

基于多模板匹配的自动人脸检测

袁超,张长水

(清华大学自动化系,北京 100084)

摘要: 本文提出了正面人脸检测的一种新算法.不同于以往将人脸看作高维空间中的单一聚类的方法,本文提出,人脸在高维空间中以多聚类形式存在.并采用最优聚类的方法将人脸分为四个聚类.以每个聚类的中心为一个模板,通过基于马氏距离(Mahalanobis distance)的多模板匹配法,检测出图像中的人脸.实验结果表明,该方法优于基于单一聚类的方法.

关键词: 人脸检测;模板匹配

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2000) 03-0095-04

Multiple Template Matching for Frontal-View Face Detection

YUAN Chao, ZHANG Chang-shui

(Department of Automation, Tsinghua University, Beijing 100084, China)

Abstract: In this paper, we present a new method for frontal-view face detection. Instead of viewing face patterns as a single cluster in the high-dimensional space as several previous approaches, we propose that face patterns can be better described by multiple clusters. We obtain four clusters by employing an optimal clustering method. Using each cluster center as a template, we adopt Mahalanobis-distance-based template matching to find the face location in an image. The experimental results show that our method outperforms the methods that are based on a single cluster.

Key words: face detection; template matching

1 引言

人脸检测的一项基本任务就是从一幅图像中找出人脸的位置,这对于后续的自动人脸鉴别工作具有重要的意义.不同于人脸鉴别,人脸检测把所有的人脸视为一类,而非人脸加以区分.人脸检测不同于一般物体检测的原因在于:准确描述人脸类是极其困难的.人脸类中的扰动包括:不同人的脸型间的差异,面部表情的变化,人的姿态的不同以及光照的影响.所以传统的相关匹配方法对于人脸检测的效果往往是不理想的.和其它的研究工作类似,本文将讨论的情况做一定的简化,即固定人脸的姿态为正面人脸,并允许少量的光照变化.在此前提下,文献[1]提出基于特征脸的人脸检测方法.它首先将人脸训练样本集作 K_L 变换,并认为人脸主要位于其主元子空间中.对任一待判断的样本,由它距主元子空间的距离作为是否是人脸的判据.文献[2]提出人脸服从高维高斯分布,从而以样本距人脸中心的马氏距离(Mahalanobis distance)作为是否是人脸的判据.不难看出,相关匹配,特征脸以及高斯分布的方法都是把人脸视作高维空间中的一个聚类.我们认为即使在单一姿态下,由其它原因如光照造成的人脸间的差异也会使人脸在高维空间中以多个聚类的形式存在.这里采用最优聚类的方法将人脸分为四个聚类,并认为每个聚类

中人脸服从高维高斯分布.通过基于马氏距离的多模板匹配法,检测出图像中的人脸.

2 人脸训练样本的预处理

为了使本文的方法更具一般性,选用了 MIT, Oxford, Yale 的标准人脸图库,我们实验室拍摄的图片,另一实验室拍摄的图片(称为 phface),以及室外照片中 160 张不同人的正常状态下的人脸(正常状态表示人脸为正面姿态,并在近似散射的光照下获得)作为训练图片.其中包含了不同表情,脸部特征(眼镜,胡子)等多种情况.对训练样本进行预处理的目的在于从一定程度上消除尺度和光照的影响,使人脸的类内距离尽可能小.具体的方法如下.首先对每一幅图像作缩放和剪裁变换:先用手工的方法标出两眼和嘴巴的位置,用 d 表示两眼间的距离,然后以这三者的中心为变换后图像的中心,以 $12/d$ 为缩放比例将原图像缩放并剪裁为 19×19 的图像(如图 1(b)所示).为了消除光照的影响,特别是明暗程度和少量侧光的影响,采用共轭梯度法用一个灰度平面(如图 1(c)所示)对变换后的图像进行拟合.并对变换后的图像减去灰度平面后的中间图像,作直方图均衡化.最后得到的 19×19 人脸图像(如图 1(d)所示)称为标准训练样本.对 160 幅图均作以上的处理,生成标准训练样本集.

收稿日期:1998-12-25;修订日期:1999-01-04

基金项目:国家“八六三”高技术项目(863-306-03-01-3)支持课题

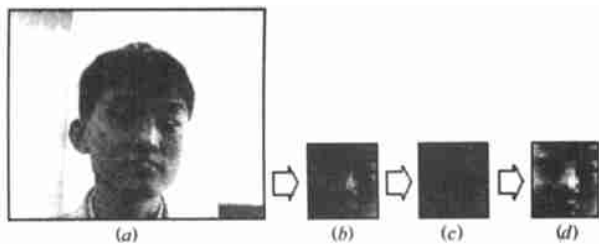


图 1 产生标准训练样本的流程图。(a)原始图像。(b)缩放并裁剪后的 19 ×19 人脸图像。(c)拟合后的灰度平面。(d)直方图均衡化后的标准训练样本图像。

3 最优聚类

令 $\{x_i\}_{i=1}^{N_T}$ 表示标准训练样本集, 其中 $N_T = 160$ 表示样本的个数。 x_i 是由标准训练样本图像从左到右, 从上到下展开形成的 1×361 维(令 $L = 361$) 的列向量。 最优聚类的思路是: 给定聚类的个数 k , 先用向量量化法产生初始聚类中心, 再用基于马氏距离的 k -均值方法将 $\{x_i\}_{i=1}^{N_T}$ 聚为 k 类, 并计算这样分类的评价函数。 对 k 的所有可能取值, 重复以上步骤, 使评价函数最大的 k 将作为最优聚类数, 其相应的聚类结果就为最优聚类结果。

3.1 用向量量化法产生初始聚类中心

在这里采用文献 [3] 所提出的 FSCL (frequency sensitive competitive learning)。 FSCL 相对于一般竞争学习的优越性在于: 当某一聚类中心受训练的次数比较多时, FSCL 将训练的机会让给其它的聚类中心。 这样能使每个聚类中心都得到充分的训练。

3.2 基于马氏距离的 k -均值算法

采用马氏距离描述一个聚类的好处在于: 允许聚类中的成员在某些分量上有大的变动, 而限制它们在另一些分量上的变动。 这是符合一般聚类性质的。 基于欧氏距离的 k -均值算法可以视为基于马氏距离的 k -均值算法的一种特例。 通过这一步, 我们将标准训练样本分给不同的聚类。 基于欧氏距离的 k -均值算法可以参考文献 [6]。

3.3 最优聚类数的判别准则

聚类的一个基本思想是: 使每个聚类的类内距离尽可能小, 使聚类间的距离尽可能大。 设 μ_c 和 Σ_c 分别表示各聚类的中心和协方差矩阵, $c = 1, 2, \dots, k$ 。 于是, 提出以下的判别准则:

(一) 对所有的 $c = 1, 2, \dots, k$, 计算聚类的平均类内马氏距离 e_c :

$$e_c = \frac{1}{N_c} \sum_{i=1}^{N_c} I(c|i) (x_i - \mu_c)^T \Sigma_c^{-1} (x_i - \mu_c) / N_c \quad (1)$$

其中 $I(c|i)$ 的意义如下: 如果 x_i 是属于第 c 个聚类, $I(c|i)$ 为 1, 否则为 0。 N_c 为属于第 c 个聚类的标准训练样本数。

(二) 对所有的 $c = 1, 2, \dots, k, p = 1, 2, \dots, k$, 计算聚类的类间距离 $D_{c,p}$:

$$D_{c,p} = \frac{1}{2} (\mu_c - \mu_p)^T (\Sigma_c^{-1} + \Sigma_p^{-1}) (\mu_c - \mu_p) \quad (2)$$

(三) 对所有的 $c = 1, 2, \dots, k, p = 1, 2, \dots, k, p \neq c$ 计算聚

类间的类内类间距离比 $R_{c,p}$:

$$R_{c,p} = (e_c + e_p) / D_{c,p} \quad (3)$$

(四) 对所有的 $c = 1, 2, \dots, k$, 计算每个聚类的类内类间距离比: $R_c = \max_{p=1,2,\dots,k, p \neq c} R_{c,p}$ (4)

(五) 计算判别函数 $J(k) = 1 / (\sum_{c=1}^k R_c / k)$ (5)

容易看出, $J(k)$ 就是聚类的平均类内类间距离比的倒数。 对 $k = 2, 3, \dots, K$ 的每一个可能的取值都按 3.1 到 3.2 的步骤进行聚类, 然后计算相应的 $J(k)$ 。 使 $J(k)$ 值最大的 k 值就是最优聚类数, 其相应的聚类结果就是最优聚类结果。 这里取 $K = 7$ 。 文献 [4] 采用一个经验的数值对人脸进行聚类, 而没有考虑类内类间距离。 从这个意义上讲, 我们的方法更合理。

3.4 在主元子空间中的最优聚类

由于实际上 $\text{rank}(\Sigma_c) < N_c < L, c = 1, 2, \dots, k$ 。 因此, Σ_c 的秩小于其阶数, Σ_c 不可逆。 所以在 3.1 ~ 3.2 中 Σ_c 的求逆是不可得的。 因此, 采用以下的近似方法进行最优聚类。 对 $\{x_i\}_{i=1}^{N_T}$ 作 KL 变换:

$$\tilde{x} = T \Sigma \quad (6)$$

其中, Σ 为 $\{x_i\}_{i=1}^{N_T}$ 的协方差矩阵, T 为 Σ 的特征向量矩阵, 是由 Σ 的特征值构成的对角阵。 令 M 表示前 M 个主特征向量组成的矩阵, 令 μ 表示 $\{x_i\}_{i=1}^{N_T}$ 的均值向量。 则 x 在前 M 维主元子空间中的分量 \tilde{x} 可由下式计算:

$$\tilde{x} = x - \mu \quad x = \tilde{T} \tilde{x} \quad (7)$$

由此得到 $\{x_i\}_{i=1}^{N_T}$ 在其主元子空间中的对应分量集 $\{\tilde{x}_i\}_{i=1}^{N_T}$ 。 在这里, 我们人脸主要存在于其主元子空间中。 于是, 用 $\{\tilde{x}_i\}_{i=1}^{N_T}$ 替换 $\{x_i\}_{i=1}^{N_T}$ 进行 3.1 ~ 3.3 小节中的最优聚类, 其中 $M = 20$ 。 对 $\{\tilde{x}_i\}_{i=1}^{N_T}$ 完成最优聚类后, 得到四个聚类。 在 L 维原始空间中, 一次性求出各个聚类的中心 μ_c 和协方差矩阵 $\Sigma_c, c = 1, 2, 3, 4$ 。 最后所得到的四个聚类中心按 19 ×19 图像显示如图 2。

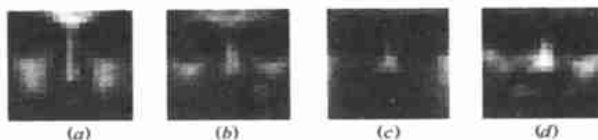


图 2 聚类中心图像。从左到右依次为第 1 ~ 4 个聚类的中心

4 基于马氏距离的多模板匹配

在第三节, 得到了描述人脸的四个聚类。 假设每个聚类中的人脸服从高维高斯分布, 并设该高斯分布的协方差矩阵为 Σ_{π_c} , 均值为已求出的 $\mu_c, c = 1, 2, 3, 4$ 。 以这四个聚类中心为四个模板, 求待检测样本 x 与每个模板的马氏距离:

$$d_c(x, \mu_c) = (x - \mu_c)^T \Sigma_{\pi_c}^{-1} (x - \mu_c) \quad (8)$$

设 Σ_{π_c} 的特征值分解的结果是: $\Sigma_{\pi_c} = \pi_c \pi_c^T$ 。 并令 $y_c = \pi_c^T (x - \mu_c)$ 。 则式 (14) 变为:

$$d_c(x, \mu_c) = y_c^T \pi_c^{-1} y_c = \sum_{i=1}^L y_{ci}^2 / \alpha_i \quad (9)$$

其中 α_i 为 π_c 的第 i 个特征值, y_{ci} 为 y_c 的第 i 个分量。 在这里, 采用文献 [2] 中的近似方法:

$$d_c(x, \mu_c) = \sum_{i=1}^{M_c} y_{ci}^2 / c_i + \sum_{i=1}^c (x_i)^2 / c_i \quad (10)$$

其中 $\sum_{i=1}^c (x_i)^2 = \sum_{i=M_c+1}^L y_{ci}^2 = x - \mu_c^2 - \sum_{i=1}^{M_c} y_{ci}^2$

$$c = \frac{1}{L - M_{c_i=M_c+1}^{c_i}} \quad (11)$$

可以近似认为 Σ_c 的前 M_c 个特征向量为 Σ_n 的前 M_c 个特征向量, Σ_c 的前 M_c 个特征值为 Σ_n 的前 M_c 个特征值. 因此可以得出 $\sum_{i=1}^c (x_i)^2$. Σ_n 的其它特征值将由 Σ_c 的特征值线性外推产生. 因此 c 可以得到. 由此 $d_c(x, \mu_c)$ 可以近似求出. 令 S 表示同一图像中所有位置所产生的待检测样本集. 对所有的 $x_j \in S$, 使下式最小的 x_j 被认为是人脸:

$$\min_{c=1,2,3,4} d_c(x_j, \mu_c) \quad (12)$$

在实际的检测过程中, 为了能检测出不同尺度的人脸, 每幅待检测图像将按比例 1/1.2 缩放为不同尺度的图像. 由此得到的所有图像中的所有 19×19 子块经第二节中的平面拟和和直方图均衡化后构成待检测样本集 S .

5 实验结果与讨论

为了检验我们方法的有效性, 这里选用了 76 张测试图



图 3 检测结果. 其中每一幅图像的右上角的数字表示检测出的人脸属于第几个聚类

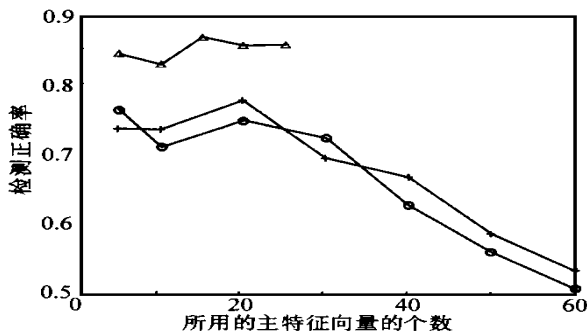


图 4 与其它方法的比较结果. 其中“ ”连成的典型表示我们方法的结果,“+”表示文献[2]的结果,“ ”表示文献[1]的结果.

(1) 从图 4 中, 可以看出, 随着主特征向量个数的增加, 文献[1,2]方法的正确率下降. 可以作如下解释: 随着主特征向量个数的增加, 一部分非人脸的子空间也被认为存在于人脸子空间中, 因此一些非人脸样本将被认为比真正的人脸更象人脸. 由此导致错误率的上升. 这说明人脸的确主要存在于其主元子空间中. 但实验表明主元子空间并不是对人脸的一种最好的描述. 本文提出的多聚类方法是对人脸一种更好, 更紧

凑的表示. 其中 20 张来自以上所提出的标准图库中除用于训练的图片以外的图片. 另外的图片来自网上, 罪犯登记照, 以及其它一些照片. 这些图片中的人脸均为正常状态下的人脸. 对图像中检测出的人脸, 用一个正方形的方框表示其位置与尺度. 一些检测结果如图 3 所示. 在这里 $M_c = 10, c = 1, 2, 3, 4$. 同时本文所提出的方法与文献[1,2]中的方法进行比较. 比较的结果如图 4 所示, 其纵坐标表示检测正确率. 横坐标表示各种方法所用的主特征向量的个数. 实验结果表明, 本文的方法优于文献[1,2]的算法. 其中本文方法的最高正确率为 86.8%. 稍加改动, 可以检测多个人脸. 方法如下: 先找出使式(12)最小的待检测样本, 如果它使式(12)小于某一阈值, 认为它是人脸, 并将它及与它在图像中相邻的待检测样本从待检测样本集中删除, 继续检测剩下的人脸, 直至无待检测样本使式(12)小于该阈值. 也可以采用文献[5]的方法. 实验结果如图 3 右上角图像所示. 同时, 用本文的方法对 Yale 库中 15 个人在左侧光和右侧光情况下共 30 幅图像进行检测. 结果有 21 幅图给出了正确的检测结果, 这说明本文的算法有一定的外推能力, 因为我们的训练样本中没有包括强侧光的情况.

针对实验的结果, 作以下讨论:

由图 4 可以直观地看出合适的主元个数应该在 20 左右, 因此, 在标准训练样本集的前 20 维主元子空间中进行最优聚类是合理的.

(2) 通过对所得到的四个聚类做进一步分析, 发现导致这样聚类结果的主要原因是不同的光照环境. 第一聚类的人脸是在有少量正面光照的环境下获得的, 因此人脸的中间比两侧亮; 第二聚类的人脸是在接近理想的散射光的环境下获得的, 因此人脸受光照的影响较均匀; 第三聚类的人脸是在有少量侧光的环境下获得的, 因此人脸的左侧或右侧比中间亮; 第四聚类的人脸是较强光线下获得的, 所以该类的人脸较之其它类的人脸整个较亮. 这种光照差异可以从图 2 的各聚类中心的图像中看出. 为了证实聚类类间的差别主要来自光照, 作了以下实验: 用本方法对同一个人在我们实验室中拍摄的图像和室外的照片进行检测(两次拍摄的时间间隔为一周, 人脸均为自然表情). 结果, 在实验室中拍摄的人脸被判为属于第三聚类, 室外照片中的人脸被判为属于第四聚类. 这说明光照的差别的确可以使人脸发生较大的变化. 我们认为, 这一结论对人脸识别具有重要的意义. 它表明: 同一个人在不同的光照下人脸的差别可能比人与人之间在同一光照下的人脸差别还要大. 因此, 可以采取基于不同光照的人脸识别. 对一幅人脸,

先判断其光照环境(属于哪一个光照下的聚类),然后再在各自光照环境下获得的人脸库中去完成识别。

(三)本方法还可以从以下几方面加以改进。首先可以引入非人脸样本,对人脸与非人脸的边界加以刻画,可以参照文献[4,5]。其次,可以引入人脸局部特征或边界信息对检测结果做进一步判断。比如在检测到人脸后,对其各个特定部位,判断是否存在眼睛,鼻子或嘴巴。第三,虽然本文的方法具有一定的外推能力,但在训练样本中不包含的光照情况下进行检测时,正确检测率有明显的降低。因此,为了使本算法更具普遍性,可以引入其它的光照环境下的训练样本。相信此时最优聚类的结果会发生变化,很有可能会产生更多的聚类。

6 小结

本文提出了正面人脸检测的一种新算法。采用最优聚类的方法将人脸分为四个聚类,并认为每个聚类中人脸服从高维高斯分布。通过基于马氏距离的多模板匹配法,检测出图像中的人脸。实验结果表明,本方法优于基于单一聚类的方法。同时,我们认为导致聚类间差别的主要原因是光照,因此提出了基于不同光照的人脸识别的想法。并将在今后的工作中加以实现。另外,所提出的基于马氏距离的最优聚类的算法适用于其它的聚类实例。

参考文献

- [1] M. Turk and A. Pentland. Eigenfaces for recognition. J. Cognitive Neuroscience, 1991, 3(1)
- [2] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. IEEE Trans. on Pattern Analysis and Machine Intell., July 1997, 19(7): 696 ~ 710
- [3] S. C. Ahalt, A. K. Krishnamurty, P. Chen and D. E. Melton. Competitive learning algorithms for vector quantization. Neural Networks, 1990, 3: 277 ~ 291
- [4] K. K. Sung and T. Poggio. Example-based learning for view-based human face detection. IEEE Trans. on Pattern Analysis and Machine Intell., Jan. 1998, 20(1): 39 ~ 51
- [5] H. A. Rowley, S. Baluja and T. Kanade. Neural network-based face detection. IEEE Trans. on Pattern Analysis and Machine Intell., Jan. 1998, 20(1): 23 ~ 38
- [6] 边肇祺等. 模式识别. 清华大学出版社, 1988



袁超 1998年在清华大学自动化系获学士学位,现为该校模式识别和智能系统专业的硕士研究生。主要研究兴趣为:图像处理、模式识别和计算机视觉。