

文章编号: 1001-0920(2005)09-1073-04

基于Gabor小波和HMM的人像鉴别算法

曹林^{1,2}, 王东峰¹, 刘小军¹, 邹谋炎¹

(1. 中国科学院电子学研究所, 北京 100080; 2 北京信息工程学院 信息与通信工程系, 北京 100101)

摘要: 探讨了利用Gabor小波和隐马尔可夫模型(HMM)进行人脸识别的方法。首先对人脸图像进行多分辨率的Gabor小波变换; 然后在图像上放置一组网格结点, 每个结点用该结点处的多尺度Gabor幅度特征描述, 采用独立元分析法对每个结点进行去相关和降维; 最后形成特征结, 把每个特征结作为观测向量, 对隐马尔可夫模型进行训练, 并将优化的模型参数用于人脸识别。ORL人脸库的实验结果表明, 该方法识别率高, 工程上易于应用。

关键词: 人脸识别; Gabor小波变换; 独立元分析; 隐马尔可夫模型

中图分类号: TP391 文献标识码: A

Face Recognition Based on Gabor Wavelets and Hidden Markov Model

CAO Lin^{1,2}, WANG Dong-feng¹, LIU Xiao-jun¹, ZOU Mou-yan¹

(1. Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China; 2 Department of Information Technology and Telecommunication Engineering, Beijing Information Technology Institute, Beijing 100101, China
Correspondent: CAO Lin, Email: charlin26@163.com)

Abstract: A new method based on Gabor wavelets transform and hidden Markov model(HMM) for face recognition is proposed. The Gabor wavelet representation of an image is the convolution of the image with a family of multi-resolution Gabor kernels. The vectors called nodes, over a dense grid of image points are formed. Each node is labeled with a set of complex Gabor wavelets coefficients. The magnitudes of the coefficients are used for recognition. Feature nodes, as observation vectors of HMM, are derived by using independent component analysis to reduce the dimensionality of nodes. A set of images representing different instances of the same person is used to train each HMM. Each individual in the database is represented by an optimal HMM face model. Experimental results on the ORL face database show that the proposed algorithm provides a high recognition rate with a good perspective.

Key words: Face recognition; Gabor wavelets transform; Independent component analysis; Hidden Markov model (HMM)

1 引言

人脸识别在视频监控、人机交互等方面具有良好的应用前景, 已成为模式识别、图像处理等领域的研究热点^[1]。Nefian^[2]采用在矩形抽样窗口进行2D-DCT变换, 以提取HMM所需要的观测向量。此后, 许多学者在利用HMM进行人脸识别时, 都采用在矩形窗口进行某种变换来提取特征。常见的变换有

KL T^[2]、小波变换^[3]、奇异值分解^[4]等。当1D-HMM识别率不高时, 可采用更复杂的嵌入式隐马尔可夫模型(EHMM)^[2]或2D-HMM^[5]。工作的重点是如何简化这些复杂模型, 而忽略了对有效特征的抽取。

本文提出了基于Gabor小波和HMM的人脸识别算法, 在人脸图像上放置一组网格结点, 每个结点的特征用该结点处的多尺度Gabor幅度特征描述,

收稿日期: 2004-09-14; 修回日期: 2005-02-04

基金项目: 中国科学院科技创新基金项目(1021-07)。

作者简介: 曹林(1977-), 男, 沈阳人, 讲师, 博士, 从事模式识别、人脸识别等研究; 邹谋炎(1941-), 男, 四川遂宁人, 研究员, 博士生导师, 从事数字图像处理、宽带无线网络等研究。

并用 1D-HMM 来描述这些结点之间的关联。实验结果表明,本文方法的识别率与EHMM 方法的识别率相当,但仅采用了 1D-HMM,所以计算复杂度较低。

2 特征提取

二维 Gabor 小波的核函数定义为^[6]

$$\psi_{v,\mu}(z) = \frac{k_{v,\mu}^2}{\sigma^2} e^{-\frac{k_{v,\mu}^2 z^2}{2\sigma^2}} \times (e^{ik_{v,\mu}z} - e^{-\sigma^2/2}), \quad (1)$$

$$k_{v,\mu} = (k_v \cos \varphi_\mu, k_v \sin \varphi_\mu). \quad (2)$$

其中: φ_μ 和 k_v 分别定义了波向量 $k_{v,\mu}$ 的方向和尺度,

• 定义了向量范数, $z = (x, y)$; $k_v = k_{\max}/f^v$, $\varphi_\mu = \pi\mu/8$; f 为频率域中的采样步长,通常取 $f = \sqrt{2}$; k_{\max} 对应最大的采样频率,取 $k_{\max} = \pi/2$; 参数 σ 决定了高斯窗的宽度与波向量长度的比率,本文取 $\sigma = 7\pi/4$ 。

用 $I(z) = I(x, y)$ 表示图像的灰度分布,则图像 I 与 Gabor 小波 $\psi_{v,\mu}$ 的卷积为

$$O_{v,\mu}(z) = I(z) \otimes \psi_{v,\mu}(z), \quad (3)$$

其中符号 \otimes 表示二维卷积。一个图像经过 40 个 Gabor 小波滤波后,形成的集合为

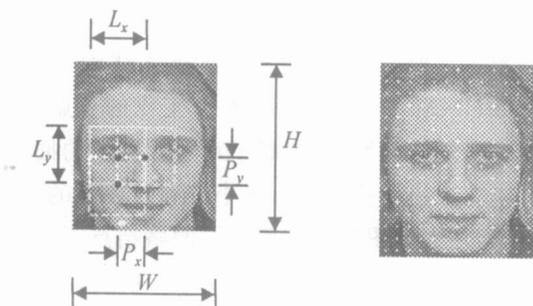
$$S = \{O_{v,\mu}(z); v \in \{0, \dots, 4\}, \mu \in \{0, \dots, 7\}\}.$$

应用卷积定理,通过快速傅里叶变换(FFT)可以计算 $O_{v,\mu}(z)$ 。

由式(3)可知,对于图像上任意一点 $z_i = (x_i, y_i)$,经过 40 个 Gabor 滤波器后形成一个特征矢量。把这个特征矢量定义为“结”,用 $J(z_i)$ 表示,即

$$J(z_i) = \{J_{v,\mu}(z_i) = |O_{v,\mu}(z)|_{z=z_i}; v \in \{0, \dots, 4\}, \mu \in \{0, \dots, 7\}\}. \quad (4)$$

本文在进行特征提取后,使用隐马尔可夫模型(HMM)表示这些结之间的关联,并用图 1 所示方法对图像进行均匀采样。



(a) 滑动窗口示意 (b) 均匀采样示意

图 1 均匀采样方案

假设人脸图像的宽度为 W , 高度为 H 。本文用宽度为 L_x 高度为 L_y 的滑动采样窗 ($L_x \times L_y$), 对图像从上到下、从左到右进行采样, 采样窗水平方向的

重叠为 P_x , 垂直方向的重叠为 P_y 。采样时只将采样窗口的中心点作为人脸图像的结, 忽略窗口内的其他点, 即一个采样窗对应一个结。采样个数由下式确定:

$$E = \left(\left\lfloor \frac{H - L_y}{L_y - P_y} \right\rfloor + 1 \right) \left(\left\lfloor \frac{W - L_x}{L_x - P_x} \right\rfloor + 1 \right). \quad (5)$$

其中 $\lfloor \cdot \rfloor$ 符号表示向下取整。本文采样窗口的大小为 7×7 , 重叠为 $P_x = 0, P_y = 0$ 。

由式(4)可知, $J(z_i) \in \mathbb{R}^D$, 其中 $D = 40$ 。文献[7]指出, 人脑总是在低维空间中完成诸如相似性判别的感知任务, 因此低维对于机器学习是非常重要的。此外, 维度的降低也减少了下一步使用 HMM 对人脸图像进行训练的计算量。独立元分析方法(ICA)是一种良好的去相关方法, 它利用数据的高阶统计量获取数据的独立成分, 从而提取随机数据集的本质特征。文献[8]采用一种最小化互信息的优化准则, 从而求出变换矩阵 F 。其主要步骤如下:

1) 对随机向量 $J(z_i)$ 进行白化变换, 去除各分量之间的相关性, 并使 $J(z_i)$ 的协方差矩阵为单位阵;

2) 利用高阶累积量(κ -statistics)对白化后的数据进行一系列旋转变换, 并用最小化互信息准则求出初始的变换矩阵 \tilde{F} 。

3) 对矩阵 \tilde{F} 中的每列进行标准化, 最后求出变换矩阵 F 。

由于 $F \in \mathbb{R}^{D \times d}$, 可将原始随机向量 $J(z_i)$ 变换成低维向量 $J_F(z_i) \in \mathbb{R}^d$, 即

$$J(z_i) = F J_F(z_i). \quad (6)$$

低维向量 $J_F(z_i) \in \mathbb{R}^d$ 表达了原始向量 $J(z_i)$ 的本质属性。把经过变换后的结 $J_F(z_i)$ 定义为“特征结”, 并把它作为下一步 HMM 的观测向量。

3 隐马尔可夫模型

3.1 基本概念

HMM 是以参数表示的用于描述随机过程统计特性的一种概率模型^[9]。HMM 的基本参数组成为: N 表示隐含状态数, S 表示隐含状态, $O = \{o_i\}$ 表示观测序列, E 表示观测序列的长度, M 表示高斯概率混合成分的个数, $A = \{a_{ij}\}$ 表示状态转移矩阵, $B = \{b_j, (o_i)\}$ 表示观测层的概率分布矩阵, $\Pi = \{\pi_i\}$ 表示初始状态概率分布。本文将 HMM 简记为 $\lambda = (A, B, \Pi)$ 。

3.2 HMM 训练

HMM 训练就是为每个类别确定一组经过优化的 HMM 参数, 每个模型可用单幅或多幅图像进行

训练 具体训练步骤如下:

1) 对人脸进行 Gabor 变换, 结合 ICA 求出特征结, 并将其作为观测向量, 即 $o_i = J_F(z_i)$.

2) 建立一个通用的模型 $\lambda = (A, B, \Pi)$, 确定模型的状态数, 允许的状态转移和高斯混合概率成分的个数

3) 将训练数据均匀分割, 并与 N 个状态对应, 计算模型的初始参数 对于状态转移矩阵 $A = \{a_{ij}\}$, 当 $j < i$ 或 $j > i + 1$ 时, 可使 $a_{ij} = 0$; 对于初始概率分布 $\Pi = \{\pi_i\}$, 可使 $\pi_1 = 0, \pi_i = 0 (i = 1)$, 即 HMM 从第 1 个状态开始

4) 用 V iterbi 分割^[9] 取代均匀分割, 并用分段 K 均值聚类方法^[9] 求出高斯混合模型的参数, 迭代调整初始模型参数

5) 采用 Baum-Welch 算法^[9] 对参数进行重新估计, 并用最后得到的优化模型参数表示人脸数据库中的某个类别

3.3 HMM 人脸识别

在人脸识别阶段, 首先对待识别的人脸图像 k 进行 Gabor 变换, 计算它的特征结, 形成观测序列 $O^{(k)}$; 然后使用前项-后项算法^[9] 计算每个训练模型 λ 产生该序列的概率 $P(O^{(k)}|\lambda)$, 最大值所对应的模型就是待识别人脸所属的类别 可用公式表达为

$$\lambda_i = \arg \max_j P(O^{(k)}|\lambda_j). \quad (7)$$

如果第 n 个模型 λ_n 产生序列 $O^{(k)}$ 的概率最大, 则将图像 k 归入第 n 类

4 实验结果及分析

本文使用 ORL 人脸数据库, 该数据库包含 40 人, 每人 10 幅图像, 共 400 幅图像 每幅图像大小为 $W = 92, L = 112$ 实验中每人取 5 幅共 200 幅图像进行训练, 用另外 200 幅进行识别

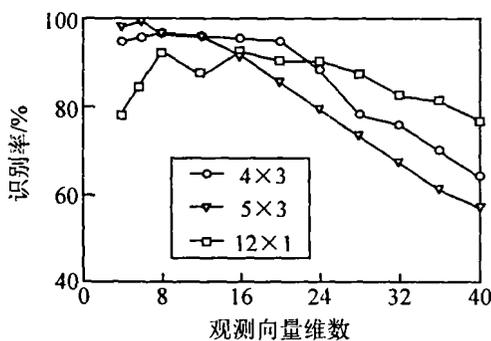


图 2 7 × 7 窗口下观测向量维数与识别率的关系

图 2 显示了 7 × 7 窗口下观测向量维数与识别率的关系 其中图例表示 HMM 状态数和高斯概率混合成分数 例如 5 × 3 表示本条曲线是在状态数 $N = 5$, 混合成分数 $M = 3$ 的条件下绘制的

从图 2 可以看出, 观测向量维数在 $d = 6$ 附近识别率较高 当 $d > 16$ 时, 随着维数的增高, 识别率降低; 当 $d < 6$ 时, 随着维数的增高, 识别率递增 这是因为维数太少, 包含的人脸特征也少, 所以识别率降低; 维数过多, ICA 所起的作用减弱, 观测向量之间的相关性增大, 干扰了 HMM 模型训练图像的能力, 因此识别率降低

图 3 显示了在 7 × 7 窗口下, 当观测向量维数 $d = 6$ 时, 状态个数 N 和高斯概率混合成分数 M 对识别率的影响 图中每条曲线都是在某个状态数下绘制的

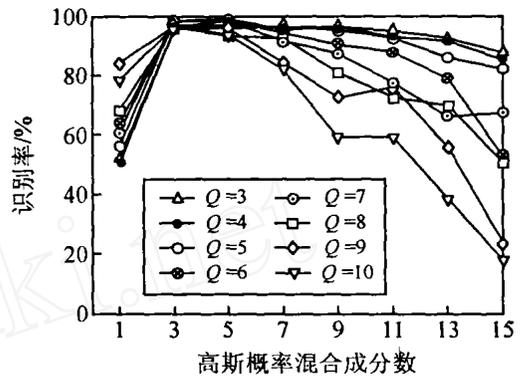


图 3 状态数和高斯概率混合成分数对识别率的影响

从图 3 可以看出, 保持状态数不变, 随着 M 的增加, 识别性能得到改善, 但当 M 增加到一定量时, 若再增加 M , 则识别率降低 由于用高斯混合模型描述概率分布矩阵 B , 而矩阵 B 描述了每个状态的概率统计属性, 适当增加 M 可以增加高斯混合模型的自由度, 能更好地表现每个状态的统计属性, 进而提高识别率 但是如果 M 增加过量, 会使高斯混合模型过分依赖于训练数据, 以至于测试数据与训练数据稍有不同, 就无法识别出测试数据 这实际上是过学习的结果

表 1 相关算法在 ORL 人脸库上识别率对比

方 法	观测向量维数	状态个数	观测序列长度	识别率 (%)
ID-HMM-L ^[10]	920	5	103	84.0
PHMM-L ^[10]	80	24	2 236	94.5
ID-HMM-D ^[2]	39	5	52	86.0
EHMM-D ^[2]	6	24	2 236	99.0
本文方法	4	7	208	99.0

注: L 表示直接取图像灰度, D 表示进行 DCT 变换

表 1 给出了相关算法在 ORL 人脸库上的识别率对比 从表中可以看出, Samaria^[10] 直接取图像灰度作为观测向量, 采用一维 HMM 识别率只有 84%, 采用伪二维 PHMM 识别率也只有 94.5%, 而且观测向量的维数高, 观测向量的个数多. Nefian^[2] 采用

2D-DCT 系数作为观测向量,降低了观测向量的维数,识别率为 86%,与 Samaria 的识别效果相当。尽管 Nefian 采用的 EHMM 具有较高的识别率,但它实际上是伪 2D-HMM 模型,观测向量数多,状态总数多,对模型进行训练所需的时间较长。本文方法的识别率达到 99%,而且仅采用 1D-HMM,观测向量数较少,状态数也少,因此整体性能较好。

5 结 语

本文提出一种基于 Gabor 变换且结合 ICA 和 HMM 的人脸识别方法,并与 Nefian 方法进行了比较。实验结果表明,本文方法识别率较高,具有良好的发展前景。但本文算法在特征提取上所需的计算量较大,在今后的研究中可采用快速算法降低复杂度,使本算法更有利于工程应用。

参考文献(References)

- [1] Gong S, Psarrou A. *Dynamic Vision: From Images to Face Recognition* [M]. London: Imperial College Press, 2000: 5-20
- [2] Nefian A. *A Hidden Markov Model-based Approach for Face Detection and Recognition* [D]. Georgia: Georgia Institute of Technology, 1999
- [3] Bicego M, Castellani U, Murino V. Using Hidden Markov Models and Wavelets for Face Recognition [A]. *Proc of the 12th Int Image Analysis and Processing* [C]. Mantova, 2003: 52-56
- [4] 刘小军, 邹谋炎. 一种基于奇异值分解和隐马尔可夫模型的人脸识别方法 [J]. *计算机学报*, 2003, 26(3): 340-344
(Liu X J, Zou M Y. An Approach for Face Recognition Based on Singular Value Decomposition and Hidden Markov Model [J]. *Chinese J of Computers*, 2003, 26(3): 340-344)
- [5] Othman H, Aboulnasr T. A Separable Low Complexity 2D HMM with Application to Face Recognition [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2003, 25(10): 1229-1238
- [6] Lades M, Vorbruggen J C, Buhmann J. Distortion Invariant Object Recognition in the Dynamic Link Architecture [J]. *IEEE Trans on Computers*, 1993, 42(3): 300-311
- [7] Tenenbaum J, Silva V, Langford J. A Global Geometric Framework for Nonlinear Dimensionality Reduction [J]. *Science*, 2000, 290(22): 2319-2323
- [8] Comon P. Independent Component Analysis: A New Concept [J]. *Signal Processing*, 1994, 36(3): 287-314
- [9] Rabiner L. A Tutorial on Hidden Markov Models and Selected Application in Speech Recognition [J]. *Proc of the IEEE*, 1989, 77(2): 257-286
- [10] Samaria F. *Face Recognition Using Hidden Markov Model* [D]. Cambridge: University of Cambridge, 1994

(上接第 1072 页)

- [5] Fei S M. An H Approach to Robust Stabilization of Nonlinear Uncertain Systems, 1995, 10(5): 390-394
- [6] Wu M, Zhang L B, Gui W H. Analysis and Synthesis for Robust Control of Uncertain Nonlinear Systems [J]. *Control Theory and Applications*, 2002, 19(2): 203-206
- [7] Isidori A. *Nonlinear Control Systems* [M]. 3rd ed. New York: Springer, 1995
- [8] Daniel Liberzon, Eduardo D Sontag, Wang Yuan. Universal Construction of Feedback Laws Achieving ISS and Integral-ISS Disturbance Attenuation [J]. *Systems and Letters*, 2002, 46(2): 111-127
- [9] Wei L. Global Robust Stabilization of Minimum-phase Nonlinear Systems with Uncertainty [J]. *Automatica*, 1997, 33(3): 453-462