

光照变化情况下的静态头部姿态估计

刘 坤, 罗予频, 杨士元

(清华大学自动化系, 北京 100084)

摘 要: 针对图像光照的变化对静态头部姿态估计的影响, 该文提出一种基于有向梯度直方图和主成分分析的姿态特征, 并利用 SVM 分类器进行分类。该算法分别在 CMU 姿态、光照、表情数据库和 CVL 人脸图像库上进行了测试。实验表明, 即使图像光照变化很大, 该算法仍可准确地估计头部姿态, 识别率达到 90% 以上。

关键词: 姿态估计; 有向梯度直方图; 光照

Static Head Pose Estimation Under Different Illumination

LIU Kun, LUO Yu-pin, YANG Shi-yuan

(Department of Automation, Tsinghua University, Beijing 100084)

【Abstract】 To solve head pose estimation problem under different illumination and expression, a new pose feature is proposed, which is based on Histogram of Oriented Gradient(HOG) feature and Principle Component Analysis(PCA), and SVM classifier is used to classify different poses. The algorithm is tested based on CMU pose, illumination and expression database and CVL database. Experiment shows that the algorithm can estimate poses accurately under different illumination, and correct ratio arrives at 90%.

【Key words】 pose estimation; Histogram of Oriented Gradient(HOG); illumination

1 概述

头部姿态估计算法在计算机视觉中具有广泛的应用。头部姿态的变化可以用在“智能”环境中分析被关注目标的注意力^[1]。在司机辅助驾驶系统中, 对司机的头部姿态估计能起到重要的作用^[2]。在人脸检测与识别中, 预先估计头部姿态便于选择合适的视角模型进行人脸的检测与识别^[3]。

头部姿态估计的算法主要分为 2 类: 基于模型的方法^[4]和基于图像表面的方法^[2-3]。前者需要对头部进行三维建模, 通过对应三维模型点和图像特征点来进行姿态估计, 其精度取决于对人脸特征定位的准确性, 姿态估计的范围通常较小。当对大范围、粗角度的头部姿态进行估计时, 通常采用基于图像表面的方法。这类方法不需要精确的定位人脸特征, 而是直接把头部姿态作为人脸图像的一种隐含参数, 即假设两者之间存在某种对应关系, 通过机器学习算法来学习这种对应关系。由于其使用的特征通常是基于图像的亮度和颜色获得的, 因此图像的光照变化会直接影响姿态估计的性能。

关于光照对图像亮度和颜色的影响, 人们已经提出了一些解决方法, 例如光照子空间方法^[5], 将不同光照条件的姿态图像作为训练图像, 但这样使得训练的样本增多, 增加了算法的时间复杂度; 或者对图像表面的光照情况进行估计, 利用光照模式参数空间估计光照模式, 然后进行针对性的光照补偿^[6]。但由于光照情况复杂多变, 对光照的估计通常限于有限的几种情况下, 或者选择对光照相对稳定的颜色常量^[7], 寻求各种光照下都稳定不变的颜色常量也是公开的难题。为此, 本文提出了一种新的姿态特征, 该特征是基于图像梯度特征和主成分分析(PCA)获得的。由于考虑到光照在图像局部范围内对图像亮度的影响可以近似为是相同的, 因此不直接提取图像的亮度值, 而是首先计算图像的梯度特征, 为了使得到的特征更具压缩性, 这里将不同姿态角度中提取

的梯度特征利用 PCA 向低维空间中投影, 将投影后的特征作为头部姿态特征, 进行分类和识别。

本文的姿态特征的优点如下:

- (1) 对图像光照的变化不敏感, 即使图像光照变化很大, 依然可以有效地估计姿态角度。
- (2) 不需要用户进行初始化, 可以抑制人脸的表情变化。
- (3) 对图像尺度和分辨率变化不敏感。
- (4) 具有很好的压缩性。

2 图像梯度特征的计算

本文提取图像的梯度分布特征与文献[8]中有向梯度直方图的计算方法类似。有向梯度直方图特征(HOG)用于行人检测时取得了非常成功的应用。与文献[8]的区别在于本文在获得图像梯度之后, 首先提取了每个图像单元的梯度变化的主方向, 该单元内的每个方向是相对于该单元的主方向获取的, 这样可以在一定范围内抑制头部姿态在其他方向的旋转带来的影响。

首先将原始图像做直方图均衡化处理, 图像的大小统一为 48×48 个像素。然后将图像分为若干个图像子区域, 每个子区域称为一个图像块, 每个图像块再划分为若干个图像单元, 在每个单元内计算像素点的有向梯度直方图。最后的图像梯度分布由每个单元内的直方图累积获得。本文将每个图像单元的大小设置为 8×8 个像素, 图像块的大小设置为 16×16 个像素。这样整幅图像就分为 3×3 个图像块, 每个图像块内有 2×2 个图像单元。图 1 给出了图像单元、图像块和原始图像的对应关系。

作者简介: 刘 坤(1980 -), 女, 博士研究生, 主研方向: 计算机视觉, 模式识别; 罗予频、杨士元, 教授、博士生导师

收稿日期: 2007-05-30 **E-mail:** liukun03@mails.thu.edu.cn

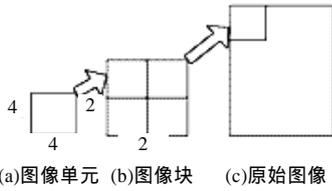


图 1 图像单元、图像块和原始图像的对应关系

图像梯度的计算方法直接影响姿态识别的性能。本文分别利用一维模板 $[-1,1]$, $[-1,0,1]$, $[1,0,-2,0,1]$, $[1,-8,0,8,-1]$ 和二维模板 $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ 对图像进行滤波,在水平方向和垂直方向计算图像的梯度。实验表明,针对头部姿态估计问题,模板 $[1,0,-2,0,1]$ 效果最好。对图像分别求水平方向和垂直方向的差分 dx , dy 之后,再根据 dx , dy 求得图像每个像素的梯度幅度和方向。梯度的方向为 dy/dx 的反正切,梯度的幅度为 $\sqrt{(dx)^2 + (dy)^2}$ 。这里,梯度的方向反映该像素点周围的灰度变化的方向,梯度的幅度反映灰度变化的大小。然后计算每个图像单元内的梯度特征在 $(-\pi/2, \pi/2)$ 范围内的以 22.5° 为间隔的梯度直方图。至于图像块之间的重叠大小,本文分别尝试了无重叠区域、重叠半个图像单元和重叠一个图像单元大小的方法。实验表明,针对头部姿态估计问题,图像块之间无重叠区域可以达到较好的效果,这时每幅图像生成324维的图像特征。图2(b)、图2(c)所示为通过上述方法得到的梯度幅值图和梯度方向图。



图 2 有向梯度直方图

有向梯度直方图是通过求图像区域内水平和垂直方向的差分获得的,利用局部差分的思想可以在一定程度上抑制光照的变化,但是当光照变化较大,或者在局部区域内光照变化不均匀时,依然会使得相同姿态的图像产生差别较大的有向梯度直方图。为了有效地抑制光照的影响,需要对有向梯度直方图做规范化处理。本文利用梯度直方图的一范数和二范数等进行归一化。结果表明,利用二范数进行归一化的效果较好。有向梯度直方图的规范化处理对于抑制光照的影响起到直接的作用。

3 利用 PCA 降维

为了得到更具压缩性和代表性的姿态特征,本文利用主成分分析将从图像中提取的梯度特征向三维空间中投影,因为人的头部姿态分布在三维空间流形上,利用投影后的向量作为姿态识别的特征。

设有 n 个 d 维的原始图像样本 x_1, x_2, \dots, x_n , $m = \frac{1}{n} \sum_{k=1}^n x_k$ 为样本的均值。 $S = \sum_{k=1}^n (x_k - m)(x_k - m)^T$ 为样本的散布矩阵。把全部样本向通过样本均值的一条直线作投影,得到代表全部样本的低维向量,该直线方程可以表示为: $x = m + \sum_{i=1}^{d'} a_i e_i$,其中, $d' < d$; e 表示这条通过样本均值的直线的单位向量,平方误

差准则函数为 $J_{d'} = \sum_{k=1}^n \left\| \left(m + \sum_{i=1}^{d'} a_i e_i \right) - x_k \right\|^2$ 。当向量 $e_1, e_2, \dots, e_{d'}$ 分别为散布矩阵的 d' 个最大特征值所对应的特征向量时,平方误差准则函数取最小值。系数 a_i 就是向量 x 对应于基 e_i 的系数,称作主成分。从几何上讲,样本点 x_1, x_2, \dots, x_n 在 d 维空间形成了一个 d 维椭圆型分布。而散布矩阵的特征向量就是这个分布的主轴。通过提取散布最大的几个方向,并将原始图像向这几个方向上投影,生成了新的姿态特征。

4 实验结果与分析

为了验证本文的特征对光照的鲁棒性,本文选用了CMU姿态、光照,表情数据库PIE^[9]和CVL人脸图像库^[10]的部分数据进行测试。PIE数据库收集了68个人的41368幅图像,分别是每个人在13种不同姿态、43种不同光照情况、4种表情下的图像,部分图像如图5(a)所示,由图示可见图像之间光照差别很大;而CVL人脸图像库内部的样本光照条件相对均匀,部分人脸图像如图5(b)所示。除此之外,还有部分数据是笔者在不同光照下拍摄的视频中截取的不同姿态图像,并利用人脸的对称性对图像进行镜像处理、模糊化等方法得到了不同姿态和分辨率下的图像。图5(c)中是经过图像压缩后得到的不同分辨率下的图像。



(a) PIE 数据库中的人脸图像



(b) CVL 数据库中的人脸图像



(c) 不同分辨率的图像

图 5 不同姿态、光照和分辨率下的图像

首先假定以被观测者为基准,其向左旋转时对应的姿态角度为负角度,其向右侧旋转时对应的姿态角度为正角度。头部姿态的变化主要指沿着水平方向的旋转变化。将头部姿态按照以上方法划分为5个姿态类别,分别对应 $-90^\circ, -45^\circ, 0^\circ, 45^\circ, 90^\circ$ 。在进行姿态估计之前,首先以PIE数据库中 45° 和 90° 两类姿态样本为例,观察样本在三维空间中的分布。分别将原始图像的灰度特征和本文使用的图像梯度特征通过PCA向三维空间中投影,投影后的特征分别如图6(a)和图6(b)所示。

从图6中可以看出,原始图像的亮度由于受光照变化的影响,使得相同姿态角度图像投影之后的特征相差很大,而与其他姿态角度相差较小,很难与其他姿态区分开来。但通过提取图像的梯度特征之后再投影后得到的特征向量,使得类内距离减小/类间距离增大,有效地抑制了光照变化带来的影响。

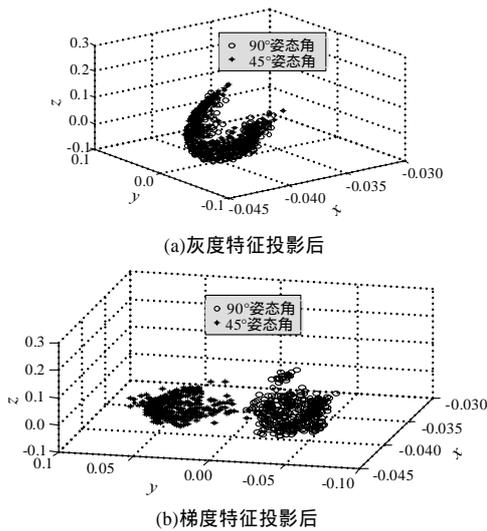


图6 原始的灰度特征和梯度特征投影后的对比

为了验证本文提出的姿态特征的有效性,本文分别比较了两种算法的结果:一种是直接利用图像的灰度特征经过PCA降维后的向量作为特征,另一种是基于图像梯度降维后的姿态特征。首先在CVL数据库中挑选570个不同姿态的样本,并对这些样本进行图像压缩,通过调整压缩比的大小,得到不同分辨率下的图像,分别利用这些图像进行训练和识别。姿态分类时采用的是SVM分类器,测试方法采用交叉验证法,统计得到的识别率如表1中数据所示。表1中压缩比分别为1, 5, 10, 15, 20。

表1 在CVL数据库上测试的识别率

算法	1	5	10	15	20
GRAY+PCA	0.863 1	0.848 9	0.825 9	0.794 5	0.761 9
HOG+PCA	0.935 0	0.897 6	0.928 6	0.928 1	0.915 2

从表1的结果中可以看出,直接利用灰度信息作为原始特征时,会受到图像分辨率的影响,当图像分辨率降低时,识别率也随之降低。而本文提出的特征对图像分辨率的变化具有很好的鲁棒性。同时,为了验证本文的姿态特征对光照的鲁棒性,本文还统计了在PIE数据库上进行训练和测试时姿态估计的识别率。表2给出了上述2种方法在PIE数据库上测试时,在不同样本数下的识别率。表中样本数分别为500, 600, 700, 800, 900, 1000。实验表明,在图像光照变化很大时,利用图像梯度特征进行降维后的特征进行姿态估计对光照的变化具有很好的鲁棒性。图7和图8为GRAY+PCA与HOG+PCA在不同条件下的对比。

表2 在PIE数据库上测试的识别率

算法	500	600	700	800	900	1000
GRAY+PCA	0.483 8	0.554 2	0.578 6	0.612 3	0.594 4	0.662 5
HOG+PCA	0.915 6	0.934 5	0.941 6	0.928 1	0.939 4	0.922 5

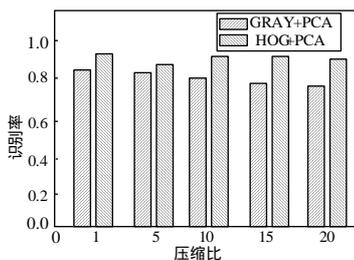


图7 GRAY+PCA与HOG+PCA在不同条件下的对比(1)

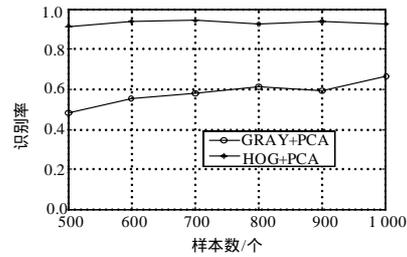


图8 GRAY+PCA与HOG+PCA在不同条件下的对比(2)

5 结束语

本文提出了一种基于图像梯度特征和主成分分析的姿态特征,对不同姿态角度的图像进行了分类和识别。实验表明,本文提出的特征即使在图像光照变化很大,依然可以准确地识别头部姿态。同时对于不同人脸表情和图像分辨率的变化也具有很好的鲁棒性。本文识别的头部姿态主要针对在水平方向上的头部姿态旋转,对同时绕不同轴旋转的头部姿态的识别,以及对更精细的姿态角度的识别是本文下一步的研究方向。

参考文献

- [1] Stiefelhagen R. Tracking Focus of Attention in Meetings[C]// Proceedings of the IEEE International Conference on Multimodal Interfaces. Pittsburgh, PA, USA: [s. n.], 2002.
- [2] Huang K, Trivedi M M, Gandhi T. Driver's View and Vehicle Surround Estimation Using Omnidirectional Video Stream[C]// Proceedings of IEEE Intelligent Vehicles Symposium. Columbus, USA: [s. n.], 2003.
- [3] Li S Z, Fu Qingding. Kernel Machine Based Learning for Multi-view Face Detection and Pose Estimation[C]// Proceedings of 8th IEEE International Conference on Computer Vision. Vancouver, Canada: [s. n.], 2001.
- [4] Alter T. 3-D Pose from 3 Points Using Weak-perspective[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(8): 802-808.
- [5] 苏宏涛, 张艳宁, 王 晶, 等. 光照变化条件下的人脸识别研究[J]. 西北工业大学学报, 2004, 22(4): 426-430.
- [6] 卿来云, 山世光, 陈熙霖, 等. 基于球面谐波基图像的任意光照下的人脸识别[J]. 计算机学报, 2006, 29(5): 760-768.
- [7] Funt B V, Finlayson G D. Color Constant Color Indexing[J]. IEEE Transaction on PAMI, 1995, 17(5): 522-529.
- [8] Dalal N. Histograms of Oriented Gradients for Human Detection[C]// Proceedings of the IEEE Conference on CVPR. San Diego, CA, USA: [s. n.], 2005.
- [9] Sim T, Baker S. The CMU Pose, Illumination and Expression Database[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(12): 1615-1618.
- [10] Computer Vision Laboratory. CVL Face Database[EB/OL]. (1999-10-20). <http://www.lrv.fri.uni-lj.si/facedb.html>.