

基于期望首达时间的形状距离学习算法

郑丹晨¹ 韩敏¹

摘 要 由于逐对形状匹配不能很好地反映形状间相似度, 因此需要引入后期处理步骤提升检索精度. 为了得到上下文敏感的形状相似度, 本文提出了一种基于期望首达时间 (Mean first-passage time, MFPT) 的形状距离学习方法. 在利用标准形状匹配方法得到距离矩阵的基础上, 建立离散时间马尔可夫链对形状流形结构进行分析. 将形状样本视作状态, 利用不同状态之间完成一次状态转移的平均时间步长, 即期望首达时间, 表示形状间的距离. 期望首达时间能够结合测地距离发掘空间流形结构, 并可以通过线性方程进行有效求解. 分别对不同数据进行实验分析, 本文所提出的方法在相同条件下能够达到更高的形状检索精度.

关键词 形状匹配, 形状距离学习, 相似度矩阵, 离散时间马尔可夫链, 期望首达时间

引用格式 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

DOI 10.3724/SP.J.1004.2014.00092

Learning Shape Distance Based on Mean First-passage Time

ZHENG Dan-Chen¹ HAN Min¹

Abstract Since pairwise shape similarity analysis can not measure the shape distance accurately, post-processing steps are introduced into shape matching process for increasing retrieval scores. In this paper, a novel shape distance based on mean first-passage time (MFPT) is proposed for resolving the problem of learning context-sensitive similarity. Given the distance matrix computed by a distance function, discrete-time Markov chains are constructed for analyzing the underlying structure of the manifold formed by shapes. With each shape in the database regarded as a state, the mean number of steps for transition between two states, named mean first-passage time, is used to measure the shape distance. The mean first-passage time induced by geodesic paths can capture the shape manifold structure, and it can be obtained by solving the linear equations. Experimental results on different databases show that shape retrieval results can be effectively achieved by using the proposed method.

Key words Shape matching, learning shape distance, affinity matrix, discrete-time Markov chain, mean first-passage time (MFPT)

Citation Zheng Dan-Chen, Han Min. Learning shape distance based on mean first-passage time. *Acta Automatica Sinica*, 2014, 40(1): 92–99

形状匹配是计算机视觉中一项重要的研究问题, 近几十年来涌现出了大量形状匹配方法, 并且已应用于目标跟踪、物体识别、医学成像等众多的领域^[1–2]. 在形状匹配方法中, 两个形状间匹配结果一般用形状距离或形状相似度表示, 形状越相似则对应的距离越小或相似度越大. 依据形状距离或相似度进行排序, 进而可以得到对应的形状检索精度, 其结果是评价形状匹配方法性能最重要的指标^[2].

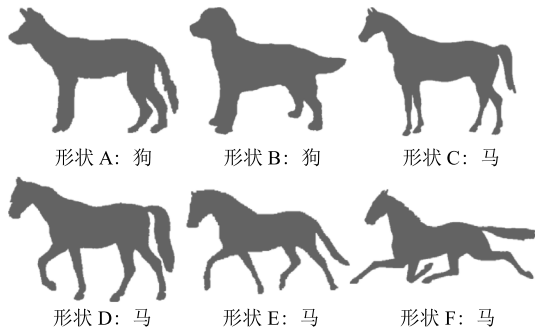
大多数形状匹配的研究都将形状目标视作孤立

的个体, 通过分析形状本身的特征计算两个目标间距离或相似度^[3–4]. 但是越来越多的研究表明, 简单比较两个形状间距离并不能准确地表达形状间的相似程度^[5–6]. 在形状特征空间中, 当相同类别形状差异较大、不同类别形状差异较小时, 仅逐对地计算形状间距离可能导致错误的匹配结果. 例如, 图 1(a) 列出了两类形状样本集合 (A 和 B 为狗; C~F 为马), 图 1(b) 列出了利用内距离形状上下文 (Inter-distance shape context, IDSC)^[4] 得到的形状匹配结果. 当选择 C 作为查询形状 (Query shape), 其余各样本作为已知形状 (Known shape) 时, C 与 A 和 B 的相似度要高于 C 与 F 的相似度. 以 d_{XY} 表示 IDSC 对应的形状距离函数, 则 C 与 A、B 和 F 间距离关系为 $d_{CB} < d_{CA} < d_{CF}$, 可以看出类内距离 (d_{CF}) 大于类间距离 (d_{CB}, d_{CA}). 因此形状检索中前 3 组结果对应的查全率为 2/3, 形状检索结果并不

收稿日期 2012-04-27 录用日期 2012-09-29
Manuscript received April 27, 2012; accepted September 29, 2012

国家自然科学基金 (61374154, 61074096) 资助
Supported by National Natural Science Foundation of China (61374154, 61074096)

本文责任编辑 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



(a) 6 个目标形状构成的 2 类样本集合

(a) A database of 6 shapes from 2 classes

	A	B	C	D	E	F
A	0	63.82	63.74	64.24	65.29	72.28
B	63.82	0	63.01	68.09	68.90	79.21
C	63.74	63.01	0	54.74	58.00	69.99
D	64.24	68.09	54.74	0	56.47	65.57
E	65.29	68.90	58.00	56.47	0	56.49
F	72.28	79.21	69.99	65.57	56.49	0

(b) 利用 IDSC 所得形状距离矩阵

(b) Shape distance matrix obtained by using IDSC

图 1 逐对的形状匹配方法例子

Fig. 1 An example of a pairwise shape matching approach

准确。

为了进一步提升形状匹配精度, 越来越多的研究尝试引入形状距离学习步骤提升形状匹配的准确度, 形状匹配流程如图 2 所示. Bai 等^[5]提出了使用半监督学习中的标记传播方法 (Label propagation) 进行形状距离学习. 由于在流形空间中的测地距离方向上标记传播速度较快, 因而可以在形状相似性度量空间中隐式地找到最优检索路径. 例如, 在图 1(a) 中, 虽然形状 C 与 F 之间因形态变化很大造成相似度较低, 但是 CD、DE、EF 之间相似度均很高, 因此由 C 通过 D、E 传递至 F 即为一条检索路径. 在构造马尔科夫链后, 标记传播方法固定查询形

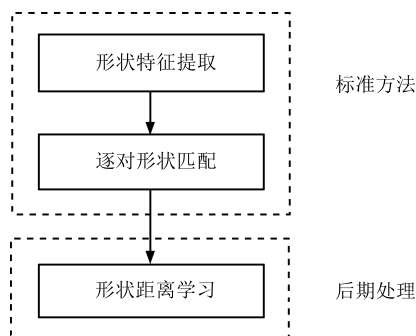


图 2 形状匹配流程

Fig. 2 Shape matching framework

状的标记函数, 迭代更新已知形状的标记函数, 并对其排序 (Ranking) 得到形状检索结果。

虽然标记传播算法能够在一定程度上提升形状检索的精度, 但是其依然存在以下问题. 首先, 算法中标记传播过程存在不平衡性. 由于查询形状的标记函数几乎总大于已知形状的标记函数, 因此查询形状与已知形状间标记传播总是较快, 而已知形状之间标记传播则较慢, 多次迭代误差累计, 难以准确获得测地距离方向上的检索路径. 同时, 标记传播算法中的形状距离学习是通过迭代更新实现的, 其时间复杂度较高, 迭代参数和终点需要通过经验确定。

标记传播方法提出后, 许多学者对形状距离学习方法进行了深入研究. Yang 等^[7]提出了一种通过局部约束扩散过程 (Locally constrained diffusion process) 进行形状距离学习的方法. 该方法通过增加构造点 (Ghost point) 对原始的样本空间集合进行扩展, 改善原始流形空间结构; 进一步在扩散图中引入局部约束条件, 利用扩散距离对形状间距离进行更新. Egozi 等^[8]也提出了一种利用元描述符 (Meta descriptor) 更新形状距离的方法. 算法通过在样本集合中选择目标形状 K 近邻形状的相似度来定义新的特征, 利用 L_1 范数重新计算形状间相似度. 最近, Bai 等^[9]进一步将协同学习 (Co-training) 方法引入形状检索当中, 综合多种方法进行形状匹配, 得到了更好的形状检索精度。

上述形状距离学习方法都借助马尔科夫链分析形状空间的流形信息, 其中形状样本和归一化后的相似度分别被视为状态和状态转移概率. 分析标记传播算法可知, 如果在形状特征空间中标记能够沿测地距离方向快速传播, 则不同状态可以在短时间内沿测地距离方向完成状态转移. 鉴于此, 本文考虑利用不同状态之间状态转移所需的平均时间步长衡量形状间距离, 该平均时间步长即为期望首达时间 (Mean first-passage time, MFPT). 在马尔科夫链中, 首达时间能够定量分析各点间移动时间步长, 其已经应用于解决目标搜索等问题^[10-11]. 对于形状距离学习问题, 期望首达时间能够很好地发掘形状样本空间的拓扑结构, 隐式地反映了相似度空间中测地距离的信息, 并可以采用线性方程进行快速求解. 通过对多组数据的仿真比较, 本文提出的形状距离学习方法与其他算法相比能够有效地提升形状检索精度。

1 构造离散时间马尔科夫链

已知由 N 个形状样本构成的集合 $S = \{s_1, s_2, \dots, s_N\}$, 通过形状匹配算法, 例如文献 [3-4] 的方

法,可以得到 $N \times N$ 距离矩阵 $D = (d_{ij})$, $d_{ij} \in [0, +\infty)$. 为了更好地发掘样本点在形状特征空间中的几何结构,将距离矩阵转化为一个相似度矩阵 (Affinity matrix) 进行分析,并构造对应的离散时间马尔科夫链.

1.1 相似度矩阵

针对已知集合 S , 可以构造图 $G = (S, W)$, 其中, 图 G 的顶点由各样本点构成; 图 G 的边表示为 s_i 与 s_j 的连接强度 $w(i, j) \in W$, $s_i \in S$, $s_j \in S$. 本文通过选择高斯核函数将 d_{ij} 转化为 $w(i, j)$:

$$w(i, j) = \exp\left(-\frac{d_{ij}^2}{\sigma_{ij}^2}\right) \quad (1)$$

其中, σ_{ij} 为核宽, 参考文献 [5] 中的方法, 根据 K 近邻的均值自适应设置核宽, 即:

$$\sigma_{ij} = \alpha \cdot \text{mean}(\{knnd(s_i), knnd(s_j)\}) \quad (2)$$

$\text{mean}(\{knnd(s_i), knnd(s_j)\})$ 表示样本 s_i 和 s_j 的 K 近邻距离的均值, 参数 K 和 α 均由经验确定. 由于形状间距离 $d_{ij} \geq 0$, 可知 $w(i, j) \in (0, 1]$, 因此图 G 是一个全连接图. 如果将 W 表示为矩阵形式, 则其为相似度矩阵. 将集合 S 视作一个状态空间 $L = \{1, 2, \dots, N\}$, 其中状态 $i \in L$ 对应样本 $s_i \in S$. 由样本的连接强度可得状态 i 和 j 间的 (一步) 转移概率可表示为

$$p_{i,j} = \frac{w(i, j)}{\sum_{k=1}^N w(i, k)} \quad (3)$$

则 (一步) 转移矩阵可以表示为 $P = (p_{i,j})$.

1.2 离散时间马尔科夫链

状态转移矩阵 P 由样本间距离矩阵 D 得到, 其中包含了形状空间的流形信息, $p_{i,j}$ 对应单位时间步长下由状态 i 转移至状态 j 的概率. 在状态空间下建立离散时间马尔科夫链. 随机变量序列 $\{X_n | n = 0, 1, 2, \dots\}$, 每个 X_n 都在状态空间 L 中取值, $\{X_n\}$ 为时齐的马尔科夫链:

$$\begin{aligned} P(X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0) = \\ P(X_{n+1} = i_{n+1} | X_n = i_n) = \\ P(X_1 = i_{n+1} | X_0 = i_n) \end{aligned} \quad (4)$$

引入条件概率

$$f_{ij}^{(n)} = P_i(X_n = j, X_m \neq j, 1 \leq m \leq n-1) \quad (5)$$

其中, $P_i(\cdot) = P(\cdot | X_0 = i)$. 若 $\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$, 则 i 是常返状态; 若 $\sum_{n=1}^{\infty} f_{ii}^{(n)} < 1$, 则 i 是非常返状态.

期望回转时间 μ_i 为

$$\begin{aligned} \mu_i = E(T_i | X_0 = i) = \\ \sum_{n=1}^{\infty} n P_i(T_i = n) = \sum_{n=1}^{\infty} n f_{ii}^{(n)} \end{aligned} \quad (6)$$

$T_i = \inf\{n \geq 0 : X_n = i\}$ 为首次到达状态 i 的转移次数.

如果 $\mu_i < +\infty$, 则 i 是正常返状态; 如果 $\mu_i = +\infty$, 则 i 是零常返状态. 因此, 仅当 i 是正常返状态时, 由状态 i 出发可以经过平均有限个时间步长返回状态 i ; 期望回转时间越小表明返回状态 i 越频繁. 由于图 G 是一个全连接图, 因此整个状态空间 L 中所有状态均是互通的, 即 L 是一个等价类; 又 L 中包含了有限个状态, 因而 L 是一个正常返的等价类, 所有状态都是正常返的^[12].

2 期望首达时间表示形状间距离

2.1 期望首达时间

首先给出期望首达时间的定义.

定义 1. 对于离散时间马尔科夫链中的状态 i 和 j , 定义由 i 转移至 j 的平均状态转移次数为期望首达时间, 记作 $\mu_{i,j}$:

$$\begin{aligned} \mu_{i,j} = E(T_j | X_0 = i) = \\ \sum_{n=1}^{\infty} n P_i(T_j = n) = \sum_{n=1}^{\infty} n f_{ij}^{(n)} \end{aligned} \quad (7)$$

由定义 1 可知期望首达时间 $\mu_{i,j}$ 反映了状态转移所需时间步长的平均值. 当 $i = j$ 时, 期望首达时间即为期望回转时间, 即期望回转时间是期望首达时间的一种特殊形式.

转移矩阵 P 包含了各样本与直接近邻样本的局部空间几何信息, 且转移概率 $p_{i,j}$ 正比于 i 和 j 之间的连接权重 $w(i, j)$. 在整个样本集合构成的流形中, 相似样本对应状态可以沿测地距离方向快速完成转移, 需要较小的时间步长; 而不相似样本之间对应状态转移相对缓慢, 需要较大的时间步长. 如图 3 所示, 状态空间中包含状态 a、b、c、d、e, 状态转移概率满足 $p_{a,b} > p_{a,e}$, 即 a 与 b 间的相似度大于 a 与 e 之间的相似度. 但是由于 a 与 d、d 与 e 之间转移概率均较大, 由 e 可以经 d 快速转移至 a; 同时, b 与 c 之间转移概率较大, 由 b 需要经多步才能转移至 a, 计算可得 $\mu_{b,a} \approx 5.67 > \mu_{e,a} = 3$. 可以看出期望首达时间能够有效结合局部空间信息, 分析空间流形中目标间的关系.

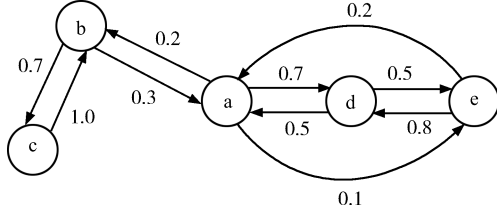


图3 期望首达时间例子

Fig. 3 An example of mean first-passage time

为了更好地分析期望首达时间的特性, 在定义 2 中引入常转移和非常转移的概念, 进一步提出了定理 1 并给出证明.

定义 2. 已知 i 和 j 分别为状态空间中的元素. 若 $\sum_{n=1}^{\infty} f_{ij}^{(n)} = 1$, 则由 i 至 j 是常转移的; 若 $\sum_{n=1}^{\infty} f_{ij}^{(n)} < 1$, 则由 i 至 j 是非常转移的.

定理 1. 已知 i 为状态空间 L 中的元素, 如果状态 i 是常返状态, 且对于任意给定的 $j \in L, j \neq i$, 均有 $p_{i,j} > 0$, 则由 j 至 i 是常转移的.

证明. 当 $n > 1$ 时

$$\begin{aligned} f_{ii}^{(n)} &= P_i(X_n = i, X_m \neq i, 1 \leq m \leq n-1) = \\ &\sum_{j \neq i} p_{i,j} P(X_n = i, X_m \neq i, \\ &2 \leq m \leq n-1 | X_1 = j) = \\ &\sum_{j \neq i} p_{i,j} P_j(X_{n-1} = i, X_m \neq i, \\ &1 \leq m \leq n-2) = \sum_{j \neq i} p_{i,j} f_{ji}^{(n-1)} \end{aligned} \quad (8)$$

可以得到:

$$\begin{aligned} \sum_{n=1}^{\infty} f_{ii}^{(n)} &= f_{ii}^{(1)} + \sum_{n=2}^{\infty} f_{ii}^{(n)} = p_{i,i} + \\ &\sum_{n=2}^{\infty} \sum_{j \neq i} p_{i,j} f_{ji}^{(n-1)} = p_{i,i} + \sum_{j \neq i} p_{i,j} \sum_{n=1}^{\infty} f_{ji}^{(n)} \end{aligned} \quad (9)$$

由于 i 是正常返状态, 又 $\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$, 因为条件概率一定满足 $\sum_{n=1}^{\infty} f_{ji}^{(n)} \leq 1$, $\sum_{j \in L} p_{i,j} = 1$, 且已知 $p_{i,j} > 0$, 则可知 $\sum_{n=1}^{\infty} f_{ji}^{(n)} = 1$, 即由 j 至 i 是常转移的. $\square$

2.2 期望首达时间的求解

假设查询形状为 $s_i \in S$, 集合 S 余下样本为已知形状, 则 i 与其他 $N-1$ 个元素间的相似度需要更新. 通过定理 1 可知在状态空间中, 对于状态空间中任意元素 $j \neq i$, 由 j 至 i 均为常转移的, 即 $\sum_{n=1}^{\infty} f_{ji}^{(n)} = 1$, 期望首达时间 $\mu_{j,i}, j \neq i$, 满足如式 (10) 所示的关系.

$$\begin{aligned} \mu_{j,i} &= \sum_{n=1}^{\infty} n f_{ji}^{(n)} = \sum_{n=1}^{\infty} (n-1) f_{ji}^{(n)} + \sum_{n=1}^{\infty} f_{ji}^{(n)} = \\ &\sum_{n=1}^{\infty} (n-1) P_j(X_n = i, X_m \neq i, \\ &1 \leq m \leq n-1) + 1 = \\ &\sum_{n=2}^{\infty} (n-1) \sum_{k \neq i} p_{j,k} P(X_n = i, X_m \neq i, \\ &2 \leq m \leq n-1 | X_1 = k) + 1 = \\ &\sum_{n=2}^{\infty} (n-1) \sum_{k \neq i} p_{j,k} P_k(X_{n-1} = i, X_m \neq i, \\ &2 \leq m \leq n-2) + 1 = \\ &\sum_{k \neq i} p_{j,k} \sum_{n=2}^{\infty} (n-1) f_{ki}^{(n-1)} + 1 = \\ &\sum_{k \neq i} p_{j,k} \mu_{k,i} + 1 \end{aligned} \quad (10)$$

对于包含 N 个元素的状态空间 L , 进一步可以得到如式 (11) 所示的线性方程组,

$$\begin{bmatrix} \mu_{1,i} \\ \vdots \\ \mu_{i-1,i} \\ \mu_{i+1,i} \\ \vdots \\ \mu_{N,i} \end{bmatrix} = \begin{bmatrix} p_{1,1} & \cdots & p_{1,i-1} & p_{1,i+1} & \cdots & p_{1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_{i-1,1} & \cdots & p_{i-1,i-1} & p_{i-1,i+1} & \cdots & p_{i-1,N} \\ p_{i+1,1} & \cdots & p_{i+1,i-1} & p_{i+1,i+1} & \cdots & p_{i+1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & \cdots & p_{N,i-1} & p_{N,i+1} & \cdots & p_{N,N} \end{bmatrix} \times \begin{bmatrix} \mu_{1,i} \\ \vdots \\ \mu_{i-1,i} \\ \mu_{i+1,i} \\ \vdots \\ \mu_{N,i} \end{bmatrix} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad (11)$$

将其表示为矩阵形式如下:

$$\boldsymbol{\mu}_i = A \boldsymbol{\mu}_i + \mathbf{1} \quad (12)$$

$$(I - A)\boldsymbol{\mu}_i = \mathbf{1} \quad (13)$$

其中, $\boldsymbol{\mu}_i = [\mu_{1,i}, \dots, \mu_{i-1,i}, \mu_{i+1,i}, \dots, \mu_{N,i}]^T$, $\mathbf{1} = [1, \dots, 1]^T$, $A = (p_{j,k})$, $j \neq i$, $k \neq i$.

通过式 (10) 可知, $\mu_i = \sum_{k \neq i} p_{i,k} \mu_{k,i} + 1$, $p_{i,k} > 0$, 又有 $0 \leq \mu_i < +\infty$, $\mu_{i,j} \geq 0$, 可知此时期望首达时间都是有界的. 方程 (13) 满足条件 $\sum_{k \in L} p_{j,k} = 1$, $p_{j,i} > 0$, 在 $I - A$ 中, 有 $1 - p_{j,j} > \sum_{k \neq j} | -p_{j,k} | = 1 - p_{j,j} - p_{j,i}$, 即 $I - A$ 为严格对角占优方阵, 其对应线性方程 (13) 存在唯一解, 因此线性方程组求得的解即为期望首达时间.

当样本集合中所包含各类别元素数目较多时, 通常不需要对全部样本进行形状距离学习, 仅需要对其少数样本进行详细分析. 已知 $s_i \in S$ 为查询形状, 仅保留与 s_i 最近邻的 M 个已知形状, 生成对应状态子空间 L' . 对于任意状态 $j \in L'$, $j \neq i$, 可知由 j 至 i 依然是常转移的, $p_{i,j} > 0$, $p_{j,i} > 0$, L' 依然是正常返等价类且满足定理 1, 可以利用其对应的线性方程求解.

综上所述, 利用期望首达时间对形状样本集合进行形状距离学习步骤如下:

1) 利用式 (1) 计算形状间相似度矩阵. 如果样本数目 N 较大, 则保留样本 i 及与之最相似的 M 个样本构成的 $(M+1) \times (M+1)$ 维的相似度矩阵.

2) 利用式 (3) 计算状态空间 L 或状态子空间 L' 对应的状态转移矩阵, 即生成对应的离散时间马尔科夫链.

3) 求解方程 (13), 求得的期望首达时间 μ_i 作为形状距离学习的结果.

在样本数目较少或已经构造状态子空间的情况下, 对于方程 (13) 可以直接一步求解; 如果构造状态子空间中依然包含大量的样本, 可以选择 Jacobi 迭代法、Gauss-Seidel 迭代法或其他方法进行近似求解.

3 数据仿真

为了验证本文所提方法的性能, 本节将给出对不同数据进行仿真的结果, 并对所得到的结果进行详细分析比较.

选择形状上下文 (Shape context, SC)^[3] 和内距离形状上下文 (IDSC)^[4] 作为标准形状匹配方法. SC 作为匹配方法时, 分别将 $\log r$ 和 θ 进行 5 等分和 12 等分; IDSC 作为匹配方法时, 分别将 $\log r$ 和 θ 进行 8 等分和 12 等分. 采样点数目均设置为 100, 利用文献 [3] 中的动态规划方法 (Dynamic programming, DP) 进行采样点的匹配, 通过计算两组采样点序列解决镜像形状的问题^[4].

在形状距离学习实验中, 选择一个形状作为查

询形状, 余下形状作为已知形状. 利用形状距离学习算法对形状距离进行更新, 并对新的形状距离进行排序, 分别对样本集合中各形状重复上述实验. 在标准形状匹配方法相同的前提下, 选择标注传播方法 (Label propagation, LP)^[5] 和基于元描述符 (Meta descriptor, MD) 方法^[8], 与本文所提出的基于期望首达时间 (MFPT) 方法进行形状距离学习的结果比较. 在利用本文方法构造相似度矩阵过程中, 参数 K 选择范围是 4~14, 形状样本数目越大时, 则对应 K 的值越大, 可以得到更准确的距离均值. 参数 α 介于 0.25 至 0.35 之间, α 通过交叉检验的方法进行确定.

首先对 Kimia-216 形状数据集^[13] 进行仿真. Kimia-216 是一个包含较少样本的数据集, 包含有 18 个类别各 12 个样本. 在计算相似度矩阵过程中, 分别设置参数 $\alpha = 0.35$, $K = 6$. 由于数据样本规模较小, 仿真中直接对全部的样本求解期望首达时间. 不同方法在 Kimia-216 数据集上仿真结果对应的 PR 曲线如图 4 所示, 表 1 列出了检索到的具体样本数目, 包括仅利用 SC 和 IDSC 进行形状匹配的结果, 以及分别选择 LP、MD 及本文提出的 MFPT 进行形状距离学习的结果. PR 曲线中包含了查准率 (Precision) 与查全率 (Recall) 的结果, 具体定义如下:

$$\text{Precision} = \frac{tp}{tp + fp} \quad (14)$$

$$\text{Recall} = \frac{tp}{tp + fn} \quad (15)$$

其中, tp 为检索结果的正确样本个数, fp 为检索结果的错误样本个数, fn 为未检索到的相关样本个数. 式 (15) 对应的查全率即为形状检索精度. 结合图 4

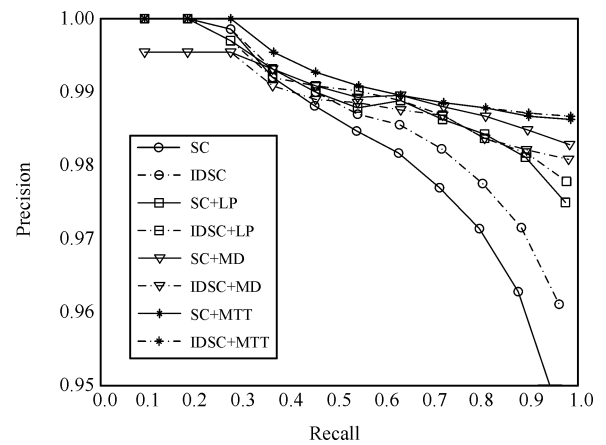


图 4 Kimia-216 数据的 PR 曲线

Fig. 4 Precision-recall curves for the Kimia-216 database

表 1 Kimia-216 数据在不同方法下检索结果比较

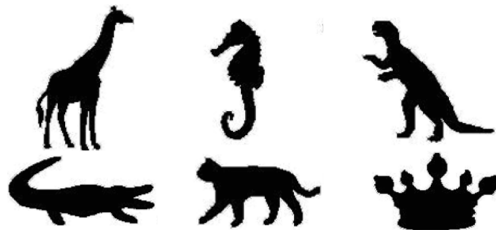
Table 1 Comparison of retrieval results for different algorithms tested on the Kimia-216 database

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

和表 1 的结果可以看出, 本文所提出的方法能够很好地解决标准形状匹配方法中存在的类内距离大于类间距离的问题, 与其他形状距离学习方法相比, 可以得到更准确的形状检索结果.

为了验证本文方法在大数据量样本下的有效性, 分别选择了 Tari-1000^[14] 形状数据集和 MPEG-7 形状数据集^[15] 两组数据进行仿真分析, Tari-1000 数据库中包含了 50 个类别各 20 个形状样本, 部分样本示例如图 5(a) 所示; MPEG-7 数据库中包含了 70 个类别各 20 组形状样本, 部分样本示例如图 5(b) 所示.

对 Tari-1000 形状数据集仿真过程中, 由于包



(a) Tari-1000 中形状样本例子

(a) Examples of shapes in the Tari-1000 database



(b) MPEG-7 中形状样本例子

(b) Examples of shapes in the MPEG-7 database

图 5 形状数据的例子

Fig. 5 Examples in shape databases

含样本数目较多, 因此每次测试过程中分别构造状态子空间进行分析, 设置 $M = 200$. 构造相似度矩阵过程中, 设置参数 $\alpha = 0.3$, $K = 12$. 不同方法在 Tari-1000 形状数据集上的仿真曲线如图 6 所示, 表 2 列出了前 12 组的检索结果, 实验中所采用的方法与之前的实验相同. 结合图 6 和表 2 可以看出, 与其他方法相比, 本文所提出的形状距离学习方法能够得到更加准确的形状检索结果.

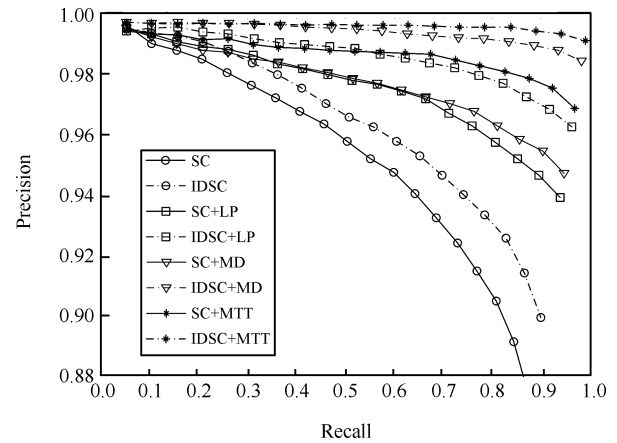


图 6 Tari-1000 数据的 PR 曲线

Fig. 6 Precision-recall curves for the Tari-1000 database

对 MPEG-7 形状数据集仿真过程中, 选择最近的 40 组中正确的结果计算查全率, 作为检索结果. 设置 $M = 300$ 构造状态子空间. IDSC 作为标准匹配方法时, 设置 $\alpha = 0.27$, $K = 13$; SC 作为匹配方法时, 设置 $\alpha = 0.3$, $K = 18$. 表 3 列出了本文方法所得形状距离学习结果, 与其他文献使用相同数据所得结果进行比较, 可以看出本文方法具有较好的结果.

利用本文所提出方法进行一次形状距离学习的

表 2 Tari-1000 数据在不同方法下检索结果比较

Table 2 Comparison of retrieval results for different algorithms tested on the Tari-1000 database

方法	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	全部
SC	996	984	983	976	954	954	949	936	928	908	893	898	11 368
IDSC	995	992	990	985	974	966	955	945	928	927	930	906	11 493
SC + LP	995	991	986	984	984	976	967	970	963	963	961	952	11 692
IDSC + LP	994	996	996	989	990	985	982	985	985	981	967	971	11 821
SC + MD	995	989	986	981	984	973	979	969	966	962	961	948	11 693
IDSC + MD	997	996	998	996	996	995	995	991	992	990	989	983	11 918
SC + MFPT	994	992	992	987	992	980	984	986	983	984	984	984	11 842
IDSC + MFPT	997	996	996	998	995	997	995	995	997	995	996	994	11 951

表 3 MPEG-7 数据在不同方法下的结果比较

Table 3 Comparison of results for different algorithms tested on the MPEG-7 database

方法	检索精度 (%)
Shape context ^[3]	76.51
MDS + SC ^[4]	84.35
IDSC ^[4]	85.40
Robust symbolic representation ^[16]	85.92
Shape-tree ^[17]	87.70
Contour flexibility ^[18]	89.31
IDSC + LP ^[5]	91.61
IDSC + LCDP ^[7]	93.32
GM + IDSC + Meta descriptor ^[8]	91.46
GM + SC + Meta descriptor ^[8]	92.51
IDSC + MFPT	92.87
SC + MFPT	94.04

时间复杂度与求解线性方程组的时间复杂度相同, 均为 $O(M^3)$, 标记传播算法进行一次形状检索的算法时间复杂度为 $O(I \times M^2)$, 其中 M 为形状距离学习的样本数目, I 为迭代次数. 通常的形状距离学习过程中, 样本规模都在 10^2 数量级, 而迭代次数往往可以达到 10^3 以上, 如文献 [5] 中对 MPEG-7 形状数据集实验设置 $I = 5 \times 10^3$, 因此本文方法与标记传播算法相比具有较低的时间复杂度. 本文方法的时间复杂度虽然高于文献 [8] 中算法的复杂度, 但是在精度上本文方法具有很大优势.

4 结论

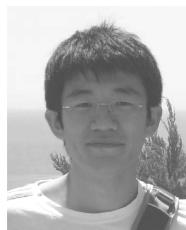
由于传统的形状匹配方法仅关注逐对的形状间

距离, 不能准确地反映形状间相似度, 本文提出了一种基于期望首达时间的形状距离学习方法. 通过标准形状匹配方法对应的距离矩阵生成离散时间马尔科夫链, 利用状态之间的平均状态转移次数作为形状距离学习的结果. 期望首达时间能够很好地发掘形状距离空间中的流形信息, 更加准确地反映形状间的距离. 通过对不同数据集样本进行仿真分析, 证明了本文所提出的方法具有良好的形状检索效果.

References

- Ding Xian-Feng, Wu Hong, Zhang Hong-Jiang, Ma Song-De. Review on shape matching. *Acta Automatica Sinica*, 2001, **27**(5): 678–694
(丁险峰, 吴洪, 张宏江, 马颂德. 形状匹配综述. 自动化学报, 2001, **27**(5): 678–694)
- 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)
- 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
- 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
- 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
- Kontschieder P, Donoser M, Bischof H. Beyond pairwise shape similarity analysis. In: *Proceedings of the 9th Asian Conference on Computer Vision*. Berlin, Heidelberg: Springer-Verlag, 2009. 655–666
- 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 Computer Society Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 357–364
- 8 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
- 9 Bai X, Wang B, Wang X G, Liu W Y, Tu Z W. Co-transduction for shape retrieval. In: Proceedings of the 11th European Conference on Computer Vision. Heraklion, Greece: Springer, 2010. 328–341
- 10 Render S. *A Guide to First-Passage Processes*. Cambridge: Cambridge University Press, 2001.
- 11 Condamin S, Bénichou O, Tejedor V, Voituriez R, Klafter J. First-passage times in complex scale-invariant media. *Nature*, 2007, **450**(7166): 77–80
- 12 He Shu-Yuan. *Stochastic Process*. Beijing: Peking University Press, 2008. 136–166
(何书元. 随机过程. 北京: 北京大学出版社, 2008. 136–166)
- 13 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
- 14 Baseski E, Baseski E, Tari S. Dissimilarity between two skeletal trees in a context. *Pattern Recognition*, 2009, **42**(3): 370–385
- 15 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. 424–429
- 16 Daliri M R, Torre V. Robust symbolic representation for shape recognition and retrieval. *Pattern Recognition*, 2008, **41**(5): 1782–1798
- 17 Felzenszwalb P F, Schwartz J D. Hierarchical matching of deformable shapes. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minnesota, USA: IEEE, 2007. 1–8
- 18 Xu C J, Liu J Z, Tang X O. 2D shape matching by contour flexibility. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(1): 180–186



郑丹晨 大连理工大学电子信息与电气工程学部博士研究生. 主要研究方向为计算机视觉和模式识别.

E-mail: danchen@mail.dlut.edu.cn

(**ZHENG Dan-Chen** Ph.D. candidate at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. His research interest covers computer vision and 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.)