

基于标签敏感最大间隔准则的人脸年龄两步估计算法

徐晓玲^{1,2} 金 忠^{1,2} 贲圣兰³

(南京理工大学计算机科学与工程学院 南京 210094)¹

(南京理工大学高维信息智能感知与系统教育部重点实验室 南京 210094)²

(南京大学电子科学与工程学院 南京 210023)³

摘 要 传统的最大间隔准则在计算类间离散度矩阵时往往忽略了类别之间的差异,但是对于人脸年龄估计,不同年龄标签之间的差异性是非常显著的。因此,在标签之间引入距离度量,提出标签敏感的最大间隔准则维数约减算法。此外,考虑到人脸变老的复杂性,提出两步的局部回归算法—— K 近邻-标签分布的支持向量回归(K Nearest Neighbors-Label Distribution Support Vector Regressor, KNN-LDSVR),以进行人脸年龄估计。在 FGNET 数据库子集上提出的人脸年龄估计方法的平均绝对误差为 4.1 岁,相对于已有的年龄估计方法,性能得到提升。

关键词 最大间隔准则,标签敏感,两步,局部回归,年龄估计

中图分类号 TP391.4 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.06.050

Facial Age Two-steps Estimation Algorithm Based on Label-sensitive Maximum Margin Criterion

XU Xiao-ling^{1,2} JIN Zhong^{1,2} BEN Sheng-lan³

(School of Computer Science and Engineering, Nanjing University of Science & Technology, Nanjing 210094, China)¹

(Key Laboratory of Intelligent Perception and System for High-Dimensional Information of Ministry of Education, Nanjing University of Science & Technology, Nanjing 210094, China)²

(School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China)³

Abstract Traditional maximum margin criterion usually ignores the differences between classes in the computation of the between-class scatter matrix. However, for facial age estimation, the differences between age labels are very significant. Therefore, this paper proposed a novel dimensionality reduction algorithm, called label-sensitive maximum margin criterion (lsMMC), by introducing a distance metric between the classes. In addition, considering the complicated facial aging process, this paper proposed a two-steps local regression algorithm named K nearest neighbors-label distribution support vector regressor (KNN-LDSVR) for age estimation. The mean absolute error of the proposed facial aging estimation method on the FGNET database subset is 4.1 years, which improves the performance compared with existing age estimation methods.

Keywords Maximum margin criterion, Label-sensitive, Two-steps, Local regression, Age estimation

1 引言

人脸作为一种丰富的信息源,反映了很多具有鉴别能力的信息,如身份、性别、种族、年龄、表情等。近年来,基于人脸图像的年龄估计成为了模式识别、人工智能、计算机视觉等领域的研究热点,在安全监控、人机交互和电子商务系统等诸多领域有着广泛的应用。如可以阻止未成年人从自动售货机购买烟酒及上网浏览不健康网页等。

人脸年龄估计存在着诸多挑战。1)由于基因、生活环境、健康状况、生活习惯等因素的影响,不同人的变老速度和方式存在差异。2)传统的多分类方法同等对待各个类别,样本发生误分时产生的代价都是相同的。但对于年龄估计而言,将

样本误分为与其真实年龄相差较小和相差较大时的代价是不同的,如将 20 岁分别判断为 21 岁和 30 岁时对人的心情的影响完全不同。3)在不同的年龄段,人脸变老的方式存在差异。

人脸年龄估计算法主要包括两个步骤:特征提取和年龄估计。特征提取是从图像中提取出能够表征人脸年龄变化的特征,稳定而有效的年龄特征能降低年龄估计算法的难度,提高系统的性能。人在变老的过程中主要包括面部形状和皮肤纹理的变化。与成人相比,儿童的面部比例结构有很大的不同,可以使用特征点的坐标表示形状特征,以区分儿童和成年人。成年后,人脸的形状不会发生太大的改变,需要借助肤色的变化及皱纹等信息来进行分类。Guo 等^[1]首次将生物启发特征用于人脸年龄估计中,该特征由简单层和复杂层组成,因

收稿日期:2017-04-18 返修日期:2017-09-04 本文受国家自然科学基金(61373063, 61375007, 61233011), 国家重点基础研究发展计划(2014CB349303)资助。

徐晓玲(1993—),女,硕士生,CCF 会员,主要研究领域为计算机视觉、人脸图像分析,E-mail: xuxiaoling1028.njust@hotmail.com;金 忠(1961—),男,博士,教授,主要研究领域为模式识别、图像分析、机器学习、计算机视觉,E-mail: zhongjin@njust.edu.cn(通信作者);贲圣兰(1982—),女,博士,主要研究领域为模式识别、人脸识别。

此可以提取皮肤的纹理信息,模拟人类视觉皮层的高级识别过程。

为了能够表示所有年龄段的特征,可以将形状特征与纹理特征相结合。主动表观模型^[2]结合了人脸的形状信息和纹理信息,被广泛应用于人脸年龄估计,但其在建模过程中使用了主成分分析,会丢失一些重要的细节信息。

此外,深度学习特征也可以用于年龄估计。Dong 等^[3]使用深度卷积神经网络来提取年龄特征,基于迁移学习对网络进行训练。尽管深度学习的性能普遍较好,但其需要大量的训练样本,且可解释性较差。

大多数现有的常规或深度学习的年龄特征提取算法都没有很好地兼顾年龄之间的有序性和相关性,因此提取到的特征的鉴别性不高。但也有部分学者对此问题进行了研究。Ma 等^[4]提出了基于类别距离的鉴别分析,对特征进行维数约减,在保持类内局部结构的同时,提高了类间的可分性。Chao 等^[5]提出使用标签敏感的局部保持投影和边缘费希尔分析对特征进行降维,以保持相邻样本之间的特征相似性和标签相关性。

年龄估计具有特殊性,既可以看作普通的多分类问题,又可以看作回归问题。与仅用单个分类或回归器相比,将两者进行组合,即用分层的年龄估计能提升年龄估计的准确性^[6]。Pontes 等^[7]首先使用支持向量机将图像划分为 4 个年龄段,再使用支持向量回归来估计具体的年龄。

与只使用图像的真实年龄作为标签相比,标签分布使得一幅图像不仅有利于自身年龄标签的学习,还有利于其相邻年龄的学习,在一定程度上弥补了人脸年龄数据库样本数较少的缺点。Geng 等^[8]为每幅图像分配一个标签分布,并提出了 IIS-LLD 和 CPNN 两种算法来学习这种标签分布。He 等^[9]提出了一个数据驱动的标签分布学习算法,可以自适应地学习图像对应的标签分布。

大多数已有的人脸年龄估计算法存在以下问题:1)提取年龄特征时未考虑类别之间的有序性和相关性;2)进行由粗到精的年龄估计时,通常是先将样本划分到某个年龄段,再基于该年龄段内所有样本构建年龄决策函数,以估计样本的具体年龄,如果在第一阶段的年龄段划分出现错误,那么后续的年龄估计会直接受到影响。为了解决上述问题,本文提出了一种基于标签敏感最大间隔准则的人脸年龄两步估计算法,具体步骤如图 1 所示。

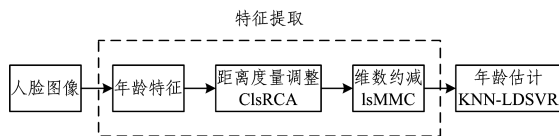


图 1 本文年龄估计算法的流程图

Fig. 1 Flow chart of age estimation algorithm in this paper

首先提取图像的年龄特征,并对其进行距离度量调整,以获得一个合适的特征空间;然后为了挖掘年龄标签之间的有序性和相关性,对新的特征空间依据标签敏感的最大间隔准则进行降维,在计算类内和类间离散度矩阵时,需要根据样本在特征空间的相似性及标签之间的距离度量分配相应的权重;最后使用标签分布和由粗到精的方式构建决策函数。针对年龄变老过程的复杂性和特征的局部统计特性,提出了一

种两步的局部回归算法—— K 近邻-标签分布的支持向量回归。在 FGNET 人脸年龄数据库^[10]的子集上进行实验,结果证明了本文方法的有效性。

2 相关工作

2.1 补偿的标签敏感相关成分分析

在模式识别中,常用的距离调整方法是对特征进行归一化,使其均值为 0,标准差为 1。但是,这种方法没有利用标签信息。对于距离度量调整和维数约减,标签敏感指计算离散度矩阵时,根据样本对特征空间的相似性和标签之间的差异程度分配相应的权重,充分挖掘特征与标签之间的关系。补偿的标签敏感相关成分分析(Compensated label-sensitive Relevant Component Analysis, ClsRCA)^[5]是一种有效的距离度量调整方法,该方法不仅充分利用了相关成分分析的有效性和鉴别性,还考虑到了年龄标签之间的有序性和相关性。

给定训练样本 $x_n \in \mathbb{R}^D$ 及相应类别标签 $l_n \in \{1, 2, \dots, C\}$ ($n=1, 2, \dots, N$), N 和 C 分别为样本数和类别数。ClsRCA 的目标是学习一个投影矩阵 $P \in \mathbb{R}^{D \times D}$,使得投影 $y_n = P^T x_n \in \mathbb{R}^D$ ($n=1, 2, \dots, N$) 后同类样本相距较近。第 c 类的类内离散度矩阵 S_c ($c=1, 2, \dots, C$) 的定义如下:

$$S_c = \frac{1}{\omega_c} \sum_{n=1}^N e_n^{(c)} (x_n - \mu_c)(x_n - \mu_c)^T \quad (1)$$

其中, $e_n^{(c)}$ 为第 n 个样本属于第 c 类的权重, ω_c 为第 c 类总权重, μ_c 为第 c 类样本的加权均值,计算式分别为:

$$e_n^{(c)} = \begin{cases} \exp\left(-\frac{(l_n - c)^2}{\rho}\right), & |l_n - c| \leq \epsilon_1 \\ 0, & \text{其他} \end{cases} \quad (2)$$

$$\omega_c = \sum_{n=1}^N e_n^{(c)} \quad (3)$$

$$\mu_c = \frac{1}{\omega_c} \sum_{n=1}^N e_n^{(c)} x_n \quad (4)$$

其中, ρ 为预先设定的系数,用于控制 $e_n^{(c)}$ 的取值; ϵ_1 用于定义标签相似样本之间的绝对年龄差的范围, ρ 和 ϵ_1 的具体取值见 4.3 节。可以看出,计算 S_c 时,不仅需要考虑到第 c 类样本,还需要考虑到与 c 年龄绝对差不大于阈值 ϵ_1 的相似样本。为了平衡各类标签的影响,定义总体类内离散矩阵为:

$$S = \frac{1}{C} \sum_{c=1}^C S_c \quad (5)$$

通过白化总体类内离散度矩阵 S 求得投影矩阵 P :

$$P = S^{-\frac{1}{2}} \quad (6)$$

2.2 最大间隔准则

Li 等^[11]提出的最大间隔准则(Maximum Marginal Criterion, MMC)是一种简单且有效的监督维数约减方法。维数约减的意义是在降低原始数据维数的同时保持其完整性,防止维数灾难的发生。在约减后的空间中执行后续程序将大大减少计算量,提高算法的效率,且能获得与原始数据基本一致的结果。费希尔鉴别分析^[12]是一种经典的监督维数约减算法,可以在最大化类间离散度的同时最小化类内离散度。但是对于小样本问题(样本数远小于样本维数),类内离散度矩阵是奇异的,此时该方法不再适用,使用 MMC 可以很好地避免这个问题。

给定训练样本 $y_n \in \mathbb{R}^D$ 及相应类别标签 $l_n \in \{1, 2, \dots, C\}$ ($n=1, 2, \dots, N$)。MMC 的目标是学习一个投影矩阵 $Q \in$

$\mathbb{R}^{D \times d}$, 使得投影 $z_n = Q^T y_n \in \mathbb{R}^d (n=1, 2, \dots, N)$ 后不同类别之间的距离最大化。类内离散度矩阵 S_w 和类间离散度矩阵 S_b 按样本对的方式定义为^[13]:

$$S_w = \frac{1}{2} \sum_{i,j=1}^N W_{i,j}^{(w)} (y_i - y_j)(y_i - y_j)^T \tag{7}$$

$$S_b = \frac{1}{2} \sum_{i,j=1}^N W_{i,j}^{(b)} (y_i - y_j)(y_i - y_j)^T \tag{8}$$

其中, $W_{i,j}^{(w)}$ 和 $W_{i,j}^{(b)}$ 分别为样本对 (y_i, y_j) 在 S_w 和 S_b 中的权重, 计算式为:

$$W_{i,j}^{(w)} = \begin{cases} \frac{1}{N_c}, & l_i = l_j = c \\ 0, & l_i \neq l_j \end{cases} \tag{9}$$

$$W_{i,j}^{(b)} = \begin{cases} \frac{1}{N} - \frac{1}{N_c}, & l_i = l_j = c \\ \frac{1}{N}, & l_i \neq l_j \end{cases} \tag{10}$$

其中, N_c 为第 c 类样本数。对于同类样本对而言, 权重与该类样本数有关, 若数据库中各类样本数分布不均衡, 各类权重之间的差异将会较大, 为了避免这种情况, 在实验时对样本数较少的年龄段进行了剔除, 具体操作见 4.1 节。通过求解以下目标函数获得最优的投影矩阵:

$$Q^* = \arg \max_{Q \in \mathbb{R}^{D \times d}} [\text{tr}(Q^T (S_b - S_w) Q)] \tag{11}$$

MMC 可以保证投影后的同类样本对相距较近, 异类样本对相距较远。

2.3 标签分布的支持向量回归

标签分布的支持向量回归 (Label Distribution Support Vector Regressor, LDSVR)^[14] 的基本思想是使用支持向量机为样本与标签分布拟合出一个 sigmoid 函数。给定训练样本 $z_n \in \mathbb{R}^d$ 及相应的标签分布 $h_n = [h_n^{(1)}, h_n^{(2)}, \dots, h_n^{(C)}] (n=1, 2, \dots, N)$, 其中 $h_n^{(c)} \in [0, 1]$ 且 $\sum_{c=1}^C h_n^{(c)} = 1$ 。 $h_n^{(c)}$ 与 z_n 的关系如下:

$$h_n^{(c)} = \frac{1}{1 + \exp(-(\omega^{(c)})^T \varphi(z_n) - b)} \tag{12}$$

其中, φ 是将 z_n 映射到高维特征空间 \mathbb{R}^V 的变换矩阵, $\omega^{(c)} \in \mathbb{R}^V$ 和 b 是模型参数。最小化下面的目标函数:

$$T(W, b) = \frac{1}{2} \sum_{c=1}^C \| \omega^{(c)} \|^2 + C_1 \sum_{n=1}^N L(u_n) \tag{13}$$

其中, $W = [\omega^{(1)}, \omega^{(2)}, \dots, \omega^{(C)}]^T$, C_1 为惩罚系数, $L(u_n)$ 为样本 z_n 的损失函数, 定义如下:

$$L(u_n) = \begin{cases} 0, & u_n < \varepsilon_2 \\ (u_n - \varepsilon_2)^2, & u_n \geq \varepsilon_2 \end{cases} \tag{14}$$

其中, u_n 的定义如下:

$$u_n = \frac{\| r_n \|^2}{4} = \frac{\sqrt{r_n^T r_n}}{4} \tag{15}$$

$$r_n = -\log\left(\frac{1}{h_n} - 1\right) - (W\varphi(z_n) + b) \tag{16}$$

3 基于标签敏感最大间隔准则的人脸年龄两步估计算法

3.1 年龄特征

本文综合使用了人脸的形状和纹理信息进行年龄估计。其中形状特征用图 2 所示的 68 个显著点表示, 其坐标为 $f_s = [x_1, y_1, \dots, x_{68}, y_{68}]$ 。为了消除旋转、平移和缩放带来的影响, 使用普鲁克分析^[15] 对 f_s 进行对齐, 使所有的形状特征都

位于同一个坐标系中。

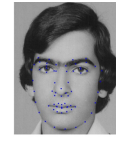


图 2 68 个显著点的位置

Fig. 2 Positions of 68 landmarks

为了提取人脸皮肤的纹理信息, 对预处理后的图像提取生物启发特征, 包括简单层 $s1$ 和复杂层 $c1$, 以模拟人类视觉皮层的高级识别过程。 $s1$ 层对应于年长动物视觉皮层中的简单细胞, 能对不同方向的条件输入作出响应。 $s1$ 层使用二维 Gabor 滤波器对图像进行处理, 计算式为:

$$G(x, y) = \exp\left(-\frac{X^2 + \xi^2 Y^2}{2\zeta^2}\right) \cos\left(\frac{2\pi X}{\lambda}\right) \tag{17}$$

其中,

$$\begin{cases} X = x \cos \theta + y \sin \theta \\ Y = -x \sin \theta + y \cos \theta \end{cases} \tag{18}$$

其中, θ 为 Gabor 滤波器旋转的角度, 介于 $0 \sim \pi$ 之间, ξ, ζ 和 λ 为预先设定的参数。

$c1$ 层能模拟灵长类动物视觉皮层的复杂细胞, 具有平移和尺度不变性。第一步是 $s1$ 层各子带, 即同一方向不同尺度的两个子图通过 MAX 操作综合成一幅图像; 第二步是将每幅综合图像划分为若干个大小为 $N_s \times N_s$ 的块, 在每个块内进行 STD 操作, 如式(19)所示。

$$std = \sqrt{\frac{1}{N_s \times N_s} \sum_{i=1}^{N_s \times N_s} (m_i - \bar{m})^2} \tag{19}$$

其中, m_i 是块内位置 i 上的值, \bar{m} 是块内的平均值。经过 $c1$ 层处理后得到生物启发特征 $f_b = [std_1, std_2, \dots, std_M]$, M 为块的个数。相关参数的设置请见文献[1]。

最后, 将形状特征 f_s 与生物启发特征 f_b 串接后得到 $f = [f_s, f_b]$, 对 f 用主成分分析得到最后的年龄融合特征 $x_n \in \mathbb{R}^D (n=1, 2, \dots, N)$ 。

3.2 标签敏感的最大间隔准则

对于第 n 个样本, 已知年龄特征 $x_n \in \mathbb{R}^D$ 及相应类别标签 $l_n \in \{1, 2, \dots, C\} (n=1, 2, \dots, N)$, 在年龄估计时, 理想的近邻样本应该在年龄上与其相近, 而不是在姿势、表情和身份上与其相似。

使用欧氏距离度量寻找近邻样本时, 若假设特征之间相互独立(通过 PCA 操作实现), 标准差越大的特征在计算欧氏距离时的影响越大。因此, 对于标准差较大的特征而言, 需要检查其与年龄是否直接相关。在 FGNET 数据库上提取的年龄融合特征的每一维的标准差如图 3 所示。可以看出, 标准差呈降序排列, 波动比较明显。

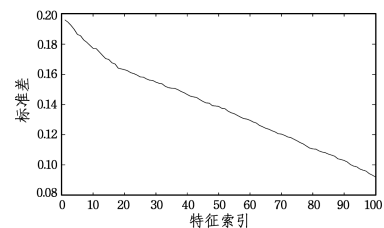


图 3 年龄融合特征每一维的标准差

Fig. 3 STD of each fused aging feature

对于一幅样本图像(见图 4(a)),只取特征的前 5 维,使用欧氏距离度量在数据库中寻找近邻,结果如图 4(b)所示。可以看出,样本图像与其近邻主要是表情和姿势上的相似,说明标准差较大的特征不与年龄直接相关。



(a) 样本图像



(b) ClsRCA 处理前的 K 近邻结果



(c) ClsRCA 处理后的 K 近邻结果

注:K=5,图像下方为样本真实年龄

图 4 ClsRCA 处理前后的样本的 K 近邻结果

Fig. 4 K neighboring results of sample picture before and after ClsRCA

为了使样本间的欧氏距离能与类别可分性直接相关,即样本间的欧氏距离越近,它们的年龄标签就越相似,使用 ClsRCA 对 x_n 进行距离度量调整,重新确定每一维特征的影响,投影矩阵为 $P \in \mathbb{R}^{D \times D}$,调整后的特征为 $y_n = P^T x_n \in \mathbb{R}^D$ ($n=1, 2, \dots, N$)。此时,样本图像的近邻结果如图 4(c)所示。可以看出,此时样本图像与其近邻的年龄比较相近,表明调整后的特征与年龄直接相关。

调整后的年龄特征维数保持不变,因此需要对其进行维数约减。MMC 是一种有监督的维数约减算法,但其忽略了年龄标签之间的差异。对于人脸年龄估计,不同年龄标签之间的差异性是非常显著的。因此,在 MMC 中通过引入标签敏感信息提出了标签敏感的最大间隔准则(label-sensitive Maximum Marginal Criterion, lsMMC)。

MMC 在计算离散度矩阵时,同类样本对的权重相同,不能很好地保持类内的局部结构。受局部费希尔鉴别分析^[15]的启发,可以将 MMC 与局部保持投影相结合,在最大化类别间距离的同时,很好地保持原始数据空间中的类内局部结构。计算离散度矩阵时,对于同类样本,根据它们之间的特征相似度赋予相应的权重,相似度越高,权重越大,投影之后的距离越近。将 MMC 中的式(9)、式(10)分别更新为:

$$W_{i,j}^{(w)} = \begin{cases} \frac{A_{i,j}}{N_c}, & l_i = l_j = c \\ 0, & l_i \neq l_j \end{cases} \quad (20)$$

$$W_{i,j}^{(b)} = \begin{cases} A_{i,j} \left(\frac{1}{N} - \frac{1}{N_c} \right), & l_i = l_j = c \\ \frac{1}{N}, & l_i \neq l_j \end{cases} \quad (21)$$

其中, $A_{i,j}$ 为权重系数,与同类样本之间的特征相似度有关,定义如下:

$$A_{i,j} = \exp\left(-\frac{\|y_i - y_j\|^2}{\sigma^2}\right) \quad (22)$$

其中, σ 为样本特征之间的平均欧氏距离。

此时,类别标签之间只有同类和异类两种关系。计算类间离散度矩阵时,所有的异类样本对的权重相同,没有考虑类别标签间的差异程度。在特征空间中相距较近的样本很容易发生误分,若此时样本与实际类别标签相差较远,将会大大降低年龄估计性能,因此需要计算类间离散度矩阵来增大它们的权重,以保证在投影后的空间中拉开它们之间的距离,将式(21)进一步更新为:

$$W_{i,j}^{(b)} = \begin{cases} A_{i,j} \left(\frac{1}{N} - \frac{1}{N_c} \right), & l_i = l_j = c \\ \frac{B_{i,j}}{N}, & l_i \neq l_j \end{cases} \quad (23)$$

其中, $B_{i,j}$ 为权重系数,定义如下:

$$B_{i,j} = \frac{A_{i,j}}{1 + \exp(-\gamma(|l_i - l_j| - \eta))} \quad (24)$$

其中, σ, γ, η 是预先设定的参数。式(24)中,分子表示样本对之间的特征相似性,分母表示标签间的差异程度。

最优投影矩阵通过求解式(11)得到:

$$(S_b - S_w)v = \lambda v \quad (25)$$

其中, S_b 由式(8)、式(23)计算得到, S_w 由式(7)、式(20)计算得到。利用求得的投影矩阵 Q 得到最终的年龄特征 $z_n = Q^T y_n \in \mathbb{R}^d$ ($n=1, 2, \dots, N$)。若 $A_{i,j}$ 和 $B_{i,j}$ 都为 1,则 lsMMC 将退化为 MMC。lsMMC 的具体步骤如算法 1 所示。

算法 1 lsMMC

输入:训练样本 $y_n \in \mathbb{R}^D$ 及对应标签 $l_n \in \{1, 2, \dots, C\}$ ($n=1, 2, \dots, N$),总样本数 N ,总类别数 C ,每类样本数 N_c ($c=1, 2, \dots, C$),参数 σ, γ, η ,投影后的空间维度 d

输出:投影矩阵 $Q \in \mathbb{R}^{D \times d}$

- Step 1 按式(22)、式(24)分别计算同类样本对和异类样本对的系数 $A_{i,j}, B_{i,j}$;
- Step 2 按式(20)、式(23)分别计算类内和类间离散度矩阵中的权重矩阵 $W_{i,j}^{(w)}, W_{i,j}^{(b)}$;
- Step 3 按式(7)、式(8)分别计算类内和类间离散度矩阵 S_w, S_b ;
- Step 4 求解式(25),得到按降序排列的特征值 $\{\lambda_i\}_{i=1}^D$ 和对应的特征向量 $\{v_i\}_{i=1}^D$;
- Step 5 最优投影矩阵 $Q = [v_1, v_2, \dots, v_d] \in \mathbb{R}^{D \times d}$ 。

值得注意的是,距离度量调整和维数约减都是在训练阶段学习一个投影矩阵,在测试阶段可以直接用其对数据进行处理,从而得到最终的年龄特征,并将其用于后续的年龄估计。

3.3 K 近邻-标签分布的支持向量回归

经过距离度量调整的特征空间是局部欧氏空间,能很好地挖掘局部统计特性。基于该特性,本文提出了一种两步的

局部回归算法(K Nearest Neighbors-Label Distribution Support Vector Regressor, KNN-LDSVR)来进行年龄估计。该方法也更符合人脸变老的特性,人脸变老是一个极其复杂的过程,在不同年龄其变老的速度和方式存在差异,与基于特定年龄段对所有样本进行回归相比,基于近邻样本进行局部回归能够产生更准确的年龄决策函数。

给定一个测试样本 $z \in \mathbb{R}^d$, KNN-LDSVR 首先在训练集中找出前 K 个欧氏近邻,然后使用高斯函数为这些近邻样本分配标签分布,具体标签分布的生成方法请参照文献[8]。基于这些近邻样本 LDSVR^[16] 训练局部回归器,最后用训练好的回归器估计出测试样本 z 的标签分布 h ,其中隶属度最高的标签即为最终的估计年龄。KNN-LDSVR 的具体步骤如算法 2 所示。

算法 2 KNN-LDSVR

输入:训练样本 $z_n \in \mathbb{R}^d$ 及对应的标签分布 $h_n \in \mathbb{R}^C$ ($n=1, 2, \dots, N$),

测试样本 z , 近邻数 K , 参数 C_1, ϵ_2

输出:测试样本 z 的年龄 l

Step 1 在训练集中寻找 z 的前 K 个近邻 $\{z_j^{KNN}, h_j^{KNN}\}_{j=1}^K$;

Step 2 用 $\{z_j^{KNN}, h_j^{KNN}\}_{j=1}^K$ 训练一个径向基核的 LDSVR 局部回归算法来预测 z 的年龄。

4 实验

4.1 FGNET 数据库

FGNET 数据库^[12] 是人脸年龄估计常用的数据库。该数据库共有 82 个对象的 1002 幅人脸图像,每幅图像有 68 个显著点的位置信息,年龄范围为 0~69 岁。数据库的分布极不均匀,各年龄段的图像数量差异很大。

考虑到 40~69 岁的图像极少(只有 69 幅),在 KNN-LDSVR 算法中为该年龄段内的样本寻找 K 近邻时,由于缺乏相近的年龄样本,这些样本的部分近邻会与其真实年龄相差很大。利用这些近邻样本训练局部回归器显然是不合理的,因此只使用小于 40 岁的 933 幅图像进行实验,这也在一定程度上避免了样本分布不均衡对离散度矩阵计算中同类样本对权重的影响。此时各年龄段图像的分布情况如表 1 所列。

表 1 FGNET 数据库子集中各年龄段图像的分布情况

Table 1 Number of images in each age group in the FGNET database subset

年龄段	0~9	10~19	20~29	30~39	40~49
样本数	371	339	144	79	933

4.2 评价准则

本文采用平均绝对误差(Mean Absolute Error, MAE)^[16] 和累积误差指数(Cumulative Score, CS)^[17] 作为年龄估计的评价准则。MAE 定义为测试集中样本的估计年龄与真实年龄的绝对误差的均值,如式(26)所示:

$$MAE = \frac{1}{N_t} \sum_{i=1}^{N_t} |\hat{l}_i - l_i| \tag{26}$$

其中, \hat{l}_i 和 l_i 分别为测试样本 z_i 的估计年龄和真实年龄, N_t

为总的测试样本数。MAE 越小,表明估计的年龄与真实年龄越接近,算法的性能越好。

CS 表示绝对误差小于阈值 δ 的测试样本占总测试样本的比例,定义为:

$$CS(\delta) = \frac{1}{N_t} \sum_{i=1}^{N_t} I[|\hat{l}_i - l_i| \leq \delta] \times 100\% \tag{27}$$

其中, $I(\cdot)$ 为逻辑函数。对于同一个阈值 δ , CS 越大,算法的性能越好。

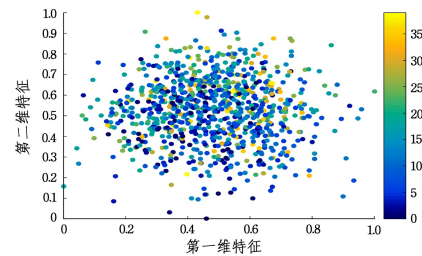
4.3 实验设置

为了避免同一个人的图像同时出现在训练集和测试集中,实验采用留一法模式,每轮只用一个人的图像来测试,其余人的图像全部用于训练。82 轮后,每幅图片都作为测试样本进行实验得到一个估计的年龄,可以在一定程度上弥补数据库图片较少的缺点。

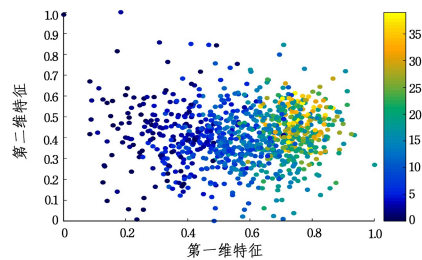
实验中的一些相关参数的设置如下: $\epsilon_1 = 5, \rho = 12, \gamma = 2, \eta = 2, C_1 = 1, \epsilon_2 = 0.1$ 。

4.4 标签有序性提升实验

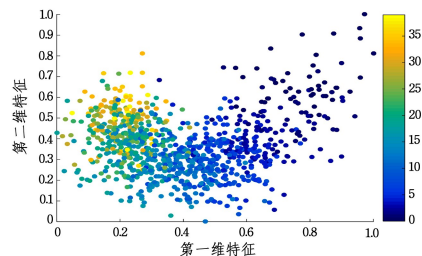
为了验证距离度量调整 ClsRCA 和标签敏感的维数约减 lsMMC 可以学习特征与标签之间的内在关系,分别对提取到的年龄融合特征及经过 ClsRCA 和 lsMMC 处理后的前两维特征进行可视化,结果如图 5 所示。



(a) 年龄融合特征可视化



(b) ClsRCA 处理后的特征可视化



(c) ClsRCA + lsMMC 处理后的特征可视化

图 5 前两维特征的可视化结果图

Fig. 5 Visualization of first two features

可以看出,年龄融合特征分布得比较杂乱,无规律可言。经过距离度量调整 ClsRCA 处理后,同类样本之间相距较近,且标签的有序性得到了初步体现:从左到右,年龄有递增的趋

势,但此时类别之间的重叠较大。经过距离度量调整 ClsRCA 和标签敏感的维数约减 lsMMC 处理后,在保持类内局部结构的同时,标签的有序性得到了进一步体现,可以保证在年龄估计时寻找的近邻样本之间有着相似的年龄标签。

4.5 算法的有效性验证

为了验证 ClsRCA 的必要性及本文所提方法中 lsMMC 和 KNN-LDSVR 两步的有效性,本实验分别对无距离度量调整和距离度量调整(ClsRCA),无维数约减、MMC 和 lsMMC, LDSVR 和 KNN-LDSVR 进行对比实验,结果如表 2 所列。

表 2 FGNET 数据库子集上算法有效性验证的 MAE 结果

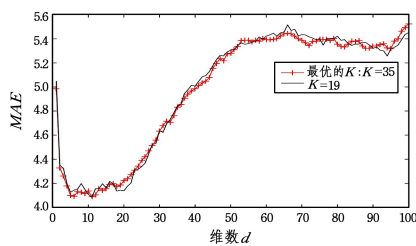
Table 2 MAE results for effectiveness validation of algorithms on FGNET database subset

年龄估计算法	距离度量调整方法	无维数约减	MMC ^[11]	lsMMC
LDSVR ^[14]	无距离度量调整	5.5970	4.6024	4.5723
	ClsRCA ^[5]	5.0311	4.5145	4.4855
KNN-LDSVR	无距离度量调整	5.2765	4.2497	4.2101
	ClsRCA ^[5]	4.9968	4.1479	4.0986

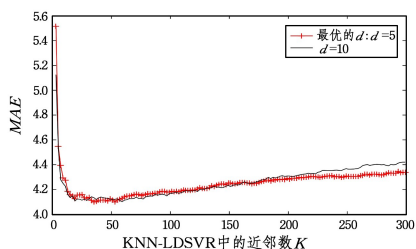
可以看出,对提取的年龄融合特征进行距离度量调整后,再用本文提出的标签敏感的维数约减和两步局部回归方法进行年龄估计的效果最好。

4.6 参数设置

图 6 分析了不同的 lsMMC 的输出维数 d 及 KNN-LDSVR 中的近邻数 K 对实验结果的影响。实验结果表明,当 d 为 5, K 取 35 左右时,效果最好。当 d 大于 5 时,MAE 值急剧增长;当 K 大于 35 时,MAE 持续增长。这说明了使用标签敏感的最大间隔准则进行维数约减的有效性及年龄估计中存在的局部统计特性。



(a) 固定近邻数 K , 改变约减维数 d 的 MAE 结果



(b) 固定约减维数 d , 改变近邻数 K 的 MAE 结果

图 6 本文算法分别在不同的约减维数和近邻数上的 MAE 结果

Fig. 6 MAE results with different choices of dimensionality and neighbors

4.7 性能比较实验

实验设备:CPU 类型为英特尔四核 i7-7700HQ, CPU 速度为 2.8 MHz, 内存容量为 8 GB; 实验环境为 MATLAB 2016

(a)。由于实验采用留一法模式,因此采用 82 轮所需的平均时间来评价算法的时间性能。在训练阶段,训练 ClsRCA 和 lsMMC 的投影矩阵平均需要 21 s;在测试阶段,对于每幅测试样本,使用 KNN 寻找近邻和 LDSVR 估计出最终年龄平均只需 0.03 s。对于年龄估计算法而言,本文算法的运行时间在可接受的范围内。

将本文提出的 lsMMC 与其他 3 种已有的标签敏感维数约减算法进行比较。为了公平起见,维数约减前都用 ClsRCA 进行距离度量调整,维数约减后用 KNN-LDSVR 进行年龄估计,不同年龄段的 MAE 结果如表 3 所列。

表 3 FGNET 数据库子集上不同的标签敏感维数约减算法在各年龄段的 MAE 结果

Table 3 MAE results in each age group of different label-sensitive dimensionality reduction algorithms on FGNET database subset

年龄段	CDDA ^[4]	ClsLPP ^[5]	ClsMFA ^[5]	lsMMC
0~9	3.7817	2.4555	2.6119	2.5310
10~19	3.7050	4.4248	4.0708	4.4307
20~29	7.0486	5.4444	4.9861	4.6875
30~39	16.962	9.1013	10.038	8.9620
0~39	5.3740	4.1951	4.1372	4.0986

从表 1 和表 3 中可以看出,年龄越大,样本数越少,年龄估计的效果越差。因此,缺乏充分且完整的训练数据是人脸年龄估计的难点之一。总的来说,与其他 3 种方法相比,lsMMC 可以取得较好的年龄估计结果。

图 7 对比了上述 4 种标签敏感维数约减方法的 CS 值。可以看出,与 CDDA 相比,lsMMC 能取得较高的 CS 值;与 ClsLPP 和 ClsMFA 相比,lsMMC 能取得相当的 CS 值。

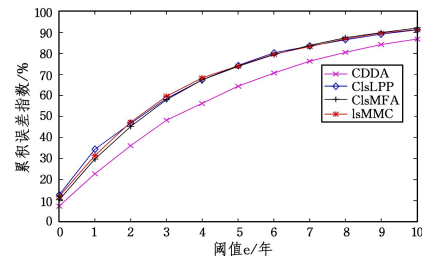


图 7 FGNET 数据库子集上不同的标签敏感维数约减算法的 CS 结果

Fig. 7 CS results of different dimensionality reduction algorithms on FGNET database subset

lsMMC 与 CDDA^[4] 的不同之处在于:1) lsMMC 基于最大间隔准则,而 CDDA 基于费希尔鉴别分析;2) 在计算类内离散度矩阵时,lsMMC 只使用了类别标签相同的样本对,而 CDDA 使用了全部样本对;3) 在计算类间离散度矩阵时,lsMMC 对特征空间中相距较近而年龄标签相差很大的样本赋予较大的权重,希望在投影后的空间中拉开它们之间的距离,此时样本对的权重与特征相似度和标签差异有关,而 CDDA 则认为年龄标签相差很小的样本对更易发生误分,需要分配较大的权重,此时样本对的权重只与标签差异有关。从表 3 的第一列和最后一列可知,相比于 CDDA,本文方法的性能得到了明显改善。

最后,为了进一步验证本文算法的有效性,将它与近年来几种典型的年龄估计算法进行比较(在整个 FGNET 数据库而不是其子集上进行实验),结果如表 4 所列。可以看出,由于考虑了年龄标签之间存在的有序性和相关性,以及特征的局部统计特性,本文方法能取得较好的年龄估计结果。

表 4 FGNET 数据库上不同年龄估计算法的 MAE 结果
Table 4 MAE results of different age estimation algorithms on FGNET database

方法	MAE	论文发表年份/年
KNN	8.24	—
SVM	7.25	—
SVR	5.91	—
AGES ^[18]	6.77	2007
RED-SVM ^[19]	6.49	2010
SVM+SVR ^[20]	5.31	2011
GM ^[21]	5.89	2012
IIS-LLD ^[8]	5.77	2013
HC-SVR ^[22]	5.28	2014
CCA-ELM ^[23]	5.55	2015
SAC-ORLS ^[24]	5.20	2016
本文方法	5.18	—

结束语 本文提出了一种基于标签敏感最大间隔准则的人脸年龄两步估计算法,其考虑了年龄标签的有序性和相关性。对图像提取融合年龄特征后,首先使用 ClsRCA 进行距离度量调整,使得标签相似的样本在特征空间中相距较近;然后使用 lsMMC 进行维数约减,保持类内的局部结构和类间的有序性;最后,考虑特征的局部统计特性,提出一种两步的局部回归算法 KNN-LDSVR 进行年龄估计。从 FGNET 数据库上的实验结果来看,本文提出的年龄估计算法与一些已有的算法相比,可以取得较好的实验结果。

在人脸年龄数据库中,各年龄的图像数差异很大,年龄越大,图像数越少,如何在训练阶段有效处理各年龄样本数不均衡的问题将是我们的下一阶段的工作方向。

参 考 文 献

- [1] GUO G, MU G, FU Y, et al. Human age estimation using bio-inspired features[C]//IEEE International Conference on Computer Vision and Pattern Recognition. 2009;112-119.
- [2] COOTES T F, EDWARDS G J, TAYLOR C J. Active appearance models[C]//European Conference on Computer Vision. 1998;484-498.
- [3] DONG Y, LIU Y, LIAN S. Automatic age estimation based on deep learning algorithm[J]. Neurocomputing, 2016, 187: 4-10.
- [4] MA B, SHAN S, CHEN X, et al. Discriminant analysis for perceptually comparable classes[C]//IEEE International Conference on Automatic Face & Gesture Recognition. 2008;1-6.
- [5] CHAO W L, LIU J Z, DING J J. Facial age estimation based on label-sensitive learning and age-oriented regression[J]. Pattern Recognition, 2013, 46(3): 628-641.
- [6] ONIFADE O F W, AKINYEMI D J. A review on the suitability of machine learning approaches to facial age estimation[J]. International Journal of Modern Education and Computer Science, 2015, 7(12): 17-28.
- [7] PONTES J K, BRITTO A S, FOOKES C, et al. A flexible hierarchical approach for facial age estimation based on multiple features[J]. Pattern Recognition, 2016, 54(C): 34-51.
- [8] GENG X, YIN C, ZHOU Z H. Facial age estimation by learning from label distributions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(10): 2401-2412.
- [9] HE Z, LI X, ZHANG Z, et al. Data-dependent label distribution learning for age estimation[J]. IEEE Transactions on Image Processing, 2017, 26(8): 3846-3858.
- [10] The FG-NET Aging Database [OL]. <http://sting.cyccollege.ac.cy/~lanitis/fgnetaging>.
- [11] LI H, JIANG T, ZHANG K. Efficient and robust feature extraction by maximum margin criterion[J]. IEEE Transactions on Neural Networks, 2006, 17(1): 157-165.
- [12] MIKA S, RATSCH G, WESTON J, et al. Fisher discriminant analysis with kernels[C]//IEEE International Conference on Signal Processing. 1999;41-48.
- [13] SUGIYAMA M. Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis[J]. Journal of Machine Learning Research, 2007, 8(5): 1027-1061.
- [14] GENG X, HOU P. Pre-release prediction of crowd opinion on movies by label distribution learning[C]//International Joint Conference on Artificial Intelligence. 2015;3511-3517.
- [15] GOODALL C. Procrustes methods in the statistical analysis of shape[J]. Journal of the Royal Statistical Society, 1991, 53(2): 285-339.
- [16] LANITIS A, TAYLOR C J, COOTES T F. Toward automatic simulation of aging effects on face images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 442-455.
- [17] GENG X, ZHOU Z H, ZHANG Y, et al. Learning from facial aging patterns for automatic age estimation[C]//ACM International Conference on Multimedia. 2006;307-316.
- [18] GENG X, ZHOU Z H, SMITH-MILES K. Automatic age estimation based on facial aging patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(12): 2234-2240.
- [19] CHANG K Y, CHEN C S, HUNG Y P. A ranking approach for human ages estimation based on face images[C]//IEEE International Conference on Pattern Recognition. 2010;3396-3399.
- [20] CHOI S E, LEE Y J, LEE S J, et al. Age estimation using a hierarchical classifier based on global and local facial features[J]. Pattern Recognition, 2011, 44(6): 1262-1281.
- [21] WU T, TURAGA P, CHELLAPPA R. Age estimation and face verification across aging using landmarks[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(6): 1780-1788.
- [22] LIU J, MA Y, DUAN L, et al. Hybrid constraint SVR for facial age estimation[J]. Signal Processing, 2014, 94(1): 576-582.
- [23] SI J, FENG J, BU Q, et al. Age estimation based on canonical correlation analysis and extreme learning machine[C]//Chinese Conference on Biometric Recognition. 2015;677-685.
- [24] TIAN Q, XUE H, QIAO L. Human age estimation by considering both the ordinality and similarity of ages[J]. Neural Processing Letters, 2016, 43(2): 1-17.