

文章编号: 1671-0444(2018)04-0503-05

3D 卡通角色的快速骨架提取算法

舒 弋, 杨旭波

(上海交通大学 软件学院, 上海 200240)

摘要: 为了简化模型的骨骼绑定工作, 设计一套自动计算骨架的方法。该算法以模型的 2D 轮廓作为输入, 从轮廓中提取候选关节并构造 2D 骨架, 再映射到相应的 3D 模型上, 并采用神经网络算法增强骨架的通用性和质量。试验证明: 使用该算法可大幅度提升骨架提取的性能, 在移动设备上可以在 1 s 内完成整个骨架计算流程, 且其精度损失基本可以忽略不计; 该算法可以处理多样化的卡通角色, 既能帮助普通用户完成骨骼绑定工作, 也可以提高专业设计师的工作效率。

关键词: 3D 卡通模型; 骨架提取; 骨骼绑定; 2D 轮廓图

中图分类号: TP 301.6 **文献标志码:** A

A Fast Skeleton Extraction Algorithm for 3D Cartoon Character

SHU Yi, YANG Xubo

(School of Software, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: To simplify the skeleton-binding procedure, an algorithm which can automatically extract skeleton from 3D cartoon model was designed. The algorithm used a 2D outline image as the input where the candidate joints were extracted from, and then the 2D skeleton was constructed. Finally, the skeleton was mapped to the 3D model. A neural network was used to improve the generality and quality of the skeleton. It is demonstrated that the algorithm greatly improves the skeleton extraction performance and maintains great quality. It can extract the skeleton in 1 s on mobile devices. This algorithm can deal with different cartoon characters and help the users or professional designers finish the skeleton-binding work more efficiently.

Key words: 3D cartoon model; skeleton extraction; skeleton-binding; 2D outline image

3D 模型需要完成骨骼绑定后才能添加相应的骨骼动画, 这是一个相对专业且繁琐的工作。因此, 本文针对卡通角色提出一种全自动的骨架生成算法, 以帮助设计师和非专业用户快速地完成骨骼绑定的工作。

本算法以一张卡通角色的图像作为输入, 将其转换为轮廓图, 从中提取出骨架后再转换到三维空间, 然后嵌入相应的 3D 模型中。该 3D 模型同样来自于输入

图像, 使用了自动建模方法 MagicToon^[1] 进行生成。

由于卡通角色本身的夸张性, 因此降维造成的略微变形基本可以忽略不计, 而骨架提取的性能却有大幅度的提升, 在降低计算量的同时将整个提取过程的耗时压缩到了 1 s 以内。

常见的骨架提取算法的核心思想是通过迭代的区域收缩将 2D 图像或 3D 模型收缩到曲线。对于 2D 图像, 这类方法被称为图像细化^[2]。对于 3D 模

收稿日期: 2017-11-14

基金项目: 国家自然科学基金资助项目(61173105, 61373085); 国家高技术研究发展计划资助项目(2015AA016404)

作者简介: 舒 弋(1993—), 女, 浙江嵊州人, 硕士研究生, 研究方向为软件工程数字艺术, E-mail: ez0247@163.com

杨旭波(联系人), 男, 教授, E-mail: yangxubo@sjtu.edu.cn

型,常见的方法有:其一,体素化方法^[3],即将模型分割成一个个体素,不断地去除边界体素直到成为线为止,其实现简单,然而鲁棒性较差;其二,基于距离场的方法^[4],即通常会构造一个中间面并从中提取极值点,以这些点作为候选点选出相应的骨架关节;其三,基于几何的方法,即以表面网格^[5]或点云^[6]作为输入。

上述方法普遍存在的缺点是耗时相对较长,无法满足实时应用的需求。因此,本文算法借鉴距离场方法中提取候选点的思想,并且使用 2D 图像而非 3D 模型作为输入,将耗时从原来的数十秒下降到平均 1 s 左右。

从图像中提取角色的骨架类似于人物的姿态估计问题。人物姿态估计是指估测照片中人的姿态,即找出照片人物的各个关节点。该领域已经有了多年的研究积累。由于卷积神经网络(CNN)在图像处理上表现出的强大潜力,近年来的研究者们往往会使用 CNN 提取关节点^[7],该方法可取得优良的效果,然而需要的计算量较大。

与人物照片不同,当前虽然存在一些手绘图像的数据集^[8],但不满足本文算法的需求。由于短时间内无法自主构建一个可用的卡通图像数据集,且需要满足移动设备的性能需求,因此本文算法没有采用 CNN,而是先从轮廓图中提取出候选关节点,以这些关节点作为神经网络的输入数据,完成初步分类后再进行下一步的提取。

本文针对卡通角色提出一种全新的骨架生成算法,兼具质量、通用性和高性能。该算法结合现有的自动建模算法^[1],可以自动从卡通图像中创建出相

应的三维模型,即输入一张 2D 卡通图像,由该卡通图像自动生成 3D 模型与骨架(如图 1 所示)。

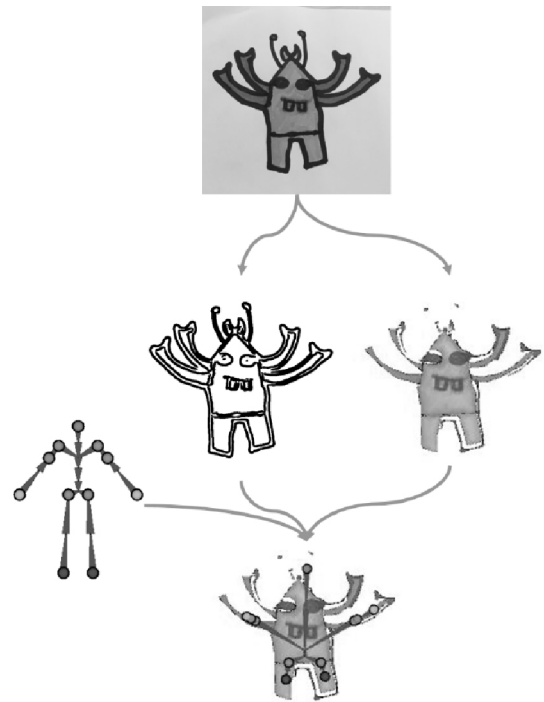


图 1 算法总览

Fig. 1 Overview of the algorithm

1 算法描述

整个算法流程可划分为 5 个阶段:(1)预处理阶段,生成轮廓图;(2)提取候选关节点;(3)神经网络初步分类;(4)识别并找出最终的关节点;(5)映射到三维空间,绑定到模型上。骨架提取流程如图 2 所示。

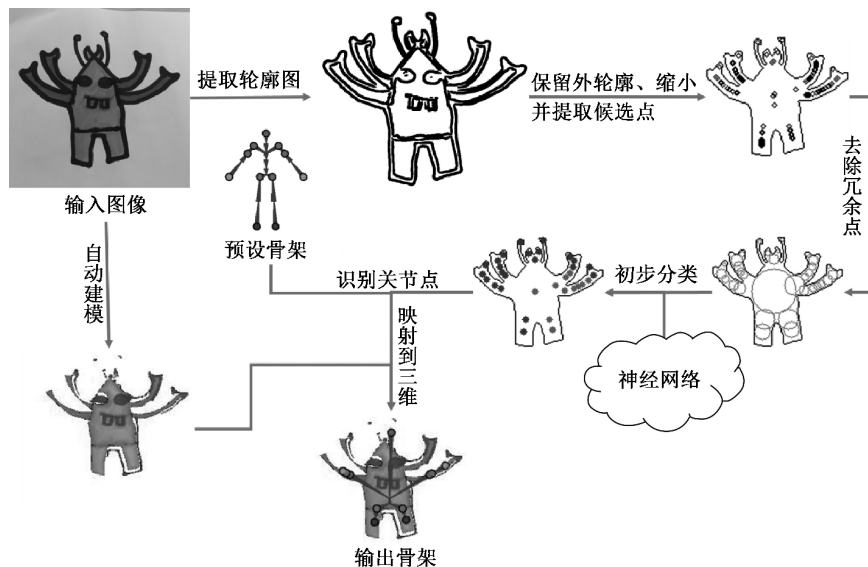


图 2 算法流程

Fig. 2 Pipeline of the algorithm

本文算法的输入是一张卡通角色图像,根据预设的人类骨架提取出相应关节点,而对应的3D卡通模型根据图像自动建模生成,最终的输出是一个已经完成骨骼绑定的3D模型。整个流程不需要用户的任何其他输入。

1.1 预处理阶段

预处理阶段的主要工作是将输入的卡通图像转换为轮廓图。

将图像转换到HSV(hue, saturation, value)颜色空间,根据明度(V)进行二值化,从而提取出线条,以水漫法区分前景和背景,在生成的轮廓图中仅保留前景的外轮廓(如图3所示)。

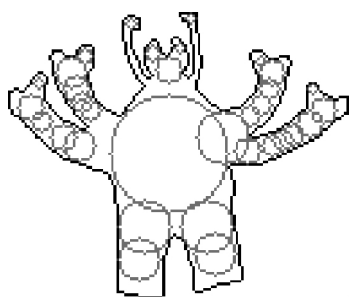


图3 轮廓图与候选点

Fig. 3 Contour and candidate points

为了保证生成的3D模型的精度,输入图像的大小为600像素×600像素。本文的骨架提取算法对图像精度的要求较低,因此在提取出前景部分后,将前景的包围盒区域依照如下的方法进行缩放:较长的边设为128像素,另一边等比例缩小。例如,原256像素×200像素的区域可缩小为128像素×100像素。

1.2 提取候选点

该阶段将依据距离从轮廓中提取出候选关节点。由于骨骼总是处于一个区域的中心位置,需要提取的候选点也应当位于各区域的中心线上。提取候选点的流程如算法1所示。

算法1:提取候选点

输入:轮廓图

输出:候选点集合C

$C = \{\}$

foreach P in Contour

 compute D // P 到轮廓距离

foreach P which $D > 0$

 if $D_P > D_{1-8}$ then

$P \rightarrow C$ // 加入候选集

 sort C // 降序排列

 foreach P in C // 去除冗余

 remove all P_i which $D_{P \rightarrow i} < D_P$

第一步,计算每个内部像素点到轮廓的距离,当一个点到轮廓的距离比周围8个点都要大时,将该点加入候选点集。第二步,按照距离大小对候选集进行排序。第三步,以降序遍历候选集,对于集中的点,以该点为圆心,以该点到轮廓的距离为半径画圆,所有位于圆内的候选点都会被移出集合。最终的候选集如图3所示。

候选点的提取质量将影响后续的结果,所有被选出的关节点必然来自候选集,如果候选集有遗漏,将会影响生成的骨架质量,且难以在后续的步骤中进行弥补。

1.3 神经网络分类

该阶段使用神经网络将候选点分为4个类别:头部、身体、手部、腿部。

神经网络的输入为候选点的 xy 坐标和到轮廓的距离 D ,网络结构为:3个输入神经元;2个隐含层,每层4个神经元;2个输出神经元以表示4个类型。需要注意的是,应当对每张图像的候选集独立做输入的归一化,而非统一进行操作。

训练阶段使用17张图片共计2124组输入作为训练集,经试验证明,该神经网络基本可以满足需求。

1.4 识别

该阶段将根据神经网络分类的结果进一步找出最终的关节点。

1.4.1 准备阶段

将候选点合并到几条线段。对于点 P_1 和 P_2 ,若两点之间的距离为 D_P ,点到轮廓的距离分别为 D_1 和 D_2 ,则当满足条件式(1)时,认为这两点在同一条线上。

$$D_P \leq D_1 + D_2 \quad (1)$$

线段划分完成后,依照式(2)计算每条线段的类别概率。

$$P_i = \frac{C_i}{C_{total}} \quad (2)$$

式中: P_i 为各个类别(头部、身体、手部、腿部)的概率; C_i 为当前线段上被标记为该类别点的个数; C_{total} 为线段上点的总数。

1.4.2 识别躯干

对于不同的卡通角色,四肢的形态变化较大,而躯干相对比较固定,所以先识别出头部和身体的关节点。

卡通角色的形态差异可能非常大,导致分类存在很大的不确定性,因此判定条件被放宽到只要头部/身体的概率大于0,即认为可能包含头部/身体的关节点。如算法2中所示,对于被标记为头部的区域,其中距离值最大的候选点将被选为头部的关

节点。对于被标记为身体的区域,选择中心点作为身体关节点。如果没有区域被标记为身体,则选择轮廓的中心作为身体关节点。如果没有区域被标记为头部,由于此时身体的中心关节点已经被选出,因此直接取距该点上方距离值最大的点作为头部点。有多个区域同时被标记为头部/身体的情况是非常罕见的,如果发生该情况,则选择平均距离值最大的区域作为标记区域。

算法 2: 识别头部和身体

输入: 被标记的区域集合 $S_{\text{head}}, S_{\text{body}}$

输出: 关节点 $J_{\text{head}}, J_{\text{body}}$

if $|S_{\text{head}}| > 1$ then

$S_{\text{head}} = \{L_{\text{head}}\}$ // 平均距离最大的区域

if $|S_{\text{head}}| = 1$ then

$J_{\text{head}} = \max \text{ distance point of } L_{\text{head}}$

// 同样方法获取 S_{body} 和 J_{body}

if $|S_{\text{body}}| = 0$ then

$J_{\text{body}} = \text{center point of contour}$

if $|S_{\text{head}}| = 0$ then

$J_{\text{head}} = \max \text{ distance point over } J_{\text{body}}$

由于身体的区域可能和四肢的区域有连接,在选出头部点和身体点后,去除位于两者之间的点,将其他点按同样的方法划分到新区域中并计算该区域的概率值。

1.4.3 识别四肢

识别四肢的方法和识别躯干的方法类似。当一个区域满足式(3)时,该区域被标记为四肢。

$$\begin{cases} P_{\text{手部}} + P_{\text{腿部}} \geq 0.5 \\ D_{\text{区域平均}} \leq 0.5 D_{\text{总体平均}} \end{cases} \quad (3)$$

式中: $D_{\text{区域平均}}$ 和 $D_{\text{总体平均}}$ 分别表示区域中各候选点到轮廓的距离平均值和所有候选点到轮廓的距离平均值。

对于所有被标记为四肢的线段区域,首先依据区域中点的个数进行排序。因为在预处理阶段去除了冗余点,所以点的个数多寡也能反映该区域的细长程度。其次,按降序遍历,根据找出的躯干关节点判断区域的可能位置。例如,若区域 A 位于中上部且在右边则是右手,如果此时右手区域还没有被找到,则将 A 标记为右手,取该线段区域两端的点作为手部和肩膀关节点,中心的点作为手肘关节点,若右手区域已经被找到,则忽略 A。

1.4.4 优化骨架

在某些特殊情况下,可能有部分的关节点无法被找到(如候选集有缺失),优化骨架阶段将指定这些未找到的关节点的位置以保证骨架的完整性。

由于信息的缺失,本阶段难以找到缺失点应在的位置,因此仅简单地根据已经找到的部分对预设

骨架进行相应的变形。

另外由于候选点过少导致关节点重合,此时可以通过对重合的点添加一个小幅度的位移将其错开。

1.5 映射与蒙皮

在识别阶段生成的骨架是二维的,加上深度值 Z 可以将其提升到三维。为简化计算,本文将骨架的深度确定在模型的中间。图像坐标与模型坐标的映射变换则由轮廓与模型的包围盒计算获得。

完成映射后,采用一个简单的基于距离场的蒙皮权重算法将骨骼绑定到 3D 模型上。至此,整个骨骼绑定流程已经完成。

2 试验结果

本文的目标是实现一个快速、通用且高质量的自动骨架生成算法,因此下文将通过 3 个方面对本算法进行分析。

2.1 性能分析

试验采用 iPhone 6 对 20 张卡通角色图像进行测试,其结果如表 1 所示。

表 1 骨架生成性能测试 (iPhone 6)

	预处理时间	提取候选点时间	总时间
最短	0.192	0.291	0.586
最长	0.413	0.430	1.012
平均	0.357	0.390	0.863

本算法的时间主要消耗在预处理阶段和提取候选点阶段,其他阶段耗时基本可忽略不计。表 1 中的总时间是指从输入图像到完成三维映射的耗时,而蒙皮阶段一般需要 1~2 s。

从表 1 可以看出,提取候选点的时间基本变化不大,该时间主要受轮廓区域大小的影响,轮廓在包围盒中所占区域的比例越大,则该阶段耗时越长。识别阶段的性能主要由候选集大小决定,同时也受到轮廓形状(区域个数)的影响。

骨架生成一般可在 1 s 内完成,即使加上蒙皮也只需要 2~3 s,可以认为当前算法已经能满足实际应用的性能需求。

2.2 质量分析

部分卡通角色骨架自动生成结果如图 4 所示。由图 4 可知,本文算法自动生成的骨架基本偏差不大。虽然有部分的误差,但由于人们对于卡通人物少许变形的接受程度较高,因此对于最终的动画呈现效果基本没有影响。

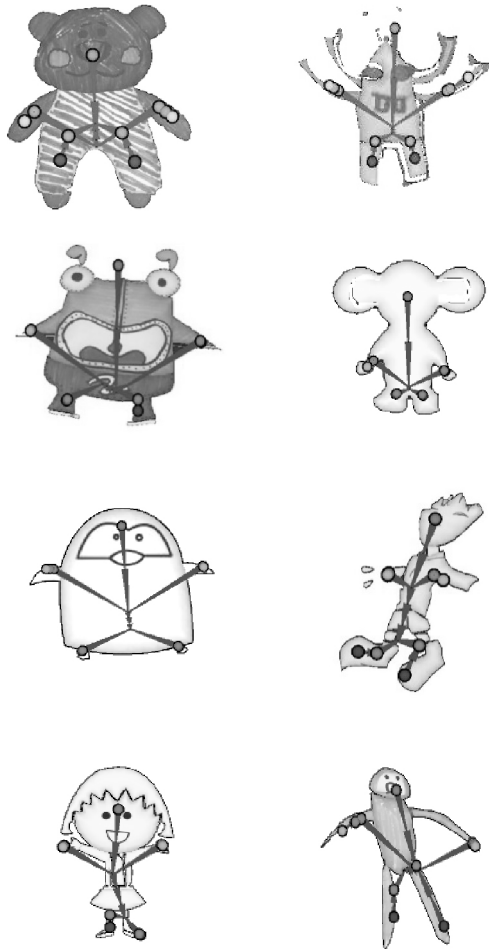


图4 部分卡通角色骨架自动生成结果

Fig. 4 Automatic extraction skeleton results of some cartoon characters

2.3 通用性分析

由于采用人形骨架,所以本文算法主要针对类人形的卡通角色。如果需要使用其他形态的骨架,则需要重新训练神经网络,并对识别阶段做略微修改。

3 总结与展望

3.1 总结

本文提出一种全新的针对卡通角色的快速骨架生成算法。该算法使用2D轮廓图作为输入,在移动设备上可以在1s内生成一个高质量的骨架,具有很好的通用性。其主要面向的用户是非专业的普通人群,也可以为专业设计师提供大致的骨架,从而减少工作量。

3.2 局限性与展望

本文的算法还存在一些不足之处与改进的空间,具体如下所述。

其一,提取候选集时以到轮廓距离的极大值点来表示各个区域中心线,此方法虽然试验证明可行,但缺少严格的数学证明,且在某些特殊情况下会产生遗漏,导致后续结果的缺失。在将来的工作中,应当通过数学推理获得更加严谨的公式,在增大候选集覆盖面的同时减少冗余点,提升质量与性能。

其二,本文仅使用了人类骨架,而该算法实际上可以使用不同形态的骨架,从而满足各种类型的卡通角色的需求。使用其他类型的骨架,只需要重新训练神经网络并对识别规则进行调整。在将来的工作中可以整合多种形态的骨架,进一步提升算法的通用性。

其三,最初期望的目标是由神经网络完成主体工作,其后再通过一些逻辑规则进行完善,但由于数据集的缺失,神经网络当前仅用于初步的分类。将来可以通过改进网络结构以及增大训练集来提高分类质量,甚至可以使用深度神经网络直接从轮廓图获取骨架。

参 考 文 献

- [1] FENG L, YANG X, XIAO S. MagicToon: A 2D-to-3D creative cartoon modeling system with mobile AR[C]// 2017 IEEE Virtual Reality. 2017: 195-204.
- [2] ZHANG T Y, SUEN C Y. A fast parallel algorithm for thinning digital patterns[J]. Communications of the ACM, 1984, 27(3): 236-239.
- [3] MA C M, WAN S Y, LEE J D. Three-dimensional topology preserving reduction on the 4-subfields[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(12): 1594-1605.
- [4] BARAN I, POPOVIC J. Automatic rigging and animation of 3D characters[J]. ACM Transactions on Graphics, 2007, 26(3): 72.
- [5] AU K C, TAI C L, CHU H K, et al. Skeleton extraction by mesh contraction[J]. ACM Transactions on Graphics, 2008, 27(3): 44.
- [6] TAGLIASACCHI A, ZHANG H, COHEN-OR D. Curve skeleton extraction from incomplete point cloud [J]. ACM Transactions on Graphics, 2009, 28(3): 1-9.
- [7] ZHOU X, ZHU M, LEONARDOS S, et al. Sparseness meets deepness: 3D Human pose estimation from monocular video [C]// IEEE Computer Vision and Pattern Recognition. 2016: 4966-4975.
- [8] SANGKLOY P, BURNELL N, HAM C, et al. The sketchy database: Learning to retrieve badly drawn bunnies[J]. ACM Transactions on Graphics, 2016, 35(4): 1-12.

(责任编辑:徐惠华)