

北京航空航天大学研究生课程考核记录

2018—2019 学年 第一学期

学号 SY1806214

姓名 陈鸿超

成绩

课程名称：《虚拟现实技术及应用》

论文题目：

基于特征点加权的三维形变模型优化

任课教师评语：

任课教师签字：

考核日期： 年 月 日

基于特征点加权的三维形变模型优化

陈鸿超

(北京航空航天大学计算机学院,北京 100191)

摘要: 三维形变模型(3DMM)是一个用于从单一 2D 图片还原 3D 面部模型的非常有用的工具,但是传统的拟合方法无法建立精确的 3D 模型。本文介绍了一种基于特征点加权的优化策略,该方法考虑到了面部不同区域的差异性,在拟合过程中根据误差程度调整不同特征点的参数修正,降低了重建误差,提高了模型的精确性。

关键词: 三维形变模型; 特征点加权; 虚拟现实

3DMM optimization based on landmark weighting

(School of computer science and technology, BeiHang University, Beijing 100191, China)

Abstract: 3D Morphable Model (3DMM) is a powerful tool for reconstructing the 3D face from a single 2D face image, but the traditional fitting methods fail to reconstruct an accurate 3D model. In this paper, we propose an optimization based on landmark weighting, this method takes into account the differences in different parts of the face, and adjusts the parameter correction of different feature points according to the degree of error during the fitting process, significantly reduces the reconstruction error and improves the authenticity of the 3D model expression.

Key words: 3DMM; landmark weighting; virtual reality

1 引言

使用 2D 图像重建 3D 人脸模型是计算机视觉和图形学的基本问题,在人脸识别、动画、游戏、虚拟现实等领域都有广泛的应用。但是,由于相机投影时的信息丢失问题,导致模型重建工作非常具有挑战性。

在面部模型重建方面,一个非常典型的解决方法就是使用三维形变模型(3DMM)[1]。该模型会先借助统计方法,对数据集进行分析,构建一个低维参数表示的平均面部模型。该模型拥有一组固定的特征点,每个特征点都使用形状参数和纹理参数进行表示。在平均面部模型中,这些参数的取值是根据数据集中所有样例计算得到,并且由一组系数进行控制。因此,通过修改这组系数,就能得到任意的面部模型。利用该平均面部形变模型,再结合输入的具体面部属性,我们就可以得到一个基本相似的 3D 面部模型。最后将 3D 模型转换为 2D 表示,与输入的具体图片进行匹配计算,根据匹配误差不断修改模型系数。这样,经过若干次迭代优化之后我们就可以获得一个比较精确的 3D 面部模型。

不过,在传统的 3DMM 模型中,对系数进行迭代优化时,所有的特征点都使用相同的误差梯度进行调整,忽略了特征点之间的差异性,这就导致传统的方法无法构建比较精确的 3D 面部模型。因此,本文基于 3DMM 模型,介绍了一种基于特征点加权的优化策略,针对面部不同区域的特征点,为其设置不同的误差权重。在迭代优化的过程中,3D 重建模型与 2D 输入图片中对应的特征点之间误差较大时,给其设置较大的误差权重,使其系数在更新时变化幅度更大,反之亦然。通过对特征点的加权处理,能够更有效的降低优化过程中重建模型的与真实图片之间的误差,建立更精确的模型。

2 相关工作

2.1 三维形变模型

人脸具有相似的全局特征,比如主要器官的位置,包括眼睛、鼻子、嘴巴等。研究证明,使用有限的参数就可以表示出人脸的基本特征。因此,借助统计方法就能够从数据集中获取面部低维模型中各参数的有效范围,缩小了参数空间的搜索范围。

3DMM 模型就是基于上述的原理设计的,首先,该模型会根据数据集收集到的人脸数据构建一个低维可变的的面部模型,也就是一个面部空间,该空间的基底是对数据库进行统计分析得到。对于任意一个具体的人

脸，只要求得各基底的相关系数，就可以在这个面部空间中构建出相应的 3D 人脸模型。同时，根据数据集也可以确定各系数之间的概率分布，防止出现奇怪的面部模型。

因此，使用 3DMM 模型进行人脸模型重建的主要步骤如下：一是通过 3D 人脸数据集构建面部空间基底，并确定各系数之间的变化关系。二就是将具体的 2D 人脸图片映射到 3D 人脸模型上，建立 2D 图片与 3D 模型之间的计算方程，求得各基底对应的系数。三是将 3D 模型还原为 2D 表示，通过与输入图片的 2D 数据进行误差计算，不断的迭代修正 3D 模型的系数。

2.2 特征点加权优化

在使用 3DMM 模型进行面部模型重建时，需要将 3D 模型转换为 2D 表示，然后不断迭代计算 3D 模型的 2D 表示与输入的 2D 图片数据中特征点之间的误差，然后更新 3D 模型中相应的系数。但由于人脸是一个具有形状和表面纹理的 3D 物体，使得这种对应关系是一种非常复杂的非线性关系，并不是所有的特征点都要采用相同的修正方式。因此，大多数传统的拟合方法都无法建立非常精确的模型。

为了解决这一问题，本文介绍了一种特征点加权的优化策略。对于不同面部组件对应的特征点，我们可以对其施加不同的权重。在计算过程中，对于误差较大的特征点，其对应的权重也较大；误差较小的特征点，其对应的权重也相应较小。这样在根据损失函数进行迭代优化的过程中，就可以依据权重的大小调整对每个特征点系数的修正程度，从而降低了模型重建的误差，提高三维模型表达的真实性。

3 传统的 3DMM 模型

3.1 3D面部空间的基底构造

对于每个 3D 面部模型，假设有 n 个面部特征点，那么就可以用两个特殊向量来表示这个 3D 面部模型。一个是形状向量 $S = ((x_1, y_1, z_1), \dots, (x_n, y_n, z_n))$ ，一个是纹理向量 $T = ((r_1, g_1, b_1), \dots, (r_n, g_n, b_n))$ 。在形状向量 S 中，每个特征点都使用 3 个值 (x, y, z) 来表示它在面部空间中的坐标。在纹理向量 T 中，每个特征点都使用 3 个颜色值 (r, g, b) 来表示其纹理信息。这样一来，所有的面部模型就都可以使用一组向量 $\{S, T\}$ 进行表示。

使用上述方法，我们就能够从面部模型数据集中获取到 m 组不同面部模型的向量表示 $\{(S_1, T_1), \dots, (S_m, T_m)\}$ 。这些向量组构成的集合 $\{(S_1, \dots, S_m), (T_1, \dots, T_m)\}$ 就是 3D 面部空间的基底。同时我们也可以计算得到一个平均面部模型的向量表示 $\{\bar{S}, \bar{T}\}$ ：

$$\bar{S} = \frac{1}{n} \sum_{i=1}^m S_i, \quad \bar{T} = \frac{1}{n} \sum_{i=1}^m T_i$$

这时，在 3D 面部空间中，任意一个面部模型都可以表示为：

$$S_{model} = \bar{S} + \sum_{i=1}^m \alpha_i (S_i - \bar{S}), \quad T_{model} = \bar{T} + \sum_{i=1}^m \beta_i (T_i - \bar{T}), \quad \sum_{i=1}^{m-1} \alpha_i = \sum_{i=1}^{m-1} \beta_i = 1$$

这里的系数 $\{\alpha_i, \beta_i\}$ 就可以用来调控 3D 面部模型中各部分的具体形状，从而实现构建任意面部模型的目的。同时，使用这种方式，面部特征中的各个部分其实也都只对应着部分系数，使用原模型划分出的独立子集就可以进行表示和调整。比如眼睛、鼻子、耳朵这些区域，就可以通过只改变 $\{S_{model}, T_{model}\}$ 的部分维度系数来进行调整，而其他区域不受到影响，如图 1 所示。

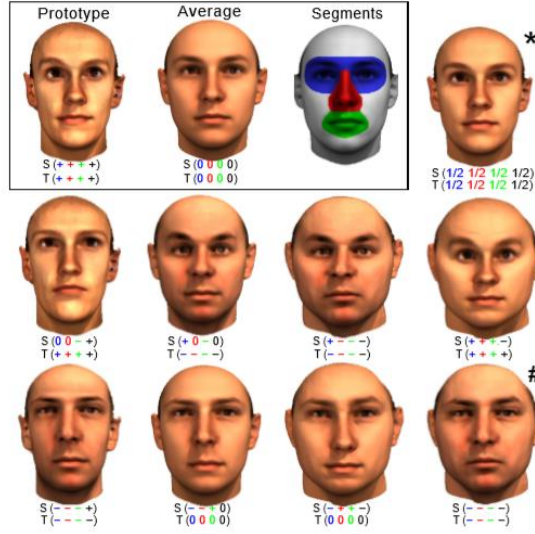


图 1 只改变原模型的部分维度即可生成若干不同的面部模型

这里我们只改变了 $\{S_{model}, T_{model}\}$ 中与相应面部属性相关的部分维度系数，而其他系数维持不变，就能够在基本模型的基础上构建出具有不同面部特征的新模型。

3.2 面部属性

通过构建 3D 面部空间的基底，我们就可以使用一组系数 $\{\alpha_i, \beta_i\}$ 来表示任意人脸模型，下一步要解决的就是如何根据输入的图片去确定这些系数的取值。

为了解决这一问题，3DMM 模型设计了一组面部相关属性的集合 $A = (a_0, \dots, a_k)$ ，包括性别、面部丰满度、眉毛粗细、双下巴等基本面部属性和表情这种可变的面部属性。然后通过对数据集中数据进行分析计算，确定这些属性与系数 $\{a_i, b_i\}$ 之间的线性关系。这样就可以通过对输入的图片进行属性提取，然后进行系数计算，初步确定系数 $\{a_i, b_i\}$ 的取值，使得 3D 模型与输入的图片在整体轮廓上基本相同。

在之前的处理中，我们已经获取到了数据集中的平均人脸模型的 $\{\bar{S}, \bar{T}\}$ ，这个平均人脸也对应着一组平均的属性 \bar{A} ，组合在一起就是 $\{\bar{S}, \bar{T}, \bar{A}\}$ 。同样，对于数据集中的任何一个人脸模型，都对应着一组 $\{S_i, T_i, A_i\}$ 。通过将 $\{S_i, T_i, A_i\}$ 与 $\{\bar{S}, \bar{T}, \bar{A}\}$ 进行运算，我们就可以得到若干组 $\{\Delta S_i, \Delta T_i, \Delta A_i\}$ ，进而计算出属性集合A中各属性与系数 $\{\alpha_i, \beta_i\}$ 之间的线性关系。其中 $\{\Delta S_i, \Delta T_i, \Delta A_i\}$ 的计算公式如下：

$$\Delta S_i = \sum_{i=1}^m \mu_i (S_i - \bar{S}), \quad \Delta T_i = \sum_{i=1}^m \mu_i (T_i - \bar{T}), \quad \Delta A_i = \{a_0 - \bar{a}, \dots, a_k - \bar{a}\}$$

当获取到各属性与系数 $\{\alpha_i, \beta_i\}$ 之间的关系之后，通过改变部分属性的值，我们就可以从原始人脸的基础上构建出多个不同的 3D 模型，如图 2 所示。

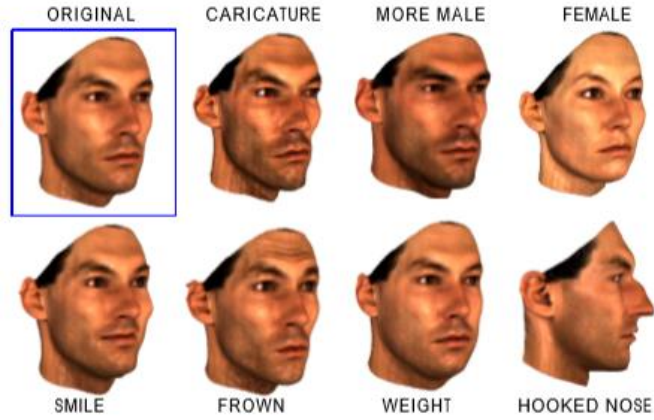


图 2 利用面部属性调节 3D 人脸模型

在第一个原始面部模型的基础上，我们分别改变不同的面部属性，比如表情、性别、体重等，就可以对部分相关系数 $\{\alpha_i, \beta_i\}$ 进行调整，而其他系数保持不变，进而实现对原始模型的精细化修正。

3.3 3D模型迭代优化

根据我们所设计 3D 形变模型，只要输入图片中人脸的面部属性，就可以得到一个基本的 3D 面部模型。在该初始模型的基础之上，我们还需要将其与输入的 2D 图片进行匹配计算和迭代优化，不断修正模型系数 $\{\alpha_i, \beta_i\}$ ，最终得到一个比较精确的 3D 面部模型，流程如图 3 所示。

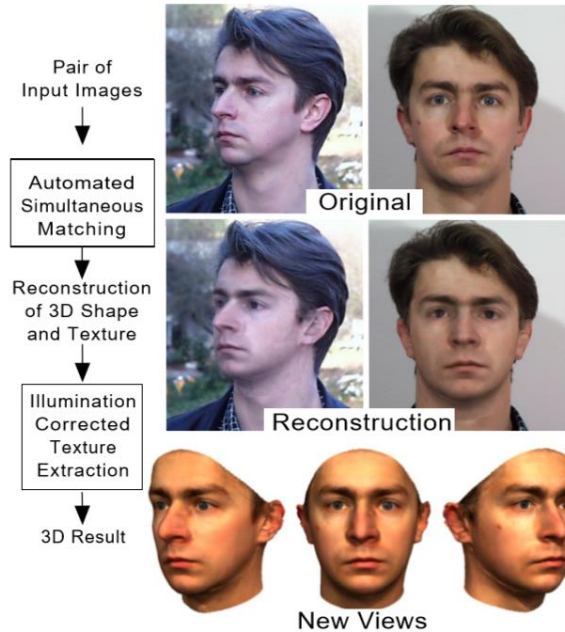


图 3 3D 模型重建流程

对于 2D 图片而言，每个像素点都可以用 3 个 RGB 值来表示，因此，一个 2D 图片可表示为 $I = ((r_1, g_1, b_1), \dots, (r_n, g_n, b_n))$ 。为了将初步构建好的 3D 模型映射到 2D 图片上，需要借助一组与图片生成相关的参数 P ，包括摄像机位置(方位角和仰角)、物体尺寸、光照强度等。根据这组参数和系数 $\{\alpha_i, \beta_i\}$ ，可以使用

透视投影技术和冯氏光照模型求得 3D 面部模型对应的 2D 图片数据 I_{model} 。

对于一个重建后的 3D 面部模型而言，该模型越准确，那么由该模型生成的 I_{model} 与原始图片 I_{input} 之间的欧几里得距离 E_I 就越小， E_I 的计算公式如下：

$$E_I = |I_{model} - I_{input}|^2 = \sum_{i=1}^n |(r_i, g_i, b_i)_{model} - (r_i, g_i, b_i)_{input}|^2$$

再考虑到一些噪声的影响，最终计算得到一个损失函数 E 。根据该损失函数计算结果，可以反向更新 $\{\alpha_i, \beta_i\}$ ，不断的减小 E 值，以求 $\min(E)$ 。在不断的迭代更新过程中， $\{\alpha_i, \beta_i\}$ 也会越来越精确。

4 基于特征点加权的模型优化

在 2D 图片和 3D 模型中，每一组对应的特征点都有其独特的语义。在迭代优化模型系数的过程中，损失函数的值对每个特征点对应系数的影响也都是不同的。因此，对每组对应的特征点之间的误差梯度进行单独分析是很有必要的。

4.1 基于特征点加权的 3DMM 模型优化策略

对于人脸的不同面部区域，比如嘴和鼻子，在做表情时，两者的变化幅度是不同的，嘴巴区域的变化幅度明显大于鼻子区域，如图 4 所示。因此，对于初步构建的 3D 面部模型，嘴巴区域的误差应该比较大，而鼻子区域的误差就比较小。而传统的优化方法将所有的特征点都等同处理，这就导致不同区域无法得到正确的误差调整，因此重建效果可能不太理想。



图 4 不同表情不同面部区域的变化差异

为了区别对待不同特征点，我们可以设计一个权重向量 $W = (w_1, \dots, w_n)$ ， w_i 表示第 i 个特征点的误差权重。这样一来，3D 模型的优化目标就从 $\min(|I_{model} - I_{input}|^2)$ 变为了 $\min(|W(I_{model} - I_{input})|^2)$ 。通过修改优化目标，我们就将不同特征点的差异化引入到了计算公式中，能够在迭代优化过程中更加合理的调控对不同面部区域的修改幅度，让模型更加精细化。

4.2 权重矩阵的计算方式

在上一节，我们已经介绍了特征点加权的意义，重点在于如何计算特征点的权重向量。这里，我们可以

记误差向量 D 为：

$$D = I_{model} - I_{input} = (d_1, \dots, d_n)$$

d_i 代表着第 i 个特征点在重建后的 3D 模型与原始 2D 图片中的差异。我们所希望正是在 d_i 较大时，相应的 w_i 也较大，这样相应特征点在 3D 模型中调整的幅度也较大，反之亦然。

因此，在特征加权模型中， w_i 的计算公式如下：

$$w_i = \frac{d_i + \varepsilon_1}{\varepsilon_2}$$

这里加入 ε_1 、 ε_2 两个参数是防止 w_i 过大或者过小，他们有多种取值，比如 $\varepsilon_1 = 3$ ， $\varepsilon_2 = 3.5$ 。这样一来，我们就可以计算出特征点的权重向量 W ，该向量在迭代的过程中根据各特征点的误差也在不断的更新。当然，为了防止过拟合，在迭代优化的过程中还可以加入正规方程，对系数进行一定约束修正。

5 总结与展望

本文主要介绍了用于面部模型重建的 3DMM 模型，以及基于特征点加权的模型优化策略，通过为每个特征点赋予不同的权重，调整其系数在迭代优化过程中的变化幅度，从而达到降低模型误差的目的。

该方法与深度学习中的注意力机制异曲同工，都是通过加权的方法区别对待不同的输入或者说是特征，进一步提高原模型的效果。不同之处在于注意力机制是借助大量数据和神经网络的训练得到的权重向量，然后与原来的输出进行融合，得到一个新的输出，并在反向传播的过程中不断更新计算权重的参数。而本文提出的方法则是根据 3D 模型转换成的 2D 数据与输入的 2D 图片数据在每个特征点上的误差计算出每个特征点的修改权重，然后根据该权重调整模型系数。

现如今，将 2D 电影自动转成 3D 电影的产业已经相当成熟了，我们相信，借鉴其中所使用的深度学习技术，对于 3D 重建技术的发展一定有更多的促进作用。

References:

- [1] Blanz V, Vetter T. A morphable model for the synthesis of 3D faces[C]//Proceedings of the 26th annual conference on Computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co., 1999: 187-194.
- [2] Blanz V, Vetter T. Face recognition based on fitting a 3D morphable model[J]. IEEE Transactions on pattern analysis and machine intelligence, 2003, 25(9): 1063-1074.
- [3] Yanga Y, Wu X J, Kittler J. Landmark Weighting for 3DMM Shape Fitting[J]. arXiv preprint arXiv:1808.05399, 2018.
- [4] Jiang L, Zhang J, Deng B, et al. 3D face reconstruction with geometry details from a single image[J]. IEEE Transactions on Image Processing, 2018, 27(10): 4756-4770.
- [5] Stylianou G, Lanitis A. Image based 3d face reconstruction: a survey[J]. International Journal of Image and Graphics, 2009, 9(02): 217-250.
- [6] Zhu X, Yan J, Yi D, et al. Discriminative 3D morphable model fitting[C]//Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on. IEEE, 2015, 1: 1-8.
- [7] Hu G, Yan F, Chan C H, et al. Face recognition using a unified 3d morphable model[C]//European Conference on Computer Vision. Springer, Cham, 2016: 73-89.