

交互标记约束的三维网格序列分割

郑磊¹ 吴俊威² 林俊勉² 潘翔²

1 浙江广播电视集团 杭州 311200

2 浙江工业大学计算机科学与技术学院 杭州 310023

(zl099@zrtg.com)

摘要 已有交互分割只能适用于单一的三维模型,文中对已有算法进行了扩展,针对三维网格序列提出了交互标记约束分割算法。首先用户任意选择一个网格模型在分割线附近进行点交互,形成交互式分割的约束标记;然后,所提算法采用测地距离构造等距映射,建立其他帧和交互网格之间的映射关系,实现约束标记的自动传递;最后考虑分割边界所具有的凹特性和平滑性,根据约束标记定义局部邻域等值线对不同帧进行交互式分割。在实验部分,文中针对不同类型的三维网格序列进行分割性能分析。实验结果表明,所提算法在姿态变化下仍能够得到一致的分割效果,并且优于已有算法。

关键词: 三维网格序列;交互式分割;点交互;标记传递;等距映射

中图法分类号 TP391

Marker-constrained Interactive Segmentation of 3D Animated Meshes

ZHENG Lei¹, WU Jun-wei², LIN Jun-mian² and PAN Xiang²

1 Zhejiang Radio and Television Group, Hangzhou 311200, China

2 College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China

Abstract Existing interactive approaches only works for single 3D meshes. In view of this, this paper proposes an interactive algorithm of segmenting 3D animated meshes based on 3D data correspondence. Firstly, users can mark some points on any one 3D mesh for interactive segmentation. Then, the algorithm can map users, marks to other meshes by geodesic distance and isometric mapping. Finally, it performs interactive segmentation of other frames by transferred markers and iso-lines. Experimental results show that the algorithm can effectively segment different kinds of 3D animations. In addition, it can effectively improve the segmenting quality and is the exiting algorithms.

Keywords 3D animated meshes, Interactive segmentation, Dot interaction, Marker transfer, Isometric mapping

随着虚拟现实、3D动画、游戏的发展,对三维模型的研究得到了更多科研人员的重视。对于三维模型,如何有效地对其进行分割从而得到语义有效的子部分结构,是一个非常重要的研究方向。分割结果可以应用于很多研究领域,包括三维物体识别^[1]、动态模型网格重构^[2]、三维模型参数化^[3]和三维物体检索^[4]。因此,三维模型分割技术在3D领域的重要性越发突出。

对于三维模型分割,一种是根据形状特征进行无监督聚类,然后得到分割。其中,最为典型的是分水岭分割算法^[5],该算法采用曲率构造高度函数,然后通过区域增长实现分割,但是该算法会导致明显的过分割。对于具有明显子部分的三维物体,测地距离和形状直径函数是两种非常有效的形状特征,通过测地距离可以确定三维模型的子部分数目,然后通过无监督聚类得到层次分割结果^[6]。类似地,形状直径函数(SDF)也可以用于聚类分割和检索^[7]。而热核信号也被广泛应用于聚类分割^[8]。考虑到对同一类模型的分割结果非常相似,如果能够对多个模型,则能够进一步提高分割结果的一致性。为此,研究人员考虑采用形状上下

文特征来提高三维网格模型序列的分割一致性^[9-10]。对于分割,也可以进一步采用深度学习来提高分割的一致性^[11]。另一种方法是在分割过程中引入用户交互,通过用户操作得到期望的分割结果。目前,针对单一模型已经提出了很多交互方式,包括点交互和草图交互。点交互是让用户在分割边界上点击,然后得到分割边界,最为典型的工作是 DotScissor^[12];而草图交互则是用户在网格上绘制轨迹,然后通过图割等算法找到最优边界。类似地,智能剪刀考虑到分割凹边界在测地距离上具有最短性,因此考虑采用最短路径闭环作为分割边界^[13]。而 painting brush 则是提出了渐进式分割逼近方法^[14]。显然,通过交互式的方式进行分割能得到正确的分割结果。

目前,已有研究在单一模型的交互式操作方面做了大量工作,但是并没有考虑如何在三维模型序列中引入交互操作来提高分割一致性。本文针对三维模型序列提出了交互分割框架,该框架是 DotScissor 的扩展,但是 DotScissor 只能应用于单一模型,而本文方法能够分割三维网格模型序列。本文通过等距映射来完成用户标记传递,从而使得用户无需在其

基金项目:浙江省自然科学基金(LY19F020031)

This work was supported by the Natural Science Foundation of Zhejiang Province, China (LY19F020031).

通信作者:吴俊威(470720501@qq.com)

他模型上进行交互标记。此外,文本针对三维模型序列分割问题采用了具有等周长特性的泊松方程进行局部优化,得到光滑的分割边界。实验中使用交互式分割算法对不同序列的三维模型进行分割,结果表明本文算法对不同类型的序列都能得到正确的语义结果,且与其他分割算法相比其性能更优。

1 算法概述

对于三维模型分割,如何解决过分割是研究需要解决的问题。对于过分割问题,交互操作是一种有效的解决方法。图1给出了交互式和非交互式方法的分割结果。可以发现,对于非交互式分割,由于算法只是考虑高曲率特征用于控制分割边界,因此分割结果是不可控制的,例如头部的耳朵和背部被分割在一起、手部分割边界不能定位在关节位置。与非交互式方法相比,交互式方法根据用户标定进行分割,使得分割结果很好地逼近用户的分割意图。

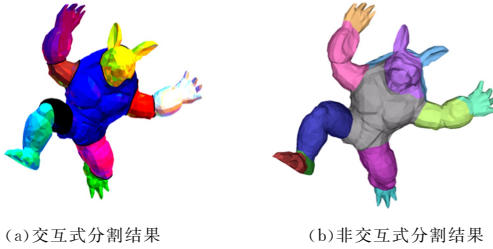


图1 交互式和非交互式分割结果的比较

Fig. 1 Comparison of interactive and non interactive segmentation results

但是,对于包含多个三维模型的网格序列,需要解决的问题是如何交互式地将分割结果映射到同一类模型中,解决三维模型序列的过分割问题。为此,本文提出了基于标记传递的三维模型序列交互式分割方法。该算法的主要思想是针对三维动画模型序列分割采用等距映射完成标记传递,并采用泊松方程完成边界优化。本文算法的流程如图2所示。

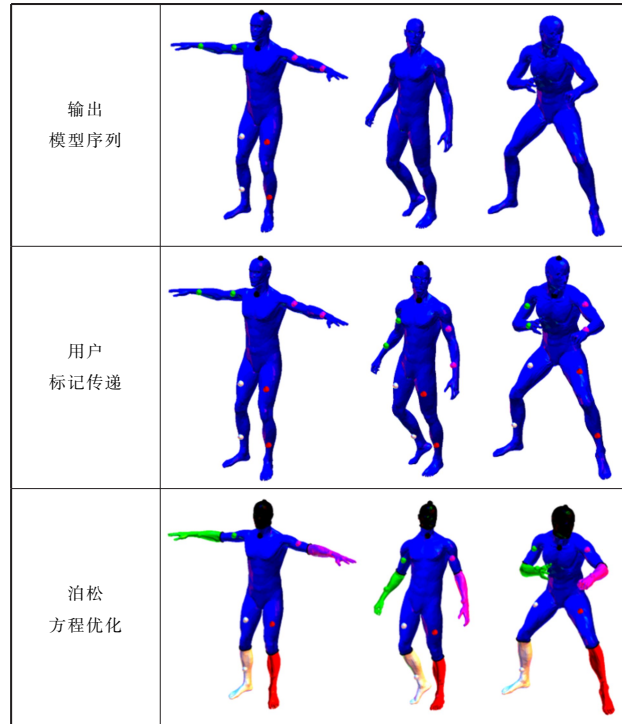


图2 本文算法的流程图

Fig. 2 Flow chart of proposed algorithm

在交互分割操作方式上,本文采用 DotScissor 提出的点交互。对于用户定义的点交互标记,等距映射可以把标记点传递到其他帧上,然后通过泊松方程完成边界优化。图2中每张图左边的第一个模型为基准模型。

另外,为了简化描述,这里给出分割的相关定义:对于一个三维网格,其顶点集合和面集合分别为 VS 和 TS , 顶点数量和面片数量分别为 N_v 和 N_f 。对于任何两个顶点 v_i 和 v_j , 其测地距离为 $g\{v_i, v_j\}$ 。分割目标是按照输入的标记将网格序列分割成语义不同的子部分 $S_1, S_2, \dots, S_m, \dots, S_n, S_K$ (K 是子部分的数量)。所有子部分满足以下条件:

$$\begin{cases} S_m \cap S_n = \varphi, m \neq n \\ S_1 \cup \dots \cup S_m \cup \dots \cup S_K = TS \end{cases} \quad (1)$$

2 标记传递

在上述算法框架上,需要分割的模型集合为:

$$\{O_1, O_2, \dots, O_i, \dots, O_j, \dots, O_K\}$$

假设用户对模型 O_i 定义了用户标记集合:

$$\{A_{i,1}, A_{i,2}, \dots, A_{i,m}, \dots, A_{i,N_i}\}$$

显然,对于被用户标记的模型 O_i 可以采用已有的分割算法来得到分割结果。但是,对于任意一个未被标记的三维模型 O_j , 需要得到其用户标记点集合:

$$\{A_{j,1}, A_{j,2}, \dots, A_{j,m}, \dots, A_{j,N_j}\}$$

因此,算法需要在任意两个模型之间建立对齐关系,从而实现约束标记点的自动传递。

$$\varphi: A_{j,m} = \varphi(A_{i,m})$$

从而在模型 O_j 中找到相应的标记点集合,通过交互式分割完成分割结果。

对于标记点传递,可以采用等距映射来完成。注意到对于等距映射,需要足够多的采样点分布在模型表面,才能保证算法的准确性。为此,需要对已有种子点进行扩展,形成采样点集合:

$$\{A_{j,1}, A_{j,2}, \dots, A_{j,m}, \dots, A_{j,N_j}\}$$

为此,算法以已有标记点为种子点,对源模型表面实现均匀采样,从而得到足够多的采样点用于构造等距映射。对于源数据模型 O_i , 其顶点集为 $\{v_1, v_2, v_3, \dots, v_n\}$, 这里 n 为顶点数目。对于模型网格顶点采样,先将网格表面所有顶点设为未访问状态,用户标记的顶点设为已访问状态。对于任意一个已标记的顶点 v_i , 把它周围在采样半径内的顶点 v_j 也设置为已经访问,即满足条件 $g(v_i, v_j) < r$ 。然后对于任一未被访问的顶点,把它加入到种子区域,并重复上述过程,从而形成一系列等面积分布的曲面片。这里, r 代表采样半径,如果半径 r 越小,则得到的采样点数目越少,标记点映射的精度也越高,但是标记点传递的耗时也更长。这里,采用如下的公式定义半径:

$$r = 0.17 \times \sqrt{\frac{A}{\pi}} \quad (2)$$

其中, A 是三维形状的面积。最后,在源模型表面选择一系列的采样点,为了保证点的均匀分布,任意两点之间的距离至少为 r 。通过以上得到的顶点就可以得到三维模型顶点的对齐结果。同理,对目标模型也进行同样的操作,从而得到采样点集合:

$$\{A_{j,1}, A_{j,2}, \dots, A_{j,m}, \dots, A_{j,N_j}\}$$

任意两个三维模型之间的匹配问题,可以转化为两个点

集之间的映射问题。设源模型和目标模型的采样点的集合分别为 $S = \{V_1, V_2, V_3, \dots, V_N\}$, $T = \{V_1, V_2, V_3, \dots, V_M\}$, 其中 $N < M$ 。

等距二分图思想是采用两个数据的采样点构造图 $G := (S \cup T, E)$, 其中边集合 E 的定义如下: 对于图 G 的任意两个顶点, 当且仅当两个顶点满足条件 $s_i \in S, t_m \in T$ 时才存在一条连接边, 其相似度为 $c(s_i, t_m)$ 。可以发现, 两个三维数据的采样点对应二分图的顶点集合, 来自于同一个三维数据的采样点不存在连接边。则对齐问题转化为集合 S 到 T 的映射 $\zeta: S \rightarrow T$, 可以通过式(3)的最大值得到最优对齐结果。

$$\arg \max_{(s_i, t_m) \in \zeta} \sum c(s_i, t_m) \quad (3)$$

本文根据两个模型的表面采样点构造二分图。其中, 对于任意一对顶点, 边代价的定义如下:

$$c(s_i, t_m) = \eta - \sum_{k=1}^K |g(p_{s_k}, s_i) - g(p_{t_k}, t_m)| \quad (4)$$

其中, $(p_{s_k}, p_{t_k}) (k=1, 2, \dots, K)$ 代表采用 SIFT 等特征匹配得到的约束锚点; η 是一个常量, 用于把欧氏距离转化为相似度。由于源模型与目标模型的采样点数目可能不等, 为了满足二分约束求解必须要求两边顶点数量相等的条件, 本文对采样点数量较少的一方进行了扩展, 得到完全二分图 G' 。其中 G' 采用如下的方法扩展得到: 在顶点数少的一侧添加顶点直到两侧顶点数目相等; 同时增加虚边(代价定义为 ∞ , 保证这些虚边不会包含在最优匹配结果中)。从而最终形成一个可用于约束求解的对齐代价矩阵: 矩阵的行表示源数据的采样点索引; 矩阵的列表示目标数据的采样点索引。矩阵中的任意一个元素代表两个采样点的相似度。

对于上述构造的对齐代价矩阵, 可以采用 Kuhn-Munkres 算法解决最大权重匹配问题, 从而得到最优对齐结果^[15]。Kuhn-Munkres 算法为每个顶点设置一个标号, 然后采用完备匹配完成求解。

图 3 给出了通过等距映射得到的标记点传递结果。可以发现, 通过等距映射在两个模型之间所形成的对齐结果, 用户标记点被有效地传递到其他模型上, 从而形成一个交互式分割问题。

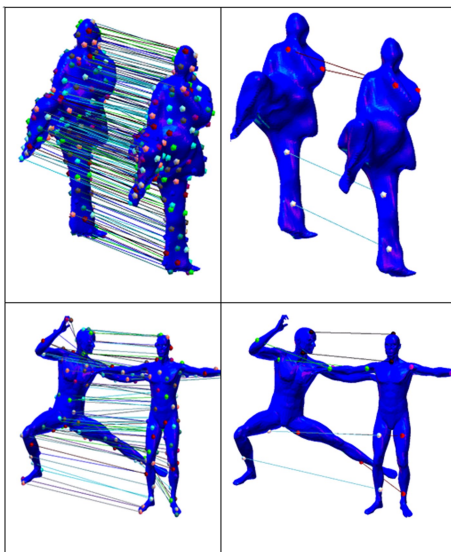


图 3 等距映射和标记传递

Fig. 3 Isometric mapping and label passing

3 边界优化

上述过程通过等距映射自动得到用户标记, 从而把模型

分割问题转化为标记点约束分割问题。但是存在一个主要问题: 分割边界粗糙并且不光滑, 需要进行优化。由于不同姿态下的三维模型具有等周长的特性, 本文采用了最小测地环作为优化目标, 同时增加形状约束来避免局部曲率噪声对分割边界的影响。图 3 给出了人体手臂分割边界优化前后的结果比较。对于交互式分割, 采用如下步骤。

(1) 局部邻域提取。对于模型上的约束标记点 $A_{j,m}$, 定义邻域半径 T_r 。同时定义最远点:

$$P_j = p_i, \max(\text{gd}(e_g, p_i)), p_i \in BR \quad (5)$$

其中, BR 为 e_g 所在的局部区域顶点集合。对于给定阈值参数 T_r , 可以得到用于分割优化的局部邻域。邻域中的顶点 p_i 满足如下条件:

$$p_i || |g(A_{j,m}, p_i)| \leq T_r \quad (6)$$

(2) 测地环提取。对于局部邻域中的顶点集合, 本文需要提取一系列测地环, 测地环采用泊松方程定义。对于邻域上的顶点, 根据约束标记可以分为上边界点集合 U 、下边界点集合 K , 形成如下的泊松方程。

$$\Delta \Phi = 0 \quad (7)$$

其中, Δ 为拉普拉斯算子。边界约束为 $\Phi(x) = 1, x \in U; \Phi(x) = 0, x \in K$ 。由于测地环具有最小测地距离和光滑特性, 因此可以通过最小二乘法来求解, 从而得到测地环集合 $\{I_1, \dots, I_i, \dots, I_z\}$, 其中 z 为分割线条数。

3) 最优分割线选择。在上述测地环的基础上, 需要通过优化求解得到最优测地环, 即分割边界。考虑到分割边界所对应的最优测地环通常在中间位置, 且测地距离最小, 因此定义最优目标函数为:

$$JU_i = e^{-\frac{(i-\frac{z}{2})^2}{2z^2}} \cdot \frac{\sum_k R(k)(2r_i - r_{i-k} - r_{i+k})}{\sum_k R(k)} \quad (8)$$

其中, $e^{-\frac{(i-\frac{z}{2})^2}{2z^2}}$ 用于约束最优分割线靠近局部邻域的中心; 第二项约束是在测地环集合中找到最短周长, 从而保证分割边界在凹区域。这里 k 是测地环索引, r_i 是测地环 I_i 的邻域半径, 定义为:

$$r_i = \frac{m_i}{2\pi}$$

其中, m_i 是测地环周长, $R(k)$ 是惩罚因子, 定义如下:

$$R(k) = \exp\left[-\frac{(k-1)^2}{8}\right] \quad (9)$$

4 实验分析

本文的实验环境为 Inter 酷睿 I5 处理器, 内存为 4 GB。实验中对多种序列的三维模型序列进行了分割, 并与其他算法的结果进行了比较, 以此保证本文算法能够适应不同类型的三维模型序列, 且表现出良好的性能。其中本文的实验数据来源分为两部分, 一部分来自于 TOSCA 数据库, 其他数据通过现实运动的数据重构而成。

4.1 标记匹配稳定性分析

本文算法采用等距匹配来得到标记在其他帧上的位置, 因此等距匹配的准确性直接决定了一致性分割的结果。注意到, 等距匹配采用测地距离作为约束条件。而测地距离在姿态变化下具有不变性。尽管这种匹配方法具有对称性问题, 但是这种对称性不影响分割。另一方面, 本文在标记匹配基础上进行局部边界优化。因此, 在匹配上允许有位置上的误

差。图 4 给出了不同序列的匹配结果。

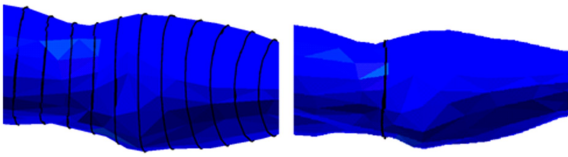


图 4 泊松方程等值线提取

Fig. 4 Isoline extraction of Poisson equation

使对于部分数据,由于等距匹配误差,使得标记匹配结果并没有到达预期位置。但是后续可以对泊松方程进行优化,使得分割结果收敛到关节凹位置。

4.2 三维网格序列的分割结果

为了验证算法的多样性和稳定性,我们采用了多种形状各异的模型序列,其中每种序列都包含不同帧下的姿态。图 5 给出了算法分割的结果。对于每个模型序列,首先在左边的源模型上添加交互标记,然后将标记依次映射到其他模型上,最后执行算法从而得到分割结果。

可以发现,等距匹配能够把标记点映射到其他帧上。即

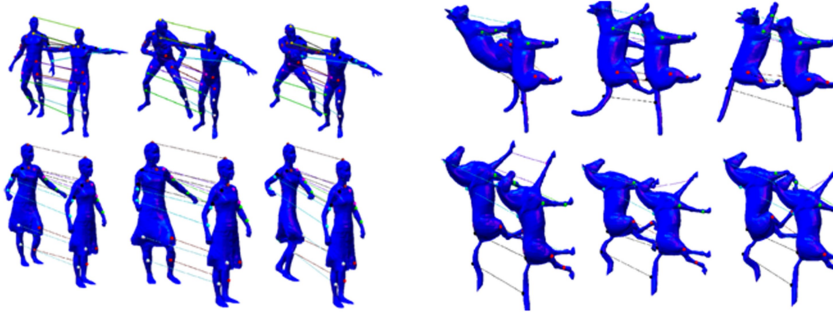


图 5 标记匹配稳定结果

Fig. 5 Stable results of label matching

4.3 分割质量比较分析

为了验证算法与其他算法的性能,选择了两种比较算法:一种是为了动态模型分割的增广轮廓方法(Augmented Silhouettes, AS)^[16]。该方法对三维动态模型在不同视点下的二维投影轮廓变化得到一致性分割结果;另一种方法

是用于单一模型分割的凸逼近方法(Approximate Convexity Analysis, ACA)^[17]。该方法首先把模型分割为一系列凸的子部分,然后通过非局部几何信号进行合并,从而得到最终的分割结果。不同三维动画序列的分割结果如图 6 所示。

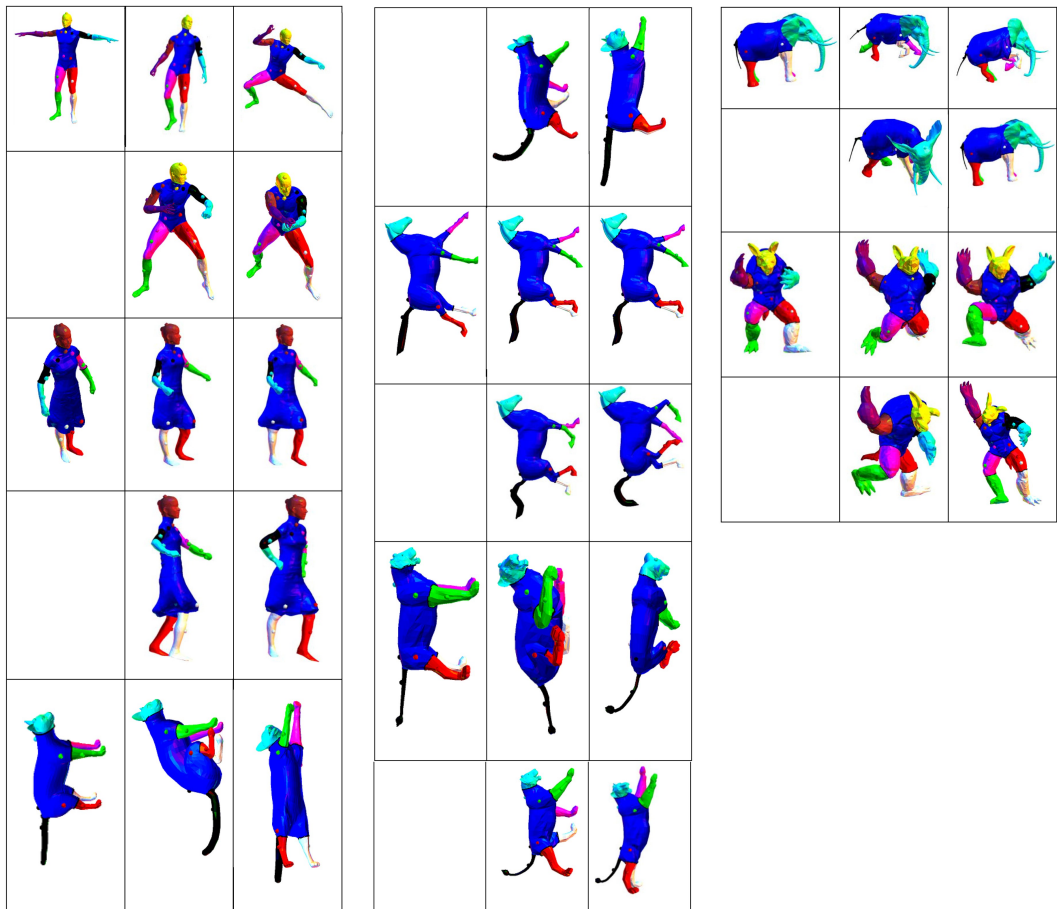


图 6 不同三维动画序列的分割结果

Fig. 6 Segmentation results of different 3D animation sequences

对于分割质量评价,本文采用海明距离评价方法(Hamming Distance, HD)。海明距离评价方法是由Chen等提出的分割基准评价方法^[18]。该方法将算法得到的分割结果(Computer-Generated, CG)和人工分割得到的基准分割结果(Human-Generated, HG)进行比较,分析算法分割结果CG和基准分割结果HG的差异性评估算法的优劣。如果CG和HG的相似度越高,则得到的分割结果越好。

HD是通过比较两种算法 S_1, S_2 在同一区域得到子部分的差异性来评估一种算法的优劣。假设 S_1 是待评估算法, S_2 是标准算法。 $S_1 = \{s_1^1, s_1^2, \dots, s_1^m\}$, $S_2 = \{s_2^1, s_2^2, \dots, s_2^n\}$,它们分别有 m, n 个分割区域。那么方向性的海明距离的公式就是:

$$DH(S_1, S_2) = \sum_i \|S_1 \setminus S_2^i\| \quad (10)$$

其中, \setminus 是差分算子, S_2^i 是 S_2 中对应 S_1^i 的分割区域, $\|x\|$ 表示区域 X 中所有网格的总面积。其中,带方向的海明距离分别为:

$$R_m(S_1, S_2) = \frac{DH(S_1, S_2)}{\|S\|}, R_f(S_1, S_2) = \frac{DH(S_2, S_1)}{\|S\|} \quad (11)$$

海明距离为:

$$HD(S_1, S_2) = \frac{1}{2}(R_m(S_1, S_2) + R_f(S_1, S_2)) \quad (12)$$

为了比较的一致性,实验中 S_2 是基准分割结果, S_1 是不同算法的分割结果,然后采用式(11)计算不同算法的分割质量。表1列出了3种算法对不同动态模型的分割准确率。可以发现,本文算法优于其他两种方法。特别是对于猫和马动态模型,本文算法分割有明显的改进,主要是AS方法采用二维轮廓进行分割是有局限性的,对于部分姿态,其二维轮廓由于遮挡等问题会明显影响分割效果。而ACA方法由于采用了凸分析,因此对于凸部分明显的Armdillo模型效果得到了较好的效果。但是由于该方法是针对单一模型的分割,没有考虑动态模型之间的相关性,因此总体效果要差于其他两种方法。

表1 不同算法的分割准确率

Table 1 Segmentation accuracy of different algorithms

动画序列	本文算法	AS方法	ACA方法
Cat	0.076376	0.17	0.19
Horse	0.083372	0.16	0.21
Elephant	0.11	0.13	0.23
Armdillo	0.20	0.34	0.17

4.4 算法的局限性分析

本文算法提出了三维模型序列的交互式分割框架。可以发现,对于模型序列的交互式分割,其算法精度主要取决于标记传递的稳定性。尽管等距映射在姿态变化下具有稳定性,但是等距映射是建立在测地距离基础上的,因此算法对拓扑变化是敏感的。图7给出了由于拓扑变化导致标记传递失败的例子。对于这一问题,在后续研究中可以考虑采用更为稳定的调谐距离进行改进。

另一方面,本文目前采用二分图匹配方法进行标记传递,这种算法需要足够多的采样点才能保证标记传递的稳定性。在后续研究中,考虑在模型序列中进行多帧标记,从而把标记点传递转化为关键帧约束下的标记点空间插值问题,避免复杂的采样和匹配问题。

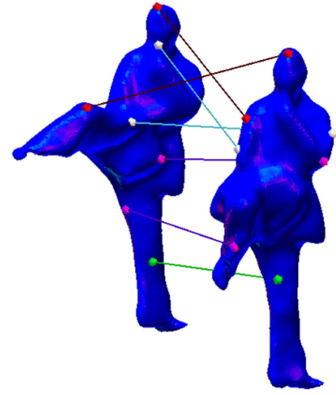


图7 拓扑变化导致等距映射失败

Fig. 7 Topology changes cause isometric mapping to fail

结束语 针对三维模型序列提出了点交互分割算法,该算法通过一帧进行标记,并将标记映射到其他帧上,从而对整个序列进行分割。实验结果表明,本文算法完全可以适应不同类型的网格序列,通过量化评价表明,其分割质量优于已有方法。

目前,本文算法并没有对分割质量差的结果进行进一步优化。在后续工作中,可以考虑建立分割的交互反馈框架,即对于分割质量差的结果,可以进一步插入用户交互的关键帧,使分割质量得到进一步的提升。

参考文献

- [1] FORSYTH D A, PONCE J. Computer vision: a modern approach[M]. Prentice Hall Professional Technical Reference, 2002.
- [2] ZHENG Q, SHARF A, TAGLIASACCHI A. Consensus skeleton for non-rigid space-time registration[J]. Computer Graphics Forum, 2010, 29(2): 635-644.
- [3] KRAEVOY V, SHEFFERA. Cross-parameterization and compatible remeshing of 3D models[J]. ACM Transactions on Graphics, 2004, 23(3): 861-869.
- [4] HAN L, CHENG Y. 3D model retrieval method combining skeleton structure and geometric details[J]. Computer applications and software, 2013, 30(11): 199-201.
- [5] MANGAN A P, WHITAKER R T. Partitioning 3D surface meshes using watershed segmentation[J]. IEEE Transactions on Visualization and Computer Graphics, 1999, 5(4): 308-321.
- [6] KATZ S, TAL A. Hierarchical Mesh Decomposition using Fuzzy Clustering and Cuts[J]. ACM Transactions on Graphics, 2003, 22(3): 954-961.
- [7] SHAPIRA L, SHALOM S, SHAMIR A, et al. Contextual part analogies in 3D objects[J]. International Journal of Computer Vision, 2010, 89(2/3): 309-326.
- [8] FANG Y, SUN M, KIM M, et al. Heat-mapping: A robust approach toward perceptually consistent mesh segmentation[C]// 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2011: 2145-2152.
- [9] SIDI O, VAN KAICK O, KLEIMANY, et al. Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering[M]. ACM, 2011.