

一种基于稀疏编码的人脸特征点检测方法

谭暑秋,陈东义,郭承刚,黄志奇,熊小锋

(电子科技大学 自动化工程学院, 四川 成都 611731)

摘 要: 为了能较好地检测遮挡的人脸形状,提出了一种基于稀疏编码的人脸特征点检测方法.在该方法中,采用基于级联回归的思想,用基于稀疏约束的重构模型来迭代搜索人脸特征点位置.首先,利用训练集中的图像学习一个具有全局性以及通用性为特点的稀疏字典,然后结合人脸特征点局部区域内的纹理信息,使用支持向量回归的方法,重构出每次迭代所需的差分形状.为了验证该方法的可行性,将提出的方法在三个公开的数据集上进行测试.进而,为了进一步验证该方法的有效性,将方法与两个性能较好的方法进行比较.实验结果表明所提方法具有可行性,且对遮挡的人脸特征点具有较高的识别率.

关键词: 人脸检测;人脸特征点检测;稀疏编码;人脸校准

中图分类号: TP391

文献标识码: A

文章编号: 1000—7180(2017)09—0007—04

A Sparse Coding-based Method for Facial Feature Point Detection

TAN Shu-qi, CHEN Dong-yi, GUO Cheng-gang,

HUANG Zhi-qi, XIONG Xiao-feng

(School of Automation Engineering, University of Electronic Science and
Technology of China, Chengdu 611731, China)

Abstract: In order to better detect the occlusion of the face shape, this paper propose a sparse coding-based method for facial feature point detection. This method adopts the idea of cascade regression, and uses the reconstruction model based on sparse constraint to iteratively find the positions of facial feature point. Firstly, a sparse dictionary is obtained during training stage, which is global and universal. Then, using the support vector regression and combining the texture information in the local region of the facial feature points, the method can reconstruct differential shape for each iteration. In order to verify the feasibility and effectiveness, this method is tested on three public face datasets, and compares with two methods. The experimental results show that our proposed method is feasible and the recognition rate of occluded face feature points is high.

Key words: face detection; facial feature point detection; sparse coding; face alignment

1 引言

人脸识别领域在近几十年间获得科研人员的关注及研究,并且产生了大量具有较高识别率的人脸识别算法,如文献[1-3].而人脸特征点检测对人脸识别至关重要,低检测率会对人脸识别的精度造成较严重的影响.现有的特征点检测与定位方法在一般情况下,已经达到了较高的检测率以及较小的定

位误差,但当在复杂条件下,如人脸姿态或表情较大,并存在部分遮挡时,人脸特征点检测与定位将会受到较严重的影响.在近几十年中,研究者们针对人脸特征点检测提出大量算法.本文将现有的人脸形状学习方法归为两类:基于局部的方法及基于整体回归的方法.前者通常依赖于局部检测,而后者则以全局的方式对形状与外观同时建模.而在基于局部的方法中,具有代表性的方法是局部约束模型^[4-5].

典型的全局方法是主动表观模型^[6]. Dollár 等人^[7]提出了级联姿态回归,在此基础上,Cao 等人^[8]使用两级级联的方式对模型进行学习;并引入了快速的基于相关性的特征选择策略对模型进行快速训练.该训练框架被 Burgos-Artizzu 等人^[9]得以进一步的改进.本文提出的算法从细处上说是属于局部的方法,但从整体上来说,本文又属于整体回归方法,与原有整体回归方法不同之处在于本文借助了回归模型与稀疏编码的思想,迭代地重构出差分人脸形状.在本文所提方法中,重构的字典是通过将对在测试阶段使用的三个人脸数据库当中提取到的大量特征点局部信息进行训练而得到的.该字典包含了丰富的构成人脸形状的特征信息,进而通过该字典可获得差分形状与稀疏系数之间的连接矩阵.故本文所提方法能在整体性能方面有较好的形状约束.

2 基于稀疏编码的人脸特征点检测方法

级联回归是指以累计方式组合回归序列.在训练期间,其主要的目的是学习回归函数 f 来最小化均方误差:

$$f = \operatorname{argmin}_f \sum_{i=1}^N \|f(I_i, S_i^0) - S_i^*\|_2^2 \quad (1)$$

在上式中 N 指的是存在的训练样本图片数, S_i^* 为真实标注的人脸形状, S_i^0 为第 i 张样本图像的初始化人脸形状.由于存在着复杂的人脸形状,一次迭代回归较难准确地检测出人脸特征点的位置信息.而将一组回归器进行级联却能较好地表达出初始形状与真实形状的映射关系.且在文献^[10]中得出线性函数是有效且高效的结论.

$$\operatorname{argmin}_{H_k} \sum_{i=1}^N \|(S_i^* - S_i^{k-1}) - H_k \Phi(I_i, S_i^{k-1})\|_2^2 \quad (2)$$

式(2)中, H_k 表示的是线性转换矩阵, Φ 是一个非线性的特征描述符.不同的特征描述符体现出不同的性能水平. SIFT^[10] 与 HoG^[11] 在人脸特征点定位方面能表现出较好的性能,因为大多数特征点都在人脸局部区域的边缘.本文采用了 SIFT 特征描述符.

为了提高模型对整体形状的约束能力,结合稀疏编码来对框架进行约束:

$$\operatorname{argmin}_{R^k} \sum_{i=1}^N \|(S_i^* - S_i^k) - R^k f(\gamma^k)\|_2^2 \quad (3)$$

在上述式子中 R^k 是连接矩阵,而 γ^k 则为 S_i^k 所对应的人脸形状的稀疏系数.而 $f(\gamma^k)$ 则可以表示为式(4):

$$f(\gamma^k) = \operatorname{argmin}_{\gamma^k} \sum_{i=1}^N \|\phi_k(I_i, S_i^k) - D\gamma^k\|_2^2, \quad \|\gamma^k\|_0 \leq T \quad (4)$$

ϕ_k 表示的是在第 k 次迭代过程中围绕当前人脸形状 S_i^k 周围提取的 SIFT 特征, D 为形状字典,为了使 γ^k 具有稀疏性,我们设定值 T 对 $\|\gamma^k\|_0$ 进行约束.由于在该方法中主要是通过提取形状的局部特征来完成人脸形状的定位,而当人脸图像存在部分遮挡时,级联回归也较难定位出被遮挡的人脸特征点位置.基于上述问题,本文将通过大量图像样本特征的提取来对稀疏字典进行学习.该字典可对不完整的图像样本进行校正,在人脸特征点定位方法中起到较重要的作用.

$$\operatorname{argmin}_D \sum_{i=1}^{N_{\text{total}}} \|\phi_k(I_i, S_i^k) - D\gamma\|_2^2, \quad \|\gamma\|_0 \leq T \quad (5)$$

式中, N_{total} 为若干个数据库中人脸图像之和.而连接矩阵 R^k 则使用支持向量回归^[12]来获得:

$$R^k = \operatorname{SVR}(S_i^* - S_i^k, \gamma^k) \quad (6)$$

基于稀疏编码的人脸特征点方法的伪代码如下:

输入: 训练图像 I_i 及相应的形状 $S_i, i = 1, 2, \dots, N$; 测试图像 I_j 及初始形状 $S_0', j = 1, 2, \dots, N'$, 稀疏编码参数集; 迭代次数 K , 初始化数 V .

输出: 测试集形状 S^K' .

通过样本数据库的 I, S 及式(5)学习字典 D ;

for $k = 1$ to K

1. 利用式(4), 计算出每次迭代中的 γ^k 值;

2. 通过式(6), 便可得出连接矩阵 R^k ;

end for

for $k = 1$ to K

for $v = 1$ to V

3. 通过式 $f(\gamma^{k'}) = \operatorname{argmin}_{\gamma^{k'}} \sum_{j=1}^{N'} \|\phi_k(I_j, S_0') - D\gamma^{k'}\|_2^2, \|\gamma^{k'}\|_0 \leq T$ 获得 $\gamma^{k'}$.

4. 将 $R^k \gamma^{k'}$ 的值赋给 $\Delta S_0'$;

5. 通过 $S_1' = S_0' + \Delta S_0'$ 计算出一次迭代后的形状估计;

6. 再将估计后的式子重新作为初始形状进行计算, 即 $S_0' = S_1'$

end for

end for

考虑到算法的整体性能,并参考文献^[10],设定了预测过程迭代的次数为 4 到 5 次.

3 实验

3.1 实验参数

在实验中,采用交替验证和局部枚举的方式来获取参数的值,最终通过实验取得较好结果的实验参数,如表 1 所示. ESR^[8]和 RCPR^[9]的参数设置与其原始文件一致. 在 ESR 中,蕨的像素和候选像素特征的数量分别为 5 和 400. 为了构建模型,该方法使用 10 和 500 阶段来训练两级框架. 在 RCPR 中,使用 15 次迭代,5 次重新启动,400 个特征和 100 个随机蕨回归.

表 1 本文中测试时所用参数的具体值

<i>D</i>	<i>T</i>	<i>K</i>	<i>V</i>
512	10	4	10

3.2 数据库

在实验中,采用了三个公开的人脸特征点数据库来对算法进行评估.

Helen 2012^[13] Helen 数据库是从户外采集到的人脸图像集,该图像集包含了任意的表情、姿态以及光照. 在该数据库中,训练集图像有 2 000 张,测试集图像有 330 张,而每张图像标注了 68 个人脸特征点以及相应的位置信息.

LFPW^[14] 由于存在网络的问题,只下载到了 811 张训练图像,以及 224 张测试图像. 在该数据库中存在着两种不同数目的特征点数据集,一种是具有 29 个人脸特征点以及相关形状的数据集,另一种是具有 68 个人脸特征点以及相关形状的数据集. 而在实验当中我们使用后者来测试实验中的方法.

COFW^[9]该数据库包含了 1 345 张训练图像以及 507 张测试图像,该数据库的较多图像都存在较大姿态的变化以及各式遮挡. 该数据库每个图像中都包含了 29 个人脸特征点以及相应的形状信息,且该数据库还提供了每个点是否被遮挡的信息.

3.3 实验标准

本文采用欧氏距离来表示检测到的人脸特征点与标注的真实人脸特征点坐标之间的校准误差,并采用两眼瞳孔之间的距离将其规一化.

$$d_{\text{error}} = \frac{\|P - P_{\text{ground-truth}}\|_2}{\|P_{\text{leye}} - P_{\text{reye}}\|_2} \tag{7}$$

在式(7)中, P 与 $P_{\text{ground-truth}}$ 分别表示检测到的人脸特征点的位置信息以及真实标注的人脸特征点的位置信息. P_{leye} 与 P_{reye} 分别表示真实标注中的左眼中心与右眼中心的位置信息.

3.4 实验结果

本文使用公式(7)对所估计的形状进行规一化

处理,而不是以像素为单位进行归一化处理. 首先我们将本文方法在数据集 COFW(29 个点)、Helen(68 个点)以及 LFPW(68 个点)上进行测试,然后将其测试的结果与 RCPR 及 ESR 进行对比分析. 除了对方法整体情况进行分析之外,我们还对人脸形状的局部区域进行测试. 类似地我们也做了三个方法的对比实验. 其分析如图 1 所示,横坐标表示的是归一化后的误差比,纵坐标表示的是方法运行时测试图像所占比. 从图 1 中可得,本文方法在数据集 COFW(29 个点)以及 LFPW(68 个点)上的累积误差分布(CED)曲线都略优于 RCPR 及 ESR 的测试结果. 而图 1(b)中,在刚开始阶段 RCPR 曲线略好于本文方法,随着实验的继续,本文方法表现出了较好的性能. 其原因是在 RCPR 方法中,其训练图像库为每张图片提供了遮挡信息,并用级联回归器对其遮挡信息进行训练,故其在存在较多遮挡情况的 Helen 数据库当中具有较好的性能. 而本文的方法具有全局稀疏约束的特性,故对多姿态以及较大遮挡的人脸形状也具有较好的重构能力.

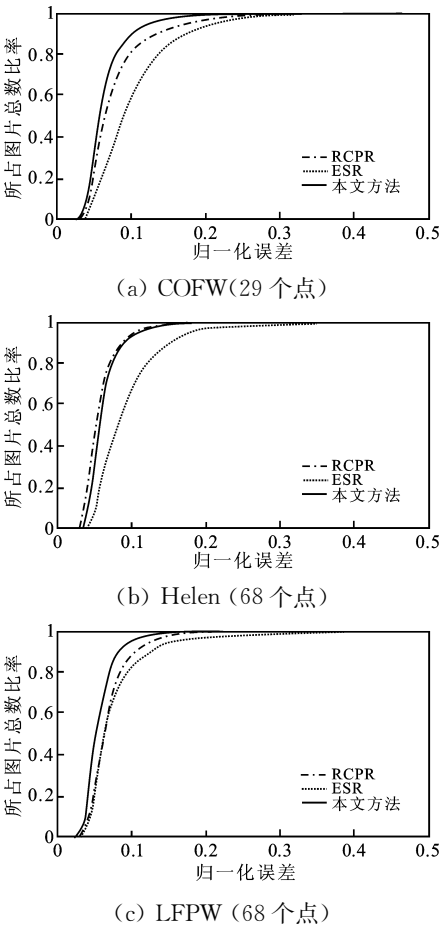


图 1 ESR、RCPR 以及本文方法在不同数据库上的累计误差曲线图

通过对人脸形状局部区域的测试,可以从另一个角度对方法进行分析,如图 2 所示. 图 2 中横坐标表示的是人脸形状区域,纵坐标表示的是归一化后的误差比. 从图 2 可知在三个人脸数据库中 RCPR、ESR 以及本文方法在人脸轮廓区域的检测精度都较差. 其主要原因是轮廓区域的点分布比较分散,且较其它区域更易形成遮挡. 相比之下,三个方法在眼睛区域产生的相对检测误差较低,其与眼睛特征点的分布有较大关系. 表 2 给出了在三个数据集上三个方法最终的人脸特征点的预测误差值,从该表所提供的结果可知本文方法可以对自然环境下的脸形状的整体以及局部进行较好地检测,并且对存在遮挡的人脸形状也能进行较好地定位.

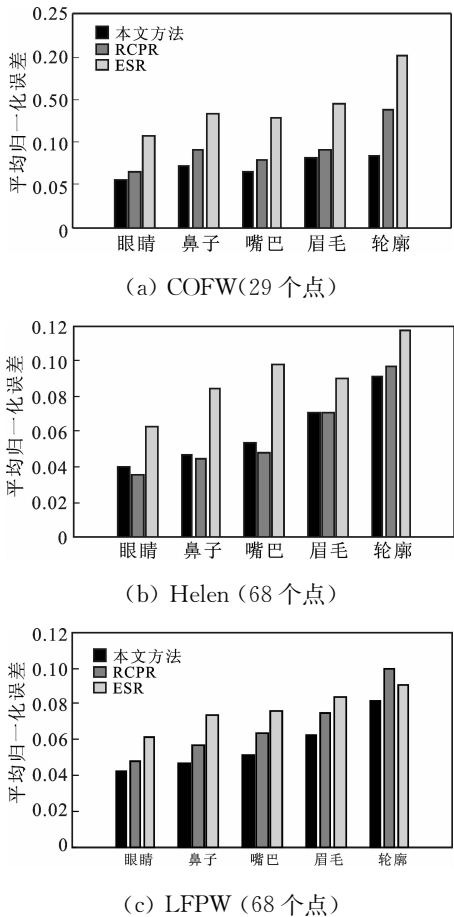


图 2 ESR、RCPR 以及本文方法在数据集上局部区域的误差柱状图

表 2 ESR、RCPR 以及本文方法在数据集上的误差值

方法	COFW (29 个点)	Helen (68 个点)	LFPW (68 个点)
ESR	0.105 5	0.093 5	0.077 9
RCPR	0.083 4	0.061 1	0.072 4
本文方法	0.065 8	0.063 4	0.057 6

4 结束语

本文提出了一种基于稀疏编码的人脸形状重建方法. 所提出的方法从训练阶段中学习到了稀疏字典以及连接矩阵,再将获得的连接矩阵用于测试图像,使其重构出差分人脸形状,进而用回归的方式估计出人脸形状的最终位置. 为了验证本文方法的可行性以及有效性,本文将提出的方法与 ESR 以及 RCPR 两个方法在公开的人脸数据库 COFW(29 个点)、Helen(68 个点)以及 LFPW(68 个点)上作了对比实验. 实验结果显示,所提出的方法具有可行性,并对遮挡的人脸形状具有较高的检测精度.

参考文献:

[1] 覃磊,李德华,周康. 基于 QR 分解与 2DLDA 的单样本人脸识别[J]. 微电子学与计算机,2015,32(2):65-68.

[2] 覃磊,李德华,周康. 基于分块 2DPCA 与 2DLDA 的单训练样本人脸识别[J]. 微电子学与计算机,2015,32(11):105-110.

[3] 蒋可心,于鸿洋. 基于摄像头 PTZ 控制的人脸识别[J]. 微电子学与计算机,2014,31(4):48-51.

[4] Cristinacce D, Cootes T F. Feature detection and tracking with constrained local models[C]//Proc. Brit. Mach. Vis. Conf., Edinburgh, UK,2006(2):929-938.

[5] Saragih J, Goecke R. A nonlinear discriminative approach to aam fitting[C]//Proceedings of the IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007:1-8.

[6] 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.

[7] Dollár P, Welinder P, Perona P. Cascaded pose regression[C]//Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. San Francisco, CA, USA, IEEE,2010:1078-1085.

[8] Cao X, Wei Y, Wen F, et al. Face alignment by explicit shape regression[J]. International Journal of Computer Vision, 2014, 107(2):177-190.

[9] Burgos-Artizzu X, Perona P, Dollar P. Robust face landmark estimation under occlusion[C]//Proceedings of the IEEE International Conference on Computer Vision,2013. Washington, 2013:1513-1520.

[10] Xiong X, De F la Torre. Supervised descent method and its applications to face alignment[C]// Proceedings of the IEEE conference on computer vision and pattern recognition, Portland, Oregon,2013:532-539.

DDOA 算法,在减少能耗以及减少延迟方面更具优势,从而延长了整个网络的生命周期.

参考文献:

[1] Gupta M, Saraswat L. Energy aware data collection in wireless sensor network using chain based PEGASIS [C]//Recent Advances and Innovations in Engineering (ICRAIE), 2014. Jaipur, India. IEEE, 2014; 1-5.

[2] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks [C]//System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on. Hawaii, IEEE, 2000(2); 10.

[3] Cheng C T, Chi K T, Lau F C M. A delay-aware data collection network structure for wireless sensor networks[J]. IEEE Sensors Journal, 2011, 11 (3): 699-710.

[4] Tan H Ö, KÖrpeolu I. Power efficient data gathering and aggregation in wireless sensor networks[J]. ACM Sigmod Record, 2003, 32(4): 66-71.

[5] 韩伟伟,降爱莲. 面向无线传感器网络数据融合的延迟优化算法[J]. 小型微型计算机系统, 2015, 35 (5): 949-953.

[6] 陶志勇,王雪. 基于分簇与自适应加权的 WSN 数据融合算法 [J]. 微 电 子 学 与 计 算 机, 2013, 31 (11): 136-138.

作者简介:

徐艮凤 女,(1990-),硕士研究生. 研究方向为无线传感器网络、嵌入式系统、计算机网络.
E-mail:2467933944@qq. com.

张曦煌 男,(1962-),教授. 研究方向为无线传感网、嵌入式系统、计算机网络、图形与图像处理、计算机分布式控制与智能控制.

(上接第 10 页)

[11] Yan J, Lei Z, Yi D, et al. Learn to combine multiple hypotheses for accurate face alignment[C]// Proceedings of the IEEE International Conference on Computer Vision Workshops. Barcelona, Spain,2013;392-396.

[12] Basak D, Pal S, Patranabis D C. Support vector regression [J]. Neural Information Processing-Letters and Reviews, 2007,11(10): 203-224.

[13] Le V, Brandt J, Lin Z, et al. Interactive facial feature localization[C]//Proceedings of European Conference on Computer Vision. Crete, Greece, 2012;679-692.

[14] Belhumeur P, Jacobs D, Kriegman D,et al. Localizing parts of faces using a consensus of exemplars [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. South Carolina, 2011;

545-552.

作者简介:

谭暑秋 女,(1982-),博士研究生. 研究方向为人脸识别.
E-mail: tanshuqiul23136@hotmail. com.

陈东义 男,(1956-),教授,博士生导师. 研究方向为计算机视觉、增强现实、身体传感网络.

郭承刚 男,(1990-),博士研究生. 研究方向为计算机视觉、增强现实.

黄志奇 男,(1977-),博士,副教授. 研究方向为人机交互、可穿戴性.

熊小锋 男,(1992-),硕士研究生. 研究方向为人机交互、可穿戴性.