

# 虚拟试衣系统服装图像匹配算法研究

耿 曼<sup>1,2</sup>, 钱国明<sup>1</sup>, 王 宇<sup>2</sup>

(1. 南京邮电大学, 江苏 南京 210000;

2. 中国科学院 上海高等研究院, 上海 200120)

**摘 要:**在目前虚拟试衣系统的研究中,服装图像的获取和处理一直是一个关键的研究问题,同时如何根据人体姿态匹配正确的姿态服装图像也一直是虚拟试衣系统中实现实时试衣的关键问题。对前人研究的以轮廓信息作为服装匹配的关键字进行了改进,结合图像识别算法去识别和返回人体的 24 个关节点的三维坐标,将骨骼信息关节点坐标作为匹配的关键字进行服装图像的转移匹配,然后使用轮廓信息提取算法进行人体轮廓的获取,使后续的转移算法利用当前图像帧轮廓检索数据库,以进行服装图像信息匹配,获取与当前帧最接近的图像帧。实验结果表明,使用人体关节点信息作为服装图像匹配的关键字可以更加快速和准确地检索到要匹配的关键帧服装图像。

**关键词:**虚拟试衣系统;服装转移算法;Kinect

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2017)01-0126-04

doi:10.3969/j.issn.1673-629X.2017.01.028

## Research on Clothing Image Matching of Virtual Fitting Room

GENG Man<sup>1,2</sup>, QIAN Guo-ming<sup>1</sup>, WANG Yu<sup>2</sup>

(1. Nanjing University of Posts and Telecommunications, Nanjing 210000, China;

2. Shanghai Academy of Sciences, Chinese Academy of Sciences, Shanghai 200120, China)

**Abstract:** At present, for virtual fitting room research, the acquisition and processing of clothing image has been a key research topics. At the same time, how to match the correct posture clothing image according to the human body posture has also been a key issue of realizing real-time fitting in virtual fitting system. The outline information as a clothing matching keywords of previous studies is improved, and 24 key points of 3d coordinates of human body are recognized and returned combined with image recognition algorithm, using the skeleton information as the matching key words to match clothing transfer image. The Kinect is also used to extract the human body contours, so that the subsequent transfer algorithm can use the current frame profile to retrieve database and match information, obtaining in touch with the current frame closest image frame. Experimental results show that the use of the human body joint point as matching keywords can be more fast and accurate retrieval to match the key frames of clothing image.

**Key words:** virtual fitting system; clothing transfer algorithm; Kinect

## 0 引 言

近年来,随着计算机图形处理技术和增强现实技术的快速发展和应用,虚拟试衣系统的研发也逐渐成熟起来。虚拟试衣系统是一种利用计算机技术将网上的服装真实地展示在试衣者身上的虚拟现实技术,对虚拟试衣系统的研究主要包含人体建模技术以及服装建模技术两方面的内容<sup>[1]</sup>。对于服装建模技术的研究,传统的虚拟试衣系统偏向于研究服装的模拟方法,如使用 2D 图像技术对服装的平面图片进行处理再贴

到人体模型上的贴图模式或者使用 3D 图像处理软件进行衣服图像三维建模再导入系统中<sup>[2]</sup>,然后转到人体身上再进行动画仿真的形式进行服装的模拟。但是在这些过程中服装材质模拟和服装三维重建问题一直是一个难点。

2011 年,Stefan 提出了一种基于图像的服装转移算法<sup>[3]</sup>,该算法避免了使用服装材质模拟和服装三维重建的问题<sup>[4-5]</sup>,转而使用其创建的离线服装数据库的方法,结合轮廓匹配法,就可以从数据库中找到与当

收稿日期:2015-10-09

修回日期:2016-02-25

网络出版时间:2017-01-04

基金项目:国家自然科学基金资助项目(61302093)

作者简介:耿 曼(1991-),女,硕士研究生,研究方向为智能信息处理;钱国明,教授,研究方向为无线通信技术和信号处理;王 宇,副研究员,研究方向为分布式计算、数据管理等。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20170104.1017.020.html>

前轮廓最匹配的衣服图像,进行服装匹配。文中对该方法进行了改进<sup>[6]</sup>,在 Stefan 的研究方法中使用人体轮廓特征点作为匹配的关键字进行轮廓匹配,即让模特穿上指定的服装进入经过特殊布置的图像信息采集房间,随意摆出各种姿势,从中采集 10 帧图像保存到本地,再对图像进行轮廓提取和抠图等后期处理。文中使用的方法是使用计算机图像处理获取硬件来替换原先需要使用的 8 台摄像机进行图像信息的获取,并运用图像识别算法计算出骨骼信息关节点坐标,将其作为服装转移算法的匹配关键字,再使用轮廓提取算法提取图像轮廓,避免了后期人工抠图带来的一些不可避免的瑕疵和巨大的工作量等问题。

### 1 离线数据库的创建

Stefan 提出的离线数据库的创建方法是在一个小房间里前后左右放置 10 个摄像头,然后让模特穿上服装在这个房间里来回地自由走动,摄像机从不同的角度对模特的影像进行采集,并将采集到的 10 幅图像保存到本地,然后再进行后期的抠图处理,得到每幅图像的人物轮廓信息,存储到数据库中。

文中方法是只使用一个 Kinect 采集图像设备(如图 1 所示),将 Kinect 放在显示器的上方,然后根据采集设备 Kinect 的放置高度和张角等信息,限制模特的最近和最远的活动距离。这种采集设备不需要一个密闭的房间,也不需要许多摄像头,使数据库的创建过程更加灵活,并且人体每幅图像的轮廓也使用 Kinect 来获取,这样可以节省后期的人工制作成本。

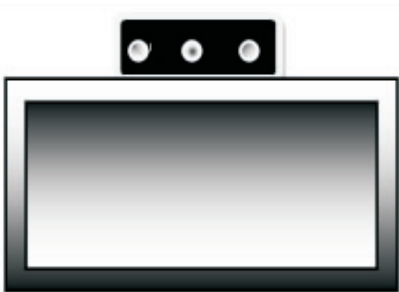


图 1 Kinect 图像采集设备

Kinect 可以为开发者提供彩色图像、深度信息和人体关节点等信息。Kinect 数据采集过程见图 2。

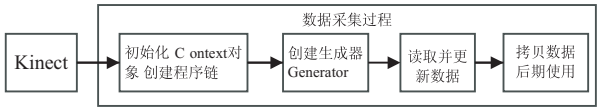


图 2 Kinect 图像采集过程

Kinect 上有一个红外发射器,当驱动 Kinect 时发射红外线扫描视场范围,红外接收器接收到物体反射回来的红外光,经过内部芯片处理后输出深度图像数据,而 RGB 图像传感器就直接采集到彩色图像数

据<sup>[7-8]</sup>。Kinect 数据采集的主要过程是首先初始化 Context 对象,创建程序链,初始化 Context 的目的是为了把 Context 作为 Create 函数的参数,这样就可以创建生成器了。Generator,depth Generator 是深度图像生成器,Imagegenerator 是彩色图像生成器。

另外,文中通过 Kinect SDK 人体骨骼追踪技术来获取人体骨骼数据<sup>[9]</sup>。骨骼追踪技术通过处理深度图像数据来建立人体各个关节点坐标,骨骼追踪能确定人体各个关节点的位置坐标,并且通过  $(x,y,z)$  三维坐标系来确定各个关节点的位置。其中, $z$  表示人体到图像感知系统的距离, $x$  与  $y$  分别为人体图像的宽度和高度。

Kinect 所能观测的人体关节点距离范围与其放置高度的关系,如表 1 所示。

表 1 Kinect 放置高度与人体距离关系表

Kinect 放置高度	最近观察距离	最远观察距离
0.8	1.762	4.167
0.9	1.831	4.184
1.0	1.988	4.119

经过实验验证表明,Kinect 放置的高度越高最近观察距离越远,但是相应的最远观察距离越小,对下半身的身体关节点的追踪能力就会下降。所以文中选择将 Kinect 放置在 0.9 m 处,这样最近和最远观察距离都能得到很好的中和,并且人相距 Kinect 的距离为 2.0 ~ 3.0 m,这样可以保证获取良好的图像质量。在 Kinect 的图像采集过程中除了要保存每一帧的图像外还要保存相应的骨骼位置信息。

编写人体骨骼数据捕获代码后,结合 Kinect 得到人体骨架图像,如图 3 所示。

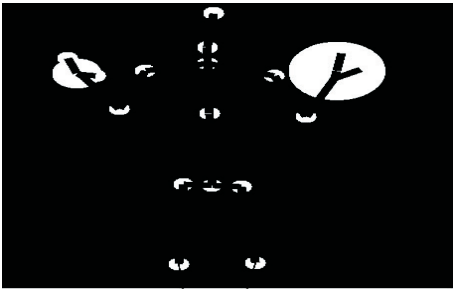


图 3 人体骨骼图像

### 2 基于骨骼信息匹配的服装图像的转移

使用开源的 OpenNI<sup>[10]</sup>来进行识别并返回人体的 24 个骨骼关节点的三维坐标。由这 24 个值组成的特征向量可以表示为 RT,其表达方式如下:

$$RT = [T_1, T_2, \cdots, T_{24}] \tag{1}$$

每一维的  $T$  均表示一个三维坐标:

$$T_i = [x_i, y_i, z_i], i \in [1, 24] \tag{2}$$

文中的算法流程主要为:

- (1) 捕获当前帧对应的图像, 通过 Kinect 获取骨骼位置信息, 构建特征向量  $RT$ 。
  - (2) 对数据库进行检索, 进行特征匹配。
  - (3) 获取目标图像, 并与当前帧进行贴合。
- 具体算法流程图如图 4 所示。

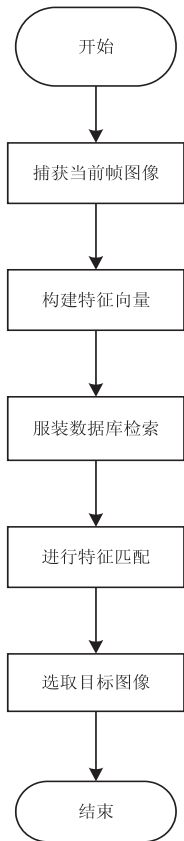


图 4 图像匹配算法流程

在整个算法流程中, 第二步是核心阶段, 在这一步中, 系统对数据库进行检索, 将得到的特征向量与数据库中的图像特征值进行匹配。文中采用距离测量的方法从数据库中选择图像, 对每一幅预处理图像进行距离计算得到距离值, 通过对距离  $F$  的控制和限定, 选出目标图像, 如式(3)所示:

$$F = \mu A(l_{i-1}, l_i) + B(l_i, i) + \lambda C(l_{i+1}, l_i) \quad (3)$$

其中,  $\lambda$  与  $\mu$  为阈值, 为常数, 具体需要通过实验来测定;  $i$  和  $l_i$  均为  $n$  维特征向量,  $n$  是经过主成分分析得到的特征点的数量,  $l_i$  是数据库中当前检索到的图像的特征向量,  $i$  是当前帧图像的特征向量;  $B(l_i, i)$  是  $i$  和  $l_i$  的欧氏距离, 具体如下:

$$i = [i_1, i_2, \dots, i_n] \quad (4)$$

$$l_i = [l_{i1}, l_{i2}, \dots, l_{in}] \quad (5)$$

$$B(l_i, i) =$$

$$\sqrt{(l_{i1} - i_1)^2 + (l_{i2} - i_2)^2 + \dots + (l_{in} - i_n)^2} = \sqrt{\sum_{j=1}^n (l_{ij} - i_j)^2} \quad (6)$$

文中使用欧氏距离作为距离测量的计算公式, 可以很好地实现各特征点之间的匹配功能, 并且为了能很好地体现出相邻帧的连贯性, 还使用了两个连贯性判别函数:  $A(l_{i-1}, l_i)$  和  $C(l_i, l_{i+1})$ 。

$$A(l_{i-1}, l_i) = \begin{cases} 0, i-1 \text{ 与 } i \text{ 为相邻帧} \\ 1, i-1 \text{ 与 } i \text{ 为非相邻帧} \end{cases} \quad (7)$$

$$C(l_i, l_{i+1}) = \begin{cases} 0, i \text{ 与 } i+1 \text{ 为相邻帧} \\ 1, i \text{ 与 } i+1 \text{ 为非相邻帧} \end{cases} \quad (8)$$

其中,  $A(l_{i-1}, l_i)$  和  $C(l_i, l_{i+1})$  的意义在于通过距离测量判别出前一帧的检索结果与当前数据库图像的关系和当前数据库图像与下一帧检索图像的关系, 通过判断前一帧的检索结果使得相邻帧的距离值小于非相邻帧, 从而拥有更大的机会成为目标图像, 而方法则是给非相邻帧加上阈值(文中实验值为 100)。用添加了连贯性判别函数的图像匹配公式, 使相邻两帧图像所匹配的正好是数据库中相邻的图像, 就可以解决短时间内由于服装的“闪动”或“突变”现象造成服装突然失常的变形现象。

### 3 人体轮廓的提取

提取出每幅图像的人体轮廓信息则是为了方便后续的转移算法利用当前帧所对应的轮廓信息来检索数据库, 然后经过信息匹配, 查找出与当前帧最接近的图像帧。目前轮廓提取方式主要有采取编写轮廓提取算法进行自动处理的方式, 这种方式效率高, 但通过算法自动提取的结果难免会有一些瑕疵; 此外, 还有另一种处理方式, 即通过手工处理获取轮廓剪影, 手工处理可以使用 Photoshop 等图像处理软件中的选区工具, 将人物的轮廓剪影选中并分离出来。但是这种方法后期的人工制作成本巨大。

为了获取到更好的人体轮廓图像同时也降低人工处理的成本, 文中使用 Kinect 进行人体轮廓的提取和处理。Kinect 具有 RGB 摄像头和红外线感应器, 可以获得所照射区域的数据信息<sup>[11]</sup>。并且, 人在照片中的信息可以通过距离信息进行识别, 这样距离图像就可以转化为距离数据的图像。假设对一个要进行人体轮廓提取的图像作这样的设定: 要提取的人体区域是彩色的, 其他区域是黑色的, 通过这个准则, 就可以得到只有人的区域图像, 这样就可以从这个图像中得到人体的轮廓, 如图 5 所示。

### 4 服装图像的转移和校准

经过前两部分的特征匹配过程后, 就可以从数据库中得到一帧需要的目标图像, 但是要将获取的目标图像准确贴合到当前图像中, 还需要经过对应比例上的校准才可以<sup>[12]</sup>。例如一件上衣, 如果要精确地确定

一件衣服的尺寸大小,需要测量肩宽、袖长、胸围、衣长,其中肩宽是最重要的上衣尺寸之一,如果肩宽不适合,衣服穿起来就不会像自己的衣服。文中通过肩宽来计算缩放比例  $scale$ ,如果要增加其他地方的尺寸缩放比例也可以按照这种方式来调整,公式如下:

$$scale = \frac{x_r - x_l}{x_{rl} - x_{ll}} \quad (9)$$

其中,  $x_r$ ,  $x_l$  分别表示当前图像的左右关节横坐标;  $x_{rl}$ ,  $x_{ll}$  分别表示搜索到的数据库中图像的左右关节的横坐标。



图5 人体轮廓图像

假设两幅图像的坐标轴以  $XOY$  坐标系为标准坐标系,则有以下几种情况发生:

(1) 当  $x_r = x_l$ ,  $x_{rl} = x_{ll}$  时,  $scale = 1$ , 这种情况下所获取的数据库图像不需要进行缩放就可以直接与获得的当前图像进行贴合。

(2) 当  $scale = 1$  时有可能出现一种特殊情况,那就是当目标图像的左右肩关节坐标正好相对于当前图像进行了同时的左移或者右移,那么即使  $scale$  为 1, 两幅图像仍然不能进行贴合,所以还应定义一个左右肩关节连线中点坐标横坐标  $x = 0$ , 其中

$$x = x_{rl} + x_{ll} \quad (10)$$

假如  $x = 0$ , 则表示两幅图像正好贴合;如果  $x > 0$ , 则表示目标图像相对于当前图像进行了右移,所以应按照得到的  $x$  的值左移  $x/2$  个坐标单位;同理,若  $x < 0$ , 则将目标图像右移  $x/2$  个单位。

(3) 当  $x_r = x_{rl}$ ,  $x_l = x_{ll}$  时,就需要对数据库中的图像进行缩放才能贴合到需要的当前图像中,这种情况下,如何缩放是主要的研究问题。当  $scale > 1$  时,表明目标图像的肩宽大于当前图像,要对目标图像进行放大,放大倍数按照给出的  $scale$  值进行。假如给出的  $scale$  值为 1.5, 则对目标图像放大 1.5 倍。当  $scale < 1$  时,就要对目标图像进行缩小,同样,缩小倍数也按照得到的  $scale$  来计算。

万方数据

## 5 结束语

文中在前人提出的基于轮廓特征的服装匹配算法的基础上,对匹配的关键字进行改进。采用体感设备进行编码计算得到骨骼关节信息作为匹配的关键字,与轮廓信息相比,具有更高的准确性,并降低了计算复杂度,且所需硬件的成本大大降低,实用性增强。

### 参考文献:

- [1] 王兆其. 虚拟人合成研究综述[J]. 中国科学院研究生院学报, 2000, 17(2): 89-98.
- [2] Hauswiesner S, Straka M, Reitmayr G. Image-based clothes transfer[C]//10th IEEE international symposium on mixed and augmented reality. Basel: IEEE, 2011: 169-172.
- [3] Nakamura R, Izutsu M, Hatakeyama S. Estimation method of clothes size for virtual fitting room with kinect sensor[C]//IEEE international conference on systems, man, and cybernetics. [s. l.]: IEEE, 2013: 3733-3738.
- [4] 刘炯宙, 李基拓, 陆国栋. 三维服装物理-几何实时混合模拟[J]. 计算机辅助设计与图形学学报, 2014, 26(12): 2244-2250.
- [5] Xu Weiwei, Umetani N, Chao Qianwen, et al. Sensitivity-optimized rigging for example-based real-time clothing synthesis[J]. ACM Transactions on Graphics, 2014, 33(4): 1-11.
- [6] Gabel M, Gilad-Bachrach R, Renshaw E, et al. Full body gait analysis with Kinect[C]//Annual international conference of the IEEE. [s. l.]: IEEE, 2012: 1964-1967.
- [7] Li Rong, Zou Kun, Xu Xiang, et al. Research of interactive 3D virtual fitting room on web environment[C]//Fourth international symposium on computational intelligence and design. Hangzhou: IEEE, 2011: 32-35.
- [8] Garcia C, Bessou N, Chadoeuf A, et al. Image processing design flow for virtual fitting room applications used in mobile devices[C]//IEEE international conference on electro/information technology. Indianapolis, IN: IEEE, 2012: 1-5.
- [9] Jia Nan, Zhou Fan, Wang Ruomei. Heat and moisture transfer computation model based on parallel simulation in clothed human[C]//5th international conference on digital home. [s. l.]: [s. n.], 2014: 412-418.
- [10] Hauswiesner S, Straka M, Reitmayr G. Virtual try-on through image-based rendering[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(9): 1552-1565.
- [11] Shimura K, Ando Y, Yoshimi T, et al. Research on person following system based on RGB-D features by autonomous robot with multi-Kinect sensor[C]//IEEE/SICE international symposium on system integration. [s. l.]: IEEE, 2014: 304-309.
- [12] 朱紫阳, 施霞萍, 陈家训. 基于部件分解的三维服装人体建模[J]. 计算机工程, 2009, 35(2): 271-273.