

模型基编码中脸部模型的自动调整

李梦东, 阮秋琦

(北方交通大学信息科学研究所, 北京 100044)

摘要: 通用模型的调整是模型基人脸图像编码的重要步骤, 本文提出了一种利用改进的变形模板提取脸部完整特征、由径向基函数内插调整模型非特征点的正面脸部模型自动调整算法。首先用矩形模板匹配确定各特征区域, 模板尺寸由先验知识设计; 在各特征区域内确定变形模板的初始位置 and 变化范围, 采用遗传算法等匹配方式, 获取变形模板的最优参数值; 最后估计输入人脸姿态, 进行模型的全局变换和用径向基函数内插调整模型非特征点, 得到输入人脸的特定模型。实验结果表明, 对于输入为简单背景的肩头像, 该算法简便快速, 可获得较好的调整效果。

关键词: 模型基编码; 模型调整; 变形模板; 反向距离内插

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 0372-2112 (2002) 10-1562-04

An Algorithm of Face Model Adaptation for Model Based Image Coding

LI Meng dong, RUAN Qiu qi

(Institute of Information Science, Northern Jiaotong University, Beijing 100044, China)

Abstract: An algorithm of automatically adapting facial wire frame model is proposed in this paper. The algorithm at first locates the positions of face and feature regions in the image with region growing and template matching, then uses deformable templates to extract the exact positions of features. Finally the algorithm does the model scaling and rotating adaptation, and then local adaptation with radial basis function interpolations. The test result is satisfactory.

Key words: face model adaptation; feature extracting; deformable template; inverse distance interpolation

1 引言

在模型基人脸图像编码中, 一项重要任务是把一个先验的人脸网格通用模型, 经过特征提取及模型调整过程, 变为输入人脸的特定模型。模型调整过程一般多为三维人脸模型的正面调整^[1~3], 这对于基本是正面图像的肩头像序列来说是适合的。由于自动性和实时性的特点, 编码系统对特征提取和模型调整方法有较高的要求。

在模型调整过程中, 多数的特征提取方法是利用变形模板匹配^[1~3], 但由于变形模板是一个多参数的优化过程, 加之光照等成像条件的影响, 使变形模板准确提取人脸特征十分困难, 算法可靠性及速度方面都有待提高。提取特征后下一步骤是利用这些特征点调整模型, 方法有弹性匹配^[4]、自由变形^[4]和散列数据内插^[5]等, 但这些方法或者适用于交互场合, 或者没能与特征提取结合为一个完整系统。因此如何自动、快速实现模型调整仍是一个需要深入研究的课题。最近 Vetter^[6]等用统计的方法由单一正面图像实现模型参数自动调整, Coote^[7]等采用主动形状模型和主动表象模型的方法实现特定人脸的表示, 但这些方法都需要较大的样本库, 而且纹理还缺少变化。

本文提出了一种用单幅人脸图像进行模型调整的算法, 该算法在变形模板提取特征的可靠性、速度等方面做了改进,

并结合径向基函数内插形成了一个完整的自动模型调整系统。首先用矩形模板匹配确定各特征匹配, 作为变形模板的初始位置。之后设计了完整特征的变形模板, 采用遗传算法等匹配过程得到特征点位置。最后估计输入人脸姿态, 利用以逆多二次式为基函数的径向基函数内插调整三维模型非特征点(深度近似)。实现结果表明, 对于简单背景的人头像图像, 该算法简便迅速, 结果准确, 具有一定的实用性。

2 脸部和特征区域定位

人脸定位有多种方法, 我们采用区域增长法定位人脸并要求输入为简单背景的灰度图像。一般人脸图像中脸部是居中的, 因此可以将图像中心作为导航点^[8], 如果碰巧遇到鼻孔等灰度值较暗的位置, 则根据增长区域的大小加以判断, 如区域过小则另选起始位置, 直至区域大小较为合理为止。由于光照、头发等因素的影响, 区域增长结果可能会出现欠分, 即少某些部分, 如额头、脸部轮廓; 还可能会过分, 即包含脖子、耳朵。对于欠分, 我们利用水平积分投影进行判断, 并按照人脸的一般比例关系, 如额头和下巴的高度近似相等加以修复。对于过分, 先用垂直积分投影去掉耳朵部分, 再用人脸的平均长宽比例, 限制脖子区域。

为了提取人脸特征, 我们首先用方框框住特征区域, 这样

可确保变形模板有一个较好的初始位置. 为了提高计算速度, 本文采用根据人脸特征区域一般比例关系设计的各特征矩形窗口, 在边缘图像和灰度图上作模板匹配. 边缘图像较清晰地反映了完整的特征区域. 嘴部为了和鼻子分开, 可加入灰度信息作为辅助, 眼眉也可利用灰度信息. 窗口模板代价函数为: 窗口内均匀分布、间隔为 n 像素的水平 and 垂直线上边缘值和灰度值的加权和^[8]:

$$S = \sum_{i,j} a_{i,j} (E_{i,j} + \alpha I_{i,j})$$

式中 $E_{i,j}$ 、 $I_{i,j}$ 分别表示像素的边缘强度和灰度值, α 可取值 1、0, $a_{i,j}$ 为负指数衰减的加权系数, 这是为了使区域中央衰减, 周围衰减大, 目的是突出中心, 有利于区分独立的特征区域与边界区域. 取样间隔 n 的大小影响搜索速度, 它与模板大小成比例, 一般取 $1/3$.

矩形模板匹配过程是在各搜索范围内进行全局搜索. 搜索过程是先确定左右眼的特征区域, 再向上确定眼眉, 向下确定嘴和鼻子, 其中右眼在定位的脸部区域右上方 $1/4$ 区域内, 左眼在左上方 $1/4$ 区域内, 嘴在下 $1/2$ 区域内. 另外匹配过程加入了一些人脸器官相互位置的约束条件, 如左右的对称性、眼眉不可能距眼睛过远以及两眼的高度差也不可能过大等, 这可有效克服头发等造成的阴影影响.

3 脸部特征的自动提取

3.1 眼睛和嘴部特征提取

(1) 眼睛变形模板和能量函数

眼睛的变形模板由两条抛物线和一个圆组成, 具体表达式见文献^[2]. 不同之处是能量函数的设计, 我们定义的眼睛变形模板能量函数为式(1)(2), 其中能量仅用灰度和边缘能量, 边缘由灰度形态学算子提取, 可避免二值化阈值的设定. 其中式(1)为眼球的能量, 式(2)中 $j = 1, 2$ 分别对应上眼帘和下眼帘的情况, $length$ 为边缘长度, $area$ 是所围的面积, b 表示边缘强度, g 表示灰度值. a_1, a_2 是加权系数, 实验中 a_1, a_2 取 1.

$$E = E_{edge} + E_{intensity} = - (a_1 \frac{\sum b_i}{length} + a_2 \frac{255 - \sum g_i}{area}) \quad (1)$$

$$E = E_{edge} + E_{intensity} = - \sum_j a_1 \frac{\sum b_i}{length^j} + a_2 \frac{255 - \sum g_i}{length^j} \quad (2)$$

(2) 嘴部变形模板和能量函数

嘴部闭合的变形模板我们采用简化的形式, 由三条抛物线组成, 分别对应上嘴唇边缘、中缝和下嘴唇边缘, 具体图示和表达式见文献^[2]. 对应于嘴部张开与闭合两种情况, 应该分别有两种形式的变形模板, 但一般嘴部是微张的, 所以目前我们都用闭合的模板表示, 要求嘴部张开距离小于下唇厚度. 嘴部变形模板的能量函数同式(2), 此时 $j = 1, 2, 3$ 分别对应上嘴唇边缘、中缝和下嘴唇边缘的能量.

(3) 变形模板的匹配过程

我们根据特征区域的定位, 确定模板参数的初始值和变换范围, 为以下的搜索过程提供搜索空间. 对于变形模板多参数的优化问题, 传统的逐个确定变形模板参数的做法, 先确定

的参数值不一定是最佳的, 故整个过程易于陷入局部最小, 为此我们采用遗传算法^[8]得到模板参数的最优解. 由于搜索范围小, 故速度较快. 对于眼睛的变形模板匹配, 为了提高速度, 首先确定眼仁的位置和大小, 之后利用遗传算法确定眼脸的参数, 包括中心位置和上下抛物线的高度、长度.

3.2 确定脸部轮廓、鼻和眼眉特征

(1) 脸部轮廓提取

M. Kampmann^[3]用四段抛物线模板, 表示下巴和脸颊形状, 这时用两个端点表示抛物线. 我们也采用四段抛物线提取脸部轮廓(下巴和脸颊), 如图 1 所示, 其中 A, B, C, D 和 E 是抛物线的端点, 模板参数即为这五点的坐标. 我们设计了新的抛物线表达式和初始位置, 下巴和脸颊四段抛物线的表达式为式(3)~(6). 其中坐标为沿两眼连线偏转后的坐标.

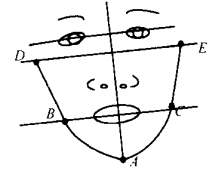


图 1 脸部完整特征模板示意图

$(x_A, y_A), (x_B, y_B), (x_C, y_C), (x_D, y_D), (x_E, y_E)$ 分别为 A, B, C, D 和 E 的坐标.

$$Y_1 = - (y_B - y_A) \left(1 - \left(\frac{X_1}{x_A - x_B}\right)^2\right) \quad (3)$$

$$Y_2 = - (y_C - y_A) \left(1 - \left(\frac{X_2}{x_C - x_A}\right)^2\right) \quad (4)$$

$$Y_3 = - \frac{y_B}{\left(1 - \frac{x_B^2}{(x_A - x_D)^2}\right)} \left(1 - \left(\frac{X_3}{x_A - x_D}\right)^2\right) \quad (5)$$

$$Y_4 = - \frac{y_C}{\left(1 - \frac{x_C^2}{(x_E - x_A)^2}\right)} \left(1 - \left(\frac{X_4}{x_E - x_A}\right)^2\right) \quad (6)$$

匹配时首先确定各端点的初始位置, A 为两眼中心和嘴部中心连线(称为中心线, 其中眼中心为眼仁中心), 与脸部区域边框的交点; B, C 在与两眼连线平行、过嘴部中心的直线上, 与中心线的距离为外眼角至中心的距离, 然后向外推测, 这样可避免背景的干扰; D, E 在与两眼连线平行, 在其下距离 g 的直线上, 与中心线的距离是外眼角至中心线的距离, 匹配时向外推测. g 可取眼睛区域到鼻子区域距离的 $1/3$. 本文中能量函数利用边缘、灰度值及线上的灰度一致性, 并利用人脸的结构特点限定特征点的变化范围.

(2) 鼻子特征提取

对于鼻部的特征, 我们采用两抛物线作为左右两个鼻翅的模板, 用两个小圆表示两鼻孔形状. 其表达式为式(7)(8). 其中 $h = H/3, w = 2h/3, r = H/6, H$ 为鼻子区域模板的高度. $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$ 分别为左鼻翅、右鼻翅、左鼻孔和右鼻孔模板的中心坐标, 它们为可变参数.

$$X_{1-} - x_1 = -w \left(1 - \left(\frac{Y_1 - y_1}{h}\right)^2\right), X_{2-} - x_2 = w \left(1 + \left(\frac{Y_2 - y_2}{h}\right)^2\right) \quad (7)$$

$$(X_3 - x_3)^2 + (Y_3 - y_3)^2 = r^2, (X_4 - x_4)^2 + (Y_4 - y_4)^2 = r^2 \quad (8)$$

匹配时由边缘和灰度能量确定其位置. 实际中设定适当的初始位置和位置约束, 如鼻梁必须在鼻孔外边、左右鼻梁和鼻孔在中心线的两侧、鼻梁长度是鼻孔直径的两倍等, 避免鼻梁和鼻孔模板相交. 有些人脸显示不出鼻孔, 但一定存在较深的灰度, 由以上过程也可近似确定鼻孔位置. 最后将两鼻梁之间中心位置作为鼻尖的位置.

(3) 眼眉特征提取

眼眉用眼眉区域边框的三点 $(x_1, (y_1 + y_2)/2)$, $((x_1 + x_2), y_1)$, $(x_2, (y_1 + y_2)/2)$ 连线表示. (x_1, y_1) , (x_2, y_2) 分别为眼眉边框的左上角和右下角坐标 (沿两眼连线偏转后的坐标). 以边框为中心, 作上下移动, 根据线上的灰度能量确定最佳位置.

4 模型的自动调整

为了使模型能够准确表达输入的人脸, 需要估计输入人脸姿态的小角度偏差, 再对模型进行旋转、平移和尺度变换, 之后再调整模型的非特征点位置. 整个模型调整过程, 可表示为下式

$$M' = S(F, \alpha \cdot R \cdot (M + T))$$

其中 M , M' 分别表示通用模型和调整后的模型. $S(f, m)$ 为模型非特征点调整变换, F 为特征点向量, R 为旋转矩阵, T 平移矩阵, α 为尺度因子. 我们规定三维模型坐标系原点为鼻尖沿 z 轴与后脑的交点, x 轴为水平方向向右为正, y 轴为垂直方向向上为正, z 轴为垂直纸面向外为正.

4.1 头部姿态的估计

变形模板本身带有偏转角, 但只是变形模板各自沿 z 轴的转动, 不能代表脸部整个的转动. 为确定输入人脸的姿态, 我们由两眼中心连线与水平线的角度判断沿 z 轴的偏转; 另外, 还可确定沿 y 轴的偏转, 可由两眼角与轮廓边 (发际) 距离的不对称来判断, 如有头发遮挡时, 用嘴角距脸廓的不对称来估计. 由于深度未知, 可近似把脸部视为圆柱体, 模型原点在脑后位置, 故深度约为脸宽; 对于沿 x 轴的转动, 由于没有可靠的依据而不易估计.

沿 z 轴的转动角为 (顺时针为正): $\theta = \arctan\left(\frac{y_1 - y_2}{x_1 - x_2}\right)$,

(x_1, y_1) , (x_2, y_2) 分别为左右眼睛中心的坐标.

沿 y 轴的转动角为 (左转为正): $\theta = (180/\pi) \cdot (\Delta d/2H)$,

Δd 为眼角到脸轮廓的不对称差, H 为脸宽.

4.2 模型的调整

特征点由上面得到的特征位置中选出, 与 MPEG-4 中标准的模型上顶点位置一致, 具体特征点位置见文献[5]. 脸部轮廓的上部由于表示头发形状困难, 目前根据特征区域位置和提取的脸颊位置, 用半圆形近似表示额头的形状. 尺度调整即为确定尺度变换系数, 可选取眼睛中心坐标与模型上对应点坐标之比作为尺度变换系数. 旋转变换是由估计的偏转角度构成旋转矩阵, 乘以特征点矢量. 非特征点调整我们采用以逆多二次式为基函数的径向基函数内插算法, 具体算法见文献[9], 其中模型顶点深度为经过尺度变换的原模型特征点深度.

5 实验结果

采用以上算法, 我们在 PC PIII-450 机上进行了实验, 其中模型的绘制采用 OpenGL 函数实现. 图 2 为一个实验实例. 图 2(a) 为 Miss American CIF 格式第 30 帧, (b) 为特征区域定位结果, (c) 为变形模板提取结果, (d) 为采用的 Waters 通用模型, (e) 为模型调整后的特定模型, (f)、(g) 为纹理映射后的模型, 纹理为包围特征点的最小边框图像区域, 且为正面纹理映射. 算法执行时间小于 2 秒. 从图 2(g) 可看出纹理与模型对得比较准, 说明模型调整结果较准确. 为进一步检验模型调整的结果, 我们用相关函数来考察重建图像的准确性. 做法是将贴上纹理的、变换为图像脸部大小的模型, 映射到图像的对对应位置求合成的图像与原图像的相关函数. 相关函数定义如下式, 计算结果 $r = 0.93$. 图 3 为另两个试验结果, 每幅中左边为原图, 右边为模型调整后的合成图.

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}} ,$$

$$\bar{A} = \text{mean}(A), \bar{B} = \text{mean}(B).$$

模型调整的准确性主要取决于特征提取的准确性. 在所有特征中眼睛是很关键的, 又是较难提取的, 因为眼睛构造复杂、受干扰大. 例如 Miss American 图像右眼部有头发的遮挡, 造成眼睛边缘模糊, 提取结果存在稍大的误差. 我们加入了左右眼对称性约束, 保证右眼提取结果的准确性. 另外, 特征点的多少也是影响调整准确性的因素, 这应该根据应用场合对准确性的要求而定. 为提高该算法的适应性, 今后还需要研究较复杂背景和成像条件下, 确保准确地定位人脸的算法, 以及边沿提取算法. 系统还应该加入错误验证的功能, 保证系统的正常运行. 为获得更为真实的重建效果, 应采用加入头发和耳朵等的完整头部模型.

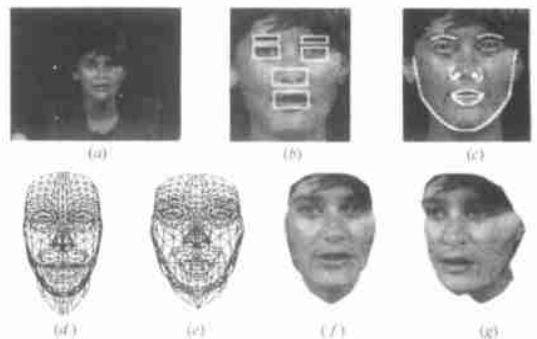


图 2 实验结果 1



图 3 实验结果 2

6 结束语

模型基人脸图像编码系统, 在分析阶段提取输入人脸的形状、运动和纹理参数, 在合成阶段利用模型和各参数组重建出原来的图像. 其中形状参数是为了获取输入人脸的特定模型, 这一任务一般是由特征提取和模型调整等步骤完成的. 我们提出了一种改进的利用变形模板提取完整特征、利用径向基函数内插调整非特征点的模型自动调整算法. 实验表明这种算法简便快速, 结果较准确, 为模型基编码中运动跟踪、表情合成等任务打下了基础. 将来的工作包括进一步提高特征提取的准确性; 利用合成图像进行分析; 由序列图像调整模型的深度.

参考文献:

- [1] M T Reinders. Model adaptation for image coding [D]. Netherland: University of Delft, 1995.
- [2] L Zhang. Automatic adaptation of a face model using action units for semantic coding of videophone sequences [J]. IEEE Trans. CSVT, 1998, 8(6): 781-795.

- [3] M Kampmann, R Farhoud. Precise Face Model Adaptation for Semantic Coding of Videophone Sequences [A]. Picture Coding Symposium (PCS' 97) [C]. Berlin: PCS, 1997.
- [4] W S Lee, N M Thalmann. Fast head modeling for animation [J]. Journal of Image and Vision Computing, 2000, 18(4): 355-364.
- [5] Fabio Lavagetto, Roberto Pockaj. The facial animation engine: toward a high level interface for the design of MPEG-4 Compliant animated faces [J]. IEEE Trans. CSVT, 1999, (2): 277-289.
- [6] T Vetter. Synthesis of novel views from a single face image [J]. International Journal of Computer Vision, 1998, 28(2): 103-116.
- [7] T F Coates. Statistical models of appearance for computer vision [R]. Tech. Report, Division of Imaging science and Biomedical Engineering, University of Manchester, 2000.
- [8] ChurHung Lin et al. Automatic facial extraction by genetic algorithms [J]. IEEE Trans. On Image Processing, 1999, 8(6): 834-845.
- [9] Li Mengdong, Ruan Qinqi. Facial wire frame model adaptation using RBF interpolation [A]. WCC2000 [C]. Beijing: WCC, 2000. 1014-1017.

作者简介:

李梦东 男, 1964 年生于山东利津, 北方交通大学信息科学研究所博士生, 研究方向为图像处理, 计算机图形学等.

软科学研究方法教程

软科学是现代自然科学、技术科学、社会科学、思维科学相互交叉、逐步发展而形成的一组具有高度综合性的新兴学科. 它以现代自然科学和社会科学的理论和方法, 特别是软学科和软技术的理论和方法为依据. 以政治、经济、科技、社会、管理、环境等广泛而复杂的问题及其内在联系(包括人、事、物系统)为研究对象, 并对其现实和发展的各个层次、各个方面进行系统、综合、全方位的分析、研究、预测和评价, 从而提出方案和战略, 制定政策和策略, 编订计划和规划, 优化组织和管理, 为决策科学化和管理现代化服务. 为了推广普及软科学研究方法, 我们率先给博士、硕士研究生班开设了“软科学研究方法”的必修课, 在总结讲授内容的基础上, 在前辈的鼓励下, 同行的支持下, 参考国内外最新参考文献集成为本教材. 本教程理论与应用并重, 学术性与实用性兼容, 定性分析与定量分析相结合, 常规方法与现代方法相结合, 专题研究与综合研究相结合, 内容丰富, 符合国情, 读者面广. 本教材可供科技工作者和软科学研究者以及大专院校管理专业等有关专业教师、研究生、本科生阅读参考. 也可作为各级党政领导、管理干部和技术干部的培训教材.

《软科学研究方法教程》由华中科技大学同济医学院医药信息系金新政主编, 湖北科技出版社 2002 年 8 月出版.

信息管理概论

全书共分 10 章, 在论述信息的基本概念和相关概念的基础上, 从源、流、用等方面全面系统地阐述了现代信息理论、知识、技术和工作方法, 详细介绍了信息源、信息传递、信息流、信息组织、信息用户及信息需求、信息服务、信息系统、信息技术及信息事业管理等知识. 本书不仅可以作为信息管理及其相关专业的教材, 同时也可供有关专业人员参考.

《信息管理概论》由华中科技大学同济医学院医药信息系金新政、陈氢主编, 华中科技大学出版社 2002 年 8 月出版.