扑邻域中心.用 h_i 表示以获胜神经元 i 为中心的拓扑邻域且包含一组合作神经元, D_i 表示获胜神经元 i 和合作神经元 l 之间的侧向距离.SOM 的一个特征是随着离散时间 n 其拓扑邻域在宽度系数 $\sigma_h(n)$ 控制下收缩,因此拓扑邻域按(2)式选择.

$$h_i(n) = \exp(-D_i^2 / \sigma_h^2(n)) \quad (2)$$

1.3 突触调节过程

为了使网络成为自组织,神经元 l 的突出权值向量 \bar{w}_l 需要随着输入向量 \bar{x} 而改变.假定时间 n 神经元 l 的权值向量为 $w_l(n)$,更新权值向量 $w_l(n+1)$ 在时间 $n+1$ 被定为式(3).

$$w_l(n+1) = w_l(n) + \eta(n)h_{li}(n)[\bar{x}(n) - \bar{w}_l(n)] \quad (3)$$

其中, η 是算法的学习效率参数.式(3)具有将获胜神经元 i 的突触权值向量 \bar{w}_i 向输入向量 \bar{x} 移动的作用.我们可以采用一幅 RGB 图像作为输入来说明一个由 12×16 个神经元组成的二维网格的 SOM 过程.如图 1(a)所示,突触权值 \bar{w}_i 选为 \bar{x} 的每个区域中心位置像素在 Lab 颜色空间中的值.SOM 算法的拓扑排序性质在图 1(b)中得到了很好的说明.尤其是算法在收敛的时候,捕捉到了如图 1(c)所示的输入图像固有的拓扑.下面的叙述中,把获胜神经元的突触称作结点.

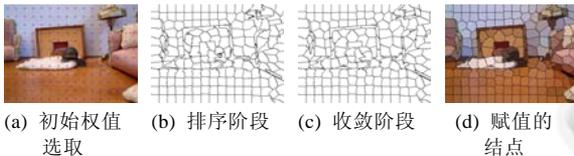


图 1 图像的自组织映射

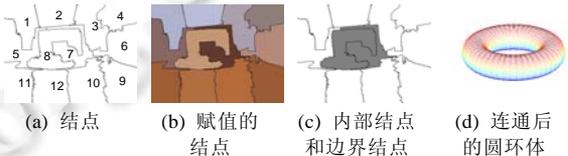


图 2 3×4 个神经元构成的连通环

2 流形学习的显著度检测

流形学习的目的是根据指定的学习样本,学习一个输入/输出的关系,对于每个输入的结点 C_j 都有一个预期的响应值 d_j 输出.每幅图像经过 SOM 后输出几百个结点,对这些结点做手工标注是非常费时费力的.然而,关注对象即便偏离了图像中心,也很少会像背景那样触及整个图像四边,这就使我们可以容易地获得部分带标注的样本:图像四边的结点因为具有较大的概率成为背景,因此把它们看成是标注的,而把内部结点当成不带类标的结点.如何利用这些带类标和不带类标的样本来训练网络,是半监督学习的关键问题.

流形学习的方法可以把图 1(c)这样的二维结点拓扑空间嵌入到高维的欧氏空间中.这样,就可以通过了解流形以及其外围空间的几何属性,设计出有效的半监督学习算法.其关键是必须知道描述输入空间内在几何结构的流形特征.本节将采用带权无向图的方法,构造流形学习的模型.

2.1 带权无向图的构造

在不失一般性下,我们用一个仅有 3×4 个神经元突触组成的网络构造包含 12 个结点的带权无向图.从图 2(a)看到,结点很好地保持了原图的轮廓.如图 2(b)所示,为每个结点赋予获胜神经元的突触权值,还能看到每个结点的颜色非常接近原图中所在区域的颜色.

一般来说,关注对象总是处于图像的非边界部分,因此区分出图像的内部和边界结点非常重要.通过给获胜神经元的每个突触标注标号,得到如图 2(a)的带标号的结点拓扑图.通过对图 2(b)边界上像素的判断,可以区分拓扑图的边界结点和内部结点.如图 2(c)所示,白色表示边界结点,灰色表示内部结点.这样,根据结点标号可以得到式(4)的关于边界和内部结点之间关系的矢量.

$$border = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1] \quad (4)$$

矢量 *Border* 的索引表示结点的编号,元素值 1 或 0 分别表示边界或者内部结点.连通性是 LRLS 学习算法的重要条件.以下从两个方面来构造带权无向图:连通边界处的结点;连通具有共享边的结点.

2.1.1 边界结点的连通

从图 2(b)中可以看出,边界结点经常具有接近的突触权值.然而,在二维的欧氏空间中,边界结点并不都是邻近的或者连通的.因此,需要进一步把边界上的结点图做连通的处理,即把二维结点图的上下两条边界连接在一起,构成一个柱状体,然后把柱状体首尾相连,得到一个如图 2(d)所示的三维圆环体,从而保证突触权值接近的图像边界结点的连通性.

2.1.2 具有共享边结点的连通

如果两个节点具有共享边,则称它们间接连通.比如,在图 2(a)中结点 8 和结点 3、6、9 没有直接连通,但它们具有共享边即结点 7,此时则认为结点 8 和结点 3、6、9 间接连通.为了把上述三维的拓扑空间嵌入到更高维的欧氏空间中,需要把间接连通结点看成是直接连通的.

如果忽视直接或者间接连通的差别,那么,如果任意两个结点 j 和 k 是连通的,则可以用 W_{jk} 表示连接结点 j 和 k 的无向边的权值.因此对于图 2 的例子,我们可以采用一个 12×12 的权值矩阵表示两两结点之间的权值.在这个矩阵中,列和行索引都表示结点的编号,某个位置比如 j 行 k 列的元素为非零值,这表明结点 j 和结点 k 是连通的,从而实现了带权无向图的构造.

2.2 权值矩阵的拉普拉斯正则化

通过上述无向图的构造实现了无向边的连通性.如果 W_{jk} 还满足对称性和非负性,则可以采用 LRLS 算法为图 2(a)中的结点自动标记类标.由于 $W_{kj} = W_{jk}$,所以满足无向边的对称性.由于 W_{jk} 是距离度量,因此无向边的非负性也可以满足.

把由 W_{jk} 构成的权值矩阵记作 \mathbf{W} .用 \mathbf{T} 表示一个 $N \times N$ 的对角阵,对角线 (k,k) 上的 t_{kk} 元素称为结点 k 的度,为权值矩阵中第 k 行或者第 k 列中所有元素的和.结点的度越大,表示这个结点越重要.因此我们可以得到拉普拉斯矩阵 $\mathbf{L} = \mathbf{T} - \mathbf{W}$.显然,由于某个结点和它自身的度为 0,因此 \mathbf{L} 矩阵对角线上的元素值为 t_{kk} ,其他位置上的元素值为 $-W_{jk}$.即 \mathbf{L} 是一个对称阵.此时,我们可以用一个正则化代价函数来表示显著度检测的误差.

$$J(S) = \sum_j [d_j - S(C_j)]^2 + \lambda_A \|\mathbf{S}\|_K^2 + \lambda_I \bar{s}^T \mathbf{L} \bar{s} \quad (5)$$

在式(5)中,等号右边第 1 项为经验项,第 2 与第 3 项为正则项. d_j 表示结点 j 期望的输出结果, $S(C_j)$ 表示实际求得的结点 j 的显著度. λ_A 和 λ_I 分别为外围和内在正则化参数, $\|\mathbf{S}\|_K^2$ 是再生核 Hilbert 空间中的范数, \bar{s} 是关于 $S(C_j)$ 的矢量,可以写成式(6)的形式.

$$\bar{s} = [S(C_1), S(C_2), \dots, S(C_N)]^T \quad (6)$$

根据广义表示定理,正则化代价函数 $J(S)$ 的优化产生一个如(7)式的扩张形式.

$$S(C_j) = \sum_{k=1}^N a_k W_{jk} \quad (7)$$

式(7)显示,正则化代价函数以及经验项即式(5)右边的第 1 项仅依赖于扩张系数 a_k 和权值矩阵 \mathbf{W} .那么,式(5)可以改写成:

$$J(\bar{a}) = \sum_j (\bar{d} - \mathbf{W}\bar{a})^T (\bar{d} - \mathbf{W}\bar{a}) + \lambda_A \bar{a}^T \mathbf{W}\bar{a} + \lambda_I \bar{a}^T \mathbf{L}\mathbf{W}\bar{a} \quad (8)$$

其中, \bar{a} 是由扩张系数 a_n 构成的矢量, \bar{d} 是由预期显著度 d_i 构成的矢量.对式(8)求关于 \bar{a} 的微分,可以求得代价函数 $J(\bar{a})$ 最小值点的 \bar{a} :

$$\bar{a}^* = (\mathbf{W} + \lambda_A \mathbf{I} + \lambda_I \mathbf{L}\mathbf{W})^{-1} \bar{d} \quad (9)$$

把求出的 \bar{a}^* 代入式(7)中,可求出相关结点的显著度.式(9)中,如果选定了 λ_A 和 λ_I ,为了求得 \bar{a}^* ,还需要知道 \bar{d} ,这正是 LRLS 算法的关键.由于 SOM 输出的结点图反映了图像固有的拓扑结构,使得我们可以把边界处的结点标注为背景,因此可以确定边界结点显著度的预期值 \bar{d} .

2.3 显著度检测

显著度检测分成 3 个步骤.第 1 步是求取边界结点的显著度,第 2 步是求取内部结点的显著度,第 3 步实现所有结点的显著度检测.首先把边界结点的显著度预期为 1,因此对于图 2 所示的例子,根据式(4)有:

$$\bar{d}_{border} = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1] \quad (10)$$

把 \bar{d}_{border} 代入式(9)中的 \bar{d} ,可求出使得式(5)代价函数取最小值的扩张系数矢量 \bar{a}^* .再把 \bar{a}^* 代入式(7),即可以求出边界结点的显著度 $S_{border}(C_j)$.然后取 $S_{border}(C_j)$ 均值 T_{border} 的 1.5 倍作为判据,以便对边界结点的显著度按式(11)二值化.

$$d(C_j) = \begin{cases} 1, & \text{if } S(C_j) \geq 1.5T_{border} \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

可以把式(11)看成是对式(10)的修正,因为式(10)是从结点的位置分布来判断一个结点是否为背景结点.而式(11)还从颜色对比度的角度去衡量它是否为背景结点.对 $d(C_j)$ 取反得到 \vec{d}_{inner} .用 \vec{d}_{inner} 替换式(9)中的 \vec{d} ,即把 \vec{d}_{inner} 当成是内部结点的预期显著度,可以得到内部结点的扩张系数矢量 \vec{a}^* .再把 \vec{a}^* 代入式(7),即可以求出内部结点的显著度 $S_{inner}(C_j)$.把 $S_{border}(C_j)$ 和 $S_{inner}(C_j)$ 归一化处理到 0~255 之间,得到如图 3(a)和(b)所示的边界和内部结点的显著度图.根据 $\vec{S}_{border} = 255 - S_{border}(C_j)$ 对 $S_{border}(C_j)$ 取反,得到如图 3(c)所示的显著度图.把 \vec{S}_{border} 和 $S_{inner}(C_j)$ 叠加并归一化,得到如图 3(d)所示所有结点的显著度图.可以看到,算法能够学习背景和边界结点的相似性,并做出了正确的检测.



图 3 显著度检测的主要步骤

图 4 边界处结点的显著度

在实验中,边界和内部结点采用相同的正则化参数.权值矩阵 \mathbf{W} 按照(12)式中的高斯函数求取.

$$W_{jk} = \exp\left(-\frac{\|w_{ij} - w_{ik}\|}{\sigma_w^2}\right) \quad (12)$$

σ_w^2 取 0.125.正则化参数 λ_A 和 λ_I 的选取非常重要,它们会直接影响实验的结果.既要考虑经验项的失真度,又要考虑正则项的粗糙度. λ_A 和 λ_I 的选择,理论上可以依据广义交叉验证函数最小化得到.经过在 ASD^[4]等数据集上反复测试,我们选定 $\lambda_A = 0.0002$, $\lambda_I = 10$.最后用算法 1 对提出的方法做一个概括.

算法 1. 自组织映射与流形学习的图像显著度检测.

- 1) 自组织映射得到结点图
- 2) 构造带权无向图并计算边的权值矩阵 \mathbf{W}
- 3) 确定边界结点预期显著度 \vec{d}_{border} 代入(9)式求取边界结点的扩张系数矢量 \vec{a}^*
- 4) \vec{a}^* 代入式(7)求取边界显著度 $S_{border}(C_j)$
- 5) 对 $S_{border}(C_j)$ 二值化得到 $d(C_j)$, 并取反得到内部结点的预期显著度 \vec{d}_{inner}
- 6) \vec{d}_{inner} 代入式(9)求取内部结点的扩张系数矢量 \vec{a}^*
- 7) \vec{a}^* 代入式(7)求取内部结点的显著度 $S_{inner}(C_j)$
- 8) $S_{border}(C_j)$ 与 $S_{inner}(C_j)$ 分别归一化
- 9) $S_{border}(C_j)$ 取反得到 $\vec{S}_{border} = 255 - S_{border}(C_j)$
- 10) 归一化 $\vec{S}_{border}(C_j) + S_{inner}(C_j)$ 为各个结点的显著度

3 实验结果

性能测试主要针对含有 1 000 幅图像的 ASD 数据集,它提供了二值化的 Ground truth(简称 GT)图^[4],便于测试 Precision 和 Recall 等指标.提出的 ML 算法与其他 8 种经典算法做比较,这些算法是 Itti 等人提出的 IT^[2],Harel 等人提出的 GB^[3],Achanta 等人提出的 FT^[4],Cheng 等人提出的 RC^[5],Goferman 等人提出的 CA^[6],Perazzi 等人提出的 SF^[7],Wei 等人提出的 GS^[8],以及 Jiang 等人提出的 CB^[11].除了 IT,GB,FT 外,其余都是近年提出的算法.

3.1 边界处结点的显著度

从图 4(a)可以看到,虽然算法第 3 步中把对象中接近底边的结点的预期显著度设置为 1,但是由于式(11)的修正作用,得到的显著度图依然可以很好地区分背景和对象.从图 4(b)可以看到,虽然图像下半部屋顶的颜色和关注对象非常一致,但由于这部分图像区域对应的结点在算法第 3 步中被预期为边界,而且含有的结点数较多,

使得式(11)无法对它修正,因此这部分区域在显著度图中被明显地抑制.

3.2 边界结点连通方式的比较

我们分析两种边界结点连通方式对检测结果的影响.一种是让对边上所有的边界结点连通,另外一种是让4条边上所有的边界结点连通.对于 ASD 中 1 000 幅图像的平均 Precision-Recall 曲线而言,对边连通的方法比4边连通大约低 0.01.然而,如图 5 所示,对某些图像而言,对边连通的视觉效果好于四边连通.



图 5 边界结点连通方式的影响

3.3 Precision-Recall与F-measure

通过与某阈值比较,可以把显著度图转换成二值图.再让二值图和 Ground truth 做集合运算,可以计算出该阈值下对象检测的 Precision-Recall 和 F-Measure 值.如果让阈值在 0~255 间变化,可获得如图 6 所示的 Precision-Recall 以及 F-Measure 曲线.可以看出,在 Recall 接近 1 的区间,GS 的 Precision 比提出的方法好.在阈值接近 255 的区间,GS 的 F-measure 也比提出的方法好.但总的来看,提出的方法的性能明显高于其他方法.

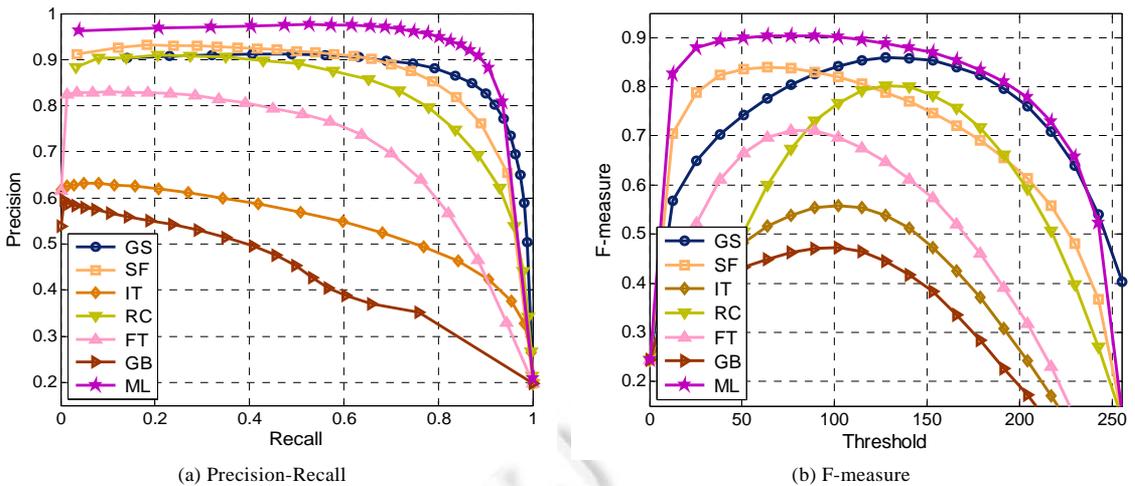


图 6 Precision-Recall 与 F-measure 比较

3.4 主观效果

图 7 给出了所提出方法及其他 6 种方法的显著度图,主要列出一些背景较复杂图像的检测结果.GB 作为一种早期的基于图的方法,非常有效地点亮了关注区域中重要的像素,特别是对对象边缘处的像素.FT 方法对于背景不很复杂的图像取得了很好的效果,因为它是基于对象含高频信息比较丰富的假设.RC 相比于 FT,在对树叶图像的检测上有了很大的提高,这是因为量化使得背景颜色趋于一致,从而体现出前景的信息量.基于上下文感知的 CA 和 SF 在有效抑制背景纹理同时,准确地点亮了对象.GS 充分考虑了图像内容的位置分布特征,属于一种先验的处理方法,它与上述几种方法相比,性能都有了很大的改善.但对背景的抑制效果还不够理想.提出的 ML 算法,在有效抑制背景的同时,准确地点亮了对象区域,尤其对于松鼠和鸟这两幅前后景颜色接近的图像.



图 7 ASD 数据集上显著度图的主观比较

如图 8 所示,在 Berkeley 大学提供的 BSD300 数据集上的测试结果也表明 ML 的性能要好于其他算法.该数据集的 Ground truth 由 Wei 等人提供^[8].

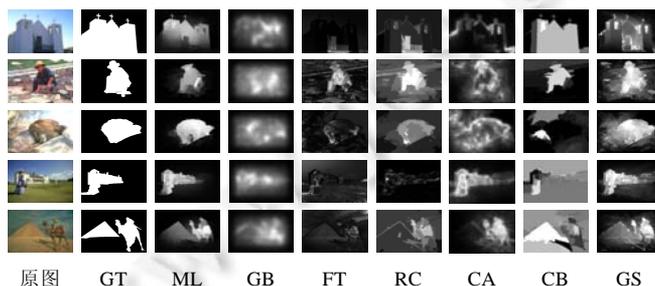


图 8 BSD 数据集上显著度图的主观比较

3.5 时间开销

与算法的其他步骤相比,求逆矩阵的计算复杂度最高,为 $O(N^3)$, N 为结点数.我们在 2.50 GHz 的 Intel i5 4200M CPU 以及 4GB RAM 的机器上运行提出的 ML 算法,运行环境为 Matlab R2012a.对于 ASD 数据集,平均每幅图像的处理时间为 1.42s.其中,自组织映射占 0.22s;由于存在大量的控制相关,无向图构造占 0.7s;流形学习的显著度检测占 0.5s.ML 实现算法的所有程序都为 .m 文件,没有采用任何交叉编译.有些算法运行环境也为 Matlab,比如 GB,虽然时间开销为 1.36s,但算法的核心部分采用了 .mex 文件.

4 结束语

作为一种半监督学习的算法,本文提出的采用流形学习的显著度检测方法取得了很好的性能,其原因首先是自组织映射这种数据预分类处理方法的良好性能;其次是无向图的构造把二维的结点图转换为高维的无向图,准确刻画流形以及其外围空间的几何属性;其三是对 LRLS 算法的灵活运用,通过边界结点成为背景的概率更大的假设,逐步检测出整幅图像的显著度.今后的研究工作应该进一步提高自组织映射的性能,使得背景和对象的结点更能反映图像的内容,并结合机器学习的最新进展,采用更好的学习算法.

References:

- [1] Treisman AM, Gelade G. A feature-integration theory of attention. *Cognitive Psychology*, 1980,12(1):97-136.
- [2] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1998,20(11):1254-1259.
- [3] Harel J, Koch C, Perona P. Graph-Based visual saliency. In: *Proc. of the Neural Information Processing Systems 2006*. 2006. 545-552.
- [4] Achanta R, Hemami S, Estrada F, Sússtrunk S. Frequency-Tuned salient region detection. In: *Proc. of the Computer Vision and Pattern Recognition 2009*. 2009. 1597-1604.
- [5] Cheng MM, Zhang GX, Mitra NJ, Huang XL, Hu SM. Global contrast based salient region detection. In: *Proc. of the Computer Vision and Pattern Recognition 2011*. Springer-Verlag, 2011. 409-416.
- [6] Goferman S, Zelnik-Manor L, Tal A. Context-Aware saliency detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012,34(10):1915-1925.
- [7] Perazzi F, Krähenbühl P, Pritch Y, Hornung A. Saliency filters: Contrast based filtering for salient region detection. In: *Proc. of the Computer Vision and Pattern Recognition 2012*. 2012. 733-740.
- [8] Wei YC, Wen F, Zhu WJ, Sun J. Geodesic saliency using background priors. In: *Proc. of the European Conf. on Computer Vision 2012*. 2012. 29-32.
- [9] He QF, Lu S, Hao QF, Li GJ. A new approach to automatic extraction of discriminant regions in image. *Chinese Journal of Computers*, 2002,25(8):801-809 (in Chinese with English abstract).
- [10] Zhang XR, Yang C, Jiao LC. Semi-Supervised SAR target recognition based on Laplacian regularized least squares classification. *Ruan Jian Xue Bao/Journal of Software*, 2010,21(4):586-596 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3538.htm> [doi: 10.3724/SP.J.1001.2010.03538]
- [11] Jiang H, Wang JD, Yuan ZJ, Liu T, Zheng NN, Li SP. Automatic salient object segmentation based on context and shape prior. *Proc. of British Machine Vision Conf.* 2011. 2011,110:1-12.

附中文参考文献:

- [9] 何清法,鲁松,郝沁汾,李国杰.一种自动抽取图像中可判别区域的新方法. *计算机学报*,2002,25(8):801-809.
- [10] 张向荣,阳春,焦李成.基于 Laplacian 正则化最小二乘的半监督 SAR 目标识别. *软件学报*,2010,21(4):586-596. <http://www.jos.org.cn/1000-9825/3538.htm> [doi: 10.3724/SP.J.1001.2010.03538]



陈加忠(1970—),男,浙江嘉善人,博士,副教授,主要研究领域为图像和视频处理,计算机视觉,多媒体数据分布式计算.



苏曙光(1975—),男,博士,讲师,主要研究领域为视觉跟踪,图像与视频处理.



曹华(1971—),男,博士,讲师,主要研究领域为图像水印,视觉跟踪.



伊斯刚(1963—),男,讲师,主要研究领域为嵌入式环境下的视觉信号处理.