

单张正面人脸图像的三维人脸重建方法研究\*

田源\*\* 黄仁 杨吉云 孔润  
(重庆大学计算机学院,重庆 400044)

**摘要:**针对基于单张正面人脸图像进行三维人脸重建时所需脸部侧面深度信息缺失的问题,提出基于 BP 神经网络快速三维重建方法。通过建立 BP 神经网络估计出正侧面人脸数据的关系,从而由输入的正面数据得到侧面数据,并对 BP 算法做出改进,加速了算法的收敛,提高了拟合的精度。然后利用获取的人脸侧面数据调整 CANDIDE-3 人脸模型,生成近似图像中人脸的目标几何模型,最后通过纹理映射生成具有真实感的特定三维人脸模型。实验证明,这种重建方法快速简便,真实感强,较好地解决了基于单张照片重建时侧面深度信息缺失的问题。

**关键词:**三维人脸重建;单张正面人脸图像;BP 神经网络;CANDIDE-3 人脸模型

**中图分类号:**TP391.41      **文献标识码:**A

3D Face Reconstruction Based on Single Frontal Face Image\*

TIAN Yuan\*\* HUANG Ren YANG Jiyun KONG Run  
(College of Computer Science, Chongqing University, Chongqing 400044)

**Abstract:** According to the problem of missing side-face depth information on 3D face reconstruction from a single frontal-face image, a quick 3D reconstruction approach is proposed based on BP Neural Network. The approximate relationship between frontal-face data and side-face data can be estimated through establishment of BP neural network. Then side-face data can be obtained from input of frontal-face data. An improved BP algorithm is proposed. Then the convergence speed and the fitted accuracy are both increased. A target geometric model approximate to the face in the image is constructed by modifying CANDIDE-3 model based on frontal-face data and side-face data obtained before. Finally an individual realistic 3D face model is generated through texture mapping. The experiments show that this approach can solve the problem of missing side-face depth information well and it is fast and efficient, the result is realistic.

**Key words:** 3D face reconstruction; single frontal-face image; BP neural network; CANDIDE-3 face model

1 引言

人脸是情感表达和身份识别的重要载体,针对特定人脸的三维重建在影视动画、游戏娱乐、即时通信、医学手术等许多领域有广阔的应用前景。

Blanz 等<sup>[1]</sup>基于人脸库的形变模型方法对单幅人脸照片进行三维建模,使用激光扫描仪获得三维原型人脸并建立形变模型,真实感较好,但计算复杂。Jiang 等<sup>[2]</sup>对形变模型方法作了改进,使用特征点代替完整人脸模型,降低了计算量,但是基于形变模型的方法需要大量特征稠密对应的人脸样本,也限制了实际应用。程日彬等<sup>[3]</sup>通过正侧面照片生成三维的人脸模型,此类方法不如三维扫描设备得到的模型精度高,但是方法简单、成本低廉、适用面广。由于并不是总能获得人脸重建所需的完备信息,通常只有一张正面的人脸照片(如身份证上的照片等),所以基于单张照片的三维模型重建仍然有相当应用价值。但由于单张照片投影时脸部深度信息丢失的问题,重建效果始终不甚理想。

针对上述问题,本文在单张中性表情人脸正面照片的基础上,通过建立 BP 神经网络得到正侧面人脸数据的统计规律,从而可由正面数据估计得到侧面数据,然后调整 CANDIDE-3 模型,使一般人脸发生形变,以近似照片中的人脸,最后通过纹理映射生成具有真实感的模型。

2 一般人脸模型及特征点的选择

2.1 一般三维人脸模型的选择

基于正面人脸图像恢复三维人脸模型,获得的信息量相对较少,为了减小不必要的点对模型的影响,选取只有脸部数据并比较简单的 CANDIDE-3 人脸模型<sup>[4]</sup>,如图 1 所示。

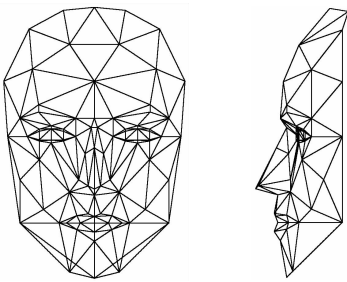


图1 CANDIDE-3 人脸模型  
Figure 1 CANDIDE-3 face model

2.2 人脸特征点的定义

2.2.1 人脸特征点的定义

根据人脸面部主要器官的特征,结合 MPEG-4 多媒体标准,选取 39 个关键特征点,如图 2 所示。

2.2.2 人脸正侧面数据测量项的定义

选取面部特征点后,分别对人脸正侧面照片中对应两组面部特征点的直线距离测量项进行定义。所选取的测量项目如表 1 所示。

\* 基金项目:重庆市自然科学基金项目(2009BA2024)  
\* \* 通讯作者:E-mail:tian\_614@yahoo.cn

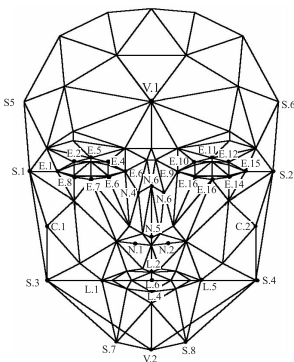


图2 人脸特征点的定义

Figure 2 Facial feature points

表1 正侧面图片中的测量项

Table 1 Measurements in the frontal and side picture					
测量项目	水平距离			垂直距离	
正面图片	E. 1-E. 13	E. 5-E. 9	E. 3-E. 11	V. 2-V. 1	V. 2-L. 5
	E. 7-E. 15	N. 1-N. 2	N. 4-N. 5	V. 2-N. 3	V. 2-N. 6
	C. 1-C. 2	S. 1-S. 2	S. 3-S. 4	V. 2-C. 1	V. 2-E. 5
	S. 5-S. 6	S. 7-S. 8	L. 1-L. 3	V. 2-S. 5	V. 2-S. 7
侧面图片	N. 3-V. 2	N. 3-L. 1	N. 3-L. 2	(侧面图片的垂直距离与正面图片的垂直距离理论上相同, 为了减少计算量, 故略去)	
	N. 3-L. 4	N. 3-L. 5	N. 3-N. 1		
	N. 3-N. 6	N. 3-E. 1	N. 3-E. 3		
	N. 3-C. 1	N. 3-S. 1	N. 3-S. 3		
	N. 3-S. 5	N. 3-S. 7	N. 3-V. 1		

3 正交人脸数据库的建立

本文使用“CAS-PEAL-R1 Face Database”人脸库中的样本,从中选择100名男性和100名女性的正侧面照片,作为正交人脸数据库的样本集。对每个样本(包括正侧面两张照片)按一定顺序手工定位39个关键特征点,使这些点在不同样本上达到位置上的一一对应。将数据库中的样本随机分为两组,一组用于训练(120个样本),一组用于测试(80个样本)。为了保证测量数据的准确性,加快神经网络的收敛速度,对正交人脸数据库中的样本作如下预处理:

- 1)头部姿态调整
  - ①正面照片,选取左右两眼内眼角 E. 5 和 E. 9,旋转照片使两点在同一水平线上;
  - ②侧面照片,仿照人体测量学头部测量的眼耳平面,选取点 E. 1 和点 S. 1,旋转使两点水平。
- 2)头部大小调整

分别测量正面照片的脸部高度 H1 和侧面照片的脸部高度 H2(即 V. 2-V. 1),得到正侧面照片的头高比例因子  $K = H1/H2$ 。通过将侧面数据乘以 K,进行头部大小比例的换算。
- 3)归一化处理

将各项数据测量项转化成与脸部高度的比值,使得到的数据全部介于(0,1]之间,再对神经网络进行训练。

4 BP 神经网络的训练

BP 神经网络模型采用误差反向传播学习算法,能将一组输入样本的 I/O 问题转变为一个非线性优化问题,以近似

实现从输入到输出的任意连续的非线性映射。

4.1 采用标准 BP 算法训练网络

为了提高拟合的精确性,对每个侧面项分别进行训练。由于每个样本的侧面拟合项共 15 项,所以需要 15 组神经网络完成训练。其中,每组网络采用输入层为 20,隐含层为 15,输出层为 1 的结构。

取误差精度  $E_p = 10^{-5}$ ,最大迭代步数为5 000步,学习速率  $\eta = 0.1$ ,采用未经优化的 BP 算法对(N. 3-V. 2)和(N. 3-L. 4)两项进行拟合。其训练情况,如表 2 所示。

表2 神经网络的训练情况			
Table 2 The training of network			
测量项	训练时间(ms)	迭代步数	E
N. 3-V. 2	69 873	5 000	$2.14 \times 10^{-2}$
N. 3-L. 4	62 385	5 000	$3.76 \times 10^{-2}$

由表 2 可知,未经优化的 BP 算法训练神经网络均是在达到最大迭代步数5 000步时停止,且实际误差 E 与期望的精度有很大差距,即不能有效、准确地拟合出侧面项。其余 13 项侧面项的拟合情况与(N. 3-V. 2)和(N. 3-L. 4)两项类似。

4.2 BP 算法的改进及试验

针对标准 BP 算法收敛速度慢的缺点,本文提出通过优化学习速率和形状因子实现对 BP 算法的改进。

由 BP 算法迭代式  $w_{k+1} = w_k - \eta \nabla E(w_k)$  可知,当学习速率满足

$$E(w_k - \eta_k \nabla E(w_k)) = E(w_k - \eta \nabla E(w_k)) \tag{1}$$

时,第 k+1 次迭代效果最佳。由此,可将求解最优学习速率等价为一维搜索过程。此外,本文选取的 Sigmoid 激励函数形式为:

$$\sigma(x) = \frac{1}{1 + e^{-\lambda x}} \tag{2}$$

其中,λ 为 Sigmoid 函数的形状因子。若 η 不变,误差函数 E 也可以看成是关于 λ 的函数。因此,求解最优形状因子同样可以等价为一维搜索过程。

$$E(\lambda_k) = \min_{\lambda > 0} E(\lambda) \tag{3}$$

三点二次插值是求解一维搜索问题效果最好的几种方法之一,可用于求解学习速率和形状因子的最优解。此方法不需要计算导数,利用 α<sub>1</sub>、α<sub>2</sub>、α<sub>3</sub> 三点的函数值 φ(α<sub>1</sub>)、φ(α<sub>2</sub>)、φ(α<sub>3</sub>)构造二次函数,得到插值公式。

$$\partial^* = \frac{1}{2} \frac{(\partial_2^2 - \partial_3^2)\varphi_1 + (\partial_3^2 - \partial_1^2)\varphi + (\partial_1^2 - \partial_2^2)\varphi_3}{(\partial_2 - \partial_3)\varphi_1 + (\partial_3 - \partial_1)\varphi_2 + (\partial_1 - \partial_2)\varphi_3} \tag{4}$$

通过三点二次插值法得到一系列点,最后得到函数一个极小点的近似。由此,提出算法 1。

- 1)算法 1——η(或 λ)加速
  - 步 0:选择要采用的误差函数及 E<sub>p</sub>,最大学习步数 K;
  - 步 1:初始化网络权向量 w<sub>0</sub>,令 k=0;
  - 步 2:对于权值向量 w<sub>k</sub> 前向计算得到误差 E<sub>k</sub>,如果 ||∇ E<sub>k</sub>|| < E<sub>p</sub>,则算法结束,否则转步 3;
  - 步 3:选取初始 α,ε,其中 α 可以是 η 或 λ;
  - 步 4:采用进退法确定初始搜索区间[α<sub>1</sub>, α<sub>2</sub>, α<sub>3</sub>],并令

$\varphi_1 = \varphi(\alpha_1), \varphi_2 = \varphi(\alpha_2), \varphi_3 = \varphi(\alpha_3)$ ;  
步5:令  $A = 2(\alpha_2 - \alpha_3)\varphi_1 + (\alpha_3 - \alpha_1)\varphi_2 + (\alpha_1 - \alpha_2)\varphi_3$ , 若  $A = 0$ , 则  $\alpha = \alpha_2$ ; 否则,  $\alpha = (-)\varphi_1 + (-)\varphi_2 + (-)\varphi_3/A$ , 且  $\varphi = \varphi(\alpha)$ ;  
步6:若  $|\alpha - \alpha_2| < \varepsilon$ , 则转步10; 否则, 若  $(\alpha - \alpha_2)(\alpha - \alpha_3) < 0$ , 则转步7, 否则转步8;  
步7:若  $\varphi < \varphi_2$ , 则令  $\alpha_1 = \alpha_2, \varphi_1 = \varphi_2$ , 转步9; 否则, 令  $\alpha_3 = \alpha, \varphi_3 = \varphi$ , 转步5;  
步8:若  $\varphi < \varphi_2$ , 则令  $\alpha_3 = \alpha_2, \varphi_3 = \varphi_2$ , 转步9; 否则, 令  $\alpha_1 = \alpha, \varphi_1 = \varphi$ , 转步5;  
步9:令  $\alpha_2 = \alpha, \varphi_2 = \varphi$ , 转步5;  
步10:得到局部最优的学习速率  $\eta$  或形状因子, 更新权值  $w_{k+1} = w_k - \eta \nabla E(w_k), k = k + 1$ , 若  $k \geq K$ , 则算法结束, 否则转步2。

此外, 在第  $k$  次权值不变的情况下, 误差可以看成是  $\eta$  和  $\lambda$  的函数, 即  $E = E(\eta, \lambda)$ 。因此可以尝试寻找  $(\eta, \lambda)$  的最优组合来加速学习过程。

采用 BFGS 算法求  $(\eta, \lambda)$  的最优解, 其中关于 Hesse 逆近似的校正为:

$$H_{k+1} = H_k + (1 + \frac{y_k^T H_k y_k}{S_k^T y_k}) \frac{S_k S_k^T}{S_k^T y_k} - \frac{S_k y_k^T H_k + H_k y_k S_k^T}{(S_k^T y_k)^2}$$
$$= (I - \frac{S_k y_k^T}{S_k^T y_k}) H_k (I - \frac{y_k S_k^T}{S_k^T y_k}) + \frac{S_k S_k^T}{S_k^T y_k}$$

(5)

其中,

$$s_k = (\eta_{k+1} - \eta_k, \lambda_{k+1} - \lambda_k)^T$$
$$y_k = \nabla E(\eta_{k+1}, \lambda_{k+1}) - \nabla E(\eta_k, \lambda_k)$$

(6)

(7)

由此可得到  $(\eta, \lambda)$  的最优解, 称其为  $\eta - \lambda$  加速法。

2) 算法2—— $\eta - \lambda$  加速

步0: 选择要采用的误差函数及  $E_p$ , 最大学习步数  $K$ ;

步1: 初始化网络权向量  $w_0$ , 令  $k = 0$ ;

步2: 对于权值向量  $w_k$  前向计算得到误差  $E_k$ , 如果  $||\nabla E_k|| < E_p$ , 则算法结束, 否则转步3;

步3: 由 BFGS 算法求出  $(\eta, \lambda)$  的最优组合, 得到学习速率  $\eta_k$  和形状因子  $\lambda_k$ , 更新权值  $w_{k+1} = w_k - \eta \nabla E(w_k), k = k + 1$ , 若  $k \geq K$ , 则算法结束, 否则转步2;

3) 数值试验

取  $E_p = 10^{-5}$ , 最大迭代步数为5 000步, 对于试验中  $\eta$  和  $\lambda$  固定的情况, 均取  $\eta = 0.1, \lambda = 0.3$ 。

(N.3-V.2)和(N.3-L.4)两项的神经网络训练加速前后比较情况, 如表3所示。由表3可知, 算法1和算法2都大大地提高了拟合(N.3-V.2)和(N.3-L.4)两项时网络收敛的速度, 同时也显著增加了学习的精度。 $\eta$ (或 $\lambda$ )加速的每次迭代都使它们达到局部最优, 有效地减少了迭代步数。尤其是 $\eta$ 加速法, 只需三步迭代就可达到误差精度的要求。 $\lambda$ 加速法对 $\lambda$ 初值的选取比较敏感, 若初值选取不当会导致效果不佳, 甚至失败, 但精度最高。 $\eta - \lambda$ 加速使用最少, 计算时间不到标准BP算法的3%, 精度也较高, 但误差函数关于 $(\eta, \lambda)$ 曲面的波动依然较大, 因此初值的选取依然比较重要。经过多次试验可知, 取迭代初值 $\eta = 0.1, \lambda = 5$ 效果较好。

表3 网络训练加速前后比较				
Table 3 The training of network before and after accelerated				
	算 法	训练时间 (ms)	迭代步数	$E$
N.3-V.2	标准 BP	69 873	5 000	$2.14 \times 10^{-2}$
	$\eta$ 加速	2 357	3	$5.83 \times 10^{-4}$
	加速	54 619	6	$3.81 \times 10^{-4}$
	$\eta$ -加速	1 988	28	$5.14 \times 10^{-4}$
N.3-L.4	标准 BP	62 385	5 000	$3.76 \times 10^{-2}$
	$\eta$ 加速	2 576	3	$4.77 \times 10^{-4}$
	加速	5 116	6	$3.17 \times 10^{-4}$
	$\eta$ -加速	1 705	25	$4.25 \times 10^{-4}$

试验发现, 此种仿真方法能快速精确地拟合样本数据, 较为真实地估计出脸部深度数据。

5 三维人脸重建

本文基于单张正面人脸图像的三维建模过程如图3所示, 主要分为两步: 1) 几何模型的重建: 调整 CANDIDE 模型上各点的位置, 获得目标人脸的几何模型; 2) 对调整后的目标几何模型进行纹理映射, 获得具有真实感的特定三维人脸。

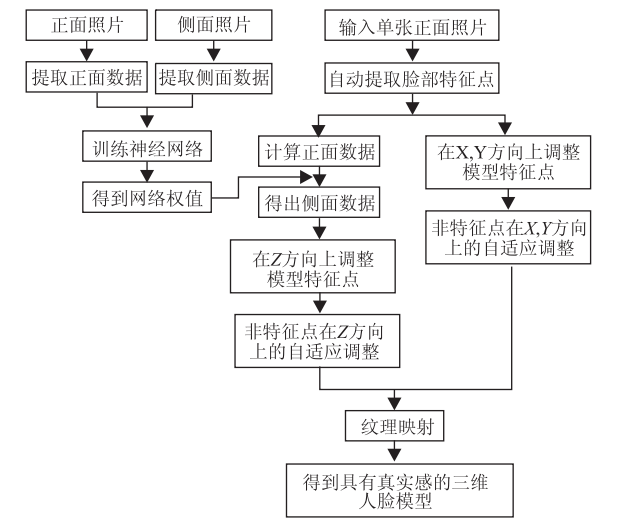


图3 三维人脸重建过程框架图

Figure 3 3D face reconstruction process

假设输入的人脸照片均为正面姿态, 均匀环境光照和中性表情, 采用改进的ASM算法<sup>[5,6]</sup>自动提取其39个特征点, 如图4所示。



图4 特征点自动提取

Figure 4 Automatic extraction of feature points

5.1 几何模型的重建

通过估计出正侧面人脸数据关系,可以在侧面方向调整一般人脸模型,因此几何模型的重建主要分为正面调整和侧面调整两步。

5.1.1 正面调整

在 X、Y 方向上调整 CANDIDE-3 模型,使其与二维目标人脸图像上的特征点相吻合。

1) 特征点的调整

CANDIDE-3 人脸模型上的任意一点 P(X,Y,Z)向其目标点 P'(X',Y',Z')作刚性运动,可以描述为:

$$[x',y',z',1]^T = [x,y,z,1]^T \cdot T(d_x,d_y,d_z) \cdot S(s_x,s_y,s_z) \cdot R(\theta_x,\theta_y,\theta_z)$$
 (8)

其中,T(d<sub>x</sub>,d<sub>y</sub>,d<sub>z</sub>)为平移矩阵,S(s<sub>x</sub>,s<sub>y</sub>,s<sub>z</sub>)为缩放矩阵,R(θ<sub>x</sub>,θ<sub>y</sub>,θ<sub>z</sub>)为旋转矩阵。特征点的调整即等价对这 9 个刚性运动参数的估计。

将 CANDIDE 模型上的特征点 P(X,Y,Z)垂直投影到目标人脸图像上可得投影点 P''(x'',y''),投影过程如图 5 所示。设 f 为三维坐标系原点与图像平面之间的距离,则:

$$x'' = f \frac{x}{z}, y'' = f \frac{y}{z}$$
 (9)

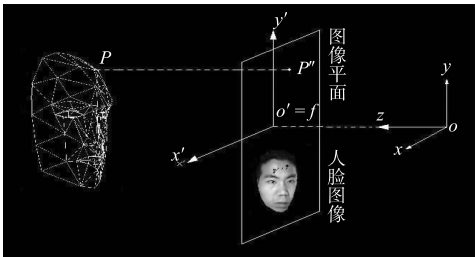


图 5 特征点的投影过程

Figure 5 Projection process of feature points

假定点 P'''(x''',y''')是在目标人脸图像上自动提取的并与 CANDIDE 模型上的点 P(X,Y,Z)对应的特征点,则可通过计算|P''-P'|得到平移量 d<sub>x</sub>、d<sub>y</sub>。此外,在 CANDIDE 模型上取左右两眼内眼角点 E.5、E.9,左右两眼中心点 O 及嘴部中心点 L.5,而 E.5'',

E.9'',O''及 L.5''为点 E.5,E.9,O 及 L.5 在目标人脸图像上的投影点,则可得比例因子 S<sub>x</sub> 和 S<sub>y</sub>:

$$\begin{cases} S_x = |E.5 - E.9| / |E.5'' - E.9''| \\ S_y = |O - L.5| / |O'' - L.5''| \end{cases}$$
 (10)

由于完成了端正姿态的预处理,故可忽略 θ<sub>x</sub>、θ<sub>y</sub>、θ<sub>z</sub> 等参数,只考虑缩放因子和平移量。另外,由于仅是正面调整,Z 方向的参数暂不考虑。

2) 非特征点的自适应调整

得到 CANDIDE 模型上特征点的空间坐标后,对于非特征点,采用散乱数据内插值<sup>[7]</sup>方法,利用特征点的变动来控制模型上非特征点的变动,以实现非特征点的自适应调整。从而对模型上的每一个点进行变换,得到正面调整后的人脸模型。

5.1.2 侧面调整

通过正面调整,CANDIDE 模型已在 X、Y 方向符合目标

人脸。在调整后的模型上取 V.1 和 V.2 两点,得到模型脸部的高度 H。根据调整后模型中特征点的坐标,可得所需的正面测量。用这些值除以 H 即可得到各测量项与脸高的比值。将这组数据输入神经网络即得到相应的拟合值,再乘以 H 即为各侧面特征点的深度数据,特征点的 Z 坐标确定,即可对模型进行侧面调整。调整结果如图 6 所示。

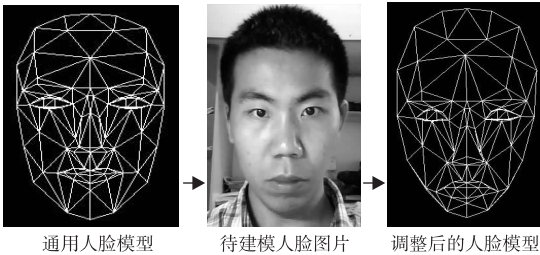


图 6 模型调整的结果

Figure 6 The result of model adjusting

5.2 真实感建模

几何调整后的模型是由许多三角形组成的网格体,缺乏真实感。针对现在纹理映射算法计算量大,算法比较复杂的缺点,选取采用江昊等<sup>[8]</sup>提出的基于特征约束点的纹理映射方法,利用三维网格剖分算法在纹理图像上建立基于特征点三维网格,进而通过求取质心坐标的方法计算出三维网格模型上所有特征点的纹理坐标,并完成整个三维网格模型的纹理映射,生成具有真实感的特定三维人脸模型。

6 实验结果与分析

6.1 实验结果

使用 VC++6.0 和 OpenGL 平台编程,选用普通单张正面中性人脸图像进行试验,生成具有真实感的三维人脸模型。通过旋转模型得到不同角度三维模型图像,并与对应角度的原始照片比较,有着较好的相似度,效果如图 7 所示。

6.2 与同类工作的比较

使用 50 幅 640×480 的真人照片进行测试,从算法实现和生成结果上与 Blanz 等<sup>[1]</sup>的方法进行比较,如表 4 所示。由表 4 可知,本文的方法是全自动的,更为快速、高效,真实感也较高。虽然输入要求较为严格,但还是比较普遍,容易满足。

表 4 方法比较

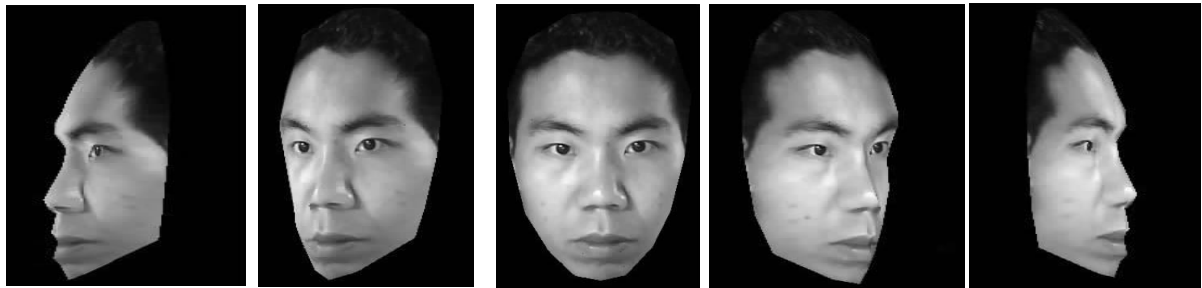
Table 4 Comparison of methods

	Blanz 等的方法 <sup>[1]</sup>	本文的方法
输入	任意姿态、光照的单张人脸照片	正面、均匀光照、中性表情单张人脸照片
自动化程度	需手工初始化	全自动
几何模型重建	用光流方法的形状误差估计来计算几何形状的主成分参数	调整主要特征点,非特征点自适应调整
纹理生成	用纹理误差估计来计算纹理主成分参数	基于特征约束点的纹理映射
系统效率	生成一张人脸需要 4.5 min(P4-2 GHz Work Station))	T3200@2.00 GHz 的 PC 上平均耗时不到 2 s

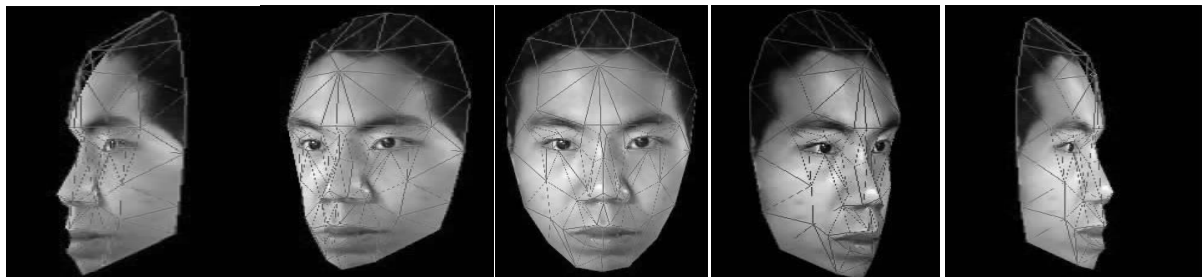




(a) 原始人脸照片



(b) 三维模型图像



(c) 带网络的三维模型图像

图 7 三维人脸重建效果

Figure 7 3D face reconstruction result

7 结束语

本文虽然沿用了基于图像重建的基本方法,但利用 BP 神经网络做了一些探索性工作,近似地找出了人脸正侧面数据的统计规律,对 BP 算法做了一定的改进,并由此得到了较好的实验效果。如果在此基础上建立的正交人脸库拥有类型更广、数量更多的样本,神经网络估计出的深度数据将更准确。

参考文献

[1] BLANZ V, VETTER T. A morphable model for the synthesis of 3D faces[C] // Proc of SIGGRAPH,1999:187-194.  
[2] JIANG D, HU Y, YAN S, et al. Efficient 3D reconstruction for face recognition [J]. Pattern Recognition,2005,38(6):787-798.

[3] 程日彬,周明全,李春龙. 基于二维图像的三维脸建模技术[J]. 计算机工程与应用,2006,42(3):33-35.  
[4] AHLBERG J. CANDIDE-3 an updated parameterized face[R]. Report No. LiTH-ISY-R-2326, Linkoping University, Sweden,2001.  
[5] COOTES T, TAYLOR C, COOPER D, et al. Active shape models-their training and application [J]. Computer Vision and Image Understanding,1995,61(1):38-59.  
[6] 范玉华,马建伟. ASM 及其改进的人脸面部特征定位算法[J]. 计算机辅助设计与图形学学报,2007,19(11):1 411-1 415.  
[7] PIGHIN F, HEEKER J, LISCHINSKI D. Synthesizeing realistic facial expressions from photographs[C] // Proc of SIGGRAPH, Orlando, FL, USA,1998:75-84.  
[8] 江昊,何明一. 基于特征约束点的纹理映射算法[J]. 计算机仿真,2009,26(1):237-239.

作者简介

田源(1985-),男,硕士研究生,主要研究方向:计算机图形学。  
黄仁(1962-),男,副教授、硕士生导师,计算机学会会员,主要研究方向:计算机图形学、图像处理。  
杨吉云(1975-),男,副教授,博士,主要研究方向:计算机图形学。  
孔润(1985-),男,硕士研究生,主要研究方向:人工智能。