

一种重要区域重采样的人脸检索方法

蒋舒卉, 钱学明

(西安交通大学信息与通信工程系, 710049, 西安)

摘要: 针对人脸中不同部位对检索和识别的区分力度的影响问题, 提出了一种基于重要区域重采样的人脸检索方法. 该方法不仅继承了图像分块后的特征提取保留局部信息的优点, 而且只需对人脸中最重要的部位进行重采样, 即以一定采样步长移动采样小块, 密集采集重要区域, 并以指定距离交叠其他小块, 从而避免了均匀分块对重要区域划分不当而丧失判别力的问题, 进而降低了识别复杂度, 提高了识别的抗干扰能力. 在人脸检索中, 所提方法实现了重要区域的加权处理, 与现有均匀分块方法相比, 在计算复杂度增加不大的情况下大大提高了人脸的识别能力. 大量的对比实验验证了所提方法的优越性.

关键词: 人脸检索; 重要区域重采样; 特征提取

中图分类号: TN911.73 **文献标志码:** A **文章编号:** 0253-987X(2012)02-0119-05

A Face Retrieval Method Based on Resampling of Interested Region

JIANG Shuhui, QIAN Xueming

(Department of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A face retrieval approach is proposed based on the re-sampling of interested region to deal with the problem that the distinguish ability of detection and retrieval is different for different parts of face. The approach inherits the advantage that the feature extraction after blocking the image can remain local information, and just re-samples important parts of human face by moving the sampling block in a certain sampling step. Thus important regions are sampled densely and meanwhile the blocks are overlapped in a specified distance. The approach not only avoids losing the judgment of improper classification in important regions caused by the uniform block, but also reduces complexity and enhances the anti-interference ability. The approach realizes the weighting of important region in face retrieval, and its performance is dramatically enhanced under the circumstance of little increase of computational complexity, compared with the existing uniformed blocking method. A large number of comparative experiments verify the conclusion.

Keywords: face retrieval; interest region resampling; feature extraction

在身份检索、门禁系统、网络安全、电子护照、案件侦破中, 人脸检索技术发挥着重要的作用. 从面部区域划分的角度出发, 人脸识别可分为基于整脸的^[1]、几何特征的^[2]、组件的(component-based)^[3-5]及分块的^[6-9]. 近年来, 基于主成分分析(PCA)^[1]、神经网络^[10]等的人脸识别方法得到了广泛应用. 这些方法主要是针对整脸信息进行提取和处理的, 而在

整图处理时易丢失局部(人脸器官)信息, 因此针对局部信息的识别方法应运而生, 即先将图像分成若干小块, 再对每一小块提取特征^[6-7]. 基于分块的方法一般为均匀分块、均匀分块加权重^[6-7]和均匀交叠^[8]等方法.

Ahonen 等^[6-7]提出, 人脸图像是由微纹理组成的, 而局部二值模式(LBP)算子是描述纹理信息的

收稿日期: 2011-06-29. 作者简介: 蒋舒卉(1989—), 女, 硕士生; 钱学明(通信作者), 男, 博士, 讲师. 基金项目: 国家自然科学基金资助项目(60903121, 61173109); 微软亚洲研究院资助项目.

网络出版时间: 2011-10-27

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20111027.0957.001.html>

最有效方法之一,他们将 LBP 特征^[11-12]运用于人脸检测的结果表明,在 FERET 数据库下,基于均匀分块 LBP 特征的相似性度量结果较 PCA、贝叶斯等方法更为有效。

一个实际的人脸检索系统需要应对光照、表情和遮蔽等问题,文献[13]将 bag of words(词袋)检索方法与人脸特性相结合,提出了基于分块的 bag of words 的人脸检索方法,在极端表情变化和有遮蔽的情况下取得了良好的效果。

均匀分块不能将人脸中重要区域突显出来。Ahonen 等^[7]又在均匀分块的基础上提出了均匀分块加权重的方法,即将图像分块按照包含信息的重要性配以权重,如眼睛的权重最高,两颊的权重较低,这样器官信息具有更高的区分度。由于重要区域经过划分后属于不同的分块,所以均匀分块加权重仍不能解决信息离散化带来的问题,而解决此类问题通常采用均匀交叠法。文献[8]对 LBP 方法进行了改进,由此提出了 VLBP 方法和 LBP-TOP 方法,并与均匀交叠分块相结合,在应对各种极端表情变化时取得了良好的效果,但在不重要区域甚至干扰区域的识别上性能有所下降且复杂度很高。

为了解决分块方法存在的问题,本文提出了一种基于重要区域重采样的人脸检索方法,该方法中人脸检索的相似性度量采用特征向量相似性度量方法。

1 方法描述

本文特征向量相似性度量方法的流程如图 1 所示,共分 7 个步骤。

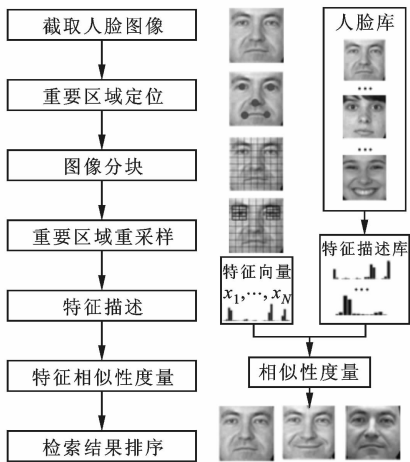


图 1 特征向量相似性度量方法的流程

步骤 1 利用训练后的分类器对人脸图像进行检

测^[9],并从整幅图中截取出所检测的人脸区域。

步骤 2 利用文献[4]方法定位图像的关键点,即用一张具有眼睛、鼻子、嘴巴集合的模板在人脸图像中移动,以匹配面部图像。该定位适用于小范围旋转变化的情况。

步骤 3 对图像进行分块,即将宽为 W 、高为 H 的图像进行分块,均匀分为 $M \times N$ 块,每个小块的宽 $w=W/M$ 、高 $h=H/N$ 。

步骤 4 对步骤 2 得到的重要区域进行重采样处理,即以一定采样步长移动采样小块,由此产生密集的和指定距离交叠的其他小块。

步骤 4 的具体操作为,通过关键点定位和五官的几何位置来确定器官所属区域的边界位置,再从边界起始位置以一定采样步长间隔地进行小块采样,直至边界终止位置与小块终止位置重合为止。本文中,采样小块的尺寸是固定的且为均匀的,采样步长为 $w/m, h/n$,其中 m, n 为均匀划分后获得的小块数。为了方便讨论采样步长对人脸检索性能的影响,将 m, n 相同时记为 hw/n 。以左眼为例的重采样描述如下。首先确定左眼区域边界的位置,记左眼中心位置为 (e_{lx}, e_{ly}) ,右眼中心位置为 (e_{rx}, e_{ry}) ,两眼中心距离 (d_x, d_y) 为 $(|e_{rx} - e_{lx}|, |e_{ry} - e_{ly}|)$ 。根据普遍的五官几何分布统计可知,左眼边界位置的上、下、左、右依次为 $(e_{ly} - d_x/4), (e_{ly} + d_x/4), (e_{lx} - d_x/4), (e_{lx} + d_x/4)$ 。由于本文方法是在整个区域内进行重采样的,因此对位置精度要求不高,这样算法简单且符合一般性要求。在得到左眼边界值之后,第一个采样小块的起始坐标应与边界坐标重合,采样小块的起始位置通式为 $(e_{lx} - (d_x/4) + (iw/m), e_{ly} - (d_x/4) + (jh/n))$,其中 i, j 分别为横、纵分块序号。

当 $e_{lx} - (d_x/4) + (iw/m) + w \geq e_{lx} - (d_x/4)$ 或者 $e_{ly} - (d_x/4) + (jh/n) + h \geq e_{ly} - (d_x/4)$ 时,重采样终止。

步骤 5 将步骤 3 对整张图像所分小块和步骤 4 得到的重要区域重采样统一进行特征提取,再经统计得到特征直方图,将所有小块的特征直方图进行拼接得到测试图像的特征描述。设每个小块的特征为 $x_i, i \in \{1, \dots, N_{\text{sum}}\}$, N_{sum} 为总分块数, $x_i \in R^d, d$ 为特征维数。经拼接后得到的特征描述 $X = \{x_1, \dots, x_{N_{\text{sum}}}\}$, $X \in R^{dN_{\text{sum}}}$ 。在均匀分块中, $N_{\text{sum}} = MN$; 在均匀交叠中, $N_{\text{sum}} = M(M-1)N(N-1)$; 在重要区域重采样中, $N_{\text{sum}} = MN + N_{\text{cross}}$, N_{cross} 表示交叠的块数。

步骤 6 将待检测图像的特征描述与测试库中经过

预处理得到的所有图像的特征描述进行比较,再通过这 2 种图像间的距离来判断它们的相似性. 本文的相似性度量准则为 CHI 平方统计法 (CHI) 和直方图相交距离法 (HI)^[7], 其表达式分别为

$$\chi^2(S,M) = \sum_i \frac{(S_i - M_i)^2}{(S_i + M_i)}$$
$$D(S,M) = \sum_i \min(S_i, M_i)$$

式中: S_i 和 M_i 分别表示 2 个人脸特征向量 S 和 M 在第 i 维的取值, $S \in \mathbf{R}^{dN_{\text{sum}}}$, $M \in \mathbf{R}^{dN_{\text{sum}}}$.

步骤 7 将步骤 6 中得到的待检测图像与测试库中所有图像进行相似性度量, 然后将相似性度量值从高到低进行排序. 当度量准则为 CHI 时, 使用升序排序, 当度量准则为 HI 时, 使用降序排序, 排序结果将作为检索结果.

2 实验结果与分析

2.1 实验测试

本文在 FERET 库下的 babjbk 文件夹中进行测试, 该库中含有 14 051 张多姿态、多光照的灰度人脸图像, 该文件夹中共有 600 张图像, 200 个人物, 每个人物有 3 张图像. babjbk 文件夹中的示例图像是 6 个人的代表图, 如图 2 所示. 从图 2 可以看出, 该文件夹下的图像仅包括人的面部区域, 人脸的角度无变化, 表情和光照有变化.



图 2 FERET 下 babjbk 文件夹中的示例图像

在带有背景图的测试中, 本文还利用了 FERET 库下 jpg_frontal 文件夹中的少量图像. 该文件夹中共有 3 368 张图像, 每个人物的图像数不等, 且拍摄情况不同, 面部表情、着装、发型、人脸在图像中所占比例均有所变化, 如图 3 所示.



图 3 jpg_frontal 文件夹中的示例图像

2.2 实验中检索性能的度量

在实验中, 首先对图像特征描述的距离进行了排序, 根据排序确定相关图像的位置, 再利用平均精度 (AP) 准则获取检索精度值 P_A . 所以有的图像 AP 均值定义为 P_{AM} , 用它可以确定分块方式的性能. P_{AM} 、 P_A 的表达式如下

$$P_{AM} = \frac{1}{n} \sum_{i=1}^n P_{Ai}$$
$$P_A = \left(\sum_{i=1}^n \left(\sum_{j=1}^i r_j / i \right) \right) / \sum_{i=1}^n r_i$$

式中: n 表示测试图中的图像数; $r_i \in \{0, 1\}$ 表示排序为 i 的图像与输入检测图像的相关性. 如果 2 个图像同为一人的, 则 2 个图像相关, $r_i = 1$; 反之, $r_i = 0$.

以 FERET 库下 babjbk 文件夹为例, 通过相似性度量及排序, 可以得到与测试图像相关的图像位置. 对于同一个人的 3 张图像位置, 若从前到后分别记为 r_1 、 r_2 和 r_3 , 则原图的位置为 r_1 且 $r_1 = 1$. 统计时, 先去除原图, 并将其余图像的位置往前移动, 这样整体的 AP 均值会大幅下降. 由于不同分块方式的 AP 均值差异较大, 所以利用它有利于比较各种分块方式的性能.

2.3 实验结果比较

本文采用 3 种 LBP 模式 (LBP^{org} 、 LBP^{u2} 、 LBP^{riu}), $(P_A, R) = (8, 1)$ 对 5×5 到 10×10 这 2 种分块数进行了 CHI 和 HI 相似性度量, 同时在 FERET 库的 jpg_frontal 文件夹中比较了本文的基于重要区域重采样 (IRB) 方法和现有的均匀分块 (NUB) 方法、均匀分块加权重 (WUB) 方法、均匀交叠 (UOB) 方法的性能, 讨论了重要区域重采样方法在分块数和交叠步长均发生变化时的性能.

实验对眉毛和眼睛以步长 hw/n 进行采样, 其中 $n > 1$, 而对其余部分以步长 hw 进行采样.

表 1 为不同 LBP 模式下 4 种分块方法的相似性比较.

方法	分块数	P_{AM}		
		LBP^{org}	LBP^{u2}	LBP^{riu}
NUB	8×8	0.211 7	0.209 1	0.212 7
WUB	8×8	0.261 7	0.266 8	0.254 5
UOB	15×15	0.193 2	0.193 3	0.191 9
IRB	152	0.291 2	0.294 1	0.285 4

从表 1 可以看出, IRB 方法的相似性比 NUB

方法提高了 8%，比 WUB 方法提高了 3%，比 UOB 方法提高了 10%。从复杂度看，IRB 方法介于 NUB 与 UOB 方法之间，可见 IRB 方法在复杂度增加不大的情况下，大大提高了性能。

图 4 为 $5\times 5\sim 10\times 10$ 分块和 LBP^{riu} 模式下的 4 种分块方法的性能比较。从图 4 可以看出，在 LBP^{riu} 模式下，IRB 方法均优于其他方法。

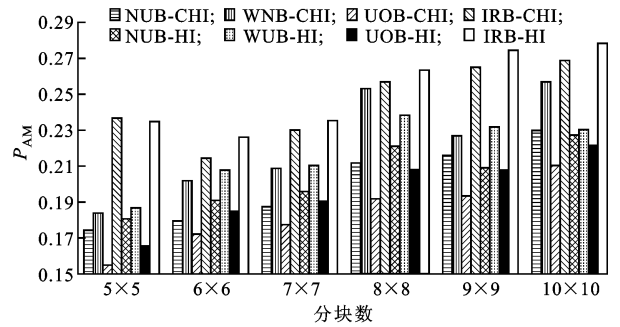


图 4 不同分块数下 4 种分块方法的性能比较

2.4 重要区域重采样方法的性能讨论

2.4.1 人脸检索性能与不同相似性度量准则的关系 采用 CHI 和 HI 的相似性度量准则在 8×8 分块的 LBP^{org} 模式下进行了 4 种分块方法的性能比较，结果如表 2 所示。从表 2 可以看出：①在各种相似性度量准则下，本文分块方法优于其他分块方法；②CHI 与 HI 相交的相似性度量结果的差异小于 2%，表明它们在评判分块的性能上区别不大。

表 2 2 种相似性度量准则下 4 种方法的性能对比

方法	P_{AM}	
	CHI	HI
NUB	0.211 7	0.228 1
WUB	0.261 7	0.275 8
UOB	0.193 2	0.208 6
IRB	0.291 2	0.309 0

2.4.2 带有背景图像的 4 种分块方法的性能比较

本文对包含背景的人脸图像进行了分块性能测试，即对于同一图像，针对全图和人脸部分分别进行了实验，测试库的文件夹为 jpg_frontal，模式为 8×8 分块的 LBP^{u2} ，采样步长为 $h\tau w/2$ ，结果如表 3 所示。从表 3 可以看出：人脸部分的分块性能比全图有所降低，原因是背景会影响区分度的性能；在人脸部位的分块性能上，本文方法的性能约下降 3%，其他方法的性能约下降 10%。这说明，在检索中本文方法能够提高重要区域的判别力。

表 3 带有背景图像的 4 种分块方法的性能对比

方法	分块数	P_{AM}	
		全图	脸部
NUB	8×8	0.481 6	0.371 9
WRB	8×8	0.492 8	0.390 3
IRB	86	0.491 2	0.465 9

2.4.3 分块数及重采样步长对检索性能的影响

在 $5\times 5\sim 10\times 10$ 分块以及交叠步长为 $h\tau w/1\sim h\tau w/5$ 的情况下，不同分块数及重采样步长对重要区域重采样方法性能的影响如表 4 所示。从表 4 可以看出：随着分块数的增加，检索性能非线性提高， 8×8 分块的性能相对于 5×5 分块明显提高， 10×10 分块的性能相对于 8×8 分块差异很小且复杂度大大提高；随着交叠次数的增加，检索性能非线性提高， $h\tau w/3$ 交叠步长的性能相对于 $h\tau w/2$ 提高得更明显， $h\tau w/5$ 交叠步长的性能相对于 $h\tau w/4$ 提高得甚少。

表 4 不同分块数及不同采样步长下的性能对比

分块数	P_{AM}				
	$h\tau w/1$	$h\tau w/2$	$h\tau w/3$	$h\tau w/4$	$h\tau w/5$
5×5	0.180 7	0.235 4	0.236 2	0.229 8	0.305 7
6×6	0.193 6	0.226 7	0.251 9	0.263 2	0.268 0
7×7	0.196 2	0.235 8	0.314 0	0.275 0	0.284 4
8×8	0.223 0	0.264 9	0.287 2	0.308 2	0.290 0
9×9	0.211 4	0.276 0	0.301 3	0.310 6	0.318 2
10×10	0.227 8	0.279 2	0.309 8	0.312 9	0.314 1

2.4.4 在人脸部位的不同区域进行重采样时采样步长对检索性能的影响

在人脸部位的不同区域进行重采样时，采样步长对检索性能的影响如表 5 所示。本文中，用 nm1~nm5 表示以 $h\tau w/1\sim h\tau w/5$ 步长对鼻子和嘴巴区域进行采样，用 eye1~eye5 表示以 $h\tau w/1\sim h\tau w/5$ 步长对眼睛和眉毛区域进行采样，分块数为 5×5 ，模式为 LBP^{riu} ，准则为 HI。

从表 5 可以看出，眼睛和眉毛区域的交叠性能优于鼻子和嘴巴区域。当面部表情发生较大变化时，嘴巴区域，特别是嘴角部分变化很大；当嘴巴区域的采样步长较大，而眼睛区域采样步长较小时，使用基于分块的检索方式可以减小面部不变或变化小的区域（如眼睛）对变化大的区域（如嘴角）的影响，但这种减小干扰的能力是有限的；当嘴角的采样步长减小时，干扰增加，从而检索性能下降。

表 5 不同区域重采样时采样步长对检索性能的影响

采样 方法	P_{AM}				
	nm1	nm2	nm3	nm4	nm5
eye1	0.180 7				
eye2	0.235 4	0.200 0			
eye3	0.256 4	0.228 6	0.218 4		
eye4	0.229 8	0.226 3	0.227 6	0.216 1	
eye5	0.303 7	0.224 0	0.233 8	0.231 2	0.227 1

3 结束语

本文提出了一种基于重要区域重采样的人脸检索方法. 该方法将面部器官,特别是眼睛眉毛区域定义为重要区域,通过对重要区域的采样来解决加权和重要信息离散化的问题. 通过大量实验对本文分块方法和传统的均匀分块、分块加权重、均匀交叠方法进行了对比,从而验证了本文方法的有效性.

参考文献:

[1] 温浩,郭崇慧. 利用粒子群优化的人脸特征提取识别算法[J]. 西安交通大学学报, 2008, 42(10): 1213-1217.
WEN Hao, GUO Conghui. Face recognition with features extraction based on particle swarm optimization [J]. Journal of Xi'an Jiaotong University, 2008, 42(10):1213-1217.

[2] SAMIR C, SRIVASTAVA A, DAOUDI M. Three-dimensional face recognition using shapes of facial curves [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(11):1858-1863.

[3] WU Zhong, KE Qifa, SUN Jian, et al. Scalable face image retrieval with identity-based quantization and multi-reference re-ranking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 33(10):1991-2001.

[4] HEISELE B, SERRE T, PONTIL M, et al. Component-based face detection [C] // Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA, USA: IEEE Computer Soc. , 2001:657-662.

[5] HEISELE B, HO P, WU J, et al. Face recognition; component-based versus global approaches[J]. Computer Vision and Image Understanding, 2003, 91(1/

2):6-21.

[6] AHONEN T, HADID A, PIETIKAINEN M. Face description with local binary patterns; application to face recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(12): 2037-2041.

[7] AHONEN T, HADID A, PIETIKAINEN M. Face recognition with local binary patterns[C]// Computer Vision — ECCV 2004. Berlin, Germany: Springer, 2004:469-481.

[8] ZHAO Guoying, PIETIKAINEN M. Dynamic texture recognition using local binary patterns with an application to facial expressions[J]. IEEE Pattern Analysis and Machine Intelligence, 2007, 29(6):915-928.

[9] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C]// Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA, USA: IEEE Computer Soc. , 2001:511-518.

[10] 许学斌,张德运,张新曼,等. 采用二代曲波变换和反向传播神经网络的人脸识别方法 [J]. 西安交通大学学报, 2010, 44(4):48-52.
XU Xuebin, ZHANG Deyun, ZHANG Xinman, et al. Method for face recognition using second-generation curvelet transform and back propagation neural network [J]. Journal of Xi'an Jiaotong University, 2010, 44(4):48-52.

[11] PIETIKAINEN M, OJALA T, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7):971-987.

[12] OJALA T, PIETIKAINEN M. Unsupervised texture segmentation using feature distributions [J]. Pattern Recognition, 1999,32: 477-486.

[13] LI Zisheng, IMAI J, KANEKO M. Robust face recognition using block-based bag of words [C]// Proceedings of the 2010 International Conference on Pattern Recognition. Los Alamitos, CA, USA: IEEE Computer Soc. , 2010:1285-1288.

[14] TURPIN A, SCHOLER F. User performance versus precision measures for simple search tasks [C]// Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA: ACM, 2006: 11-18.

(编辑 苗凌)