

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G06K 9/00 (2006.01)



# [12] 发明专利说明书

专利号 ZL 200610117050.3

[45] 授权公告日 2008 年 9 月 3 日

[11] 授权公告号 CN 100416596C

[22] 申请日 2006.10.12

[21] 申请号 200610117050.3

[73] 专利权人 上海交通大学

地址 200240 上海市闵行区东川路 800 号

[72] 发明人 杜春华 杨杰 张田昊 陈鲁

王华华 吴证 袁泉

[56] 参考文献

CN1731416A 2006.2.8

CN1786980A 2006.6.14

US20030206171A1 2003.11.6

CN1797420A 2006.7.5

审查员 王涛

[74] 专利代理机构 上海交达专利事务所

代理人 王锡麟 张宗明

权利要求书 3 页 说明书 5 页 附图 2 页

[54] 发明名称

用贝叶斯网络分类器图像判别特征点位置的方法

[57] 摘要

一种人脸识别领域的用贝叶斯网络分类器搜索特征点位置的方法，包括如下步骤：(1) 建立 ASM 模型；(2) 通过人脸检测和眼睛定位初始化 ASM 搜索的初始位置；(3) 为人脸上的每个特征点生成对应的样本；(4) 对于每一个特征点，用其对应的样本获得一个贝叶斯网络分类器；(4) 用 ASM 搜索的初始位置作为起始位置，并使用贝叶斯网络分类器进行特征点定位。本发明提出的涉及了眼睛检测、分类器训练、ASM 特征点定位的人脸特征点定位方法可以进一步应用于人脸识别、性别识别、表情识别、年龄估计等方面，其具非常高的精度。



1. 一种用贝叶斯网络分类器搜索特征点位置的方法，其特征在于，包括如下步骤：

- (1) 建立 ASM 模型；
- (2) 通过人脸检测和眼睛定位初始化 ASM 搜索的初始位置；
- (3) 为人脸上的每个特征点生成对应的样本；
- (4) 对于每一个特征点，用其对应的样本获得一个贝叶斯网络分类器；
- (5) 用 ASM 搜索的初始位置作为起始位置，并使用贝叶斯网络分类器进行特征点定位。

2. 根据权利要求 1 所述的用贝叶斯网络分类器搜索特征点位置的方法，其特征是，所述的步骤(1)，是指：首先在训练集的每一个训练样本图像上手工标定  $k$  个人脸的特征点，这  $k$  个特征点组成的形状由一个向量  $x(i)=[x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_k]$  来表示，具有相同编号的特征点在各图像中代表了相同的特征， $n$  个训练样本图像就对应  $n$  个形状向量，校准这  $n$  个向量从而使得它们所表示的形状在大小、方向和位置上最为接近，然后对  $n$  个校准后的形状向量进行 PCA 处理，最终任何一个形状都表示为  $x = \bar{x} + Pb$ ，其中  $b = P^T \cdot (x - \bar{x})$ ， $b$  代表了前  $t$  个最大的模式的变化情况，这样就训练好了 ASM 模型；

3. 根据权利要求 1 所述的用贝叶斯网络分类器搜索特征点位置的方法，其特征是，所述的步骤(2)，是指：用 ababoost 方法在图像上找到人脸区域，然后在人脸图像上用模板匹配的方法找到两眼位置，并设定两眼连线的中点坐标为  $[a, b]$ ，对上述求得的平均形状模型  $\bar{x}$ ，分别计算左右眼球周围的四个特征点的中心作为左右眼睛位置，从而得到两眼连线的中点坐标  $[c, d]$ ，然后把整个平均形状模型  $\bar{x}$  平移  $[a-c, b-d]$ ，这样就得到了 ASM 搜索的初始位置。

4. 根据权利要求 1 所述的用贝叶斯网络分类器搜索特征点位置的方法, 其特征是, 所述的步骤(3), 是指: 对于人脸上的每个特征点, 在垂直于该特征点的前后两个特征点连线方向上两边各选择 1 个像素, 计算这  $2 \times 1 + 1$  个像素的灰度值导数并归一化, 这样就得到了一个一维向量, 然后从前往后在该一维向量中依次选择  $2 \times m$  ( $0 < m < 1$ ,  $m$  取整数值) 个元素组成一个一维子向量, 总共有  $2 \times (1 - m) + 1$  个这样的子向量, 并依次为这些一维子向量标以相应的类别号,  $1, 2, \dots, 2 \times (1 - m) + 1$ , 如此, 一幅人脸图像上的一个特征点就生成了  $2 \times (1 - m) + 1$  个不同类的训练样本,  $n$  个训练图像, 每个特征点就对应  $n \times (2 \times (1 - m) + 1)$  个训练样本。

5. 根据权利要求 1 所述的用贝叶斯网络分类器搜索特征点位置的方法, 其特征是, 所述的步骤(4), 是指: 对于每个特征点, 用其对应的  $n \times (2 \times (1 - m) + 1)$  个训练样本训练其对应的贝叶斯网络分类器, 这样就得到  $k$  个贝叶斯网络分类器。

6. 根据权利要求 1 所述的用贝叶斯网络分类器搜索特征点位置的方法, 其特征是, 所述的步骤(5), 是指: 以模型的初始位置作为起始位置在新的图像中搜索目标形状, 这个搜索过程主要是通过仿射变换和参数  $b$  的变化来实现。

7. 根据权利要求 6 所述的用贝叶斯网络分类器搜索特征点位置的方法, 其特征是, 所述的步骤 (5), 具体通过反复如下两步来实现:

1) 计算每个特征点的新位置

首先把初始模型覆盖在图像上, 对于模型中第  $j$  个特征点, 在垂直于其前后两个特征点连线方向上以其为中心两边各选择 1 个像素, 计算这  $2 \times 1 + 1$  个像素的灰度值导数并归一化, 这样就得到了一个一维向量, 然后从前往后在该一维向量中依次选择  $2 \times m$  个元素组成一个一维子向量, 总共有  $2 \times (1 - m) + 1$  个这样的子向量, 其中  $m < 1$ , 把这  $2 \times (1 - m) + 1$  个子向量送入第  $k$  个分类器进行分类, 并把类别号被判别为  $l + 1$  的子向量的中心设定为第  $j$  个特征点的新位置, 同时计算该特征点位置的变化  $dX_j$ , 对每个特征点都进行这样的计算就得到  $k$  个位置变化  $dX_i, i = 1, 2, \dots, k$ , 并组成一个向量  $dX = (dX_1, dX_2, \dots, dX_k)$ ;

## 2) 仿射变换中的参数和 b 的更新

对参数作如下更新： $X_c = X_c + w_t dX_c$ ， $Y_c = Y_c + w_t dY_c$ ， $\theta = \theta + w_\theta d\theta$ ，

$b = b + W_b db$ ，式中  $w_t, w_\theta, w_s, W_b$  是用于控制参数变化的权值，这样由式  $x = \bar{x} + Pb$  得到新的形状。

## 用贝叶斯网络分类器图像判别特征点位置的方法

### 技术领域

本发明涉及的是一种人脸识别领域中的人脸特征点定位方法，具体是一种用贝叶斯网络分类器图像判别特征点位置的方法。

### 背景技术

人脸识别技术是众多生物特征识别核心中最为实用的技术，其包括表情识别、性别识别、年龄估计、姿态估计等，而人脸特征点定位是这些研究领域中的核心技术，最终人脸识别的精度在很大程度上取决于人脸特征点定位的精确性。因而，精确地定位大量的人脸特征点可以极大地提高人脸识别的精度。目前，最实用的人脸特征定位方法是全局特征点定位方法。在这类方法中，ASM（活动形状模型）方法可以同时定位很多人脸特征点，且其速度快，对光照、背景的变化具有较强的鲁棒性，因而被广泛地应用于特征点定位。

经对现有技术文献的检索发现，Cootes, T. F. 等在《Computer Vision and Image Understanding》（计算机视觉与图像理解）（1995年第一期第38页）发表的（“Active shape models—their training and application”）（活动形状模型—其训练及应用），该文中提出了活动形状模型方法，方法中，当搜索特征点新位置时，在垂直于前后两个特征点连线的方向上的一维轮廓上找到使马氏距离最小的子轮廓的中心并设定该中心位置为当前特征点的新位置，其不足点在于：这种操作是在假设特征点周围像素的灰度值服从正态分布的前提下完成的，事实上，因为图像判别样本的数量不可能非常多以使得特征点的轮廓完全服从正态分布，而且，特征点周边像素的灰度值也不一定完全服从正态分布，特别是在背景比较复杂，光照不均的情况下，特征点周围的像素就完全不服从正态分布，这样就严重地影响了特征点定位的精度。同时这种方法效果并不理想。

## 发明内容

本发明针对 ASM 方法中图像判别不精确性等缺陷，提出了一种用贝叶斯网络分类器图像判别特征点位置的方法，使其把图像判别特征点新的位置问题转换为基于机器图像判别问题来解决，从而可以提高判别特征点的精度。

本发明是通过以下技术方案实现的，包括如下步骤：

- (1) 建立 ASM 模型；
- (2) 通过人脸检测和眼睛定位初始化 ASM 搜索的初始位置；
- (3) 为人脸上的每个特征点生成对应的样本；
- (4) 对于每一个特征点，用其对应的样本获得一个贝叶斯网络分类器；
- (5) 用 ASM 搜索的初始位置作为起始位置，并使用贝叶斯网络分类器进行特征点定位。

所述的步骤(1)，是指：首先在训练集的每一个训练样本图像上手工标定  $k$  个人脸的特征点，这  $k$  个特征点组成的形状可以由一个向量  $x(i)=[x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_k]$  来表示，具有相同编号的特征点在不同的图像中代表了相同的特征， $n$  个训练样本图像就对应  $n$  个形状向量，然后校准这  $n$  个向量从而使得它们所表示的形状在大小、方向和位置上最为接近。然后对  $n$  个校准后的形状向量进行 PCA（主元分析）处理，最终任何一个形状都可以表示为  $x = \bar{x} + Pb$ ，其中  $b = P^T \cdot (x - \bar{x})$ ， $b$  代表了前  $t$  个最大的模式的变化情况，这样就训练好了 ASM 模型。

所述的步骤(2)，是指：用 ababoost 方法在图像上找到人脸区域，然后在人脸图像上用模板匹配的方法找到两眼位置，并设定两眼中点坐标为  $[a, b]$ ，对上述求得平均形状模型  $\bar{x}$ ，分别计算左右眼球周围的四个特征点的中心作为左右眼睛位置，从而得到两眼的中点坐标  $[c, d]$ ，然后把整个平均形状模型  $\bar{x}$  平移  $[a-c, b-d]$ ，这样就得到了 ASM 搜索的初始位置。

所述的步骤(3)，是指：对于人脸上的每个特征点，在垂直于该特征点的前后两个特征点连线方向上两边各选择 1 个像素，计算这  $2 \times 1 + 1$  个像素的灰度值导数

并归一化，这样就得到了一个一维向量，然后从前往后在该一维向量中依次选择  $2 \times m (m < 1)$  个元素组成一个一维子向量，总共有  $2 \times (1-m) + 1$  个这样的子向量，并依次为这些一维子向量标以相应的类别号， $1, 2, \dots, 2 \times (1-m) + 1$ 。如此，一幅人脸图像上的一个特征点就生成了  $2 \times (1-m) + 1$  个不同类的训练样本， $n$  个训练图像，每个特征点就对应  $n \times (2 \times (1-m) + 1)$  个训练样本。

所述的步骤(4)，是指：对于每个特征点，用其对应的  $n \times (2 \times (1-m) + 1)$  个训练样本训练其对应的贝叶斯网络分类器，这样就可以得到  $k$  个贝叶斯网络分类器。

所述的步骤(5)，是指：用步骤(2)得到的初始位置在新的图像中搜索目标形状，这个搜索过程主要是通过仿射变换和参数  $b$  的变化来实现。具体通过反复迭代如下两步来实现：

#### 1) 搜索特征点的新位置

首先用步骤(2)得到的初始位置作为起始位置，对于模型中第  $j$  个特征点，在垂直于其前后两个特征点连线方向上以其为中心两边各选择 1 个像素，计算这  $2 \times 1 + 1$  个像素的灰度值导数并归一化，这样就得到了一个一维向量，然后从前往后在该一维向量中依次选择  $2 \times m (m < 1)$  个元素组成一个一维子向量，总共有  $2 \times (1-m) + 1$  个这样的子向量，把这  $2 \times (1-m) + 1$  个子向量送入第  $j$  个贝叶斯网络分类器进行分类，并把类别号被判别为  $m+1$  的子向量的中心设定为第  $j$  个特征点的新位置，同时计算该特征点位置的变化  $dX_j$ ，对每个特征点都进行这样的计算就得到  $k$  个位置变化  $dX_i, i = 1, 2, \dots, k$ ，并组成一个向量  $dX = (dX_1, dX_2, \dots, dX_k)$ 。

#### 2) 仿射变换中的参数和 $b$ 的更新

由公式  $X = M(s, \theta)[x] + X_c$  得： $M(s(1+ds), (\theta + d\theta))[x + dx] + (X_c + dX_c) = (X + dX)$ ，  
 $M(s(1+ds), (\theta + d\theta))[x + dx] = M(s, \theta)[x] + dX + X_c - (X_c + dX_c)$ ，同时由式  
 $x = \bar{x} + Pb$ ，期望找到  $db$  使得  $x + dx = \bar{x} + P(b + db)$ ，由式  $x = \bar{x} + Pb$  可得  $db = P^{-1}dx$ ，  
 这样就可以对参数作如下更新： $X_c = X_c + w_i dX_c$ ， $Y_c = Y_c + w_i dY_c$ ， $\theta = \theta + w_\theta d\theta$ ，

$b = b + W_b \Delta b$ , 式中  $w_1, w_\theta, w_s, W_b$  用于控制参数变化的权值, 这样就可以由式  $x = \bar{x} + Pb$  得到新的形状。

本发明提出的人脸特征点定位方法具有非常高的精度。用拍摄的人脸库对比本发明提出的用贝叶斯网络分类器进行的特征点定位方法和原始 ASM 方法中用一维轮廓进行特征点定位的方法, 前者的特征点定位的平均误差分别是 2.8 个像素, 而后者特征点定位的平均误差分别是 4.5 个像素, 实验表明本发明提出的用贝叶斯网络分类器定位特征点的方法在精度上比其他人脸特征点定位方法有很大的改善。

#### 附图说明

图 1 为标有特征点的人脸图像。

图 2 为眼睛定位的结果。

图 3 为用初始模型并利用贝叶斯网络分类器进行 ASM 搜索得到的结果。

#### 具体实施方式

以下结合一个具体的实施例对本发明的技术方案作进一步详细的描述。

实施例采用的图像来自上海交通大学拍摄的人脸图像库。整个实现过程如下:

1. 从人脸库中选择 600 个标好特征点的人脸图像建立 ASM 模型。标好特征点的人脸图像, 如图 1 所示。即首先在训练集的每一个训练样本图像上选择 60 个特征点, 这 60 个特征点组成的形状可以由一个向量  $x(i) = [x_1, x_2, \dots, x_{60}, y_1, y_2, \dots, y_{60}]$  来表示, 具有相同编号的特征点在不同的图像中表示相同的特征, 600 个训练样本图像就有 400 个形状向量, 然后对这 400 个向量进行校准操作, 使这些形状向量所表示的形状在大小、方向和位置上最为接近。然后对 400 个校准后的形状向量进行 PCA (主元分析) 处理, 这样任何一个形状都可以表示为  $x = \bar{x} + Pb$ , 式中  $b = P^T \cdot (x - \bar{x})$ ,  $b$  的值表示了前 26 个模式的变化情况, 这样就训练好了 ASM 模型。

2. 用 adaboost 方法在图像中找到人脸区域并用模板匹配方法找到两眼的位置, 如图 2 所示, 并得到两眼中心点坐标为 (234, 251), 然后求得 ASM 初始模型上两眼中心点坐标 (113, 145), 然后把该模型在 X 方向上平移 121 (234-113) 个像素, 在 Y 方



向上平移 106 (251-145) 个像素, 这样就得到了 ASM 模型的初始位置。对于人脸上的 60 个特征点, 在垂直于该特征点的前后两特征点连线方向上两边各选择 16 个像素, 计算这 33 (2\*16+1) 个像素的灰度值导数并归一化, 这样就得到了一个一维向量, 然后从前往后在该一维向量中依次选择 16 个元素组成一个一维子向量, 总共有 17 个这样的子向量, 并依次为这些一维子向量标以相应的类别号, 1, 2, ..., 17。这样一幅人脸图像上的一个特征点就生成了 17 个不同类的训练样本, 600 个训练图像, 每个特征点就对应 10200 (600\*17) 个训练样本。

3. 对于每个特征点, 用其对应的 10200 个训练样本训练贝叶斯网络分类器, 这样就可以得到 60 个分类器。

4. 以模型的初始位置作为起始位置并用贝叶斯网络分类器进行 ASM 搜索得到人脸特征点的位置, 即把初始模型覆盖在图像上, 对模型中第  $j$  个特征点, 在垂直于其前后两个特征点连线方向上以其为中心两边各选择 16 个像素, 这样就形成了一个长度为 33 (2\*16+1) 的一维向量, 计算这 33 个像素的灰度值导数并归一化, 这样就得到了一个一维向量, 然后从前往后在该一维向量中依次选择 16 个元素组成一个一维子向量, 总共有 17 个这样的子向量, 把这 17 个子向量送入第  $j$  个分类器进行分类, 并把类别号被判别为 9 的子向量的中心设定为第  $j$  个特征点的新位置, 同时计算该特征点位置的变化  $dX_j$ , 对每个特征点都进行这样的计算就得到 60 个位置变化  $dX_i, i=1, 2, \dots, 60$ , 并组成一个向量  $dX = (dX_1, dX_2, \dots, dX_{60})$ 。

5. 然后再根据上述  $dX$  计算仿射变换参数和  $b$  的变化, 经过 19 步迭代就可以最终定位 60 个特征点, 如图 3 所示。

从以上实施例可以看出, 本发明提出的涉及了人脸检测、眼睛检测、贝叶斯网络分类器训练、ASM 特征点定位的人脸特征点定位方法可以进一步应用于人脸识别、性别识别、表情识别、年龄估计等方面, 其具非常高的精度。



图 1



图 2

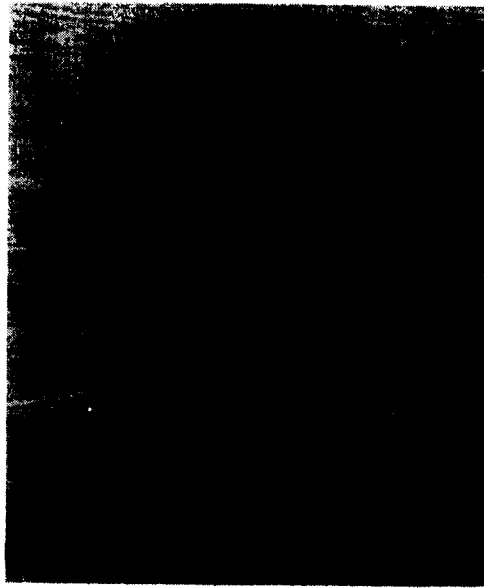


图 3