

稀疏样本人脸检索方法研究

沈旭东, 蒋 婷, 朱 明

(中国科学技术大学 自动化系, 安徽 合肥 230027)

摘 要: 提出一种将 Gabor, LGBP, LPQ 三种特征融合的人脸特征描述方法. 将人脸按五官进行分块, 分别提取 Gabor, LGBP 和 LPQ 特征, 再通过融合得到最终的融合特征来表征每一个人脸. 实验表明, 在 FERET, AT&T 数据集以及在网吧采集的实际人脸数据集上, 该融合特征检索性能与当前常用的特征相比有了一定程度的提高.

关键词: 特征融合; 人脸检索; Gabor; LGBP; LPQ

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2017)01-0081-04

Approach on Sparse Samples Face Retrieval

SHEN Xu-dong, JIANG Ting, ZHU Ming

(Department of Automation, University of Science and Technology of China, Hefei 230027, China)

Abstract: Face retrieval training samples is difficult to obtain in many situations. Therefore, an effective way to describe the human face becomes more important. We have fused Gabor feature, LGBP feature and LPQ feature to describe face in this paper. Faces are divided into blocks according to features, than fusion features obtained to characterize each individual face. Taken together, our results suggest that fused features for data retrieval have a certain improvement compared with the current popular features in FERET, AT&T database and ours dataset.

Key words: feature fusion; face retrieval; Gabor; LGBP; LPQ

1 引言

人脸检索由于自身独有的自然性和无感知等特点, 在视频监控、信息安全、刑侦等方面有着广泛的应用. 而现实场景下, 大数量的训练样本的采集依然非常困难, 使得深度网络特征的人脸检索性能很难满足要求. 针对上述问题, 研究者们提出了多种方法, 其一是生成虚拟样本, 增加训练样本的个数^[1]. 其二是利用事先收集好的辅助数据集进行训练^[2]. 其三是充分利用已有的少量训练样本, 找到较好的人脸描述方法^[3]. 本文正是根据第三点思路, 提出一种融合多种特征的人脸描述方法进行人脸检索.

2 融合特征人脸检索方法

本文的主要贡献是找到一种有效的多特征融合方法. 考虑到各种人脸局部特征描述方法的特点以

及实际使用过程中的经验, LGBP, Gabor, LPQ 三种特征比其他特征表现出了更加优秀的性能, 本文选择这三种特征作为基础特征, 再根据本文的融合方法, 得到一种融合特征, 利用这种融合特征来表征人脸进行检索.

2.1 人脸图像预处理

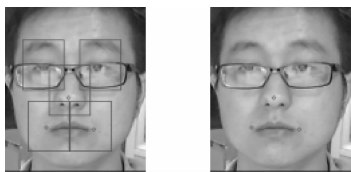
现实环境下, 人脸图像往往受到光照、姿态、遮挡的影响, 为了减小上述因素对人脸检索性能的影响, 需要事先对人脸进行预处理.

本文采用了光照均衡和人脸对齐两种方法对人脸图像进行预处理. 采用了文献[4]中提出的 SMMML 方法进行光照均衡. 人脸对齐则是先在人脸图像中找到人脸的位置, 再使用 ASM 算法检测出人脸两个眼球中心、鼻尖和两个嘴角位置共 5 个关键点. 一般认为人类两个眼球中心位置在一条水平线上, 这样就可以根据两个眼球连线与水平方向的夹角, 来校正人脸

的倾角,再根据连线的距离将人脸归一化到相同的大小(本文采用 60×60)。如图 1(a)所示。



(a)不同姿态人脸关键点定位



(b)基于五官的人脸分块(c)预处理后的人脸图片

图 1 人脸预处理

2.2 人脸特征抽取

人脸预处理减小了人脸光照、姿态等对人脸检索性能的影响。但是,检索性能的提高,更加关键的是找到一种有效的人脸特征描述方法。基于这一点,本文采用的方法是首先分别提取人脸图片的 LGBP, Gabor, LPQ 三种特征。再分别将三种特征按照不同的权值融合后生成新的融合特征(Fusion Feature, FF)。

2.2.1 LGBP

LGBP 是为了抽取采样后的 Gabor 幅度图像的局部空间变化信息,对采样 Gabor 幅度图像计算每个点的 LBP 值。由于人脸遮挡会对人脸检索的结果产生影响,而解决此类问题的途径是提取局部特征。因此,本文在提取 LGBP 特征时,将人脸按照五官位置分成 5 个小块,如图 1(b)所示,分别提取每一个小块作为该人脸的 5 个局部特征,同时提取整个人脸的一个全局特征,6 个部分特征串联再得到了整个人脸的 LGBP 特征。

LBP 算子采用 3×3 的窗口,编码形式如式(1)所示。

$$\text{LBP}(x_c, y_c) = \sum_{n=0}^7 2^n S(i_n - i_c) \quad (1)$$

2.2.2 Gabor

Gabor 小波变换是利用不同尺度和不同方向的 Gabor 滤波函数与二维图像求卷积,它能够提取人脸图像在不同尺度不同方向上的特征。本文采用的是如下形式的 Gabor 滤波函数。

$$\varphi_{\mu, \nu}(z) = \frac{\|k_{\mu, \nu}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu, \nu}\|^2 \|z\|^2}{2\sigma^2}} [e^{ik_{\mu, \nu} \cdot z} - e^{-\frac{\sigma}{2}}] \quad (2)$$

式中, μ 和 ν 是 Gabor 核函数的方向和波长, $\|\cdot\|$ 是范数操作,小波矢量 $k_{\mu, \nu} = k_\nu (\cos \varphi_\mu, \sin \varphi_\mu)$, 这里 $k_\nu = \frac{k_{\max}}{f^\nu}$, $\varphi_\mu = \frac{\pi \mu}{8}$, $k_{\max}$ 是最大频率, f 是不同频域核函数之间的距离系数。把滤波器设置为 5 个波长 8 个方向,得到 40 个 Gabor 幅度图像,采样降低图像维度后,串联各图像作为最终的 Gabor 特征。

2.2.3 LPQ

现实场景下,很多采集到的图片有一定程度的模糊,在样本数量有限的情况下,如何利用好这些模糊图像获得好的性能是稀疏样本人脸检索问题的关键之一。

鉴于此,在融合特征中加入了局部相位量化(Local Phase Quantization, LPQ)特征。LPQ 特征能够有效地对图像的纹理进行分析,该方法认为观察到的图像是原始图像经过一个模糊变换过程后的结果。如式(3)所示。

$$g(x) = f(x) \times h(x) \quad (3)$$

式中, $g(x)$ 是观察到的图像, $f(x)$ 是原图像, $h(x)$ 是模糊函数。式(3)的傅里叶变换形式为

$$G(u) = F(u) \times H(u) \quad (4)$$

为了有效反映人脸的局部信息,LPQ 特征通常计算局部图像的傅里叶变换。同样的,本章选取人脸 5 个关键点的邻域提取 LPQ 特征最终得到一个 6 400 维的 LPQ 特征。

2.3 特征融合

根据上述方法,本文完成了 3 种基本特征的提取。然而,如何完成 3 种特征的融合,并且提高融合特征的检索性能是本文需要研究的重点问题。

实验发现,3 种特征在不同的检索条件下各有优缺点,比如 Gabor 特征对光照变化较为鲁棒,而经过本文改进后的 LGBP 特征则同时对光照和姿态性能优良,对于较为模糊的人脸,LPQ 则更为优秀。因此,本文将 3 种特征加权后串联在一起作为特征融合的方法,具体融合思路如图 2 所示。图中的 $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ 称为融合向量。融合向量是通过训练生成的。

根据示意图我们发现,随着三种特征维度的增加,3 个融合向量 $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ 的维度也会相应增加,这就意味着需要训练的参数增加,而本文考虑的是稀疏样本情况下的人脸检索问题,训练样本的数量是有限的。对于这一问题,本文采用的方法是共享参数。对于每一个人脸,上文中将之分成了 5 个块,则融合矩阵对每一个块里面的权值参数采用同一个

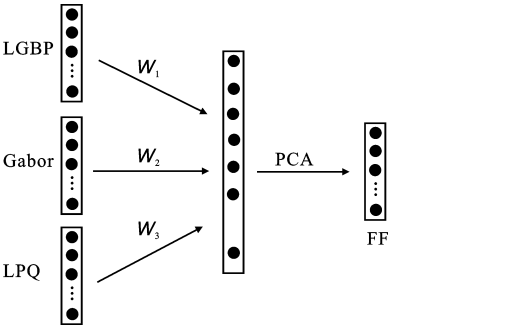


图 2 特征融合示意图

值,如此,权值参数就会大幅下降,使训练过程更加容易收敛。

综上所述,本文的人脸检索算法流程如下所示。

Step1. 人脸图像预处理。

Step2. 人脸面部区域按主要器官分块。

Step3. 根据分块后的人脸,分别计算人脸的 LGBP,Gabor,LPQ 特征。

Step4. 初始化权值向量 W_1,W_2,W_3 为全 1。

Step5. 根据特定的权值向量 W_1,W_2,W_3 ,融合上述人脸特征进行训练,代价函数如式(3)所示。

$$J(W;x,y)=\frac{1}{2}||f_{w,b}(x)-y|| \tag{5}$$

式中, W 为权值向量, x 为融合特征, y 为样本标签。

Step6. 判断是否满足收敛条件,满足则结束训练,否则利用梯度下降法更新 W ,再转到 S4。

Step7. 使用权值向量 W 得到融合特征后利用相似距离度量方法进行人脸检索。

3 实验结果分析

为验证本文方法的有效性,分别在 2 个通用人脸数据集 FERET, AT&T 和另一个我们在与合肥市公安局合作时在现实场景下采集的部分可以公开的人脸数据集(Public Faces Dataset)中进行实验。本文将实验结果与当下比较常用的 Eigenfaces,Uniform Pursuit(UP)^[5], Adaptive Discriminant Analysis(ADA)^[6] 三种特征进行了比较,所有参数都按照原论文中的参数进行设置。

3.1 FERET 人脸数据集

本文选择 FERET 人脸库^[7]中的 200 个人,每个人 7 张照片进行实验。实验参数选择如表 1 所示。其中 train_num 是训练图像的个数,test_num 是测试图像的个数,range 表示检索最终输出的结果数。

表 1 FERET 实验参数设定

实验参数	1	2	3
train_num	200	600	1 000
test_num	1 200	800	400
range	2	2	2

注:表中所有人脸图像数对应每一个人的数目。

实验结果如图 3 所示。从图中可以看出,LG-BP,Gabor,LPQ 比 Eigenfaces,UP,ADA 三种特征具有更加优秀的性能,而我们的融合特征比上述 3 种特征的性能更加优越,特别是在训练样本只有 1 时(即单样本人脸检索问题)更是比最好了 LPQ 算法高出了 5 个百分点。因此,在 FERET 上,本文提出的融合特征具有很好的效果。

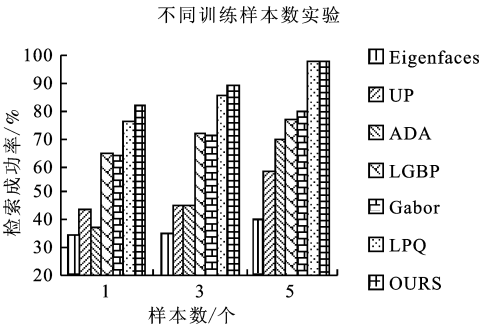


图 3 FERET 上实验结果

3.2 AT&T 人脸库

AT&T 人脸库^[8]中包含了 40 个人,每人 10 张图像,共 400 张人脸图像。实验参数选择如表 2 所示。

表 2 AT&T 实验参数设定

实验参数	1	2	3	4	5
train_num	40	80	120	200	280
test_num	360	320	280	200	120
range	2	2	2	2	2

实验结果如图 4 所示,实验比较了 LGBP,Gabor,LPQ 与本文融合特征的性能。在 AT&T 人脸库上,本文的融合特征依然具有比较优秀的性能,在 train_num=1 时,由于参与训练的样本过少,只有 40 张,因此,融合特征的检索成功率比 LPQ 特征略差,而优于 LGBP 和 Gabor。随着训练样本数量增加,融合特征检索成功率不断提高。因此,本文提出特征融合方法比较适合于有一定量训练样本的稀疏样本人脸检索问题。

3.3 网吧采集数据集

为了验证本文提出的方法在实际环境下的效果,我们与合肥市公安局合作在合肥市的多家网吧部署了监控摄像头人脸采集系统。由于涉及到个人隐私问题,本文不方便公开真实采集到的人脸。因此,我们选择了 40 位志愿者分批次,分时段在实验网吧参与采

集,图 5 展示了部分经过预处理的人脸图像。

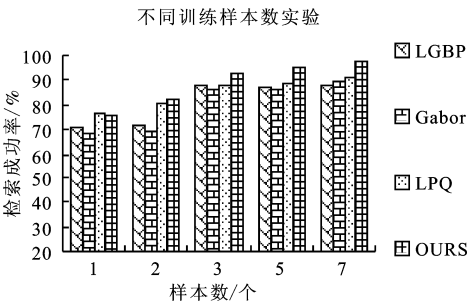


图 4 AT&T 上实验结果

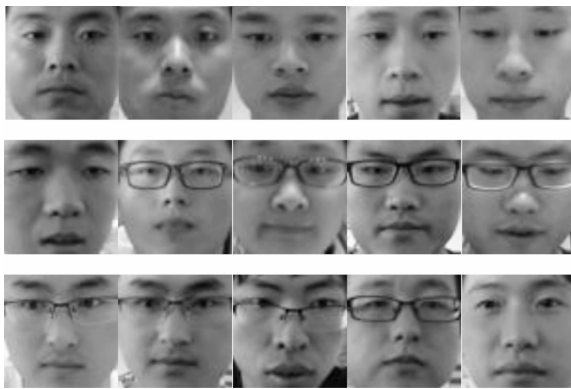


图 5 网吧采集数据集

实验采用了网吧采集数据集中 300 个人共计 3 000张图片。实验参数如表 3 所示。实验结果如图 6 所示,从图中可以看出,在现实场景下,本文提出的融合算法,与 LGBP,Gabor,LPQ 三种特征比较,优势依然比较明显,尤其在训练样本数量较多时。

表 3 网吧采集数据集参数设定

实验参数	1	2	3	4	5
train_num	300	600	900	1 500	2 100
test_num	2 700	2 400	2 100	1 500	900
range	2	2	2	2	2

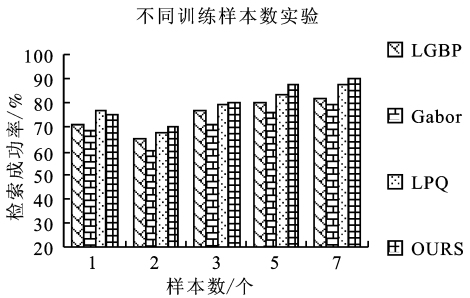


图 6 网吧采集数据集实验结果

4 结束语

本文提出一种稀疏样本人脸检索特征描述新方法,将常用的 LGBP,Gabor,LPQ 人脸特征相互融合,取

得了优于上述三个特征的检索性能。同时经过在实际采集人脸数据集下的实验,证明本文提出的特征在现实条件下也具有一定的实用价值。然而,由于在样本稀疏的条件下,较大的光照、姿态等条件的变化,会使本文的算法性能下降明显,因此,如何表征外部条件变化较大时的人脸将成为今后一段时间的研究重点。

参考文献:

[1] Wang Z, Zhu M. Sparse representation-based face recognition for single example image per person[C]//Informatics in Control, Automation and Robotics. Berlin Heidelberg:Springer ,2011.

[2] Majumdar A ,Ward R. Pseudo-fisherface method for single image per person face recognition [C]//Acoustics, Speech and Signal Processing (ICASSP 2008), IEEE International Conference on. LasVegas, NV: IEEE,2008:989-992.

[3] Turk M,Pentland A. Eigenfaces for recognition [J]. Journal of Cognitive Neuroscience,1991,3(1):71-86.

[4] Ji S, Zhang B, Du D, et al. Similarity Michaelis-Menten law pre-processing descriptor for face recognition[C]//Neural Networks (IJCNN), 2014 International Joint Conference on. Beijing: IEEE, 2014: 1272-1277.

[5] Deng W, Hu J, Guo J, et al. Robust, accurate and efficient face recognition from a single training image: A uniform pursuit approach [J]. Pattern Recognition, 2010, 43(5): 1748-1762.

[6] Kan M, Shan S, Su Y, et al. Adaptive discriminant analysis for face recognition from single sample per person[C]// Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. Santa Barbara:IEEE, 2011: 193-199.

[7] Phillips P J, Wechsler H, Huang J, et al. The FERET database and evaluation procedure for face-recognition algorithms [J]. Image and vision computing, 1998, 16(5): 295-306.

[8] Samaria F, Young S. HMM-based architecture for face identification [J]. Image and vision computing, 1994, 12(8): 537-543.

作者简介:

沈旭东 男,(1989-),硕士研究生.研究方向为人脸识别、深度学习、图像处理. E-mail: xdshen@mail. ustc. edu. cn.

蒋 婷 女,(1991-), 硕士研究生.研究方向为安卓开发、图像处理.

朱 明 男,(1965-),教授,博士生导师.研究方向为机器学习、数据挖掘、人脸识别.