

基于 MDL Shape Model 及 EFD 的行人 轮廓 2D + time 表示

王绍宇¹⁾ 戚飞虎²⁾ 夏小玲¹⁾

¹⁾(东华大学计算机学院, 上海 201620) ²⁾(上海交通大学计算机科学与工程系, 上海 200240)

摘要 针对智能监控的行人轮廓特征提取, 先利用 MDL shape model 得到紧致但能充分表征轮廓几何特征标志点之间的对应关系, 再利用椭圆傅里叶分解把行人轮廓表示为一系列不同频率下椭圆傅里叶系数组成的向量, 最后得到行人 2D + time 轮廓的描述向量。实验结果表明, 此方法不但可以在几何上直观地表示行人轮廓, 而且大大降低了轮廓向量的维数。

关键词 最小描述长度模型 椭圆傅里叶分解 主分量分析 procrustes 对齐

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2008)10-1898-04

2D + time Representation of Pedestrian Contour Based MDL Shape Model and EFD

WANG Shao-yu¹⁾, QI Fei-hu²⁾, XIA Xiao-ling¹⁾

¹⁾ (School of Computer Science and Technology, Donghua University, Shanghai 201620)

²⁾ (Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200240)

Abstract To extract the contour features of pedestrian under surveillance, this paper applies Minimum Description Length Shape to get compact point correspondences, which can capture important shape characteristics, of 2D + time contour, and then uses Elliptic Fourier Decomposition to a series of EFT coefficients at different frequency level to compose a contour vector. Experiments indicate our method is efficient in contour representation and vector dimension reduction.

Keywords MDL shape model, elliptic Fourier decomposition, PCA, procrustes alignment

1 引言

面向行人的智能监控系统中, 需要对人的轮廓进行提取, 从而进一步分析场景中行人的姿态。轮廓的表示方法对于后续的步态识别和异常行为检测至关重要。

国内外的不少学者, 对轮廓的表示方法做了深入的研究。Zahn 首先使用傅里叶描述子来识别物

体的形状特征^[1]。但是传统的傅里叶分解方法在形状的描述上以预估半径和周期的方式构造相似形状来逼近目标轮廓。这种方法采用的基函数是正弦和余弦函数, 在形状边界的快速傅里叶变换和去掉边界起始点位置在傅里叶描述子中的相位影响等方面仍存在一定的问題。

基于以上原因, 本文提出了一种基于最小描述长度形状模型(MDL shape model)和椭圆傅里叶分解(EFD)的行人时空轮廓表示方法。该方

基金项目: 上海市科技攻关计划(06511011)

收稿日期: 2008-07-11; 改回日期: 2008-07-30

第一作者简介: 王绍宇(1973 ~), 男, 讲师。2006年获上海交通大学计算机应用技术专业博士学位。主要研究方向为计算机视觉及智能视频处理。E-mail: wang@sjtu.org

法利用 MDL shape model 获取行人相同标志点在不同轮廓间的点对应关系,然后利用傅里叶的正交分解和周期性的特点,使用椭圆函数作为基函数来做周期性的运算。从而把行人的 2D + time 轮廓用一系列的椭圆族系数组成的向量来进行表示。

通过以上方法,不但能够以比较直观的方式来表示某个特定动作的 2D + time 轮廓,而且使得轮廓向量的维数大大降低。

2 行人轮廓标志点对应问题求解

2.1 行人轮廓的检测与提取

先使用背景减法提取场景中的行人目标。为了克服场景中如光照和风等因素的干扰,使用色度、亮度空间的多个分布模型来建立背景模型。在对背景模型更新时,将均值、方差的更新速率和多个模型的更替速率分开。对像素值属于多个分布模型的情况,用最小相似距离确定要更新的模型^[2]。在得到二值化行人目标图像后,再使用形态学算子过滤噪声,并采用 8 连通区域的轮廓跟踪法来提取出行人的原始轮廓。

2.2 MDL Shape Model 模型

现有的行人轮廓标志点大都通过手工标注或者沿着轮廓按等距离法确定,不但费时费力、主观性强,而且精度也不高。为了有效地自动对行人轮廓特征进行提取,不同姿态下的轮廓不但要有相同的标志点数目,而且相同序号的标志点之间也要保持正确的点对应关系。

假设某个行人某一姿态的视频包括 t 帧图像,相对应的轮廓集合 S 包含 t 个轮廓样本,每个轮廓 $\{S_i: i=1,2,\dots,t\}$ 由一个维数为 $2n_p$ 的形状向量 $\{X_i: i=1,2,\dots,t\}$ 表示,其中 n_p 为每个轮廓标志点的个数。通过主分量分析(PCA)把训练样本从原始的笛卡尔坐标空间转换到特征空间,得到 $t-1$ 个相互正交的特征向量 $\{P_m: m=1,2,\dots,(t-1)\}$,由这些特征向量张成一个 $t-1$ 维的形状空间, S_i 对应的形状向量 X_i 在这个特征空间可以被线性表示为

$$x_i = \bar{x} + \sum_{m=1}^{t-1} p_m b_{m,i} \quad (1)$$

式中, $b_{m,i}$ 为第 i 个形状向量在第 m 个主方向上的投影值。由每个特征向量对应的特征值 λ_m (特征值

按降序排列)可得到在第 m 个特征向量方向上的数据方差 $\sigma_m = \sqrt{\lambda_m}$ 。在此特征空间中,第 i 个轮廓 S_i 在第 m 个方向 P_m 上投影得到的坐标为 $y_{m,i}$ 。平均形状向量 \bar{X} 和特征向量 $\{P_m: m=1,2,\dots,(t-1)\}$ 的描述长度对于给定的轮廓集合是一个常量,因此只需要计算 $\{y_{m,i}: m=1,2,\dots,(t-1); i=1,2,\dots,t\}$ 的描述长度^[3]。通过 PCA 得到的特征向量之间是相互正交的,所以可以分别计算形状空间在 $t-1$ 个主方向上的描述长度,最后得到 MDL 目标函数^[4]:

$$DL = \sum_{\lambda \geq \lambda_{cut}} \left(1 + \log \frac{\lambda_m}{\lambda_{cut}}\right) + \sum_{\lambda < \lambda_{cut}} \frac{\lambda_m}{\lambda_{cut}} \quad (2)$$

式中, λ_{cut} 为 MDL 模型精度阈值。在对标志点参数进行初始化的时候,可以先均匀地沿着轮廓分配标志点。在后续的 PCA 统计分析以及 MDL 目标函数计算及优化处理之前,需要将每个轮廓都对齐到统一的度量框架中,以去掉轮廓的缩放、旋转和位移因素。采用 procrustes alignment 方法^[5],通过最小化每个轮廓与平均形状向量之间的误差平方和 $D = \sum_{i=1}^t |x_i - \bar{x}|^2$ 来对齐轮廓。

2.3 点对应关系求解

为了提高算法的稳定性,考虑到行人轮廓关键标志点(如头部顶点和手足末端点)的曲率在变形过程中保持一定规律性的特点,在 MDL 目标函数中增加 CurvatureCost 项:

$$CurvatureCost = C \frac{1}{n_p} \frac{1}{t} \sum_{i,r} (k_{ir} - k_i^{mean})^2 \quad (3)$$

式中, n_p 为每个轮廓上标志点的个数, t 为轮廓的数目, k_{ir} 为标志点在第 r 个形状第 i 个标志点的曲率, k_i^{mean} 为所有样本第 i 个标志点的平均曲率, C 为曲率代价的权重。

算法过程分为 3 步:第 1 步为初始化,每个形状由均匀分布在轮廓上的固定数目的标志点来定义;第 2 步是对所有训练轮廓进行 Procrustes 对齐;第 3 步是标志点位置的更新。通常 MDL 算法没有必要在所有的标志点位置进行优化,只需要在给定层次上的控制节点上进行,每个控制节点的位置由位置参数和随机变化步长确定,在每个训练样本上的控制节点的位置在优化过程中变化后,计算对应的 MDL 目标函数,根据目标函数的值来控制节点位置朝着优化的方向移动。

3 基于 EFD 的 2D + time 轮廓表示

轮廓点对应关系确定后,任一轮廓上第 p 个点的坐标可以表示为复数 $(x_p + jy_p)$ 。如果让 p 点沿着轮廓一直移动,该点会回到起点的位置。 $x(t_p)$ 和 $y(t_p)$ 就形成了一个周期为 T 的周期函数。对于周期函数可以利用 Fourier 函数表示为

$$x(t_p) = \sum_{n=1} \left(a_n \cos \frac{2n\pi t_p}{T} + b_n \sin \frac{2n\pi t_p}{T} \right)$$

$$y(t_p) = \sum_{n=1} \left(c_n \cos \frac{2n\pi t_p}{T} + d_n \sin \frac{2n\pi t_p}{T} \right)$$

可推导出下面的椭圆方程:

$$ax_n^2(t_p) + bx_n(t_p)y_n(t_p) + cy_n^2(t_p) = d \quad (4)$$

式中, $a = c_n^2 + d_n^2$, $b = -2(a_n c_n + b_n d_n)$, $c = a_n^2 + b_n^2$, $d = |a_n d_n - b_n c_n|$, 由此可知,任何一个形状的轮廓都可以通过一族椭圆来逼近。图 1 为椭圆傅里叶分解的示意。

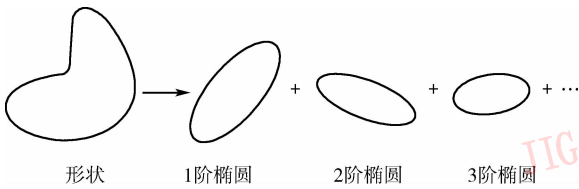


图 1 轮廓的椭圆傅里叶分解效果

Fig. 1 Demonstration of elliptic Fourier decomposition

对于 N 阶椭圆傅里叶分解,除去 3 个常椭圆傅里叶系数,剩下的椭圆傅里叶系数可组成一长度为 $4N - 3$ 的 1 维向量 $\mathbf{EFC}(\mathbf{S}, N)$:

$$[d_1, a_2, b_2, c_2, d_2, \dots, a_N, b_N, c_N, d_N]^T$$

其中, \mathbf{S} 为在笛卡尔坐标空间中用 x 和 y 的坐标值来表示的轮廓向量, N 为椭圆傅里叶分解的阶数^[6]。

由于通过 MDL shape model 唯一确定了 2D + time 各个轮廓上各标志点的对应关系,行人的 2D + time 轮廓向量 \mathbf{V} 可以表示为

$$\mathbf{V}(\mathbf{S}, N, t) = \begin{bmatrix} d_1^1 & a_2^1 & b_2^1 & \dots & a_N^1 & b_N^1 & c_N^1 & d_N^1 \\ d_1^2 & a_2^2 & b_2^2 & \dots & a_N^2 & b_N^2 & c_N^2 & d_N^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ d_1^t & a_2^t & b_2^t & \dots & a_N^t & b_N^t & c_N^t & d_N^t \end{bmatrix}^T \quad (5)$$

总的算法流程如图 2 所示。

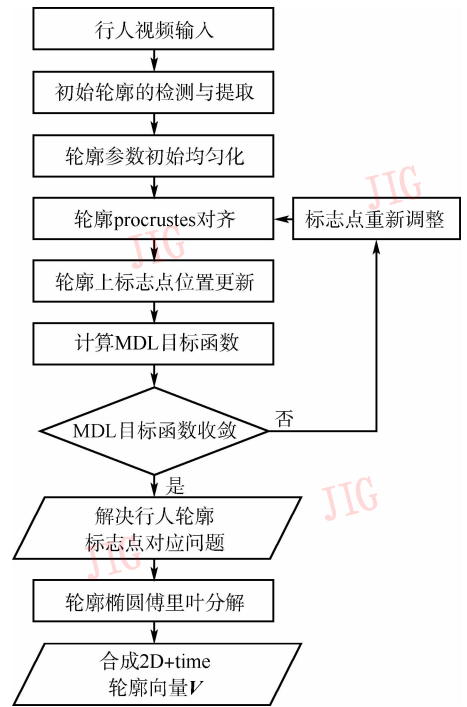


图 2 2D + time 轮廓表示算法示意图

Fig. 2 Illustration of 2D + time contour descriptor algorithm

4 实验结果

采集 75 帧的行人步行视频,通过背景差等算法获取原始轮廓后,使用 8 个控制节点来进行轮廓参数化,每个训练轮廓上最后选取的标志点个数 n_p 为 64,轮廓数目 t 为 75。对于 2D + time 轮廓,在 MDL 优化过程中, λ_{cut} 的值设为 0.003,循环的最大次数为 40。图 3 为经过 80 000 次 MDL 目标函数值的更新后,使用 MDL shape model 得到的主要 8 个控制节点的点对应关系:

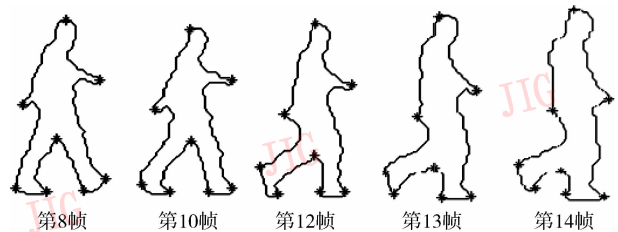


图 3 2D + time 轮廓的点对应结果

Fig. 3 Point correspondences of a 2D + time contour

确定轮廓上所有 64 个点在时间方向上的点对应关系后,由式(5)可分别求出轮廓在不同阶数的椭圆傅里叶分解后的 2D + time 轮廓向量。由椭圆傅里叶重建结果(图 4)可以看出,随着椭圆分解阶

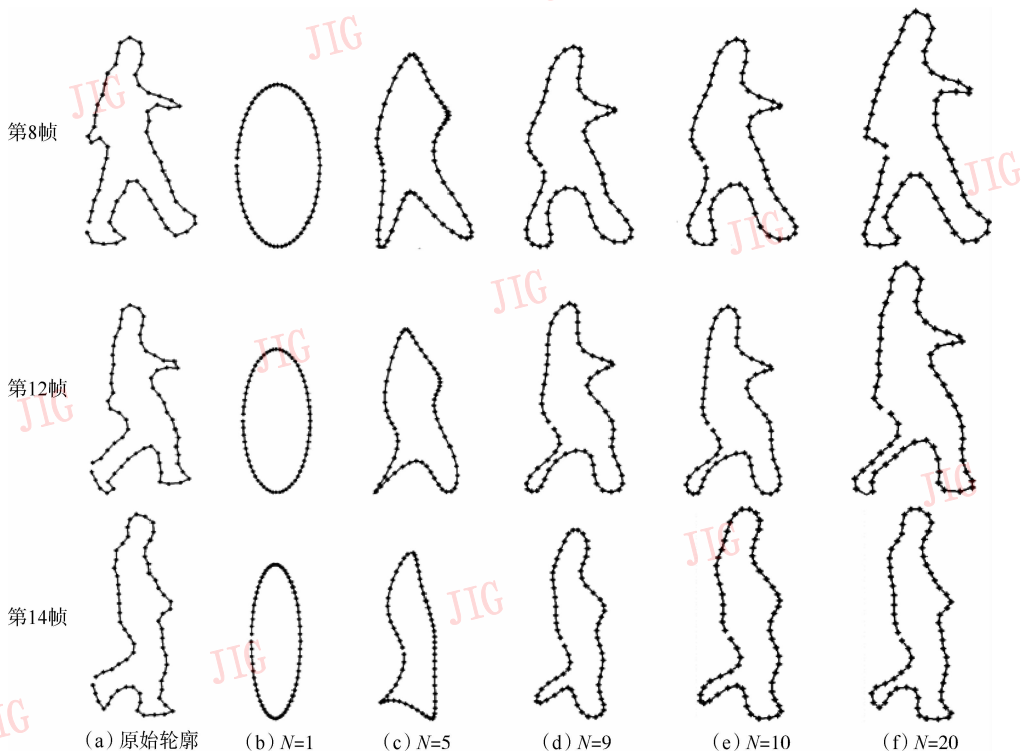


图 4 2D + time 轮廓重建

Fig. 4 2D + time contour reconstruction by increasing harmonics' number from 1 to 20

数 N 的增加,重建的精度也在增加,当 N 为 10 时,就可以比较精确地重建原始的轮廓。

如果用传统的 x 和 y 坐标值来表示行人时空轮廓,对应的向量的维数为 $9\ 600(2n_p \times t)$,通过椭圆傅里叶分解,高阶椭圆代表形状的高频信息,即轮廓的局部特征信息。低阶椭圆代表形状的低频信息,也就是重要的全局信息,这部分信息在重建过程中得到了保留。2D + time 轮廓向量的维数从传统方法的 9 600 降到 2 775。

5 结 论

基于 MDL shape model 解决了轮廓标志点的对应关系,在此基础上利用 EFD 主要考虑低频的全局轮廓信息,忽略高频的轮廓局部细节信息。使得行人时空轮廓向量的维数从 $2n_p \times t$ 降到了 $(4N - 3) \times t$ 维,其中 n_p 远大于 $(4N - 3)$ 。对于复杂轮廓,因为需要更多的标志点,压缩效果更好。相对于传统的方法,2D + time 轮廓表示方法随着 N 的增加,对轮廓的描述从全局到局部分别用不同频率的椭圆傅里叶系数表示,具有直观的物理意义,同时也提高了整体时空轮廓特征的分析能力,可用于面向行人的视

频监控和步态识别。

下一步的工作包括获取 2D + time 轮廓在单个步态周期内的表示及动态时间归整(DWT)等。

参考文献 (References)

- 1 Zahn C T, Roskies R Z, et al. Fourier descriptors for plane closed curves [J]. IEEE Transactions on Computers, 1972, 21 (3): 269 ~ 281.
- 2 Cen Feng, Qi Fei-hu, Chen Mao-lin. Multi-distribution model for Background Subtraction in long term video surveillance system [J]. Journal of Infrared and Millimeter Waves, 2002, 21 (1): 59 ~ 63. [岑峰, 戚飞虎, 陈茂林. 长期视频监控系统的多分布模型背景差方法 [J]. 红外与毫米波学报, 2005, 21 (1): 59 ~ 63.]
- 3 Davies R H, Twining C J, Cootes G H. A minimum description length approach to statistical shape modeling [J]. IEEE Transactions on Medical Imaging, 2002, 21 (5): 525 ~ 537.
- 4 Thodberg H H. Adding curvature to minimum description length shape models [A]. In: Proceedings of 14th British Machine Vision Conference [C], Norwich, UK, 2003, 2: 251 ~ 260.
- 5 Ericsson A, Karlsson J. Aligning shapes by minimising the description length [A]. In: Proceedings of Scandinavian Conference on Image Analysis [C], Joensuu, Finland, 2005: 709 ~ 718.
- 6 Wang Shao-yu, Qi Fei-hu, Li Hua-qing. Minimum description length shape model based on elliptic fourier descriptors [A]. In: Proceedings of 3rd International Symposium on Neural Networks [C], Chengdu, China, 2006, 2: 646 ~ 651.