



(12) 发明专利申请

(10) 申请公布号 CN 104598936 A

(43) 申请公布日 2015. 05. 06

(21) 申请号 201510090459. X

(22) 申请日 2015. 02. 28

(71) 申请人 北京畅景立达软件技术有限公司
地址 100080 北京市海淀区海淀中街 16 号
中关村公馆 F 座 1509 室

(72) 发明人 李军 张军 程东岳

(51) Int. Cl.

G06K 9/66(2006. 01)

G06K 9/46(2006. 01)

G06K 9/00(2006. 01)

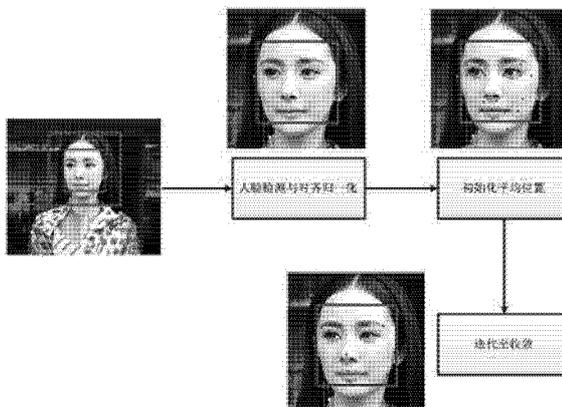
权利要求书4页 说明书9页 附图2页

(54) 发明名称

人脸图像面部关键点的定位方法

(57) 摘要

本发明公开了一种人脸图像面部关键点的定位方法,包括如下步骤:基于互联网数据,标注形成一个具有多个面部关键点的人脸图像训练集;基于人脸检测器检测到的人脸检测框将训练集中标注了面部关键点的人脸图像进行对齐,得到对齐后的每张人脸图像上面部关键点的位置和所有人脸图像面部关键点的平均位置;分别学习每个面部关键点的判别特征表示,将每个面部关键表示为一个 10 维的特征向量;基于偏最小二乘回归学习面部关键点的回归定位模型;对于输入的测试人脸图像,经过人脸检测和对齐、归一化、面部关键点位置初始化以及迭代回归处理得到最终的面部关键位置并输出。



1. 一种人脸图像面部关键点的定位方法,其特征在于:

- (1) 设置人脸图像训练集;
- (2) 对人脸图像训练集中的人脸图像进行面部关键点位置标注;
- (3) 对人脸图像训练集中标注后的每张人脸图像进行检测和对齐处理;
- (4) 对人脸图像训练集中对齐后的每张人脸图像进行归一化处理;
- (5) 对面部关键点进行判别特征学习;
- (6) 为每张人脸图像 t_n 进行面部关键点特征提取;
- (7) 训练得到面部关键点定位模型;
- (8) 对输入的人脸图像利用定位模型定位得到面部关键点位置。

2. 如权利要求 1 所述的方法,其中步骤 (1) 具体为:收集 N 张互联网人脸图像,形成人脸图像训练集 $T = \{t_n, n = 1, 2, \dots, N\}$ 。

3. 如权利要求 2 所述的方法,其中步骤 (2) 具体为:为人脸图像训练集中的每张人脸图像进行面部关键点位置标注并得到关键点坐标,其中标注的面部关键点包括人脸器官和面部轮廓。

4. 如权利要求 3 所述的方法,其中的人脸器官包括眉毛中心、眼睛中心、鼻子中心和嘴巴中心,面部轮廓为下面部轮廓,共标记 28 个面部关键点。

5. 如权利要求 4 所述的方法,其中以每张人脸图像 t_n 的左上角为原点 O ,以水平右向为 X 轴正方向,以垂直下向为 Y 轴正方向,建立直角坐标系,得到标注的每个面部关键点 i 在该坐标系中的位置 $(x_{n,i}, y_{n,i})$,其中。

6. 如权利要求 5 所述的方法,其中步骤 (3) 中的检测和对齐处理具体为:利用人脸检测器对每张人脸图像进行人脸检测,基于人脸检测器检测到的人脸检测框对训练集中的每张人脸图像进行对齐处理。

7. 如权利要求 6 所述的方法,其中基于人脸检测器检测到的人脸检测框对训练集中的每张人脸图像进行对齐处理具体为:人脸检测器在每张人脸图像上检测到 $L_n * L_n$ 像素大小的矩形人脸检测框,其中心在每张人脸图像 t_n 的直角坐标系中的坐标为 $(w_{n,x}, w_{n,y})$,每张人脸图像的人脸检测框的大小随其对应的人脸图像改变,以该人脸检测框的中心为参考点,在每张人脸图像 t_n 上抠取 $1.5L_n * 1.5L_n$ 像素大小的矩形区域。

8. 如权利要求 7 所述的方法,其中步骤 (4) 中的归一化处理具体为:对于抠取的 $1.5L_n * 1.5L_n$ 像素大小的矩形区域,将其归一化到 $150 * 150$ 像素大小矩形区域,以人脸检测框的中心为原点 O ,水平方向为 X 轴,垂直方向为 Y 轴,则面部关键点 i 的坐标归一化为:

$(X_{n,i}, Y_{n,i}) = (\frac{x_{n,i} - w_{n,x}}{100/L}, \frac{y_{n,i} - w_{n,y}}{100/L})$, 对齐归一化后的每张人脸图像 t_n 的所有面部关键点的

位置记为 I_n , 则 $I_n = [X_{n,1}; Y_{n,1}; \dots; X_{n,i}; Y_{n,i}; \dots; X_{n,28}; Y_{n,28}] \in R^{56}$ 。

9. 如权利要求 8 所述的方法,其中步骤 (5) 中的判别特征学习具体为:在对齐归一化后的人脸图像训练集上为每个面部关键点 i 单独地进行判别特征学习。

10. 如权利要求 9 所述的方法,其中在对齐归一化后的人脸图像训练集上为每个面部

关键点 i 单独地进行判别特征学习具体为：

(5.1) 分别获得每个面部关键点 i 的正例样本和反例样本；

(5.2) 对每个面部关键点 i 得到的所有正例样本和所有反例样本分别提取 59 维的 Uniform LBP 特征；

(5.3) 计算每个面部关键点 i 得到的所有样本的均值 Γ_i 、正例样本均值 $\Gamma_{U,i}$ 和反例样本的均值 $\Gamma_{D,i}$ ；

(5.4) 计算每个面部关键点 i 正例样本和反例样本的散度矩阵 $\Sigma_{U,i}$ 和 $\Sigma_{D,i}$ 以及类间散度矩阵 Φ_i ；

(5.5) 基于 Fisher 准则学习每个面部关键点 i 的判别投影矩阵 V_i 。

11. 如权利要求 10 所述的方法, 其中步骤 (5.1) 具体为: 在对齐归一化后的每幅人脸图像 t_n 上, 为每个关键点 i 截取 $7*7$ 像素大小的窗体作为正例, 在距离该关键点 i 水平或垂直方向距离为 3 到 14 个像素的范围内随机采样四个 $7*7$ 像素大小的窗体反例, 为每张人脸图像的每个关键点 i 采样得到了一个正例样本和四个反例样本。

12. 如权利要求 11 所述的方法, 其中步骤 (5.2) 具体为: 关键点 i 的所有 N 个正例样本的 Uniform LBP 特征记为: $U_i = [u_{1,i}, u_{2,i}, \dots, u_{n,i}, \dots, u_{N,i}] \in R^{59 \times N}$, 其中 $u_{n,i}$ 为人脸图像 t_n 上的关键点 i 的正例样本的 Uniform LBP 特征, 所有反例样本的 Uniform LBP 特征记为: $D_i = [d_{1,i,1}, d_{1,i,2}, d_{1,i,3}, d_{1,i,4}, \dots, d_{n,i,m}, \dots, d_{N,i,4}] \in R^{59 \times 4N}$, 其中 $d_{n,i,m}$ 为人脸图像 t_n 的关键点 i 的第 m 个反例样本的 Uniform LBP 特征。

13. 如权利要求 12 所述的方法, 其中步骤 (5.3) 具体为:

$$\Gamma_i = \frac{\sum_{n=1}^N u_{n,i} + \sum_{n=1, m=1}^{N, m=4} d_{n,i,m}}{5N}, \quad \Gamma_{U,i} = \frac{\sum_{n=1}^N u_{n,i}}{N}, \quad \Gamma_{D,i} = \frac{\sum_{n=1, m=1}^{N, m=4} d_{n,i,m}}{4N}.$$

14. 如权利要求 13 所述的方法, 其中步骤 (5.4) 具体为: ,

$$\Sigma_{U,i} = \sum_{n=1}^N (u_{n,i} - \Gamma_{U,i})(u_{n,i} - \Gamma_{U,i})^T, \quad \Sigma_{D,i} = \sum_{n=1, m=1}^{N, m=4} (d_{n,i,m} - \Gamma_{D,i})(d_{n,i,m} - \Gamma_{D,i})^T, \\ \Phi_i = N * (\Gamma_{U,i} - \Gamma_i)(\Gamma_{U,i} - \Gamma_i)^T + 4N * (\Gamma_{D,i} - \Gamma_i)(\Gamma_{D,i} - \Gamma_i)^T.$$

15. 如权利要求 14 所述的方法, 其中步骤 (5.5) 具体为: 基于 Fisher 准则, 最大化判别

投影的可分性, 其优化目标为 $\max_{v_i} \frac{\text{Trace}(v_i^T \Phi_i v_i)}{\text{Trace}(v_i^T (\Sigma_{U,i} + \Sigma_{D,i}) v_i)}$, s.t. $v_i^T v_i = I_{59}$, 其中 v_i 的最优

解为 $\tilde{\Phi}_i = (\Sigma_{U,i} + \Sigma_{D,i})^{-1} \Phi_i$ 的最大特征值对应的特征向量, 将上述过程重复 10 次, 则每个关键点 i 得到 10 个投影向量, 记为投影矩阵 V_i : $V_i = [v_{i,1}, v_{i,2}, \dots, v_{i,10}] \in R^{59 \times 10}$, 为所有 28 个面部关键点重复以上学习过程, 共得到 28 个投影矩阵。

16. 如权利要求 15 所述的方法, 其中步骤 (6) 具体为: 对于对齐归一化后的人脸图像 t_n , 分别截取以 28 个面部关键点为中心的 $7*7$ 像素大小的窗口, 为每个窗口提取 59 维 Uniform LBP 特征, 则在关键点 i 处提取的 Uniform LBP 特征记为: $\text{LBP}(X_{n,i}, Y_{n,i}) \in R^{59}$

, 关键点 i 的判别特征 $f_{n,i}$ 为: $f_{n,i} = V_i^T LBP(X_{n,i}, Y_{n,i}) \in R^{10}$, 人脸图像 t_n 的 28 个面部关键点可表示为 $f_n = [f_{n,1}; f_{n,2}; \dots; f_{n,i}; \dots; f_{n,28}] \in R^{280}$, 为一个 280 维的列向量, 用函数 $G(l_n)$ 表示对对齐归一化后的人脸图像 t_n 在 28 个面部关键点坐标为 l_n 时所提取出的 280 维特征向量, 则对齐归一化后的人脸图像 t_n 的面部关键点特征提取为:

$$G(l_n) = [f_{n,1}; f_{n,2}; \dots; f_{n,i}; \dots; f_{n,28}]。$$

17. 如权利要求 16 所述的方法, 其中步骤 (7) 具体为:

(7.1) 将对齐归一化后的训练集中所有人脸图像的 28 个面部关键点位置的平均值记为 l_0 , 则

$$l_0 = \left[\frac{\sum_{n=1}^N X_{n,1}}{N}, \frac{\sum_{n=1}^N Y_{n,1}}{N}, \dots, \frac{\sum_{n=1}^N X_{n,i}}{N}, \frac{\sum_{n=1}^N Y_{n,i}}{N}, \dots, \frac{\sum_{n=1}^N X_{n,28}}{N}, \frac{\sum_{n=1}^N Y_{n,28}}{N} \right] \in R^{56}$$

, 将训练集中 M 幅对齐归一化后的人脸图像的平均位置记为 X_0 , 则 $X_0 = [l_0; l_0; \dots; l_0] \in R^{56 \times N}$, 将训练集中 M 幅对齐归一化后的人脸图像的人工标注的面部关键点标记位置记为 X_g , 则 $X_g = [l_1; l_2; \dots; l_n; \dots; l_N] \in R^{56 \times N}$;

(7.2) 利用迭代回归学习方式建立面部关键点的定位模型, 令 $\Delta_k = X_g - X_{k-1}, k \geq 1$, 其中 Δ_k 表示训练集中对齐归一化后的人脸图像的人工标注的面部关键点位置 X_g 与第 $k-1$ 次迭代结果 X_{k-1} 之间的差异, 令 $G(X_k)$ 表示在第 k 次迭代时, 训练集中所有对齐归一化后的人脸图像在面部关键点位置处提取出的特征向量, 则 $G(X_k) = [[G(l_{k,1})], \dots, [G(l_{k,n})], \dots, [G(l_{k,N})]]$, 其中 $l_{k,n}$ 表示训练集中的对齐归一化后的人脸图像 t_n 在第 k 次迭代时面部关键点的位置, $G(l_{k,n}) = [f_{k,n,1}; f_{k,n,2}; \dots; f_{k,n,i}; \dots; f_{k,n,28}] \in R^{280}$, 其中 $f_{k,n,i}$ 表示对齐归一化后的人脸图像 t_n 在第 k 次迭代时面部关键点 i 的判别特征;

(7.3) 利用迭代回归学习来学习 Δ_k 与 $G(X_{k-1})$ 之间的回归函数, 将面部关键点定位建模为: $\Delta_k = A_k * G(X_{k-1}) + [b_k; b_k; \dots; b_k]$, 该迭代回归学习的目标函数为: $\min_{A_k, b_k} \|\Delta_k\|_F$, 用偏最小二乘求解模型参数: $A_k \in R^{56 \times 280}$ 和 $b_k \in R^{56}$, 训练过程共迭代 M 次, 其中 $M = 5$, 通过 M 次迭代回归学习确定定位模型参数 A_k 和 $b_k, k = 1, 2, 3, 4, 5$, 其中每次迭代得到一组对应的模型参数 (A_k, b_k) , 通过训练得到 5 组模型参数。

18. 如权利要求 17 所述的方法, 其中步骤 (8) 具体为:

(8.1) 对于输入的人脸图像 t_n 进行检测和对齐处理, 利用人脸检测器对该人脸图像进行人脸检测, 基于人脸检测器检测到的人脸检测框对该人脸图像进行对齐处理, 对齐处理

具体为使用人脸检测器检测得到的 $L_e * L_e$ 像素大小的矩形人脸检测框,以该人脸检测框的中心为参考点,在该人脸图像上抠取 $1.5L_e * 1.5L_e$ 像素大小的矩形区域;

(8.2) 对于抠取的该 $1.5L_e * 1.5L_e$ 像素大小的矩形区域,将其归一化到 150*150 像素大小矩形区域;

(8.3) 将该对齐归一化后的人脸图像 I_e 上面部关键点的初始位置设为由人脸图像训练集得到的平均位置 I_0 ;

(8.4) 初始化 $k=1$, 计算 $\Delta_{i_e,k} = A_k * G(I_{k-1}) + b_k$, $I_k = I_{k-1} + \Delta_{i_e,k}$ 并使得 $k=k+1$, 其中对齐归一化后的人脸图像 I_e 在第 k 次迭代后所有面部关键点的位置记为 I_k , 判断 $k \leq 5$ 是否成立, 如果不成立, 进入步骤 (8.5), 如果成立, 重复计算, 其中每次计算使用一组 k 值对应的训练得到的模型参数;

(8.5) 输出 I_5 为该输入对齐归一化后的人脸图像 I_e 的面部关键点位置;

(8.6) 设输出的 $I_5 = (X_{e,1}, Y_{e,1}, \dots, X_{e,i}, Y_{e,i}, \dots, X_{e,28}, Y_{e,28})$, 则将

$$L_e = (x_{e,1}, y_{e,1}, \dots, x_{e,i}, y_{e,i}, \dots, x_{e,28}, y_{e,28}), (x_{e,i}, y_{e,i}) = \left(\frac{100X_{e,i}}{L} + w_{e,x}, \frac{100Y_{e,i}}{L} + w_{e,y} \right) \text{ 作为输}$$

入人脸图像 I_e 的面部关键点位置, 其中 $(w_{e,x}, w_{e,y})$ 为该人脸检测框在输入人脸图像 I_e 建立的直角坐标系上的坐标, 其中以该人脸图像左上角为原点, 以水平右向为 x 轴正方向, 以垂直下向为 y 轴正方向, 建立直角坐标系。

人脸图像面部关键点的定位方法

技术领域

[0001] 本发明属于计算机视觉和图像处理技术领域，特别的涉及人脸图像的面部关键点的定位方法。

背景技术

[0002] 计算机人脸识别将根据人脸来辨别未知人物身份的能力赋予了计算机系统，使得计算机系统基于已知的人脸样本库，利用计算机分析和模式识别技术从静态或者动态的场景中，识别或验证一个或多个个人脸，该技术已经被广泛的应用于公共安全，身份识别等场合。人脸识别系统包括有人脸检测、面部关键点定位、识别和跟踪等步骤，而面部关键点定位是解决这些问题的关键，其定位的精确程度极大地影响了人脸识别的准确性，错误的定位会导致提取出的人脸描述特征的严重变形。现有技术中，面部关键点的定位方法主要包括在对包含人脸图像的图片进行分解变换和对称变换后利用几何分布关系得到的估算定位，以及基于 ASM 和 AMM 模型对面部关键点进行的定位。然而，这些技术没有很好的解决在人脸光照条件、人脸姿态以及人脸表情多变且不确定的非理想情况下的定位精度下降问题。

[0003] 为克服现有技术的上述不足之处，本发明提出了一种基于面部关键点判别特征学习以及关键点回归定位模型学习的定位方法，利用 Fisher 准则学习了面部关键点的判别投影矩阵，从而得到每个关键点的判别特征表示，在此基础上，基于偏最小二乘回归学习了面部关键点的回归定位模型，进而实现对未知的人脸图像的面部关键点定位。本发明提出的基于两次模型学习的面部关键点定位方法，具有更好的鲁棒性，能够更加有效地应对人脸的局部变化，稳定而准确地实现对面脸图像的面部关键点定位。

发明内容

[0004] 本发明要解决的问题是提供一种人脸图像面部关键点的定位方法，以解决现有技术中面部关键点定位不够精确及稳定的缺陷。

[0005] 为了实现以上目的，本发明提供了一种人脸图像面部关键点的定位方法，其特征在于：

- (1) 设置人脸图像训练集；
- (2) 对人脸图像训练集中的人脸图像进行面部关键点位置标注；
- (3) 对人脸图像训练集中标注后的每张人脸图像进行检测和对齐处理；
- (4) 对人脸图像训练集中对齐后的每张人脸图像进行归一化处理；
- (5) 对面部关键点进行判别特征学习；
- (6) 为每张人脸图像进行面部关键点特征提取；
- (7) 训练得到面部关键点定位模型；
- (8) 对输入的人脸图像利用定位模型定位得到面部关键点位置。

[0006] 优选的，其中步骤 (1) 具体为：收集 N 张互联网人脸图像，形成人脸图像训练集

$$T = \{t_n, n=1, 2, \dots, N\}.$$

[0007] 优选的,其中步骤(2)具体为:为人脸图像训练集中的每张人脸图像进行面部关键点位置标注并得到关键点坐标,其中标注的面部关键点包括人脸器官和面部轮廓。

[0008] 优选的,其中的人脸器官包括眉毛中心、眼睛中心、鼻子中心和嘴巴中心,面部轮廓为下面部轮廓,共标记 28 个面部关键点。

[0009] 优选的,其中以每张人脸图像 t_n 的左上角为原点 O ,以水平右向为 x 轴正方向,以垂直下向为 y 轴正方向,建立直角坐标系,得到标注的每个面部关键点 i 在该坐标系中的位置 $(x_{n,i}, y_{n,i})$,其中。

[0010] 优选的,其中步骤(3)中的检测和对齐处理具体为:利用人脸检测器对每张人脸图像进行人脸检测,基于人脸检测器检测到的人脸检测框对训练集中的每张人脸图像进行对齐处理。

[0011] 优选的,其中基于人脸检测器检测到的人脸检测框对训练集中的每张人脸图像进行对齐处理具体为:人脸检测器在每张人脸图像上检测到 $L_n * L_n$ 像素大小的矩形人脸检测框,其中心在每张人脸图像 t_n 的直角坐标系中的坐标为 $(w_{n,x}, w_{n,y})$,每张人脸图像的人脸检测框的大小随其对应的人脸图像改变,以该人脸检测框的中心为参考点,在每张人脸图像 t_n 上抠取 $1.5L_n * 1.5L_n$ 像素大小的矩形区域。

[0012] 优选的,其中步骤(4)中的归一化处理具体为:对于抠取的 $1.5L_n * 1.5L_n$ 像素大小的矩形区域,将其归一化到 $150 * 150$ 像素大小矩形区域,以人脸检测框的中心为原点 O ,水平方向为 x 轴,垂直方向为 y 轴,则面部关键点 i 的坐标归一化为:

$$(X_{n,i}, Y_{n,i}) = \left(\frac{x_{n,i} - w_{n,x}}{100/L}, \frac{y_{n,i} - w_{n,y}}{100/L} \right),$$

对齐归一化后的每张人脸图像 t_n 的所有面部关键点的位置记为 I_n ,则 $I_n = [X_{n,1}, Y_{n,1}, \dots, X_{n,i}, Y_{n,i}, \dots, X_{n,28}, Y_{n,28}] \in R^{56}$ 。

[0013] 优选的,其中步骤(5)中的判别特征学习具体为:在对齐归一化后的人脸图像训练集上为每个面部关键点 i 单独地进行判别特征学习。

[0014] 优选的,其中在对齐归一化后的人脸图像训练集上为每个面部关键点 i 单独地进行判别特征学习具体为:

(5.1) 分别获得每个面部关键点 i 的正例样本和反例样本;

(5.2) 对每个面部关键点 i 得到的所有正例样本和所有反例样本分别提取 59 维的 Uniform LBP 特征;

(5.3) 计算每个面部关键点 i 得到的所有样本的均值 Γ_i 、正例样本均值 $\Gamma_{O,i}$ 和反例样本的均值 $\Gamma_{D,i}$;

(5.4) 计算每个面部关键点 i 正例样本和反例样本的散度矩阵 $\Sigma_{O,i}$ 和 $\Sigma_{D,i}$ 以及类间散度矩阵 Φ_i ;

(5.5) 基于 Fisher 准则学习每个面部关键点 i 的判别投影矩阵 V_i ;

优选的,其中步骤(5.1)具体为:在对齐归一化后的每幅人脸图像 t_n 上,为每个关键点 i 截取 $7*7$ 像素大小的窗体作为正例,在距离该关键点 i 水平或垂直方向距离为3到14个像素的范围内随机采样四个 $7*7$ 像素大小的窗体反例,为每张人脸图像的每个关键点 i 采样得到了一个正例样本和四个反例样本。

[0015] 优选的,其中步骤(5.2)具体为:关键点 i 的所有 N 个正例样本的Uniform LBP特征记为: $U_i = [u_{1,i}, u_{2,i}, \dots, u_{n,i}, \dots, u_{N,i}] \in R^{59 \times N}$,其中 $u_{n,i}$ 为人脸图像 t_n 上的关键点 i 的正例样本的Uniform LBP特征,所有反例样本的Uniform LBP特征记为: $D_i = [d_{1,i,1}, d_{1,i,2}, d_{1,i,3}, d_{1,i,4}, \dots, d_{n,i,m}, \dots, d_{N,i,4}] \in R^{59 \times 4N}$,其中 $d_{n,i,m}$ 为人脸图像 t_n 的关键点 i 的第 m 个反例样本的Uniform LBP特征, $m=1,2,3,4$ 。

[0016] 优选的,其中步骤(5.3)具体为:

$$\Gamma_i = \frac{\sum_{n=1}^N u_{n,i} + \sum_{n=1, m=1}^{N, m=4} d_{n,i,m}}{5N}, \quad \Gamma_{U,i} = \frac{\sum_{n=1}^N u_{n,i}}{N}, \quad \Gamma_{D,i} = \frac{\sum_{n=1, m=1}^{N, m=4} d_{n,i,m}}{4N}.$$

[0017] 优选的,其中步骤(5.4)具体为:

$$\Sigma_{U,i} = \sum_{n=1}^N (u_{n,i} - \Gamma_{U,i})(u_{n,i} - \Gamma_{U,i})^T, \quad \Sigma_{D,i} = \sum_{n=1, m=1}^{N, m=4} (d_{n,i,m} - \Gamma_{D,i})(d_{n,i,m} - \Gamma_{D,i})^T,$$

$$\Phi_i = N * (\Gamma_{U,i} - \Gamma_i)(\Gamma_{U,i} - \Gamma_i)^T + 4N * (\Gamma_{D,i} - \Gamma_i)(\Gamma_{D,i} - \Gamma_i)^T.$$

[0018] 优选的,其中步骤(5.5)具体为:基于Fisher准则,最大化判别投影 v_i 的可分性,其优化目标为 $\max_{v_i} \frac{\text{Trace}(v_i^T \Phi_i v_i)}{\text{Trace}(v_i^T (\Sigma_{U,i} + \Sigma_{D,i}) v_i)}$, s.t. $v_i^T v_i = I_{59}$,其中 v_i 的最优解为

$\bar{\Phi}_i = (\Sigma_{U,i} + \Sigma_{D,i})^{-1} \Phi_i$ 的最大特征值对应的特征向量,将上述过程重复10次,则每个关键点 i 得到10个投影向量,记为投影矩阵 $V_i: V_i = [v_{i,1}, v_{i,2}, \dots, v_{i,10}] \in R^{59 \times 10}$,为所有28个面部关键点重复以上学习过程,共得到28个投影矩阵。

[0019] 优选的,其中步骤(6)具体为:对于对齐归一化后的人脸图像 t_n ,分别截取以28个面部关键点为中心的 $7*7$ 像素大小的窗口,为每个窗口提取59维Uniform LBP特征,则在关键点 i 处提取的Uniform LBP特征记为: $LBP(X_{n,i}, Y_{n,i}) \in R^{59}$,关键点 i 的判别特征 $f_{n,i}$ 为: $f_{n,i} = V_i^T LBP(X_{n,i}, Y_{n,i}) \in R^{10}$,人脸图像 t_n 的28个面部关键点可表示为 $f_n = [f_{n,1}; f_{n,2}; \dots; f_{n,i}; \dots; f_{n,28}] \in R^{280}$,为一个280维的列向量,用函数 $G(l_n)$ 表示对对齐归一化后的人脸图像 t_n 在28个面部关键点坐标为 l_n 时所提取出的280维特征向量,则对齐归一化后的人脸图像 t_n 的面部关键点特征提取为:

$$G(l_n) = [f_{n,1}; f_{n,2}; \dots; f_{n,i}; \dots; f_{n,28}].$$

[0020] 优选的,其中步骤(7)具体为:

(7.1) 将对齐归一化后的训练集中所有人脸图像的28个面部关键点位置的平均值记

为 l_0 , 则

$$l_0 = \left[\frac{\sum_{n=1}^N X_{n,1}}{N}; \frac{\sum_{n=1}^N Y_{n,1}}{N}; \dots; \frac{\sum_{n=1}^N X_{n,i}}{N}; \frac{\sum_{n=1}^N Y_{n,i}}{N}; \dots; \frac{\sum_{n=1}^N X_{n,28}}{N}; \frac{\sum_{n=1}^N Y_{n,28}}{N} \right] \in R^{56}$$

, 将训练集中 M 幅对齐归一化后的人脸图像的平均位置记为 X_0 , 则 $X_0 = [l_0, l_0, \dots, l_0] \in R^{56 \times N}$, 将训练集中 M 幅对齐归一化后的人脸图像的人工标注的面部关键点标记位置记为 X_g , 则 $X_g = [l_1, l_2, \dots, l_n, \dots, l_N] \in R^{56 \times N}$;

(7.2) 利用迭代回归学习方式建立面部关键点的定位模型, 令 $\Delta_k = X_g - X_{k-1}, k \geq 1$, 其中 Δ_k 表示训练集中对齐归一化后的人脸图像的人工标注的面部关键点位置 X_g 与第 $k-1$ 次迭代结果 X_{k-1} 之间的差异, 令 $G(X_k)$ 表示在第 k 次迭代时, 训练集中所有对齐归一化后的人脸图像在面部关键点位置处提取出的特征向量, 则 $G(X_k) = \left[[G(l_{k,1})], \dots, [G(l_{k,n})], \dots, [G(l_{k,N})] \right]$, 其中 $l_{k,n}$ 表示训练集中的对齐归一化后的人脸图像 t_n 在第 k 次迭代时面部关键点的位置, $G(l_{k,n}) = [f_{k,n,1}; f_{k,n,2}; \dots; f_{k,n,i}; \dots; f_{k,n,28}] \in R^{280}$, 其中 $f_{k,n,i}$ 表示对齐归一化后的人脸图像 t_n 在第 k 次迭代时面部关键点 i 的判别特征;

(7.3) 利用迭代回归学习来学习 Δ_k 与 $G(X_{k-1})$ 之间的回归函数, 将面部关键点定位建模为: $\Delta_k = A_k * G(X_{k-1}) + [b_k; b_k; \dots; b_k]$, 该迭代回归学习的目标函数为: $\min_{A_k, b_k} \|\Delta_k\|_F$, 用偏最小二乘求解模型参数: $A_k \in R^{56 \times 280}$ 和 $b_k \in R^{56}$, 训练过程共迭代 M 次, 其中 $M = 5$, 通过 M 次迭代回归学习确定定位模型参数 A_k 和 $b_k, k=1, 2, 3, 4, 5$, 其中每次迭代得到一组对应的模型参数 (A_k, b_k) , 通过训练得到 5 组模型参数。

[0021] 优选的, 其中步骤(8)具体为:

(8.1) 对于输入的人脸图像进行检测和对齐处理, 利用人脸检测器对该人脸图像进行人脸检测, 基于人脸检测器检测到的人脸检测框对该人脸图像进行对齐处理, 对齐处理具体为使用人脸检测器检测得到的 $L_c * L_c$ 像素大小的矩形人脸检测框, 以该人脸检测框的中心为参考点, 在该人脸图像上抠取 $1.5L_c * 1.5L_c$ 像素大小的矩形区域;

(8.2) 对于抠取的该 $1.5L_c * 1.5L_c$ 像素大小的矩形区域, 将其归一化到 $150 * 150$ 像素大小矩形区域;

(8.3) 将该对齐归一化后的人脸图像 t_n 上面部关键点的初始位置设为由人脸图像训练集得到的平均位置 l_0 ;

(8.4) 初始化 $k=1$, 计算 $\Delta_{l_{i,k}} = A_k * G(l_{k-1}) + b_k, l_k = l_{k-1} + \Delta_{l_{i,k}}$, 其中对齐归

一化后的人脸图像 t_e 在第 k 次迭代后所有面部关键点的位置记为 L_k , 并使得 $k=k+1$, 判断 $k \leq 5$ 是否成立, 如果不成立, 进入步骤 (8.5), 如果成立, 重复计算, 其中每次计算使用一组 k 值对应的训练得到的模型参数;

(8.5) 输出 L_5 为该输入对齐归一化后的人脸图像 t_e 的面部关键点位置;

(8.6) 设输出的 $L_5 = (X_{e,1}, Y_{e,1}, \dots, X_{e,i}, Y_{e,i}, \dots, X_{e,28}, Y_{e,28})$, 则将

$$L_5 = (x_{e,1}, y_{e,1}, \dots, x_{e,i}, y_{e,i}, \dots, x_{e,28}, y_{e,28}), (x_{e,i}, y_{e,i}) = \left(\frac{100X_{e,i}}{L} + w_{e,x}, \frac{100Y_{e,i}}{L} + w_{e,y} \right)$$

作为输入人脸图像 t_e 的面部关键点位置, 其中 $(w_{e,x}, w_{e,y})$ 为该人脸检测框在输入人脸图像 t_e 建立的直角坐标系上的坐标, 其中以该人脸图像左上角为原点, 以水平右向为 x 轴正方向, 以垂直下向为 y 轴正方向, 建立直角坐标系。

附图说明

[0022] 图 1 为 人脸图像面部关键点标注示例图

图 2 为 人脸图像对齐归一化示意图

图 3 为 左眼中心正例和反例样本的采样方式示意图

图 4 为 人脸图像面部关键点的定位方法流程图。

具体实施方式

[0023] 为了使本发明的目的、技术方案和优点更加清楚, 下面将结合附图及具体实施例对本发明做出进一步地详细描述。

[0024] 步骤 1: 设置人脸图像训练集

收集一个互联网人脸图像训练集 $T = \{t_n, n=1, 2, \dots, N\}$, 在该训练集中具有 N 张人脸图像。

[0025] 步骤 2: 对训练集中的每张人脸图像进行面部关键点位置标注

为该训练集中的每张人脸图像进行面部关键点的位置标注。图 1 为面部关键点的标注示意图。标注的面部关键点包括人脸器官和面部轮廓, 其中人脸器官包括眉毛中心、眼睛中心、鼻子中心和嘴巴中心, 面部轮廓为下面部轮廓, 共标记 28 个面部关键点。将人脸图像 t_n 的左上角记为坐标原点 O , 以水平右向为 x 轴正方向, 以垂直下向为 y 轴正方向, 建立直角坐标系, 将第 i 个面部关键点在该人脸图像坐标系中的坐标记为 $(x_{n,i}, y_{n,i})$, 其中 $i=1, 2, \dots, 28$ 。

[0026] 步骤 3: 对训练集中标注后的每张人脸图像进行检测和对齐处理。

[0027] 利用人脸检测器对人脸图像进行人脸检测, 基于人脸检测器检测到的人脸检测框的结果对训练集中的每张人脸图像进行对齐。设人脸检测器在每张人脸图像 t_n 上检测

到的人脸检测框为 $L_n * L_n$ 像素大小的矩形,其中心在人脸图像 t_n 的直角坐标系中的坐标为 $(w_{n,x}, w_{n,y})$,其中每张人脸图像的人脸检测框大小随着其对应的人脸图像而改变,对齐方式为:以该人脸检测框的中心为参考点,在人脸图像 t_n 上抠取 $1.5L_n * 1.5L_n$ 像素大小的矩形区域,如图2所示。

[0028] 步骤4:对对齐后的每张人脸图像进行归一化处理

将该抠取出的图像进行归一化处理,得到 $150*150$ 像素大小的矩形区域。以人脸检测框的中心为原点 O ,水平方向为 X 轴,垂直方向为 Y 轴,则经过归一化处理后的面部关键点 i 的坐标归一化为:

$$(X_{n,i}, Y_{n,i}) = \left(\frac{X_{n,i} - w_{n,x}}{100/L}, \frac{Y_{n,i} - w_{n,y}}{100/L} \right) \quad \text{公式(1)}$$

将对齐归一化后的人脸图像 t_n 的所有面部关键点的位置记为 I_n :

$$I_n = [X_{n,1}, Y_{n,1}, \dots, X_{n,i}, Y_{n,i}, \dots, X_{n,28}, Y_{n,28}] \in R^{56} \quad \text{公式(2)}$$

步骤5:面部关键点判别特征学习

在对齐归一化后的人脸图像训练集上为每个面部关键点单独地学习判别特征表示,包括如下步骤:

(1) 分别获得每个面部关键点 i 的正例样本和反例样本。图3为左眼中心正例和反例样本的采样方式示意图,其中的红色矩形框表示正例样本,绿色矩形框表示随机选择的反例样本。以左眼中心为例:在对齐归一化后的每幅人脸图像 t_n 上,以左眼中心为中心,截取 $7*7$ 像素大小的窗体作为正例,在距离左眼中心水平或垂直方向距离为3到14个像素的范围内随机采样四个 $7*7$ 像素大小的窗体反例,即对左眼中心这一面部关键点来说,采样得到了一个正例样本和四个反例样本。对训练集中的所有 N 张人脸图像的左眼中心都进行上述正例和反例样本的提取。对训练集中的所有 N 张人脸图像的所有关键点 i 都进行上述正例和反例样本的提取。

[0029] (2) 对每个面部关键点 i 得到的 N 个正例和 $4N$ 个反例样本分别提取59维的Uniform LBP特征。以左眼中心为例:

将左眼中心的所有正例样本的Uniform LBP特征记为:

$$U_i = [u_{1,i}, u_{2,i}, \dots, u_{n,i}, \dots, u_{N,i}] \in R^{59 \times N} \quad \text{公式(3), 其中 } u_{n,i} \text{ 为人脸图像 } t_n$$

上的关键点 i 的正例样本的Uniform LBP特征。

[0030] 所有反例样本的Uniform LBP特征记为:

$D_i = [d_{1,i,1}, d_{1,i,2}, d_{1,i,3}, d_{1,i,4}, \dots, d_{n,i,m}, \dots, d_{N,i,4}] \in R^{39 \times 4N}$ 公式(4), 其中 $d_{n,i,m}$ 为人脸图像 t_n 的关键点 i 的第 m 个反例样本的Uniform LBP特征, $m=1,2,3,4$ 。

[0031] (3) 计算每个面部关键点 i 所有样本的均值 Γ_i 、正例样本均值 $\Gamma_{U,i}$ 和反例样本的均值 $\Gamma_{D,i}$:

$$\Gamma_i = \frac{\sum_{n=1}^N u_{n,i} + \sum_{n=1}^N \sum_{m=1}^4 d_{n,i,m}}{5N}, \quad \Gamma_{U,i} = \frac{\sum_{n=1}^N u_{n,i}}{N}, \quad \Gamma_{D,i} = \frac{\sum_{n=1}^N \sum_{m=1}^4 d_{n,i,m}}{4N} \quad \text{公式(5)}$$

(4) 计算每个面部关键点 i 正例和反例样本的散度矩阵 $\Sigma_{U,i}$ 和 $\Sigma_{D,i}$ 以及类间散度矩阵 Φ_i :

$$\Sigma_{U,i} = \sum_{n=1}^N (u_{n,i} - \Gamma_{U,i})(u_{n,i} - \Gamma_{U,i})^T, \Sigma_{D,i} = \sum_{n=1, n \neq i}^{N, n \neq i} (d_{n,i} - \Gamma_{D,i})(d_{n,i} - \Gamma_{D,i})^T$$

$$\Phi_i = N^*(\Gamma_{U,i} - \Gamma_i)(\Gamma_{U,i} - \Gamma_i)^T + 4N^*(\Gamma_{D,i} - \Gamma_i)(\Gamma_{D,i} - \Gamma_i)^T \quad \text{公式 (7)}$$

(5) 基于 Fisher 准则学习每个关键点 i 的判别投影矩阵 V_i :

学习过程基于 Fisher 准则, 最大化判别投影 v_i 的可分性, 其优化目标如下:

$$\max_{v_i} \frac{\text{Trace}(v_i^T \Phi_i v_i)}{\text{Trace}(v_i^T (\Sigma_{U,i} + \Sigma_{D,i}) v_i)}, \quad \text{s.t. } v_i^T v_i = I_{59} \quad (8)$$

其中的最优解为 $\Phi_i = (\Sigma_{U,i} + \Sigma_{D,i})^{-1} \Phi_i$ 的最大特征值对应的特征向量, 将上述过程重复 10 次, 则每个关键点得到 10 个投影向量, 记为投影矩阵 V_i :

$$V_i = [v_{i,1}, v_{i,2}, \dots, v_{i,10}] \in R^{59 \times 10} \quad (9)$$

为所有 28 个面部关键点重复以上学习过程, 共得到 28 个投影矩阵。

[0032] 步骤 6: 为每张人脸图像进行面部关键点特征提取

对于对齐归一化后的人脸图像 I_n , 分别截取以当前 28 个面部关键点为中心的 7×7 像素大小窗口, 为其提取 59 维 Uniform LBP 特征, 将第 i 个关键点处提取的 Uniform LBP 特征记为:

$$\text{LBP}(X_{n,i}, Y_{n,i}) \in R^{59}, i = 1, 2, \dots, 28 \quad \text{公式(10)}, \text{则第 } i \text{ 个关键点判别特征 } f_{n,i}$$

计算方式如下:

$$f_{n,i} = V_i^T \text{LBP}(X_{n,i}, Y_{n,i}) \in R^{10} \quad \text{公式(11)}$$

所有 28 个面部关键点可表示为 $f_n = [f_{n,1}; f_{n,2}; \dots; f_{n,i}; \dots; f_{n,28}] \in R^{280}$, 为一个 280 维的列向量。记函数 $G(l_n)$ 表示对对齐归一化后的人脸图像在 28 个面部关键点坐标为 l_n 时提取出的 280 维特征向量, 其中 l_n 的表示见公式(2), 则

$$G(l_n) = [f_{n,1}; f_{n,2}; \dots; f_{n,i}; \dots; f_{n,28}]$$

[0033] 步骤 7: 训练面部关键点定位模型, 得到模型参数

(1) 将训练集中对齐归一化后的所有人脸图像的 28 个面部关键点位置的平均值记为 l_0 , 计算方法为:

$$l_0 = \left[\frac{\sum_{n=1}^N X_{n,1}}{N}, \frac{\sum_{n=1}^N Y_{n,1}}{N}, \dots, \frac{\sum_{n=1}^N X_{n,i}}{N}, \frac{\sum_{n=1}^N Y_{n,i}}{N}, \dots, \frac{\sum_{n=1}^N X_{n,28}}{N}, \frac{\sum_{n=1}^N Y_{n,28}}{N} \right] \in R^{56}$$

(12)

将训练集中 M 幅人脸图像对齐归一化后的平均位置记为 X_0 :

$$X_0 = [l_0, l_0, \dots, l_0] \in R^{56 \times N} \quad (13)$$

将训练集中 M 幅人脸图像对齐归一化后的人工标注的面部关键点标记位置记为 X_g :

$$X_g = [l_1, l_2, \dots, l_n, \dots, l_N] \in R^{56 \times N} \quad (14)$$

(2) 利用迭代回归学习方式建立面部关键点的定位模型, 记

$$\Delta_k = X_g - X_{k-1}, k \geq 1 \quad (15)$$

Δ_k 表示人工标记的面部关键点位置 X_g 与第 $k-1$ 次迭代位置结果之间的差异。记 $G(X_k)$ 表示在第 k 次迭代时, 训练集中所有人脸图像对齐归一化后的关键点 X_k 处提取出的特征向量:

$$G(X_k) = [[G(l_{k,1})], \dots, [G(l_{k,n})], \dots, [G(l_{k,N})]] \quad (16)$$

$$G(l_{k,n}) = [f_{k,n,1}, f_{k,n,2}, \dots, f_{k,n,i}, \dots, f_{k,n,28}] \in R^{280} \quad (17)$$

其中 $l_{k,n}$ 表示训练集中第 n 幅对齐归一化后的人脸在第 k 次迭代时面部关键点的位置。

[0034] (3) 利用迭代回归学习来建模 Δ_k 与 $G(X_{k-1})$ 之间的回归函数, 即定位模型为:

$$\Delta_k = A_k * G(X_{k-1}) + [b_k; b_k; \dots; b_k] \quad (18)$$

该迭代回归学习的目标函数如下: $\min_{A_k, b_k} \|\Delta_k\|_F$

公式(18)可以用偏最小二乘求解 $A_k \in R^{56 \times 280}$ 和 $b_k \in R^{56}$, 训练过程共迭代 M 次, 优选的, 设置 $M = 5$ 。通过 M 次迭代回归学习确定定位模型参数 A_k 和 b_k , 其中每次迭代得到一组对应的模型参数 (A_k, b_k) , 通过训练得到 5 组模型参数。

[0035] 步骤 8 :对输入人脸图像利用定位模型定位得到面部关键点位置。

[0036] 图 4 给出了面部关键点的定位方法流程图。

[0037] (1) 对于作为测试样本的人脸图像 t_e , 经过和训练集人脸图像相同的人脸检测、对齐和 150×150 矩形大小的归一化处理;

(2) 将该对齐归一化后的人脸图像 t_e 上面部关键点的初始位置设为由公式(13)得到的平均位置 l_0 ;

(3) 初始化 $k=1$, 计算 $\Delta_{t_e, k} = A_k * G(l_{k-1}) + b_k$, $l_k = l_{k-1} + \Delta_{t_e, k}$ 并使得 $k=k+1$, 其中对齐归一化后的人脸图像 t_e 在第 k 次迭代后所有面部关键点的位置记为 l_k , 判断 $k \leq 5$ 是否成立, 如果不成立, 输出 l_5 为该测试样本人脸图像面部关键点的最终位置, 如果成立, 重复计算, 每次计算使用一组 K 值对应的训练得到的模型参数。由此, 通过 5 次循环计算, 利用定位模型参数得到了输入人脸图像对齐归一化后的面部关键点的最终位置;

(4) 设输出的 $l_5 = (X_{e,1}, Y_{e,1}, \dots, X_{e,i}, Y_{e,i}, \dots, X_{e,28}, Y_{e,28})$, 则将

$$l_5 = (x_{e,1}, y_{e,1}, \dots, x_{e,i}, y_{e,i}, \dots, x_{e,28}, y_{e,28}), (x_{e,i}, y_{e,i}) = \left(\frac{100X_{e,i}}{L} + w_{e,x}, \frac{100Y_{e,i}}{L} + w_{e,y} \right)$$

作为输入人脸图像 I_i 的面部关键点位置。其中 $(w_{e,x}, w_{e,y})$ 为该人脸检测框在输入人脸图像 I_i 建立的直角坐标系上的坐标。

[0038] 显然,本领域的技术人员可以对本发明进行各种改动和变形而不脱离本发明的精神和范围,这样,倘若本发明的这些修改和变形属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变形在内。

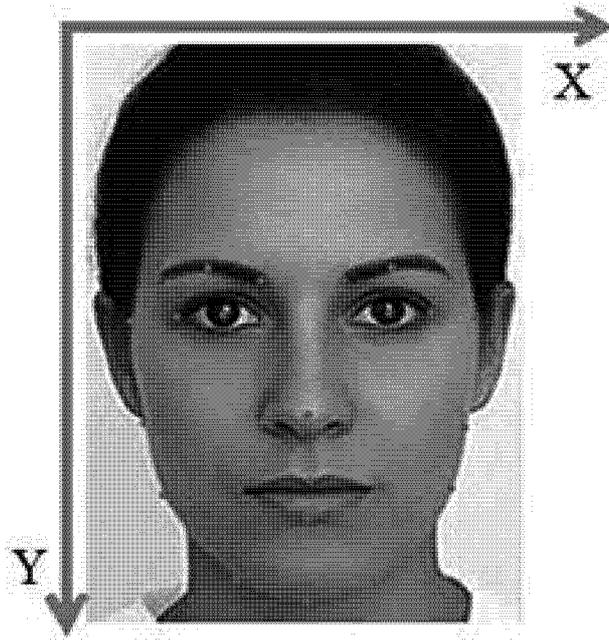


图 1

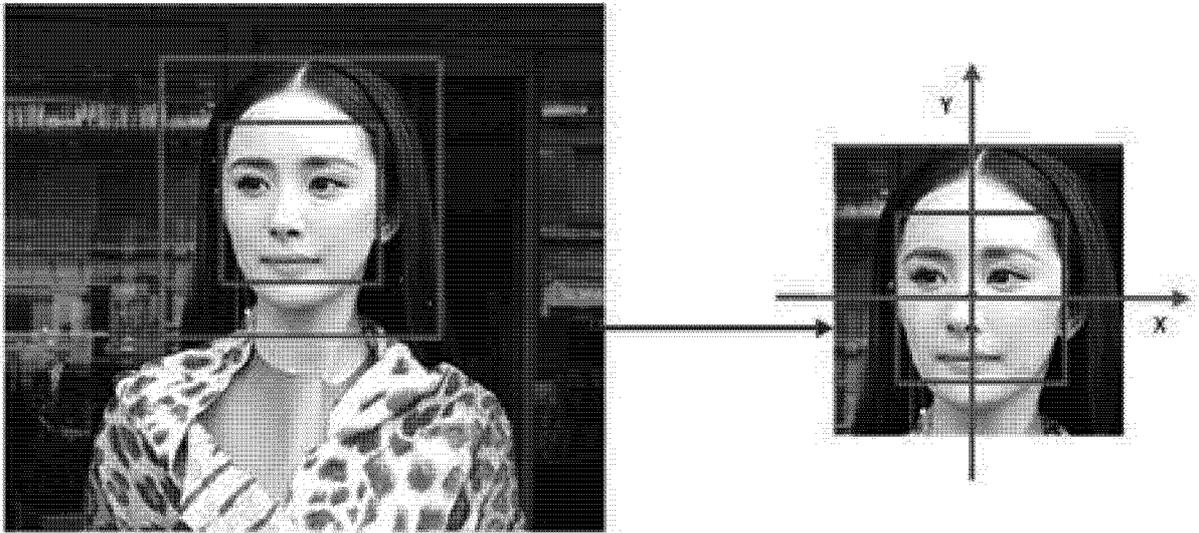


图 2

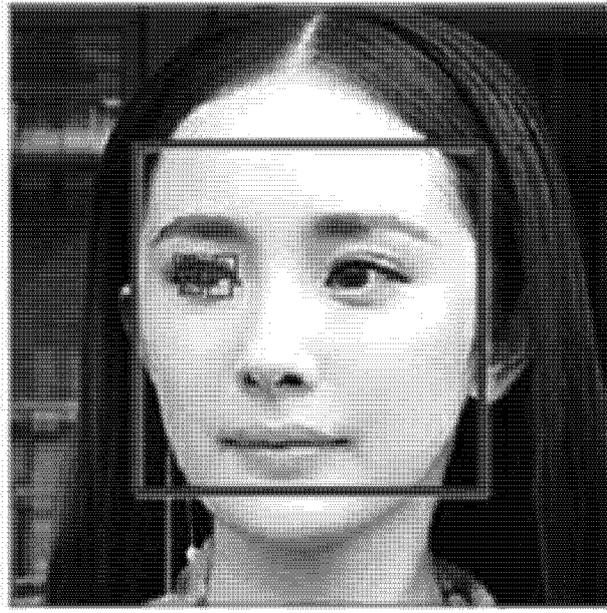


图 3

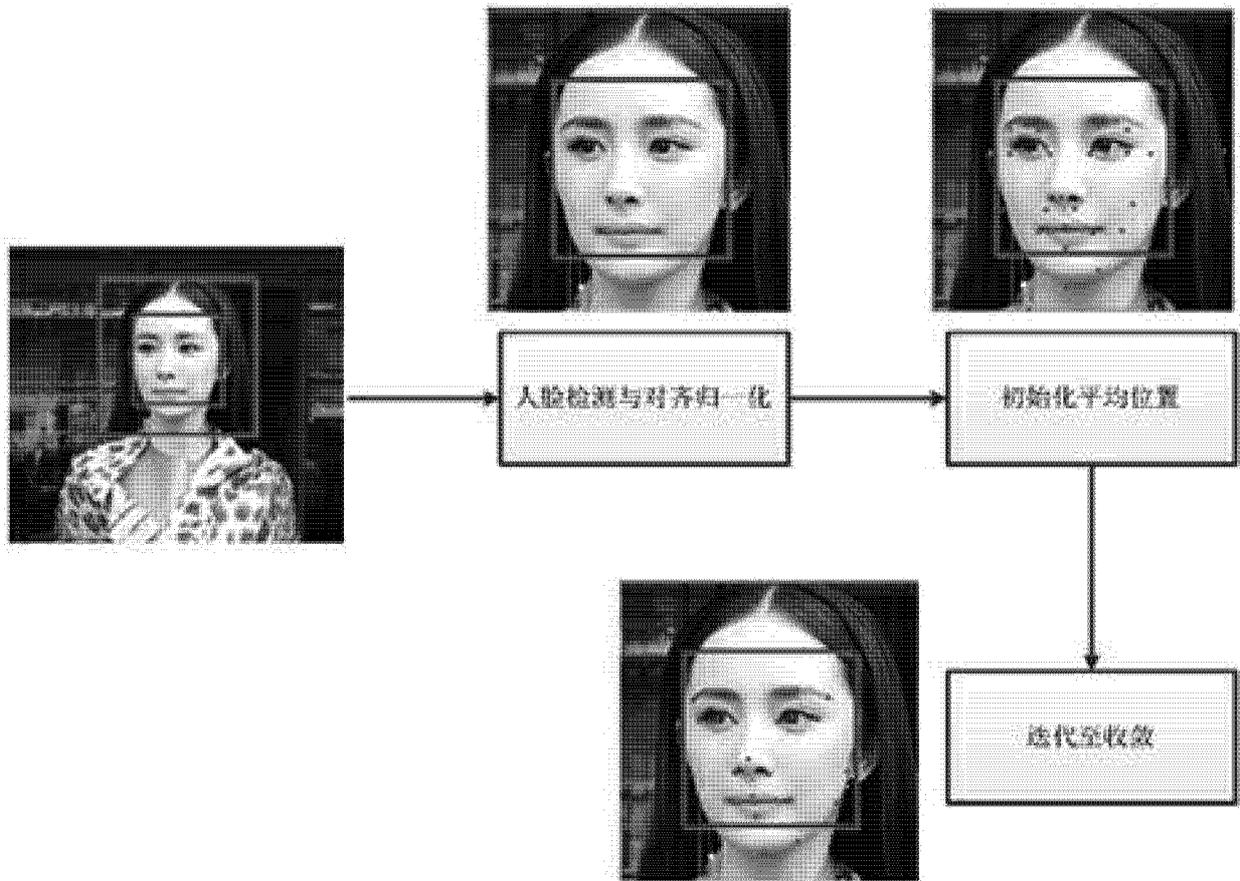


图 4