

中图分类号: TP391.9 文献标志码: A 文章编号: 1006-8961(2011)03-0449-05

论文索引信息: 张抗, 樊养余, 吕国云, 马元媛, 王毅: 基于机构学和肌肉模型的单视频驱动人脸情感表达仿真 [J]. 中国图象图形学报, 2011, 16(3): 449-453

基于机构学和肌肉模型的单视频驱动人脸情感表达仿真

张抗, 樊养余, 吕国云, 马元媛, 王毅

(西北工业大学电子信息学院, 西安 710072)

摘要: 描述一种3维人脸模型的单视频驱动方法。该方法在传统肌肉模型的基础上, 根据嘴部运动特性采用机构学原理建立嘴部运动控制模型, 通过视频图像序列跟踪得到特征点的运动规律曲线, 进而驱动眼部、嘴部及面部其他部分的网格点运动, 产生具有真实感的面部表情动作。仿真结果表明, 采用此方法可以得到逼真的人脸表情模拟动画。

关键词: 机构学; 肌肉模型; 情感表达; 人脸仿真

Drive facial animation by using single video based on the mechanism theory and muscle model

Zhang Kang, Fan Yangyu, Lü Guoyun, Ma Yuanyuan, Wang Yi

(Institute of Electronic Information Engineering, Northwestern Polytechnical University, Xi'an 710072 China)

Abstract: A method to drive 3D facial model using single video was presented. Based on the movement characteristics, the mouth was controlled by using a mechanism model. The motion trace of the feature points was acquired by tracking video image sequence. Then we drive the grid nodes on the facial model to obtain a realistic facial animation. The results show that this method could simulate the facial movements and expressions very well.

Keywords: theory of mechanism; muscle model; expressing feelings; face emulation

0 引言

人脸表情是人类情感传递的一个重要途径, 是人类进行相互交流的基础。3维人脸建模和表情仿真一直是计算机图形学领域研究的热点和难点问题, 具有真实感的3维人脸动画在人机交互、影视广告、游戏制作、视频会议等方面都有广泛的应用。因此, 对于人脸表情的研究具有重要的理论意义和实际应用价值。

自从1974年Parke^[1]首次采用多边形网格模型建立虚拟人脸, 实现了一种基于参数的人脸模型以来, 研究人员在该领域进行了大量的研究工作: Waters^[2]提出了一种物理肌肉模型, 采用多边形网格进行近似抽象的肌肉控制, 通过计算某一点的角位移和径向位移控制网格点伸缩, 取得了良好的效果; Lee等人^[3]提出了基于解剖学的多层人脸物理模型, 将人脸模型分为皮肤层(包括表皮、真皮、皮下组织和筋膜)、肌肉层和骨骼层, 通过计算影响人脸运动中的多种力, 驱动模型运动, 该模型的运动真

收稿日期: 2009-06-08; 修回日期: 2009-08-30

基金项目: 国家高技术研究发展计划(863)项目(2007AA01Z324)。

第一作者简介: 张抗(1984—), 男, 西北工业大学信号与信息处理专业硕士研究生, 主要研究方向为虚拟现实、图像处理、模式识别。E-mail: m4-szk@163.com。

实感较好,但计算量较大,难以满足实时性要求; Pighin 等人^[4-5]提出了一个基于图像的真实感人脸建模和动画系统,该系统通过不同角度的人脸图像和一个一般人脸模型交互地建立特定的人脸几何模型,然后再通过各个角度的图像合成该几何模型的纹理,人脸动画通过关键帧技术实现人脸模型在 3 维空间上的变形;姜大龙等人^[6]通过改进 MPEG-4 中已知 FAP 值和 FaceDefTables,计算 FAP 控制区内点的坐标的算法,实现了一种以 FAP 作为驱动参数的 3 维人脸动画;张青山等人^[7]通过将人脸模型划分为若干个运动相对独立的功能区,利用基于加权狄里克利自由变形 DFFD (Dirichlet free-form deformation) 和刚体运动模拟的混合技术模拟功能区的运动,同时,通过交叉的运动控制点模拟功能区之间运动的相互影响,实现了一种具有高度真实感和良好计算性能的人脸动画仿真算法;孔德慧等人^[8]通过采用参数模型和肌肉模型相结合的方法控制模型网格点的变化,生成人脸动画,同时引入人脸功能区的划分并定义了人脸动画表,可以在 FAP 流的驱动下生成理想的 3 维人脸动画。

以往的方法主要通过几何模型或肌肉模型来驱动人面部运动,只考虑了模型网格点的运动量或肌肉的收缩量,对网格点的运动速度没有过多考虑,本文在传统肌肉模型的基础上,结合机构学原理^[9]提出了一种新的面部动画仿真算法,该方法同时考虑了面部运动时肌肉的收缩量和各点运动速度与运动方式。文中所述的方法已在计算机上进行了模拟,实验证明该方法建模简单,物理概念清楚,易于驱动,能够模拟真实的人眼、嘴部运动,具有较好的仿真效果。

1 3 维人脸动画模型

1.1 肌肉模型

肌肉模型根据人体解剖学原理,最初由 Waters^[1]提出,采用多边形网格进行近似抽象的肌肉控制,通过计算某一点的角位移和径向位移控制网格点伸缩,取得了良好的效果。人面部肌肉主要包括提上唇肌、颧大肌、笑肌、颈阔肌、降口角肌、颊肌、降下唇肌、咬肌、口轮匝肌和颊肌,其中 7 种对面部表情影响较大,分别为提上唇肌、颧大肌、笑肌、降口角肌、降下唇肌、咬肌和口轮匝肌。根据肌肉神经

与肌肉被拉伸或压缩方向之间的关系,可以将肌肉分为窄线性肌(提上唇肌,颧大肌,笑肌,降口角肌,咬肌)、括约肌(口轮匝肌)、宽线性肌(降下唇肌),其中加入一个虚拟肌肉——上唇方肌,它和提上唇肌共同作用于上提嘴唇的动作。

1.1.1 窄线性肌模拟

窄线性肌又称为向量肌,即此种肌肉可用一个向量表示。根据解剖学原理,线性肌的起点固定,附着在骨头上,止点嵌入皮肤下层,随着肌肉收缩向起点点方向运动。图 1 表示了 2 维空间中的线性肌矢量模型。

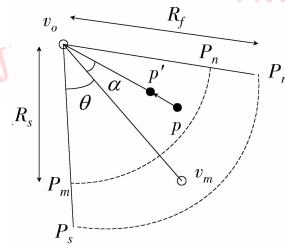


图 1 线性肌示意图

Fig. 1 Schematic diagram of the linear muscle

其中 v_o , v_m 分别代表肌肉的起点与止点, R_s 和 R_f 代表了起点和止点的衰减半径,扇形区域 $v_o P_s P_r$ 是肌肉的影响区域,在影响区域内的任意一点 p 沿向量 pv_o 朝向 v_o 运动到 p' 点,则

$$p' = p + k \times \alpha_g \times r \times \frac{pv_o}{d_{pv_o}} \times \Delta l \quad (1)$$

式中, k 是肌肉弹性系数, $d_{pv_o} = \|pv_o\|$, 代表 p 点到起点 v_o 的距离, α_g 是角度因子, r 是径向位移衰减因子, Δl 是肌肉的收缩量, α_g 和 r 分别定义如下

$$\alpha_g = \cos\left(\frac{\alpha}{\theta} \times \frac{\pi}{2}\right) \quad (2)$$

此处 α 是矢量 pv_o 与 $v_o v_m$ 间的夹角, θ 是影响区域的边界矢量与 $v_o v_m$ 之间的夹角。 pv_o 越靠近扇形区域边界 ($v_o P_r$, $v_o P_s$), α_g 值越小, pv_o 越靠近肌肉矢量 $v_o v_m$, α_g 值越大。

$$r = \begin{cases} \cos\left(\left(1 - \frac{d_{pv_o}}{R_s}\right) \times \frac{\pi}{2}\right) & d_{pv_o} \leq R_s \\ \cos\left(\frac{d_{pv_o} - R_s}{R_f - R_s} \times \frac{\pi}{2}\right) & R_s < d_{pv_o} \leq R_f \end{cases} \quad (3)$$

在扇形控制区域 $v_o P_s P_r$ 内,当 $p \in v_o P_m P_n$ 时,径向位移因子 r 随 d_{pv_o} 的增大由 0 开始逐渐增大,当 p

点位于弧线 $P_m P_n$ 附近 ($d_{pv_o} = R_s$) 时, r 达到最大, 当 $p \in P_m P_s P_r P_n$ 时, r 随 d_{pv_o} 的增大逐渐减小, 在弧线 $P_s P_r$ 附近减小到 0。

1.1.2 宽线性肌及括约肌的模拟

宽线性肌,括约肌的模拟与窄线性肌模拟类似, 见参考文献[10]。

1.2 机构模型

由于真实人嘴部张合运动并不是简单的匀速圆周运动而是一个变角速度的转动,故当虚拟人嘴部张合的时候,可以用空间杆机构模拟。

图2所示的摆杆机构运动简图,其中 OA 模拟虚拟人的上颚, OB 为下颚, α 为虚拟人嘴部张开最大的角度, β 为当前嘴部张开角度, BC 为空间假想杆且围绕 C 点旋转,易知当 BC 旋转到与 PC 重合时,虚拟人嘴部为闭合状态,当 $\beta = 2\alpha$ 为嘴部所能张开的最大角度。由于转动杆 BC 以周期为 2π 做圆周运动,故考虑一个周期。在 t 时刻, $\theta_2 = \omega t$, 由余旋定理

$$OB' = \sqrt{OC^2 + BC^2 - 2 \times OC \times BC \times \cos\left(\frac{\pi}{2} - \alpha + \theta_2\right)} \quad (4)$$

$$\text{则 } \theta_1 = \arccos\left(\frac{OB'^2 + OC^2 - BC^2}{2 \times OB' \times OC}\right) \quad (5)$$

$$\beta = \alpha - \theta_1 \quad (6)$$

对于人嘴部非连续的张合运动,可根据 FDP 特征点取上下嘴皮中央的点,对这两点进行跟踪,以它们的距离为纵坐标,时间为横坐标,进行曲线拟合,然后归一化。对曲线求导,当导数符号改变的时候假想杆的角速度 ω 取 $-\omega$,嘴部张合的速度由 ω 大小控制。

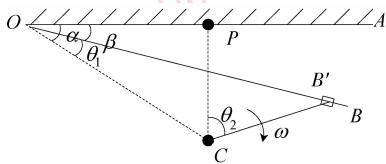


图2 空间杆机构模型

Fig.2 Space-bar linkage model

对于人嘴部连续的张合运动,模型更加简单,只需给定一个假想转动杆的角速度 ω ,本实验取 $\omega = \frac{2}{3}\pi$ 可以较逼真地模拟虚拟人嘴部连续的张合运动。

2 单视频驱动人脸的情感表达

根据解剖学及大量视频序列的分析,面部肌肉运动和眼部肌肉运动是两个相对独立的运动,故分别驱动面部肌肉和眼部肌肉的运动,可以获得逼真的人脸情感表达。

2.1 坐标系及特征点选取与跟踪

在视频截取的人面部头像中,以人物的鼻尖为基准,建立局部坐标系 $X'Y'Z'$,水平方向为 X 轴,垂直方向为 Y 轴, Z 轴垂直于 XOY 平面向外,3个轴方向符合右手法则。

结合 MPEG-4 中 84 个 FDP 特征点,经过考虑选取其中各个感情表达影响较大的 14 个特征点 P_i ($1 \leq i \leq 10$),分别为上嘴唇点 P_1 ,下嘴唇点 P_2 ,左(右)嘴角点 $P_3 P_4$,左(右)上眼皮点 $P_5 P_6$,左(右)下眼皮点 $P_7 P_8$,左(右)眼角的点 $P_9 P_{10}$,左(右)眉毛的中点 $P_{11} P_{12}$,左(右)眉毛的外角点 $P_{13} P_{14}$,如图3。

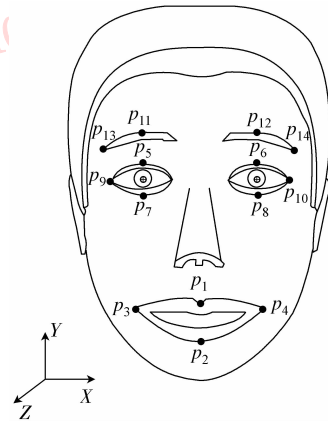


图3 人脸取点示意图

Fig.3 Face take point diagram

接下来对一段人脸视频逐帧截取图像,跟踪每一帧图像中的特征点 P_i ($1 \leq i \leq 14$) 在坐标系 $X'Y'Z'$ 中的平面 $X'O'Y'$ 中的坐标值,第 j 帧特征点 P_i ($1 \leq i \leq 14$) 的坐标值为 p_{ij} ($1 \leq i \leq 14, 1 \leq j \leq n$),其中 n 为图像总帧数。以视频段中的第 1 帧的无表情标准视频为基准,其余帧相对第一帧特征点的位移为 $d_{ij} = p_{ij} - p_{i0}$ 。特别的,对于 $P_1 P_2, P_3 P_7, P_6 P_8$ 这 3 对特征点,取 $A_{1j} = |p_{1j} - p_{2j}|, A_{2j} = |p_{5j} - p_{7j}|, A_{3j} = |p_{6j} - p_{8j}|$,令 $B_{mj} = |A_{mj} - A_{m1}|$ ($1 \leq m \leq 3$) 最后分别对 d_{ij} 和 B_{mj} 归一化,得到 d'_{ij}, B'_{mj} 。

2.2 生成感情表达曲线

在得到了 $d'_{ij} (1 \leq i \leq 14, 1 \leq j \leq n)$, 及 $B'_{mj} (1 \leq m \leq 3, 1 \leq j \leq n)$ 后, 以帧数 j 为横坐标, B'_{mj} 为纵坐标, 描绘出这些离散点集的分布图, 进而使用最小二乘法进行曲线拟合, 得到 3 条曲线 $l_m (1 \leq m \leq 3)$, 这 3 条曲线分别驱动嘴部张合运动和左右眼部的张闭运动。

第 j 帧嘴角和眼角特征点的归一化位移为 $d'_{ij} (i = 3, 4, 9, 10, 1 \leq j \leq n)$,

$$\text{令 } F_j = \frac{d'_{3j} + \alpha \times d'_{4j} + \beta \times d'_{9j} + \gamma \times d'_{10j}}{4}, \text{ 其中 } \alpha, \beta,$$

γ 为加权系数。

类似的, 以帧数 j 为横坐标, F_j 为纵坐标, 得到曲线 l_4 用以驱动人脸面部肌肉运动。

同理, 第 j 帧左右眉毛中间特征点的归一化位移分别为 $d'_{11j}, d'_{12j}, d'_{13j}, d'_{14j}$ 其中 $(1 \leq j \leq n)$, 对 $d'_{11j}, d'_{12j}, d'_{13j}, d'_{14j}$ 进行加权处理然后曲线拟合得到曲线 l_5 以驱动人眼部周围肌肉运动。

最终, 用上述 5 条曲线驱动真实感人脸的情感表达, 得到与视频序列基本一致的图像序列。

3 仿真结果

为了验证上述方法的有效性, 在 PC 机上进行了模拟, 操作系统为 WindowsXP, 所用编程软件为 VS2005, 驱动视频为一段 6.24 s 的大笑视频, 系统模型采用了机构模型和 Waters 肌肉模型相结合的方式进行了模拟。表 1 为每隔 5 帧采样嘴唇中间距离跟踪后处理结果。

表 1 部分采样数据示例表

Tab. 1 Examples of some of the sampling data table

帧号	1	6	11	16	21	...
B_{mj}	0	1	69	93	96	...
B'_{mj}	0	0.010 3	0.711 3	0.958 8	0.989 7	...

图 4 为一副面部情感表达曲线图, 图 5 为嘴部张合曲线, 所用的拟合函数为 15 阶多项式进行曲线拟合, 出现函数小于 0 或大于 1 的情况令函数值为 0 和 1。面部感情表达曲线加权系数 $\alpha = 1, \beta = \gamma$ 。图 6 为对驱动视频前 30 帧共 6 幅图像视频仿真结果, 真实视频和仿真结果上下对应。

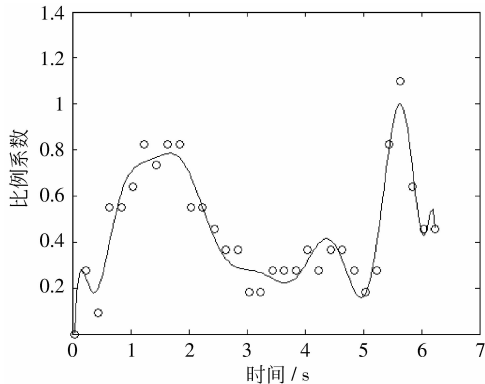


图 4 面部情感表达曲线图

Fig. 4 Curve of emotional facial expression

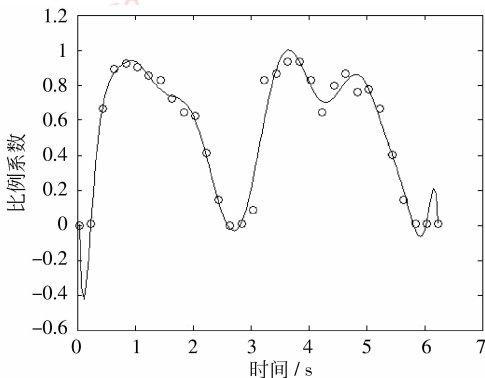


图 5 嘴部张合曲线图

Fig. 5 Curve of mouth open-closed

4 结论

描述了一种基于机构学和肌肉模型的单视频驱动人脸情感表达的方法, 该方法基于传统的肌肉模型并结合机构学原理进行人面部运动的模拟, 面部肌肉分别由提上唇肌、颧大肌、笑肌、降口角肌、降下唇肌等 7 条肌肉控制, 嘴部张合运动由机构模型进行模拟, 通过控制肌肉的收缩量 Δl_i , 和摆杆机构转动杆运动角速度 ω , 控制面部各点的位移量与嘴部的张合运动, 实现人面部运动的模拟。通过对视频中人头像上特征点 P_i 的跟踪, 可以生成 4 条曲线, 从而实现单视频驱动人面部情感的表达。

该方法已经在计算机上进行模拟, 从实验结果可以看到, 通过传递 16×5 个多项式系数就可以比较完美的对真实视频进行模拟。特别地, 对嘴部张合进行模拟时, 由于使用机构模型, 使得表情更加逼真, 无呆滞感。

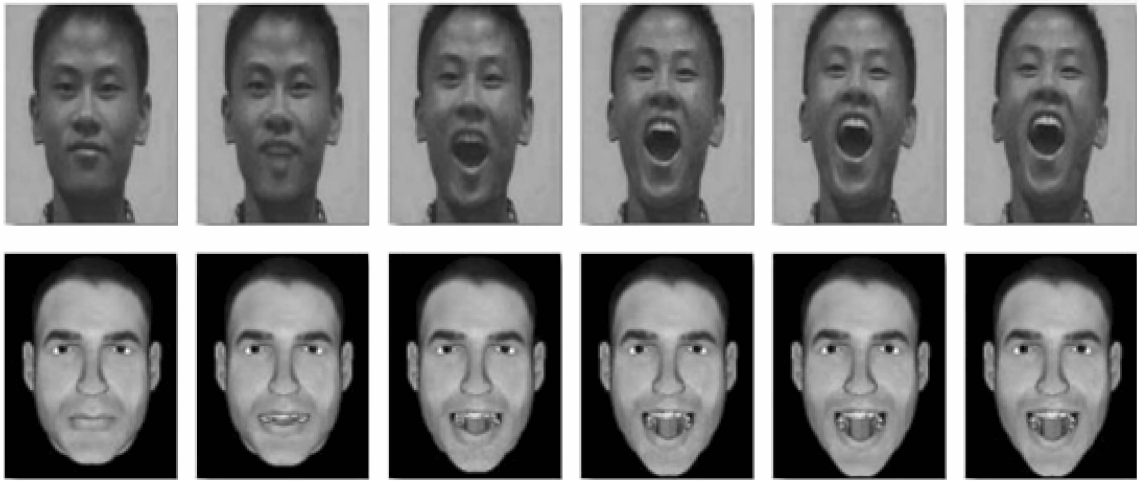


图6 仿真结果

Fig.6 Simulation results

更细腻的情感表达,如跟踪更多的特征点,对情感表达曲线进行更好的综合,将是下一步研究的重点。随着计算机技术的发展和虚拟人脸技术研究的深入,真实感人脸运动及情感表达必将有着更加广阔的发展前景。

参考文献 (References)

- [1] Parke F I. Computer Generated Animation of Faces [D]. Salt Lake City: University of Utah, 1972.
- [2] Waters K. A muscle model for animation three-dimensional facial expressions [J]. Computer Graphics, 1987, 21(4):17-24.
- [3] Lee Y C, Terzopoulos D, Waters K. Realistic modeling for facial animation [C]//Proceedings of the SIGGRAPH'95. Los Angeles, USA: ACM Press, 1995: 55-62.
- [4] Pighin F, Hecker J, Lischinski D, et al. Synthesizing realistic facial expressions from photographs [C]//Proceedings of the SIGGRAPH'98. Los Angeles, USA: ACM Press, 1998: 75-84.
- [5] Joshi P, Tien W C, Desbrun M, et al. Learning controls for blend shape based realistic facial animation [C]//Proceedings of the ACM SIGGRAPH Symposium on Computer Animation. Los Angeles, USA: Eurographics Assoc, 2003: 35-42.
- [6] Li Zhiguo, Jiang Dalong, Gao Wen, et al. An automatic MPEG-4 based arbitrary topology 3D facial animation system [J]. Journal of Computer-Aided Design & Computer Graphics, 2003, 15(07):863-869. [李治国,姜大龙,高文,等.全自动的基于MPEG-4的任意拓扑结构人脸动画 [J].计算机辅助设计与图形学学报, 2003, 15(07):863-869.]
- [7] Zhang Qingshan, Chen Guoliang. Realistic three-dimension facial animation [J]. Journal of Software, 2003, 14(03):643-650. [张青山,陈国良.具有真实感的三维人脸动画 [J].软件学报, 2003, 14(03):643-650.]
- [8] Jiang Xiufeng, Pu Xiaorong, Zhang Yi. Realization of 3-D facial animation based on MPEG-4 [J]. Journal of University of Electronic Science and Technology of China, 2007, 36(3):569-572. [蒋秀凤,蒲晓蓉,章毅.基于MPEG-4的三维人脸动画 [J].电子科技大学学报, 2007, 36(3):569-572.]
- [9] Sun Huan, Chen Zuomo, Ge Wenjie. Theory of Mechanisms [M]. 5th ed. Beijing: Higher Education Press, 1996: 137-186. [孙桓,陈作模,葛文杰.机械原理 [M].第五版.北京:高等教育出版社, 1996: 137-186.]
- [10] Zeng Dan, Cheng Yimin, Ge Shiming, et al. 3D muscle-controlled eye model [J]. Journal of Computer-Aided Design & Computer Graphics, 2006, 18(11): 1710-1716. [曾丹,程义民,葛仕明,等.人眼3D肌肉模型控制 [J].计算机辅助设计与图形学学报, 2006, 18(11): 1710-1716.]