

一种基于图像表观的鲁棒姿态估计方法*

马丙鹏^{1,2,3}, 山世光^{1,2+}, 陈熙霖^{1,2}, 高文^{2,4}

¹(中国科学院 智能信息处理重点实验室, 北京 100190)

²(中国科学院 计算技术研究所, 北京 100190)

³(中国科学院 研究生院, 北京 100049)

⁴(北京大学 信息科学技术学院, 北京 100871)

Robust Appearance-Based Method for Head Pose Estimation

MA Bing-Peng^{1,2,3}, SHAN Shi-Guang^{1,2+}, CHEN Xi-Lin^{1,2}, GAO Wen^{2,4}

¹(Key Laboratory of Intelligent Information Processing, The Chinese Academy of Sciences, Beijing 100190, China)

²(Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)

³(Graduate University, The Chinese Academy of Sciences, Beijing 100049, China)

⁴(School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China)

+ Corresponding author: E-mail: sgshan@jdl.ac.cn

Ma BP, Shan SG, Chen XL, Gao W. Robust appearance-based method for head pose estimation. *Journal of Software*, 2009,20(6):1651–1663. <http://www.jos.org.cn/1000-9825/3372.htm>

Abstract: This paper proposes a new pose estimation method based on the appearance of 2D head image. First, the 1D Gabor filters are used to extract the features on the raw images. Compared with the traditional 2D Gabor represents, the 1D Gabor represents are more closely related to the head pose, while the advantages of computation and storage are obvious. Second, for the extracted features, a new method, named kernel local fisher discriminant analysis, is applied to eliminate the multimodal problem, while at the same time enhance the discrimination ability. Experimental results show that the proposed method is effective for pose estimation. It must be pointed out that the generalizability of the proposed method is illustrated by the impressive performance when the training dataset and the testing dataset are heterogeneous.

Key words: head pose estimation; Gabor transform; local Fisher discriminant analysis

摘要: 提出一种利用图像的表现特征进行头部姿态估计的方法.该方法首先使用了一维 Gabor 滤波器对头部图像进行特征提取,然后对提取得到的一维 Gabor 特征进一步使用了基于核函数的局部费舍尔判别分析方法增强特征的判别能力.与传统二维 Gabor 特征相比,一维 Gabor 特征除了在计算速度和存储空间上具有明显的优势以外,更与姿态紧密相关.而基于核函数的局部费舍尔判别分析方法,能够解决姿态问题中存在的非线性问题和多模态问题.

*Supported by the National Natural Science Foundation of China under Grant No.60673091 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant Nos.2006AA01Z122, 2007AA01Z163 (国家高技术研究发展计划(863)); the 100 Talents Program of the CAS (中国科学院百人计划); the ISVISION Technology Co. Ltd. (上海银晨智能识别科技有限公司)

Received 2007-08-02; Revised 2007-11-19; Accepted 2008-04-15

大量的实验结果表明,该算法对于姿态估计问题是有效的.特别需要指出的是,该算法具有良好的推广能力,在训练数据和测试数据异质时,该算法的性能明显高于其他对比算法的性能.

关键词: 头部姿态估计;Gabor 变换;局部费舍尔判别分析

中图法分类号: TP391 文献标识码: A

统计表明,大约 75%的人脸图像为非正面的人脸图像^[1].对于人脸识别及其相关问题来说,姿态估计是一个重要的前处理过程.在过去的十几年里,虽然人脸识别问题的研究取得了巨大的进步,但是在姿态变化时进行鲁棒的人脸识别仍然十分困难.由于准确估计输入图像的姿态是解决这些问题的先决条件,因而姿态估计受到了越来越多的关注.

人脸姿态估计是指从输入的图像中确定人脸在三维空间中姿态的过程.头部的旋转通常分成 3 种类型:左右旋转(yaw rotation)、上下旋转(pitch rotation)和平面内旋转(roll rotation).由于平面内旋转可以根据眼睛的相对位置计算得到,因而姿态估计方法通常集中在估计头部的左右旋转和上下旋转上.姿态估计的方法通常可以分成两种类型^[2]:基于模型的方法(model-based method)^[2-6]和基于表观的方法(appearance-based method)^[7-14].

基于模型的方法利用某种几何模型或结构表示人脸的结构和形状,并通过提取某些特征,在模型和图像之间建立起对应关系,然后通过几何或者其他方法实现姿态的估计.Nikolaidis 等人提出通过眼睛和嘴之间建立的等腰三角形来估计头部姿态^[3].Ji 等人提出,2D 图像中的头部可以用椭圆来近似,而头部姿态可以通过椭圆的纵横比来计算得到^[2].Xiao 等人利用圆柱表示 3D 的头部,并根据透视投影来估计头部的姿态^[4].基于模型的方法有一些共同的缺点:首先,简单的几何模型或结构不能准确地描述人脸的真实形状和特征;其次,基于模型的方法往往对面部特征点的定位敏感,而准确的面部特征点定位仍然是一个尚未解决的问题;最后,基于模型的方法对图像的质量要求较高,很多实际应用中的图像难以达到要求.

由于基于表观的方法能够克服基于模型方法的这些缺点,因而越来越受到重视.基于表观的方法假定在 3D 的头部姿态和 2D 人脸图像的某些性质之间存在一定的联系,并且认为通过大规模训练集上的学习可以找到这种联系.基于表观的姿态估计方法通常使用统计学习理论提取特征,例如支持向量机^[7]、核主成分分析^[8]、独立主成分分析(independent component analysis,简称 ICA)^[9]、多视角特征空间^[10]、最优 Gabor 滤波器的特征空间^[11]、流形学习^[12]和弹性图匹配^[13]等等.其中,Gabor 特征的弹性图匹配和最优 Gabor 滤波器方法均取得了不错的实验结果.弹性图匹配方法对头部图像中部分关键特征点进行 Gabor 变换,并将头部描述为以这些特征点位置为顶点,以其 Gabor 变换系数为顶点属性,以其关键点位置关系为边属性的属性图,从而将姿态问题转化为图匹配问题.而基于 Gabor 滤波器选择的方法,通过研究认为,不同姿态的头部图像对 Gabor 频率的响应不同,因此对不同姿态选用不同频率的 Gabor 特征建立特征空间,根据测试样本在特征空间的投影来确定输入图像的最终姿态类别.这两种方法的共同点在于均将 Gabor 滤波器作为主要的特征提取方法.除了在姿态估计中以外,近年来在人脸识别领域中,对图像灰度进行多尺度、多方向的 Gabor 小波变换也逐渐成为主流的思路之一.Gabor 小波之所以在多个领域均受到重视,主要是因为 Gabor 小波可以很好地模拟大脑皮层中单细胞感受野的轮廓,捕捉突出的视觉属性,例如空间定位、方向选择等.特别是 Gabor 小波可以提取图像特定区域内的多尺度、多方向空间频率特征,像显微镜一样放大灰度的变化,这样,人脸图像中的眼睛、鼻子和嘴以及其他一些局部特征被放大.因此,采用 Gabor 变换来处理姿态图像可以增强一些关键特征来提高分类不同姿态图像时的准确率.虽然这些基于 Gabor 特征的方法取得了不错的效果,但是它们也都有共同的困难.一般来说,基于 Gabor 的方法在对图像进行多尺度、多方向的变换时,极大地增加了计算量,从而降低了算法的速度.一般来说,由于头部姿态估计是人脸识别系统或其他应用的一个中间步骤,为人脸识别或者 3D 重建提供有效的信息,因此姿态估计的精度固然重要,但是姿态估计的速度也同样非常重要.因此,虽然弹性图匹配方法取得了较好的实验结果,但并没有成为姿态估计的主流技术.虽然基于最优 Gabor 滤波器的方法可以根据不同的姿态选用不同的 Gabor 滤波器来减少计算量,但在连续的姿态空间中选择每个姿态所合适的 Gabor 滤波器,同样是一个困难的问题.因此,如何在姿态估计问题中快速、有效地利用 Gabor 特征来提高准确率,是一个困难而又非常重要的问题.

除了特征提取以外,如何正确认识姿态样本的分布也是姿态估计中的一个重要问题.一般认为,姿态样本存在于高度非线性的流形空间中.因而传统的线性技术,如主成分分析(principal component analysis,简称 PCA)^[15]、线性费舍尔判别分析(fisher discriminant analysis,简称 FDA)^[16]等,无法直接在姿态问题中取得理想的实验结果.为了解决姿态估计中的非线性问题,目前研究者们提出了两类方法:一种是原有线性技术的非线性化,如基于核函数的主成分分析(KPCA)^[17]和基于核函数的费舍尔判别分析(KFDA)^[18],被用于姿态估计中;另一种为基于流形学习的算法,如局部线性嵌套^[19]、等度规映射^[20]、拉普拉斯特征图^[21]和局部保持投影(locality preserving projections,简称 LPP)^[22]等.拉普拉斯特征图提出了利用局部保持性的问题,文献[23]分析了拉普拉斯特征图与其他线性判别分析方法之间的关系.虽然这些方法能够解决姿态估计中的非线性问题,但是它们忽略了多姿态问题中的另一个问题,即多模态问题.一般来说,多模态是指每类的样本分布在多个簇中,不同类的簇之间有重叠的区域.在姿态估计问题中,由于头部的旋转同时有上下、左右、平面内旋转 3 种情况.具有相同上下旋转的图像,可能具有不同的左右旋转和平面内旋转.在单独识别一种旋转时,具有相同的左右(上下)旋转的图像,其上下(左右)旋转的角度很可能会有较大的差别.因此,姿态估计问题是典型的多模态问题.只有解决了多模态问题,才能进一步提高姿态估计的最终性能.

为了解决姿态估计中的多模态问题并提高特征提取速度,本文提出了基于 1D Gabor 特征的局部判别方法.该方法具有以下两个方面的优点:一方面,在特征提取时,使用了 1D Gabor 滤波器做特征提取的方法.与其他基于 2D Gabor 特征提取的方法相比,该方法在具有计算量小这一优点的同时,能够使得特征与姿态旋转的角度更紧密地相关,从而提高姿态估计的性能.另一方面,通过使用基于核函数的局部费舍尔判别分析方法增强了特征的判别能力,从而能够解决姿态估计问题中存在的多模态问题和非线性问题.

为了检验提出算法的有效性,本文使用了简单的最近质心分类器(nearest centroid,简称 NC)在两个数据库上进行了两种实验.在第 1 种实验中,训练数据和测试数据均来自同一个数据库.在这种实验条件下,本算法取得了最好的实验结果.在第 2 种实验中,训练数据和测试数据来自不同的数据库.一般认为,这种条件下的实验更能检验方法的推广性.在这种实验条件下,本文提出的方法具有更加明显的优势,验证了本文提出的方法具有良好的推广性.

1 基于 1D Gabor 特征的局部判别方法

由于 1D Gabor 特征在较低的图像质量时具有较强的表述能力,而局部费舍尔判别分析能够解决姿态估计中的多模态问题和非线性问题,因此,通过将 1D 的 Gabor 特征和局部费舍尔判别分析相结合,本文提出一种新的姿态估计方法:基于 1D Gabor 特征的局部判别方法(Gabor based local discriminant method,简称 GLDM),算法 1 给出了该方法的计算流程.

算法 1. 基于 1D Gabor 特征的局部判别方法的流程.

输入:2D 的图像矩阵 $\mathbf{X}_i = (x_{i1}, x_{i2}, \dots, x_{iw}, \dots, x_{iw})$, 图像的尺寸为 $w \times h$.

1. 将图像的每一行当作一个向量.此时 $\mathbf{X}_i = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_h)$, 其中 $\mathbf{x}_j = (x_{j1}, x_{j2}, \dots, x_{jw})$.
2. 对 \mathbf{x}_j 提取 1D Gabor 特征,得到 $\mathbf{g}_j = (g_{j1}, g_{j2}, \dots, g_{jw})$.
3. 将图像所有的 1D 特征 $\{\mathbf{g}_j\}$ 向量化,得到图像的新 $\mathbf{G}_i = (\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_h)$.
4. 利用核函数计算 \mathbf{G}_i 和训练样本的特征 \mathbf{G}_i^{train} 在高维空间中的度量:

$$\mathbf{Gk}_i = (K(\mathbf{G}_i, \mathbf{G}_1^{train}), K(\mathbf{G}_i, \mathbf{G}_2^{train}), \dots, K(\mathbf{G}_i, \mathbf{G}_n^{train})).$$

5. 根据局部费舍尔判别分析计算得到投影矩阵 \mathbf{T}_{LFDA} .
6. 根据 \mathbf{T}_{LFDA} 得到最终的特征 $\mathbf{X}'_i = \mathbf{T}_{LFDA} \times \mathbf{Gk}_i$.
7. 将 \mathbf{X}'_i 送入分类器,得到样本估计的类别 y .

输出: y .

1.1 1D Gabor特征表述

1946年, Gabor 提出了可以到达时频测不准关系下界的 Gabor 函数, 它实际上是高斯函数在频域中的平移. Gabor 函数是唯一能够到达测不准关系下界的函数, 它能够很好地在时域和频域中兼顾对信号分析的分辨率要求. Gabor 函数这一特性使其在信号处理中得到了广泛的应用. 2D Gabor 函数较好地描述了哺乳动物初级视觉系统中一对简单视觉神经源的感受特性. Daugman 在 20 世纪 80 年代最早将其用于计算机视觉领域. 在本文中用到的 1D Gabor 滤波器如公式(1)所示:

$$g_{\mu}(r) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{r^2}{2\sigma^2}} e^{i(2\pi\mu r)} \quad (1)$$

其中, μ 是调制频率, σ 是决定高斯窗宽度的尺度参数. 1D 信号的 Gabor 表述是由 1D 信号和 1D Gabor 滤波器族的卷积得到. 设 $s(r)$ 是图像的灰度信号, 则该信号和 Gabor 滤波器的卷积按照公式(2)来定义:

$$O_{\mu}(r) = s(r) * g_{\mu}(r) \quad (2)$$

其中 * 是卷积符号, $O_{\mu}(r)$ 是在频率 μ 下的卷积结果.

基于 1D Gabor 特征的局部判别方法的第 1 步是对图像每行上的灰度特征上进行 1D Gabor 特征提取. 为了提高计算速度, 可以依据卷积理论通过快速傅里叶变换和逆傅里叶变换完成信号和 Gabor 核函数的卷积:

$$F\{O_{\mu}(x_j)\} = F\{s(x_j) * g_{\mu}(x_j)\} = F\{s(x_j)\} F\{g_{\mu}(x_j)\} \quad (3)$$

根据信号理论, 对于 1D 信号来说, 对于经过傅里叶变换后得到的复信号, 其低频部分包含信号的主要信息, 高频部分通常是信号的细节和噪声. 对于头部图像来说, 傅里叶变换后的细节通常与输入图像的人的身份相关, 而与姿态无关. 因此, 我们能够安全地去掉 $F\{s(x_j)\}$ 的高频部分, 这既降低了特征的维数, 又减少了与姿态无关的信息, 从而提高姿态估计的准确度. 在实际操作中, 我们保留了 $F\{s(x_j)\}$ 整体能量的 97% 的低频部分. 在对 $F\{O_{\mu}(x_j)\}$ 降维后, 再利用逆傅里叶变换得到信号的最终 Gabor 特征 $O_{\mu}(x_j)$:

$$O_{\mu}(x_j) = F^{-1}\{F\{O_{\mu}(x_j)\}\} = F^{-1}\{F\{s(x_j)\} F\{g_{\mu}(x_j)\}\} \quad (4)$$

本文使用 1D Gabor 特征的原因, 除了如前所述 Gabor 变换的优良特性外, 还在于与 2D Gabor 特征相比, 1D Gabor 特征在姿态估计问题中更具优势. 虽然在人脸识别中 2D Gabor 滤波器得到了广泛的应用, 并取得了不错的实验结果, 但是我们认为 1D Gabor 滤波器更适用于姿态估计问题. 首先, 与 2D 的图像块相比, 1D 的行信号更与姿态紧密相关, 因此使用 1D Gabor 滤波器是有理论依据的. 为了说明这个问题, 在图 1 中, 画出了对姿态估计具有重要意义两条线. 其中, 间隔线为 2D 图像的中心线, 连续线为 3D 头部的对称面和面部表面的交线. 由于在切割人脸图像时固定了眼睛在图像中的位置, 因此图像的中心线和两个眼睛之间的中心线在 2D 图像上是重合的. 而在 3D 空间中, 交线通过鼻尖. 从图 1 中可以很容易地看出这两条线随姿态变换的关系. 在正面图像时, 两条线是重合的. 随着姿态从正面变化到半侧面, 两条线之间的间隔越来越大. 两条线之间的间隔是头部从 3D 到 2D 的投影时造成的, 并与姿态紧密相关. 这种间隔可以通过垂直于两条垂直线的水平方向上的灰度信号来反映. 从信号的观点来看, 如果将正面姿态时的人脸图像的一行灰度值当成信号, 当姿态从正面变换到侧面时, 信号的一半在空域上被放大, 而另外一半被压缩. 而这种放大或者压缩的程度是与姿态紧密相关的. 因此将图像的每一行当成信号, 提取与姿态紧密相关的特征对姿态估计问题是具有意义的.



Fig.1 Relationship between pose variations, the head across planes and the image center lines

图 1 头部的交平面、图像中心线和姿态变化的关系

其次, 1D Gabor 滤波器与 2D Gabor 滤波器相比具有明显的存储空间和计算时间上的优势. 考虑到 2D Gabor 滤波器通常有 5 个尺度、8 个方向共 40 个滤波器, 而 1D Gabor 滤波器只有 5 个不同尺度的 5 个滤波器, 因此在计算时间和存储空间上, 1D Gabor 特征均具有明显的优势. 在实际操作中, 由于仅仅保留了信号在频域空间中的整体能量的 97% 的低频部分, 因此特征的维数进一步降低. 统计结果表明, 当 1D 信号为 32(64) 维时, 大约 9(16) 个

最低频率的能量就能达到整体能量的 97%.此时,对于 32×32 (64×64) 的图像来说,提取的 1D Gabor 特征的维数仅为 1440(5120).与 2D Gabor 特征相比,维数下降了 28(30)倍.

最后,1D Gabor 特征虽然表述细节的能力要差于 2D Gabor 特征,但是这种劣势在低分辨率、低质量的图像上并不明显.具有特定方向参数的 2D Gabor 滤波器仅能捕捉特定方向上的纹理变化和局部结构.由于不能确定图像中纹理的方向,因此在实际实验中需要使用多个方向的 2D Gabor 变换,才能较好地建模图像的局部纹理变化.而局部的纹理变化对于正确的人脸识别是非常重要的,因此,2D Gabor 在人脸识别中得到了广泛应用,并在姿态估计中也得到了一定的应用.对于 1D Gabor 来说,其对局部纹理的提取要明显差于 2D Gabor.因此,在人脸识别和姿态估计中,1D Gabor 并没有得到广泛关注.但在实际的姿态估计问题中,由于使用的头部图像多为 20×20 或者 32×32 的低分辨率图像或者低质量的图像,此时,与 2D Gabor 特征相比,1D Gabor 特征对于局部纹理的表述其劣势并不明显.

考虑到与 2D Gabor 滤波器相比的巨大优势,本文使用了 1D Gabor 滤波器对头部图像进行特征提取.这种特征虽然对左、右方向姿态估计更有意义,但它们并不局限于用来估计左、右方向姿态.事实上,由于 Gabor 滤波与原始图像的灰度特征相比,在一定程度上可以消除光照和噪声的影响,因此也能提高上下方向姿态估计的准确率.

1.2 局部费舍尔判别分析

局部费舍尔判别分析方法(local fisher discriminant analysis,简称 LFDA)^[24]是将局部保持投影与线性费舍尔判别分析有效结合的降维方法.其中,线性费舍尔判别分析是有效的线性监督学习的降维方法,其主要思想是寻找一个使得类间散度矩阵最大化和类内散度最小化的变换.局部保持投影是 He 和 Niyogi 提出的一种非常有效的降维方法^[22],是 Laplace Beltrami 算子特征函数的一个线性估计,其目标是保持数据之间的相似关系,即在原始空间中相邻的数据点在投影空间中也保持相应的邻近关系.虽然线性费舍尔判别分析和局部保持投影都是有效的降维方法,但是它们都存在一定的缺点.具体来说,线性费舍尔判别分析在处理样本为多个子簇时效果较差.而对于局部保持投影来说,由于其无监督性,对于在原始高维空间中接近但类别不同的样本,在局部保持投影的投影空间中可能重合.为了说明它们的局限性,我们在理想数据上进行了实验,如图 2 所示.图中 2D 的两类样本嵌入在 1D 空间中.图中用“+”号标记的样本为第 1 类样本,用“o”号标记的为第 2 类样本.在图 2 的左图中,第 2 类样本具有明显的多模态问题.而在右图中,两类样本均具有明显的多模态问题.图 2 分别给出了线性费舍尔判别分析和局部保持投影的投影方向.可以看出,无论线性费舍尔判别分析还是局部保持投影均不能对样本进行正确的投影.

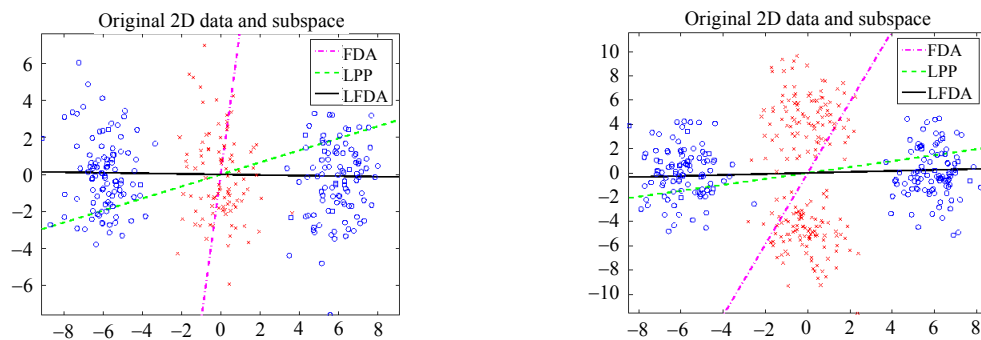


Fig.2 The projection of FDA, LPP and LFDA

图 2 FDA,LPP 和 LFDA 的投影示意

为了克服线性费舍尔判别分析和局部保持投影的局限性,Sugiyama 提出了局部费舍尔判别分析方法.局部费舍尔判别分析方法通过在局部方式下重新计算类内散度和类间散度,得到了最大的类间可分性和类内局部保持性.在局部费舍尔判别分析中,局部保持投影寻找一个嵌入变换,该变换使得在嵌入空间和原始空间中的数

据邻域关系接近,这样,数据的局部结构就能得到保持.设 $\mathbf{x}_i \in R^d (i=1,2,\dots,n)$ 为 d 维空间的样本, $y_i \in \{1,2,\dots,c\}$ 为样本所对应的类别标号, n 为样本的总数, c 为类别的总数, n_i 为属于类 l 的样本的个数.在局部费舍尔判别分析中,局部类内散度矩阵 $\tilde{\mathbf{S}}^w$ 和局部类间散度矩阵 $\tilde{\mathbf{S}}^b$ 分别通过公式(5)和公式(6)进行计算:

$$\tilde{\mathbf{S}}^w \equiv \frac{1}{2} \sum_{i,j=1}^n \tilde{W}_{i,j}^w (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T \quad (5)$$

$$\tilde{\mathbf{S}}^b \equiv \frac{1}{2} \sum_{i,j=1}^n \tilde{W}_{i,j}^b (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T \quad (6)$$

其中,

$$\tilde{W}_{i,j}^w = \begin{cases} A_{ij}/n_i, & \text{if } y_i = y_j = l \\ 0, & \text{if } y_i \neq y_j \end{cases} \quad (7)$$

$$\tilde{W}_{i,j}^b = \begin{cases} A_{ij}(1/n - 1/n_i), & \text{if } y_i = y_j = l \\ 1/n, & \text{if } y_i \neq y_j \end{cases} \quad (8)$$

这里, $A_{i,j}$ 为用于说明同类的样本 \mathbf{x}_i 和 \mathbf{x}_j 组成的样本对对 $\tilde{\mathbf{S}}^w$ 和 $\tilde{\mathbf{S}}^b$ 影响的权重.对于同类中的两个样本,它们之间的距离越小, $A_{i,j}$ 越大,它们组成的样本对对于 $\tilde{\mathbf{S}}^w$ 和 $\tilde{\mathbf{S}}^b$ 的影响也就越大.在局部费舍尔判别分析中, $A_{i,j}$ 可以采取多种定义方式.在本文中,我们采用了较为常用的方式:

$$A_{i,j} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\tilde{\sigma}_i \tilde{\sigma}_j}\right) \quad (9)$$

这里, $\tilde{\sigma}_i$ 是样本 \mathbf{x}_i 的局部标度,设 $\mathbf{x}_i^{(k)}$ 是样本 \mathbf{x}_i 的第 k 个邻域点,则 $\tilde{\sigma}_i$ 定义为

$$\tilde{\sigma}_i = \|\mathbf{x}_i - \mathbf{x}_i^{(k)}\| \quad (10)$$

最终,局部变换矩阵 T_{LFDA} 可以通过求解如下的优化问题得到:

$$T_{LFDA} \equiv \operatorname{argmax}_{T \in R^{d \times r}} \left[\operatorname{tr} \left((T^T \tilde{\mathbf{S}}^w T)^{-1} T^T \tilde{\mathbf{S}}^b T \right) \right] \quad (11)$$

关于局部费舍尔判别分析的详细介绍可以参见文献[24].图2中给出了相同数据下的局部费舍尔判别分析的投影方向.从图中可以看出,局部费舍尔判别分析的方法能够很好地解决多模态问题.

在基于1D Gabor特征的局部判别方法中,在提取图像的1D Gabor特征之后,使用了局部费舍尔判别分析的方法解决姿态估计中的多模态问题.多模态问题是影响姿态估计中的一个重要问题.一般来说,多模态是指每类的样本分布在多个簇中,不同类的簇之间有重叠的区域.例如在图3中,相同列上的图像均具有相同的左右旋转角度,而相同行上的图像均具有相同上下旋转的角度.在进行估计头部的左右旋转时,相同列上的图像由于具有相同的左右旋转角度而被视为相同的一类,但是,由于上下旋转的角度有较大的差异,因此同类之间的样本具有较大的类内差异.同时,对于相同行上的相邻两个样本,虽然它们因为具有不同的左右旋转角度而被视作不同的类,但它们之间的差异较小,甚至小于它们与同类样本之间的差异.因此可以认为,姿态估计中存在较为明显的多模态问题.而根据前述的局部费舍尔判别分析的介绍可知,局部费舍尔判别分析能够很好地解决多模态问题,因此在本文中,在提取图像的1D Gabor特征之后,进一步使用局部费舍尔判别分析来提高特征的判别能力.

虽然局部费舍尔判别分析能够有效地解决姿态估计中的多模态问题,但是由于局部费舍尔判别分析是线性变换的技术,而姿态图像是存在于高度非线性的流形空间中,因此,直接将局部费舍尔判别分析技术应用到姿态估计中并不能得到很好的效果.为了解决姿态估计中的非线性问题,本文中使用了基于kernel核函数的局部费舍尔判别分析方法.由于传统的高斯核函数在许多领域中得到了应用,并且在文献[24]中也使用了高斯核函数,因此在本文中,我们也使用了高斯核函数:

$$K(\mathbf{G}_i, \mathbf{G}_j) = \exp\left(-\frac{\|\mathbf{G}_i - \mathbf{G}_j\|^2}{2\gamma^2}\right) \quad (12)$$

其中, $\gamma > 0$.在实际应用中,核函数也可以选择其他形式.由于核函数的参数 γ 的选择对于特征的可分性有着重要的

影响,而选择最优的 γ 也是一个非常困难的问题,因此为了解决该问题,本文利用基于高斯核函数的 SVM 分类器在小数据集(从训练样本中随机抽取)上进行实验,并使用最优识别率时的 γ 值作为最终的 γ 取值.在本文中, γ 取值为 2^{-10} .

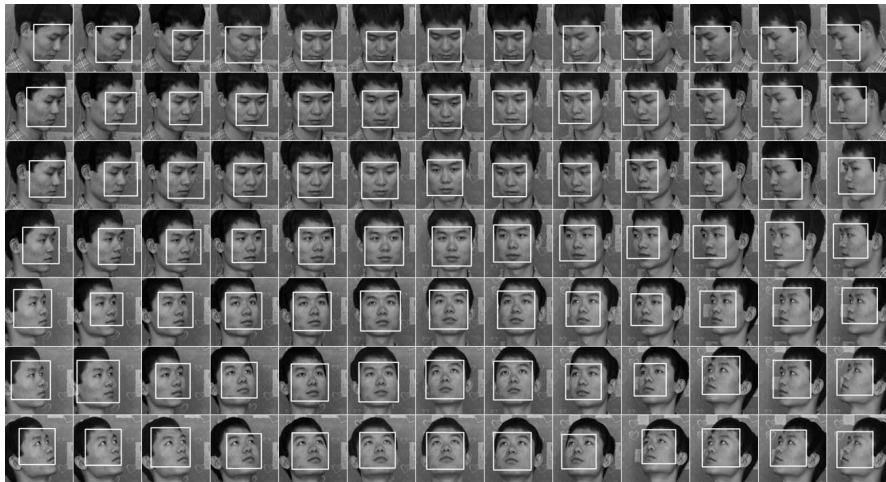


Fig.3 Images in the multi-poses database and the results of face detection

图3 多姿态库中的人脸图像和人脸检测结果示意

1.3 在姿态估计系统中的应用

由于本文的方法是特征提取的方法,不涉及分类器设计,因此本文将姿态估计问题看作是姿态分类问题,并使用了最简单的最近质心分类器来验证算法的性能.对于每类的训练样本,本文使用 k 均值聚类算法得到该类的 k 个中心.对于测试样本,根据与其距离最近的一个中心的类别作为该测试样本的最终类别.

在实际的姿态估计系统中,为了得到连续的姿态角度,可以很容易地用支持向量回归(support vector regression)^[25]、相关向量机(relevance vector machine)^[26]等具有回归能力的算法代替最近质心分类器.

2 实 验

为了验证基于 1D Gabor 特征的局部判别分析方法在姿态估计问题中的有效性,本节将其与以下算法进行了对比,它们是主成分分析(PCA)、线性判别分析(FDA)、独立主成分分析(ICA)、基于 2D Gabor 特征的方法(2D Gabor)、基于 2D Gabor 特征的判别分析方法(gabor fisher classifier,简称 GFC)、基于 2D Gabor 特征的局部判别分析方法(2D Gabor_LFDA)、基于 1D Gabor 特征的方法(1D Gabor)和基于 1D Gabor 特征的判别分析方法(1D Gabor_FDA).像在人脸识别中一样,PCA 也被认为是基于表观的姿态估计方法中的基准算法.由于最近 ICA 在姿态估计中取得了不错的实验结果^[9],本文使用 ICAFace^[27]的 ICA 代码进行了结果对比.为了说明 1D Gabor 特征的有效性,现将我们的算法与多个使用 2D Gabor 特征的方法,如 2D Gabor,GFC 和 2D Gabor_LFDA 进行对比.在这些算法中,我们使用了 5 个尺度、8 个方向,共 40 个 Gabor 滤波器来提取 2D Gabor 特征.为了说明局部判别分析的效果,我们与同样使用 1D Gabor 特征的算法 1D Gabor_FDA 进行了对比.

为了减少计算量,在 PCA,FDA,ICA 等方法中,本文首先使用 PCA 的方法对图像向量进行降维,然后在降维后的向量上用不同的方法进行特征提取.对于使用 1D Gabor 特征和 2D Gabor 特征的方法,本文对图像进行 Gabor 特征提取后再对新的特征进行 PCA 降维.考虑到 2D Gabor 特征的巨大维数,为了能够有效地降低存储空间,我们对 2D Gabor 特征进行了 4×4 邻域上的下采样.在所有使用 PCA 的方法中,均保留了特征值能量的 95%.由于 FDA 中有降维后的特征维数小于类别数减 1 的限制,虽然在局部费舍尔判别分析中没有维数的限制,但为了和 FDA 有可比性,所有方法提取的特征的维数均为类别数减 1.例如,在 CAS-PEAL 数据库上进行实验时,特

征的维数均为 $6(=7-1)$, 在多姿态库上进行实验时, 特征的维数为 $100(=101-1)$.

本文中进行了两种类型的实验. 在第 1 种类型的实验中, 训练数据和测试数据均来自同一个数据库, 此时可以认为训练数据和测试数据同质. 在第 2 种类型的实验中, 训练数据和测试数据来自不同的人脸库, 此时训练数据和测试数据异质, 从而能够更好地检验算法的推广性能.

2.1 数据库

本文使用了两个数据库来检验算法的性能. 第 1 个数据库为公开的 CAS-PEAL 数据库^[28]. 在 CAS-PEAL 库中, 每人共有 21 张图像, 其中头部的俯仰旋转包含上、中、下 3 种情况, 在水平方向上姿态包含 $-45^\circ, -30^\circ, -15^\circ, 0^\circ, 15^\circ, 30^\circ$ 和 45° 共 7 种情况. 由于俯仰旋转有 3 种不同的情况, 因此这相当于在估计水平方向上的姿态时, 每个类别均有 3 个子中心, 因此在利用 CAS-PEAL 库进行姿态估计时, 存在多模态问题. 为了减少计算量, 本文使用了 CAS-PEAL 库的一个子集, 该子集包含身份标号从 401~600 的 200 个人全部 4200 张图像. 在图 4 中给出了 CAS-PEAL 数据库中一些人的示例图像.



Fig.4 Images in the CAS-PEAL database

图 4 CAS-PEAL 数据库图像示意

与 CAS-PEAL 数据库离散的姿态不同, 我们采集了一个具有连续深度旋转的多姿态数据库(multi-pose database). 该数据库中包含 102 个人的 7 731 张图像. 为了使图像之间能够尽量仅有姿态的变化, 在采集数据时我们控制了光照条件和背景变化. 在多姿态库中, 头部左右旋转的角度以 1° 间隔的方式从 -90° 连续变化到 $+90^\circ$, 同时, 俯仰角度从 -45° 连续变换到 $+45^\circ$. 由于采集条件的限制, 每个人的样本数略有不同. 为了减少计算量, 本文使用了多姿态库中的一个子集. 该子集的图像是从原始的多姿态库中随机抽取得到. 为了使子集中的图像更具一般性, 在随机抽取时作了如下限制: 1) 左右方向上的姿态角度限制在 $-50^\circ \sim +50^\circ$ 之间; 2) 水平方向上的任意姿态的样本个数均为 30; 3) 每个人贡献的样本个数大致相等. 图 3 为数据库中某个人的一些图像组成的示例图像.

对于所有的人脸图像, 本文均使用了全自动人脸检测技术来定位人脸区域的位置, 并将检测到的人脸区域作为姿态估计方法的输入图像. 在这种情况下, 姿态估计方法可以看作是全自动姿态估计方法, 并能嵌入到全自动的姿态估计系统中. 图 3 中给出了人脸检测算法在多姿态库上的检测结果. 在实验中, 自动检测到的人脸区域均被缩小到 32×32 , 同时使用了直方图均衡化技术以消除光照的影响.

2.2 实验1: 训练集和测试集同质

本文首先在训练集和测试集均来自相同的数据库的情况下进行了实验. 此时训练图像和测试图像是同质的. 在实验中, 为了避免过学习的问题, 本文使用了 3 份的交叉验证技术, 即将数据库中的全部数据根据被采集人的身份标号分成 3 份, 在实验中, 选取其中的 2 份作训练集, 另外 1 份作测试集, 这样, 训练集和测试集中的人物没有重叠. 重复实验 3 次, 这样每份均曾作为测试集. 最终实验结果为所有测试集上结果的均值.

2.2.1 CAS-PEAL 数据库上的实验结果

首先在 CAS-PEAL 数据库上进行了实验. 对于最近质心分类器, 使用了欧式距离(Euclidean distance)和余弦距离(cosine distance)两种距离度量方式, 实验结果在表 1 中给出. 从表中可以看出, 首先, 基于 1D Gabor 特征的局部判别方法的性能要远好于其他方法的性能. 其次, 对于各种方法, 两种不同的距离度量方式的实验结果相似. 因此, 在接下来的实验中, 均只使用了欧式距离这一种距离度量方式.

在实验中, 每类的训练样本数大概有 $400(=600/3 \times 2)$ 个. 我们认为, 对某些算法来讲, 由于并不能解决多模态

问题,此时仅有 1 个类中心不能很好地代表该类样本的分布,因而结果较差.而随着类中心个数的增加,这些算法的性能能够得到提升.为了验证该想法,本文使用了 k 均值聚类算法,使得每类的样本有 k 个类中心,然后使用最近质心分类器对测试样本进行分类.在不同类中心的个数下,其识别率如图 5 所示.

Table 1 Accuracy on the CAS-PEAL database (%)

表 1 CAS-PEAL 数据库上的准确率(%)

	PCA	ICA	FDA	2D Gabor	GFC	2D Gabor_LFDA	1D Gabor	1D Gabor_FDA	GLDM
Euclidean	65.06	83.35	86.38	64.97	88.66	90.80	79.65	91.42	93.46
Cosine	64.51	83.66	86.35	65.26	88.66	88.71	79.43	91.01	93.54

从图 5 中我们可以看到如下几点: (1) 对于不同类中心的个数,本文提出的基于 1D Gabor 的局部判别方法均能够取得最好的实验结果.这说明我们提出的算法对于姿态估计问题是有效的; (2) 在姿态估计中,1D Gabor 特征要优于 2D Gabor 特征.从图中可以看出,1D Gabor,1D Gabor_FDA 和 GLDM 的实验结果分别优于 2D Gabor,GFC 和 2D Gabor_LFDA.这说明,在姿态估计中,1D Gabor 特征能够与姿态紧密相关,从而提高姿态估计的准确率; (3) GLDM 算法的结果要优于 1D Gabor_FDA 算法的结果.这说明通过使用局部费舍尔判别分析算法能够有效地解决姿态问题中的多模态问题; (4) 对于不同的类中心个数,基于 1D Gabor 的局部判别方法的结果基本保持稳定.这说明,基于 1D Gabor 的局部判别方法能够很好地将多模态的样本有效地降维; (5) 对于 PCA,ICA,1D Gabor 和 2D Gabor 等算法,当类中心的个数较小时,随着类中心个数的增加,分类准确率有较为明显的上升.而对于 FDA 和 GFC 算法,随着类中心个数的增加,分类准确率略有下降.这说明这些算法均不能很好地处理多模态问题,造成每类的样本在低维空间投影后,仍然较为分散; (6) 2D Gabor 特征的结果甚至要差于 PCA 的结果,我们认为其原因在于,一方面,图像分辨率较低,另一方面,为了降低 2D Gabor 特征的维数进行了下采样.

2.2.2 多姿态数据库上的实验结果

类似地,本文在多姿态库上测试了基于 1D Gabor 的局部判别方法的性能.较之姿态分类的准确率,在实际问题中,往往更关心图像的真实姿态和预测得到的姿态的误差值.由于多姿态库可以看作是具有近似连续姿态的数据库,因此我们在多姿态库上使用了误差均值,而不是分类准确率作为衡量方法性能指标.真实姿态和预测姿态之间的误差均值 m 可以通过公式(13)计算得到:

$$m = \frac{1}{N} \sum_{i=1}^N |p'_i - p_i| \tag{13}$$

其中, p_i 和 p'_i 分别是第 i 个样本的真实姿态和预测得到的姿态, N 为总样本的个数.在图 6 中给出了不同的算法在不同类中心个数时的误差均值.需要指出的是,在多姿态库中,由于每类的样本个数大约只有 $20(=30/3 \times 2)$ 个,因此,在实验中设定的最大的类中心个数为 7,这与 CAS-PEAL 数据库的最大类中心个数为 10 不同.从图中可以看出,与 CAS-PEAL 数据库上的结果类似,基于 1D Gabor 的局部判别方法要优于其他方法,并且对于不同的类中心的个数保持鲁棒.同时,基于 1D Gabor 特征的算法分别优于对应的基于 2D Gabor 特征的算法,这说明 1D Gabor 特征在姿态估计问题中更加有效.同时,在使用相同的 Gabor 特征的情况下,局部判别分析的结果要优于费舍尔判别分析的结果.这说明局部判别分析方法能够有效地解决姿态估计中存在的多模态问题.

从两个数据库中的实验中可以看出,基于 1D Gabor 的局部判别方法能够有效地解决姿态估计问题.本文将归结为所采用的 1D Gabor 特征能够准确地对图像提取与姿态相关的信息,同时消除了噪声等干扰;而基于核

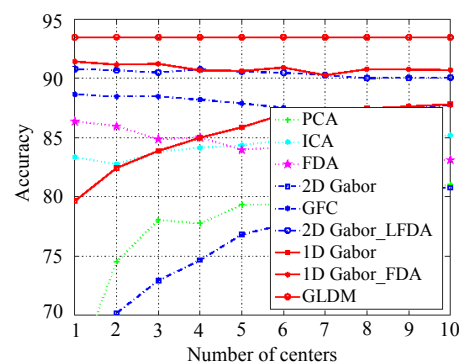


Fig.5 Accuracy(%) on the CAS-PEAL database

图 5 CAS-PEAL 数据库上的准确率(%)

函数的局部费舍尔判别分析方法能够解决姿态估计中的非线性问题和多模态问题,从而使性能得到进一步的提升.

2.2.3 对图像尺寸的鲁棒性实验

为了研究算法对于图像尺寸的鲁棒性,本文在 CAS-PEAL 数据库上在图像尺寸为 64×64 时重新进行了实验.在图 7 中,给出了不同的算法在不同类中心个数时的姿态估计的准确率.为了与 32×32 时的实验结果对比,在表 2 中,给出了在两种不同的图像尺寸情况下,不同算法在类中心个数从 1~10 时的最大准确率(MAR)和最大准确率的下降值.

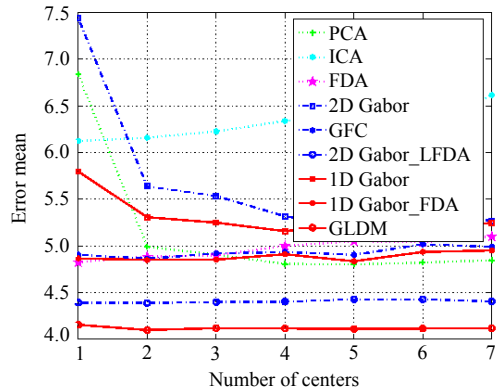


Fig.6 Error means ($^{\circ}$) on the multi-pose database

图 6 Multi-Poses 数据库上的误差均值($^{\circ}$)

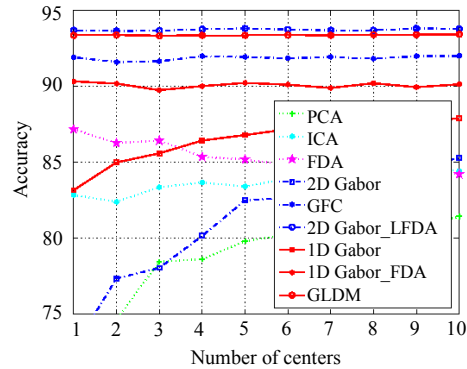


Fig.7 Accuracy on the CAS-PEAL databases, image size is 64×64

图 7 CAS-PEAL 数据库上的准确率,图像尺寸为 64×64

Table 2 Accuracy under different image sizes (%)

表 2 不同图像尺寸下的准确率(%)

	PCA	ICA	FDA	2D Gabor	GFC	2D Gabor_LFDA	1D Gabor	1D Gabor_FDA	GLDM
MAR 32×32	81.04	85.49	86.38	80.78	88.66	90.80	87.82	91.42	93.49
MAR 64×64	81.43	84.77	87.17	85.27	92.00	93.83	87.89	90.32	93.42
Difference	0.38	-0.72	0.79	4.49	3.34	3.03	0.07	-1.10	-0.07

从表 2 和图 6、图 7 中,我们可以得出 3 个简单的结论.首先,对于使用 2D Gabor 特征的 2D Gabor,GFC 和 2D Gabor_LFDA 来说,图像为 64×64 时的结果要明显优于 32×32 时的结果,基于 1D Gabor 的局部判别方法的结果要略差于 2D Gabor_LFDA,而 GFC 算法的结果要远优于 1D Gabor_FDA.我们认为,就像我们在本文第 2.1 节中介绍的那样,2D Gabor 需要更多的图像的细节信息,并随着图像质量的上升,2D Gabor 滤波器能够更好地提取局部特征.虽然基于 2D Gabor 特征算法在较高的图像质量时有不错的性能,但这也严重影响了算法的计算时间.对于姿态估计问题来说,由于姿态估计在很多实时系统中作为前端处理的过程,对计算时间有严格的要求,而 2D Gabor 特征需要较高的图像质量,因此实验结果再次说明了 2D Gabor 特征不适合在姿态估计问题中应用.最后,对于基于 1D Gabor 的局部判别方法来说,姿态估计的准确率对于图像尺寸基本稳定.这是因为本文的特征更多地与姿态本身相关,并且可以看作是利用行方向上的姿态信息进行姿态估计.因此可以认为,本文的算法受图像的细节影响较小,因而对于图像尺寸具有一定的鲁棒性.利用算法的这个性质,本文的方法可以在低分辨率、低质量的图像上取得不错的实验结果.同时,对于高质量的图像,本文可以将图像缩小到低分辨率,从而在保持姿态估计准确率的基础上显著地提高姿态估计的速度.

2.2.4 样本点投影空间中的分布分析

因为本文的主要目的是解决姿态估计问题,同时,在基于 1D Gabor 的局部判别方法中使用局部费舍尔判别分析的目的是为了解决姿态估计问题中存在的多模态问题,因此,特征在低维空间中对于不同姿态的可分性是关注的一个重点.为了展示基于 1D Gabor 的局部判别方法对于可分性的有效性,我们设计了下面的实验.在该实

验中,对于 CAS-PEAL 库中的数据,研究了不同姿态的样本在低维特征空间的分布情况.本文将对图像进行特征提取后的特征作为高维样本,然后利用全部的样本建立低维投影空间,在图 8 中,给出了 0° 、 15° 、 30° 和 45° 的 4 种姿态的 2 400 个样本在低维空间中的分布.为了可视化,只画了投影空间的前 3 维,实圆环、实三角、实菱形和实方块分别代表 0° 、 15° 、 30° 和 45° 的样本.

从图 8 中可以看出,在基于 1D Gabor 的局部判别方法的投影空间中,4 种不同姿态的样本按顺序排列,可分性最好,而其他方法的样本在低维空间中高度重叠.从数据降维后的分布可以看出,本文提出的算法能够有效地解决姿态估计问题中的多模态问题.

2.3 实验2:训练集和测试集异质

虽然本文提出的算法在两个数据库上均能取得不错的实验结果,但是上面实验中的训练数据和测试数据都来自同一个数据库,数据间具有较大的相似性.而在实际应用系统中,因为系统的研发者几乎不可能知道在实际应用中系统要处理何种类型的测试数据,因此测试数据和训练数据往往是异质的.因此,为了能够准确地反映不同的算法在实际姿态估计系统中的性能,在研究算法时进行训练数据和测试数据异质时的实验,是非常有意义的.本质上讲,这类实验的计算结果与模式分类的推广性高度相关.基于这个目的,我们设计了下面的实验.

在数据异质的实验中,本文使用多姿态库的图像作为训练数据,CAS-PEAL 数据库的图像作为测试数据.由于多姿态库中的图像是通过 SONY EVI-D31 摄像头采集得到的,而 CAS-PEAL 库中的图像通过普通的 USB 摄像头采集得到,同时不可避免地,这两个数据库是建立在不同人、不同的图像条件下的,因此,这两个数据库在摄像机参数、光照条件、表情、背景和采集人等方面是完全异质的.这非常有利于上述的数据异质的实验目的.

图 9 中给出了训练数据和测试数据异质时的实验结果,即不同的方法在不同的类中心个数下的误差均值 m .从图中我们可以看出,数据异质对姿态估计的结果有极大的影响.在数据同质时,所有方法的误差均值均不超过 7.5° .而在数据异质时,具有最优的实验结果的基于 1D Gabor 的局部判别分析方法的误差均值也大于 7.9° .这说明当数据异质时,训练数据和测试数据间的巨大差异并不能完全得到消除.同时我们可以发现,不同的方法受到数据异质影响的程度不同.与数据同质时的各种算法的结果较为近似相比,在数据异质时不同方法的性能有很大的差异.在数据同质时,不同方法在最优误差均值上的差异不超过 2° .而与此对应的是,在数据异质时,不同方法在最优误差均值上的差异可以达到 10° .一般认为,基于图像本身的统计特征的方法,比如主成分分析,更容易受到数据异质的影响,而只有能够提取与姿态相关的信息的方法才能减少因数据异质所产生的影响.我们提出的方法由于特征来自于行方向的信息,能够与姿态信息紧密相关,因此受到数据异质的影响相对较小,我们的方法能够提高姿态估计算法的推广性.总之,本实验结果表明,基于 1D Gabor 的局部判别分析的方法的性能优势在数据异质时能够更好地体现出来.由于数据异质时的结果更多地可以反映算法的推广性,因此可以认为,本文的算法具有较好的推广性,在实际姿态估计系统中,本文提出的算法更能体现出优越性.

总之,上述的实验结果支持如下的观察:基于 1D Gabor 的局部判别方法在未知的异质数据上能够得到更好的推广性.与其他方法相比,基于 1D Gabor 的局部判别方法对于光照、被采集人、背景、摄像机参数等更加鲁棒.这说明本文的方法提取的 1D Gabor 特征与姿态信息更加紧密相关.

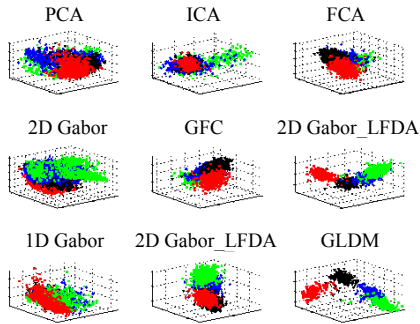


Fig.8 3D embedding of different features

图8 不同的特征在3D空间的嵌入

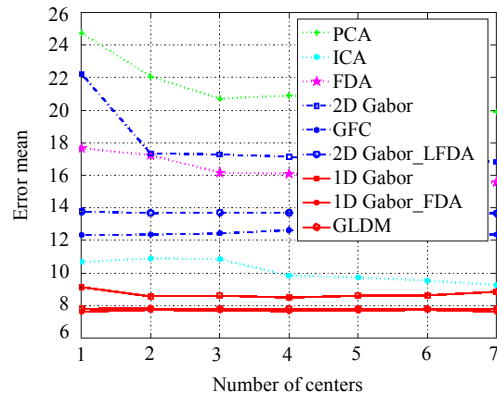


Fig.9 Error means (°) when the training dataset and the testing dataset are heterogeneous

图9 训练数据集和测试数据集异质时的误差均值(°)

3 结论和展望

本文提出了一种能够有效进行姿态估计的方法.通过使用 1D Gabor 特征,可以使得特征与姿态更加紧密相关,通过使用基于核函数的局部费舍尔判别分析,可以消除姿态估计存在的非线性和多模态问题.实验结果表明,本文提出的方法能够提高姿态估计的准确率,并具有很好的推广性能.

在接下来的工作中,有些方面的问题值得进一步探讨.首先,本文提出的算法主要是针对左右方向的头部旋转,如何让算法能够估计头部的任意旋转仍然需要进一步探讨.其次,为了强调面部不同区域具有不同的重要性,对面部的不同区域如何选用合适的 1D Gabor 的参数也是一项值得研究的工作,同时,这也能减少特征的维数.再次,虽然我们提出的算法的有效性通过实验得到了验证,但还是缺乏数学上的定量分析.因此,更深入的理论分析和扩展也将是本文的后续工作之一.最后,由于为了解决非线性问题,使用了高斯核函数,因此在计算速度上不可避免地受到影响.如何进一步提高算法的速度,仍然是一个值得研究的问题.

References:

- [1] Kuchinsky A, Pering C, Creech ML, Freeze D, Sera B, Gwizdka J. FotoFile: A consumer multimedia organization and retrieval system. In: Williams MG, ed. Proc. of the SIGCHI Conf. on Human Factors in Computing Systems: The CHI is the Limit. New York: ACM, 1999. 496–503.
- [2] Ji Q, Hu R. 3D face pose estimation and tracking from a monocular camera. Image and Vision Computing, 2002,20(7):499–511.
- [3] Nikolaidis A, Pitas I. Facial feature extraction and determination of pose. In: Proc. of the NOBLESSE Workshop on Nonlinear Model Based Image Analysis. 1998. 257–262. http://poseidon.csd.auth.gr/papers/PUBLISHED/CONFERENCE/pdf/Nikolaidis_A98a.ps.Z
- [4] Xiao J, Moriyama T, Kanade T, Cohn JF. Robust full-motion recovery of head by dynamic templates and re-registration techniques. Int'l Journal of Imaging Systems and Technology, 2003,13:85–94.
- [5] Cootes TF, Walker K, Taylor CJ. View-Based active appearance model. In: Crowley JL, ed. IEEE Int'l Conf. on Automatic Face and Gesture Recognition. Washington: IEEE Computer Society, 2000. 227–232.
- [6] Braathen B, Bartlett MS, Littlewort-Ford G, Movellan JR. 3D head pose estimation from video by nonlinear stochastic particle filtering. In: Proc. of the Annual Joint Symp. on Neural Computation. La Jolla: Institute for Neural Computation. 2001. <ftp://ftp.cnl.salk.edu/pub/marni/mplab2001.05.pdf>
- [7] Li YM, Gong SG, Liddell H. Support vector regression and classification based multi-view face detection and recognition. In: Crowley JL, ed. Proc. of the Int'l Conf. on Automatic Face and Gesture Recognition. Washington: IEEE Computer Society, 2000. 300–305.
- [8] Chen LB, Zhang L, Hu YX, Li MH, Zhang HJ. Head pose estimation using fisher manifold learning. In: Azada D, ed. IEEE Int'l Workshop on Analysis and Modeling of Faces and Gestures. Washington: IEEE Computer Society, 2003. 203–207.

- [9] Li SZ, Lu XG, Hou XW, Peng XH, Cheng QS. Learning multiview face subspaces and facial pose estimation using independent component analysis. *IEEE Trans. on Image Processing*, 2005,14(6):705–712.
- [10] Srinivasan S, Boyer KL. Head pose estimation using view based eigenspaces. In: Kasturi R, Laurendeau D, Suen C, eds. *Proc. of the Int'l Conf. on Pattern Recognition*. Washington: IEEE Computer Society, 2002. 302–305.
- [11] Wei YC, Fradet L, Tan TN. Head pose estimation using gabor eigenspace modeling. In: Reibman A, Knox K, eds. *Proc. of the Int'l Conf. on Image*. Washington: IEEE Computer Society, 2002. 281–284.
- [12] Raytchev B, Yoda I, Sakaue K. Head pose estimation by nonlinear manifold learning. In: Goldfarb L, ed. *Proc. of the Int'l Conf. on Pattern Recognition, Vol.4*. Washington: IEEE Computer Society, 2004. 462–466.
- [13] Krüger N, Pötzsch M, Malsburg CVD. Determination of face position and pose with a learned representation based on labeled graphs. *Image and Vision Computing*, 1997,15(8):665–673.
- [14] Krüger V, Sommer G. Gabor wavelet networks for efficient head pose estimation. *Image and Vision Computing*, 2002,20(9-10):665–672.
- [15] Turk M, Pentland A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 1991,3(1):71–86.
- [16] Belhumeur PN, Hespanha JP, Kriegman DJ. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997,20(7):711–720.
- [17] Schölkopf B, Smola A, Müller KR. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 1998,10:1299–1319.
- [18] Mika S, Rätsch G, Weston J, Schölkopf B, Müller KR. Fisher discriminant analysis with kernels. In: Larsen, J, Hu YH, Wilson E, Douglas S, eds. *Proc. of the Int'l Workshop on Neural Networks for Signal*. 1999. 41–48.
- [19] Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000,290:2323–2326.
- [20] Tenenbaum JB, Silva VD, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science*, 2000,290:2319–2323.
- [21] Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Dietterich TG, Becker S, Ghahramani Z, eds. *Proc. of the Advances in Neural Information Systems*. Cambridge: MIT Press, 2001. 585–591.
- [22] He XF, Niyogi P. Locality preserving projections. In: Thrun S, Saul LK, Schölkopf B, eds. *Proc. of the Advances in Neural Information Systems*. Cambridge: MIT Press, 2003. 153–160.
- [23] Liu J, Chen SC. Discriminant common vectors versus neighbourhood components analysis and Laplacianfaces: A comparative study in small sample size problem. *Image and Vision Computing*, 2006,24(3):249–262.
- [24] Sugiyama M. Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis. *Journal of Machine Learning Research*, 2007,8:1027–1061.
- [25] Cortes C, Vapnik V. Support vector network. *Machine Learning*, 1995,20:273–297.
- [26] Tipping ME. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 2001,1:211–244.
- [27] Bartlett MS. *Face Image Analysis by Unsupervised Learning*. Kluwer Academic Publishers, 2001.
- [28] Gao W, Cao B, Shan SG, Chen XL, Zhou DL, Zhang XH, Zhao DB. The CAS-PEAL large-scale chinese face database and baseline evaluations. *IEEE Trans. on Systems, Man, and Cybernetics, Part A*, 2008,38(1):149–161.



马丙鹏(1978—),男,河南睢县人,博士生,主要研究领域为模式识别,机器学习,计算机视觉.



陈熙霖(1965—),男,博士,研究员,博士生导师,CCF高级会员,主要研究领域为多模式人机接口,图像处理,模式识别,计算机视觉,数字电视技术.



山世光(1975—),男,博士,副研究员,主要研究领域为计算机视觉,模式识别,机器学习,生物特征识别,人脸识别.



高文(1956—),男,博士,教授,博士生导师,CCF高级会员,主要研究领域为计算机视觉,多媒体,图像处理,多模式接口,人工智能,虚拟现实.