

两种近似 EMD 的图像检索方法 *

宋和平, 杨群生, 战荫伟

(广东工业大学 计算机学院, 广东 广州 510006)

摘要: 相似度量是图像检索的关键, EMD 是一种有效的度量距离, 但其计算比较复杂, 而且依赖于基本距离的选择。采用 Lloyd 聚类算法对图像进行高斯混合建模, 并以聚类失真作为基本距离, 提出了两种近似 EMD 的方法计算相似度。实验结果验证了该方法的有效性, 其检索效率与 EMD 方法接近, 而且计算复杂度比 EMD 方法低, 基本距离的选择敏感。

关键词: 图像检索; 劳埃德聚类; 推土机距离; 最小元素法; 伏格尔法

Two methods approximating EMD for image retrieval

SONG He Ping, YANG Qun Sheng, ZHAN Yin Wei

(Faculty of Computer, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Similarity measure is crucial for content based image retrieval, EMD is an efficient distance measurement, but it costs high computation complexity, and relies on efficient ground distance. In this paper, the images are modeled with Lloyd clustering, and two methods to approximate EMD with clustering distortion as ground distance are proposed. Experiment results demonstrate the effectiveness of the proposed methods. With similar performance compared with EMD, the proposed methods not only have lower computation complexity, but aren't sensitive to the ground distance.

Key words: image retrieval; Lloyd clustering; EMD; minimum factors method; vogel method

随着数码设备的普及和互联网的兴起, 每天都将产生海量数字图像。为了有效地存储、管理图像数据库, 需要对图像库进行索引, 按特定的需求检索图像。以往的图像检索模式是基于文本的, 采用关键字的方法, 需要大量的人工注释, 而且注释内容也存在很大的主观差异性, 往往不能反映图像的本质内容。基于内容的图像检索(CBIR)克服了传统方法的缺陷, 直接利用图像的内容如颜色、纹理、形状、空间关系等进行检索。特征提取和相似度量是 CBIR 的两个关键步骤, 特征提取是用颜色等特征按一定的方式概括图像内容, 从而获得图像的特征分布。相似度量是计算特征分布间的距离, 并以此作为图像间的相似度。常用的相似度量有 Minkowski 距离度量、直方图相交度量、Jeffrey 散度量、K-L 散度量等^[1]。

EMD(Earth Mover's Distance)是一种反映计算机视觉感知相似性的距离度量, 被广泛用于计算机视觉、模式识别、机器学习等领域。图像特征分布聚类后得到称为签名(Signature)的聚类中心及相应的权值。EMD 考虑了

不同签名的重要性, 使总的签名间距离最小。EMD 方法可以计算具有不同签名个数的图像间距离, 是一种多对多的匹配方法, 所以能计算部分匹配。如果签名间的距离即基本距离(Ground Distance)是一种度量(metric), 那么 EMD 也是一种度量。但 EMD 计算比较复杂, 不同应用需根据要求选择有效的的基本距离^[2]。本文提出两种近似 EMD 方法(最小元素法(MFM)和 Vogel 法)计算图像间的相似度, 其计算复杂度比 EMD 方法低。在本文图像检索框架下, 两种近似 EMD 的方法对基本距离的选择不敏感。

本文首先采用 Lloyd 聚类算法^[3]对图像进行高斯混合建模, 并以 Lloyd 聚类失真作为基本距离, 然后提出两种近似 EMD 的方法计算图像间的相似度, 最后根据图像间的相似度大小返回检索结果。

1 图像检索框架

图像检索首先要提取图像特征向量, 对图像进行建模, 然后度量图像间的相似度, 最后根据相似度大小返回检索结果。

1.1 图像建模

高斯混合模型具有良好的统计特性, 被广泛用于统

* 国家自然科学基金项目(60572078); 广东省自然科学基金项目(05006349);
武汉大学测绘遥感信息工程国家重点实验室基金项目(WKL(03)0101)

计模式识别、统计信号处理等领域。

高斯混合模型的概率密度函数为:

$$p(x) = \sum_{i=1}^L w_i g(x; u_i, \Sigma_i) \quad (1)$$

式中, x 是 k 维特征向量, L 是高斯混合成份个数, w_i 表示第 i 个高斯混合成份的权值且 $w_i=1$, 第 i 个高斯混合成份表示为:

$$g(x) = \frac{1}{(2\pi)^{k/2} |\Sigma_i|^{1/2}} e^{-\frac{1}{2}(x-u_i)^T \Sigma_i^{-1} (x-u_i)} \quad (2)$$

式中, u_i 、 Σ_i 分别是高斯混合成份的均值向量、协方差矩阵。

本文采用 Lloyd 聚类算法对图像进行高斯混合建模, 估计其参数。算法步骤如下:

(1) 初始化: 初始化高斯混合成份 $\{g_m, m=1, \dots, L\}$, 记迭代次数为 n 、初始失真为 D_0 和阈值为 T 。

(2) 寻找最小失真, 满足:

$$d(x_i, g_m^{(n)}) \leq d(x_i, g_j^{(n)}) \quad \forall j \quad m \quad (3)$$

计算总失真:

$$D_n = \sum_{m=1}^L \sum_{x_i \in g_m} d(x_i, g_m^{(n)}) \quad (4)$$

(3) 更新:

$$w_m^{(n)} = \frac{k_m}{N} \quad (5)$$

$$u_m^{(n)} = \frac{\sum_{x_i \in g_m} x_i}{k_m} \quad (6)$$

$$\Sigma_m^{(n)} = \frac{1}{k_m} \sum_{x_i \in g_m} (x_i - u_m^{(n)})(x_i - u_m^{(n)})^T \quad (7)$$

式中, k_m 是特征向量 x_i 聚类到混合成份 g_m 的个数, N 是特征向量总数。

(4) 如果 $|D_{n-1} - D_n|/D_{n-1} < T$, 则迭代停止, 否则返回第(2)步。

$d(x_i, g_m)$ 是特征向量 x_i 与高斯混合成份 g_m 间的距离, 采用参考文献[3]所用的平方误差失真 SED (Squared Error Distortion) 和量化错匹失真 QMD (Quantizer Mismatch Distortion) 度量:

$$d_{SED} = (x_i - u_m)^T (x_i - u_m) \quad (8)$$

$$d_{QMD} = \frac{1}{2} \{ \ln(2\pi) + \ln |\Sigma_m| + (x_i - u_m)^T \Sigma_m^{-1} (x_i - u_m) - \ln w_m \} \quad (9)$$

对图像进行 Lloyd 聚类后, 图库中的每一幅图像可以用高斯混合成份表示, 得到高斯混合成份参数。完成图像高斯混合建模后, 下一步是度量图像间的相似度。

1.2 EMD 相似度量

EMD 度量是 Rubner 等人提出的一种相似度量, 它把运筹学的运输问题引入到图像检索中, 采用最优化解最小运输成本的方法来度量图像间的相似性^[1]。

EMD 度量的数学模型描述^[4]: 设某产品有 m 个产地 A_1, \dots, A_m , 供应量分别为 w_{a1}, \dots, w_{am} ; n 个销地 B_1, \dots, B_n 的需求量分别为 w_{b1}, \dots, w_{bn} ; 产品从产地 A_i 运输到销地 B_j 的

单位运价为 d_{ij} , 求怎样分配从产地 A_i 到销地 B_j 的运输量 f_{ij} , 才能使总运输成本最小。图 1 是 $m=3, n=2$ 的 EMD 模型。

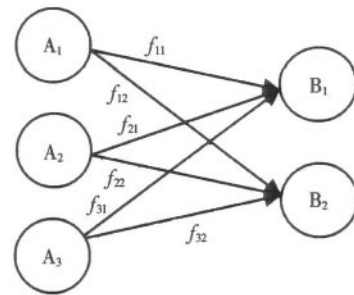


图 1 EMD($m=3, n=2$)模型

目标函数为:

$$\min \sum_{i=1}^m \sum_{j=1}^n f_{ij} d_{ij} \quad (10)$$

约束条件有:

$$f_{ij} \geq 0 \quad (i=1, \dots, m; j=1, \dots, n) \quad (11)$$

$$\sum_{j=1}^n f_{ij} \leq w_{ai} \quad (i=1, \dots, m) \quad (12)$$

$$\sum_{i=1}^m f_{ij} \leq w_{bj} \quad (j=1, \dots, n) \quad (13)$$

$$\sum_{i=1}^m \sum_{j=1}^n f_{ij} = \min \left(\sum_{i=1}^m w_{ai}, \sum_{j=1}^n w_{bj} \right) \quad (14)$$

定义 EMD:

$$EMD(A, B) = \frac{\sum_{i=1}^m \sum_{j=1}^n f_{ij} d_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}} \quad (15)$$

式(15)中的分母是规范化因子。

在图像检索中, 利用 EMD 计算图像间相似度时, d_{ij} 对应图像高斯混合成份间的距离(在参考文献[2]中称为基本距离), 可以通过 d_{SED} 或 d_{QMD} 来计算; w_{ai} 、 w_{bj} 对应图像高斯混合成份的权值。

2 近似 EMD 方法

EMD 方法的数学模型是一个线性规划问题, 参考文献[2]采用的是单纯形法求解, 其计算复杂度为 $O(n^3 \log n)$, 其中, n 是图像高斯混合成份个数。在图像检索中, w_{ai} 、 w_{bj} 分别对应高斯混合成份的权值, 公式(12)、公式(13)变为等式, 而且有:

$$\sum_{i=1}^m w_{ai} = \sum_{j=1}^n w_{bj} = 1 \quad (16)$$

则 EMD 方法简化为产销平衡问题, f_{ij} 有 $m \times n$ 个决策变量, $m+n$ 个约束条件, 而且满足公式(16), f_{ij} 系数矩阵的值小于等于 $m+n-1$ 。考虑到在图像检索中, 权值系数矩阵 f_{ij} 的特殊性, 可以通过表上作业法^[4]计算 f_{ij} 。本文采用最小元素法(MFM)和近似 EMD 的 Vogel 法, 这两种方法类似 Kruskal 最小生成树聚类算法^[5], 符合计算机视觉中的感知相似性。由最小生成树性质可知 f_{ij} 非零元素个数

《电子技术应用》2008 年第 3 期

为 $m+n-1$ 。

在图像检索中,表上作业法的产销平衡表和运价表如表 1 和表 2 所示,分别对应权值分配表和高斯混合成份间的距离表。下面详述这两种近似 EMD 方法。

表 1 产销平衡表^[4]

	销地 B_1	销地 B_2	...	销地 B_n	产量
产地 A_1	f_{11}	f_{12}	...	f_{1n}	w_{a1}
产地 A_2	f_{21}	f_{22}	...	f_{2n}	w_{a2}
...
产地 A_m	f_{m1}	f_{m2}	...	f_{mn}	w_{am}
销量	w_{b1}	w_{b2}	...	w_{bn}	

表 2 运价表^[4]

	销地 B_1	销地 B_2	...	销地 B_n
产地 A_1	d_{11}	d_{12}	...	d_{1n}
产地 A_2	d_{21}	d_{22}	...	d_{2n}
...
产地 A_m	d_{m1}	d_{m2}	...	d_{mn}

2.1 最小元素法 (MFM)

在产销平衡表中,尽量满足运价表中最小元素 d_{ij} 对应的 f_{ij} ,算法步骤如下:

- (1) 初始化产销平衡表, $f_{ij} = 0$ 。
- (2) 在运价表中找出最小元素 d_{ij} 。
- (3) 在产销平衡表中,找出 d_{ij} 对应的 f_{ij} , $f_{ij} = \min\{w_a, w_b\}$, 如果 $w_a > w_b$, 在运价表中划去 d_{ij} 所在的第 j 列, $w_a = (w_a - w_b)$; 否则在运价表中划去 d_{ij} 所在的第 i 行, $w_b = (w_b - w_a)$ 。

- (4) 返回第(2)步,直至运价表中所有元素被划去。

规范化 $m=n$, 第(3)步最差的情况是交叉地划去运价表中的行、列,划去行后查找最小元素 d_{ij} 循环 (i^2-i) 次,再划去列后查找最小元素 d_{ij} 循环 i^2 次,则算法最多的循环次数为:

$$\sum_{i=1}^n \{(i^2-i)+i^2\} = \sum_{i=1}^n (2i^2-i) = 2 \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2} = \frac{n(n+1)(4n-1)}{6} = O(n^3) \quad (17)$$

上述算法的计算复杂度为 $O(n^3)$ 。

2.2 Vogel 法

在产销平衡表中,尽量满足运价表中行(列)最小、次小元素差额最大的最小元素 d_{ij} 对应的 f_{ij} ,算法步骤如下:

- (1) 初始化产销平衡表, $f_{ij} = 0$ 。
- (2) 在运价表中,找出行(列)最小元素与次小元素之差最大所在的行(列),得该行(列)的最小元素 d_{ij} 。
- (3) 在产销平衡表中,找出 d_{ij} 对应的 f_{ij} , $f_{ij} = \min\{w_a, w_b\}$, 如果 $w_a > w_b$, 在运价表中划去 d_{ij} 所在的第 j 列, $w_a = (w_a - w_b)$; 否则在运价表中划去 d_{ij} 所在的第 i 行, $w_b = (w_b - w_a)$ 。
- (4) 返回第(2)步,直至运价表中所有元素被划去。

类似最小元素法,规范化 $m=n$, 第(3)步最差的情况是交叉地划去运价表中的行、列,划去行后查找最小、次小元素差额最大的最小元素 d_{ij} 循环 $[i+(i-1)+1](i-1)+[(i-1)+(i-2)+1]i=4(i^2-i)$ 次,再划去列后查找最小次小元素差额最大的最小元素 d_{ij} 循环 $[i+(i-1)+1]2i=4i^2$ 次,那么算法最多的循环次数为:

$$\sum_{i=1}^n \{4(i^2-i)+4i^2\} = \sum_{i=1}^n (8i^2-4i) = 8 \frac{n(n+1)(2n+1)}{6} - 4 \frac{n(n+1)}{2} = \frac{2n(n+1)(4n-1)}{3} = O(n^3) \quad (18)$$

上述算法的计算复杂度为 $O(n^3)$ 。

根据最小元素法和 Vogel 法计算 f_{ij} , 则图像 A、B 间的相似度定义为:

$$D(A, B) = \sum_{i=1}^m \sum_{j=1}^n f_{ij} d_{ij} \quad (19)$$

3 实验结果与分析

本文实验采用 Corel 图像库,从中选取非洲、海滩、建筑、汽车、恐龙、大象、花、马、雪山、食物共 10 类,每类 100 幅图像。将图像从 RGB 颜色空间转化到 CIE-Luv 颜色空间^[6],考虑到像素间的空间关系,把图像划分为不相交的 8×8 子块^[7],提取颜色和纹理特征^[8]。利用 Lloyd 聚类算法^[9]对图像特征向量进行高斯混合建模,以及利用 EMD、MFM、Vogel 三种方法度量图像间的相似性。检索效率采用查准率-查全率^[9]评价,查准率是返回的相关图像数与总的返回图像数的比例,查全率是返回图像数与图库总数的比例。三种方法的效率比较如图 2 所示,在两种基本距离下,MFM 法和 Vogel 法检索效率与 EMD 法接近。图 3、图 4、图 5 分别是以各自图中的第一幅图像作为例子以利用 EMD、MFM、Vogel 方法检索返回的前 20 幅图像。

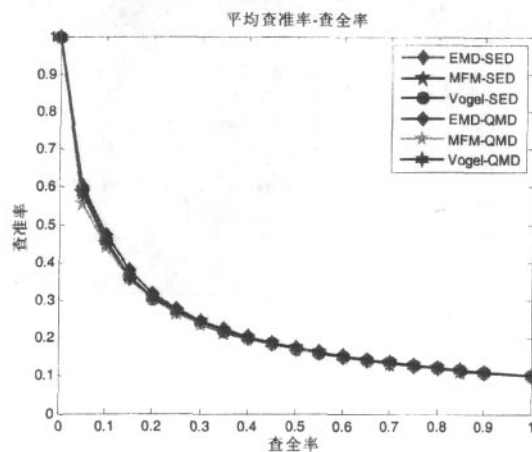


图 2 三种方法的检索效率比较

从图 2 可以看出,EMD-QMD 与 EMD-SED 检索效率接近。本文图像检索框架对基本距离的选择不敏感,而 L_1 (Manhattan 距离)与 L_2 (欧氏距离)在图像检索中的效率相似^[10],可以采用计算更为简单的 L_1 作为基本距离。当采用 SED 度量时,EMD、MFM、Vogel 方法实际上变

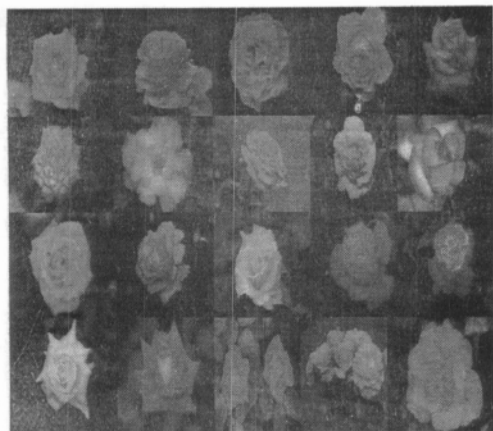


图 3 EMD 方法的检索结果

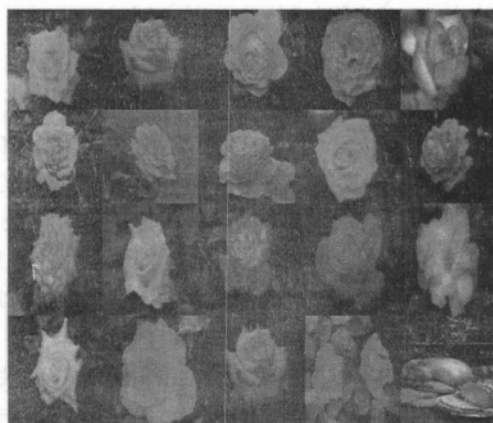


图 4 最小元素法的检索结果

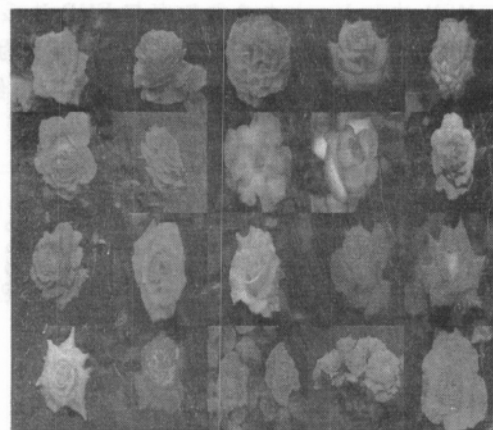


图 5 Vogel 法的检索结果

成了二次距离, 类似 Mahalanobis 距离, 不同的 Mahalanobis 距离的加权矩阵是其协方差矩阵^[10], 本文只是在加权时采用不同的策略。三种相似度量算法权值分配的策略分别是: EMD 是从整体高斯混合考虑, 使加权距离最小; MFM 考虑局部高斯混合成份间的距离最小, 使行(列)最小元素优先; Vogel 也是从局部高斯混合成份考虑, 只是采用的是行(列)最小与次小元素差额距离最大的最小元素优先, 而且 Vogel 更接近 EMD。

EMD 是一种有效的相似度量, 本文把原 EMD 模型简化为产销平衡问题, 提出两种权值分配方法近似 EMD 应用于图像检索时, 能达到与 EMD 接近的检索效率, 而且对基本距离的选择不敏感。最小元素法、Vogel 法在权值分配时, 采用最小元素优先, 即最相似优先, 比 EMD 更符合人的感知, 而且计算复杂度从原来的 $O(n^3 \log n)$ 降到 $O(n^3)$, 在一些实时计算要求较高的情况下, 最小元素法更能体现其优势。鉴于 EMD 在计算机视觉、模式识别、机器学习的广泛应用, 最小元素法、Vogel 法也可以应用于相关的领域, 如图像分类、识别、分割、聚类等。

参考文献

- [1] RUBNER Y, PUZICHA J, TOMASI C, et al. Empirical evaluation of dissimilarity measures for color and texture. *Computer Vision and Image Understanding*, 2001,(84): 25- 43.
- [2] AIYER A, PYUNB K, HUANG Y, et al. Lloyd clustering of gauss mixture models f- or image compression and classification. *Signal Processing: Image Communication*, 2005,(20):459- 485.
- [3] 孙麟平. 运筹学[M]. 北京:科学出版社,2005.
- [4] THEODORIDIS S, KOUTROUMBAS K. Pattern recognition. 2nd ed.[S. l.]:Academic Press, 2003.
- [5] WYSZECKI G, STILES W S. Color science: Concepts and methods, quantitative data and formulae. 2nd ed. Wiley, 2000.
- [6] JEONG S, WON C S, GRAY R M. Image retrieval using color histograms generated by gauss mixture vector quantization. *Computer Vision and Image Understanding*, 2004,(94):44- 66.
- [7] LIAPIS S, TZIRITAS G. Color and texture image retrieval using chromaticity histograms and wavelet frames. *IEEE Trans. Multimedia*, 2004,6:676- 686.
- [8] SMITH J R, CHANG S F. Tools and techniques for color image retrieval. In: *Proc. of SPIE: Storage and Retrieval for Image and Video Database*, 1996:426- 437.
- [9] ANDROUTSOS D, PLATANIOTIS K N, VENETSANOPOULOS A N. A novel vect- or based approach to color image retrieval using a vector angular- based distance measure. *Computer Vision and Image Understanding*, 1999, 75(1/2):46- 58.
- [10] ZHANG D, LU G. Evaluation of similarity measurement for image retrieval. *IEEE Int. Conf. Neural Networks and Signal Processing*, 2003,(2):928- 931.

(收稿日期: 2007- 08- 31)