

约束条件下的人脸五官替换算法

高 岩¹⁾ 许建中²⁾ 王长波¹⁾ 邱兆文³⁾ 马利庄⁴⁾

¹⁾ (华东师范大学软件学院, 上海 200063) ²⁾ (南昌陆军学院, 南昌 330000) ³⁾ (东北林业学院计算机系, 哈尔滨 150040)

⁴⁾ (上海交通大学计算机系, 上海 221000)

摘 要 针对医学美容以及数字娱乐领域中的特定应用, 提出了基于图像的人脸五官替换算法。该算法将人脸特征提取与图像复合相结合, 无需 3 维人脸模型重建, 自动合成具有源图像主要五官特征的结果图像。系统允许用户调整权重, 在源五官与目标五官间进行插值, 控制结果图像与二者的相似程度。进一步地, 以鼻子附近的拉普拉斯算子为约束, 求解泊松方程以去除阴影影响。实验结果表明, 该算法效率高, 合成效果达到了实用的要求。

关键词 Poisson 方程 主动形状模型 五官替换 图像混合

中图分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2010)03-502-05

A Algorithm for Human Face Fusion Under Constraints

GAO Yan¹⁾, XU Jianzhong²⁾, WANG Changbo¹⁾, QIU Zhaowen³⁾, MA Lizhuang⁴⁾

¹⁾ (Institute of Software and Engineering, East China Normal University, Shanghai 200063)

²⁾ (Institute of Computer Science and Technology, Nanchang Army College, Nanchang 330000)

³⁾ (Institute of Computer Science and Technology, Northeast Forestry University, Harbin 150040)

⁴⁾ (Institute of Information and Computer Engineering, Shanghai Jiaotong University, Shanghai 221000)

Abstract This paper presents an image-based human face fusion algorithm, which can be used in medicinal beauty and digital entertainment areas. This paper combines facial features extraction with image fusion algorithm, so that we can automatically obtain synthesis results without using 3D facial models. Our system allows users to adjust the interpolation weights between source image and target image to control the similarity degree. Furthermore, we use Laplacian as constraints to reduce shadows near the nose. The experiment results show that our algorithm is efficient enough for practical demands.

Keywords Poisson equation, active shape model, facial fusion, image blending

0 引 言

图像视频编辑是近来计算机图形学领域的研究热点。为便于对特定对象进行编辑操作, 用户通常需要手工辅助完成图像分割。Perez 等人首先由用户交互指定 Trimap 区域, 并利用 Poisson 梯度域编

辑的方法, 自动将源图像无缝克隆到目标图像指定区域^[1]。Agawala 等人整合了图分割优化方法与梯度域混合方法, 用户只需在前景区和背景区绘制若干线条即可完成交互式的图像编辑^[2]。Jia 等人自动计算 Poisson 编辑方法的最优分割边界, 进一步减弱了用户的输入限制^[3]。以上方法都是针对图像中的通用物体进行编辑, 需要用户提供部分干预, 无

基金项目: 国家重点基础研究发展计划 (973) 项目 (2006CB303105); 国家高技术研究发展计划 (863) 项目 (2009AA01Z334); 国家自然科学基金项目 (60603076); 上海市科委自然科学基金项目 (07ZR14035); 上海市青年科技启明星计划项目 (08QA14025); 黑龙江省科技攻关重点项目 (GB06A107); 黑龙江省青年科学技术专项资金项目 (QC06C003)

收稿日期: 2008-10-20 改回日期: 2008-12-17

第一作者简介: 高 岩 (1973—), 男, 讲师。2006 年获上海交通大学计算机博士学位, 研究方向为计算机动画、真实感绘制、图像处理等。E-mail: yga@sei.ecnu.edu.cn

法全自动实现。

人脸是重要的信息交流通道,如何合成逼真的人脸图像是图形学领域的重要课题。虽然 3 维人脸能够结合姿态、光照、视点的变换,但其建模过程过于复杂。在很多场合下,输入只有单张人脸正面照片。尽管目前存在算法(如文献[4])可以从单张图像中得到 3 维人脸模型,但从本质上这一问题是欠约束问题,无法得到精确的深度信息。在医学美容或预测后代相貌的场合,需要利用特定的源五官(明星或本人)来替换目标五官图像(本人或明星)。如果采用通用的图像编辑替换算法,需要用户大量的手工交互;此外,分割算法难以对脸部区域产生理想的分割结果。另一方面,在这些场合,输入相片可以被严格限定为人脸的正面照,无需考虑姿态的不同,视点和光照条件差异也较小。这表明,无需对人脸图像进行 3 维重建,即可直接利用基于图像的方法得出问题的简化解。

提出了基于图像的人脸五官替换算法,可应用于医学美容以及数字娱乐领域。该算法以 MPEG-4 框架作为人脸模型标准,利用主动形状模型(ASM)算法^[5]自动提取脸部特征点。在此基础上,以邻近特征点作为约束边界,计算出需替换的五官区域;最后,对提取的特征点进行仿射对齐,利用特征保留的图像混合技术得出了具有不同五官混合权重的人脸图像。

1 系统框架

图 1 为五官替换系统的流程图。

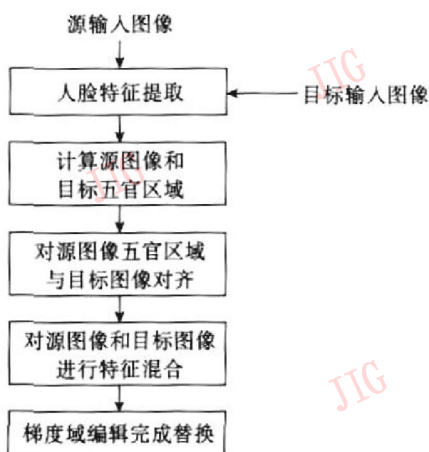


图 1 系统流程图

Fig. 1 System flow chart

文中的人脸正面照片来自中国科学院计算技术研究所的 PEAL 大规模人脸数据库^[6],部分小孩人脸照片下载自互联网。

2 特征提取

人脸面部特征定位一直是人脸识别领域中一个具有挑战性的课题。ASM 是常用的获得人脸形状的方法。MPEG-4 是 MPEG (moving picture experts group) 提出的新一代多媒体数据压缩标准,其中定义了人脸模型标准。该标准定义了人脸定义参数(FDP)和人脸动画参数(FAP)用来描述人脸的几何、纹理以及表情运动。由于该标准具有很强的兼容性,基于 MPEG-4 的人脸动画得到了广泛地应用。为了更加细致地描述五官区域,在前额和脸颊增加了一些特征点,修订出符合需求的 118 个特征点。这些新增点的坐标计算方法详细描述见文献[7]。图 2 中给出了本文提取的特征点结果。

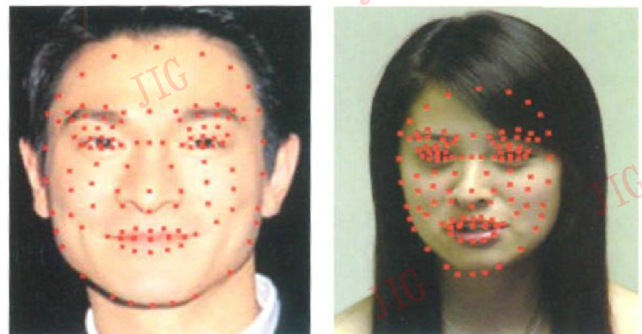


图 2 点为自动检测出的特征点

Fig. 2 The points are detected feature points

3 五官替换

在提取出输入图像的特征点后,可以根据规则自动计算出五官区域的位置,并采用梯度域编辑方法将源五官区域克隆到目标五官区域。对美容领域的应用,源图像为客户提供的明星照片,目标图像为客户自身;对后代相貌预测问题,源图像为客户自身,目标图像为从小孩人脸库中检索出来的脸形相似的小孩图像。在实现过程中,对算法做了部分简化,从男孩库中或女孩库中随机搜索一幅图像作为目标图像返回。

3.1 梯度域编辑

不同来源的图像具有不同的光照条件、纹理特

征,直接将源图像拷贝至目标图像将产生明显的接缝,采用低通滤波进行后处理将会导致图像的模糊化。由于人眼相对于灰度的变化比较不敏感, Perez 等人提出了梯度域编辑的思想^[1]——即拷贝图像的梯度信息,这将导致更加自然的结果。

该问题以目标图像的灰度作为边界约束,求解使梯度域保留的泊松方程。方程可由 Gauss-Seidel 迭代求解,具体过程见文献[1]。

3.2 计算五官区域

在决定人的面部相似程度方面,眼睛、鼻子以及脸轮廓起着关键性作用。脸的轮廓编辑需要准确提取脸的轮廓线,然后利用图像变形技术进行变形。由于存在头发、背景的影响(如图 2(b)),对脸轮廓作变形将带动背景区域的变形,这将导致难以预测的结果。出于简化的目的,本文只考虑眼睛和鼻子的替换(嘴的替换过程与眼睛类似,但未考虑嘴的张合影响)。

提取的脸部特征点虽然精度较高(具体结果见图 2),但由于人脸图像中姿态、背景、光照等信息存在着不确定性,任何特征点定位算法都无法达到像素级的精度。作为结果,在确定人脸五官区域时,要对相应特征点组成的封闭轮廓线进行放大,使整个眼睛或鼻子区域能够包括进来,不会由于拷贝不完整而导致不自然的结果。简单地将每个相关特征点向外缩放若干个像素阈值的方法难以保证通用性。为保证算法的健壮性,以周围的特征点来作为扩大轮廓线的依据。以眼部为例,其轮廓具体见图 3。

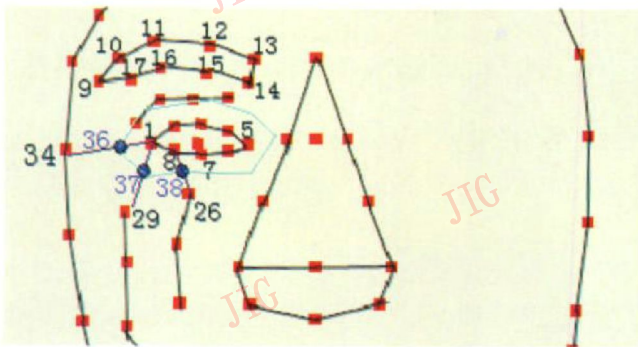


图 3 右眼区域示意图(图中编号仅为示意)

Fig. 3 The sketch map of right eye area

在图 3 中红色矩形点为检测出的脸部特征点。利用脸轮廓线、眉毛、鼻子、脸颊上的点作为约束,确定左眼拷贝区域。例如,在内眼点 1 与脸轮廓线点 34 连线上线性插值(系数 $1/3$)可得到点 36 在内眼点 1 与脸颊点 29 连线上线性插值(系数 $1/2$)可得

到点 37……连接这样的点,就得到最终右眼拷贝区域(青色线条所示)。图 4 给出了实际结果。与眼睛处理的思路类似,鼻子的拷贝区域是根据周围的眉毛、脸颊、眼睛上的对应点线性插值而来(由于只需保证能将鼻子区域整个包括即可,故对应点以及插值系数的选取没有严格的要求,对拷贝结果影响很小)。

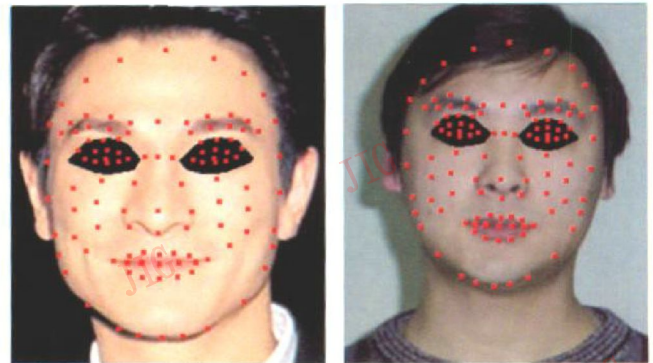


图 4 提取出的源图像和目标图像眼部区域

Fig. 4 The extracted eye areas of the source and target images

3.3 替换算法

现在需要将计算出的源图像的区域克隆到目标图像的对应区域。与通常的图像拷贝算法不同,目标区域与源区域大小不同,需要首先进行对齐操作,再进行拷贝。

3.3.1 区域对齐

在五官拷贝问题中,用户也许会提出比完全拷贝更进一步的要求:希望拥有这样的鼻子,他有 80% 像刘德华, 20% 像我。这就涉及到一个对图像作参数化混合的操作。眼睛区域具有眼睑、睫毛,由于特征点定位不可能准确到像素级,对它们进行混合很难产生自然的结果,本文只考虑对鼻子作参数化混合操作。

眼睛区域无需作图像混合,只需保证待拷贝的源图像区域能够覆盖目标区域即可。首先计算两瞳孔连线作为脸部的近似旋转偏移;接着计算源区域和目标区域的包围盒;对源区域作旋转和缩放变换,使其能够正确地覆盖目标区域。

对鼻子区域,需要对其中的特征点进行对齐处理。该过程包括两个步骤: 1) 粗配准,即求出全局平移、旋转和缩放变换,使得变换后对应顶点距离和最小。这一问题可归结为最小二乘问题,具有解析解。 2) 精配准,即求出源三角形与其对应的目标三

角形之间的仿射变换。假设有 6 个仿射变换参数, 3 个顶点共有 6 个方程, 可以解析求解出参数。

3.3.2 区域克隆

在将源五官区域与目标五官区域对齐后, 可以进行图像拷贝操作。图 5 给出了部分实验结果。



图 5 将刘德华的鼻子克隆到小孩脸上的结果

Fig. 5 The result of cloning Liu-DeHua's nose to the child

对鼻子的参数化混合算法进行测试, 图 6 给出了部分结果, 权重 ω 越大表明混合结果中源图像所占比例越大。

从图 6 可以看出, 在结果图像的鼻孔附近, 出现了明显的虚影。这是因为, 源图像和目标图像鼻子区域的特征点无法完全对齐, 当把源图像的鼻孔处像素 (一般近似为黑色) 与目标图像鼻子处的像素 (一般近似为肤色) 进行混合操作时, 将会产生虚影现象。



图 6 将刘德华的鼻子按照不同权重克隆到小孩脸上的结果

Fig. 6 The result of cloning Liu-DeHua's nose to the child according to different blending weights

3.3.3 约束情况下的替换

根据第 3.3.2 节对虚影成因的分析, 可以看出, 由于鼻孔附近的像素存在较大变化, 在做像素混合操作时, 特征点位置略有偏差即会导致不自然的结果。这种情况难以通过对齐操作来消除, 甚至在手工精心调节源图像和目标图像的特征点后, 虚影现象仍然可能发生。

以鼻孔附近的原始图像作为约束消除虚影现

象, 具体处理如下: 对鼻孔区域, 只采用来自源图像或目标图像的内容, 即尽可能不对鼻子下半部分作混合处理, 使其保持与初始图像相似。问题的关键在于如何找到鼻孔位置。由于在鼻孔附近, 像素存在跳变, 用图像处理中的边缘检测算法检测潜在的候选点。回顾在利用 Gauss-Seidel 迭代求解泊松方程的过程中, 需要计算源图像和目标图像的拉普拉斯算子, 并在迭代过程中尽量使它们一致。当拉普拉斯算子相差过大时, 难以保证收敛。这样, 当源图像或目标图像上某点的拉普拉斯算子与图像拉普拉斯场均值之差超过阈值时 (阈值选为该点周围 5×5 区间的方差的 k 倍, 这里 k 取 1.4), 即认为其位于鼻孔位置, 再根据混合权重大小分别选择对源图像或目标图像进行复制。具体算法如下:

1) 计算源图像和目标图像的拉普拉斯算子, 设其分别为 f_1 和 f_2 , 按照混合权重计算混合后的拉普拉斯算子 $f = \omega f_1 + (1 - \omega) f_2$, 将其作为指导 Gauss-Seidel 迭代中的源拉普拉斯场;

2) 计算源图像和目标图像中人脸特征点所包围区域的边界矩形, 取矩形的下半部分作为待检测区域 (根据先验知识, 鼻孔区域位于人脸中间偏下部位); 分别计算源图像和目标图像的拉普拉斯算子的平均值以及方差, 若任一图像中某像素方差值大于阈值, 则标记为潜在鼻孔区;

3) 对标记为潜在鼻孔区的像素, 根据混合权重选择源图像或目标图像。若 $\omega \geq 0.5$ 则取源图像的拉普拉斯场替代第一步计算所得的拉普拉斯场, 替代范围为当前像素周围的 $n \times n$ 区域 (本文取 $n = 7$); 反之, 若 $\omega < 0.5$ 则取目标图像的拉普拉斯场替代第 1) 步计算所得的拉普拉斯场;

经过上述步骤处理, 计算出的拉普拉斯场在非鼻孔位置处取源图像和目标图像拉普拉斯算子的加权混合, 在鼻孔位置则直接取源图像或目标图像的拉普拉斯算子。这一性质满足了保持鼻孔外形不变的约束条件。

然而, 这样计算出的拉普拉斯场仍然存在一些小的缺陷。由于在非鼻孔位置和鼻孔位置采用了两个不同来源的数据, 接触边界位置会产生抖动现象, 需要对这一现象作平滑处理。

上述平滑问题可以看作从数据源 2 到数据源 1 的过渡。使用 Gao 等人^[8] 的混合函数 $\alpha(t) = 2t^3 - 3t^2 + 1$ 来消除过渡间的不连续。该函数具有如下性质:

$$\alpha(0) = 1, \alpha(1) = 0, \frac{d\alpha}{dt}(0) = \frac{d\alpha}{dt}(1) = 0$$

这些性质保证了图像区域过渡的平滑性。

4 实验结果

系统环境是 CPU Pentium 4 1.8 GHz 内存 512 M, 开发环境是 VC++ 6.0。对算法的性能做了测试, 测试样例为 43 幅图像, 图像平均大小为 284×410 平均特征定位时间为 5.3 s, 五官替换算法时间为 0.35 s。进一步测试了约束情况下的五官替换算法, 实验结果见图 7。

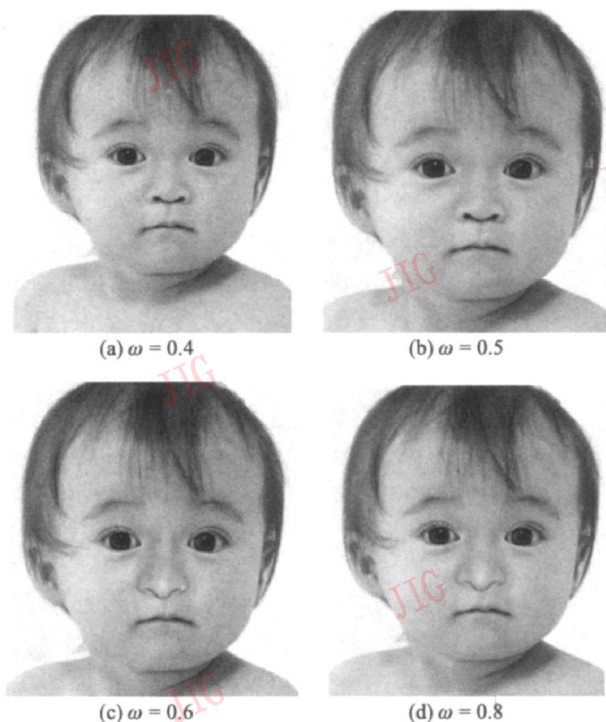


图 7 在约束条件拉普拉斯域编辑下, 将刘德华的鼻子按照不同权重克隆到小孩脸上的结果

Fig. 7 Cloning Liu Dehua's nose to the child with different blending weights under constrained Laplacian editing

从图 7 可以看出, 约束情况下的五官替换算法消除了鼻孔附近的虚影现象。同时, 鼻子上部仍然是由源图像和目标图像按权重混合而成, 在边界处没有明显痕迹, 表明过渡算法是可行的。

然而, 本文算法仍然有些缺陷。为去除鼻孔附近的阴影, 由权重来决定鼻孔附近的拉普拉斯场, 这导致了混合空间的不均匀, 即在 $\omega < 0.5$ 处, 鼻孔附近采用的仍是源图像的梯度场, 而在 $\omega > 0.5$ 处, 鼻孔附近采用的仍是目标图像的梯度场。这样, 在 $\omega = 0.5$ 附近, 结果图像的鼻孔部位将分别偏重于源图像和目标

图像, 导致了图 7(b) 和图 7(c) 的不连续。虽然在一幅图像内部, 用过渡函数消除了不平滑, 但对整个混合空间, 仍然存在着跳变。但本文针对的是娱乐领域, 并非尝试构造严格数学意义上的混合空间。

5 结论

提出了约束条件下的人脸五官替换算法。该算法针对人脸图像, 自动提取脸部特征点并计算五官区域, 最后利用 Poisson 方程将源五官克隆到目标人脸。为取得更好的结果, 进一步提出权重克隆的思想, 用户交互调节权重, 混合结果同时具有源五官和目标五官的特征; 为消除虚影现象, 对混合过程施加了约束。

对提出算法进行了大量实验, 结果表明, 本文算法效率较高, 合成结果自然, 具有一定的实用价值。尽管如此, 本文的技术一旦被不法分子利用, 可能会应用于不同身份人脸照片的篡改, 造成不良的社会影响, 这一点必须引起人们充分的注意。

参考文献 (References)

- [1] Perez P, Gangnet M, Blake A. Poisson image editing [J]. ACM Transactions on Graphics, 2003, 22(3): 313-318.
- [2] Agarwala A, Dontcheva M, Agarwala M, et al. Interactive digital photomontage [C] // Proceedings of SIGGRAPH 2004. Los Angeles, CA, USA: ACM Press, 2004: 294-302.
- [3] Jia J, Sun J, Tang C K, et al. Drag-and-drop pasting [J]. ACM Transactions on Graphics, 2006, 25(3): 631-637.
- [4] Jiang Dalong. Research on Realistic 3D Human Facial Animation [D]. Beijing: Institute of Computing Technology of the Chinese Academy of Sciences, 2005. [姜大龙. 真实感 3 维人脸合成方法研究 [D]. 北京: 中国科学院计算技术研究所, 2005.]
- [5] Cootes T F, Taylor C J, Cooper D. Active shape models—their training and application [J]. Computer Vision and Image Understanding, 1995, 61(1): 38-59.
- [6] Gao W, Cao B, Shan S. The CAS-PEAL Large Scale Chinese Face Database and Evaluation Protocols. JDL TR 04 FR 001 [R]. Beijing: Institute of Computing Technology, Chinese Academy of Sciences, 2004.
- [7] Zhou Renqin, Zhou Jingye, Chen Yiqiang, et al. Caricature generation based on facial feature analysis [J]. Journal of Computer-Aided Design & Computer Graphics, 2006, 18(9): 1362-1366. [周仁琴, 周经野, 陈益强, 等. 基于特征发现的卡通人脸肖像生成 [J]. 计算机辅助设计与图形学学报, 2006, 18(9): 1362-1366.]
- [8] Gao Y, Ma L, Chen Z, et al. Motion normalization: The preprocess of motion data [C] // Proceedings of the ACM Symposium on Virtual Reality Software and Technology. New York, USA: ACM Press, 2005: 253-256.