

人类面部属性估计研究:综述^{*}

曹 猛[†], 田 青[†], 马廷淮¹, 陈松灿²

¹(南京信息工程大学 计算机与软件学院, 江苏 南京 210044)

²(南京航空航天大学 计算机科学与技术学院, 江苏 南京 210016)

通讯作者: 田青, E-mail: tianqing@nuist.edu.cn



摘 要: 近年来,人脸属性估计因其广泛的应用而得到了大量的关注和研究,并且很多估计方法被提了出来.主要对现有相关工作进行归纳总结,为研究者提供相关参考.首先,根据是否考虑人脸性别、年龄、人种等不同属性间的内在关联,将现有的人脸面部属性研究方法划分成朴素的研究方法和自然的研究方法这两大类进行总结介绍.然后,从单一人脸数据库标记不完备、现有方法未能完备利用多属性联合估计、现有方法未能很好地利用各面部属性间关系这3个方面阐述当前方法的不足.最后,给出关于人脸面部属性估计进一步的研究方向.

关键词: 面部属性估计;人脸面部估计;年龄估计;性别判别;人种识别;多属性联合估计

中图法分类号: TP391

中文引用格式: 曹猛,田青,马廷淮,陈松灿.人类面部属性估计研究:综述.软件学报,2019,30(7):2188–2207. <http://www.jos.org.cn/1000-9825/5837.htm>

英文引用格式: Cao M, Tian Q, Ma TH, Chen SC. Human facial attributes estimation: A survey. Ruan Jian Xue Bao/Journal of Software, 2019, 30(7): 2188–2207 (in Chinese). <http://www.jos.org.cn/1000-9825/5837.htm>

Human Facial Attributes Estimation: A Survey

CAO Meng[†], TIAN Qing[†], MA Ting-Huai¹, CHEN Song-Can²

¹(School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China)

²(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Over the past decades, human facial attributes (e.g. gender, age, and race) estimation has received large amount of attention and research due to its potential applications, and variety of methods have been proposed to address it. This article is devoted to review related works comprehensively and give references for researchers. Firstly, in accordance with whether exploiting the potential correlations between these facial attributes, the existing approaches are classified into naïve and natural groups and they are reviewed within each group. Then, in terms of incompleteness of annotated labels, considered attributes, and correlations utilization, the drawbacks of existing methods are analyzed. Finally, future research directions are provided at the end of this work.

Key words: facial attributes estimation; human facial estimation; age estimation; gender recognition; race classification; multi-attributes joint estimation

近年来,在计算机视觉与模式识别等领域,人脸相关研究得到了越来越多的关注.这主要由于人脸拥有丰富

[†].共同第一作者.

^{*} 基金项目: 国家自然科学基金(61702273, 61472186, 61672281); 江苏省自然科学基金(BK20170956); 江苏省高校自然科学研究面上项目(17KJB520022)

Foundation item: National Natural Science Foundation of China (61702273, 61472186, 61672281); National Natural Science Foundation of Jiangsu Province, China (BK20170956); Natural Science Foundation of the Jiangsu Higher Education Institutions of China (17KJB520022)

收稿时间: 2018-08-08; 修改时间: 2018-12-27; 采用时间: 2019-03-16; jos 在线出版时间: 2019-04-10

CNKI 网络优先出版: 2019-04-09 17:32:30, <http://kns.cnki.net/kcms/detail/11.2560.TP.20190409.1732.007.html>

且稳定的特征信息,且易于被摄像头等仪器捕捉,进而被广泛应用于人机交互^[1]、智能消费推荐^[2,3]、访问控制^[4-6]和辅助身份识别^[1,7]等领域。

当前,人脸属性估计相关工作主要集中在人种估计、性别识别、年龄估计等几个方面。在相关研究方向上,分别已有大量工作被提出。例如,文献[8,9]利用集成学习进行人种分类估计;文献[10-25]借助贝叶斯分类器和SVM分类器等进行人脸性别识别;文献[26-61]将人类年龄估计分别视为分类问题、回归问题或者混合问题进行研究;文献[62]基于字典学习进行人脸老化合成研究等。然而,相关研究^[26,63-65]表明,不同面部属性之间存在关联性影响,并且某一属性估计受到其他属性的影响。因此,依据是否考虑人脸不同属性之间的内在关联和影响,本文将现有人脸面部属性研究工作划分为两大类,如图1所示:一是未考虑不同属性间关联影响的独立属性估计方法,这与朴素贝叶斯理论^[66]中不考虑属性间关联、独立对待不同属性的假设相类似,因此本文称其为朴素的(naïve)方法;二是考虑并利用不同属性之间关联影响的属性估计方法(例如,男女之间存在老化差异,该差异会对人类面部年龄估计准确性产生影响^[26]),在建模估计面部属性时,考虑面部其他属性对当前属性估计的影响,能够更好地迎合人脸属性演化的自然规律,因此,本文将该类方法称为自然的(natural)方法。

尽管目前已有大量的人脸属性估计相关工作被提出,然而研究者们通常基于 FG-NET^[6]、Morph^[3]、Adience^[67]、CACD 2000^[68]等单一数据集进行研究。而单一的相关数据集普遍存在属性标记不完备(如 FG-NET、CACD 2000 仅有年龄标签)、样本缺失(如 Morph 样本的年龄标记空间为 16 岁~77 岁,缺少 0~15 岁样本)、样本不均衡(如 Morph 含有 55 000 多张人脸样本,其中男性样本占 84.6%;黑种人与白种人样本占总体的 96.4%,黄种人和其他人种仅占 3.6%)等固有缺陷,这将严重影响属性估计器的识别准确性,并限制其泛化性能。因此,可基于多属性跨库标记以扩充训练数据集,并构建人脸年龄、性别和人种等多属性联合估计模型,探索属性之间和属性内在的关联,以实现更可靠的属性联合估计算法。

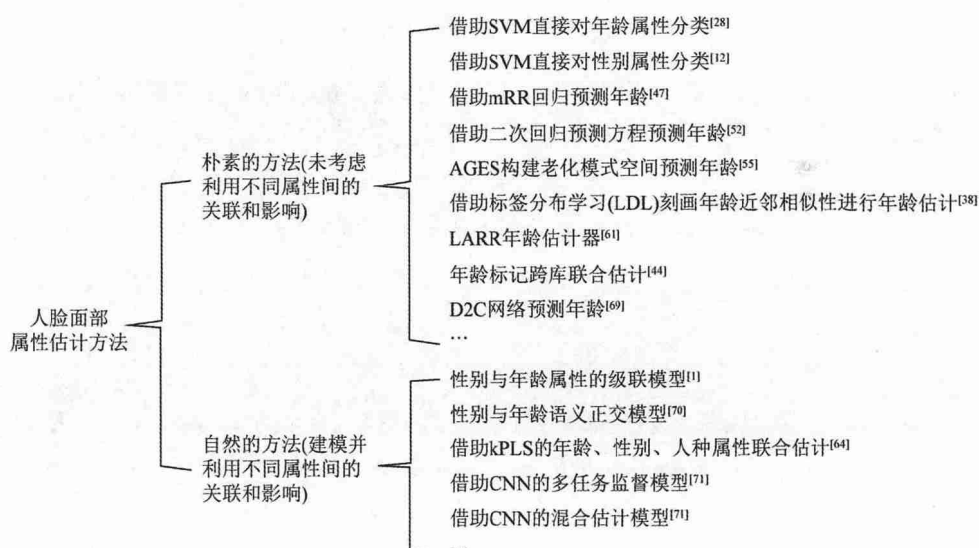


Fig.1 Research of naïve or natural human facial attributes estimation

图1 朴素的与自然的人脸面部属性估计方法

本文第1节总结并分析朴素的人脸年龄、性别、人种等面部属性估计相关工作。第2节回顾自然的人脸年龄、性别和人种等面部属性识别相关工作。第3节探索未来的面部属性估计研究方向。最后,对本文进行总结。

1 朴素的人脸面部属性识别

本节主要回顾朴素的人脸年龄、性别、人种等面部属性识别相关工作,即完全不考虑属性间关联影响的单

一属性识别.

1.1 人种分类

人种分类(race classification,简称 RC),顾名思义是对给定的人脸图像识别出其属于哪一类人种,例如白种人、黑种人、黄种人等人种类型,如图2所示.



Fig.2 Face image with corresponding race

图2 人脸图像及对应人种属性示意图

相对于性别和年龄属性,人脸人种估计相关工作较少,代表性工作如下.

(1) 文献[8,9],对于给定的人脸图像,先抽取 Haar-like 特征(简单来看,该特征反映了图像的灰度变化情况,能够描述简单的图形结构,对边缘与线段较为敏感),然后采用 Adaboost 集成分类器进行人种估计,并在 FERET 等人脸数据集上获得了较好的人种分类结果.

(2) 文献[72],首先提取灰度图像的人脸外观特征,然后使用多尺度描述算子提取人脸特征,并利用主成分分析(principal component analysis,简称 PCA)进行特征降维,最后借助线性判别分析(linear discriminant analysis,简称 LDA)实现测试人脸的人种分类.

1.2 性别识别

所谓性别识别(gender recognition,简称 GR),即对给定的人脸图像估计判断其为男性还是女性,如图3所示.



Fig.3 Face image with corresponding gender (upper row for female; lower row for male)

图3 人脸图像及对应性别属性示意图(上排为女性;下排为男性)

不同于 RC 的多分类问题,人脸性别估计为二值分类问题(非男即女).其代表性实现方法总结如下.

(1) 贝叶斯分类器(Bayesian).文献[10,11]先将给定人脸图像分割成多块互不重叠的网格块,然后通过学习出来的公共数据库 L 获取相似网络块表示,最后使用这些相似网络块进行后验概率估计,获得性别估计结果.

(2) 支持向量机(SVM)分类.文献[12-17]首先提取脸部特征,经过降维获取主要面部特征,最后使用 SVM 分类器得到性别识别结果.其中,文献[13,14]直接借助线性 SVM 分类器;为应对人脸样本的线性不可分问题,文献[12]采用高斯径向基函数(radial basis function,简称 RBF)核化的非线性 SVM 进行人脸性别分类;并且,为了更好地融合利用人脸不同部件的语义判别信息,文献[15-17]通过集成头部、面部、眼部和嘴(如图4所示)等不同部位的性别识别结果,以得到最终的性别识别结果.

Fig.4 Human facial part^[16] and starting from the left: Head, face, periocular and mouth图4 人脸不同部件示意图^[16](从左至右):头部、面部、眼部、嘴

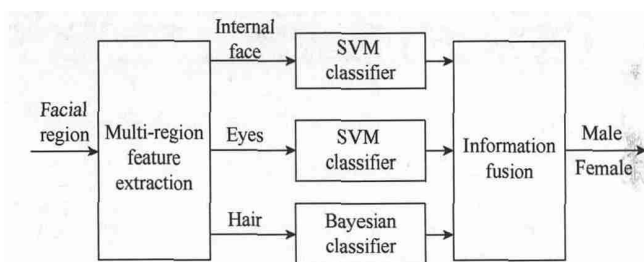
(3) 贝叶斯与 SVM 融合分类.为应对复杂的人脸获取环境并充分利用人脸区域的特有信息,文献[18]提出人脸多区域 SVM 性别识别融合的 GR.其算法流程如图 5 所示,通过贝叶斯准则(见公式(1))融合人脸不同区域的决策结果获取 GR.

$$D = \text{sgn} \left(w_0 + \sum_{i=1}^{N_i} (w_i \cdot d_i) \right) \quad (1)$$

其中, $\text{sgn}(x)$ 表示符号判别函数,当 $x > 0$ 时, $\text{sgn}(x) = 1$, 反之, $\text{sgn}(x) = -1$; $w_0 = \log \frac{P_1}{P_0}$, 其中, P_1 和 P_0 分别表示男性和女性的先验概率且 $P_1 + P_0 = 1$, d_i 表示人脸第 i 个区域性别判别的决策结果(用 ± 1 表示, -1 表示为男性, $+1$ 表示为女性), 权重 w_i 定义如下:

$$w_i = \delta(d_i - 1) \times \log \frac{1 - P_{FN_i}}{P_{FP_i}} + \delta(d_i + 1) \times \log \frac{1 - P_{FP_i}}{P_{FN_i}} \quad (2)$$

其中, P_{FP} 和 P_{FN} 分别表示分类器判别当前样本为女性的正确概率和错误概率, $\delta(x)$ 表示当 x 为 0 时, 函数值为 1, 反之为 0.

Fig.5 Illustration of gender classifier based on information fusion^[18]图5 混合贝叶斯与 SVM 性别判别方法流程图^[18]

(4) 决策树. 为挖掘利用三维人脸的空间分布信息, 文献[19]首先提取人脸的三维梯度和空间距离信息, 然后利用随机森林(random forest)法对提取的面部特征进行性别分类.

(5) 极速学习机. 文献[20]采用极速学习机(extreme learning machine, 简称 ELM)构建性别分类网络. 与传统的神经网络不同, ELM 不需要迭代地调整学习参数, 在设定好网络结构后, 仅需随机初始化输入层到隐藏层的权值, 隐藏层到输出层的权值即可直接解出, 进一步提高了人工神经网络的训练效率. 典型的 ELM 由输入层、隐藏层、输出层构成, 其网络输出可表示为

$$Y_{predict} = W_2 \sigma(W_1 X) \quad (3)$$

其中, W_1 表示输入向量到隐藏节点层的随机权重矩阵, σ 表示激活函数, W_2 为隐藏节点层到输出向量的权重矩阵, $Y_{predict}$ 表示网络的预测输出. 当 W_1 随机初始化后, W_2 可由最小二乘法计算得出(见公式(4)、公式(5)).

$$\min_{W_2} \|Y_{real} - Y_{predict}\|_F^2 \quad (4)$$

$$\Rightarrow W_2 = \sigma(W_1 X)^+ Y_{real} \quad (5)$$

其中, Y_{real} 表示期望输出, $\sigma(W, X)^+$ 表示矩阵的 Moore-Penrose 广义逆.

(6) 神经网络. 鉴于深度卷积学习(convolutional neural network, 简称 CNN)在计算机视觉等领域的成功应用, 研究人员开始将 CNN 用于性别识别之中. 为此, 文献[21–25]首先提取归一化的人脸样本集, 然后借助 CNN 提取人脸语义表示并采用“端-端(end-to-end)”学习方式训练性别分类器, 在 Morph 等人脸数据库上获得了更准确的 GR 结果. 其流程如图 6 所示.

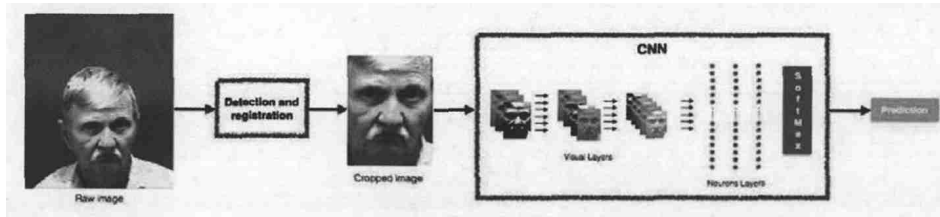


Fig.6 CNN-based human gender recognition^[22]

图 6 借助 CNN 的人脸性别判别模型流程示意图^[22]

1.3 年龄估计

所谓年龄估计(age estimation, 简称 AE), 就是根据给定的人脸图像估计其年龄值(或年龄范围), 如图 7 所示.

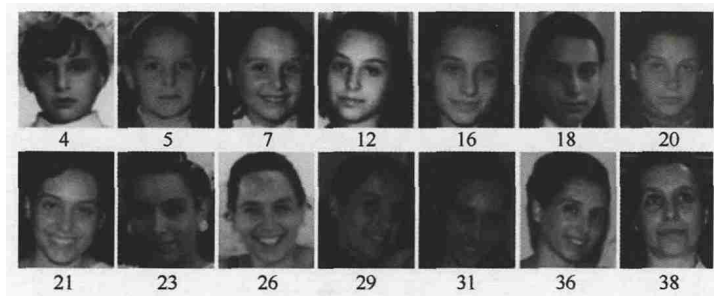


Fig.7 One person's face image (numeric below each image represents age)

图 7 同一人不同时期人脸图像以及对应年龄属性示意图(人脸图像下对应数字, 单位: 岁)

截止目前, 已有大量 AE 工作被提了出来. 依据这些工作的方法类型, 本文将其归纳为基于分类的 AE、基于回归的 AE、基于分类与回归混合的 AE 这 3 种类型.

1.3.1 基于分类的 AE

若将年龄(诸如 0 岁、5 岁、23 岁等)看成互不相关的离散类, 则 AE 可视为多分类问题.

年龄相关特征表示在 AE 中的利用. 文献[35,36]分别采用非负矩阵分解(构建图像局部信息)和独立成分分析(获取图像本质结构特征)的方法, 构造低维子空间获取人脸特征, 然后采用浅层神经网络作为年龄估计器进行 AE. 随后, 文献[26–34]将更丰富的特征描述方法应用至 AE 当中, 主要包括主动外观模型(active appearance model, 简称 AAM)、生物启发特征(bio-inspired feature, 简称 BIF)、局部二值模式(local binary pattern, 简称 LBP)等, 然后借助 SVM 进行年龄分类.

然而, 上述研究方法仅将年龄看作是互不相关的离散类, 并没有将年龄的有序信息(例如, 20 岁大于 10 岁, 却小于 25 岁)考虑其中.

考虑年龄有序性的 AE. 文献[37]提出了一种 LBP 特征的代价敏感版本(cost-sensitive local binary feature learning, 简称 CS-LBPL), 将代价敏感学习融入 LBP 特征当中, 然后使用超平面序列排序器(ordinal hyperplanes ranker, 简称 OHRanker)借助年龄的有序性进行 AE 分类估计. 与文献[37]相异, 文献[73]考虑直接在特征上利用年龄的有序性, 该文作者提出一种用于保持人脸图像局部流型结构与年龄间序列特性的特征提取方法, 具体模型见公式(6).

$$\left. \begin{aligned} \max J_1(y_1, y_2, \dots, y_M) &= \sum_{u=1}^M w_u^L y_u + \alpha \sum_{u=1}^M w_u^R y_u \\ \text{s.t. } y_u &\in \{0, 1\}, u=1, 2, \dots, M \\ \sum_{u=1}^M y_u &= d \end{aligned} \right\} \quad (6)$$

其中, M 表示原始图像特征维度, d 表示获取后的特征子空间维度, y_u 表示是否利用原始特征中的第 u 维特征. 在公式(6)中, 第 1 项用于获取人脸图像的局部流型结构信息(用来刻画 M 维原始特征各自的权重), 第 2 项用于保持年龄间的有序信息(利用 w_u 维护), 超参数 α 用于平衡两种信息之间的重要程度. 通过最大化这一模型, 可以获得一个较小维度的既保持序列信息又保有面部局部的流型结构, 同时也保持了人脸特征子空间的相对完整性, 最后通过该特征子空间, 借助 OHRanker 获取 AE 结果.

考虑年龄近邻相似性的 AE. 年龄属性不仅具有序列性, 同时还具有近邻相似性. 如图 7 所示, 对于同一个人的人脸图像, 18 岁时, 与其 16 岁和 20 岁时的面部较为相似. 为此, 文献[38]提出采用标记分布学习(label distribution learning, 简称 LDL)方法刻画年龄属性的近邻相似特性. 对于年龄标记, LDL 并不认为其为单一的标签, 而是将其转成一个标记分布向量(包含临近年龄的描述度), 并将标记分布向量作为训练集中的类别信息. 例如, 20 岁的年龄图像同样也可以作为标记为 21 岁的样本, 只是描述度相比描述 20 岁时要低, 即真实年龄的描述度最高. 当前样本辅助描述(标记)的年龄距离真实年龄越远, 描述度越低, 且所有描述度之和为 1. 文献[38]通过条件概率假设, 利用最大熵模型来构建映射关系, 采用 KL 散度作为衡量预测分布和真实分布距离的损失函数, 获得了比较好的效果. 然而, 这些年龄标签分布需要基于某些先验假设从训练样本中获得, 年龄近邻关系的描述必然不是十分准确的. 同时, 由于人类各阶段年龄变化速度不同, 各年龄阶段相似度描述也就相异(如图 8 所示), 文献[39]提出一种自适应版本的 LDL(adaptive label distribution learning, 简称 ALDL)来解决上述问题. 鉴于文献[38]使用固定的标记分布以及文献[39]对于不同样本集分布描述不灵活的缺陷, 文献[74]提出基于数据依赖的标记分布算法.

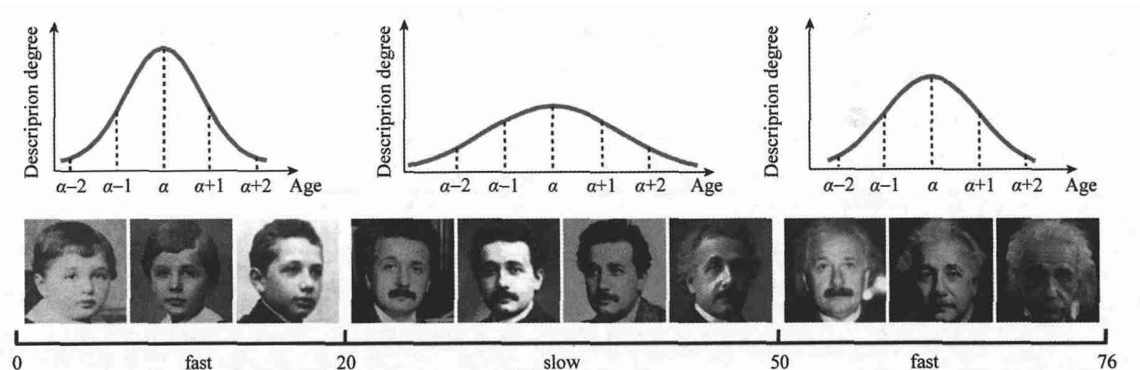


Fig.8 Facial appearance aging of Albert Einstein in different age ranges^[39]

图 8 爱因斯坦在不同年龄段的面部外表变化及其年龄标签分布示意图^[39]

同时利用年龄有序性和近邻相似性的 AE. 为建模利用年龄类的有序性和近邻相似性, 文献[42]利用马氏距离度量, 构建了如下度量目标函数, 以学习有序度量空间来实现 AE:

$$\left. \begin{aligned} \max J(A) &= \sum_{i,j} w(i,j) d_A^2(x_i, x_j) + \alpha \sum_l \varepsilon_l \\ \text{s.t. } d_A^2(x_i, x_j) + \varepsilon_l &= d_{ij}^2 \text{ if } \eta_{ij} = 1 \\ A &\geq 0 \\ \varepsilon &\geq 0 \end{aligned} \right\} \quad (7)$$

其中, $d_A(x_i, x_j) = \|x_{ij}\|_A = \sqrt{x_{ij}^T A x_{ij}}$ 刻画 x_i 和 x_j 之间的马氏距离, $w(i, j)$ 代表权重因子, 其借助两个样本标签的差异来加以表示, ϵ_i 表示松弛项, d_{ij} 表示两个样本之间的欧式距离. 在公式(7)中, 利用 $w(i, j)$ 惩罚每个输入和所有其他非相同标签输入之间的距离, 以保护年龄组之间的有序信息, 并通过约束项保证样本之间的近邻关系不发生改变. 然而, 文献[43]认为, 最大化加权距离的总和并不足以完全满足全局的序列关系, 该文作者借助马氏距离度量通过类似构造 LDA(类内距离尽量小, 类间距离尽量大)的方式, 构造有序度量特征空间, 以保证全局的序列关系, 最后通过 K 近邻分类器(K -nearest neighbor, 简称 KNN)对年龄进行类别估计.

跨库年龄类别互补的 AE. 为解决单一年龄库存在部分年龄样本不完备的问题(即某些年龄类别无对应训练样本), 文献[44]尝试跨越多个年龄数据库构建混合表示空间(如图 9 所示), 利用典型相关分析(canonical correlation analysis, 简称 CCA)合并样本空间, 并在所得样本表示空间通过样本联合表示构建年龄估计器, 并实验验证了其优越性.

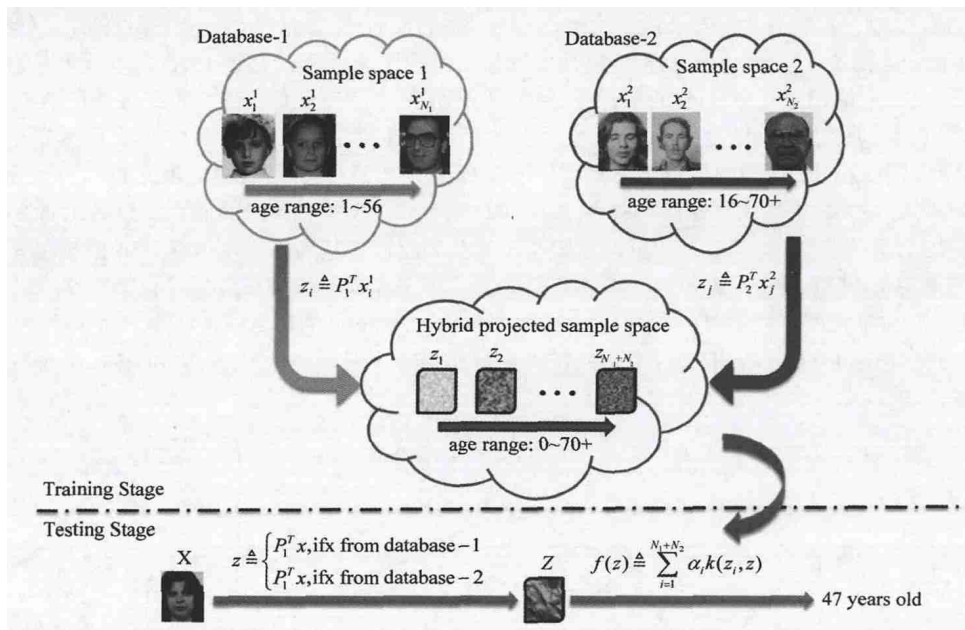


Fig.9 Illustration of the co-representation based cross-heterogeneous-database age estimation^[44]

图 9 跨年龄数据库的年龄估计流程示意图^[44]

借助深度神经网络的 AE. 近年来, 深度学习在计算机视觉、模式识别等领域得到大量成功的应用. 相较于以上传统方法将特征表示(feature representation)与目标决策(decision making)分离学习的模式, 深度学习通常以 end-to-end 学习方式将二者融合到一起同时学习, 并可在特定任务上取得更优的性能. 受此启发, 文献[75~78]等借助 CNN 或 VGG 网络抽取图像特征, 然后采用 Softmax 函数等分类方法预测年龄实现 AE. 其中, 文献[75, 76]采用较深的 VGG-16 网络结构以获取更高维的特征表示. 为防止训练过拟合, 同时也为了利用年龄属性间的近邻关系, 文献[40, 41]借助高斯分布生成年龄分布标记, 并采用 IMDB-WIKI 数据集微调网络以获取更好的泛化能力.

1.3.2 基于回归的 AE

基于分类的 AE 方法将不同的年龄模式视作相互独立的普通的多类问题进行处理, 然而人脸面部老化是一个连续和渐变的过程. 因此, 将 AE 视作连续的回归问题加以处理似乎更为合理. 为此, 文献[13, 45~51]尝试从人脸图像中提取不同的特征表示, 然后借助回归器(如支持向量回归, SVR)进行年龄回归预测. 其中, 文献[45]采用先提取人脸不同区域的特征, 然后借助贝叶斯准则融合特征以获取分层的人脸特征表示, 如图 10 所示.

文献[47,49,51]采用多元岭回归模型(multi-ridge regression,简称 mRR)来拟合年龄回归函数.mRR 优化目标如下所示:

$$\min_W \sum \|y_i - (W^T x_i + b)\|_F^2 + \lambda \|W\|_F^2 \quad (8)$$

其中, y_i 表示期望输出, W 表示投影矩阵, b 表示偏置项.在公式(8)中,第1项用于估计拟合损失,第2项用于控制模型复杂度,当 λ 取 0 时,模型退化为普通的最小二乘线性回归问题.然而,文献[52-54]认为,相对于线性回归模型,二次回归模型具有更好的年龄估计拟合能力,并通过如下优化目标实现对年龄估计器 $f(x_i)$ 的学习:

$$\min \sum \|y_i - f(x_i)\|_2^2 \quad (9)$$

其中, $f(x_i) = \alpha_0 + \alpha_1^T x_i + \alpha_2^T x_i^2$ 表示二次年龄回归模型(α_i 表示回归系数), y_i 表示真实年龄标签.
















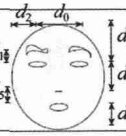





	Image	Topology	Geometry	Photometry			Configuration
Global appearance		[0,0,1,0,0]					—
		—					—
Components		[0,0,0,1,0]					
Skin zones		[1]					—

Fig.10 Four types of features at three levels for age estimation^[45]

图 10 在 3 种不同等级之下提取的 4 种特征^[45]

文献[55]认为:(1) 人脸老化模式不可控;(2) 存在个人因素;(3) 人脸老化模式是随时间而变化的.基于以上考虑,文献[55]将人脸按时间顺序排列构建年龄模式空间(aging pattern subspace,简称 AGES)(如图 11 所示).鉴于所用数据库中每个个体年龄标记存在部分缺失的问题,该文作者提出一种编码再解码的方式重建缺失的年龄模式,并构建年龄模式空间进行 AE.具体来说,使用相同年龄图像均值初始化缺失部分 x^l ,并与已存在部分 x^a 用 PCA 降维后,再重新变换回原始空间,利用重建后的 x_{new}^a 与 x^a 进行误差估计;反复上述过程,直至重建误差低于预定值.

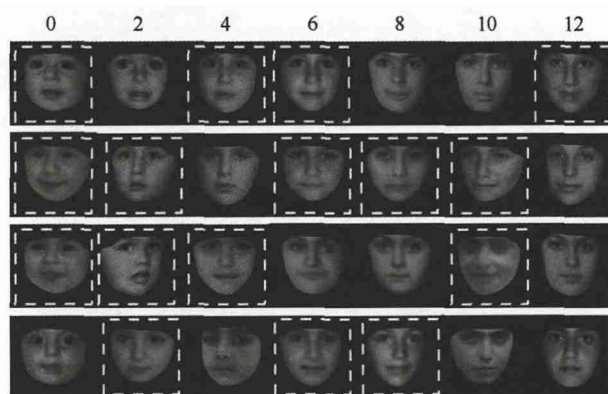


Fig.11 Illustration of aging pattern subspace^[55]

图 11 年龄模式空间示意图^[55]

文献[56]为了保证局部语义相似,借助马氏距离度量,重新调整同一邻域内样本对的距离,同时扩大邻域外部的样本距离,构建空间度量模型 mkNN,并将其用于 AE 度量.mkNN 目标函数如下:

$$\left. \begin{aligned} \max_A \quad & \sum_{i,j} d_A(x_i, x_j) - \lambda \sum_{i,j} \varepsilon_{ij} \\ \text{s.t.} \quad & d_A(x_i, x_j) + \varepsilon_{ij} = d_{ij}^2 \text{ if } \eta_{ij} = 1 \\ & A \geq 0 \\ & \varepsilon_{ij} \geq 0 \end{aligned} \right\} \quad (10)$$

其中, $d_A(x_i, x_j)$ 表示马氏距离, λ 为非负超参数, ε_{ij} 为松弛变量, $\eta_{ij} = 1$ 表示样本 x_i 与 x_j 是近邻样本, $d_{ij} = \left(\frac{L(x_i, x_j) + \gamma}{C - L(x_i, x_j)} \right)^p \times d(x_i, x_j)$ 用于刻画样本的近邻关系. $L(x_i, x_j)$ 表示两个样本标签之间差异的绝对值, γ 由标签噪声决定, p 为预设值, $C = \max \{L(x_i, x_j)\} + \theta, \theta > 0$ 以确保分母不为 0, 而 $d(x_i, x_j)$ 表示样本之间的欧式距离. 获取有序度量空间后, 其借助简单的回归 KNN 策略, 实现人脸年龄模式的回归预测.

文献[47]为利用年龄类的有序性和近邻相似性, 将累积属性编码(cumulative attribute coding, 简称 CA)融入 mRR 模型当中, 保证人脸的累积变化特性. 其中, CA 编码的编码长度与当前所有样本的最大年龄相关(如, 最大年龄为 100 岁, 编码长度为 101), 具体表示为

$$l_i = \begin{cases} 1, & l_i \leq y \\ 0, & l_i > y \end{cases} \quad (11)$$

其中, y 表示样本年龄, l_i 表示第 i 位编码. 为利用 CA 编码间的关系, 文献[49]借助差分运算将 CA 编码的 0 阶关系与 1 阶关系融入年龄估计模型, 获得了性能更优的 AE 结果.

在深度学习方面, 文献[79]在 CNN 网络中融入有序回归思想, 在 Morph II 数据集上取得较好的 AE 结果. 文献[71]借助同一网络模型主体架构, 对比了 5 种方案在 AE 问题上的识别性能, 发现基于拉普拉斯损失的回归型方法略优于其他对比方法. 这 5 种方案分别为基于交叉熵损失的分类型方法(分别采用 One-Hot 编码和 LDL 编码两种标号编码方式)、基于最小平方损失和基于拉普拉斯损失的回归型方法以及基于有序损失的回归型方法. 与此同时, 文献[69]在 CNN 网络中借助 CA 编码的累积特性以解决训练样本类分布不均衡问题, 提出采用对比信号来刻画人脸有序性以增强特征表示的判别能力, 并构建 D2C 网络实现 AE. D2C 网络结构如图 12 所示.

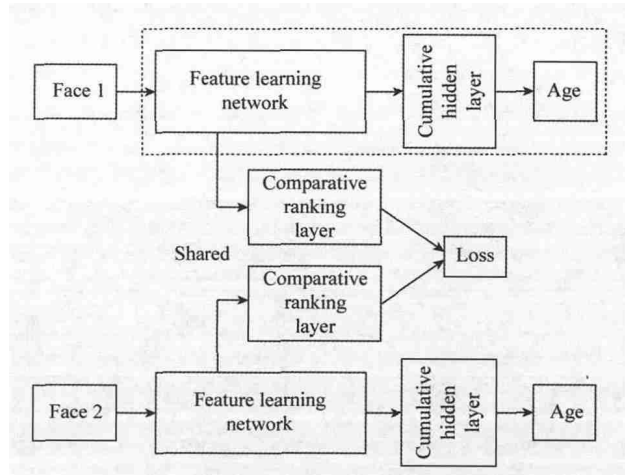


Fig.12 Illustration of D2C network architecture

图 12 D2C 网络基本结构示意图

其中, CRL(comparative ranking layer)用于学习样本的排列函数 f (即若输入样本 x_i 年长于 x_j , 则 $f(\varphi_i) > f(\varphi_j)$, φ_i 与 φ_j 分别表示样本 i, j 的特征向量). D2C 网络整体损失函数定义为

$$Loss_{ij} = \sum_{m=i,j} loss_m^{age} + \alpha \sum_{m=i,j} loss_m^{CHL} + \beta loss_{ij}^{Rank} \quad (12)$$

其中, α 与 β 为超参数, 第 1 项对应年龄预测的损失项, 第 2 项对应 CA 编码损失, 第 3 项辅助约束样本排序. 文献[69]认为, 年龄的有序性比单一的年龄标签信息更加稳定. 更为重要的是, 通过 CRL 构建样本的序列关系可实现重复利用训练样本, 以减缓样本类不平衡的状况.

1.3.3 基于分类与回归混合的 AE

为了获得性能更优的年龄估计器, 文献[48, 57–61]采用分类和回归相结合的策略实现 AE. 其中, 文献[61]认为, SVR 回归预测年龄不精确, 仅能预测全局的老化趋势(如图 13(a)所示). 为此, 该文提出借助局部调整策略来构建 LARR(locally adjusted robust regressor)估计器, 并利用其在测试人脸样本上进行年龄识别. 具体而言, 采用局部偏移(如图 13(b)所示)策略, 使得预测年龄 $y_{predict}$ 更接近真实年龄 y_{real} (即期望 $y_{real} \in [y_{predict}-d, y_{predict}+d]$); 为了获取每类偏移方向, 文献[61]还采用二叉树构造多分类 SVM 进行最后的决策. 之后, 很多对 LARR 的改进工作被提了出来, 如文献[60]将流型学习融入 LARR; 文献[59]对 LARR 进行了概率化扩展, 并在 UIUC-IFP-Y 等数据集上取得了比 LARR 更优的 AE 识别结果. 在深度学习方面, 文献[57]采用 CNN 和 LDL 分别对人脸图像进行特征表示与年龄编码, 通过集成两套网络(一套基于分类、一套基于回归)的预测结果实现 AE.

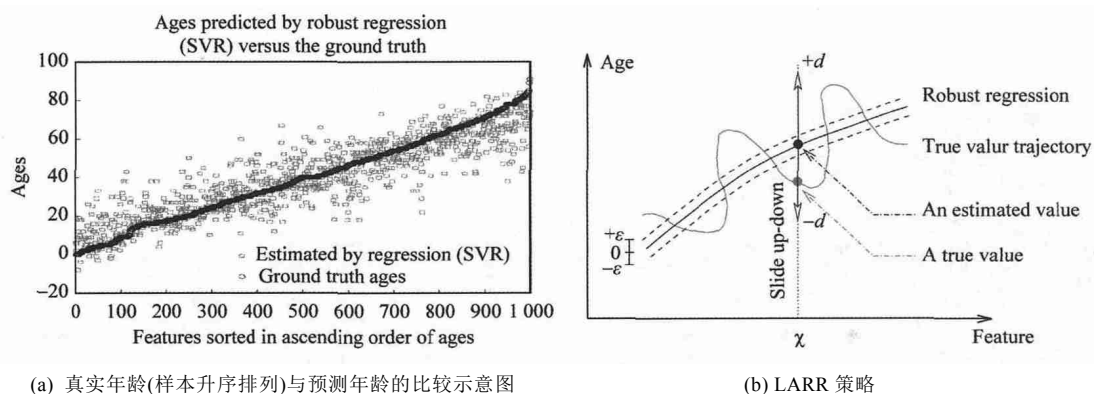


Fig.13 Illustration of age estimation with LARR^[61]

图 13 借助 LARR 策略拟合真实年龄示意图^[61]

1.4 性能对比

本节通过列表对比形式对有代表性的人脸单一属性估计工作给出总结性性能对比. 具体见表 1.

Table 1 Summary and comparison of naïve-group human facial attributes estimation techniques on different databases

表 1 在不同数据集上的人脸单一属性估计方法的对比与总结

方法	数据集	特征表示或图像输入	应对任务	模型性能		
				预测器	MAE	Acc
Wu 等人, 2004 ^[9]	FERET	Haar-like	RC	Adaboost (C)	N/A	95.00% (Mongoloid) 96.10% (Caucasian) 93.50% (African)
Lu 等人, 2004 ^[72]	NLPR	多尺度灰度图	RC	LDA (C)	N/A	96.00% (Asian) 96.60% (Non-Asian)
Aghajanian 等人, 2009 ^[11]	Web	原始图像分割	GR	贝叶斯分类器(C)	N/A	89.00%
Patel 等人, 2016 ^[13]	LFW	CoLBP	GR	SVM (C)	N/A	95.37% (M) 73.32% (F) 89.70%
Dantcheva 等人, 2016 ^[12]	FERET	融合特征 (Dynamic, Appearance)	GR	SVM+RBF 核(C)	N/A	99.14% (M) 86.24% (F)
Lin 等人, 2016 ^[18]	[18]	[18]	GR	贝叶斯与 SVM 融合(C)	N/A	91.10% (M) 98.60% (F)
Liao 等人, 2016 ^[19]	Texas 3D	人脸特征点距离	GR	随机森林(C)	N/A	90.18% (M) 88.42% (F)

Table 1 Summary and comparison of naïve-group human facial attributes estimation techniques on different databases (Continued)

表 1 在不同数据集上的人脸单一属性估计方法的对比与总结(续)

方法	数据集	特征表示或图像输入	应对任务	模型性能		
				预测器	MAE	Acc
Coco 等人,2016 ^[22]	Morph	原始图像	GR	Softmax (C)	N/A	96.80% (M) 95.70% (F) 96.30%
Mahmood 等人,2017 ^[20]	LEW	LBP	GR	R-ELM (C)	N/A	89.92%
Geng 等人,2006 ^[55]	FG-NET	AGES	AE	(R)	6.77	N/A
Guo 等人,2008 ^[61]	FG-NET U-I-Y/M U-I-Y/F	灰度图 60×60	AE	LARR (H)	5.07	N/A
					5.38	
					5.25	
Guo 等人,2008 ^[58]	FG-NET U-I-Y/M U-I-Y/F	灰度图 60×60	AE	PFA (H)	4.97	N/A
					5.12	
					5.11	
Guo 等人,2009 ^[33]	YGA/M YGA/F	BIF	AE	SVM (C)	3.47	N/A
					3.91	
Liang 等人,2011 ^[51]	FG-NET	GLOH	AE	RR (R)	5.45	N/A
Li 等人,2012 ^[42]	FG-NET	AAM	AE	kNN (C)	4.51	N/A
Lu 等人,2012 ^[73]	FG-NET	Gabor	AE	OHRanker (C)	4.82	N/A
Chen 等人,2013 ^[47]	FG-NET Morph	AAM	AE	SVR+RBF 核(R)	4.67	N/A
					5.88	
Lu 等人,2013 ^[54]	Morph	灰度图	AE	(R)	5.24 (W) 4.18 (B)	N/A
Geng 等人,2014 ^[39]	Morph	BIF	AE	(C)	4.34	N/A
Lu 等人,2015 ^[36]	FG-NET	CS-LBFL	AE	OHRanker (C)	4.43	N/A
Tian 等人,2015 ^[49]	FG-NET Morph	AAM	AE	SVR+RBF 核(R)	4.37	N/A
					4.69	
Liu 等人,2015 ^[57]	Apparent age	[57]	AE	(H)	3.33 (V)	N/A
Rothe 等人,2015 ^[75]	IMDB-WIKI	分辨率 256×256	AE	Softmax (C)	3.22	N/A
Huo 等人,2016 ^[40]	IMDB-WIKI	分辨率 224×224	AE	Softmax (C)	3.12	N/A
Tian 等人,2016 ^[44]	FG-NET Morph	AAM BIF	AE	(C)	<5.85 <5.80	N/A
Niu 等人,2016 ^[79]	Morph AFAD	分辨率 60×60	AE	(R)	3.27 3.34	N/A
Xing 等人,2017 ^[71]	Morph WebFace	分辨率 227×227	AE	(R)	3.40 6.12	N/A
Li 等人,2017 ^[69]	Morph WebFace	分辨率 227×227	AE	RR (R)	3.16 6.12	N/A

注:RC 表示人种估计,GR 表示性别估计,AE 表示年龄估计.MAE(mean asolute error, $MAE = \frac{1}{N} \sum_{i=1}^N |y_i^r - y_i^p|$.其中, N 表示测试样本数目, y_i^r 表示真实年龄标签, y_i^p 表示预测年龄值.可见,MAE 越小,预测器性能越好)表示平均绝对误差,Acc(uracy)表示预测精度.C(classification)表示基于分类的方法,R(egression)表示基于回归的方法,H(ybrid)表示基于分类和回归的混合方法.M(ale)表示男性,F(emale)表示女性.N/A 表示未获取数据.W(hite)表示白种人,B(lack)表示黑种人.文献[57]所提 AgeNet 网络未能获取在测试集上 AE 性能的 MAE 结果,表中数据为在其验证集上结果,其最终 AE 性能采用 MNE(mean normalized error, $MNE = \frac{1}{N} \sum_{i=1}^N \left[1 - \exp \left(\frac{-(y_i^r - y_i^p)^2}{2\sigma_i^2} \right) \right]$.其中, σ_i 表示预测第 i 张人脸图像的标准差,其余符号定义同 MAE.可见,MNE 越小,预测器性能越好)表示,结果约为 0.27,位居 ICCV 2015 LAP (looking at people)竞赛年龄估计任务第 2 名

2 自然的人脸面部属性识别

上述朴素的人脸年龄、性别和人种等面部属性识别方法未考虑人脸年龄、性别和人种等不同属性间的潜在关联和影响,而是将它们视作独立的问题分别加以考虑.相关研究^[26,63-65]表明,不同面部属性之间存在关联性影响,并且某一属性估计受到其他属性的影响.为了与上述朴素的人脸年龄、性别和人种等面部属性识别相区别,我们称此类考虑属性间关系的识别方法为自然的人脸年龄、性别和人种等面部属性识别.

2.1 考虑性别、人种属性对AE的影响

为探究性别属性是否对 AE 产生影响,文献[31]分别在按性别拆分的样本集与在相同样本集混合情况(即性别当成未知)下,借助 SVM 进行 AE.通过比较实验结果可知,性别属性显著地影响 AE 的准确率.文献[26]采用相同的方式在 Morph II 数据集上(划分见表 2)研究性别、人种属性对 AE 是否存在关系.

鉴于同一年龄层内,男性和女性的人脸在年龄特征上的差异较大,文献[1]首先基于提取的人脸 Gabor 特征进行 GR,然后依据性别分类结果再借助 SVM 进行 AE 以提高 AE 的性能,具体流程图如图 14 所示.

Table 2 Distribution on Morph Ablum II database^[26]

表 2 Morph II 数据集使用情况^[26]

	Female	Male	Total
White	2 570	7 960	10 530
Black	2 570	7 960	10 530
White and Black	5 140	15 920	21 060

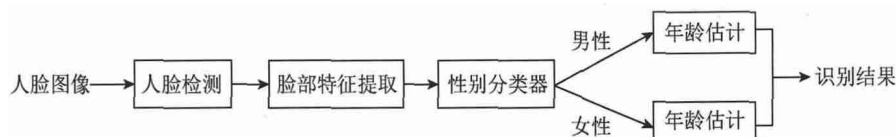


Fig. 14 Cascaded classification of gender and age^[1]

图 14 性别识别与年龄估计的级联模型^[1]

与文献[55]相异,为构建人脸老化空间,同时考虑到性别间的差异与不同年龄段人脸相似性差异不同,文献[62]划分年龄段,并将同一人的训练样本填充至相邻年龄段中构建老化空间.鉴于当前数据集的固有缺陷,同一个人的训练样本不可能完全填充至每一个年龄段中,文献[62]在保证至少一组相邻年龄含有同一人样本的情况下,借助耦合字典学习构建空间,其学习目标函数如下:

$$\min_{\{D^g, A^g, P^g\}_{g=1}^{G-1}} \left\{ \sum_{g=1}^{G-1} \left\{ \|X^g - H^g D^g A^g - P^g\|_F^2 + \gamma \|P^g\|_F^2 + \|Y^g - H^{g+1} D^{g+1} A^g - P^g\|_F^2 + \lambda \|A^g\| \right\} \right\} \quad (13)$$

$$\text{s.t. } \|D^g(:, d)\|_2 \leq 1, \|D^{g+1}(:, d)\|_2 \leq 1, \forall d \in \{1, \dots, k\}, \forall g \in \{1, \dots, G-1\}$$

其中, X^g 表示第 g 个年龄段中所有人的样本图像, Y^g 表示在 $g+1$ 个年龄段中对应人的样本图像, $D^g(:, d)$ 表示第 g 个年龄段第 d 幅图像特征.在公式(13)中,第 1 项用于度量第 g 个年龄段的损失,第 3 项用于度量第 $g+1$ 个年龄段的损失,第 2 项用于保持人脸的个性细节,第 4 项为稀疏项;通过第 1 项与第 3 项的级联,可以利用人脸对之间存在的关系促进人脸老化关系的学习.然而,文献[62]为获取结果需要依赖中间层输出,中间层结果存在误差,必然导致实验效果降低.为此,文献[80]将老化字典学习模型层次化,以获得更优的结果.

考虑到性别、人种属性对 AE 的影响,文献[81]借助 LDA 的思想将来自不同性别和人种的同一年龄类的年龄模式相融合进行研究.其主要思想为:同年龄类内散度小,同人种不同类间尽量分离,同类不同人种之间尽可能接近(如图 15 所示).经过映射后,来自不同群体的老化模式被投射到共同空间中,在该空间中不同群体的老化模式可以产生相似或更接近的分布.

$$w^* = \arg \max_w \frac{w^T (S_{P,b} + S_{Q,b}) w + \beta \Phi_{P,Q}}{w^T (S_{P,t} + S_{Q,t}) w} \quad (14)$$

其中,若不考虑 $\Phi_{P,Q}$ 项,该模型即为 LDA 模型($S_t = S_b + S_w$)用于描述不同人种内部的约束关系之间; $\Phi_{P,Q} =$

$\sum_{k=1}^c (\theta_P^{(k)})(\theta_Q^{(k)})^T$ 用于刻画不同人种之间,相同类别的距离关系($\theta_P^{(k)}$ 表示 P 人种内第 k 类样本均值).文献[82]考虑到性别与人种的交叉影响,并利用年龄的有序性,借助深度网络模型,提出一种保持序列关系的代价敏感网络模型进行 AE.

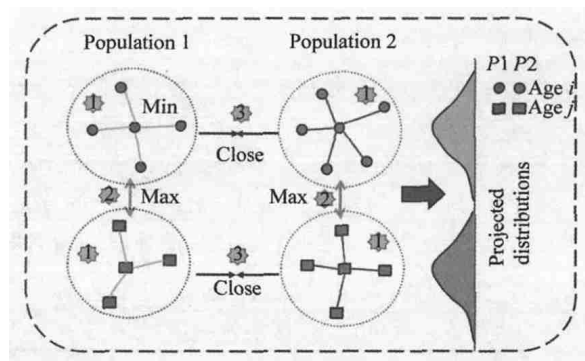


Fig.15 Three measures are considered in Ref.[81]: (1) Within-class variation; (2) Between-class scatter; (3) Between-population correlation

图 15 构建空间的 3 个度量:(1) 类内散度;(2) 类间散度;(3) 种群间相关性^[81]

2.2 考虑年龄属性分别对RC、GR的影响

为探究年龄等属性对 RC 的影响,文献[65]在 Morph II 图像集上首先进行年龄和性别属性的人群划分(同文献[26]),然后针对每类人群进行 RC,并发现 RC 准确率在不同人群间存在较大差异.

为探究年龄等属性对 GR 的影响,文献[83]对 YGA 等人脸集进行“老”“中”“青”这 3 类人群年龄组的分类,然后针对每个分组进行 GR;文献[84]首先将人脸集划分为“未成年”和“成年”两类,然后针对每个子类进行 GR,并发现后者的 GR 准确率高于前者.

2.3 考虑性别属性和人种属性对GR和RC的相互影响

针对人脸面部属性关系的建模利用问题,同时考虑性别属性与人种属性相互关系的有关研究十分稀少.为了将性别和人种两种属性间内在关系融入到 GR 和 RC 中,文献[85]借助 SVM 分别对不同模态的 3D 图像进行预测,最后进行决策融合以实现性别与人种的联合估计.尽管文献[86]借助架构较浅的神经网络模型进行性别与人种的联合估计,却获得了较好的识别性能,其具体网络结构如图 16 所示.整体网络除输入输出层外,仅包含 4 层卷积层,前 3 层卷积层之间由 2×2 大小的最大池化层连接进行特征选择,而第 3 层卷积层的输出采用 ReLU 函数激活送往第 4 层卷积层,最后借助 Softmax 函数实现联合分类估计.

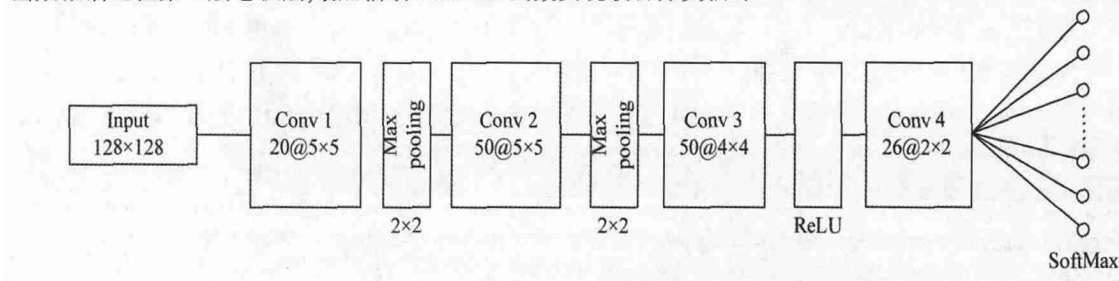


图 16 融合性别与人种属性的 VGG 网络结构

2.4 考虑年龄属性和性别属性对AE和GR的相互影响

为了将年龄属性和性别属性两者内在关系融入到 AE 和 GR 中,文献[87–89]借助深度卷积网络(CNN)获取性别和年龄相关的语义特征表示,再基于年龄和性别联合输出编码实现 AE 和 GR.

在传统方法中,文献[70,90]认为:(1) 文献[1]提出的性别识别与年龄估计的级联模型不能避免性别分类器对整个模型带来的误差累计效应(即初始性别分类错误,造成年龄估计误差增大);(2) 文献[63]虽然联合了人脸多种属性,但其标签连接不仅可能混淆性别和年龄之间的语义关系,还忽略了男性和女性之间的老化差异.为此,文献[76]提出一种语义空间中近邻正交的方案,对性别与年龄属性内在关系联合建模,同时研究 AE 和 GR,

其基本模型如下:

$$\min_{w_g, w_a} L_g(w_g; X, Y_g) + \frac{\lambda_1}{2} \|w_g\|_F^2 + L_a(w_a; X, Y_a) + \frac{\lambda_2}{2} \|w_a\|_F^2 + \frac{\lambda_3}{2} R_{NOSSpaces} \quad (15)$$

其中,第1项表示性别分类损失项,第3项表示有序回归方法下的年龄估计损失项,第2项、第4项用于控制模型复杂度,第5项定义为 $R_{NOSSpaces} = (w_a^T w_g)^2$,期望在语义空间中,不同性别间样本尽可能地分离。

2.5 同时考虑年龄、性别和人种三者属性之间的内在关联

偏最小二乘(partial least square,简称PLS,见公式(16))与典型相关分析(CCA,见公式(17))可以在输入样本集 X 与输出集 Y 之间建模,获取两者最大相关的子成分,其中, w_x 与 w_y 表示投影向量.因此,文献[63,64,91]借助PLS与CCA及其变种,将年龄、性别、人种整合为三元集标签,采用回归策略在样本空间与联合属性间建立关系,实现AE、GR和RC的联合估计。

$$\left. \begin{aligned} \max \quad & \text{cov}(w_x^T X, w_y^T Y) \\ \text{s.t.} \quad & |w_x| = |w_y| = 1 \end{aligned} \right\} \quad (16)$$

$$\left. \begin{aligned} \max \quad & w_x^T X Y^T w_y \\ \text{s.t.} \quad & w_x^T X X^T w_x = 1 \\ & w_y^T Y Y^T w_y = 1 \end{aligned} \right\} \quad (17)$$

为了同时利用三者属性信息,文献[71,92]借助深度模型CNN实现对年龄、性别和人种这三者属性的同时估计.文献[71]认为,AE、GR、RC这3种属性估计任务难度并不相同,因此,各任务所需特征表示的精细程度相异.该文通过共享底层以实现三者属性之间的公共特征表示,然后为各属性分别构建网络以满足不同的特征表示需求.网络基本结构如图17所示,损失函数定义为

$$L(X) = L^a + \alpha L^g + \beta L^r \quad (18)$$

其中, L^a 、 L^g 与 L^r 分别表示年龄估计损失、性别分类估计损失、人种分类估计损失, α 与 β 为非负超参数,用于平衡3种任务间的重要程度.然而,对于AE,此框架仅利用了性别与人种信息,没有兼顾到任务之间的内在关系.尽管此后,文献[71]提出一种混合模型以利用任务之间的内在关联,并在Morph II数据集上获得MAE结果首次低于3的优异成绩,但其本质与文献[1]提出的级联模型相同,仍然存在分类器分类错误带来的累计误差等问题,还有提升空间。

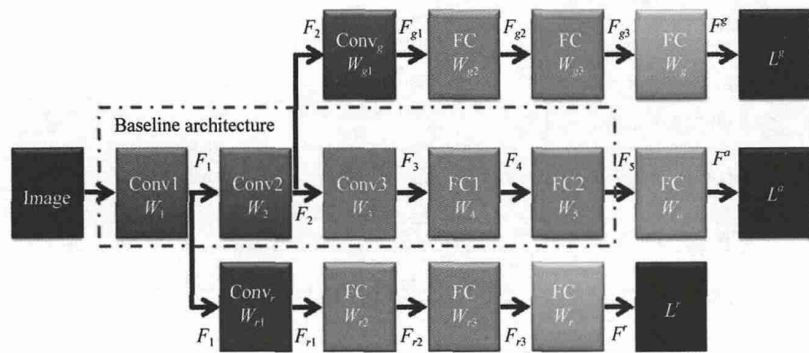


Fig.17 Deeply supervised multi-task learning architecture^[71]

图17 多任务深度监督学习框架^[71]

2.6 性能对比

本节通过列表对比形式对有代表性的人脸多属性联合估计工作给出总结性性能对比.具体可见表3.

Table 3 Summary and comparison of natural-group human facial attributes estimation techniques on different databases

表 3 在不同数据集上的人脸多属性联合估计方法的对比与总结

方法	数据集	特征表示或图像输入	应对任务	模型性能		
				预测器	MAE	Acc
Lu 等人, 2010 ^[1]	[1] (年龄组跨度:10)	Gabor	GR⇒AE	SVM (C)	N/A	85.98% (AE/M) 83.84% (AE/F)
Shu 等人, 2015 ^[62]	FG-NET (男女共分 18 类)	[62]	GR⇒AE	(C)	N/A	91.47% (AE)
Guo 等人, 2010 ^[26]	Morph	BIF	GR、RC ⇒AE	SVM (C)	4.33(AE)	98.90% (GR) 97.80% (RC)
Guo 等人, 2014 ^[81]	Morph	BIF	GR、RC ⇒AE	CpDA (C)	7.73(BM→BF) 8.73(BM→WF) 6.70(WM→BF) 7.67(WM→WF)	N/A
Li 等人, 2018 ^[82]	Morph	分辨率 227×227	GR、RC ⇒AE	DCP (C)	3.75(BM→BF) 3.18(BM→WF) 3.90(WM→BF) 3.13(WM→WF)	N/A
Guo 等人, 2010 ^[65]	Morph	BIF	AE⇒RC	SVM (C)	N/A	98.30% (RC/B) 97.10% (RC/W)
Guo 等人, 2009 ^[83]	YGA	BIF	AE⇒GR	SVM (C)	N/A	92.25% (GR)
Wang 等人, 2010 ^[84]	FG-NET	AAM	AE⇒GR	SVM (C)	N/A	78.10% (GR/0-10) 80.79% (GR/0-18) 92.16% (GR/19-55) 88.89% (GR/56-69)
Lu 等人, 2006 ^[85]	[85]	[85]	GR⇔RC	SVM (C)	N/A	93.30% (RC/Asian) 97.30% (RC/Non-Asian) 91.00% (GR/M) 75.50% (GR/F)
Narang 等人, 2016 ^[86]	WVU and LDHF	分辨率 128×28	GR⇔RC	Softmax (C)	N/A	99.04% (RC/I) 78.98% (RC/II) 96.41% (GR)
Levi 等人, 2015 ^[88]	Adience	分辨率 227×227	AE⇔GR	Softmax (C)	N/A	86.80% (AE) 84.70% (GR)
Ozbulak 等人, 2016 ^[87]	Adience	CNN 特征	AE⇔GR	SVM (C)	N/A	57.90% (AE) 92.00% (GR)
Tian 等人, 2018 ^[21]	FG-NET Morph	AAM BIF	AE⇔GR	SVM(C/GR) SVOR(R/AE)	≈3.50(AE) ≈5.10(AE)	≈72.00% (GR) ≈87.00% (GR)
Guo 等人, 2011 ^[91]	Morph	BIF	H	KPLS (R)	4.18(AE)	98.85% (RC) 98.20% (GR)
Guo 等人, 2014 ^[64]	Morph	BIF	H	rKCCA (C)	3.98(AE)	99.00% (RC) 98.45% (GR)
Yi 等人, 2014 ^[92]	Morph	[92]	H	Multi-scale CNN (R/AE+C/RC and GR)	3.63(AE)	98.60% (RC) 97.90% (GR)
Xing 等人, 2017 ^[71]	Morph WebFace	分辨率 227×227	H	Hybrid Net (R/AE+C/RC and GR)	2.96(AE) 5.75(AE)	99.20% (Morph/RC) 98.70% (Morph/GR) 92.30% (WebFace/GR)

注:AE 表示年龄估计,GR 表示性别估计,RC 表示人种估计,MAE 表示平均绝对误差,Acc 表示预测精度.C 表示基于分类的方法,R 表示基于回归的方法.N/A 表示未能获取实验结果. $X \Rightarrow Y$ 表示借助 X 属性的 Y 属性估计, $X \Leftrightarrow Y$ 表示 X 、 Y 属性相互作用的属性估计方法,H(ybrid)表示年龄、人种、性别三者相互作用的属性估计方法.M 表示男性,F 表示女性.B 表示黑种人,W 表示白种人,BF 表示黑种人女性,BM 表示黑人男性,WF 表示白种人女性,WM 表示白种人男性. $X \rightarrow Y$ 表示 X 作为训练集, Y 作为测试集. X/Y 表示在 X 上具体 Y 部分,如:R/AE 表示基于回归方法的 AE,GR/M 表示男性性别估计

3 分析与展望

目前为止,学者们提出了多种方法利用单一的人脸属性和不同程度地联合多属性进行研究,以提高 AE、RC 和 GR 的准确率或者精度.文献[1,26,63–65]的研究表明,男性和女性在面部年龄老化方面存在差异,即性别因素

对年龄估计存在影响;白种人、黄种人和黑种人等不同人种的年龄老化规律和性别外观差异也不尽一致.同时,不同年龄阶段的男性和女性的面部性别外观差异也不尽相同.能否有效地利用多属性内在关系将成为影响实际应用的主要因素.

然而,现有方法模型训练通常基于单一的相关数据集(如 FG-NET、Morph Album、Adience、CACD 2000 等),这些数据集通常存在以下两个严重问题:(1) 单一数据集在某些属性上的标记分布不完备.如表 4 所示, Morph 数据集上样本的年龄标记主要集中在 16 岁~70 岁范围;而 FG-NET 数据集上样本的年龄标记主要分布在 0~70 岁.(2) 单一数据集在某些属性上的标记缺失.如表 4 所示, Morph 数据集上的每张人脸图像同时标记了年龄、性别和人种这 3 种属性信息;而 FG-NET 集合中的人脸图像仅提供了年龄属性等标记信息.人脸不同属性在现有单个脸库的标记不完备,使得对其联合估计变得困难.同时,单个数据库样本的缺失导致单独基于该数据集获得的属性估计器泛化能力存在严重缺陷.

Table 4 Label distributions between Morph Album II and FG-NET

表 4 相关属性在不同人脸数据集上的标记信息对比(以 Morph 和 FG-NET 为例)

	年龄属性			性别属性			人种属性		
	有无标记	标记分布(岁)	标记分布是否完备	有无标记	标记分布	标记分布是否完备	有无标记	标记分布	标记分布是否完备
Morph	有	16~70	否(0~15 岁样本缺失)	有	男性、女性	是	有	黄、黑、白和棕等人种	是
FG-NET	有	0~70	是	无	—	—	无	—	—

本文中所指自然的人脸属性估计仅是相对朴素的方法而言,第 2 节所回顾的相关工作仅是一一定程度上考虑并利用了属性之间的关系,然而利用得并不完善.除文献[63,64,71,91,92]等所做的工作之外,本文上述总结回顾的绝大部分人脸年龄、性别和种族等面部属性识别相关工作都未同时实现以上多属性的联合估计.然而,即便是文献[63,64,71,91,92]等给出的联合估计模型,也未能很好地利用年龄、性别和种族等面部属性关系.例如,文献[63]、文献[70]指出其标签连接的方法不仅可能混淆性别和年龄之间的语义关系,而且忽略了男性和女性的老化差异;而文献[70]也仅利用了性别与年龄之间的内在关系,并未考虑人种属性的影响.

在零星的跨语义库研究中,尽管文献[44]实现了跨库训练的年龄单种属性估计,同时通过实验表明跨语义库进行 AE 的效果优于单一数据集,但其并未考虑面部性别和种族等其他属性对年龄估计的影响.此外,该方法亦未实现人脸多个属性的联合估计.

基于以上考虑,人脸面部属性未来的研究方向应基于多属性跨库标记搜集训练数据集,并构建人脸年龄、性别和种族等多属性联合估计模型,探索并建模利用属性间内在关联,以实现更可靠的属性联合估计算法.

4 总 结

就基于人脸图像的性别、年龄、人种等面部属性估计,本文首先系统地回顾了朴素的(不考虑属性之间的相关关联和影响)和自然的(利用了属性之间的相关性)人脸属性相关研究工作.通过对现有工作的回顾和总结,发现现有工作的以下问题和不足:单一训练库属性标记不完备致使估计器泛化性有缺陷;未对年龄、性别和人种等面部多属性进行完全估计;未很好地利用年龄、性别和人种等面部属性间关系.为应对以上问题以获得更好的面部属性估计结果,本文认为进一步工作可基于多属性跨库标记以扩充现有面部属性数据集,并构建人脸年龄、性别和人种等多属性联合估计模型,探索并建模利用属性之间和属性内在的关联,以实现更可靠的属性联合估计算法.

References:

- [1] Lu L. A study of gender classification and age estimation based on face image [Ph.D. Thesis]. Shanghai: Shanghai Jiaotong University, 2010 (in Chinese with English abstract).
- [2] Linoff GS, Berry MJ. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. 3rd ed., New York: John Wiley & Sons, 2011.

- [3] Romano NC, Fjermestad J. Electronic Customer Relationship Management. New York: M.E. Sharpe, 2006.
- [4] Ramanathan N, Chellappa R. Face verification across age progression. *IEEE Trans. on Image Processing*, 2006,15:3349–3361. [doi: 10.1109/TIP.2006.881993]
- [5] Lanitis A, Draganova C, Christodoulou C. Comparing different classifiers for automatic age estimation. *IEEE Trans. on Systems Man & Cybernetics Part B: Cybernetics*, 2004,34(1):621–628. [doi: 10.1109/TSMCB.2003.817091]
- [6] Guo G, Fu Y, Dyer CR, *et al.* Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Trans. on Image Processing*, 2008,17:1178–1188. [doi: 10.1109/TIP.2008.924280]
- [7] Wu Y, Wei Z, Deng RH. Attribute-based access to scalable media in cloud-assisted content sharing networks. *IEEE Trans. on Multimedia*, 2013,15:249–253. [doi: 10.1109/TMM.2013.2238910]
- [8] Wu B, Ai H, Huang C. LUT-based adaboost for gender classification. In: Kittler J, Nixon MS, eds. *Proc. of the Audio- and Video-Based Biometric Person Authentication. Lecture Notes in Computer Science*, Guildford, 2003. 104–110. [doi: 10.1007/3-540-44887-X_13]
- [9] Wu B, Ai H, Huang C. Facial image retrieval based on demographic classification. In: *Proc. of the 17th IEEE Int'l Conf. on Pattern Recognition*. 2004. 914–917. [doi: 10.1109/ICPR.2004.1334677]
- [10] Wang S, Gao Z, He S, *et al.* Gender recognition from visible and thermal infrared facial images. *Multimedia Tools and Applications*, 2016,75(14):8419–8442. [doi: 10.1007/s11042-015-2756-5]
- [11] Aghajanian J, Warrell J, Prince SJD, *et al.* Patch-based within-object classification. In: *Proc. of the 12th IEEE Int'l Conf. on Computer Vision*. 2009. 1125–1132. [doi: 10.1109/ICCV.2009.5459352]
- [12] Dantcheva A, Bremond F. Gender estimation based on smile-dynamics. *IEEE Trans. on Information Forensics and Security*, 2016, 12:719–729. [doi: 10.1109/TIFS.2016.2632070]
- [13] Patel B, Maheshwari RP, Raman B. Compass local binary patterns for gender recognition of facial photographs and sketches. *Neurocomputing*, 2016,218:203–215. [doi: 10.1016/j.neucom.2016.08.055]
- [14] Zhang W, Smith M, Smith L, *et al.* Gender and gaze gesture recognition for human-computer interaction. *Computer Vision & Image Understanding*, 2016,149:32–50. [doi: 10.1016/j.cviu.2016.03.014]
- [15] Castrillón-Santana M, Lorenzo-Navarro J, Ramón-Balmaseda E. Descriptors and regions of interest fusion for in- and cross-database gender classification in the wild. *Image and Vision Computing*, 2017,57:15–24. [doi: 10.1016/j.imavis.2016.10.004]
- [16] Castrillón-Santana M, Lorenzo-Navarro J, Ramón-Balmaseda E. Multi-scale score level fusion of local descriptors for gender classification in the wild. *Multimedia Tools & Applications*, 2016,76:1–17. [doi: 10.1007/s11042-016-3653-2]
- [17] Castrillón-Santana M, Lorenzo-Navarro J, Ramón-Balmaseda E. On using periocular biometric for gender classification in the wild. *Pattern Recognition Letters*, 2015,82:181–189. [doi: 10.1016/j.patrec.2015.09.014]
- [18] Lin G, Chang M, Chang Y, *et al.* A gender classification scheme based on multi-region feature extraction and information fusion for unconstrained images. *Multimedia Tools and Applications*, 2016,16:1–21. [doi: 10.1007/s11042-015-2797-9]
- [19] Liao G, Chen W, Wu Y. Facial features for gender recognition. In: *Proc. of the 35th Chinese Control Conf.* 2016. 27–29. [doi: 10.1109/ChiCC.2016.7554002]
- [20] Mahmood SF, Marhaban MH, Rokhani FZ, *et al.* FASTA-ELM: A fast adaptive shrinkage/thresholding algorithm for extreme learning machine and its application to gender recognition. *Neurocomputing*, 2017,219(5):312–322. [doi: 10.1016/j.neucom.2016.09.046]
- [21] Liew S, Khalil-Hani M, Syafeeza A, *et al.* Gender classification: A convolutional neural network approach. *Turkish Journal of Electrical Engineering & Computer Sciences*, 2016,24:1248–1264. [doi:10.3906/elk-1311-58]
- [22] Coco MD, Carcagni P, Leo M, *et al.* Assessment of deep learning for gender classification on traditional datasets. In: *Proc. of the 13th IEEE Int'l Conf. on Advanced Video and Signal Based Surveillance*. 2016. 271–277. [doi: 10.1109/AVSS.2016.7738061]
- [23] Juefeixu F, Verma E, Goel P, *et al.* DeepGender: Occlusion and low resolution robust facial gender classification via progressively trained convolutional neural networks with attention. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*. 2016. 68–77. [doi: 10.1109/CVPRW.2016.24]
- [24] Mansanet J, Albiol A, Paredes R. Local deep neural networks for gender recognition. *Pattern Recognition Letters*, 2016,70:80–86. [doi: 10.1016/j.patrec.2015.11.015]
- [25] Antipov G, Berrani SA, Dugelay JL. Minimalistic CNN-based ensemble model for gender prediction from face images. *Pattern Recognition Letters*, 2016,70:59–65. [doi: 10.1016/j.patrec.2015.11.011]
- [26] Guo G, Mu G. Human age estimation: What is the influence across race and gender? In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition Workshops*. 2010. 71–78. [doi: 10.1109/CVPRW.2010.5543609]
- [27] Geng X, Smithmiles K. Facial age estimation by multilinear subspace analysis. In: *Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*. 2009. 865–868. [doi: 10.1109/ICASSP.2009.4959721]

- [28] Kim K, Kang S, Chi S, *et al.* Human age estimation using multi-class SVM. In: Proc. of the Int'l Conf. on Ubiquitous Robots and Ambient Intelligence. 2015. 370–372. [doi: 10.1109/URAI.2015.7358911]
- [29] Zhang C, Guo G. Age estimation with expression changes using multiple aging subspaces. In: Proc. of the IEEE Int'l Conf. on Biometrics: Theory, Applications and Systems. 2013. 1–6. [doi: 10.1109/BTAS.2013.6712720]
- [30] Li Z, Fu Y, Huang T. A robust framework for multi-view age estimation. In: Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition Workshops. 2010. 9–16. [doi: 10.1109/CVPRW.2010.5543813]
- [31] Guo G, Mu G, Fu Y, *et al.* A study on automatic age estimation using a large database. In: Proc. of the 12th IEEE Int'l Conf. on Computer Vision. 2009. 1986–1991. [doi: 10.1109/ICCV.2009.5459438]
- [32] Wang X, Guo G. A study on human age estimation under facial expression changes. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition. 2012. 2547–2553. [doi: 10.1109/CVPR.2012.6247972]
- [33] Guo G, Mu G, Fu Y, *et al.* Human age estimation using bio-inspired features. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2009. 112–119. [doi: 10.1109/CVPR.2009.5206681]
- [34] Torrisi A, Farinella GM, Puglisi G, *et al.* Selecting discriminative CLBP patterns for age estimation. In: Proc. of the IEEE Int'l Conf. on Multimedia & Expo Workshops. 2015. 1–6. [doi: 10.1109/ICMEW.2015.7169755]
- [35] Yu Q, Du JY. Age estimation of facial image based on an improved non-negative matrix factorization algorithm. Journal of Image and Graphics, 2008,13(10):1865–1868 (in Chinese with English abstract). [doi: 10.11834/jig.20091029]
- [36] Du JY, Zhai CM, Shang L. Age automation estimation based on a ICA coefficient sparse description algorithm. Journal of Image and Graphics, 2009,14(10):2102–2105 (in Chinese with English abstract). [doi: 10.11834/jig.20091029]
- [37] Lu J, Liong VE, Zhou J. Cost-sensitive local binary feature learning for facial age estimation. IEEE Trans. on Image Processing, 2015,24(12):5356–5368. [doi: 10.1109/TIP.2015.2481327]
- [38] Geng X, Yin C, Zhou ZH. Facial age estimation by learning from label distributions. IEEE Trans. on Pattern Analysis & Machine Intelligence, 2013,35(10):2401–2412. [doi: 10.1109/TPAMI.2013.51]
- [39] Geng X, Wang Q, Xia Y. Facial age estimation by adaptive label distribution learning. In: Proc. of the 22nd IEEE Int'l Conf. on Pattern Recognition. 2014. 4465–4470. [doi: 10.1109/ICPR.2014.764]
- [40] Huo Z, Yang X, Xing C, *et al.* Deep age distribution learning for apparent age estimation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops. 2016. 17–24. [doi: 10.1109/CVPRW.2016.95]
- [41] Yang X, Gao B, Xing C, *et al.* Deep label distribution learning for apparent age estimation. In: Proc. of the IEEE Int'l Conf. on Computer Vision Workshop. 2015. 344–350. [doi: 10.1109/ICCVW.2015.53]
- [42] Li C, Liu Q, Liu J, *et al.* Learning distance metric regression for facial age estimation. In: Proc. of the 21st IEEE Int'l Conf. on Pattern Recognition. 2012. 2327–2330.
- [43] Tian Q, Chen S, Qiao L. Ordinal margin metric learning and its extension for cross-distribution image data. Information Sciences, 2016,349:50–64. [doi: 10.1016/j.ins.2016.02.033]
- [44] Tian Q, Chen S. Cross-heterogeneous-database age estimation with co-representation among them. In: Proc. of the Int'l Conf. on Pattern Recognition. 2016. 1333–1338. [doi: 10.1109/ICPR.2016.7899822]
- [45] Suo J, Wu T, Zhu S, *et al.* Design sparse features for age estimation using hierarchical face model. In: Proc. of the 8th IEEE Int'l Conf. on Automatic Face & Gesture Recognition. 2008. 1–6. [doi: 10.1109/AFGR.2008.4813314]
- [46] Chen C, Yang W, Wang Y, *et al.* Learning gabor features for facial age estimation. In: Sun Z, Lai Z, Chen X, Tan T, eds. Biometric Recognition. Lecture Notes in Computer Science, Beijing, 2011. 204–213. [doi: 10.1007/978-3-642-25449-9_26]
- [47] Chen K, Gong S, Xiang T, *et al.* Cumulative attribute space for age and crowd density estimation. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition. 2013. 2467–2474. [doi: 10.1109/CVPR.2013.319]
- [48] Zhu Y, Li Y, Mu G, *et al.* A study on apparent age estimation. In: Proc. of the IEEE Int'l Conf. on Computer Vision Workshop. 2015. 267–273. [doi: 10.1109/ICCVW.2015.43]
- [49] Tian Q, Chen S. Cumulative attribute relation regularization learning for human age estimation. Neurocomputing, 2015,165: 456–467. [doi: 10.1016/j.neucom.2015.03.078]
- [50] Fernández C, Huerta I, Prati A. A comparative evaluation of regression learning algorithms for facial age estimation. In: Ji Q, Moeslund TB, Hua G, Nasrollahi K, eds. Face and Facial Expression Recognition from Real World Videos. Lecture Notes in Computer Science, Stockholm, 2015. 133–144. [doi: 10.1007/978-3-319-13737-7_12]
- [51] Liang Y, Liu L, Xu Y, *et al.* Multi-task GLOH feature selection for human age estimation. In: Proc. of the 18th IEEE Int'l Conf. on Image Processing. 2011. 565–568. [doi: 10.1109/ICIP.2011.6116611]
- [52] Lu J, Tan Y. Fusing shape and texture information for facial age estimation. In: Proc. of the IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing. 2011. 1477–1480. [doi: 10.1109/ICASSP.2011.5946772]

- [53] Lu J, Tan Y. Ordinary preserving manifold analysis for human age estimation. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops. 2010. 90–95. [doi: 10.1109/TSMCC.2012.2192727]
- [54] Lu J, Tan Y. Ordinary preserving manifold analysis for human age and head pose estimation. IEEE Trans. on Human-machine Systems, 2013,43(2):249–258. [doi: 10.1109/TSMCC.2012.2192727]
- [55] Geng X, Zhou ZH, Zhang Y, *et al.* Learning from facial aging patterns for automatic age estimation. In: Proc. of the 14th ACM Int'l Conf. on Multimedia. 2006. 307–316. [doi: 10.1145/1180639.1180711]
- [56] Xiao B, Yang XK, Xu Y, *et al.* Learning distance metric for regression by semidefinite programming with application to human age estimation. In: Proc. of the 17th ACM Int'l Conf. on Multimedia. 2009. 451–460. [doi: 10.1145/1631272.1631334]
- [57] Liu X, Li S, Kan M, *et al.* AgeNet: Deeply learned regressor and classifier for robust apparent age estimation. In: Proc. of the IEEE Int'l Conf. on Computer Vision Workshop. 2015. 258–266. [doi: 10.1109/ICCVW.2015.42]
- [58] Guo G, Fu Y, Dyer C, *et al.* A probabilistic fusion approach to human age prediction. In: Proc. of the IEEE Computer Society Conf. on Computer Vision & Pattern Recognition Workshops. 2008. 1–6. [doi: 10.1109/CVPRW.2008.4563041]
- [59] Zhang Y, Zhou ZH. A new age estimation method based on ensemble learning. ACTA Automatica Sinica, 2008,34(8):997–1000 (in Chinese with English abstract). [doi: 10.3724/SP.J.1004.2008.00997]
- [60] Guo G, Fu Y, Dyer C, *et al.* Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Trans. on Image Processing, 2008,17(7):1178–1188. [doi: 10.1109/TIP.2008.924280]
- [61] Guo G, Fu Y, Huang T, *et al.* Locally adjusted robust regression for human age estimation. In: Proc. of the IEEE Workshop on Application of Computer Vision. 2008. 1–6. [doi: 10.1109/WACV.2008.4544009]
- [62] Shu X, Tang J, Lai H, *et al.* Personalized age progression with aging dictionary. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2015. 3970–3978. [doi: 10.1109/ICCV.2015.452]
- [63] Guo G, Mu G. Joint estimation of age, gender and ethnicity: CCA vs. PLS. In: Proc. of the 10th IEEE Int'l Conf. and Workshops on Automatic Face and Gesture Recognition. 2013. 1–6. [doi: 10.1109/FG.2013.6553737]
- [64] Guo G, Mu G. A framework for joint estimation of age, gender and ethnicity on a large database. Image & Vision Computing, 2014,32(10):761–770. [doi: 10.1016/j.imavis.2014.04.011]
- [65] Guo G, Mu G. A study of large-scale ethnicity estimation with gender and age variations. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition Workshops. 2010. 79–86. [doi: 10.1109/CVPRW.2010.5543608]
- [66] Andrew YN, Michael IJ. On discriminative vs. generative classifiers: A comparison of logistic regression and naïve Bayes. In: Proc. of the Neural Information Processing Systems. 2002. 605–610.
- [67] Eidinger E, Enbar R, Hassner T. Age and gender estimation of unfiltered faces. IEEE Trans. on Information Forensics & Security, 2014,9(12):2170–2179. [doi: 10.1109/TIFS.2014.2359646]
- [68] Chen BC, Chen CS, Hsu WH. Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset. IEEE Trans. on Multimedia, 2015,17(6):804–815. [doi: 10.1109/TMM.2015.2420374]
- [69] Li K, Xing J, Hu W, *et al.* D2C: Deep cumulatively and comparatively learning for human age estimation. Pattern Recognition, 2017,66:95–105. [doi: 10.1016/j.patcog.2017.01.007]
- [70] Tian Q, Chen S. Joint gender classification and age estimation by nearly orthogonalizing their semantic spaces. Image and Vision computing, 2018,69:9–21. [doi: 10.1016/j.imavis.2017.10.003]
- [71] Xing J, Li K, Hu W, *et al.* Diagnosing deep learning models for high accuracy age estimation from a single image. Pattern Recognition, 2017,66:106–116. [doi: 10.1016/j.patcog.2017.01.005]
- [72] Lu X, Jain A. Ethnicity identification from face images. Int'l Society for Optics and Photonics, 2004,5404:114–123.
- [73] Lu H, Liu J, Li C, *et al.* Learning ordinal discriminative features for age estimation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2012. 2570–2577. [doi: 10.1109/CVPR.2012.6247975]
- [74] He Z, Li X, Zhang Z, *et al.* Data-dependent label distribution learning for age estimation. IEEE Trans. on Image Processing, 2017, 26(8):3846–3858. [doi: 10.1109/TIP.2017.2655445]
- [75] Rothe R, Timofte R, Gool LV. DEX: Deep expectation of apparent age from a single image. In: Proc. of the IEEE Int'l Conf. on Computer Vision Workshop. 2015. 252–257. [doi: 10.1109/ICCVW.2015.41]
- [76] Rothe R, Timofte R, Gool LV. Deep expectation of real and apparent age from a single image without facial landmarks. Int'l Journal of Computer Vision, 2018,4:144–157. [doi: 10.1007/s11263-016-0940-3]
- [77] Lou Z, Alnajjar F, Alvarez JM, *et al.* Expression-invariant age estimation using structured learning. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2018,40(2):365–375. [doi: 10.1109/TPAMI.2017.2679739]
- [78] Li Y, Peng Z, Liang D, *et al.* Facial age estimation by using stacked feature composition and selection. The Visual Computer, 2016,32(12):1525–1536. [doi: 10.1007/s00371-015-1137-4]

- [79] Niu Z, Zhou M, Wang L, *et al.* Ordinal regression with multiple output CNN for age estimation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 4920–4928. [doi: 10.1109/CVPR.2016.532]
- [80] Shu X, Tang J, Li Z, *et al.* Personalized age progression with bi-level aging dictionary learning. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017,40:905–917. [doi: 10.1109/TPAMI.2017.2705122]
- [81] Guo G, Zhang C. A study on cross-population age estimation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2014. 4257–4263. [doi: 10.1109/CVPR.2014.542]
- [82] Li K, Xing J, Su C, Hu W, *et al.* Deep cost-sensitive and order-preserving feature learning for cross-population age estimation. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition. 2018. 399–408. [doi: 10.1109/CVPR.2018.00049]
- [83] Guo G, Dyer CR, Fu Y, *et al.* Is gender recognition affected by age? In: Proc. of the 12th IEEE Int'l Conf. on Computer Vision Workshops. 2009. 2032–2039. [doi: 10.1109/ICCVW.2009.5457531]
- [84] Wang Y, Ricanek K, Chen C, *et al.* Gender classification from infants to seniors. In: Proc. of the IEEE Int'l Conf. on Biometrics: Theory Applications and Systems. 2010. 1–6. [doi: 10.1109/BTAS.2010.5634518]
- [85] Lu X, Chen H, Jain AK. Multimodal facial gender and ethnicity identification. In: Zhang D, Jain AK, eds. Advances in Biometrics. Lecture Notes in Computer Science, Hong Kong, 2006. 554–561. [doi: 10.1007/11608288_74]
- [86] Narang N, Bourlai T. Gender and ethnicity classification using deep learning in heterogeneous face recognition. In: Proc. of the Int'l Conf. on Biometrics. 2016. 1–8. [doi: 10.1109/ICB.2016.7550082]
- [87] Ozbulak G, Aytar Y, Ekenel HK. How transferable are CNN-based features for age and gender classification. In: Proc. of the Int'l Conf. of the Biometrics Special Interest Group. 2016. 1–6. [doi: 10.1109/BIOSIG.2016.7736925]
- [88] Levi G. Age and gender classification using convolutional neural networks. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops. 2015. 34–42. [doi: 10.1109/CVPRW.2015.7301352]
- [89] Uricar M, Timofte R, Rothe R, *et al.* Structured output SVM prediction of apparent age, gender and smile from deep features. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops. 2016. 730–738. [doi: 10.1109/CVPRW.2016.96]
- [90] Tian Q, Chen S, Tan X. A unified gender-aware age estimation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 1–8.
- [91] Guo G, Mu G. Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition. 2011. 657–664. [doi: 10.1109/CVPR.2011.5995404]
- [92] Yi D, Lei Z, Li SZ. Age estimation by multi-scale convolutional network. In: Cremers D, Reid I, Saito H, Yang MH, eds. Proc. of the Computer Vision-ACCV 2014. Lecture Notes in Computer Science, Singapore, 2014. 144–158. [doi: 10.1007/978-3-319-16811-1_10]

附中文参考文献:

- [1] 陆丽.基于人脸图像的性别识别与年龄估计研究[博士学位论文].上海:上海交通大学,2010.
- [35] 余庆,杜吉祥.基于一种改进的 NMF 算法的人脸年龄估计方法.中国图像图形学报,2008,13(10):1865–1868.
- [36] 杜吉祥,翟传敏,尚丽.基于 ICA 系数稀疏表示的年龄自动估计.中国图像图形学报,2009,14(10):2102–2105.
- [59] 张宇,周志华.基于集成的年龄估计方法.自动化学报,2008,34(8):997–1000.



曹猛(1994—),男,江苏扬州人,学士,主要研究领域为机器学习,模式识别.



田青(1983—),男,博士,副教授,主要研究领域为机器学习,模式识别,机器视觉.



马廷淮(1974—),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为数据挖掘,数据共享与存储,社会网络,隐私保护.



陈松灿(1962—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为机器学习,模式识别.