

目标识别的相似性测量方法^①

李国祥, 夏国恩

(广西财经学院 教务处, 南宁 530003)

摘要: 相似性测量是图像后期特征匹配的关键步骤, 本文深入研究了特征向量顺序比对和交叉比对方法的差异, 并以形状上下文提取的图像二维直方图特征为基础, 分别结合 x^2 距离、推土机距离和扩散距离三种不同的相似性测量方法对多种图像标准库进行了识别验证。

关键词: 推土机距离; 扩散距离; 相似性测量; 形状上下文

Similarity Measurement Method of Target Recognition

LI Guo-Xiang, XIA Guo-En

(Department of Academic Affairs, Guangxi University of Finance and Economics, Nanning 530003, China)

Abstract: Similarity measurement is a key step in image matching, this paper studies the differences between feature vector sequence alignment and cross comparison method, and based on two-dimensional histogram features which are extracted by shape context, respectively combined with x^2 distance, earth mover's distance and the diffusion distance these three different similar measurements, it carries out the identification and verification of a variety of image standard library.

Key words: web information extraction; knowledge presentation; data intensive web pages; ontology-based keyword library

随图像的检索和匹配是机器视觉重要的组成部分, 由于细节特征能够有效的对图像进行精确分类, 越来越多的局部形状描述子得以提出进行图形的识别与匹配^[1-3]. 这些局部描述子针对图形局部的特点分别构建特征向量来对整个图形进行标注, 其中 shape context 利用对数极坐标直方图对特征点一定域内采样点物理分布情况进行特征描述, 在众多领域得到了广泛的应用^[4,5]. 同时众多学者也对其进行了改进, Ling, H. B.^[6] 对于 shape context 方法进行了优化改进, 利用内距离代替欧式距离计算样本点的上下文, 增强了对于图形复杂性的识别. Jun Xie^[7]利用图形中轴来分段匹配形量方法也相继提出. 目前对于高维特征向量的匹配主要分为顺序比对法和交叉比对法. 顺序相似度主要包括 L1 距离、L2 距离、 x^2 统计和 KL 分歧等, 交叉比对主要包括线性规划的 Earth Mover's Distance, 基于热

扩散原理 Diffusion Distance 和在此之上延伸而得的状上下文. 另外还有一些局部算子结合的方法^[8]. 但随之而来的高维细节特征向量, 又对配准的时间复杂度和配准率提出了新的挑战, 为此, 一系列的相似性测 Spectral Distance^[9]等. Julien Rabin 针对于部分二三维图像研究了不同相似性测量方法的有效性^[10]. 顺序比对运算简单、迅速, 但较少考虑直方图间空间关系. 交叉比对能够较好的反映直方图间空间位置关系, 但是时间复杂度较高. 因此后期特征的相似性测量方法的选择也成为衡量整个算法优劣的重要组成部分, 然而目前的研究, 仅仅停留在将交叉比对方法与局部描述子结合应用的层面, 并没有深入探讨不同方法的有效和适用范围. 因此如何有效的将相似性测量方法与模式识别算子进行结合, 将是一个亟待解决的问题.

基于以上问题, 本文分别将扩散距离和推土机距离与形状上下文结合, 深入探讨了以二者为代表的交

① 基金项目: 广西高校科学技术研究项目(kY15YB267); 广西高等学校高水平创新团队及卓越学者计划

收稿时间: 2015-12-09; 收到修改稿时间: 2016-01-14 [doi:10.15888/j.cnki.csa.005282]

又比对方法和经典的 x^2 距离比对的差异, 并通过大量实验验证了不同方法的执行效率和适用范围.

1 典型的形状上下文(Shape Context)

形状上下文(Shape Context)是一种通过描述特征点周围采样点物理分布的局部特征表示方法. 设目标采样点集合 $P = \{p_1, p_2, \dots, p_n\}$ 共 n 个采样点, 设带匹配特征点为 p_i , 以其为中心, 建立对数极坐标, 并根据所求特征细程度划分栅格大小, 构建直方图 h 表示剩余采样点物理分布特征.

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\}$$

其中, q 表示余下的采样点, 其中 K 为直方图栅格划分数目.

设待匹配特征点 p, q 分别属于不同的形状. 令 $C_{ij} = C(p_i, q_j)$ 表示两点的匹配代价, 在经典的形状上下文当中, 普遍使用顺序比对的 χ^2 距离测量方法:

$$C_{ij} = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) + h_j(k)]^2}{h_i(k) + h_j(k)}$$

其中 $h_i(k), h_j(k)$ 分别表示点 p, q 的 K 栅格归一化直方图分布.

由于在形状的检索中普遍使用轮廓点集来作为采样点集, 因而可以使用轮廓点集来约束匹配序列 π , 进而使用动态匹配算法替代匈牙利算法, 提高执行效率和匹配精度. 对于图像的检索过程实际上转化为待检索图像采样点集匹配代价的最小化过程:

$$H(\pi) = \sum_i C(p_i, q_{\pi(i)})$$

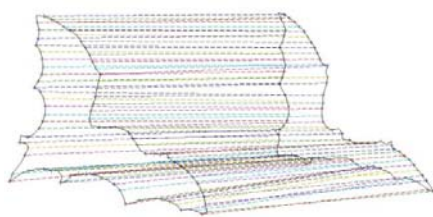


图 1 基于形状上下文的两幅图像匹配

2 推土机距离(Earth Mover's Distance)

Ruber^[11]将图像检索的特征向量比对问题转化为线性规划的运输问题. 设两组特征向量集合分别为: $P = \{(p_1, \omega_{p1}), \dots, (p_m, \omega_{pm})\}$, $Q = \{(q_1, \omega_{q1}), \dots, (q_n, \omega_{qn})\}$, 分别包含 m 和 n 个聚类中心, 其中 ω 表示其聚类权重. $D = [d_{ij}]$ 表示特征向量间的测度矩阵, d_{ij} 表示向量聚类 p_i 和 q_j 中心间的测度距离, 其距离的定义根据所求问题而具体化.

通过找到最佳运输方案 $F = [f_{ij}]$, f_{ij} 表示供给 i 到需求 j 的运输量, 从而使运输费用最小.

$$E(H, K) = \min_{F=\{f_{ij}\}} \frac{\sum_{i=1}^N \sum_{j=1}^N d_{ij} f_{ij}}{\sum_{i=1}^N \sum_{j=1}^N f_{ij}}$$

满足以下约束条件

$$\begin{aligned} \sum_{j=1}^N f_{ij} &\leq p_i, \quad 1 \leq i \leq N \\ \sum_{i=1}^N f_{ij} &\leq q_j, \quad 1 \leq j \leq N \\ f_{ij} &\geq 0, \quad 1 \leq i \leq N, 1 \leq j \leq N \\ \sum_{i=1}^N \sum_{j=1}^N f_{ij} &= \min(\sum_{i=1}^N p_i, \sum_{j=1}^N q_j) \end{aligned}$$

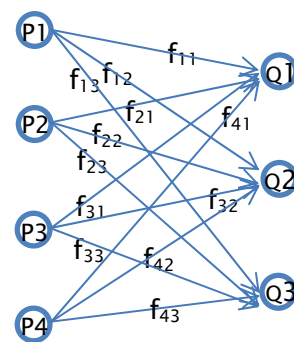


图 2 经 EMD 转化后的运输问题模型

针对于二维的直方图而言, 设二维直方图 $H: m \times n$ 个栅格, 栅格坐标集合定义为, $I = \{(i, j), 1 \leq i \leq m, 1 \leq j \leq n\}$, 运输流方向表示为 $J = \{(i, j, k, l) : (i, j) \in I, (k, l) \in I\}$, 即从栅格 (i, j) 运输至 (k, l) , 待匹配直方图矩阵 $P = \{p_{ij} : (i, j) \in I\}, Q = \{q_{ij} : (i, j) \in I\}$, 对 P, Q 归一化: $p_{ij} = \frac{P_{ij}}{P}, q_{ij} = \frac{Q_{ij}}{Q}$ 可得约束条件 $\sum_{i,j} p_{ij} = 1, \sum_{i,j} q_{ij} = 1$

从而直方图矩阵的匹配问题转化为线性规划的运输问题:

$$\begin{aligned} E(H, K) &= \min_{F=\{f_{ij}\}} \sum_{i,j,k,l} f_{i,j,k,l} d_{i,j,k,l} \\ s.t. &\begin{cases} \sum_{(k,l) \in I} f_{ijkl} = p_{ij} & \forall (i, j) \in I \\ \sum_{(i,j) \in I} f_{ijkl} = q_{kl} & \forall (k, l) \in I \\ f_{i,j,k,l} \geq 0 & \forall (i, j, k, l) \in J \end{cases} \end{aligned}$$

一般而言距离 d_{ij} 可以定义为任何距离, 通常根据所测量对象进行选择. 在对彩色图像或者纹理图像进行对比时, 一般定义 $d_{ij} = 1 - e^{-\alpha \|p_i - q_j\|}$, $\alpha = \left\| [\delta_1 \dots \delta_{\text{dim}}]^T \right\|$, 其中是 i th 维数特征部分的标准差, dim 是特征向量维度. 在对直方图特征进行对比时, 一般采用为 L_p 距离

$$d_{ij} = (|i - k|^p + |j - l|^p)^{1/p}$$

令 $p = 1$, LING H B^[12] 将直方图中两点间的测度距离 L_1 分解为相邻点的累积和, 定义集合 $J_s = \{(i, j, k, l) : (i, j, k, l) \in I, d_{i,j,k,l} = 1\}$, 通过将长距离的两点测度转化为相邻点的距离之和, 即 $d_{i,j,k,l} = d_{i,j,i,l} + d_{i,j,k,l} = \sum_{j \leq x < l} d_{i,x,i,x+1} + \sum_{i \leq y < k} d_{y,l,i,y+1}$, 进一步减少变量数目和约束条件, 降低运算复杂度, 最终得优化后的 EMD 模型:

$$E(H, K) = \min_{G = \{g_{i,j,k,l} : (i,j,k,l) \in J_s\}} \sum_{J_s} g_{i,j,k,l}$$

$$s.t \begin{cases} \sum_{(k,l) \in J_s} g_{ijkl} - g_{klji} = b_{ij} & \forall (i, j) \in I \\ g_{i,j,k,l} \geq 0 & \forall (i, j, k, l) \in J_s \end{cases}$$

3 扩散距离^[13] (Diffusion Distance)

对于 m 维直方图分布 $h_1(x), h_2(x)$, 其中 $x \in R^m$, 设存在温度场 $T(x, t)$, 其温度变化符合热扩散方程:

$$\frac{\partial T}{\partial t} = \nabla^2 T \quad \text{其具有唯一解:}$$

$$T(x, t) = T_0(x) * \phi(x, t)$$

其中 $\phi(x, t)$ 为高斯滤波器

$$\phi(x, t) = \frac{1}{(2\pi)^{m/2} t} \exp\left\{-\frac{X^T X}{2t}\right\}$$

因为扩散过程中差异性区域的均值为 0, 因此 $T(x, t)$ 可以看做是随着时间维度的增加而在直方图特征 $h_1(x), h_2(x)$ 中扩散差异的过程, 即特征向量 $h_1(x), h_2(x)$ 的均衡化过程. 对整个扩散过程的差异性测量即为扩散距离:

$$\hat{K}(h_1, h_2) = \int_0^t k(|T(x, t)|) dt$$

实际运算中特征向量都是离散变量, 需要将连续过程离散化, 由于高斯金字塔中图像的平滑子采样不会产生混淆现象, 因此使用高斯金字塔对连续扩散过程离散化. 可得:

$$K(h_1, h_2) = \sum_{l=0}^L k(|d_l(x)|)$$

其中

$d_0 = h_1(x) - h_2(x), d_l(x) = [d_{l-1}(x) * \phi(x, \delta)] \downarrow_2$ 其中 $l = 1, \dots, L$ 和 \downarrow_2 分别表示高斯金字塔的下采样层数和大小, k 为测量距离, 类似于 EMD 模型, 这里同样选择 $L1$ 距离, 进而将上式简化为: $K(h_1, h_2) = \sum_{l=0}^L |d_l(x)|$

各相似性测量方法的时间复杂度对比如下:

表 1 不同相似性测量方法的时间复杂度对比

名称	χ^2	EMD	DD
时间复杂度	$O(n)$	$O(n^2)$	$O(n)$

4 实验结果

为了验证不同相似性测量方法的有效性, 本文针对于不同图像特点, 分别选择 Tool shapes^[14], Gopalan 数据库^[15], Kimia dataset 2^[16]和 MPEG-7 作为实验数据库. 本文统一使用 shape context 局部描述算子. 设置形状上下文特征中栅格参数为, 采样点数目 $n = 100$. 采用 Bullseye score^[6]进行识别率度量, 即对于每一幅图像与其他图像进行交叉验证, 选取该类所含幅图像的两倍数量作为候选图像, 最终的识别率以候选图像中正确的匹配与所有可能的匹配之比来衡量. 同时按照匹配度排序选取与待检索图像最为相似的前若干幅图像, 分别以 1st, 2nd...等来进行标注, 来进行检索精度的度量.

4.1 Tool shapes 数据库

该数据库包含 35 幅图像共 7 类工具图像, 如图所示. 同类图形中的不同图像是由其弯曲形变而来, 不同图像间具有较高相似性.

表 2 Tool shapes 数据库的检索结果

方法名称	1st	2nd	3th	4th	Bulleyescore
SC+ χ^2	24	10	10	14	73.1%
SC+DD	28	24	20	10	85.7%
SC+EMD	31	24	25	17	85.7%



图 3 Tool shapes 数据库

通过实验可以看出, 相比于顺序特征比对, 交叉比对能够更为有效的区分细节特征. 推土机距离最优化的匹配方式效果最好.

4.2 Kimia dataset 2 数据库

该数据库包含 99 幅图像共 9 类, 如图 4 所示. 这里我们同时对不同相似性测量方法的时间复杂度进行

了验证. 实验结果见表 3.



图 4 Kimia 数据库

表 3 Kimia dataset 2 的检索结果

识别方法	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Time/s
SC+ χ^2	97	91	88	85	84	77	75	66	56	37	21.2
SC+DD	95	92	90	89	88	80	73	60	59	37	42.8
SC+EMD	96	93	93	90	90	82	85	76	65	53	90.9

表 4 Gopalan 数据库的检索结果

识别方法	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	Score
SC+ χ^2	40	34	28	29	24	25	20	30	23	82.2%
SC+DD	30	33	23	24	12	20	13	18	18	66.4%
SC+EMD	35	32	19	21	13	17	16	10	16	65.2%

4.3 Gopalan 数据库

该数据库包含 5 类共 50 幅图像, 其中大部分图像包含分割不完整的空洞部分, 给图像的分类带来更大的难度. 部分图像见图 5, 实验结果结果见表 4.



图 5 部分 Gopalan 数据库图像

4.4 Mpeg 7 数据库

Mpeg 7 数据库包含 70 类, 每类 20 幅共 1400 幅图像. 部分图像见图 6, 实验结果见图 7



图 6 部分 MPEG7 数据库图像

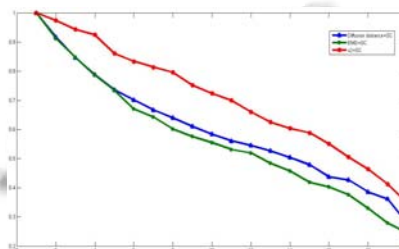


图 7 Mpeg 7 数据库的检索结果

5 结论

通过实验可以看出, χ^2 适合常规图像的区分, 而 EMD, DD 距离更容易区分相似性程度更高的图像. 在 Tool shapes 数据库中, 包含了大量高度相似的弯曲变形图像, EMD 能够较好的区分不同特征, DD 次之. 在 kimia 的 99 幅图像的检索中, EMD 和 DD 略优于 χ^2 距离. 但是在时间复杂度方面, 由于 EMD 实际上是将特征向量的比对转化为线性规划的运输问题, 因此其计算时长最大为 90.9s, DD 和 χ^2 分别为 42.8s 和 21.2s. 在对象包含一定干扰的时候, 如 Gopalan 图像中这些

分割不完全的空洞,影响了EMD, DD特征的交叉比对的优化策略,抗噪性相反不如 χ^2 的顺序比对,其检索率也不如 χ^2 距离.在最后的Mpeg7的大样本测试中, χ^2 的检索曲线下降平滑,检索结果仍然优于EMD和DD, Bulleye score 分别为78.8%, 66.3%, 64.3%.

由此可见,虽然以EMD和DD为代表的交叉比对方法能够更好的兼顾不同直方图特征空间位置特性,但是并非在所有情况会优于传统的 χ^2 距离.对于相似性较高的图像区分,交叉比对识别效果较好,但是由于其在兼顾不同特征向量空间位置的过程中,容易错误匹配不同空间噪声而产生误识别,而顺序比对恰好能够规避这一点,所以其鲁棒性相比于顺序方法较差.并且对于视觉上有较大差距的图像, χ^2 却不失为一种计算简单,耗时少,效果较好的方法.

因此在相似性测量方法的选择上,应该以简单实用为原则,摆脱交叉比对优于顺序比对的传统观念,针对于待检索图像的性质和实时性进行综合考虑,对于数据量大或实时性要求较高的数据集应优先考虑顺序法,相似性程度较高的数据集则以交叉比对为主.同时顺序比对方方法的鲁棒性优于交叉比对法,在对大规模复杂的图像检索时,在充分考虑不同局部描述子特点的情况下,可以充分结合二者的优势,利用顺序方法进行粗比对,进而结合交叉方法区分相似性较高的图像,从而提高整体检索效率和细节区分度.

参考文献

- 1 Netz A, Osadchy M. Recognition using specular highlights. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2013, 35(3): 639–652.
- 2 Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509–522.
- 3 Marszalek M, Schmid C. Accurate object recognition with shape masks. International Journal of Computer Vision, 2012, 97(2): 191–209.
- 4 Xudong L, Wang ZJ. Perceptual image hashing based on shape contexts and local feature points. IEEE Trans. on Information Forensics and Security, 2012, 7(3): 1081–1093.
- 5 Premachandran V, Kakarala R. Perceptually motivated shape context which uses shape interiors. Pattern Recognition, 2013, 46(8): 2092–2102.
- 6 Ling HB, Jacobs DW. Shape classification using the inner-distance. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(2): 286–299.
- 7 Xie J, Heng PA, Shah M. Shape matching and modeling using skeletal context. Pattern Recognition, 2008, 41(5): 1756–1767.
- 8 Gui Y, Su A, Du J. Point-pattern matching method using SURF and Shape Context. Optik, 2013, 124(14): 1869–1873.
- 9 Bronstein MM, Bronstein AM. Shape recognition with spectral distances. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2011, 33(5): 1065–1071.
- 10 Rabin J, Peyre G, Cohen LD. Geodesic shape retrieval via optimal mass transport. In: Daniilidis K, Maragos P, Paragios N, eds. Computer Vision-Eccv 2010, Pt V. 2010: 771–84.
- 11 Rubner Y, Tomasi C, Guibas LJ. The earth mover's distance as a metric for image retrieval. International Journal of Computer Vision, 2000, 40(2): 99–121.
- 12 Ling H, Okada K. An efficient earth mover's distance algorithm for robust histogram comparison. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(5): 840–53.
- 13 Haibin L, Okada K. Diffusion distance for histogram comparison. Proc. of the Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, F 17–22 June, 2006.
- 14 Bronstein AM, Bronstein MM, Bruckstein AM, et al. Analysis of two-dimensional non-rigid shapes. International Journal of Computer Vision, 2007, 78(1): 67–88.
- 15 OF AIR, Shapes NP. Articulation-invariant representation of non-planar shapes. Europ Conf Computer Vision, 2010, 3: 286–299.
- 16 Sebastian TB, Klein PN, Kimia BB. Recognition of shapes by editing their shock graphs. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004, 26(5): 550–571.