

文章编号:1001-9081(2008)07-1738-04

权重马氏距离高斯核在谱分割中的应用

陈应良, 王士同

(江南大学 信息工程学院, 江苏 无锡 214122)

(wxchenyingliang@126.com)

摘要:为了使经典谱分割的 Nyström 采样快速算法得到更清晰的结果,将权重马氏距离高斯核应用于其中,相对于常用的马氏距离高斯核,得到了更好的分割效果。结果表明,使用权重马氏距离高斯核更能准确的反映两个向量的相似度,从而实现准确的分割。

关键词:谱分割;聚类;Ncut;Nyström 估计;权重马氏距离

中图分类号: TP391.41 **文献标志码:** A

Application of WMD Gaussian kernel in spectral partitioning

CHEN Ying-liang, WANG Shi-tong

(College of Information Engineering, Jiangnan University, Wuxi Jiangsu 214122, China)

Abstract: To obtain a better segmentation result, this paper used Weighted Mahalanobis Distance (WMD) Gaussian kernel for Nyström-Ncut segmentation. It proves that weighted Mahalanobis distance Gaussian kernel is more appropriate for spectral graph theoretic methods than Mahalanobis distance, because weighted Mahalanobis distance can compute the similarity between two pixels more accurately.

Key words: spectral graph partition; clustering; Ncut; Nyström approximation; Weighted Mahalanobis Distance (WMD)

0 引言

图像的谱分割技术近年来在国际上的图像分割领域内表现出了良好前景^[1],但在国内相关的报道却很少。其传统方法是:将原图像映射为带权的无向图,把像素视为节点,用相应的核函数计算任意两个节点的距离,视为两点之间的权值,再利用最小剪切准则得到图像的最佳分割。其中具体的实现需要计算亲和力矩阵,该矩阵元素分别表示任意两像素点之间的相似度距离,所以对 $m \times n$ 的图像必须计算 $(m \times n) \times (m \times n)$ 维的亲和力矩阵,针对这种耗时的传统方法,有研究者提出了用 Nyström 方法对亲和力矩阵进行估计,并能直接依此估计出应用于最后分割的亲和力矩阵的主特征向量,大大地减少了计算时间,却不影响分割结果。但是在传统方法和 Nyström 估计方法对亲和力矩阵构造和估计的时候,使用的核函数都比较单调。本文旨在引进一种新的核函数——权重马氏距离的高斯核,这种核函数兼顾了传统的马氏距离高斯核,而且考虑到原像素点所在类别大小对相似度的影响,因此更能准确地构造和估计相似度矩阵,使最终的分割效果有所改善。

1 谱分割基本理论及常用的核函数

1.1 谱分割基本理论及其一种快速算法

基于图论的图像分割与图像的谱分割在本质上是一致的,该方法是将图像映射为带权无向图,把像素视作节点,利用最小剪切准则得到图像的最佳分割。

设原图像有 $m \times n$ 个像素点,将其映射到一个带权的无向图 $G = (V, E)$,像素集被看作节点集 G ,边缘集被看作边集

E ,其中两个像素点之间的权值是通过相似度距离公式(权函数)得到的,可以存放在亲和力矩阵 W 中。

为了实现图像 I 的最佳分割,可以根据目前的最优割集准则: Minimum cut、Average cut、Normalize cut^[2]、Min-Max Cut^[3]、Ratio cut^[4]、Bcut^[5]、Nested cut^[6]、Foreground-cut 等来完成。

图像的谱分割方法是直接对原图像构造亲和力矩阵 W 或 Laplacian 矩阵 L ,然后求解矩阵的特征矢量(是最优割集的一种实现方式)并以此为基础直接或进一步构造特征矢量指导分割,可以参照常用的谱方法^[7-8];其中,经典的谱分割使用 Ncut(Normalize cuts)割集的方法^[2],具体思想可以参考原文。

谱分割的分割效果和发展前景都很不错,但仍有许多不足之处。最明显的就是随着图像的增大,亲和力矩阵成幂级数增长,不但会引起计算机内存不够,而且会浪费大量的计算时间,更谈不上计算亲和力矩阵的特征矢量了。

研究表明,通过 Nyström 采样方法来估计亲和力矩阵及其主特征向量,就很好地解决了上述问题^[9-11],基本思路与经典的谱分割比较:经典的谱分割用的是图像的 L 氏矩阵 $L = D^{-1/2} (D - W) D^{-1/2} = I - D^{-1/2} W D^{-1/2}$,求此矩阵的第二最小特征值对应的特征向量对图像进行二值指导分割,可以使用递归的方法对图像进行多次分割,以实现最终目的^[4]。而基于 Nyström 采样方法的谱分割使用的特征式系统是 $D^{-1/2} W D^{-1/2}$ (W, D 是通过 Nyström 采样估计得到的),求此式的 k 个主特征值对应的特征矢量,按列排好后,把每个像素嵌入到一个 k 维的行向量,再对这些向量使用 k -means 方法聚类,从而实现了图像的分割^[9]。

收稿日期:2008-01-07;修回日期:2008-03-24。

作者简介:陈应良(1983-),男,安徽怀宁人,硕士研究生,主要研究方向:人工智能、模式识别;王士同(1964-),男,江苏扬州人,教授,博士生导师,主要研究方向:人工智能、模式识别、神经模糊系统、生物信息学。

1.2 常用的核函数

虽然 Nyström-NCut 方法才提出时,只能限定使用正定核,以至于 one-shot 就能估计出主特征向量,后来有研究者做出改进,对核函数不需做特殊限制,只要 two-step 就能估计出主特征向量^[11]。以下列举常用的核函数,同时分析各自的优点和使用场合。

Gaussian weighed distance:

$$W_{ij} = e^{-1/2(x_i-x_j)^T \Sigma^{-1}(x_i-x_j)}$$

其中: x_i, x_j 分别表示两个像素点的特征向量,此函数为正定核,是最常用的衡量两个像素特征向量相似度的核函数,应用于 Nyström-NCut 时可以用 one-shot 就能估计出主特征向量,同时马氏距离使用了向量间的协方差,消除了向量之间各个指标之间的相互影响,比欧氏距离更准确。

Histogram comparison using the χ^2 test:

$$\chi_{ij}^2 = \frac{1}{2} \sum_{k=1}^K \frac{(h_i(k) - h_j(k))^2}{h_i(k) + h_j(k)}, W_{ij} = e^{-\frac{\chi_{ij}^2}{\alpha}}$$

其中: $h_i(k), h_j(k)$ 表示两个像素的直方图,则 W_{ij} 表示这两个直方图的相似度,同时此核函数也是正定的,高斯权值 χ^2 核函数适合彩色纹理图像像素相似度的度量。

Intervening contour:

令两像素点所有路径能量最小者为 r_{ij} , 则衡量两点相似

性的核函数为 $W_{ij} = e^{-\frac{r_{ij}^2}{\alpha}}$, 此函数目前没有证明是否正定,适合于对图像边界提取来分割图像。

One minus squared distance:

$$W_{ij} = 1 - \frac{r_{ij}^2}{\alpha}$$

其中: r_{ij} 表示两点特征向量的距离,此核不正定,但与 Nyström-NCut 时使用 two-step 估计出的主特征向量能得到一个比较满意的分割效果。

2 权重马氏距离的高斯核

通过上文的分析可以看出, Nyström-NCut 方法中可以使用多种核函数,但是每个核函数有各自的优点和适用场合,结合本文后面灰度图像分割的实验,这里只对高斯核这种最常用的核函数做出相应的改进,其他的核函数可以借鉴这种方法。

计算两个像素点的相似度时,使用的核函数都是基于两个点的特征向量的距离,以下分析从欧氏距离到马氏距离,再到本文提出的加权马氏距离(可参照 IEEE2007 收录的文章 Weighted Mahalanobis Distance Kernels for Support Vector Machines)对度量两个像素点特征向量距离的影响。

图像的谱分割中,像素点的特征向量就是图像的样本点。设 $X = (x_1, \dots, x_p)'$, $Y = (y_1, \dots, y_p)'$, $Z = (z_1, \dots, z_p)'$, 则定义距离一般要满足下面三个公理:

- 1) $d(X, Y) \geq 0, d(X, Y) = 0 \Leftrightarrow X = Y$
- 2) $d(X, Y) = d(Y, X)$
- 3) $d(X, Z) \leq d(X, Y) + d(Y, Z)$

距离公式中常用欧氏距离 $d_v(X, Y) =$

$\sqrt{\sum_{k=1}^{k=q} (x_k - y_k)^2}$, 由于每个向量中的各个指标的量纲不同,而且指标之间因不一定独立而相互影响,所以得到的距离结果与实际往往出现很大的偏差,但是马氏距离就能很好地克服这两个方面的缺点。

马氏距离比欧氏距离表示了更多的信息,假设 X, Y 属于同一个类,则其可以表示为

$$d_M(X, Y, \Sigma) = \sqrt{(X - Y)^T \Sigma^{-1} (X - Y)}$$

其中 Σ 表示样本的协方差矩阵,如果 Σ 不是满秩的,则 Σ 的逆矩阵 Σ^{-1} 不存在,但是我们可以用 Σ 的伪逆 Σ^+ 代替 Σ^{-1} 。秩为 d 的半正定实对称矩阵 Σ 可以对角化为 $\Sigma = A^T \Lambda A$, 其中 Λ 是对角矩阵,只包含 Σ 的正特征值, A 的行向量与特征值对应,且单位正交。伪逆 Σ^+ 可以通过以下方式得到 $\Sigma^+ = A^T \Lambda^{-1} A$, 计算伪逆的时间复杂度为 $O(p^3)$, 这种情况下的马氏距离可以表示为 $d_M(X, Y, \Sigma) = \sqrt{(X - Y)^T A^T \Lambda^{-1} A (X - Y)}$ 。以上的马氏距离两种情况是针对 X, Y 两个向量属于同一个类,服从同一个分布,谱分割也就是使用这种近似估计来使用马氏距离的,但是当 X, Y 不属于同一个类时,我们可以使用加权的马氏距离(简称 WMD)来度量这两个不属于同一个类的向量。

假设 X, Y 属于两个不同的类别 S_X, S_Y , 那么我们可以定义加权的马氏距离为

$$d_{wM}(X, Y) =$$

$$\sqrt{\frac{|S_X|}{|S_X| + |S_Y|} \cdot d_M^2(X, Y, \Sigma_X) + \frac{|S_Y|}{|S_X| + |S_Y|} \cdot d_M^2(X, Y, \Sigma_Y)}$$

其中 Σ_X, Σ_Y 分别表示 S_X, S_Y 的协方差矩阵,此公式很好地反映了结构思想。加权的马氏距离不一定满足三角不等式,但是它是非负对称的,且 $d_{wM}(X, X) = 0$ 。

从公式我们可以看出, WMD 是 MD(马氏距离)的泛化,如果向量 X, Y 属于同一个类, WMD 就是 MD。但 X, Y 不属于同一个类时,类规模大的样本对距离的影响比类规模小的样本影响要大得多。

从上面的分析得到,基于欧氏距离或者马氏距离的核,我们可以直接用前面的加权马氏距离公式代替其中的欧/马氏距离,但是有些核是基于内积距离的,可以同样找到相应的加权马氏距离的形式代替内积形式。内积可以表示为 $I(X, Y) = X \cdot Y = \frac{1}{2} (\|X - 0\|^2 + \|Y - 0\|^2 - \|X - Y\|^2)$,

则相应的加权马氏距离可以表示为

$$d_{wM}^I(X, Y) = \frac{1}{2} (d_M^2(X, 0, \Sigma_X) + d_M^2(Y, 0, \Sigma_Y) - d_{wM}^2(X, Y))$$

如前所示,本文因需要使用权重马氏距离的高斯核,所以用基于欧氏距离的 WMD。提出了核的形式为

$$W_{XY} = e^{-\frac{1}{2} \left(\frac{|S_X|}{|S_X| + |S_Y|} \cdot d_M^2(X, Y, \Sigma_X) + \frac{|S_Y|}{|S_X| + |S_Y|} \cdot d_M^2(X, Y, \Sigma_Y) \right)}$$

其他核可以根据需要结合上面 WMD 的两种形式进行组合。

3 实验描述和分析

按照本文的开始部分介绍可知,经典的 NCut 算法比 Nyström-NCut 的时间复杂度和空间复杂度都要高很多,例如对一个 128×128 图像分割,经典的 NCut 算法与 Nyström-NCut 算法的时间空间分析比较如表 1。由表 1 可知,如果使用经典的 NCut 算法对权重马氏距离核函数做实验,不但要耗费大量的时间,而且对内存有较高的要求。所以我们选取 Nyström-NCut 采样快速算法的谱分割方法对我们提出的权重马氏距离进行实验,因为这种算法的时间空间复杂度比较低,

而且能得到比较理想的分割。

下面的两个实验从不同角度来测试权重马氏距离高斯核在谱分割中的性能和优点,实验一使用权重马氏距离高斯核的 Nyström-NCut 谱方法对图像 Lena 分割,在使用相同的主特

征向量数情况下观测后期使用的 k -means 分成不同的聚类数的分割效果。实验二用四张不同的图像在相同的主特征向量数和聚类数的情况下,对比使用马氏距离高斯核和加权马氏距离高斯核的 Nyström-NCut 谱分割效果。

表 1 两种算法时间空间消耗对比

对比项	经典的谱分割(Ncut)	Nyström-NCut 采样快速算法的谱分割
时间	通过计算求得 16384×16384 维的亲力和矩阵,而每个矩阵元素是通过复杂的距离公式计算得到,同时在计算 L 氏矩阵时必须计算 16384×16384 的对角矩阵 D ,除此之外,还要计算 L 氏矩阵的特征值和特征向量以求得第二小的特征值对应的特征向量	只需要计算矩阵 A 和 B ,其维数分别为 $100 \times 100, 100 \times 16284$,由于高斯权值的正定性,就可以直接估计出 $D^{-1/2} W D^{-1/2}$ 的主特征向量
空间	必须在内存中存放一个庞大数组 16384×16384 ,而且还要通过计算存放它的特征向量	不需要存放亲和力矩阵 W ,对它的估计只是一个理论证明过程,并不计算,更不必存放在内存中,仅仅只需要存放估计的 $D^{-1/2} W D^{-1/2}$ 的主特征向量即可

文献[3]证明,在 Nyström-NCut 采样快速算法的谱分割方法中只需对原图像采样 1% 即可以得到较高的估计质量与较好的分割效果。实验一使用 128×128 的灰度图像,提取三维的特征值向量,分别为像素点的横纵坐标和灰度值,从 $128 \times 128 = 16384$ 个像素点随机抽取 100 个,首先根据像素的灰度值对要分割的图像预分成几个类,再选用加权马氏距离的高斯核

$$W_{ij} = e^{-\frac{1}{2} \left(\frac{|s_i|}{|s_i| + |s_j|} \cdot d_M^2(\mathbf{X}, \mathbf{Y}, \Sigma_i) + \frac{|s_j|}{|s_i| + |s_j|} \cdot d_M^2(\mathbf{X}, \mathbf{Y}, \Sigma_j) \right)}$$

其中: x_i, x_j 表示像素点 i 和 j 的特征值向量, Σ_i, Σ_j 分别表示像素点 i 和 j 所属类的协方差矩阵,在此可以采用 3 个主特征向量做指导分割(主特征向量个数可以自己根据具体图像选定,后面所有实验和算法都是以 3 个主特征向量为基础,其他情况与此类似)。

实验一显示了使用权重马氏距离高斯核的 Nyström-NCut 谱方法对图像分割的结果,得到一个比较理想的分割效果,但是随着后期聚类数的不同,得到不同的分割效果,而且分割效果很理想。为了不使人工调节聚类中心数对分割结果造成很大影响,可以在后期聚类过程中使用自动聚类算法,不用指定中心数,研究表明,能够针对地使用一些特定的算法,不属于本文讨论内容。

图 1 中:图(a)为原始图像 Lena,图(b)是对图(a)预划分三个区域结果(根据像素值),图(c)~(f)使用权重马氏距离高斯核的 Nyström-NCut 谱方法对图像分割,使用 k -means 最后聚类数目分别为 2,3,5,10。



图 1 实验一

实验二选择几张标准图像(child, baboon, camera, plane),

每个图像分别分成两块和三块来比较使用马氏距离高斯核和加权马氏距离高斯核的 Nyström-NCut 谱方法的分割效果。

实验结果如图 2 所示,MD 的 Nyström-NCut 谱分割将图像分割得不彻底,而 WMD 的 Nyström-NCut 谱分割分割得很清晰。其中 WMD 的 Nyström-NCut 方法中对图像做了初步处理:根据灰度将图像分成三个类,再对每个类求协方差矩阵,而 MD 的 Nyström-NCut 的方法把所有像素点的特征值看作同分布的,只计算一个协方差矩阵,图 2 所对应图像的协方差矩阵如表 2 所示。



图 2 实验二

Nyström-NCut 谱分割是通过计算部分像素间权值来估算整个图像的亲力和矩阵,因此它所使用的核函数尤为重要,理所当然,就传递到核函数所使用的距离公式很重要。马氏距离消除了欧氏距离中各个指标量纲的不同和各个量纲相关性

对计算机结果的影响。然而,在计算两个特征向量之间的距离的时候,马氏距离只粗略地认为两个向量属于同一个类,同分布,没有考虑到两个向量不属于同一个类时,结果依赖于类规模大小的情况,所以,本文提出的加权马氏距离更能贴切地

反映两个向量之间的相似度。所以,如图2,WMD的Nyström-NCut方法比MD的Nyström-NCut方法分割出来的图像更精确,并经反复实验得到,分割的块数越小这种效果越明显,这是受预分割块数的影响。

表2 用于两种分割方法的图像的协方差矩阵

图像名称	MD 中的协方差			WMD 预分割的三个协方差								
child	1365.300	0.000	1736.800	[980.23	114.79	7.3925;	114.79	2058.1	106.37;	7.3925	106.37	541.62]
	0.000	1365.300	-102.500	[977.18	-42.411	368.13;	-42.411	673.11	-52.917;	368.13	-52.917	531.81]
	1736.800	-102.500	6247.500	[61.727	-11.593	-0.079622;	-11.593	790.33	17.32;	-0.079622	17.32	14.838]
baboon	1365.300	0.000	336.230	[1238.7	315.87	98.827;	315.87	1666.4	153.23;	98.827	153.23	300.57]
	0.000	1365.300	44.474	[1339.7	-119.21	123.59;	-119.21	1273.7	-37.782;	123.59	-37.782	654.04]
	336.230	44.474	1421.900	[254.7	-19.976	2.4503;	-19.976	227.74	4.0667;	2.4503	4.0667	7.976]
camera	1365.300	0.000	-819.520	[716.62	-72.908	149.23;	-72.908	567.38	290.14;	149.23	290.14	746.85]
	0.000	1365.300	513.630	[1555.1	126.5	-693.29;	126.5	1550.2	-156.43;	-693.29	-156.43	523.66]
	-819.520	513.630	3621.300	[778.55	-88.759	55.679;	-88.759	140.06	-40.24;	55.679	-40.24	164.13]
plane	1365.300	0.000	-143.060	[267.45	206.56	42.732;	206.56	907.58	25.562;	42.732	25.562	363.81]
	0.000	1365.300	218.050	[1434.6	221.35	-464.47;	221.35	1024.2	-51.367;	-464.47	-51.367	908.75]
	-143.060	218.050	1448.400	[1380.4	-289.71	57.238;	-289.71	1684.4	83.883;	57.238	83.883	43.638]

4 结语

图像的谱分割本质上是把图像映射到一个带权的无向图,再把对图分割的问题转化为一种聚类问题。该方法在国际图像分割领域内已是一个研究热点,因为它是一种图像分割的新方法,具有良好的分割性能,而且表现出了很好的应用前景。但是涉及到的理论知识很多,具体的应用还处在初级阶段,在国内做这方面研究和应用的都很罕见。本文不但简要的介绍了谱分割的大致情况和基本知识,详细阐述和实现了谱分割的一种快速算法 Nyström-NCut 方法,并且总结了适用于本方法的一些核函数,尤其重点提出了一种加权马氏距离高斯核在 Nyström-NCut 方法中的应用,通过实验证明了这种方法的效果和优势。

参考文献:

- [1] 闫成新, 桑农, 张天序. 基于图论的图像分割研究进展[J]. 计算机工程与应用, 2006, 42(5): 11-14.
- [2] SHI J, MALIK J. Normalized cuts and image segmentation[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905.
- [3] DING C, HE XIAO-FENG, ZHA HONG-YUAN, *et al.* Spectral min-max cut for graph partitioning and data clustering[C]// Proceeding of the IEEE International Conference on Data Mining. Florida: IEEE Computer Society Press, 2001: 107-14.

- [4] WANG S, SISKIND J M. Image segmentation with ratio cut[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2003, 25(6): 675-690.
- [5] SCANLON J, FLORIDA N. Graph-theoretic algorithms for image segmentation[C]// Proceeding of the IEEE International Symposium on Circuits and Systems. Florida: Orlando, 1999, 6: 141-144.
- [6] VEKSLER O. Image segmentation by nested cuts[C]// IEEE Conference on Computer Vision and Pattern Recognition. Seattle: IEEE Press, 2000: 339-344.
- [7] ANDREW Y N, MICHAEL I J, YAIR W. On spectral clustering: Analysis and an algorithm[J]. Neural Information Processing Systems, 2001, 14(2): 849-856.
- [8] WEISS Y. Segmentation using eigenvectors: A unifying view[C]// Proceedings IEEE International Conference on Computer Vision. Bangkok: IEEE Press, 1999, 2: 975-982.
- [9] FOWLKES C, BELONGIE S, CHUNG F, *et al.* Spectral grouping using the Nyström method[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(2): 214-225.
- [10] FOWLKES C, BELONGIE S, MALIK J. Efficient spatio-temporal grouping using the Nyström method[C]// Proceeding IEEE Conference Computer Vision and Pattern Recognition. Hawaii: IEEE Press, 2001: 231-238.
- [11] BELONGIE S, FOWLKES C, CHUNG F, *et al.* Spectral partitioning with indefinite kernels using the Nyström extension[J]. European Conference on Computer Vision, 2002(3): 531-542.

(上接第 1725 页)

所以该算法是自适应调节的。同时,它在处理局部区域时没有引入复杂的计算工具,在效率和效果上取得了折中。进一步,该算法对映射的数据区间只进行了常规划分,对于亮度跨度不同的图像如何划分它们的映射区域,还没有完全考虑数据的统计特征,并且渐变系数决定曲线参数的方法的合理性还值得进一步的分析验证,在今后的研究中,希望该算法能够实现自动调节区间大小的功能。

参考文献:

- [1] TUMBLIN J, RUSHMEIER H. Tone reproduction for realistic image [J]. IEEE Computer Graphics and Application, 1993, 13(6): 42-

48.

- [2] YOURGANOV G, STUERZLINGER W. Tone-mapping for high dynamic range images [R]. Department of Computer Science, York University, 2001.
- [3] DRAGO F, MYSZKOWSKI K, ANNEN T, *et al.* Adaptive logarithmic mapping for displaying high contrast scenes[C]// Proceeding of Eurographics 2003. Granada, Spain: Computer Graphics Forum, 2003, 22(3): 419-426.
- [4] 李晓光, 沈兰荪, 林健文. 一种高动态范围图像可视化算法[J]. 计算机应用研究, 2007, 24(11): 303-305.