

文章编号 1000 5013(2004) 03 0244 03

视频流分析的隐藏马尔可夫链算法

黄 建 新

(华侨大学数学系, 福建 泉州 362011)

摘要 视频流分析是多媒体技术研究的重要内容, 具有广泛的应用领域. 应用马尔可夫链模型, 研究视频中的人体运动目标位置的跟踪. 通过对跟踪的第一帧视频中的人体区域进行手工标注, 建立马尔可夫链模型. 然后结合卡尔曼估计算法, 跟踪出下一帧的人体位置. 实验证明, 这种算法可以有效地解决在复杂的运动背景下, 人体运动目标的跟踪问题.

关键词 视频分析, 马尔可夫链模型, 卡尔曼估计, 人体运动跟踪

中图分类号 TP 391.41: O 211.62 文献标识码 A

视频流分析是目前多媒体技术研究的热门话题, 其成果可以广泛应用在数字图书馆、视频动画、人机交互和实时监控等领域. 运动目标的跟踪是视频流分析的主要对象之一, 至今人们提出了许多关于人体运动目标的跟踪算法. 下列其代表性: (1) Yamamoto 提出的光流法^[1]; (2) Yamada 利用统计模型提取人体的轮廓^[2]; (3) Wren 使用 blob 跟踪人体的运动位置^[3]; (4) Blake 使用基于变形活动模板 (snake) 的跟踪算法^[4]. 这些算法主要应用在静止背景下人体运动目标的跟踪, 而本文提出了一种可以在复杂背景下跟踪人体运动目标的算法. 隐藏马尔可夫链模型 (Hidden Markov Model, 简称 HMM) 最早在 19 世纪 60 年代末和 70 年代初就开展了许多研究, 并在语音识别中获得了应用^[5]. 而真正的广泛应用是在最近 20 多年期间. 本文试图使用 HMM 跟卡尔曼估计算法结合, 跟踪视频流中人体运动的目标. 它的基本思想是在被跟踪视频流的第一帧图像中, 对人体所在的区域进行手工标注. 利用该区域的采样集, 建立 HMM. 然后, 使用卡尔曼算法估计下一帧图像中人体所在区域的中心点位置, 对预测中心点的附近范围内进行搜索和匹配, 求出最佳匹配位置. 最后, 使用匹配区域采样的数据重新训练 HMM. 如此重复, 就可以跟踪视频流中人体运动的轨迹.

1 隐藏马尔可夫链模型

马尔可夫链是一个独立增量的随机过程, 而 HMM 是一种状态未知的马尔可夫链. 定义为 $\lambda = (A, B, \pi)$, 其中 $A = \{a_{ij}\}$ 为概率转移矩阵, $a_{ij} = P[q_{t+1} = S_j | q_t = S_i]$, 即在 t 时刻状态为 S_i 的条件下 $t+1$ 时刻状态为 S_j 的概率. $B = \{b_j(k)\}$ 是观察概率矩阵, $b_j(k) = P[v_k | q_t = S_j]$ 为在状态 S_j 下 t 时刻的观察概率分布. $\pi = [\pi_i]$ 是初始状态概率矩阵, $\pi_i = p[q_1 = S_i]$ 为状态 S_i 初始概率. HMM 存在许多形式, 我们在人体运动目标跟踪中使用多态观察密度分布的从左到右模型, 状态数为 4. 状态模型使用多态高斯分布函数, 如图 1 所示. 其模型参数定义为

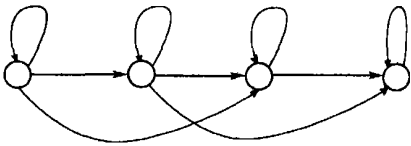


图 1 HMM 左右模型

$$b_j(0) = \sum_{m=1}^M C_{jm} \pi(0, \mu_m, \sigma_m), \quad 1 \leq j \leq N. \tag{1}$$

在式 (1) 中, M 为多态分布数目, N 为状态数, $\pi(0, \mu_m, \sigma_m)$ 是多态高斯分布函数. 其均值为 μ_m , 而协方差则为 σ_m .

2 HMM 的学习算法

HMM 的学习使用前向-背向算法^[6,7]. 这种算法, 具有计算量少和收敛速度快的特点.

2.1 前向概率方法

假设部分观察序列为 $O_1 O_2 \dots O_t$, 则定义 $\alpha_t(i) = P(O_1 O_2 \dots O_t, q_t = S_i | \lambda)$ 为在 t 时刻状态为 S_i 的观察序列在模型 λ 下的概率. 使用递归方法计算, 获得

$$P(O | \lambda) = \sum_{i=1}^N \alpha_t(i). \tag{2}$$

2.2 后向概率方法

类似地, 后向概率方法采用后向传播方法. 即定义 $\beta_t(i) = P(O_{t+1} O_{t+2} \dots O_T | q_t = S_i, \lambda)$ 为在 t 时刻状态为 S_i 时, 部分观察概率 $O_{t+1} O_{t+2} \dots O_T$ 在模型 λ 下的概率. 有

$$\beta_t(i) = \sum_{j=1}^N \alpha_j b_j(O_{t+1}) \beta_{t+1}(j). \tag{3}$$

2.3 模型参数更新

利用前向后向概率计算方法, 我们可进行模型参数更新. 为此, 首先定义 $\gamma_t(i) = P(q_t = S_i | O, \lambda)$ 为观察序列在 t 时刻状态为 S_i 的概率, 及 $\zeta(i, j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda)$ 为观察序列在 t 和 $t+1$ 时刻状态分别为 S_i, S_j 的概率. 那么, 模型参数更新公式为

$$\bar{C}_{jk} = \frac{\sum_{t=1}^T \gamma_t(j, k)}{\sum_{t=1}^T \sum_{k=1}^M \gamma_t(j, k)}, \tag{4}$$

$$\bar{\mu}_{jk} = \frac{\sum_{t=1}^T \gamma_t(j, k) \cdot O_t}{\sum_{t=1}^T \gamma_t(j, k)}, \tag{5}$$

$$\bar{\sigma}_{jk} = \frac{\sum_{t=1}^T \gamma_t(j, k) \cdot (O_t - \mu_{jk}) \cdot (O_t - \mu_{jk})'}{\sum_{t=1}^T \gamma_t(j, k)}, \tag{6}$$

$$\bar{\pi}_i = \gamma_1(i), \tag{7}$$

$$\alpha_{ij} = \frac{\sum_{t=1}^{T-1} \zeta_t(i, j)}{\sum_{t=1}^T \gamma_t(i, j)}, \quad b_j(k) = \frac{\sum_{t=1, O_t=v_k}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}. \tag{8}$$

3 匹配准则和跟踪过程

3.1 匹配准则

假设观察区域的采样序列为

$$O = \{O_1, O_2, \dots, O_T\},$$

我们计算该观察序列在最好的状态序列 $Q = \{q_1, q_2, \dots, q_T\}$ 下的概率值. 即

$$p = P(q_1, q_2, \dots, q_T, O_1, O_2, O_T | \lambda),$$

其计算方法采用 Viterbi 算法^[8]. 在搜索过程中, 概率值最大者为匹配区域.

3.2 跟踪过程

(1) 步骤 1. 对要跟踪的视频中第 1 帧人体所在区域进行手工标注, 用矩形框表示, 如图 2(a) 所示. 由于 HMM 输入是一个时间序列的向量, 因此我们将标注矩形框根据设定的隐藏状态数按列进行等分采样. 使用 K 均值聚类方法^[9], 求出各个状态的参数. 然后采用 Viterbi 算法, 对 HMM 模型进行训练. (2) 步骤 2. 使用卡尔曼估计算法预测下一帧的运动位置, 具体算法见^[10]. 这里, 跟踪对象是区域的中心位置. (3) 步骤 3. 对估计的附近区域进行匹配搜索. 每一步搜索都要根据 HMM 模型计算区域的概率值, 详见节 3.1. 最大概率者为匹配区域. (4) 步骤 4. 对匹配区域重新采样. 根据节 2.3 更新 HMM 模型. (5) 步

骤 5. 返回步骤 2, 继续跟踪下一帧图像中的人体运动位置.

4 实验结果

我们使用了本文所述的方法, 对一段视频中的人体运动位置进行了跟踪. 在视频中背景是一直在变化的, 所有的算法都在 Matlab 环境下实现, 跟踪结果如图 2 所示. 图(a) 是手工标注人体区域的第一帧图

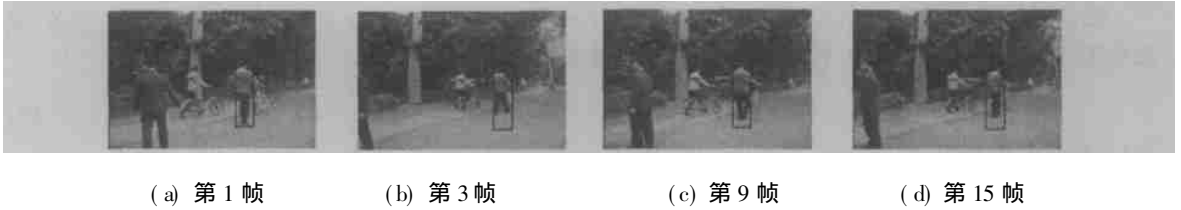


图 2 人体运动目标跟踪结果

像, 用矩形框表示. 图(b) , (c) , (d) 分别是第 3, 9, 15 帧图像的跟踪结果.

5 结束语

通过实验证明, 使用 HMM 模型结合卡尔曼估计算法跟踪, 具有运动背景的视频流中人体运动位置是有效的. 我们下一步工作将解决在遮挡情况下的人体运动, 其跟踪问题、跟踪结果进行姿势估计以及行为识别.

参 考 文 献

- 1 Yamamoto M, Koshikawa K. Human motion analysis based on a robot arm model[J]. CVPR, 1991, 6: 664~ 665
- 2 Yamada M, Ebihara K, Ohya J. A new robust real time method for extracting human silhouettes from color image[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(3): 226~ 239
- 3 Wren C R, Azarbayejani A, Darrell T, et al. Real time tracking of the human body[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence. 1997, 19(7): 780~ 785
- 4 Blake A, Isard M. Active contours[M]. Birlin: Springer, 1998. 3~ 50
- 5 Baker J K. The dragon system an averview[J]. IEEE Trans. Acoust. Speech Signal Processing, 1975, 23(1): 24~ 29
- 6 Baum L E, Egon J A. An inequality with applications to staticistical stimation for probabilistic functions of a markov process and to a model for ecology[J]. Amer. Meteorol. Soc. , 1967, 73(3): 360~ 363
- 7 Baum L E, Sell G R. Growth functions for transformations on manifolds[J]. Pac. J. Math, 1968, 27(2): 211~ 227
- 8 Viterbi A J. Error bounds for convolutional codes and an asymptotically optimal decoding algorithm[J]. IEEE Trans. Informat. Theory, 1967, 13(4): 260~ 269
- 9 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2002. 235~ 237
- 10 黄建新. Kalman 滤波的人体运动位置跟踪算法[J]. 华侨大学学报(自然科学版) , 2003, 24(3): 254~ 256

Applying Hidden Markov Chain Algorithm to Video Stream Analysis

Huang Jianxin

(Dept. of Math., Huaqiao Univ., 362011, Quanzhou, China)

Abstract Tracking a moving target is the main topic in video stream analysis which is a hotspot in the study of multimedia technology. In this paper, the hidden Markov chain algorithm is applied to tracking moving target position of human body in motion video. A Markov chain model is built at first by marking manually the area of human body to be tracked at the first frame of video. And then, the position of human body at the next frame of video is tracked down by combining with kalman algorithm of estimate. As proved by experiment, this algorithm is able to settle effectively the problem of tracking moving target of human body against a complex moving background.

Keywords video analysis, Markov chain model, kalman estimate, tracking moving human body