

浙江大学

本科生毕业论文(设计)



题目 基于步态的人体识别

姓名与学号 冯继雄 3090101374

指导教师 于慧敏 (教授)

年级与专业 09 信息与通信工程

所在学院 信息与电子工程学系

摘要

步态即人走路的姿态，步态识别是通过步态来识别人体。本文采用隐马尔可夫模型（Hidden Markov Model, HMM）作为主要方法进行识别，同时和其他两种方法进行比较。HMM 是一种具有较强的时间序列建模能力的信号模式处理工具，在语音识别中获得了广泛应用。由于步态序列具有时序周期性，可以看做是“隐藏的状态”，故把 HMM 引入步态的识别中。训练时，首先提取待训练数据库人物的轮廓点到质心的距离作为步态特征；其次引入 CHMM（Continuous HMM, 连续 HMM）处理连续数据，采用高斯混合模型（Gaussian Mixture Model, GMM）对 CHMM 进行参数初始化，用 Baum-Welch (BW) 算法实现参数重估，建立每个人的 CHMM 模型；然后提取测试序列的步态特征；最后通过前向算法求测试序列的最大似然概率值来识别人体。实验结果证明了该方法具有较好的识别效果。

关键词： 步态识别 CHMM 轮廓质心距 模式识别

Abstract

Gait means the posture of a walking person. Gait recognition uses gait to recognize human bodies. This paper uses HMM as a major method for recognition. HMM is a kind of signal pattern processing tool with strong ability to modeling time series, which has been used widely in speech recognition. As gait sequences have timing periodic property, which can be seen as the hidden states, this paper introduces HMM for gait recognition. When training, extract contour centroid distance as gait characteristics; then, use CHMM to process continuous data, GMM to initialize CHMM's parameters and Baum-Welch algorithm to re-estimate parameters, in order to build up CHMM for each person; finally get the maximum probability of test sequences by forward algorithm to recognize each person. The experiment shows a good recognition result.

Keywords: Gait Recognition, CHMM, Contour Centroid Distance, Pattern Recognition

目录

第一章 引言	1
1.1 背景	1
1.2 已往的研究	1
1.3 最新的研究成果	3
1.4 本文的主要内容	4
第二章 识别实验过程	5
2.1 实验模型简介	5
2.2 实验数据的选取	8
2.3 提取图像的步态特征	8
2.4 初始化 CHMM 参数	9
2.5 训练步态数据	10
第三章 识别结果和优化	13
3.1 识别结果	13
3.2 其他方法对比	16
3.3 结果优化	17
第四章 总结	19
参考文献	20
致谢	22

第一章 引言

1.1 背景

人的步态即人走路时的姿态。步态很早就是医学和心理学的研究课题，早期的医学研究指出，人的步态中有 24 种不同的成分，如果把这 24 种成分都考虑到，则步态是为个体所特有的^[1]。近年来随着计算机视觉和生物特征识别领域技术的发展，自动步态识别成为一个备受关注的研究方向。与其他生物特征识别技术(例如指纹识别，人脸识别，虹膜识别等)相比，步态识别的优势在于远距离，不接触，很难伪装。基于上述特点，步态识别在安防系统、医疗诊断和刑侦等领域具有广泛的应用前景和经济价值，因此激发了国内外广大科研工作者的研究热情。美国国防部高级研究计划署 2000 年重大研究项目 HID（远距离身份识别）就包括步态识别，这个项目对步态识别的研究产生了深远的影响^[2]。

步态识别在理论上取得进展的同时，在实际应用中也有产品出现，比如芬兰一个研究中心研制成功基于步态识别的电脑放到装置^[3]。然而，当前步态识别仍然处于初期阶段，步态识别中的各种变化因素比如时间、服饰、行走路面、视角和鞋帽都会对识别率产生影响，实际中的情况往往更复杂，因此这方面的研究还需要更多投入。

1.2 已往的研究

步态识别通常对人走路时的图像序列进行分析处理，通常包含运动检测、周期检测、特征提取、识别算法 4 个过程。

运动检测是从图像序列中提出运动的目标（比如正在行走的人），这样就可以把研究的关注点放在人体的姿势上而忽略背景。通常，运动检测受扰动、光照、影子的因素的干扰较大，因此目标的精确提取也是一个可以深入的研究方向。有的步态数据库比如中科院自动化所的 CASIA^[4]中自带背景图片，可以用背景减除法方便处理。

通过目标检测和其他辅助手段处理（比如滤波、二值化、数学形态学处理）后，可以从图像序列中提取出人体区域。随后，就要进行下一个环节：周期检测。

考虑到计算复杂度和实时处理的需求,当前步态周期检测主要是通过分析人体宽度信号或人体区域像素点数来实现的。一般检测的周期图 1.1 所示^[5]。

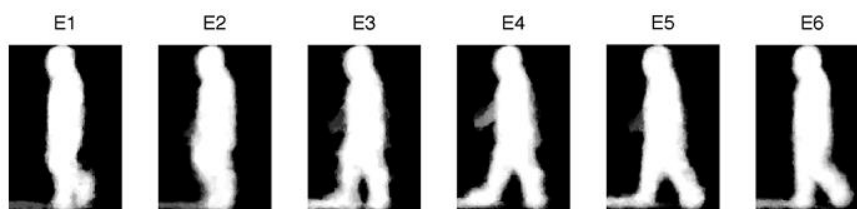


图 1.1 检测的步态周期

特征提取即提取每个人走路时候的特征,根据提取方法的差异可将步态识别的方法大致分为两类:基于模型的方法和非模型的方法。基于模型的方法首先要建立人体结构模型或者是人体运动模型,然后获取特征,主要有以下几种主要的模型:椭圆模型^[6](图 1.2)、钟摆模型^[7](图 1.3)、3 维模型^[8](图 1.4)。



图 1.2 椭圆模型

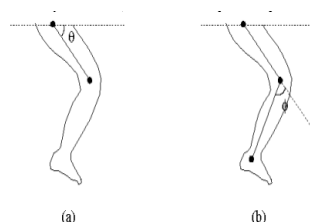


图 1.3 钟摆模型

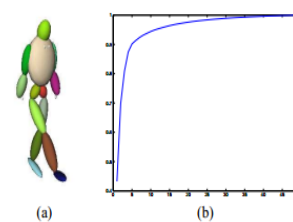


图 1.4 3D 模型

非模型的方法无需建立这些模型,而是从其他方面提取特征,比如 Wang 等人^[9]则提取轮廓点到人体质心的距离向量(图 1.5), Boulgouris 等人^[10]把轮廓图像以 $3\sim 5^\circ$ 的角度分割,然后计算每个区域的平均距离(图 1.6)。

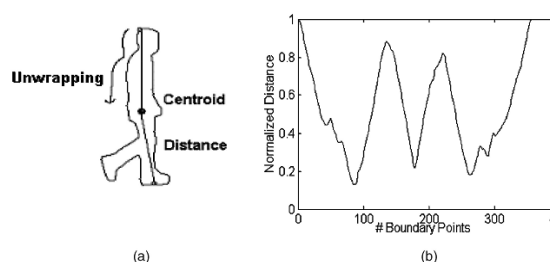


图 1.5 轮廓到质心距离

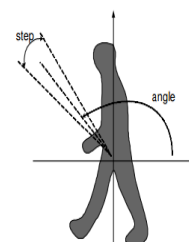


图 1.6 角度区域平均距离

得到步态特征之后，就可以进行识别步骤，一般是事先对要识别的人提取样本进行训练，然后输入待测序列，完成识别任务。当前步态识别研究中采用的两个主要方法是模板匹配和统计方法。模板匹配方法广泛应用于模式识别领域，是一种基本的特征匹配方法。比如归一化欧氏距离度量和统计形状信息分析和动态时间规整（Dynamic Time Warping, DTW），其中 DTW 能够对存在全局或局部扩展、压缩或变形的模式进行匹配，解决动态模式的相似度和匹配问题。

模板匹配实际是一种基于距离度量的方法，在度量的基础上增加了匹配的步骤，但是在步态识别中使用这种方法就忽略了隐含在步态序列中的姿态之间的联接关系，而统计方法中的隐马尔可夫模型^[11](HMM)可以很好地融合这些信息。步态周期可以看作一个具有隐含过程和表面过程的双重随机过程，其中隐含过程用姿态之间的转移来描述，而表面过程则用特定姿态下的图像特征刻画。通过对样本库中每个目标调整模型参数 $\lambda = (\pi, A, B)$ ，使得同一目标的测试序列的 $P(O|\lambda)$ 达到最大。文献^[12]中使用半个步态周期的 5 个关键姿态（图 1.7）的宽度向量作为特征，通过计算给定图像与样本姿态的距离产生低维观测向量，并以此训练 HMM 的参数进行步态识别。



图 1.7 轮廓方法检测的 5 个周期

1.3 最新的研究成果

目前大部分文献中采用的数据库以及训练测试方法都不尽相同，难以有确切的性能评测。但为了对目前典型方法的实验结果有较为直观的认识，表 1.1 汇集了步态识别领域中的部分典型方法的实验结果^[13]。

表 1.1 步态识别领域典型方法及其识别结果

方法名称	识别率	时间
Wang 等人静态+动态特征融合算法	87.5%~91.25% (CASIA) ^①	2004
Kale 等人基于 HMM 的算法	54.1%~95.8% (CMU Mobo)	2004
Sarkar 等人 Baseline 算法	26.0%~87.0% (USF)	2005
Han 等人的 GEI 算法	31.0%~100% (USF)	2006
Boulgouris 等随机变换+LDA 降维算法	19.0%~87.0% (MIT AI)	2007
Zhang 等五连杆双足人体模型算法	61% (USF), 96% (CMU Mobo)	2007
Geng 等人脸和步态特征融合算法	85.0%~95.0% (CASIA)	2008
Lu 等人的 MPCA+LDA+Boosting 算法	30.0%~100% (USF)	2009
Kusakunniran 等 GEI+TSVD 多视角算法	90% (CASIA)	2009
Kellokumpu 等纹理模式+Adaboost 算法	16.0%~89.0% (CMU Mobo)	2010
Lam 等人的 GFI 算法	3% ~ 89% (USF)	2011
Yoo 等人的 2D 人体杆状模型算法	84% (Soton Large Database)	2011

注：① 括号里的内容表示实验时采用的步态数据库，这里用到步态数据库有 CASIA, CMU Mobo, USF, MIT AI 和 Sonton Large Database。

1.4 本文的主要内容

本文采用隐马尔可夫模型（HMM）作为主要方法进行步态图像的识别，并和其他两种较为简单的方法（最小欧氏距离、傅里叶描绘子）的结果进行比较，分析各自的优劣势。为了让实验更具有说服性和可信度，文中的实验都在 Matlab 软件上进行，并且将关键的步骤加以阐明。

下面，文章将就训练、识别和优化三个内容详细展开说明。

第二章 步态数据的预处理

2.1 实验模型简介

HMM 是一种具有较强的时间序列建模能力的信号模式处理工具，在语音识别中获得了广泛应用^[11]。由于步态序列具有时序周期性，可以看做是“隐藏的状态”，故把 HMM 引入步态的识别中。由于步态的特征数据是连续的，因此需要用到 CHMM（Continuous HMM，连续 HMM）^[14]。

假设有 K 个观测矢量序列如下

$$O = \{O_1, O_2, \dots, O_K\} \quad (2.1)$$

式中， O_k ——第 k 个观测矢量序列

$$O_k = \{O_{k1}, O_{k2}, \dots, O_{kT}\} \quad (2.2)$$

O_t ——D 维第 t 个观测序列

为了得到矢量序列的统计特征，选择 N 个状态和 M 个混合数如下

$$S = \{S_1, S_2, \dots, S_N\} \quad (2.3)$$

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & x_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NM} \end{bmatrix} \quad (2.4)$$

式中， x_{nm} ——状态 S_n 的第 m 混合数

则 CHMM 可以表示如下

$$\lambda = [\pi, A, C, \mu, U] \quad (2.5)$$

(1) $\pi = [\pi_1, \pi_2, \dots, \pi_N]$ 为初始状态分布，用于描述观测序列 O 在 t=1 时刻所处状态 q_1 的概率分布，即

$$\pi_i = P(q_1 = S_i) \quad i = 1, 2, \dots, N \quad \sum_{i=1}^N \pi_i = 1 \quad (2.6)$$

(2) $A = \{a_{ij} | i, j = 1, 2, \dots, N\}$ 为状态转移矩阵，这里只考虑一阶 HMM，当前所处

状态 q_i 只与前一时刻所处状态 q_{i-1} 有关，即

$$a_{ij} = P(q_i = S_j | q_{i-1} = S_i) \quad \sum_{j=1}^N a_{ij} = 1 \quad (2.7)$$

式中， a_{ij} ——从状态 q_i 到状态 q_j 的概率

(3) C 为混合权重

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1M} \\ c_{21} & c_{22} & \cdots & c_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & \cdots & c_{NM} \end{bmatrix} \quad c_{nm} \geq 0 \quad \sum_{m=1}^M c_{nm} = 1 \quad (2.8)$$

式中， c_{nm} ——状态 n 中第 m 个高斯密度混合权重

(4) μ 为均值矢量

$$\mu = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1M} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{N1} & \mu_{N2} & \cdots & \mu_{NM} \end{bmatrix} \quad (2.9)$$

式中， μ_{nm} ——在状态 n 中第 m 个混合高斯密度的 D 维均值矢量

(5) U 为协方差矩阵

$$U = \begin{bmatrix} U_{11} & U_{12} & \cdots & U_{1M} \\ U_{21} & U_{22} & \cdots & U_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ U_{N1} & U_{N2} & \cdots & U_{NM} \end{bmatrix} \quad (2.10)$$

式中， U_{nm} ——在状态 n 中第 m 个高斯密度的 $D \times D$ 维协方差矩阵。

对于 CHMM 模型来说，其观测序列通常都是假设由高斯概率密度函数产生的，观测序列概率密度函数的形式采用混合高斯概率密度函数 (Mixture Gaussian PDF)

$$\begin{aligned} b_{t,n}(o) &= \sum_{m=1}^M c_{nm} \times b_{t,nm} \\ &= \sum_{m=1}^M c_{nm} \times \frac{1}{(2\pi U_{nm})^{\frac{1}{2}}} \exp \left[-\frac{1}{2U_{nm}} (O_t - \mu_{nm})^T (O_t - \mu_{nm}) \right] \end{aligned} \quad (2.11)$$

式中， $b_{t,nm}$ ——第 n 状态中第 m 个分量的单高斯概率密度函数

在本文的实验中，以上模型的参数选取如表 2.1 所示：

表 2.1 本实验中 CHMM 参数的选择

K 观测序列数	T 图像帧数	D 特征点数	N 状态数	M 高斯混合数
1	48	300	5	3

与 DHMM 相同，CHMM 模型为了在实际中能够得到更好的应用，必须要解决三个基本问题。

1、HMM 的概率推理问题 (Forward-Backward 算法)：给定观测序列 $O=\{O_1, O_2, \dots, O_K\}$ 和模型 λ ，怎样有效地计算观测变量序列在给定模型下的概率 $P(O|\lambda)$ ？

2、HMM 的解码问题 (Viterbi 算法)：给定观测序列 $O=\{O_1, O_2, \dots, O_K\}$ 和模型 λ ，怎样选择一个相应的状态序列 $q=\{q_1, q_2, \dots, q_T\}$ 能够在某种意义上达到最优？

3、HMM 的训练问题 (Baum-Welch 算法)：如何调整模型参数使 $P(O|\lambda)$ 为更大？即模型的训练问题。

本文中关于 HMM 算法的使用情况如表 2.2 所示。

表 2.2 本论文中用到的算法及其用途

用到的算法	用途
Forward 算法	用于步态识别测试序列
Viterbi 算法	用于展示步态图像的状态周期变化
Baum-Welch 算法	用于训练步态数据库

计算 HMM 和 GMM 的相关参数时使用了 MIT 的 HMM Toolbox for Matlab^[15]，在此向工具箱的作者表示感谢。

2.2 实验数据的选取

采用中科院 CASIA Dataset A^[4]里面的 Silhouette 轮廓图像，轮廓图像是彩色图像经过背景减除、灰度化、滤波、二值化和数学形态学处理后的图像，同时精心选取处理效果较好的 8 个人（fyc, hy, ljg, lqf, lsl, ml, rj, zjf），每个人取 48 帧图像，并且用手工修补了一下有缺陷的图像，作为最终输入的训练和测试图像，处理过程如图 2.1 所示。

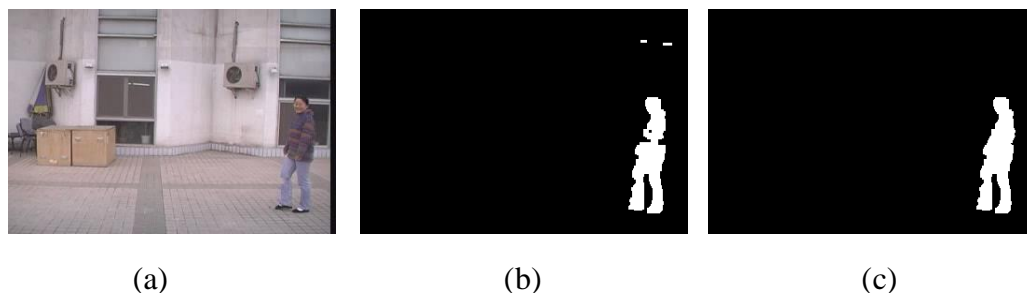


图 2.1 (a)是原始图像，(b)是二值化之后的图像，(c)是手工修补后的图像

2.3 提取图像的步态特征

步态特征有很多种，本文中取轮廓质心距^[16]，轮廓起始点开始作为参考起点，沿顺时针方向将轮廓边界展开为边界像素点 (x_i, y_i) 到质心 (x_c, y_c) 的距离所组成的一维信号 $d = (d_1, d_2, \dots, d_{300})$ ，其中

$$d_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad i=1, 2, \dots, 300 \quad (2.12)$$

这个距离信号间接地在一维空间中表达原始的二维轮廓形状，轮廓质心距是如图 2.2 所示

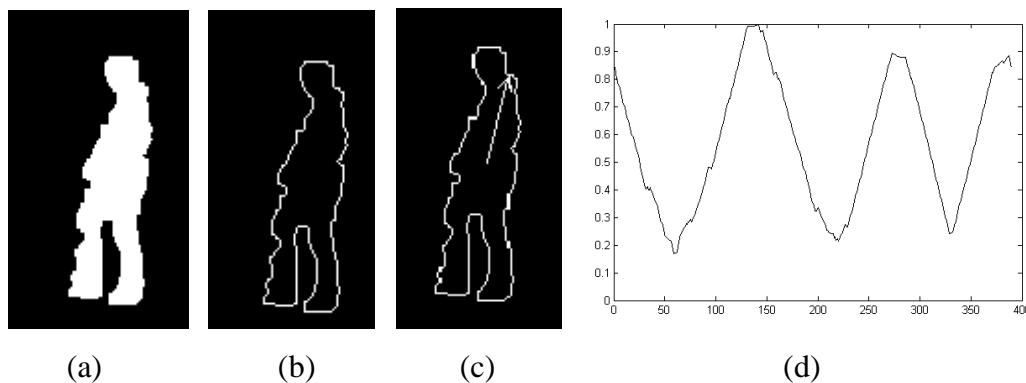


图 2.2 (a)是输入图像, (b)为提取的轮廓, (c)是计算轮廓质心距的示意图, (d)是得到的轮廓质心距

提取的距离数据长度不等, 为了便于处理只取其前 300 个点, 这样, 一个人的步态数据变为一个 48×300 的特征向量矩阵。

2.4 初始化 CHMM 参数

由训练数据得到 CHMM 的参数时, 一个重要的问题就是初始模型参数的选取。一般认为, 参数 π 和状态转移概率矩阵 A 对初始值的选取影响不大, 可以任意选取。重要的是参数 B 的选取。这里对 CHMM 的混合密度初估计, 采用 K-means 算法来实现。

本文中, 取初始状态分布

$$\pi = (0.2, 0.2, 0.2, 0.2, 0.2)^T \quad (2.13)$$

取从左到右型 HMM^[17] (如图 2.3 所示)

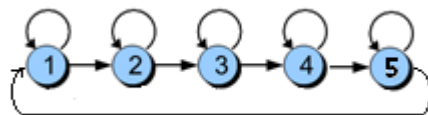


图 2.3 从左到右型 HMM 示意图

其状态转移矩阵 A 为

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.14)$$

参数混合权重 C_{NM} 随机产生,但是需要符合条件 $c_{nm} \geq 0 \sum_{m=1}^M c_{nm} = 1$, $N=5, M=3$ 。

对于均值矢量 \mathbf{u} 和协方差矩阵 \mathbf{U} 需要用 GMM 初始化^[18], 具体过程是: 先用 K-means 算法把数据聚为 $N \times M = 5 \times 3 = 15$ 类, 我们把隐含类别标签用 z^t 表示, 即 $z^t = j$ 表示 O_t 属于第 j 个类别 (对应的均值是 u_j , 协方差矩阵是 U_j)。找出第 j 个类的中心点设为均值向量 u_j , 然后计算协方差矩阵 U_j 。

$$U_j = \frac{\sum_{t=1}^T \mathbf{1}\{z^t = j\} (O_t - u_j)(O_t - u_j)^T}{\sum_{t=1}^T \mathbf{1}\{z^t = j\}} \quad (2.15)$$

根据得到的 u_j 和 U_j , 返回去计算 z^t , 如此循环下去直到收敛, 得到初始化的均值矩阵 μ_{NM} 和协方差矩阵 U_{NM} 。

2.5 训练步态数据

完成 CHMM 初始化后, 用 Baum-Welch 算法对 CHMM 模型进行参数重估。

在前述的前向-后向算法和 Baum-Welch 算法中, 都有前向变量 $\alpha_t(i)$ 和后向变量 $\beta_t(i)$ 的递归计算, 由于所有的量都小于 1, 因此 $\alpha_t(i)$ (随着 t 的增加) 和 $\beta_t(i)$ (随着 t 的减少) 都迅速地趋向于零, 为了解决这种下溢问题, 采取增加比例因子的方法对有关算法进行修正^[14]。

前向变量 α 和变量因子 α^*

初始变量:

$$\alpha_1(i) = \pi_i b_i(O_1) \quad 1 \leq i \leq N$$

$$g_1 = \frac{1}{\sum_{i=1}^N \alpha_1(i)} \quad 1 \leq i \leq N$$

$$\alpha_1^*(i) = g_1 \alpha_1(i)$$

递推方程:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t^*(i) a_{ij} \right] b_j(O_{t+1}) \quad (2.16)$$

$$g_t = \frac{1}{\sum_{j=1}^N \alpha_{t+1}(j)} \quad t = 1, 2, \dots, T-1$$

$$\alpha_{t+1}^*(i) = g_t \alpha_{t+1}(j) \quad 1 \leq j \leq N$$

后向变量 β 和后向变量因子 β^*

$$\beta_t(i) = 1 \quad 1 \leq i \leq N$$

$$\beta_t^*(i) = 1 \quad 1 \leq i \leq N$$

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}^*(j) \quad t = 1, 2, \dots, T-1 \quad (2.17)$$

$$\beta_t^*(i) = g_t \beta_t(i) \quad 1 \leq i \leq N$$

后向变量与前向变量相同，都引入了中间变量和比例因子，在运算上和前向变量相同。

对 α 和 β 做了上述处理之后，为了保持与原来公式的计算结果一致，必须在概率 $P(O|\lambda)$ 的计算公式中作相应的处理，消去比例因子的影响。

由的处理过程推出

$$\alpha_t^* = \alpha_t(i) g_1 \cdot g_2 \cdot \dots \cdot g_t$$

即

$$\sum_{j=1}^N \alpha_t(j) = g_1 \cdot g_2 \cdot \dots \cdot g_t \quad (2.18)$$

或

$$\lg P(O|\lambda) = - \sum_{t=1}^T \lg g_t$$

经过这样处理之后，前向概率和后向概率通过分子和分母中的约分，保证了和未引入变量因子时算法的一致性，同时又避免了计算的溢出问题。由此得到的是概率 $P(O|\lambda)$ 的对数值，而不是单纯概率值，避免了计算出的概率 $P(O|\lambda)$ 值太小。

定义三个变量

$$\begin{aligned}
\gamma_{t,n} &= P[q_t = S_n | O, \lambda] = \frac{\alpha_t^* \beta_t^*}{\sum_{i=1}^N \alpha_t^* \beta_t^*} = \sum_{t=1}^N \xi_t = \sum_{m=1}^M \gamma_{t,nm} \\
\xi_{t,ij} &= P[q_t = S_i, q_{t+1} = S_j | O, \lambda] = \frac{\alpha_t^* \beta_t^*}{\sum_{i=1}^N \alpha_t^* \beta_t^*} a_{ij} b_{t+1,j} \\
\gamma_{t,nm} &= P[q_t = S_n, r_t = x_{nm} | O, \lambda] = \frac{\alpha_t^* \beta_t^*}{\sum_{i=1}^N \alpha_t^* \beta_t^*} \times \frac{c_{nm} b_{t,nm}}{\sum_{m=1}^M c_{nm} b_{t,nm}}
\end{aligned} \tag{2.19}$$

引入三个变量的目的是为了使后面的 CHMM 参数重估更加简单。其中

$\gamma_{t,n}$ 表示 t 时刻 $q_t = S_n$ 的概率； $\xi_{t,ij}$ 表示 t 时刻 $q_t = S_i$ 并且 t+1 时刻 $q_{t+1} = S_j$ 的概率； $\gamma_{t,nm}$ 表示 t 时刻 $q_t = S_n$ 并且混合数为 x_{nm} 时的输出概率。

则重估公式如下

$$\begin{aligned}
\bar{\pi}_n &= \frac{\sum_{k=1}^K \gamma_1}{K} \quad \bar{a}_{ij} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \xi_t}{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \gamma_t} \quad \bar{c}_{ij} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_{t,nm}}{\sum_{k=1}^K \sum_{t=1}^{T_k} \sum_{m=1}^M \gamma_{t,nm}} \\
\bar{\mu}_{nm} &= \frac{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_{t,nm} O_t}{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_{t,nm}} \quad \bar{U}_{nm} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_{t,nm} (O_t - \mu_n)^2}{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_{t,nm}}
\end{aligned} \tag{2.20}$$

第三章 识别结果和优化

3.1 识别结果

8 个人的模型都建好以后，就可以进行识别了。实验中的样本图像序列如图 3.1 所示，实验用的测试序列也来自 CASIA Database A，拍摄角度也相同，是同一场景下另一次拍摄，如图 3.2 所示（实验用的是 Silhouettes 轮廓图，为了便于观察，用彩色图）。



图 3.1 训练用的样本图像序列



图 3.2 测试用的测试图像序列

可以从两组图像中看出，训练的图像和测试的图像在时间上是错开的，这也对步态识别提出了较高的要求，单纯的“匹配法”或者“最小欧氏距离”很难做出正确的识别。

用 CHMM 识别时，先提取测试图像的特征（过程同 2.2 节），得到 48×300 的特征矩阵，然后根据 Forward 算法计算每个人的 CHMM 模型下该特征矩阵所得的似然概率，取最大概率对应的人即为识别结果。似然概率的结算公式由下式给出

$$P(O_t | \lambda) = \alpha_T(N) = \sum_{i=2}^{N-1} \alpha_T(i) a_{iN} \quad (3.1)$$

其中， $\alpha_t(i)$ 表示第 t 个观测序列在状态 i 时的概率，而 a_{ij} 表示从状态 i 转移

到状态 j 的转移概率。

将 8 个人的测试序列分别计算 8 个似然概率，得到一个 8×8 的概率矩阵，如表 3.1 所示为一次计算结果

表 3.1 一次实验的计算结果

样本 测试	1	2	3	4	5	6	7	8
1	256	-1497	-Inf	-Inf	-5437	-Inf	-1036	-Inf
2	-1116	317	-165	-Inf	-1229	-2240	-929	-4288
3	-2740	-Inf	-256	-Inf	-682	-520	-2183	-2275
4	-6920	-3827	-2926	782	-587	486	-8463	-2862
5	-6172	-3629	-3119	-1338	-707	-1387	-6244	-4011
6	-3949	-Inf	-1279	-Inf	-901	874	-3867	-1237
7	-Inf	-Inf	-Inf	-Inf	-4948	-11603	-1363	-8434
8	-Inf	-Inf	-759	-Inf	-2617	-Inf	-Inf	-345

注①：表的行表示训练过的样本，表的列表示测试序列，表中数据为对数似然概率

$\ln[P(O_i | \lambda)]$;

注②：对数概率可以大于零是因为数据是连续分布；

注③：有背景色的单元格是一行里面最大的，Matlab 中的 -Inf 表示 $-\infty$ ；

注④：如果有背景色的单元格对应的行数等于列数，则为正确识别的一个结果

根据表格中的概率值，选择最大概率对应的人即识别结果。此表格的识别结果是 100%（没有一个人被识别错误）。

由于识别过程中的均值矩阵 u 是通过 K-means 算法随机产生的，在迭代的情况下趋于局部最优（而不是全局最优），所以每次实验的识别率是随机的。笔者重复计算 100 次，去掉无效的 6 次（因为程序原因出错）一共 94 次，画出识别率频度直方图如下图所示。

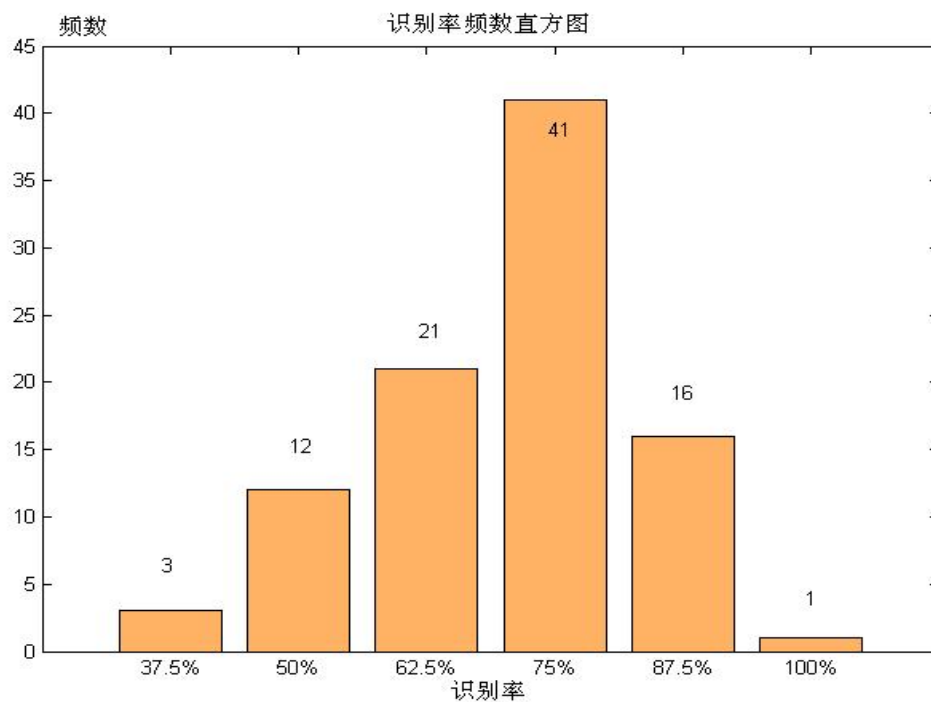


图 3.3 识别率频数直方图

由图 3.3 可知，识别率在 37.5%~100% 范围，以 75% 最常见。

表 3.2 实验结果中，测试图像及其对应的状态

图 像				
	4	5	1	1
图 像				
	1	1	1	2
图 像				
	3	4	5	5

通过 Viterbi 算法还可以得到图像序列对应的“状态”。截取几个测试图像的结果如表 3.2 所示，上一行是图像，下一行是对应的状态。（实验用的是 Silhouettes 轮廓图，为了便于观察，用彩色图）

从状态和图像序列来看，似乎并不太合理，这也是后面需要优化的地方。

3.2 其他方法对比

为了说明 HMM 方法在时间序列上具有的优势，本文又采用其他两种算法来作为对比。

表 3.3 “最小欧氏距离法”实验结果

模型 测试	1	2	3	4	5	6	7	8
1	2699	3812	2990	4063	4006	3369	2956	3823
2	3348	3754	2228	3299	3307	3750	3233	3201
3	3839	4276	3276	3517	3725	2446	3355	3628
4	3367	3728	3726	3578	4428	2956	3678	3494
5	4263	3283	3318	3188	2676	4350	3762	2961
6	3376	3832	3428	3862	4328	2361	3906	3487
7	3648	3489	3773	4373	3694	3717	4398	3866
8	3001	3772	2734	3309	3641	2732	3052	3059

第一种是“最小欧氏距离”法，特征向量仍然是 2.2 节中的 48 x 300 轮廓质心距，但是识别方法是求测试矩阵与每个人的特征矩阵的“距离”，取最小距离对应的人作为识别结果。距离的定义如下

$$\text{distance} = |A_{T \times D} - B_{T \times D}| = \sum_{t=1}^T \sum_{d=1}^D |a_{ij} - b_{ij}| \quad (3.2)$$

测试结果如表 3.3 所示。表 3.3 是测试序列的距离矩阵，选取最小距离对应的人作为识别结果，有背景色的是一行中数据最小的，如果在对角线上则说明识别正

确。

由表 3.3 可知,识别率固定在 37.5%(8 个人里面正确识别 3 个),用时 14.5s。

第二种方法是“傅里叶描绘子”法^[19]。如果把一个步态图像序列中每帧图像的轮廓坐标表示为二维数组 $S=\{(x_i, y_i)\}$, $i=1,2,\dots,N$, N 为边界的点数,本文中取 $N=300$, 如果不够则补零。令 $x(i)=x_i$, $y(i)=y_i$, 则每个坐标对可以当作一个复数来处理, 从而有:

$$s(i) = x(i) + jy(i) \quad (3.3)$$

由此可以得到 $s(k)$ 的离散傅里叶变换 (DFT) 为:

$$a(u) = \sum_{k=0}^{N-1} s(k) \exp(-j2\pi uk / N), \text{ 其中 } k=0, 1, \dots, N-1 \quad (3.4)$$

复系数 $a(u)$ 直接与边界曲线的形状相关的,称之为边界的傅立叶描绘子 (Fourier Descriptor, FD)。当系数项取到足够阶次的时候,它可以将物体的形状信息完全提取并恢复过来。变换后得到的 $|a(u)|$ 有很多接近 0 的值,就可以把这些值舍去,只取前 32 点和后 32 点共 64 点,这样,就把二维的特征点变为一维,同时可以减小冗余。得到 48×64 的特征矩阵,然后用“最小欧氏距离”识别,识别率为 25% (8 个人里面正确识别 2 个),用时 5s。三种方法的对比见表 3.4。

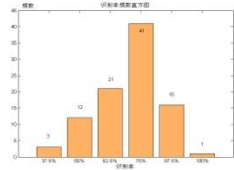
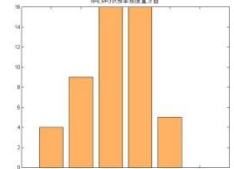
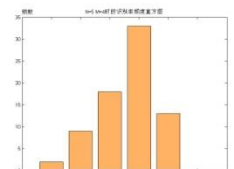
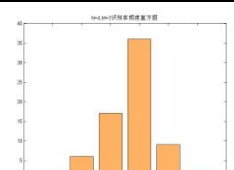
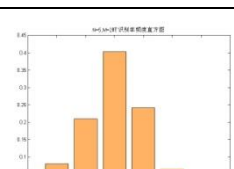
表 3.4 步态识别的三种方法对比

方法	对数据库自身识别率	对测试序列识别率	耗时
CHMM	100%	37.5%~100%	90s 左右
最小欧氏距离	100%	37.5%	14.5s
傅里叶描绘子	100%	25%	5s

3.3 结果优化

基于 CHMM 模型中, u , U 和 C 的初始参数对结果影响较大,同时这些参数在计算时又是随机赋值的,本文重复 100 次识别操作,选取识别结果最好的一组初值,这样就可以使识别率基本固定在 100%。

表 3.5 状态数 N 和混合数 M 的优化实验结果

识别率 初始 条件	37.5%	50%	62.5%	75%	87.5%	100%	识别率 频率直方图
N=5 M=3 (默认)	3.2%	12.8%	22.3%	43.6%	17.0%	1.1%	
N=6 M=3	8%	18%	32%	32%	10%	0	
N=5 M=4	2.7%	12%	24%	44%	17.3%	0	
N=4 M=3	1.4%	8.6%	24.3%	51.4%	12.9%	1.4%	
N=5 M=2	8.1%	21.0%	40.3%	24.2%	6.5%	0	

此外，CHMM 的初值 N（状态数）和 M（混合数）也可能会影响识别率，由此，分别将 N 和 M 的值+1 和-1，来观察 N 和 M 的变化对识别率的影响。同样由于实验结果的随机性，每次实验均重复 50~100 次来查看频率直方图，如上表所示。由表 3.5 可以看出，N=4，M=3 是局部最优，识别率更集中于 75% 附近。

第四章 总结

本文采用 CHMM 作为主要步态识别方法,较为详细地介绍了识别的实验过程。实验中,提取待训练数据库人物轮廓点到质心的距离作为步态特征,用 GMM 对 CHMM 进行参数初始化,用 Baum-Welch 算法实现参数重估,建立每个人的 CHMM 模型,通过前向算法求测试序列的最大似然概率值来识别人体。实验取得了较好的成果,同时和最小欧氏距离和傅里叶描绘子法进行比较,证明了 CHMM 在处理步态序列中强大的时序处理能力。

实验最后,用重复试验的办法获取重要参数如均值矩阵 μ 和协方差矩阵 Σ 的初始值,由每次随机的 37.5%~100%识别率变为基本固定的 100%识别率,同时,用重复试验的办法来优化 CHMM 的其他参数比如状态数 N 和混合数 M ,使实验结果得到了较好的优化。

当然,由于时间有限,实验也存在一些局限和有待进一步优化的地方。实验中,参与识别的人数可以增加来检验数据库大小对识别率的影响,同时用更好的自动修补算法来替代人工修补,还有一点是计算时间较久,达 80~90 秒 (Matlab, 2GHz 主频),尚难以达到实时识别,识别的快速算法也有待进一步研究。

参考文献

- [1] Murray M P, Drought A B, Kory R C. Walking patterns of normal men[J]. The Journal of Bone & Joint Surgery, 1964, 46(2): 335-360.
- [2] 王科俊, 侯本博. 步态识别综述[J]. 中国图象图形学报, 2007, 12(7): 1152-1160.
- [3] 芬兰发明手机电脑防盗装置 通过步态识别主人 <http://news.sina.com.cn/w/2005-10-15/11277175830s.shtml>
- [4] 中科院计算所 CASIA 步态识别数据库 <http://www.cbsr.ia.ac.cn/china/Gait%20Databases%20CH.asp>
- [5] Kale A, Sundaresan A, Rajagopalan A N, et al. Identification of humans using gait[J]. Image Processing, IEEE Transactions on, 2004, 13(9): 1163-1173.
- [6] Lee L, Grimson W E L. Gait analysis for recognition and classification[A]. In: Proceedings of the Fifth IEEE International Conference on Automatic Face Gesture Recognition [C], Washington, DC, USA, 2002: 155~162.
- [7] Cunado D, Nixon M S, Carter J. Using gait as a biometric, via phase2weighted magnitude spectra [A]. In: Proceedings of International Conference on Audio and Video-based Biometric Person Authentication[C], Crans-Montana, Switzerland, 1997: 95~102.
- [8] Urtasun R, Fua P. 3D Ttracking for gait characterization and recognition[A]. In: Proceedings of the Sixth IEEE International Conference on Automatic Faceand Gesture Recognition[C], Seoul, Korea, 2004: 17~22.
- [9] Wang Liang, Tan Tieniu, Ning Huazhong, et al. Silhouetteanalysis2 based gait recognition for human identification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(12): 1505~1518.
- [10] Boulgouris N V, Plataniotis K N, Hatzinakos D. An angular transform of gait sequences for gait assisted recognition[C]//Image Processing, 2004. ICIP' 04. 2004 International Conference on. IEEE, 2004, 2: 857-860.
- [11] Rabiner L R. A tutorial on hidden Markov models and selected applications

- in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2): 257-286.
- [12] Kale A, Rajagopalan A, Cuntoor N, et al. Gait based recognition of humans using continuous hmms [A]. In: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition [C], Washington, DC, USA, 2002:336~341.
- [13] 柴艳妹, 夏天, 韩文英, 等. 步态识别研究进展[J]. 计算机科学, 2012, 39(6): 10-15.
- [14] 宋雪萍, 马辉, 毛国豪, 等. 基于 CHMM 的旋转机械故障诊断技术[J]. 机械工程学报, 2006, 42(5).
- [15] HMM toolbox for Matlab 主页 <http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html>
- [16] 兰叶敏, 李轶, 张聪, 等. 基于图像边界中心距和傅立叶描绘子相结合的多维步态特征识别研究[Z]. 第四届立体图象技术专业委员会学术研讨会, 沈阳, 2010-05-21.
- [17] Jurafsky D, Martin J H, Kehler A. Speech and language processing: an introduction to natural language processing, computational linguistics and speech recognition[M]. New Jersey: Prentice Hall, 1999.
- [18] 混合高斯模型(Mixtures of Gaussians)和 EM 算法, <http://www.cnblogs.com/jerrylead/archive/2011/03/05/1971867.html>
- [19] 冈萨雷斯, 伍兹, 埃丁斯, 等. 数字图像处理: MATLAB 版[M]. 电子工业出版社, 2005.

致谢

本课题在选题及研究过程中得到于慧敏教授的悉心指导。于老师多次询问研究进程，并为我指点迷津，帮助我开拓研究思路，精心点拨、热忱鼓励，在此表示衷心感谢。

毕业论文完成之际，我的大学生活也即将结束。四年真快，正如四年前初到杭州时酷暑难耐到现在已经基本习惯，四年的时间里带给我太多的改变。期间，失落过也振奋过，犹豫过也执着过，大多时候是一点一滴的小事，虽模糊远去，但也许正是某件小事中的人和物让我曾经觉得温暖和感动。

感谢我的爸爸妈妈，儿行千里母担忧，在外读书的日子里，总能从家中得到最坚定的支持和一如既往的鼓励；

感谢我的老师们，他们的一言一行，让我看到了对学术的认真执着，对知识的敬畏和对学生的殷切期望，这些都深深地感染和激励着我；

感谢我的学长学姐们，在我刚入校有任何困难的时候，首先想到的就是他们。不论在学业上，生活上还是发展规划上，他们对我的帮助和影响很大；

感谢我的同学和朋友们，和他们在一起很快乐，有种说法是“在大学里，对自己影响最大的是同学”，我很认同这种观点，从他们的身上能看到自己的长处和不足，时时刻刻鞭策着我努力前进。

最后，感谢我的亲人和生活中遇到的陌生人，是他们让我懂得什么是亲情、人情和世情，能够更加冷静地看待自己所处的社会。

本科生毕业论文（设计）任务书

一、题目：_____基于步态的人体识别_____

二、指导教师对毕业论文（设计）的进度安排及任务要求：

第 1-6 周：查阅至少 5 篇相关多人步态识别的论文和实验室相关的技术文档。

针对人体步态识别算法设计展开调研。

第 7-8 周：在现有文献阅读的和调研的基础上，撰写研究报告，要求针对算法的设计思路，国内外研究现状提交报告。

第 9-11 周：完成开题报告，要求给出毕设的技术方案，明确毕设任务、要求和进度安排。

第 12-22 周：按开题报告确定的进度安排执行。

第 23-24 周：撰写毕业设计，答辩。

起讫日期 2012 年 12 月 日至 2013 年 6 月 日

指导教师（签名）_____职称_____

三、系或研究所审核意见：

同意指导教师的安排与要求。

负责人（签名）_____

年 月 日

毕 业 论 文（设计） 考 核

一、指导教师对毕业论文（设计）的评语：

冯继雄查阅了相关文献，认真研究了现有的相关技术。论文采用隐马尔可夫模型（Hidden Markov Model, HMM）方法实现了基于步态的人体识别，设计合理，并通过实验证明了该方法的有效性。
论文书写规范，论述清楚，反映了作者有较好的专业基础知识和一定的开发能力。

指导教师(签名)_____

年 月 日

二、答辩小组对毕业论文（设计）的答辩评语及总评成绩：

论文研究了基于步态的人体识别算法设计，论文选题具有很好的工程应用价值。作者通过对相关文献的阅读，论文采用隐马尔可夫模型（Hidden Markov Model, HMM）作为主要方法实现，实现了基于步态的人体识别。实验结果证明了该方法具有较好的识别效果。
论文书写规范，论述基本清楚，反映了作者有一定的专业基础知识和科研能力，论文达到了本科学士学位论文水平。答辩过程中，讲述较清楚，回答问题基本正确。答辩小组一致同意通过学士学位论文答辩。

成 绩 比例	文献综述 占（10%）	开题报告 占(20%)	外文翻译 占（10%）	毕业论文（设计）质 量及答辩 占（60%）	总 评 成绩
分 值					

答辩小组负责人（签名）_____

年 月 日