

基于隐马尔科夫模型的人脸识别

余龙华,王 宏,钟洪声

(电子科技大学 电子工程学院,四川 成都 611731)

摘 要:对于人脸识别系统来说,人脸图像的特征提取和匹配是决定人脸识别系统性能的关键所在。文中提出基于隐马尔科夫模型的人脸识别方法。首先,根据人脸的特点建立马尔科夫模型,然后对图像进行预处理,再利用采样窗对人脸图像进行采样并进行离散余弦变换,提取变换后的系数作为观察向量。最后对人脸图像进行HMM训练,训练结束后即建立了一个人的HMM。基于DCT系数的二维隐马尔科夫模型由于充分利用了人脸图像的二维统计特性,具有较高的识别率。实验结果证明此方法在准确性方面具有良好的性能。

关键词:隐马尔科夫模型;光照补偿;人脸识别

中图分类号:TP31

文献标识码:A

文章编号:1673-629X(2012)02-0025-04

Face Recognition Based on Hidden Markov Model

YU Long-hua, WANG Hong, ZHONG Hong-sheng

(School of Electronic Engineering, University of Electronic Science and Technology, Chengdu 611731, China)

Abstract: Extracting face features from facial images and facial image matching are the most important part in the face recognition system. A face recognition approach based on Hidden Markov Model was presented in this paper. Firstly, this method establishes the Hidden Markov Model according to the features of face, then preprocessing the image, sample the face image by sampling window and process DCT (Discrete Cosine Transform) transformation, get the DCT coefficient as observation vector. Finally train the face image, after trained the image, get the HMM parameters of a people. The face recognition algorithm based on DCT coefficient makes the full use of face image 2-D statistical property, with high rate in different condition. The experimental results demonstrate that the proposed algorithm has excellent performance.

Key words: HMM; illumination compensation; face recognition

0 引言

基于生物特征的识别技术是人工智能、计算机视觉等领域的极具挑战性的研究课题之一,其中的人脸识别是当前最热门的模式识别研究课题,与指纹识别、虹膜识别等相比它具有非接触性、非侵犯性^[1],所以它是一种比较友好的生物特征识别技术。目前的人脸识别算法众多,归纳起来主要分为基于几何特征的方法^[2]、模板匹配方法^[3]、神经网络法、弹性图匹配方法及隐马尔科夫方法。基于几何特征的方法具有存储量小、对光照变化不敏感的优点,但是存在对表情和姿态变化的鲁棒性较差、忽略了局部细微特征等缺点;模板匹配的方法的优点是充分利用了特征区域的全局信息,缺点是计算时间长;神经网络法的优点是避免了复

杂的特征提取工作,缺点是神经元数目多、训练时间长、识别速度慢;弹性图匹配方法的优点是对光照、位移、旋转及尺度变化都不敏感,缺点是对每个存储的人脸需要计算其模型图,计算量大、存储量大。而基于HMM(Hidden Markov Model)的人脸识别算法^[4],既充分考虑了人脸各器官的不同特征,又考虑了他们间的相互联系,该方法鲁棒性较好,对表情姿态变化不太敏感,计算量适中。

1 基于HMM的人脸识别

1.1 HMM简介

马尔科夫模型是马尔科夫过程的模型化。一个HMM可简记为 $\lambda = (N, M, \pi, A, B)$ 。它由两部分组成,一部分是马尔科夫链,由 A 描述,刻画状态的转移,产生的输出是状态序列;另一部分是一个随机过程,刻画状态和观察值之间的统计对应关系,由 B 描述,产生的输出为观察值序列。在这个随机过程中,并不能直接观察到随机过程内部的状态,而只能看到观察值序列,通过这个随机过程去感知状态的存在和特征^[5]。

收稿日期:2011-07-04;修回日期:2011-10-18

基金项目:云南省科技计划项目(2009CA021);中央高校基本科研业务费项目(ZYGX2010J023)

作者简介:余龙华(1986-),男,安徽安庆人,硕士研究生,研究方向为人脸检测和识别;钟洪声,教授,硕士生导师,研究方向为信号处理。

HMM 由下述五个参数描述^[5]。

1) N : 马尔科夫链的状态数目, 设 N 个状态分别记为 $\theta_1, \theta_2, \dots, \theta_N$, 设 t 时刻马尔科夫链所在状态记为 q_t , 显然 $q_t \in (\theta_1, \theta_2, \dots, \theta_N)^{[5]}$ 。

2) M : 每个状态对应的观察值数目, 设 M 个观察值分别为 V_1, V_2, \dots, V_M , 设 t 时刻的观察值为 O_t , 其中 $O_t \in V_1, V_2, \dots, V_M^{[5]}$ 。

3) $\pi = (\pi_1, \pi_2, \dots, \pi_N)$: 初始状态概率矢量, 其中 $\pi_i = P(q_1 = \theta_i), 1 \leq i \leq N$ 。

4) $A = (a_{ij})_{N \times N}$: 状态转移概率矩阵, 其中 $a_{ij} = P(q_{i+1} = \theta_j | q_i = \theta_i), 1 \leq i, j \leq N$ 。

5) $B = (b_{jk})_{N \times M}$: 观察概率矩阵, 其中 $b_{jk} = P(O_j = V_k | q_i = \theta_j), 1 \leq j \leq N, 1 \leq k \leq M$ 。

1.2 HMM 三个基本问题及解决办法

问题 1: 给定观察序列 $o = o_1, o_2, \dots, o_n$ 和模型 $\lambda = (\pi, A, B)$ 计算 $P(o | \lambda)$ 。(评估问题)

问题 2: 给定观察序列 $o = o_1, o_2, \dots, o_n$ 和模型 λ , 如何选择一个合适的状态序列 $S = q_1, q_2, \dots, q_n$, 使 S 能最合理解释观察值序列^[6]。(解码问题)

问题 3: 如何调整模型参数 $\lambda = (\pi, A, B)$ 使得 $P(o | \lambda)$ 最大。(学习问题)

问题 1 解决方法: 用前向-后向算法, 定义前向变量 $a_t(i) = P(o_1, o_2, \dots, o_n, q_t = \theta_i | \lambda), 1 \leq t \leq T$, 初始化: $a_1(i) = \pi_i b_j(o_1), 1 \leq t \leq T$; 递归: $a_{t+1}(j) = [\sum_{i=1}^N a_t(i) a_{ij}] b_j(o_{t+1}), 1 \leq t \leq T-1, 1 \leq j \leq N$; 终结: $P(o | \lambda) = \sum_{i=1}^N a_T(i)$, 后向算法和前向算法刚好反过来。

问题 2 解决方法: 使用 Viterbi 算法, 此算法能够解决给定观察序列 $o = o_1, o_2, \dots, o_n$ 以及模型 $\lambda = (\pi, A, B)$, 选择一个合适的状态序列 S , 使 S 能最合理地解释观察序列 o ^[6]。

问题 3 解决方法: 用 Baum-Welch 算法(模型训练算法), 给定观察值序列 o , 通过计算确定一个模型 λ 使得 $P(o | \lambda)$ 最大。算法步骤:

第一步: 初始化模型(待训练模型 λ_0);

第二步: 基于 λ_0 以及观察值序列 o , 训练模型 λ ;

第三步: 如果 $\log P(X | \lambda) - \log P(X | \lambda_0) < \Delta$, 说明训练已经达到预期效果。算法结束。否则, 令 $\lambda_0 = \lambda$ 继续第二步工作。

1.3 图像预处理

对于人脸识别系统来说预处理是一个非常重要的步骤, 通过预处理可以减小光照、距离等条件的影响, 文中主要进行光照补偿。

光照补偿可以通过调整人脸灰度图像的均值和方

差来完成。图像的均值 μ 和方差 σ 的定义如下:

$$\mu = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H I(x, y) \quad (1)$$

$$\sigma^2 = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H (I(x, y) - \mu)^2 \quad (2)$$

其中 $I(x, y)$ 表示图像像素的灰度值, W 表示图像的宽度, H 表示图像的高度。

光照越强均值越大, 光照越弱均值越小; 整个图像的亮度变化越大, 则方差越大, 整个图像的亮度变化越小, 则方差越小。因此, 要减少光照强度的影响, 可以通过对人脸图像的每一个像素的灰度值进行公式(3)的运算。

$$I_i(x, y) = (I(x, y) - \mu) \times \sigma_0 / \sigma + \mu_0 \quad (3)$$

其中 $I_i(x, y)$ 为变换后相应像素的灰度值, σ_0, μ_0 为设定的均值和方差, σ_0, μ_0 的均值要接近样本库中大部分图像的均值和方差。文中取 σ_0, μ_0 的值分别为 33, 140。这样经过补偿后的人脸图像都具有相同的均值和方差。其处理效果如图所示, 其中图 1 为光照较弱的原始图片及进行处理后的图像; 图 2 为光照较强的原始图像及进行处理后的图像, 从中可以看出, 无论是光照较强还是较弱, 经过处理后的图像的亮度比较接近。

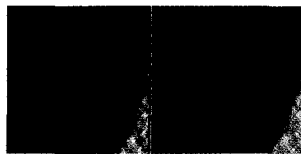


图 1 光照较弱的原始图像及处理后效果



图 2 光照较强的原始图像及处理后图像

1.4 人脸 HMM 建立

对于人脸来说, 其在垂直方向上具有相对稳定的结构, 即人脸从上到下分别由额头、鼻子、眼睛、嘴巴以及下巴构成, 将这 5 个部分分别抽象成 HMM 的状态^[7]。同样由于人脸在水平方向上也具有相对固定的空间结构, 所以, 可将垂直方向划分的状态进行扩充, 扩充为包含水平方向变化的 HMM^[7]。由于这种模型嵌入了水平方向状态, 故将其称为嵌入式隐马尔科夫模型(E-HMM)。由于 E-HMM 从水平和垂直方向上反映了人脸的二维结构特征, 所以可以更加细微地刻画人脸的个人特征。E-HMM 的人脸二维模型如图 3 所示, 共取 5 个超状态, 5 个超状态分别对应嵌入 3、6、6、6、3 共 24 个子状态。状态之间的转移从第一个超状态开始, 并从上向下在相邻两个超状态之间转移。

1.5 特征提取

提取人脸的观察向量的原则是具有强的抗干扰性和鲁棒性。用来提取 HMM 的观察值序列的方法主要有以下几种,一是 Sammaria 提出的直接用图像像素灰度值作为观察向量^[8],二是 Nefian 提出的用二维离散余弦变换(2D-DCT)的变换系数作为观察向量^[9],以及基于 SVD(singular value decomposition)分解的方法^[10]。文中采用 DCT 系数作为观察序列,一是可以降低观察向量的维数,二是可以降低对噪声和光照变化的敏感度。其具体实现过程如下:用大小为 $L_x \times L_y$ 的采样窗对人脸图像(设人脸图像的宽度为 W ,高度为 L)进行遍历采样,垂直方向的相邻采样窗之间的重叠部分为 p_x ,水平方向的相邻采样窗之间的重叠部分为 p_y ,其示意图如图 4 所示。文中采用 12×12 像素采样窗进行采样,采用 75% 的重叠扫描,可以获得较好的识别效果^[11]。因此采样窗每次移动 3 个像素。由于 DCT 系数主要集中在低频部分,因此可以用三角窗取 DCT 系数矩阵的前 9 个系数得到一个观察向量,以降低观察向量的维数。

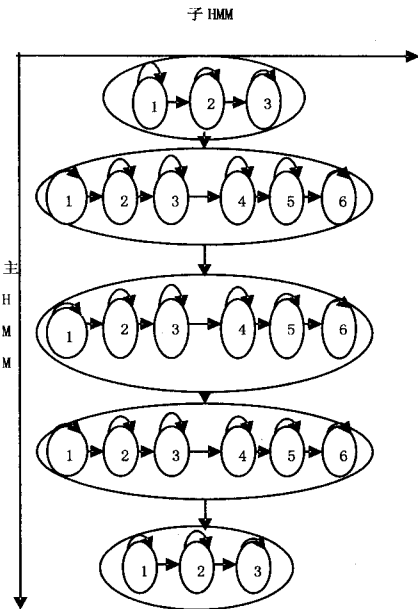


图 3 人脸 E-HMM 模型

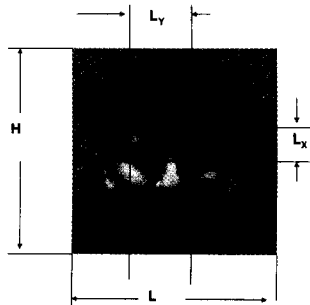


图 4 特征提取

1.6 HMM 模型的训练

为了识别人脸,需要先建立人脸库,人脸库的建立

就是在数据库中保存人脸的 HMM 模型参数,这个过程称为 HMM 训练。为了使训练结果对光照和姿态具有一定的鲁棒性,一般采用同一个人的不同光照、姿态的样本图片进行训练。

人脸 HMM 训练按以下步骤进行^[12]:

- 1) 首先需要建立 HMM 模型 $\lambda = (N, M, \pi, A, B)$, 然后对人脸图像进行采样,提取采样窗的 DCT 系数,将其作为观察向量。设定 E-HMM 的超状态数目以及嵌入的状态数目,对状态之间的转移矩阵 A 进行初始化,对观察概率矩阵 B 进行初始化,设定 $b_{jk} = 1/M, 1 \leq j \leq N, 1 \leq k \leq M$ 。
- 2) 对观察序列实行双重 Viterbi 算法,调整初始化参数。
- 3) 使用 Baum-Welch 算法对模型参数进行重估,迭代调整模型参数,使其达到最优化,当两次迭代的误差小于指定的阈值时,训练结束。

训练结束后,人脸库中就建立了一个人的 HMM。这个过程的 HMM 参数即代表数据库中的人脸。

1.7 人脸识别

在建好人脸库之后,就可以进行身份识别,其具体过程如下所示:

- 1) 用训练相同的方法提取待被识别人脸的 HMM 参数。
- 2) 将待识别人脸的 HMM 参数与人脸库中已经存在的 HMM 参数进行比较,求出各项的相似概率。
- 3) 对这些概率进行排序,输出相似概率最大项。

整个系统的框架图如图 5 所示。

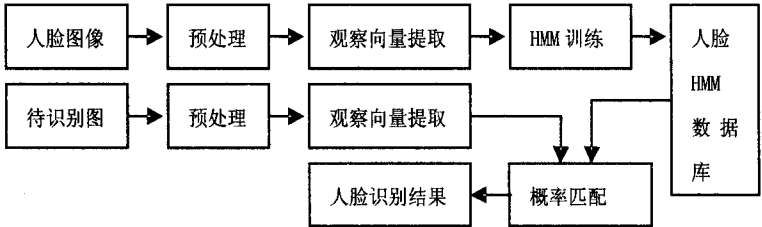


图 5 系统框架图

2 实验结果与分析

整个系统采用 C++ 编程实现。为了检测基于 E-HMM 人脸识别方法的效果,文中对 Yale 人脸库以及自建的人脸库进行了测试,Yale 人脸库包括 15 个人,每个人 11 张图片,表情、姿态、光照都有一定的变化,取前 4 张用来训练,后面 7 张用来识别。自建人脸库一共采集了 5 组,每组 10 个人,每个人 14 张图片,每组采集的条件和时间段都不一样,其中自建人脸库 1 的采集条件为下午采集,室外光线较好;人脸库 2 为晚上采集;人脸库 3 为白天采集的照片,采集条件为阴天,室外光线较暗;人脸库 4 采集条件为晚上采集且在

人脸面前加一个打开的台灯;人脸库 5 的采集条件为白天采集且在人脸面前加一个打开的台灯。测试结果如表 1 所示。

表 1 测试结果

	Yale 人脸库	人脸库 1	人脸库 2	人脸库 3	人脸库 4	人脸库 5
识别图片总数	105	100	100	100	100	100
识别正确数	85	92	98	95	93	89
识别率	81%	92%	98%	95%	93%	89%

通过表 1 的实验结果可以看出,E-HMM 算法与文献[13]算法相比^[13],具有较高准确率,对光照、表情和姿态具有一定的鲁棒性,但是光照变化太剧烈时识别率会急剧下降。

3 结束语

文中利用隐马尔科夫模型对人脸进行识别,取得了较好的效果。对光照、表情和姿态具有一定的鲁棒性。其他很多算法进行人脸识别的时候都有很多的限制,例如要求光照一定、不能佩戴饰物、面部表情单一、头部转动角度一定等,而基于 E-HMM 的人脸识别方法受这些条件的限制较小。但是通过实验发现,光照对识别结果有一定的影响,下一步的工作是怎样通过图像预处理来减少光照对系统的影响。

参考文献:

- [1] 王映辉. 人脸识别-原理、方法与技术[M]. 北京:科学出版社,2010:18-19.

(上接第 24 页)

值的经验模态分解改进算法提高了分解的精度和可靠性。

参考文献:

- [1] Huang N E, Shen Z, Long S R, et al. The Empirical Mode decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis[J]. Proceedings of the Royal Society of London Series A, 1998, 454: 903-995.
- [2] Deering R, Kaiser J F. The use of a masking signal to improve empirical mode decomposition[C]//ICASSP. [s. l.]: [s. n.], 2005: 485-489.
- [3] Hong Hong, Wang Xinlong, Tao Zhiyong. Local Integral Mean-Based Sifting for Empirical Mode Decomposition[J]. IEEE Signal Processing Letters, 2009, 16(10): 841-844.
- [4] 汝振, 李昕, 陈飞, 等. 一种基于 HHT 的语音增强算法研究与仿真[J]. 计算机技术与发展, 2010, 20(8): 116-119.
- [5] Huang N E, Shen Z, Long S R. A new view of nonlinear water waves: the Hilbert spectrum[J]. Annual Review of Fluid Me-

- [2] Brunelli R, Poggio T. Face Recognition through Geometrical Features[C]//Proceedings of ECCV 92. Santa Margherita Ligure, Italy: [s. n.], 1992: 792-800.
- [3] Yuille A L, Cohen D S, Hallinan P W. Feature extraction from faces using deformable templates[J]. International Journal of Computer Vision, 1992, 8(2): 99-111.
- [4] 王东李. 一种新的人脸识别算法[J]. 计算机技术与发展, 2009, 19(5): 147-148.
- [5] 陈盛映, 刘盛. 基于 OpenCV 的计算机视觉技术实现[M]. 北京: 科学出版社, 2008: 283-284.
- [6] 谈昌彬, 李一民. 基于 EHMM 的人脸识别[J]. 云南民族大学学报, 2006, 15(4): 285-286.
- [7] 赵栋. 基于 HMM 的人脸识别研究与实现[D]. 北京: 北京化工大学, 2009.
- [8] Samaria F. Face recognition using hidden markov model[D]. Cambridge: University of Cambridge, 1994.
- [9] Nefian A V, Hayes M H. Hidden markov models for face recognition[C]//Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing. Seattle, Washington: [s. n.], 1998: 2721-2724.
- [10] 赵振勇, 王宝华, 王力, 等. 人脸图像的特征提取[J]. 计算机技术与发展, 2007, 17(5): 223-224.
- [11] 徐毅琼, 李弼程, 王波. 基于隐马尔可夫模型的自动人脸识别方法[J]. 计算机应用, 2004, 24(12): 226-227.
- [12] Kohir V V, Desai U B. Face Recognition Using A DCT-HMM Approach[C]//Applications of Computer Vision, 1998. [s. l.]: [s. n.], 1998: 229-230.
- [13] 尹飞, 冯大政. 基于 PCA 算法的人脸识别[J]. 计算机技术与发展, 2008, 18(10): 32-33.

chanics, 1999, 31: 417-457.

- [6] Zhao Jinping, Huang Daji. Mirror extending and circular spline function for empirical mode decomposition method[J]. Journal of Zhejiang University, 2001, 2(3): 247-252.
- [7] 苏玉香, 刘志刚, 李科亮, 等. 一种改善 EMD 端点效应的新方法及其在谐波分析中的应用[J]. 电工电能新技术, 2008, 27(2): 33-36.
- [8] 顾小丽, 李培良, 谭海涛, 等. 基于 RBF 神经网络的 EMD 方法在海平面分析中的应用[J]. 海洋与湖沼, 2009, 40(5): 532-534.
- [9] 朱金龙, 邱晓晖. 正交多项式拟合在 EMD 算法端点问题中的应用[J]. 计算机工程与应用, 2006(23): 72-74.
- [10] 黄先祥, 李胜朝, 谢建. 新型经验模式分解端点效应消除方法[J]. 机械工程学报, 2008, 44(9): 1-5.
- [11] 胡劲松, 杨世锡. EMD 方法基于 AR 模型预测的数据延拓与应用[J]. 振动、测试与诊断, 2007, 27(2): 116-121.
- [12] 程军圣, 于德介, 杨宇. Hilbert-Huang 变换端点效应问题的探讨[J]. 振动与冲击, 2006, 24(6): 40-47.
- [13] 舒忠平, 杨智春. 抑制经验模分解边缘效应的极值点对称延拓法[J]. 西北工业大学学报, 2006, 24(5): 639-643.