

文章编号: 2095-4980(2016)02-0282-06

基于 PSO 的人脸姿态估计

雷 欢^{a,b}, 陈 虎^{a,b}, 吴志红^{a,b*}

(四川大学 a. 计算机学院; b. 视觉合成图形图像技术国防重点学科实验室, 四川 成都 610065)

摘 要: 提出了一种人脸姿态估计的新方法。根据三维扫描预先生成一个通用的模型, 基于三维形态模型采用粒子群优化算法对人脸姿态进行初步估计, 由特征点检测确定姿态大致范围, 再在初步估计的结果上进行修正迭代, 从而对人脸姿态进行精确估计。实验表明, 该方法在简化数学运算的基础上, 具有较好的估计效果, 平衡了计算复杂度与结果精确度。

关键词: 人脸姿态估计; 三维形态模型; 粒子群优化; 迭代

中图分类号: TN911.73

文献标识码: A

doi: 10.11805/TKYDA201602.0282

Face pose estimation based on PSO

LEI Huan^{a,b}, CHEN Hu^{a,b}, WU Zhihong^{a,b*}

(a. College of Computer; b. National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu Sichuan 610065, China)

Abstract: The face pose estimation based on Particle Swarm Optimization(PSO) is proposed. Firstly, a common model is generated by 3-D scanning. Then, PSO algorithm is used to estimate the face pose preliminarily based on the 3-D morphable model to acquire the pose range approximately with a detection of feature points. Finally, the rectification and iteration are applied with the preliminary result to estimate the face pose correctly. Experimental results show that the proposed method can obtain satisfying result by simplifying the math arithmetic, which keeps a balance between the complexity and the precision, and has an good effect on further studying of face identification.

Key words: face pose estimation; 3-D morphable model; Particle Swarm Optimization; iteration

人脸姿态估计指通过输入一幅或多幅人脸图像来确定该人脸在三维空间中姿态位置的过程。姿态问题涉及到头部在三维坐标系中绕 3 个轴旋转而造成的面部变化, 其中垂直于图像平面的 2 个方向的深度旋转会丢失部分面部信息, 增加人脸识别的难度。因此, 姿态估计在人脸识别领域也显得尤其重要。现有的人脸姿态估计方法大体可分为两类^[1]: 基于模型的方法和基于外观特征的方法。后者通过三维姿态与二维图像的某些特征(如灰度、纹理、光照等)之间的关系, 匹配输入图像与姿态样本的信息来得到姿态结果。本文使用前者进行姿态估计, 利用三维模型表示人脸形状, 并建立模型与图像的对应关系, 然后基于此对应关系求解姿态参数。

姿态估计中, 通常需要迭代多次进而得到一个最优解。一般的迭代优化算法有线性回归^[1]、基于 Mahalanobis 距离的迭代优化^[2]、牛顿迭代^[3]、ICP(Iterative Closest Point)算法^[4]等。本文所采用的迭代优化算法为粒子群优化(PSO)算法, 该算法于 1995 年由 Kennedy 和 Eberhart 共同开发, 源于对鸟群觅食行为的研究^[4]。较之其他算法, PSO 算法使用参数少, 不需要深厚的数学知识功底, 实现起来简单且通用性强。本文通过三维扫描构建人脸模型, 而后基于通过主成分分析^[5]所得到的模型参数, 使用 PSO 算法进行初步迭代, 得到人脸姿态基本位置, 最后基于该基本位置进行修正迭代, 从而得到人脸姿态的最优位置。

1 三维人脸模型

三维模型基于人脸的矢量空间表示, 这个矢量空间的定义使得样本集任何凸面结合的形态向量 s_i 都能描述一张真实的人脸:

收稿日期: 2015-04-21; 修回日期: 2015-05-23

基金项目: 国家自然科学基金资助项目(61202160)

*通信作者: 吴志红 email: wuzhihong@scu.edu.cn

$$S = \sum_{i=1}^m \alpha_i s_i \quad (1)$$

式中 α 为三维人脸模型形态系数。

本文所需三维模型的数据库包括年龄在 8 岁~62 岁不同种族的 100 名女性和 100 名男性。通过激光扫描仪扫描数据库中的人脸,得到三维模型。扫描通过以人头中心线为垂直轴线的圆柱坐标的形式来呈现人脸形态^[6],可表示为: $I(h, \phi) = r(h, \phi)^T$, 其中 ϕ 为角度转动变量, h 为垂直移动变量, r 为半径。

在人脸和参照脸之间,建立稠密的点到点的一致性构建三维形态模型的关键。这种关系通过矢量域 $V(h, \phi) = (\Delta h(h, \phi), \Delta \phi(h, \phi))^T$ 使得 2 次扫描中的点一一对应。参照脸 I_0 上所有 n 个点的坐标构成形状向量:

$$s_0 = (x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)^T \quad (2)$$

对参照脸的形态向量 s_i 集合进行主成分分析(本数据库中共有 200 人, $i=1,2,\dots,200$), 可得:

$$S = \bar{s} + \sum_{i=1}^{199} \alpha_i s_i \quad (3)$$

式中: $\bar{s} = \frac{1}{200} \sum_{i=1}^{200} s_i$; $\alpha_i = s_i - \bar{s}$; $A = (\alpha_1, \alpha_2, \dots, \alpha_{200})$; $C = \frac{1}{200} A A^T = \frac{1}{200} \sum_{i=1}^{200} \alpha_i \alpha_i^T$ 。计算 C 的特征向量: $s_1, s_2, \dots$,

C 的方差(从大到小)记为: $\sigma_{s,1}^2 \geq \sigma_{s,2}^2 \geq \dots$ 。

2 姿态估计

姿态估计的目的是找到描述三维人脸模型形态系数 α_i , 使得参照脸 R_p 生成的图像 I_{model} 与输入图像 I_{input} 尽可能相同, 从模型构建到姿态估计的整体流程如图 1 所示, 图中向量 S 即为式(3)中向量 S 的简易表示。

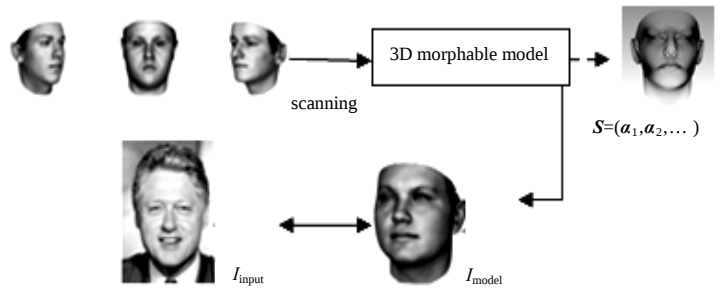


Fig.1 Process of building a 3D model

图 1 三维模型构建流程

2.1 坐标定义

在初始化时,手动定义 7 个人脸特征点(眼角、鼻尖等,根据所视区域的差异选取合适的特征点位置) $k_j(j=1,2,\dots,7)$,其位置二维坐标为 (p_{x,k_j}, p_{y,k_j}) ,该点对应在输入图像 I_{input} 上的坐标为 $(q_{x,j}, q_{y,j})$ 。对于轮廓线上的点,给予特殊处理。

设 n 个点中第 k 个点为物体中心,其三维坐标 (x_k, y_k, z_k) ,相机中坐标 $(\omega_{x,k}, \omega_{y,k}, \omega_{z,k})$,二维坐标 $(p_{x,k}, p_{y,k})$ 之间的相互转换关系如下:

$$(\omega_{x,k}, \omega_{y,k}, \omega_{z,k})^T = \gamma \theta \phi (x_k, y_k, z_k)^T + t_\omega \quad (4)$$

$$p_{x,k} = P_x + f \frac{\omega_{x,k}}{\omega_{z,k}}, p_{y,k} = P_y - f \frac{\omega_{y,k}}{\omega_{z,k}} \quad (5)$$

式中: γ 为物体绕相机轴线角度; θ 为物体与水平轴夹角; ϕ 为物体与垂直轴夹角; t_ω 为空间移动矢量; (p_x, p_y) 为光轴与图像平面的交点坐标; f 为相机焦距。

通过一张新的人脸图像及相关模型参数,经过初始化可拟合重建出该人脸的三维图像,具体过程如图 2 所示。

2.2 拟合

在拟合过程中,形态系数向量 α 和姿态角度变量 γ, θ, ϕ , 空间移动矢量 t_ω , 以及焦距 f 都得到优化。

2.2.1 代价函数

通过一张图像上每个点的位置坐标可得:输入图像 $I_{\text{input}}(x, y)$, 拟合重建得到图像 $I_{\text{model}}(x, y)$, 函数:

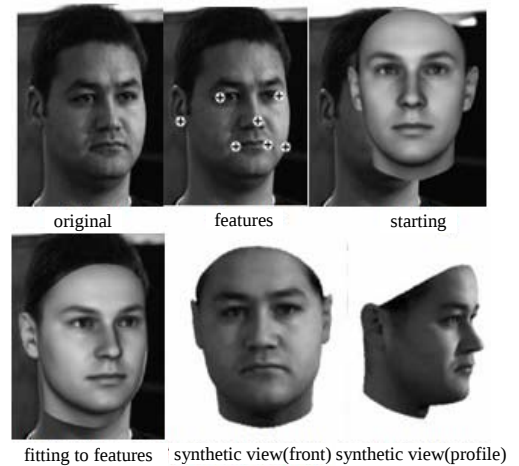


Fig.2 Reconstruction from a single face image

图 2 一张人脸图像的重建过程

$$E_I = \sum_{x,y} \|I_{\text{input}}(x,y) - I_{\text{model}}(x,y)\|^2 \quad (6)$$

匹配一张人脸图像的最终目的是使代价函数 E_I 最小。

首次迭代从手动定义的 7 个特征点开始, 即:

$$E_F = \left\| \begin{pmatrix} q_{x,j} \\ q_{y,j} \end{pmatrix} - \begin{pmatrix} p_{x,k_j} \\ p_{y,k_j} \end{pmatrix} \right\|^2 \quad (7)$$

给定输入图像 I_{input} , 特征点 F , 在迭代的过程中找到满足最佳拟合度 $p(\alpha, \rho | I_{\text{input}}, F)$ 的模型参数, 根据贝叶斯规则:

$$p(\alpha, \rho | I_{\text{input}}, F) \sim p(I_{\text{input}}, F | \alpha, \rho) \cdot p(\alpha, \rho) \quad (8)$$

式中 ρ 为一个容纳了本节开始介绍的除 α 之外的其他 5 个参数的向量。

若不考虑这些参数变量之间的联系, 则有:

$$p(I_{\text{input}}, F | \alpha, \rho) \cdot p(\alpha, \rho) \sim p(I_{\text{input}} | \alpha, \rho) \cdot p(F | \alpha, \rho) \cdot p(\alpha) \cdot p(\rho) \quad (9)$$

式中 $p(\alpha), p(\rho)$ 服从正态分布, 定义相应的平均值 $\bar{\alpha}_i, \bar{\rho}_i$, 标准差 $\sigma_{S,i}, \sigma_{R,i}$, 即: $p(\alpha) = e^{-\frac{1}{2} \sum_i \frac{(\alpha_i - \bar{\alpha}_i)^2}{\sigma_{S,i}^2}}$,

$p(\rho) = e^{-\frac{1}{2} \sum_i \frac{(\rho_i - \bar{\rho}_i)^2}{\sigma_{R,i}^2}}$ 。 $p(I_{\text{input}} | \alpha, \rho)$ 在每个像素上服从高斯噪声分布, 即: $p(I_{\text{input}} | \alpha, \rho) = e^{-\frac{1}{2\sigma_I^2} E_I}$; 同理,

$p(F | \alpha, \rho) = e^{-\frac{1}{2\sigma_F^2} E_F}$, σ_I, σ_F 分别为输入图像和特征点在高斯分布中的标准差, 即有:

$$p(\alpha, \rho | I_{\text{input}}, F) = e^{-\frac{1}{2\sigma_I^2} E_I} e^{-\frac{1}{2\sigma_F^2} E_F} e^{-\frac{1}{2} \sum_i \frac{(\alpha_i - \bar{\alpha}_i)^2}{\sigma_{S,i}^2}} e^{-\frac{1}{2} \sum_i \frac{(\rho_i - \bar{\rho}_i)^2}{\sigma_{R,i}^2}}$$

定义总代价函数:

$$E = -2 \ln p(\alpha, \rho | I_{\text{input}}, F)$$

即:

$$E = \frac{1}{\sigma_I^2} E_I + \frac{1}{\sigma_F^2} E_F + \sum_i \frac{(\alpha_i - \bar{\alpha}_i)^2}{\sigma_{S,i}^2} + \sum_i \frac{(\rho_i - \bar{\rho}_i)^2}{\sigma_{R,i}^2} \quad (10)$$

2.2.2 PSO 优化

PSO 算法的基本思想是: 将所优化问题的每一个解称为一个微粒, 每个微粒在 d 维搜索空间中以一定的速度飞行, 通过适合度(代价)函数来衡量微粒的优劣, 微粒根据自己的飞行经验以及其他微粒的飞行经验, 来动态调整飞行速度, 以期向群体中最好微粒的位置飞行, 从而得到所优化问题的最优解^[7-8]。

标准 PSO 算法描述为: 假设搜索空间为 d 维, 种群中有 $PNum$ 个粒子, 则群体中的粒子 i 在第 t 次迭代的位置表示为一个 d 维向量 $X_i(t) = (x_{i1}(t), x_{i2}(t), \dots, x_{id}(t))$ 。粒子的速度定义为位置的方向改变, 用向量 $V_i(t) = (v_{i1}(t), v_{i2}(t), \dots, v_{id}(t))$ 表示。粒子 i 的速度和位置更新通过式(11)和(12)可得。

$$v_{ij}(t+1) = wv_{ij}(t) + c_1 rand1_{ij}(pbest_{ij}(t) - x_{ij}(t)) + c_2 rand2_{ij}(gbest_{ij}(t) - x_{ij}(t)) \quad (11)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (12)$$

式中: w 为惯性系数; $rand1, rand2 \in [0, 1]$, 为服从均匀分布的 2 个独立随机数; c_1, c_2 为加速系数。

在使用 PSO 优化之前, 本文预先自动进行人脸检测, 给出标记点以初步确定人脸轮廓, 如图 3 所示。

随后开始进行 PSO 优化初步迭代以及修正迭代, 具体流程如图 4 所示。

本文在姿态估计中, 抽象地用一个 d 维向量来表示一张人脸, 每次迭代更新当前粒子的 d 个参数 $x_{i1}, x_{i2}, \dots, x_{id}$, 得到下一次迭代粒子群体, 计算每次迭代粒子群中各粒子的代价函数 E , 选择最佳适合度粒子作为当前粒子。以图 3 的



Fig.3 Face detection
图 3 人脸检测

这张正面人脸(0°, 0°)为例, 其迭代滤波过程中各变量的变化状态如表 1 所示。

表 1 迭代滤波过程(初次)

times	cost function	pitch/(°)	yaw/(°)
100	-1 881.780	6.46	5.68
200	-1 684.230	5.94	4.24
300	-1 626.190	5.88	3.73
400	-1 576.750	4.62	3.55
2 400	-453.809	4.96	1.32

由于迭代次数较多, 相邻 2 次迭代之间变化极小, 为便于分析数据, 本实验以 100 为间隔显示迭代数据结果, 表 1 给出了初步估计时前 4 次以及最后 1 次的显示数据结果。

迭代终止策略:

迭代终止的情况有 2 种: 第 1 种, 达到最大迭代次数 K ; 第 2 种, 由于 PSO 算法自身决定了局部最优解 FitBest 会随着迭代次数的增加不断趋近于全局最优解 GBestIndex, 直到 2 个值重合(迭代次数足够大)。所以在迭代的后期, 相邻 2 次粒子代价函数值 E 的差值会越来越小, 甚至趋于 0。所以当其差值 ΔE 小于某个阈值 g 时, 可以近似认为寻优进展停止, 便可在未达到最大迭代次数 K 之前终止迭代。

因此, 基于 PSO 的人脸姿态估计优化过程如下:

- 1) 确定初始粒子群及其 d 维参数;
- 2) 计算所有粒子的代价函数值 E , 选取 E 值最小(最佳适合度)的粒子作为当前粒子, 根据终止条件判断是否继续迭代, 满足则终止迭代;
- 3) 根据式(11)和(12)及当前粒子更新位置和速度, 并根据代价函数值 E 确定局部最优 FitBest 和全局最优 GBestIndex, 得到下一粒子群, 转到步骤 2);
- 4) 迭代结束, 全局最优解对应的即为最佳人脸姿态位置。

3 实验结果

实验在 Intel(R) Core(TM)i5CPU, 3.40 GHz, 8 GB 内存, VS2012 环境下编程所得。本实验测试采用的数据库包括处于不同年龄层次的 25 名男性, 25 名女性, 每人 10 种不同角度的姿态, 共 500 张 816×624 的彩色人脸图像, 在三维形态模型的基础上, 通过 PSO 迭代优化计算出人脸姿态位置。

实验的通用参数设定: a) 初始种群粒子数设定为 24; b) 粒子的速度限制在 $V_{\max}=0.05$; c) 初步姿态估计迭代次数设定为 2 500; d) 精确姿态估计迭代此处设定为 4 000。

实验采用数据库人脸部分呈现如图 5 所示。

根据实际情况中人脸的一般姿态位置, 本实验主要从 x 轴和 y 轴 2 个方向对人脸姿态角度进行估计^[9-10], 为便于统计分析, 把 x 轴和 y 轴的角度转动划分为 2 个区域:

x 轴(俯仰角度): 微转 $[-15^\circ, 15^\circ]$; 偏转 $[-30^\circ, -15^\circ]$, $[15^\circ, 30^\circ]$;

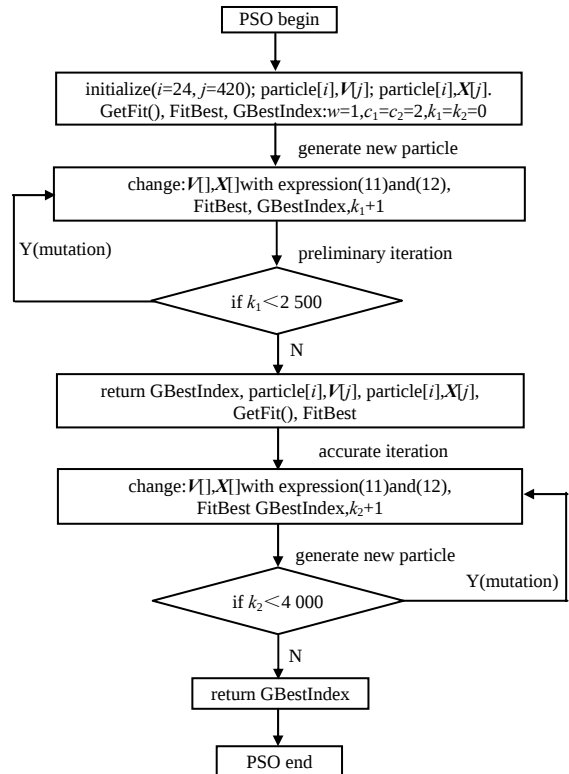


Fig.4 Process of PSO algorithm
图 4 PSO 算法流程



Fig.5 Partial experimental data
图 5 部分实验数据

表 2 姿态对应角度平均误差
Table2 Average error of different poses

yaw	average error/(°)	
	pitch $[-15^\circ, 15^\circ]$	pitch $[-30^\circ, -15^\circ][15^\circ, 30^\circ]$
$[-30^\circ, 30^\circ]$	2.29	1.87
$[-60^\circ, -30^\circ][30^\circ, 60^\circ]$	4.72	3.17

y 轴(旋转角度): 微转 $[-30^\circ, 30^\circ]$; 偏转 $[-60^\circ, -30^\circ]$, $[30^\circ, 60^\circ]$ 。

得到实验数据如表 2 所示。

综合表 1 及其他实验结果来看, 基于本文所构建的三维模型, 用 PSO 算法进行姿态估计时, x 轴方向上的角度变化较小, 因此本文将 y 轴角度偏转对精确度的影响描述如图 6 所示。

在微转区域内, 线性回归算法的平均误差精确度在 $1.58^\circ \sim 2.80^\circ$ 之间^[1], 牛顿方法^[3]没有给出具体角度区域的误差精确度, 其在正面方向上的误差为 1.6° 。在此区域内, 文中算法精确度优势不明显, 但在数学运算复杂度方面较之其他算法有显著优势, 如本文算法在更新每一次迭代时只需通过式(11)和式(12)来计算, 牛顿方法^[3]则需要求一个庞大的二次偏微分方程才能得出下一次迭代的参数, 且较之偏转区域, 文中算法有一个较好的估计结果。

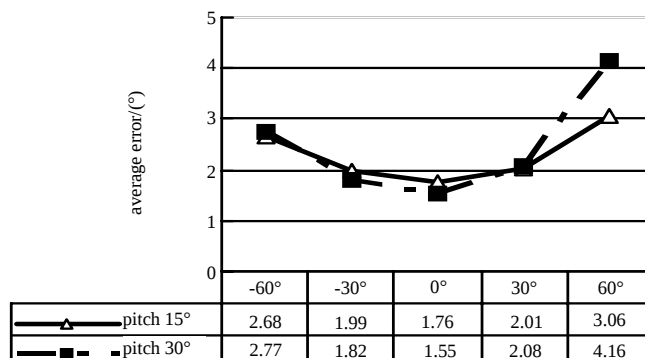


Fig.6 Effect of yaw to average error

图 6 y 轴角度对平均误差的影响

4 结论

本文成功地将粒子群优化算法应用到人脸姿态估计中, 在进行 PSO 优化过程中, 通过初步和修正 2 次迭代过程得到最优估计结果。为了克服粒子群优化算法的早熟收敛问题, 本文对迭代过程中的全局最优解进行了变异。算法提取可见特征点, 不依赖于纹理信息, 对光照条件无特殊要求, 在简化运算复杂度的基础上, 估计结果在小角度偏转区域稍优于其他算法, 在较大偏转角度区域有明显优势, 具有一定的实用价值。对于实验中出现的 y 轴偏角对精确度影响较大的问题, 在今后的研究中还有待对该算法进行进一步改进提升以得到更精确的估计。本文仅作姿态估计, 未对纹理及光照信息进行说明, 在以后的研究工作中有待在这方面加强, 逐渐完善整个人脸识别过程。

参考文献:

- [1] 牛晓霞. 基于三维模型的人脸姿态估计[J]. 微处理机, 2014,12(6):62-65. (NIU Xiaoxia. Face pose estimation based on 3D model[J]. Microprocessors, 2014,12(6):62-65.)
- [2] 曾慧,穆志纯,袁立. 基于三维模型的人脸姿态估计方法[J]. 计算机工程, 2011,37(9):1-3. (ZENG Hui,MU Zhichun, YUAN Li. Face pose estimation method based on 3-D model[J]. Computer Engineering, 2011,37(9):1-3.)
- [3] BLANZ V,VETTER T. Face recognition based on fitting a 3-D morphable model[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013,25(9):1063-1074.
- [4] 梁国远,查红彬,刘宏. 基于三维模型和仿射对应原理的人脸姿态估计方法[J]. 计算机学报, 2005,28(5):792-800. (LIANG Guoyuan,ZHA Hongbin,LIU Hong. Face pose estimation based on 3D models and affine correspondences[J]. Chinese Journal of Computers, 2005,28(5):792-800.)
- [5] 李冠楠,李强. 一种基于人脸核心特征的 PCA 人脸识别算法及应用[J]. 电子器件, 2012,35(5):607-610. (LI Guannan, LI Qiang. Implementation and application of one PCA face recognition algorithm based on the core features of the face[J]. Chinese Journal of Electron Devices, 2012,35(5):607-610.)
- [6] RUANGYAM P,COVAVISARUCH N. 3-dimensional face pose and shape estimation based on relaxed model fitting optimization[C]// Proceedings of 2010 IEEE 17th International Conference on Image Processing. Hong Kong:IEEE, 2010:1533-1536.
- [7] 唐俊. PSO 算法原理及应用[J]. 计算机技术与发展, 2010,20(2):213-216. (TANG Jun. Principle and application of PSO algorithm[J]. Computer Technology and Development, 2010,20(2):213-216.)
- [8] 李晓鹏. 基于 PSO 训练 SVM 的人脸识别[J]. 电脑知识与技术, 2007,2(8):498-500. (LI Xiaopeng. Face recognition based on Support Vector Machine trained with Particle Swarm Optimization[J]. Computer Knowledge and Technology, 2007,2(8):498-500.)
- [9] LIANG Xiaoni,TONG Weiqing. Face pose estimation using near-infrared images[C]// 2012 International Conference on Communication Systems and Network Technologies. Rajkot:IEEE, 2012:216-220.