

基于 CPR 和 CLM 的多视角人脸特征点定位方法

石正权,赵启军,陈 虎

(四川大学 计算机学院,四川 成都 610065)

摘 要:准确定位人脸特征点在人脸识别、三维人脸模型重建等领域都有重要作用。目前,针对正面人脸的特征点定位已经相当成熟;但是,当姿态偏转角较大时,准确定位人脸特征点依然是一个有待解决的难题。文中针对姿态偏转比较大的特征点定位,提出了一种多视角人脸特征点定位方法。在训练阶段,针对不同姿态角度,分别定义其特征点模板,并通过训练得到用于特征点搜索的 CPR(级联姿态回归)模型和用于特征点模板选择的 CLM(约束局部模型)模型。在测试阶段,利用每个模型分别对测试样本进行特征点搜索,然后,利用 CLM 模型计算各特征点模板的拟合度,选择拟合度最高的模型作为最终结果。在 FERET 公开库上,与当前比较先进的算法进行的对比实验表明,文中方法有效提高了较大偏转姿态下人脸特征点定位的准确性。

关键词:人脸识别;多视角;人脸特征点定位;模型选择

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2015)11-0001-05

doi:10.3969/j.issn.1673-629X.2015.11.001

A Multi-view Facial Landmark Localization Method Based on CPR and CLM

SHI Zheng-quan, ZHAO Qi-jun, CHEN Hu

(School of Computer Science, Sichuan University, Chengdu 610065, China)

Abstract: Facial landmark localization plays an important role in many face-related applications such as face recognition and 3D face reconstruction. Although existing methods already achieve promising results on frontal and near-frontal face images, their performance on face images with large pose angles is still far from being satisfactory. A multi-view facial landmark localization method is proposed in this paper. It divides head pose angles into a number of non-overlapping ranges. During training, for each range of head pose angles, a facial landmark template is constructed by using the CPR (Cascaded Pose Regression) method, and a corresponding texture model is established by using the CLM (Constrained Local Model) method. During testing, given a new face image, all the templates are applied to it, each resulting in a set of facial landmarks. The associated texture models are then used to compute the fitness values of them, from which the one with the maximum fitness is chosen as the final result. Experiments on the FERET database with comparison to a state-of-the-art method prove the effectiveness of the proposed method in localizing the facial landmarks on face images with large pose angles.

Key words: face recognition; multi-view; facial landmark localization; model selection

1 概 述

人脸特征点在与人脸相关的很多问题中都有着非常重要的作用,比如在人脸识别中人脸特征点广泛用于人脸的对齐、尺度归一化和特征模板抽取^[1],在人脸表情分析中人脸特征点定义的面部形状是表情变化的一个重要依据^[2-3]。因此,在过去的十多年中,人脸特征点定位吸引了大批研究人员的注意,各种不同的方法被纷纷提出。

Taylor 和 Cootes 等提出的主动形状模型 (Active Shape Model, ASM)^[4]是使用最广泛的人脸特征点定位方法。ASM 利用人脸特征点之间的约束关系,通过训练模型对形状参数加以限制,使得搜索到的特征点在一个合理的形状约束之下,从而可以较为精确地进行特征点定位。在 ASM 的基础上, Cootes 和 Edwards 等进一步提出了主动表观模型 (Active Appearance Model, AAM)^[5]。AAM 将形状变化模型与纹理变化

收稿日期:2015-01-06

修回日期:2015-04-16

网络出版时间:2015-11-04

基金项目:国家自然科学基金资助项目(61202160,61202161);科技部重大仪器专项(2013YQ49087904)

作者简介:石正权(1991-),男,硕士研究生,研究方向为模式识别、计算机视觉;赵启军,副教授,硕士生导师,研究方向为模式识别、机器学习、计算机视觉等;陈 虎,讲师,硕士生导师,研究方向为模式识别、图像处理等。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20151104.0952.050.html>

模型结合起来,通过先验知识去指导整个优化过程,使得形状和纹理都具有较高的逼近精度。

在 ASM 和 AAM 的基础上,David C 和 Cootes 提出了约束局部模型 (Constrained Local Model, CLM)^[6]。CLM 跟 ASM 和 AAM 具有相同的形状模型,但是,其纹理模型不同;CLM 以每个特征点为中心选择一个邻域块并将其归一化,然后所有的邻域块拉伸连接成一个灰度向量,作为图像的纹理向量。每一个特征点的局部块均训练一个线性 SVM 来识别它,并且应用到后续的搜索中。

近年来,有不少其他新颖的人脸特征点定位方法被提出,包括基于回归的方法^[7]和基于深度网络的方法^[8]等。在这些方法中,基于级联的姿态回归 (Cascaded Pose Regression, CPR) 方法^[7]在特征点邻域的纹理特征与特征点位置的调整量之间建立回归关系,从而能够从特征点初始位置出发经过多次迭代回归和调整收敛于特征点的正确位置。该方法易于实施,且检测效果也明显优于传统的基于 ASM 和 AAM 的方法。然而,所有这些方法都只考虑了正面或者接近正面的情形,其检测精度在人脸姿态角度增大时明显下降。

针对姿态变化范围比较大的人脸的特征点定位,目前的解决方案大致可以分为三类:基于模板选择的多视角人脸特征点定位方法^[9-10]、基于姿态预估计的多视角人脸特征点定位方法^[11]和基于贝叶斯的人脸特征点定位方法^[12]。其中,第二类方法依赖于姿态预估计的精度,而根据单张二维人脸图像估计人脸姿态本身亦是一个非常困难的问题;第三类方法在人脸特征点模型中引入特征点的可见性属性,将不同姿态下的人脸特征点模型进行统一,并利用贝叶斯模型来预测输入的任意姿态人脸图像上的特征点。不同于这两类方法,第一类方法针对不同的姿态分别定义特征点模型,通过合适的模型选择方法来确定最终的特征点定位结果。这类方法相对直观、复杂度低。因此,文中重点考虑第一类方法。

为了进一步提高现有方法在不同姿态 (尤其是大偏转角度) 下的特征点定位效果,文中采用基于 CPR 的特征点定位方法生成多个有限视角范围的基本特征点模板,采用基于 CLM 的纹理模型作为模板选择的依据确定最终的特征点定位结果,提出一种新的多视角人脸特征点定位方法。与传统的多视角方法相比,文中方法具有下列两个特点:

(1) 采用对姿态变化更加鲁棒的特征点定位基本方法,有效提高特定姿态下的特征点定位效果;

(2) 采用鲁棒性更好、区分能力更强的纹理模型,提高不同姿态特征点模板选择的精度。

得益于上述两个特点,文中方法不但在较小姿态角的人脸图像上取得了和现有方法相当的精度,而且在较大姿态角的人脸图像上也能达到很好的特征点定位精度。

2 相关研究

CPR 方法是形状估计视为一个回归的问题来解决。利用训练所得的回归器就可以直接预测物体的形状。CPR 方法的关键在于计算形状索引特征和训练回归器。文献^[7,13]运用了深度为 5 的随机森林回归器,并且采用形状索引来刻画特征点的特征。形状索引特征首先由 Fleuret 等提出^[14],其输出依赖于输入图片和当前的形状估计值。为了获取对形状变化和姿态角变化具有鲁棒性的特征,文献^[15]中提出通过在两个特征点之间进行线性插值来获得形状索引特征。该形状索引特征能更好地适应形状的变化。

为了解决特征点定位结果对于初值的敏感性问题,并在性能和速度之间做一个权衡,文献^[15]中还提出了重新开始策略:即给定一张图片以及一组不同的初始化值,最初只执行 10% 的级联次数;然后检测预测值的变化,如果变化小于给定的阈值,那么剩下 90% 的级联能够继续执行;否则,就选择一个新的初始值重新开始。

针对多姿态人脸特征点检测的问题,文献^[9]中提出了一种基于多模板 AAM 的人脸特征点检测方法。该方法首先将人脸姿态划分为若干类,为每一类姿态训练一个特定的 AAM 模板;然后对测试样本,利用每个 AAM 模板搜索特征点,并计算搜索结果与模板的相似度,选择相似度最大的 AAM 模板的搜索结果作为最终的特征点检测结果。该方法使用 AAM 的全局纹理模型计算特征点检测结果和模板之间的相似度,对于表情等变化的鲁棒性较差。此外,训练特定姿态的特征点模板时,只采用了该姿态下的数据,这样所得到的模板对于姿态的冗余度较差,而实际应用中输入图像可能具有任意姿态,即并不一定是训练模板所针对的特定姿态角。文中提出的方法能很好地克服以上这些缺点。

该方法能解决水平偏转的人脸特征点定位问题,但其在训练和利用每个 AAM 搜索特征点时需要花费大量时间,影响特征点检测的效率。

3 多视角人脸特征点定位方法

为了解决特征点定位中受人脸姿态角变化的影响,提出了多视角人脸特征点定位的方法。在训练阶段,考虑了水平偏转角度范围为 -90° 到 90° ,分别在水平偏转角为 90° 、 45° 、 0° 、 -45° 、 -90° 进行训练。首先,

采用级联回归定位的方法,对不同姿态角度的训练图片进行训练,获得每个角度下用于特征点搜索的形状模型;然后,采用 CLM 的方法,对不同角度的训练图片选取区分度较高的特征点进行训练(如:鼻尖、嘴角、眼角),以选取的每个特征点为中心,生成正负样本用于训练支持向量机(SVM),正样本以特征点位置(或附近)为中心块,负样本为特征点位置偏移一定距离后以此为块,训练后得到一组参数矩阵。该矩阵与正负样本卷积得到 SVM 分值,所有的 SVM 分值以及相应的标签(1 或 0,1 为正样本,0 负样本),再进行逻辑回归训练得到逻辑回归方程的参数。在测试

阶段,对于给定的测试图片(已经通过人脸检测器检测出人脸区域),采用训练获得的不同角度下的形状模型进行特征点搜索,获得不同角度下的特征点序列;对不同角度下的特征点序列选取与 CLM 训练选取的特征点进行模型选择:分别以每个选取的特征点为中心提取相应的块,经过 SVM 特征点检测和逻辑回归方程的计算得到该点与 CLM 中定义的特征点的相似度,测试图片与该模型的相似度为该模型下由于模型选择的特征点相似度之和;最后挑选相似度最高的模型作为最佳模型输出。整体框图如图 1 所示。

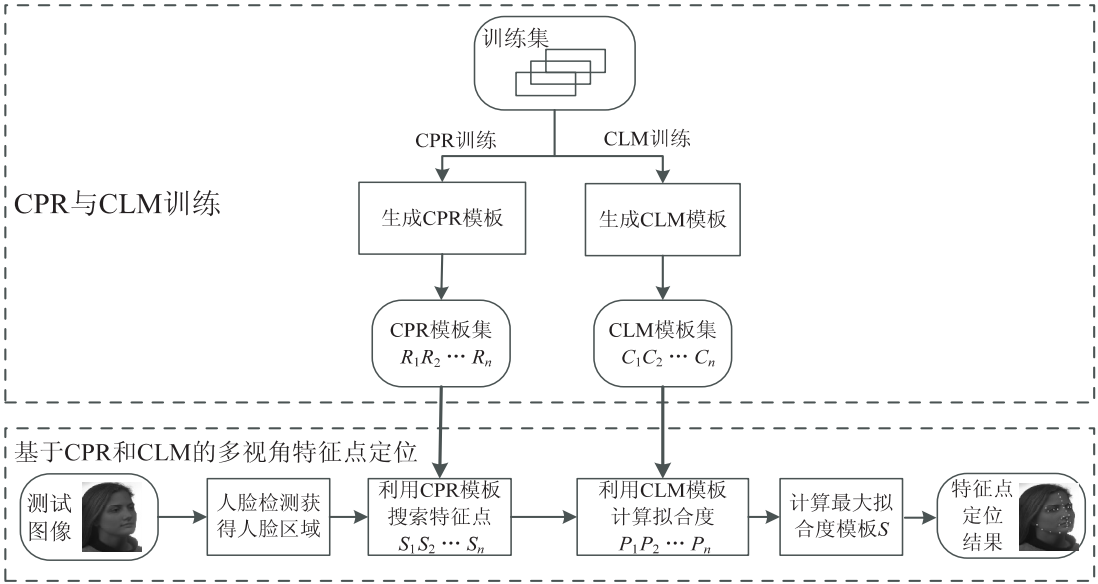


图 1 基于 CPR 和 CLM 的多视角人脸特征点定位方法的流程图

图 1 中,通过训练获得 CPR 模板集 $R_1R_2\cdots R_n$ 和 CLM 模板集 $C_1C_2\cdots C_n$;在测试阶段,首先利用 CPR 模板集搜索特征点 $S_1S_2\cdots S_n$,然后利用 CLM 模板和搜索到的特征点计算模板拟合度,最后选择拟合度最大模板 S 输出。基于特征对形状变化和姿态角变化的鲁棒性考虑,在定义特征时,文中采用了文献[15]中定义形状索引特征的方法:通过在两个特征点之间进行线性插值来获得形状索引特征,形状索引特征能更好地适应形状的变化。为了在性能和速度之间做一个权衡,采用文献[15]中提出的重新开始策略。

4 实验

4.1 实验数据及细节

文中使用的实验数据为 FERET 人脸数据库,该数据库包含 1 万多张多姿态和光照的人脸图像,是人脸识别领域应用最广泛的人脸数据库之一。其中的多数人是西方人,每个人所包含的人脸图像的变化比较单一。基于实验的需要,只考虑存在水平偏转的情况,从 FERET 数据库中挑选了不同姿态角的图片进行实验。

在不同角度的图片下手工标注了不同数量的特征点,图片角度、各个角度下的训练图片数以及图片对应的特征点个数,如表 1 所示。

表 1 不同角度的图片与其特征点个数关系

偏转角/°	训练图片/张	特征点数/个
0	1 345	29
45	521	29
-45	625	29
90	978	25
-90	571	25

因为偏转角度的影响,在不同的水平偏转角度下,人脸的可见部分有很大的差异,所以,用于训练的图片中,每张图片手工标注的特征点个数,以及特征点的位置都是不一样的。为了便于分析误差,测试数据也通过手工标注特征点,不同角度下的测试图片标注的特征点个数和顺序跟该角度下的训练集中标注的特征点个数和顺序一一对应;同时,选取的测试图片不包括在训练集中。

为了对实验结果进行观察和分析,实验数据中包

括:每张图片的像素信息、人脸区域(通过人脸检测获得)、特征点信息。训练分为两个部分进行:首先,采用级联回归估计的方法,应用准备好的训练数据,在每个角度下单独训练,通过这部分训练获得不同角度下的形状模型,用于特征点定位;其次,采用 CLM 的方法,同样是在不同角度下应用相应的训练数据单独训练,获得不同角度下的 CLM 模型,用于模型选择。在测试阶段,由于不知道测试图片的偏转角度,所以,对测试图片,首先分别应用训练获得的 5 个形状模型进行估计,产生 5 个预测的形状模型;然后,将这 5 个预测的形状模型作为 CLM 的输入,从这 5 个预测模型中选择一个最佳模型作为最后的估计模型。

4.2 实验结果

首先,为了验证多视角训练模型在不同姿态偏转角下的预测效果,人为选择不同角度的测试数据,并且选择相对应的形状模型进行测试。在水平偏转角分别为 90°、45°、0°、-45°、-90°的情况下分别用 47、40、47、28、44 张图片进行测试。图 2 展示了不同角度下的部分结果图。

在图 2 中,上面的三张图片从左至右分别是水平偏转角度为 90°、0°、-90°时选择的测试图片;下面的三张图片从左至右分别是水平偏转角度为 45°、0°、-45°时选择的测试图片。因此,在知道测试图片的水平偏转角度的情况下,人为选择相应偏转角度下训练获得形状模型进行特征点定位的准确性是有保证的。

表 2 每张测试图片与各模型的相似度及最佳模型

图片姿态角度	图片编号	90°	45°	0°	-45°	-90°	最佳模型
0°	1	0.462 3	0.640 4	0.843 0	0.650 2	0.547 0	3
-90°	2	0.265 2	0.380 3	0.247 9	0.550 9	0.862 9	5
-90°	3	0.376 2	0.327 3	0.487 4	0.574 8	0.802 7	5
0°	4	0.489 2	0.829 6	0.860 3	0.790 0	0.541 7	3
90°	5	0.731 0	0.489 2	0.302 7	0.320 4	0.284 1	1
45°	6	0.707 0	0.836 0	0.734 9	0.372 0	0.454 9	2
-45°	7	0.422 2	0.443 2	0.476 3	0.879 0	0.586 2	4
-45°	8	0.490 4	0.544 1	0.592 5	0.787 1	0.543 4	4
45°	9	0.624 0	0.881 8	0.820 1	0.509 0	0.419 9	2
90°	10	0.907 0	0.804 9	0.513 6	0.384 5	0.306 0	1
-45°	11	0.494 9	0.452 9	0.648 0	0.855 2	0.630 3	4
90°	12	0.723 8	0.533 5	0.370 5	0.284 6	0.293 6	1

表 2 中每行分别代表该测试图片跟每个模型的相似度及应用 CLM 方法进行模型选择后获得的最佳模型,最佳模型用编号表示,在 90°、45°、0°、-45°、-90°的模型编号分别对应于 1、2、3、4、5。

图 3 是经过模型选择后对应的测试结果图。观察图 3 可以看到,基于多视图模型训练的方法,对姿态偏



图 2 不同角度下的测试结果

转但是对于任意的测试图片,其偏转角度是未知的,对其进行特征点定位就需要应用文中提出的多视角的方法。首先要利用训练获得的 5 个形状模型进行特征点搜索,获得各个角度下的特征点序列;然后将这五个特征点序列作为 CLM 算法的输入进行模型选择,分别计算该测试图片属于各个模型下的相似度,选择相似度最大者作为最佳模型输出。对任意的测试图片,经过 CLM 方法进行模型选择,其中的 12 张图片与各个模型的相似度如表 2 所示,对应的测试图片如图 3 所示(其中,第一行代表编号为 1 到 4 的测试图片,第二行代表编号为 5 到 8 的测试图片,最后一行则代表编号为 9 到 12 的测试图片)。

转具有很好的鲁棒性,测试图片在不同偏转姿态下(水平偏转从-90°到 90°)都可以比较准确地定位到各特征点。作为对比,采用文献[15]中的方法对图 3 中的图片进行测试,结果如图 4 所示。

从图 4 中可以看出,在姿态偏转角度较小时,采用文献[15]中的方法得到的特征点比较准确,然而,随

着姿态偏转角度的增大,特征点定位越来越不准确,特别是在姿态偏转角为 90° 或 -90° 时,定位的特征点完全错误。



图 3 经过模型选择的测试结果



图 4 文献[15]中的方法的测试结果

5 结束语

针对水平偏转角度对人脸特征点定位的影响,文中提出了一种多视角人脸特征点定位的方法。该方法的主要特点是:为减少水平偏转角度对特征点定位准确性的影响,采用回归的方法在不同水平偏转角度下获得相应的形状模型,然后利用 CLM 方法从候选形状模型中获得最佳形状模型作为特征点定位的结果。实

验结果表明,文中提出的方法有效提高了在水平偏转角度下人脸特征点定位的准确性。

参考文献:

- [1] Li S Z, Jain A K. Handbook of face recognition[M]. Berlin: Springer, 2011.
- [2] Martinez A, Du S. A model of the perception of facial expressions of emotion by humans; research overview and perspectives[J]. JMLR, 2012, 13: 1589–1608.
- [3] Ekman P, Friesen W. Facial action coding system[M]. Palo Alto: Consulting Psychologists Press, 1978.
- [4] Cootes T F, Taylor C J, Cooper D H, et al. Active shape models—their training and application[J]. Computer Vision and Image Understanding, 1995, 61(1): 38–59.
- [5] Cootes T F, Edwards G J, Taylor C J. Active appearance models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(6): 681–685.
- [6] Cristinacce D, Cootes T F. Feature detection and tracking with constrained local models[C]//Proc of BMVC. Edinburgh: BMVA, 2006: 929–938.
- [7] Dollar P, Welinder P, Perona P. Cascaded pose regression[C]//Proc of IEEE conference on computer vision and pattern recognition. [s. l.]: IEEE, 2010: 1078–1085.
- [8] Sun Y, Wang X, Tang X. Deep convolutional network cascade for facial point detection[C]//Proc of IEEE conference on computer vision and pattern recognition. [s. l.]: IEEE, 2013: 3476–3483.
- [9] 呼月宁, 张艳宁, 朱宇, 等. AAM 在多姿态人脸特征点检测中的应用[J]. 计算机工程与应用, 2010, 46(12): 161–165.
- [10] 付海晶, 王民, 王稚慧. 一种基于偏转角度的 AAM 人脸特征点定位方法[J]. 计算机技术与发展, 2012, 22(9): 25–28.
- [11] 傅由甲. 多视角下多模板人脸特征定位方法[J]. 计算机工程与设计, 2014, 35(1): 267–270.
- [12] Zhou Y, Zhang W, Tang X, et al. A Bayesian mixture model for multi-view face alignment[C]//Proc of IEEE conference on computer vision and pattern recognition. [s. l.]: IEEE, 2005.
- [13] Cao X, Wei Y, Wen F, et al. Face alignment by explicit shape regression[C]//Proc of IEEE conference on computer vision and pattern recognition. [s. l.]: IEEE, 2012: 177–190.
- [14] Fleuret F, Geman D. Fast face detection with precise pose estimation[C]//Proc of IEEE conference on computer vision and pattern recognition. [s. l.]: IEEE, 2002: 235–238.
- [15] Burgos-Artizzu X P, Perona P, Dollar P. Robust face landmark estimation under occlusion[C]//Proc of IEEE international conference on computer vision. [s. l.]: IEEE, 2013: 1513–1520.

基于CPR和CLM的多视角人脸特征点定位方法

作者：[石正权](#)，[赵启军](#)，[陈虎](#)，[SHI Zheng-quan](#)，[ZHAO Qi-jun](#)，[CHEN Hu](#)
作者单位：[四川大学 计算机学院, 四川 成都, 610065](#)
刊名：[计算机技术与发展](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2015, 25 (11)

引用本文格式：[石正权](#). [赵启军](#). [陈虎](#). [SHI Zheng-quan](#). [ZHAO Qi-jun](#). [CHEN Hu](#) [基于CPR和CLM的多视角人脸特征点定位方法](#)[期刊论文]-[计算机技术与发展](#) 2015 (11)