一种基于广义期望首达时间的形状距离学习算法

郑丹晨 [1] 杨亚飞 [1] 韩 敏 [1]

摘 要 形状距离学习是形状匹配框架中引入的后处理步骤, 能够有效改善逐对计算得到的形状间距离. 利用期望首达时间分析形状间相似度可能导致距离更新不准确, 针对这一问题提出了一种基于广义期望首达时间 (Generalized mean first-passage time, GMFPT) 的形状距离学习方法. 将形状样本集合视作状态空间, 广义期望首达时间表示质点由一个状态转移至指定状态集合所需的平均时间步长, 本文将其视作更新后的形状间距离. 通过引入广义期望首达时间, 形状距离学习方法能够有效地分析上下文相关的形状相似度, 显式地挖掘样本空间流形中的最短路径, 并消除冗余上下文形状信息的影响. 将所提出的方法应用到不同形状数据集中进行仿真实验, 本文方法比其他方法能够得到更准确的形状检索结果.

关键词 形状匹配, 形状距离学习, 离散时间马尔科夫链, 期望首达时间, 广义期望首达时间

引用格式 郑丹晨, 杨亚飞, 韩敏. 一种基于广义期望首达时间的形状距离学习算法. 自动化学报, 2016, **42**(2): 246−254

DOI 10.16383/j.aas.2016.c150105

A Shape Distance Learning Algorithm Based on Generalized Mean First-passage Time

ZHENG Dan-Chen[1] YANG Ya-Fei[1] HAN Min[1]

Abstract With the help of shape distance learning introduced into shape matching framework as a post-processing procedure, shape distances obtained by pairwise shape similarity analysis can be improved effectively. A novel shape distance learning method based on generalized mean first-passage time (GMFPT) is proposed to solve the problem of inaccurate matching results caused by mean first-passage time. Given a set of shapes as the state space, the generalized mean first-passage time, which is regarded as the updated shape distance, is used to represent the average time step from one state to a certain set of states. With the generalized mean first-passage time introduced into the distance learning algorithms, context-sensitive similarities can be evaluated effectively, and the shortest paths on the distance manifold can be explicitly captured without redundant context. Simulation experiments are carried out on different shape datasets with the proposed method, and the results demonstrate that the retrieval score can be improved significantly.

Key words Shape matching, shape distance learning, discrete-time Markov chain, mean first-passage time, generalized mean first-passage time

Citation Zheng Dan-Chen, Yang Ya-Fei, Han Min. A shape distance learning algorithm based on generalized mean first-passage time. *Acta Automatica Sinica*, 2016, **42**(2): 246−254

形状匹配一直都是计算机视觉领域中的重要研究问题, 其经过几十年发展已经得到了很多成熟有效的方法, 在目标识别、图像检索、医学成像分析等众多领域都得到了广泛的应用[1−2]. 形状匹配结果一般表示为形状距离, 一对相似形状对应的形状距离通常较小, 反之两个不相似形状对应的距离往往较大[3].

收稿日期 2015-03-02 录用日期 2015-10-28
Manuscript received March 2, 2015; accepted October 28, 2015
国家自然科学基金 (61374154), 中央高校基本科研业务费专项资金 (DUT14RC(3)128) 资助
Supported by National Natural Science Foundation of China (61374154) and Fundamental Research Funds for the Central Universities (DUT14RC(3)128)
本文责任编委 贾云得
Recommended by Associate Editor JIA Yun-De
1. 大连理工大学电子信息与电气工程学部 大连 116023
1. Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116023

逐对形状匹配方法由形状特征提取和形状特征匹配两个步骤组成, 其专注于分析两个待匹配形状间的关系来计算距离[4]. 但是很多情况下, 仅考虑一对形状间相似程度往往难以分析得到准确的匹配结果[5−6]. 图 1 给出了一组利用逐对形状匹配方法可能产生错误匹配结果的示例 (A 为狗, B 和 C 为马). 由于 A 与 B 的姿态比较接近, 逐对形状匹配方法通常分析得到 AB 间距离小于 BC 间距离, 与实际分类结果不符. 当相同类别形状差异较大、不同类别形状差异较小时, 通过分析数据集中蕴含的流形信息可以较为有效地改善匹配结果. 鉴于此, 近些年很多研究都在逐对形状匹配方法后引入距离度量学习作为后处理步骤, 经过该步骤后可以得到新的 "相似度"、"距离" 或 "排序" 结果, 从而改善逐对形状匹配方法得到的距离. 改进后得到形状匹配方法框架如图 2 所示.

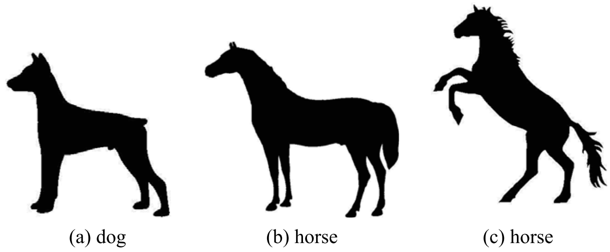


图 1　逐对形状匹配方法可能导致错误结果的示例

Fig. 1　An example of misunderstanding of objects caused by pairwise shape matching methods

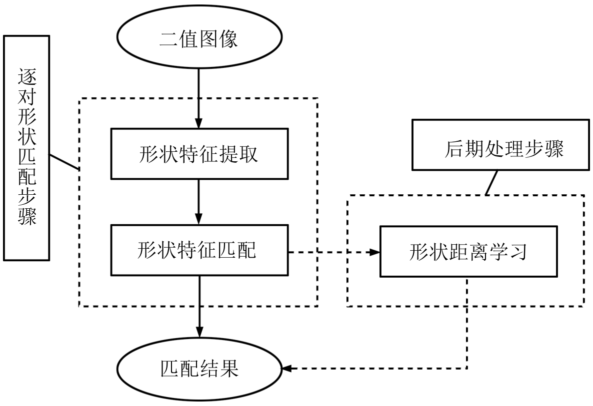


图 2　形状匹配方法框架

Fig. 2　The framework for shape matching methods

在形状距离学习的研究中, Bai 等[5,7] 首先引入半监督学习中的标记传播算法 (Label propagation, LP) 对逐对形状匹配的结果进行改进. 在样本空间中, 形状标记可以沿测地线路径方向快速传播, 因此标记传播方法能够隐式地分析样本空间中的流形信息. Yang 等[6] 提出了基于局部约束扩散过程 (Locally constrained diffusion process, LCDP) 的形状距离学习算法. 该方法首先对原始形状集合进行扩展, 在此基础上建立具有约束条件的扩散图, 进而选择扩散距离作为距离更新结果. 文献 [8−9] 也应用该算法对不同的形状匹配结果进行距离学习, 取得了较好结果. Egozi 等[10] 也提出了一种基于元描述符 (Meta descriptor, MD) 的形状距离学习方法. 元描述符是以相似度空间中 K 近邻权重生成的矢量特征, 利用 L_1 范数计算可得更新后的形状间距离. 文献 [11] 提出在张量积图 (Tensor product graph, TPG) 的基础上分析扩散过程, 由于引入了更高阶的信息, 因此得到了更可靠的目标间相似度. Donoser 等[12] 对距离学习方法进行了总结回顾, 并在扩散过程框架下提出了更加有效的距离学习方法. 文献 [13] 引入期望首达时间 (Mean first-passage time, MFPT) 对逐对形状匹配结果进行分析, 以完成状态转移所需的平均时间步长作为形状距离学习结果, 能够较为有效地挖掘空间流形结构信息.

上述距离学习方法都尝试隐式地挖掘样本空间拓扑结构, 虽然能较好地分析局部空间信息, 但很难准确描述相距较远的同类样本间相似程度. 当样本空间中一对同类别目标距离较远时, 空间流形对应的检索路径上通常会包含许多样本点. 这可能导致标记信息在传播过程中衰减较快, 难以有效地影响到距离较远的同类样本. 对于期望首达时间而言, 这一情形可能导致完成状态转移所需的平均时间步长较大, 不能得到理想的形状距离更新结果.

针对此类问题, Wang 等[14] 提出了一种最短路径传播算法 (Shortest path propagation, SPP), 通过显式地挖掘样本空间中最优路径来实现距离学习, 但是其无法克服标记传播算法存在的不平衡性问题. 为了更好地分析形状距离学习问题, 提出一种基于广义期望首达时间 (Generalized mean first-passage time, GMFPT) 的形状距离学习算法. 本文在期望首达时间概念基础上引申定义了广义期望首达时间, 其反映了质点从状态空间中一个状态转移至某个状态集合所对应的平均时间步长. 已知逐对形状匹配方法得到的距离矩阵, 通过构造稀疏连接图以去除不相关样本间的连接, 可以一定程度上避免质点在非同类样本间的状态转移, 从而消除不相关样本在距离学习过程中的干扰. 在进行形状距离学习的过程中, 考虑利用查询样本及距离较近的样本构造样本子集合, 利用广义期望首达时间来表示更新后的形状间距离, 通过多次迭代显式地挖掘得到最优路径. 本文分别选择了不同的形状数据集进行仿真实验, 对比其他文献中形状距离学习方法, 所提方法能够得到更好的形状检索结果.

1　基于期望首达时间的形状距离学习

本节首先介绍期望首达时间的相关概念, 进一步对文献 [13] 中提出的形状距离学习算法进行回顾. 在逐对形状匹配的基础上, 该方法借助期望首达时间来分析挖掘样本空间流形结构, 对形状匹配结果进行更新.

对于形状样本集合 $\{s_1, s_2, \cdots, s_N\}$, 可以构造状态空间 $Space = \{1, 2, \cdots, N\}$, 样本 s_k 对应元素 $k \in Space$. 对于任意的元素 $i, j \in Space$, (一步) 状态转移概率满足 $p_{ij} > 0$, 状态空间对应的状态转移矩阵为 $P = [p_{ij}]$. 随机序列 $\{X_n : n = 0, 1, 2, \cdots\}$ 是时齐的离散时间马尔科夫链

$$P\left(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \cdots, X_0 = i_0\right) = P\left(X_{n+1} = j | X_n = i\right) = P\left(X_1 = j | X_0 = i\right) = p_{ij} \tag{1}$$

其中, X_n 在状态空间 $Space$ 中取值.

在状态空间中, 期望首达时间表示质点由一个状态转移至另一个状态所需的平均转移次数. 由状态 i 至状态 j 的期望首达时间定义为 μ_{ij}

$$\mu_{ij} = \mathrm{E}\left(T_j|X_0=i\right) = \sum_{n=1}^{\infty} nP\left(T_j=n|X_0=i\right) \tag{2}$$

其中, $T_j = \min\{n>0: X_n=j\}$ 表示质点首次到达状态 j 所对应的转移次数. 对于集合中同类形状样本而言, 质点可以沿检索路径快速完成状态转移.

标记传播算法固定查询样本标记, 将标记的传播过程作为研究重点. 对比而言, 以期望首达时间更新形状距离则重点关注质点完成状态转移所需的时间步长, 可以避免标记传播算法的不平衡性问题. 其使同类形状样本的分布更加紧凑, 进而更好地挖掘局部空间的几何信息[12].

假设 s_N 和 s_k 分别为查询样本和目标样本, $1 \le k \le K$, $K=N-1$. 以期望首达时间 μ_{Nk} 作为二者间更新后的距离, 可以得到如下方程形式

$$(I-A)\boldsymbol{\mu}_i = \mathbf{1}_K \tag{3}$$

其中, $\boldsymbol{\mu}_i = [\mu_{1i}, \cdots, \mu_{Ki}]^{\mathrm{T}}$, $A=[p_{ij}]$, $1\le i\le K$, $1\le j\le K$, $\mathbf{1}_K=[1,\cdots,1]^{\mathrm{T}}$ 包含了 $K=N-1$ 个元素. 由于 A 中元素满足 $p_{ij}>0$, $I-A$ 中有 $1-p_{ii} > \sum_{j\neq i}|-p_{ij}| = 1-p_{ii}-p_{iN}$. 因此, $I-A$ 是严格对角占优方阵, 式 (3) 对应的线性方程可以直接进行求解, 通过对 $\boldsymbol{\mu}_i$ 中各元素进行排序 (Ranking) 即可得到形状检索结果.

2 广义期望首达时间

为了能够分析状态空间中更加复杂的状态转移问题, 本节在期望首达时间的基础上提出了更具一般意义的广义期望首达时间, 为后续形状距离学习提供相关的理论支持.

已知状态空间 $Space$ 下的时齐离散时间马尔科夫链 $\{X_n\}$, 对于状态 $i\in Space$ 和集合 $L\subset Space$, 引入条件概率

$$f_{iL}^{(n)} = P_i\left(X_n\in L, X_m\notin L, 1\le m\le n-1\right) \tag{4}$$

其中, $P_i(\cdot) = P(\cdot|X_0=i)$. 在式 (4) 中, f_{iL} 表示质点从状态 i 出发, 第 n 步到达集合 L 的概率. 这里进一步给出广义期望首达时间的定义.

定义 1. 定义由 i 转移至 L 的平均状态转移次数为广义期望首达时间, 记作 η_{iL}

$$\eta_{iL} = \mathrm{E}\left(T_L|X_0=i\right) = \sum_{n=1}^{\infty} nP_i\left(T_L=n\right) = \sum_{n=1}^{\infty} nf_{iL}^{(n)} \tag{5}$$

其中, $T_L=\min\{n>0: X_n\in L\}$ 表示质点首次到达集合 L 的转移次数. 比较式 (2) 中期望首达时间和式 (5) 中广义期望首达时间的定义可以看出, 如果集合 L 中仅包含元素 $j\in Space$, 则 $\eta_{iL}=\mu_{ij}$, 即期望首达时间是广义期望首达时间对应的一种特殊情况.

从式 (4) 可知, f_{iL} 满足如下关系

$$\begin{aligned} f_{iL}^{(n)} &= P_i\left(X_n\in L, X_m\notin L, 1\le m\le n-1\right) = \\ &\sum_{j\notin L} p_{ij} P\left(X_n\in L, X_m\notin L, 2\le m\le n-1|X_1=j\right) = \\ &\sum_{j\notin L} p_{ij} P_j\left(X_{n-1}\in L, X_m\notin L, 1\le m\le n-2\right) = \\ &\sum_{j\notin L} p_{ij} f_{jL}^{(n-1)} \end{aligned} \tag{6}$$

结合式 (5) 中广义期望首达时间的定义, 可得如下所示的表达式

$$\begin{aligned} \eta_{iL} = &\sum_{n=1}^{\infty} nf_{iL}^{(n)} = \sum_{n=2}^{\infty}(n-1)f_{iL}^{(n)} + \sum_{n=1}^{\infty} f_{iL}^{(n)} = \\ &\sum_{j\notin L} p_{ij}\sum_{n=2}^{\infty}(n-1)f_{jL}^{(n-1)} + \sum_{n=1}^{\infty} f_{iL}^{(n)} = \\ &\sum_{j\notin L} p_{ij}\eta_{jL} + \sum_{n=1}^{\infty} f_{iL}^{(n)} \end{aligned} \tag{7}$$

为方便讨论, 将状态空间划分为两个集合 $U=\{1,\cdots,M\}$ 和 $L=\{M+1,\cdots,N\}$, 二者分别包含 M 个元素和 $N-M$ 个元素, $Space=U\cup L$. 状态转移矩阵 P 可以写成如下式所示的分块矩阵

$$P = \begin{bmatrix} P_{UU} & P_{UL} \\ P_{LU} & P_{LL} \end{bmatrix} \tag{8}$$

其中, $P_{UU}=[p_{ij}]$, $P_{UL}=[p_{ik}]$, $P_{LU}=[p_{lj}]$, $P_{LL}=[p_{lk}]$, $1\le i\le M$, $1\le j\le M$, $M+1\le k\le N$, $M+1\le l\le N$. 进一步可知 $f_{iL}^{(n)}=\sum_{j=1}^{M} p_{ij} f_{jL}^{(n-1)}$, $f_{iL}^{(1)}=\sum_{j=M+1}^{N} p_{ij}$, 满足如下关系

$$\begin{bmatrix} f_{1L}^{(n)} \\ \vdots \\ f_{ML}^{(n)} \end{bmatrix} = P_{UU}\times\begin{bmatrix} f_{1L}^{(n-1)} \\ \vdots \\ f_{ML}^{(n-1)} \end{bmatrix} \tag{9}$$

$$\begin{bmatrix} f_{1L}^{(1)} \\ \vdots \\ f_{ML}^{(1)} \end{bmatrix} = P_{UL}\times\mathbf{1}_{N-M} \tag{10}$$

通过式 (9) 和式 (10) 可得

$$\begin{bmatrix} \sum\limits_{n=1}^{\infty} f_{1L}^{(n)} \\ \vdots \\ \sum\limits_{n=1}^{\infty} f_{ML}^{(n)} \end{bmatrix} = \left(I + P_{UU} + P_{UU}^2 + \cdots\right) \times P_{UL} \times \mathbf{1}_{N-M} = \sum_{k=0}^{\infty} (P_{UU})^k \times P_{UL} \times \mathbf{1}_{N-M} \tag{11}$$

在状态空间划分下, 式 (7) 可写作 $\eta_{iL} = \sum_{j=1}^{M} p_{ij}\eta_{jL} + \sum_{n=1}^{\infty} f_{iL}^{(n)}$, 结合式 (11) 可得

$$\boldsymbol{\eta}_L = P_{UU} \times \boldsymbol{\eta}_L + \begin{bmatrix} \sum\limits_{n=1}^{\infty} f_{1L}^{(n)} \\ \vdots \\ \sum\limits_{n=1}^{\infty} f_{ML}^{(n)} \end{bmatrix} \tag{12}$$

$$(I - P_{UU}) \times \boldsymbol{\eta}_L = \sum_{k=0}^{\infty} (P_{UU})^k \times P_{UL} \times \mathbf{1}_{N-M} \tag{13}$$

其中, $\boldsymbol{\eta}_L = [\eta_{1L}, \cdots, \eta_{ML}]^{\mathrm{T}}$. 从式 (13) 可以看出, 如果矩阵 $I - P_{UU}$ 可逆, 直接求解线性方程即可得到广义期望首达时间.

3　基于广义期望首达时间的形状间距离分析

已知逐对形状匹配方法所得到的形状间距离, 本节将首先通过形状间距离矩阵构造状态转移矩阵, 进一步借助广义期望首达时间对形状间距离进行学习和更新.

3.1　构造状态转移矩阵

已知形状样本集合 $\{s_1, \cdots, s_N\}$, d_{ij} 是利用逐对形状匹配方法得到的 s_i 和 s_j 间距离, 距离矩阵为 $D = [d_{ij}]$. 相似度 r_{ij} 可由 d_{ij} 计算得到, 将其对应表达式定义为

$$r_{ij} = \exp\left(-\frac{d_{ij}^2}{\sigma_{ij}^2}\right) \tag{14}$$

其中, σ_{ij} 对应高斯核函数的核宽. 利用 $kd\,(i)$ 表示样本 s_i 的 K_d 个近邻形状样本集合. 参考文献 [5, 10] 中的核宽设置方法, σ_{ij} 可以表示为

$$\sigma_{ij} = \alpha \frac{1}{2K_d} \left(\sum_{s_k \in kd(i)} d_{ik} + \sum_{s_l \in kd(j)} d_{jl} \right) \tag{15}$$

其中, 参数 K_d 由数据集中形状样本数目决定, 样本数目越多 K_d 的取值越大, 即能够得到更合适的距离均值. 参数 α 的取值通常介于 0.2～0.4 之间, 往往需要结合逐对形状匹配方法的选择来确定.

在得到样本间相似度的基础上, 建立稀疏连接的无向图 G. 以 $ks\,(i)$ 表示 s_i 的 K_s 个近邻样本, 各顶点连接权重 w_{ij} 定义为

$$w_{ij} = \begin{cases} r_{ij}, & s_i \in ks\,(j),\ s_j \in ks\,(i) \\ 0, & \text{其他} \end{cases} \tag{16}$$

其中, 参数 K_s 与数据集中样本分布情况相关, 同类样本间越接近则 K_s 值越小, 反之 K_s 值越大, 其可以根据交叉验证方法确定. 通过式 (16) 可得连接权重矩阵 $W = [w_{ij}]$. 图 G 是通过连接少数近邻样本得到的, 相距较远的样本间没有连接, 其对状态转移过程进行了约束. 由于参考点与近邻样本大多属于同一类别, 因此可以保证质点仅在同类别样本对应状态间进行转移, 能有效地避免不相关样本间的影响, 极大地消除冗余信息.

由于数据集中样本分布存在不平衡性, G 可能是由多个连通分量构成的非连通图, 可以通过深度优先遍历等方法得到各连通分量[15]. 为了方便后续的分析, 本小节假设 G 是一个连通图, 对于非连通图的情形将在第 3.3 节进行说明.

在形状样本集合对应的状态空间 $Space$ 中, 由状态 i 至状态 j 的状态转移概率定义为

$$p_{ij} = \frac{w_{ij}}{\sum\limits_{k=1}^{N} w_{ik}} \tag{17}$$

可得状态转移矩阵为 $P = [p_{ij}]$. 由于 G 是连通图, 且 $p_{ii} > 0$, 进而可知一定存在 $\lambda \geq 1$, 使得 $p_{ij}^{(\lambda)} > 0$, $P^{(\lambda)} = \left[p_{ij}^{(\lambda)}\right]$.

3.2　利用广义期望首达时间更新距离

状态转移矩阵 P 是由形状距离矩阵 D 构造生成的, 其中包含了丰富的样本空间局部几何信息. 借助期望首达时间进行形状距离学习时, 流形空间中的检索路径可以隐式地挖掘得到. 但是, 当两个同类样本在样本空间中相距较远时, 二者之间必定间隔了数目较多的同类别样本. 因此质点在状态转移过程中, 必将经过多个同类别样本对应的状态, 对应的期望首达时间较长.

图 3 中给出了一个由 7 个状态构成状态空间的示例. b 和 c 属于一个类别, a、d、e、f 和 g 属于另一个类别, 状态间连接权重与样本间距离成反比. 如果将 a 视作查询样本, 通过分析期望首达时间进行距离学习, 可以得到 $\mu_{ba} > \mu_{ea}$, 但是存在 $\mu_{ba} < \mu_{fa}$

和 $\mu_{ba} < \mu_{ga}$. 可以看出, 虽然期望首达时间能够有效分析相距较近的样本间关系, 却难以准确描述相距较远的样本间相似程度.

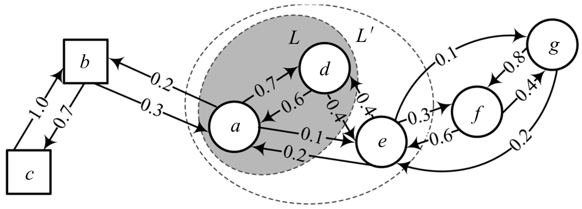


图 3 由 2 个类别样本对应的 7 个状态构成状态空间的示例

Fig. 3 An example of state space consisting of 7 states which corresponds to the samples from 2 categories

为了减少质点在同类样本对应状态间完成转移所需时间步长, 考虑显式地挖掘空间中的检索路径, 分析计算质点由各状态转移至检索路径对应集合的时间步长, 即将上一节定义的广义期望首达时间作为距离更新结果. 在图 3 的示例中, 借助广义期望首达时间分析的结果, 依次将 d 和 e 划入检索路径集合得到对应的 L 和 L', 可知 $\eta_{\mathrm{b}L} > \eta_{\mathrm{f}L}$ 和 $\eta_{\mathrm{b}L'} > \eta_{\mathrm{g}L'}$. 不难发现, 结合广义期望首达时间分析样本间距离能够得到更准确的检索结果.

通过此前分析可知广义期望首达时间满足线性方程 (13), 这里进一步分析如何进行求解. 定义 $(P_{UU})^{\lambda} = [q_{ij}]$, 可得

$$q_{ij} = \sum_{k_1=1}^{M} \cdots \sum_{k_{\lambda-1}=1}^{M} p_{ik_1} \dots p_{k_{\lambda-1}j} \leq \sum_{k_1=1}^{N} \cdots \sum_{k_{\lambda-1}=1}^{N} p_{ik_1} \dots p_{k_{\lambda-1}j} = p_{ij}^{(\lambda)} \tag{18}$$

由式 (18) 的关系可知

$$\sum_{j=1}^{M} q_{ij} \leq \sum_{j=1}^{M} p_{ij}^{(\lambda)} = 1 - \sum_{j=M+1}^{N} p_{ij}^{(\lambda)} < 1 \tag{19}$$

结合式 (19) 可知谱半径满足如下关系

$$\left\| (P_{UU})^{\lambda} \right\|_{\infty} = \max_{1 \leq i \leq M} \sum_{j=1}^{M} |q_{ij}| < 1 \tag{20}$$

$$\rho\,(P_{UU}) \leq \left\| (P_{UU})^{\lambda} \right\|_{\infty}^{\frac{1}{\lambda}} < 1 \tag{21}$$

据此, 进一步可以得到 $I - P_{UU}$ 可逆且 $\sum_{k=0}^{\infty} (P_{UU})^{k} = (I - P_{UU})^{-1}$, 式 (13) 对应的线性方程表示为

$$(I - P_{UU}) \times \boldsymbol{\eta}_L = (I - P_{UU})^{-1} \times P_{UL} \times \mathbf{1}_{N-M} \tag{22}$$

又 $\sum_{j=1}^{N} p_{ij} = 1$, 可知 $P_{UU} \times \mathbf{1}_M + P_{UL} \times \mathbf{1}_{N-M} = \mathbf{1}_M$, 进而

$$(I - P_{UU})^{-1} \times P_{UL} \times \mathbf{1}_{N-M} = \mathbf{1}_M \tag{23}$$

线性方程最终形式表示为

$$\boldsymbol{\eta}_L = (I - P_{UU})^{-1} \times \mathbf{1}_M \tag{24}$$

因此, 线性方程具有唯一解, 广义期望首达时间可以通过不同的线性方程求解方法得到. 进而可知, 其对应计算过程的时间复杂度与求解线性方程组的时间复杂度相同, 均为 $\mathrm{O}\,(M^3)$.

3.3 形状距离学习算法流程

此前为方便讨论, 假设 G 是一个连通图, 但是很多情况下 G 可能是非连通图, 质点无法在隶属于不同连通分量的状态和集合间完成状态转移. 以 C 表示包含查询样本的连通分量, 可以将其视作状态空间 $Space$ 并分析广义期望首达时间. 此外, 考虑到同类样本可能隶属于不同的连通分量, 因此进一步将其他连通分量中样本与查询样本间相似度进行分析排序.

综合此前的分析, 给出基于广义期望首达时间的形状距离学习算法步骤如下:

步骤 1. 利用形状间距离矩阵构造状态转移矩阵. 如果样本数目 N 很大, 则选择距离查询样本最近的 N' 个样本构造图 G. 进一步求取查询样本对应的连通分量 C 作为状态空间 $Space$, 并计算状态转移矩阵. 将查询形状样本视作集合 L, C 中余下样本作为集合 U.

步骤 2. 求解式 (24), 得到 U 中对应各样本至集合 L 的广义期望首达时间 $\boldsymbol{\eta}_L$.

步骤 3. 对 $\boldsymbol{\eta}_L$ 中各元素按照升序进行排序. 如果未达到终止条件 I_e 且 $\#U > K_\eta$, 则将排序中前 K_η 个状态从集合 U 中移除并依次加入到集合 L 中, 跳转至步骤 2; 否则, 算法停止, 将 $\boldsymbol{\eta}_L$ 的排序结果依次添加至集合 L 中, L 中的样本排序即为检索结果.

步骤 4. 对于包含在 G 中但不属于 C 的样本, 依照其与查询样本的相似度关系进行排序, 列于检索结果的最后.

在上述步骤中, 包含 3 个参数需要设定, 分别是 N'、K_η 和 I_e. 参数 N' 负责对检索过程中数据集规模进行限制, 通常设置 $N' \ll N$, 本文实验中选择 20% 左右的样本进行形状距离学习. 参数 K_η 负责控制每次迭代过程中加入检索路径中的样本数目, 通常为一个较小正数, 本文实验中设置 $K_\eta = 3$. I_e 负责对迭代次数进行约束, 待检索的相关样本数目越多则 I_e 越大. 本文实验中设置 $I_e = \mathrm{floor}\,(rs/K_\eta)$, 其中 rs 表示待检索的相关样本数目.

4　数据仿真

为了验证本文所提方法的有效性, 本节将对不同形状数据集进行仿真实验分析比较. 在实验过程中, 首先选择已有的几种比较成熟的逐对形状匹配方法分析形状间距离, 用到的方法包括形状上下文 (Shape context, SC)[16]、内距离形状上下文 (Inner-distance shape context, IDSC)[17] 和连接不变表示 (Articulation-invariant representation, AIR)[18]. 在选择相同逐对形状匹配方法的基础上, 分别利用本文提出的广义期望首达时间 (GMFPT) 方法和其他文献中的方法进行形状距离学习, 使用到的方法包括标记传播 (LP)[5]、元描述符 (MD)[10]、期望首达时间 (MFPT)[13] 等.

在分析实验结果的过程中, 分别将样本集合中每个样本依次作为查询形状, 分析其与各样本间的距离并将其按升序排列, 将距离查询样本较近的几组结果作为检索结果. 在检索得到的全部样本中, 相关样本数目与全部样本数目之比即为检索精度. 在对实验结果进行验证的过程中, 将主要以检索结果和检索精度作为评价方法性能的指标.

首先选择 Kimia-216 数据集[19] 进行实验比较, 该数据集中含了 18 个类别各 12 个形状样本, 分别选择形状上下文 (SC) 和内距离形状上下文 (IDSC) 计算形状距离矩阵, 此部分实验参数根据文献 [17] 中给出的方法进行设置. 由于数据集规模较小, 可以直接对样本集合中全部形状进行距离学习. 在两组形状距离学习实验中, 设置 $\alpha = 0.38$, $K_d = 12$, $K_s = 8$. 表 1 中列出了不同方法在 Kimia-216 数据集上得到的形状检索结果, 表 1 中各列分别表示与查询样本近邻的检索结果所属类别正确与否, 其中包含了最接近的 1～11 组相关样本数目及合计相关样本数目. 从表 1 中的结果中可以看出, 与其他形状距离学习方法相比, 本文的方法在相同条件下能够达到最好的形状检索效果.

为了进一步说明提出的形状距离学习方法在大数据集下的有效性, 本文分别选择了 Tari-1000 数据集[20] 和 MPEG-7 数据集[21] 进行实验. Tari-1000 数据集中包含了 1 000 个目标剪影, 其对应 50 个类别, 每个类别各 20 个形状样本, 图 4 列出了该数据集部分类别中一个剪影的示例. MPEG-7 数据集中包含了 1 400 个目标剪影, 其对应 70 个形状类别, 每个类别各 20 个样本, 图 5 列出了该数据集部分类别中一个剪影的示例.

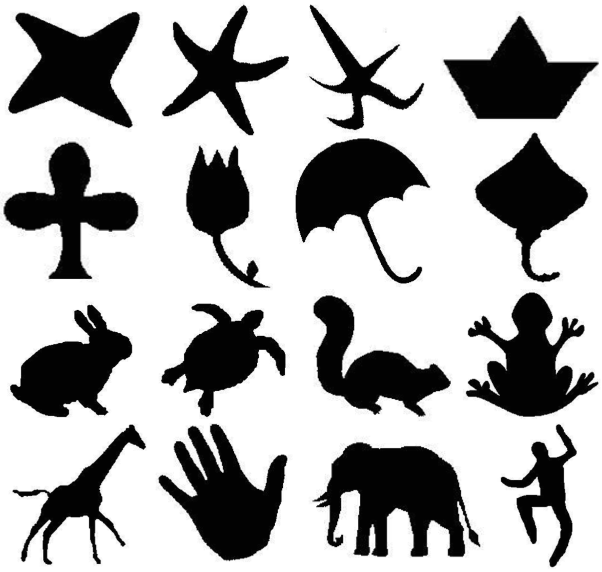


图 4　Tari-1000 数据集中部分类别形状样本示例

Fig. 4　Examples of shapes from different categories in Tari-1000 database

表 1　Kimia-216 数据集在不同方法下检索结果比较

Table 1　Comparison of retrieval rates for different algorithms tested on Kimia-216 database

方法	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	全部
SC	216	216	215	210	210	209	208	204	200	191	175	2 254
IDSC	216	216	215	211	211	210	211	207	203	198	185	2 283
SC + LP	216	216	214	212	211	211	215	209	209	206	197	2 316
IDSC + LP	216	216	214	211	213	213	212	210	207	208	203	2 323
SC + MD	215	215	215	213	212	212	214	211	211	209	208	2 335
IDSC + MD	215	215	215	211	212	213	212	212	207	209	209	2 330
SC + MFPT	216	216	216	212	212	212	212	212	212	211	212	2 343
IDSC + MFPT	216	216	216	212	212	212	212	212	212	212	212	2 344
SC + GMFPT	216	216	216	216	216	216	216	216	216	216	216	2 376
IDSC + GMFPT	216	216	216	216	216	216	216	216	216	216	216	2 376

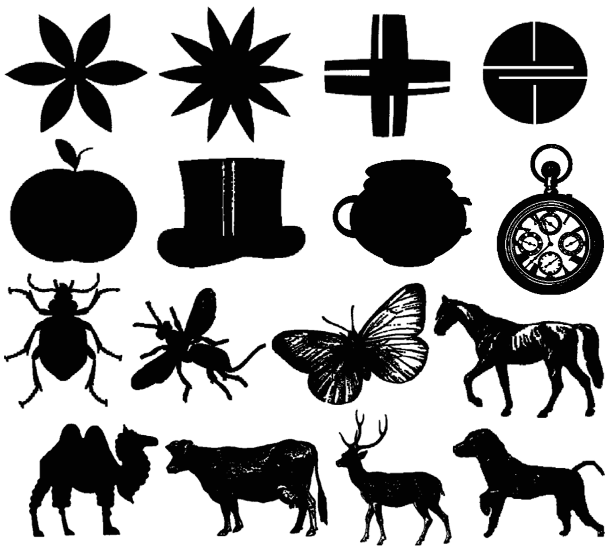


图 5 MPEG-7 数据集中部分类别形状样本示例
Fig. 5 Examples of shapes from different categories in MPEG-7 database

对 Tari-1000 数据集仿真过程中, 分别选择形状上下文 (SC) 和内距离形状上下文 (IDSC) 计算形状距离矩阵. 查询样本与全部样本匹配后, 分析计算距离最近的前 20 组结果对应的检索精度. 由于数据集包含样本数目较多, 因此设置参数 $N' = 200$. 选择 SC 计算距离矩阵时, 设置 $\alpha = 0.28$, $K_d = 22$, $K_s = 30$; 选择 IDSC 计算距离矩阵时, 设置 $\alpha = 0.32$, $K_d = 24$, $K_s = 30$. 表 2 中列出了不同方法在 Tari-1000 数据集上进行实验得到的形状检索精度. 从表中的结果可以看出, 与其他形状距离学习方法相比, 本文提出的方法在相同条件下能够得到最好的形状检索精度.

表 2 Tari-1000 数据集在不同方法下的结果比较
Table 2 Comparison of results for different algorithms tested on Tari-1000 database

方法	检索精度 (%)
SC	88.01
IDSC	90.43
SC + LP	94.22
IDSC + LP	96.44
SC + MD	94.98
IDSC + MD	98.49
SC + MFPT	97.02
IDSC + MFPT	99.11
SC + GMFPT	97.15
IDSC + GMFPT	99.27

对 MPEG-7 数据集进行实验过程中, 选择连接不变表示 (AIR)[18] 计算样本间距离矩阵, 该逐对形状匹配方法在 MPEG-7 数据集上具有最高的匹形状配精度. 在分析检索精度时, 各查询样本都与全部样本进行匹配, 统计距离最近的前 40 组结果中相关样本数目并计算查全率, 这种评价规则被称为 Bullseye 方法. 由于 MPEG-7 数据集中同类目标间差异大而不同类目标间差异较小, 形状匹配研究一直都将该数据集下的检索精度作为重点关注问题[21]. 由于包含样本数目较多, 检索过程中仅保留与查询形状临近的 300 个样本进行分析, 即 $N' = 300$. 余下参数分别设置为 $\alpha = 0.38$, $K_d = 18$, $K_s = 14$. 表 3 中列出了近年来文献中给出方法和本文方法在 MPEG-7 数据集下的形状检索精度. 从表中可以看出, 与其他形状距离学习方法相比, 本文提出的方法能够有效地提升形状匹配结果, 并能够得到 100% 的形状检索精度, 与文献中公布的最好结果相当.

表 3 MPEG-7 数据集在不同方法下的结果比较
Table 3 Comparison of results for different algorithms tested on MPEG-7 database

方法	检索精度 (Bullseye) (%)
IDSC + LP[5]	91.61
SC + GM + Meta Descriptor[10]	92.51
IDSC + LCDP[6]	93.32
IDSC + Mutual Graph[22]	93.40
SC + MFPT[13]	94.04
ASC + LCDP[8]	95.96
ASC + TPG Diffusion[11]	96.47
SC + IDSC + Co-transduction[23]	97.72
IDSC + SSC+LCDP[9]	98.85
AIR + TPG Diffusion[11]	99.99
AIR + Generic Diffusion Framework[12]	100.00
AIR + GMFPT	100.00

为了对不同形状距离学习方法进行综合比较, 这里进一步分析探讨各方法对应的时间复杂度. 已知求解广义期望首达时间的时间复杂度为 $\mathrm{O}(M^3)$, 且 M 与查询样本对应连通分量的样本数目 N_C 较为接近. 通过第 3.3 节的算法流程可知本文所提方法的时间复杂度为 $\mathrm{O}(I_e N_C^3)$. 表 4 中列出了本文方法与其他距离学习方法的时间复杂度, I_t 表示矩阵运算迭代次数. 从表中结果可以看出, 本文所提方法的时间复杂度高于 LP 和 MFPT 两种方法, 与文献 [6]、文献 [11] 和文献 [12] 所提算法复杂度大致相当.

但与上述方法相比, 本文方法能够得到更好的形状检索结果.

表 4 不同方法下的时间复杂度比较

Table 4 Comparison of the time complexities for different algorithms

方法	时间复杂度
LP[5]	$O(I_t N'^2)$
LCDP[6]	$O(I_t N^3)$
MFPT[13]	$O(N'^3)$
TPG diffusion[11]	$O(I_t N^3)$
Generic diffusion framework[12]	$O(I_t N^3)$
GMFPT	$O(I_e N_C^3)$

5 结论

考虑到逐对形状匹配方法难以准确反映目标间相似程度, 本文提出了基于广义期望首达时间的形状距离学习方法以改善形状匹配结果. 广义期望首达时间是在期望首达时间的基础上推广得到的, 其反映了质点由状态转移至状态集合所需的平均时间步长. 借助广义期望首达时间能够更好地反映流形结构, 提出的方法可以显式地挖掘样本空间信息, 所得结果能够更加准确地反映样本间的距离关系. 分别在不同的形状数据集下进行实验比较, 所提方法在相同条件下能够得到更好的形状检索结果.

References

1 Hu R X, Jia W, Ling H B, Zhao Y, Gui J. Angular pattern and binary angular pattern for shape retrieval. *IEEE Transactions on Image Processing*, 2014, **23**(3): 1118−1127

2 Hong B W, Soatto S. Shape matching using multiscale integral invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(1): 151−160

3 Zhou Yu, Liu Jun-Tao, Bai Xiang. Research and perspective on shape matching. *Acta Automatica Sinica*, 2012, **38**(6): 889−910
(周瑜, 刘俊涛, 白翔. 形状匹配方法研究与展望. 自动化学报, 2012, **38**(6): 889−910)

4 Hasanbelliu E, Sanchez G L, Principe J C. Information theoretic shape matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(12): 2436−2451

5 Bai X, Yang X W, Latecki L J, Liu W Y, Tu Z W. Learning context-sensitive shape similarity by graph transduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(5): 861−874

6 Yang X W, Koknar-Tezel S, Latecki L J. Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 357−364

7 Yang X W, Bai X, Latecki L J, Tu Z W. Improving shape retrieval by learning graph transduction. In: Proceedings of the 10th European Conference on Computer Vision. Marseille, France: Springer, 2008. 788−801

8 Ling H B, Yang X W, Latecki L J. Balancing deformability and discriminability for shape matching. In: Proceedings of the 11th European Conference on Computer Vision. Crete, Greece: Springer, 2010. 411−424

9 Premachandran V, Kakarala R. Perceptually motivated shape context which uses shape interiors. *Pattern Recognition*, 2013, **46**(8): 2092−2102

10 Egozi A, Keller Y, Guterman H. Improving shape retrieval by spectral matching and meta similarity. *IEEE Transactions on Image Processing*, 2010, **19**(5): 1319−1327

11 Yang X W, Prasad L, Latecki L J. Affinity learning with diffusion on tensor product graph. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(1): 28−38

12 Donoser M, Bischof H. Diffusion processes for retrieval revisited. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 1320−1327

13 Zheng Dan-Chen, Han Min. Learning shape distance based on mean first-passage time. *Acta Automatica Sinica*, 2014, **40**(1): 92−99
(郑丹晨, 韩敏. 基于期望首达时间的形状距离学习算法. 自动化学报, 2014, **40**(1): 92−99)

14 Wang J Y, Li Y P, Bai X, Zhang Y, Wang C, Tang N. Learning context-sensitive similarity by shortest path propagation. *Pattern Recognition*, 2011, **44**(10−11): 2367−2374

15 Hopcroft J, Tarjan R. Efficient algorithms for graph manipulation. *Communications of the ACM*, 1973, **16**(6): 372−378

16 Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(4): 509−522

17 Ling H B, Jacobs D W. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(2): 286−299

18 Gopalan R, Turaga P, Chellappa R. Articulation-invariant representation of non-planar shapes. In: Proceedings of the 11th European Conference on Computer Vision. Crete, Greece: Springer, 2010. 286−299

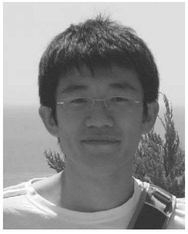
19 Sebastian T B, Klein P N, Kimia B B. Recognition of shapes by editing their shock graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(5): 550−571

20 Baseski E, Erdem A, Tari S. Dissimilarity between two skeletal trees in a context. *Pattern Recognition*, 2009, **42**(3): 370−385

21 Latecki L J, Lakamper R, Eckhardt T. Shape descriptors for non-rigid shapes with a single closed contour. In: Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head, USA: IEEE, 2000, **1**: 424−429

22 Kontschieder P, Donoser M, Bischof H. Beyond pairwise shape similarity analysis. In: Proceedings of the 9th Asian Conference on Computer Vision. Xi′an, China: Springer, 2010. 655−666

23 Bai X, Wang B, Yao C, Liu W Y, Tu Z W. Co-transduction for shape retrieval. *IEEE Transactions on Image Processing*, 2012, **21**(5): 2747−2757



郑丹晨 大连理工大学电子信息与电气工程学部讲师. 主要研究方向为计算机视觉和模式识别.
E-mail: dcjeong@dlut.edu.cn
(**ZHENG Dan-Chen** Lecturer at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. His research interest covers computer vision and pattern recognition.)

杨亚飞 大连理工大学电子信息与电气工程学部硕士研究生. 主要研究方向为模式识别.
E-mail: yangyafei@mail.dlut.edu.cn
(**YANG Ya-Fei** Master student at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. His main research interest is pattern recognition.)

韩 敏 大连理工大学电子信息与电气工程学部教授. 主要研究方向为模式识别, 复杂系统建模与分析及时间序列预测. 本文通信作者.
E-mail: minhan@dlut.edu.cn
(**HAN Min** Professor at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. Her research interest covers pattern recognition, modeling and analysis of complex system, and time series prediction. Corresponding author of this paper.)