

DOI: 10.13382/j.jemi.2017.01.005

适用于老年公寓的动作识别设计方法*

潘月浩¹ 宋执环¹ 杜往泽¹ 吴乐刚²

(1 浙江大学 控制科学与工程学院 杭州 310027; 2 杭州三云科技有限公司 杭州 310013)

摘要:为帮助老年公寓监护人员及时发现老年人摔倒等动作,提出了一种基于视频监控的动作识别方法。对监控视频,首先通过基于 HSV 空间的混合高斯背景建模方法提取前景图像,然后利用所提出的运动特征和形态特征相结合的方式特征提取,最后通过具有高斯输出的 HMM 模型实现动作类型的识别。提出的方法能够适应光照变化影响,对不同动作的动作方向和动作幅度变化具有很好的鲁棒性,实验中动作的识别准确率达到 90%。结果表明,本方法能够满足老年公寓动作识别的基本要求,具有一定的实用价值。

关键词:视频监控;混合高斯背景建模;运动特征;形态特征;HMM

中图分类号: TN801;TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Design method for action recognition applied to senile apartment

Pan Yuehao¹ Song Zhihuan¹ Du Wangze¹ Wu Legang²

(1. College of Industrial Control, Zhejiang University, Hangzhou 310027, China;

2. San Cloud Technology Co. Ltd., Hangzhou 310013, China)

Abstract:To help nursing staff in senile apartment find the elderly fall and other actions timely, an action recognition method based on video surveillance is proposed. Firstly, the foreground images are extracted by the GMM background modeling method in HS color space. Feature extraction is performed by combining the motion features and morphological features. And action recognition can be achieved by HMM with Gaussian output. The method proposed in this paper can adapt to the changes of illumination. The method also has good robustness to the change of motion direction and motion range, and the recognition accuracy rate reaches 90%. The result shows that the method can meet the basic requirements of action recognition and the method has certain practical value.

Keywords: video surveillance; GMM; motion features; morphological features; HMM

1 引言

由于老年人身体条件的特殊性,日常生活中摔倒等现象极易给老年人的安全造成威胁,因此及时发现老人摔倒并给予救助是老年人看护很重要的工作^[1-2]。对于生活在老年公寓中的老人或独居老人,单纯依靠护理人员看护,不仅需要付出极大的人力资源,且看护过程容易产生疏漏。因此,寻求一种智能监控方案,自动识别生活中人体摔倒等动作,是现实社会的迫切需求。

当前国内外有很多关于室内智能监控或者动作识别

的研究^[3-5],此外,也有很多研究专注于摔倒检测^[6-8]。Foroughi 等人^[9]提出了一种基于历史运动图像的方法,通过将连续视频帧的人体前景部分以不同的权重叠加到一起形成历史运动图像,利用特征空间技术实现降维并与标定的运动图像进行相似度判断来判别未知动作类型,由于动作的变化差异以及特征空间维度选择造成信息的损失,该方法鲁棒性及效果均有一定局限性。Yu 等人^[7]提出了一种基于前景直方图和 DAGSVM 的识别方法,该方法只采用单帧图像信息作为分类特征,忽略运动的连续型特征,对具有相似特征帧的不同动作不能很好的区分,且不能充分利用动作过程中的运动信息。

Nait-Charif^[10]提出的方法采用离散型隐马尔科夫作为识别方法,需要完成对每帧图像与观测状态的近似转换。

本文在前人工作的基础上,提出了一种针对室内环境的动作识别方法。利用基于 HSV 空间的混合高斯背景建模方法,结合光照情况下的阴影消除方法,实现人体的检测;采用运动特征和形态特征相结合的特征提取方式获取特征向量,该特征向量能够适应不同的运动方向情况;通过具有高斯输出的隐马尔科夫模型对不同动作建模,实现对未知类别的动作识别。

2 本文算法

本文算法的主要包括人体检测、特征提取和动作识别 3 个主要部分,算法的流程如图 1 所示。

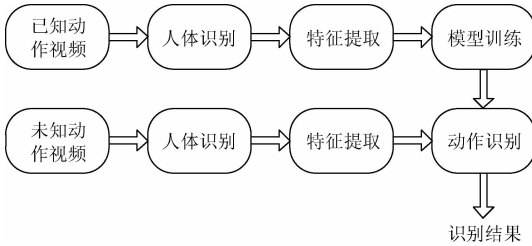


图 1 算法框架

Fig. 1 Framework of the algorithm

2.1 人体检测

由于室内环境的复杂性变化性以及存在光照的影响,人体检测算法必须能够适应背景的变化,能够消除因光照而产生的人体阴影影响。混合高斯背景模型通过多高斯模态来模拟背景,通过参数的自适应更新可以应对背景参数的变化^[11]。在 RGB 空间中,3 个颜色分量相互耦合,光照的变化会对 3 个分量均产生较大的影响,文献[12]中证明在 HSV 空间中,当光照变化时, H 分量不随光照变化而变化, S 分量有较小的变化,而 V 分量有相应的变化。因此,本文借助于 H 、 S 分量,建立混合高斯背景模型实现初步人体检测,借助于 V 分量实现进一步的阴影检测和消除。

首先,定义高斯模式的概率密度函数如下:

$$\eta(\mathbf{X}_t, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \left(\frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}_i|^{1/2}} \right) \cdot e^{-\frac{1}{2}(\mathbf{X}_t - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_i)} \quad (1)$$

式中: $\mathbf{X}_t = [H \ S]^T$, $\boldsymbol{\mu}_i$ 是均值, $\boldsymbol{\Sigma}_i$ 是协方差矩阵,且 $\boldsymbol{\Sigma}_i = \sigma^2 \mathbf{I}$ 。则有高斯背景模型:

$$P(\mathbf{X}_t) = \sum_{i=1}^K \omega_{i,t} \cdot \eta(\mathbf{X}_t, \boldsymbol{\mu}_{i,t}, \boldsymbol{\Sigma}_{i,t}) \quad (2)$$

模型的参数更新主要包括以下过程。

1) 匹配测试:待测试像素 \mathbf{X}_t 分别与 K 个高斯模式进

行匹配,若满足式(3),则测试像素与当前模式匹配。

$$\|\mathbf{X}_t - \boldsymbol{\mu}_{i-1}\| \leq 2.5\sigma_{i-1} \quad (3)$$

2) 更新均值和标准差:未匹配模式的均值和标准差不变,匹配模式的均值和标准差按式(4)、(5)更新。

$$\boldsymbol{\mu}_i = (1 - \alpha)\boldsymbol{\mu}_{i-1} + \alpha\mathbf{X}_t \quad (4)$$

$$\sigma_i^2 = (1 - \alpha)\sigma_{i-1}^2 + \alpha(\mathbf{X}_t - \boldsymbol{\mu}_i)^T (\mathbf{X}_t - \boldsymbol{\mu}_i) \quad (5)$$

3) 更新模式权重如下:

$$\omega_{k,t} = (1 - \alpha)\omega_{k,t-1} + \alpha \cdot M_{k,t} \quad (6)$$

其中模式匹配时 $M_{k,t} = 1$, 否则 $M_{k,t} = 0$ 。

在混合高斯模型中,背景像素所属的模式具有大权重、小方差的特性。因此,利用权重与方差之比 ω/σ^2 对各高斯模态进行降序排序,若前 N 个模式的权重达到或超过背景像素所占阈值 T ,则认为前 N 个模式属于背景模型,若待测试像素根据式(3)不匹配背景模型,则认为是前景像素。

在光照的作用下,检测到的前景图像往往包含因光照产生阴影;在老人摔倒过程中,由于光线的变化对前景的提取会产生较大的影响。因此,对已经提取出来的前景像素,通过阴影检测程序进一步判断像素是否为阴影。阴影像素与实际前景像素不同之处在于阴影像素与对应位置背景像素值有一定的对应关系:阴影像素的色调 H 分量和饱和度 S 分量与实际背景像素值相差不大,而明度 V 会降低^[13]。同时考虑到阴影区域的连贯性,利用下述方法对初步标识为前景的像素进行阴影像素判断。

$$S(x, y) = \begin{cases} 1, & \alpha \leq \frac{P^V(x, y)}{B_k^V(x, y)} \leq \beta \quad \text{and} \\ & |P^S(x, y) - B_k^S(x, y)| \leq \tau_s \quad \text{and} \\ & |(P^H(x, y) - B_k^H(x, y))| \leq \tau_u \\ 0, & \text{其他} \end{cases} \quad (7)$$

式中: $P(x, y)$ 为待测试像素, $B_k(x, y)$ 为背景模型中权重较大的背景像素值。

后期利用必要的形态学处理去掉离散点和小的轮廓可以得到人体的二值图像。

2.2 特征提取

特征提取的目的是从视频图像中提取出能描述人体运动状态的信息。本文采用的特征主要由两部分组成:1) 基于不同视频每帧间运动信息的运动特征,2) 基于单帧图像人体前景主要形态变化的形态特征。

2.2.1 运动特征

老人在发生摔倒、坐下等动作时,身体的运动速度、角度、高度以及身体的形态都会发生显著的变化。本文首先以椭圆拟合人体来提取运动信息, t 帧与 $t-1$ 帧图像之间的基于椭圆描述的一些特征信息量,如图 2 所示。

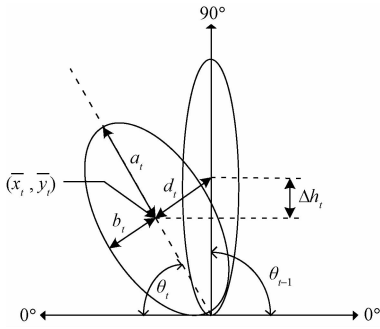


图2 椭圆拟合人体原理

Fig.2 Principle of ellipse fitting body

其中 $2a_t, 2b_t$ 分别表示第 t 帧图像拟合椭圆的长轴和短轴, (\bar{x}_t, \bar{y}_t) 表示 t 帧椭圆中心, d_t 表示中心位置的变化, Δh_t 表示重心高度的变化, θ_t 表示椭圆的角度, 取值范围为 $0^\circ \sim 90^\circ$ 。

对二值图像 $f(x, y)$, 图像矩和中心矩的定义分别为式(8)和(9)。

$$m_{pq} = \sum_{x,y} x^p y^q f(x, y) \quad p, q = 0, 1, 2, \dots \quad (8)$$

$$u_{pq} = \sum_{x,y} (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad p, q = 0, 1, 2, \dots \quad (9)$$

利用一阶图像矩和中心矩, 可以计算椭圆的中心 (\bar{x}, \bar{y}) 、椭圆旋转角度, 式(11)中角度 θ 经过不同象限的对应变换可化为图2中 $0^\circ \sim 90^\circ$ 形式的角度值。

$$\bar{x} = m_{10}/m_{00}; \quad \bar{y} = m_{01}/m_{00} \quad (10)$$

$$\theta = \frac{1}{2} \arctan\left(\frac{2u_{11}}{u_{20} - u_{02}}\right) \times \frac{180}{\pi} \quad (11)$$

利用二阶中心矩可以计算椭圆的长轴 $2a$ 、短轴 $2b$ 和椭圆率 α 。

$$a = (4/\pi)^{1/4} [(I_{\max})^3 / I_{\min}]^{1/8} \quad (12)$$

$$b = (4/\pi)^{1/4} [(I_{\min})^3 / I_{\max}]^{1/8}$$

$$\text{式中: } I_{\max} = \frac{u_{20} + u_{02} + \sqrt{(u_{20} - u_{02})^2 + 4u_{11}^2}}{2}, \quad I_{\min} = \frac{u_{20} + u_{02} - \sqrt{(u_{20} - u_{02})^2 + 4u_{11}^2}}{2}.$$

$$\alpha = a/b. \quad (13)$$

根据椭圆拟合原理以及人体运动的性质, 选取5个特征作为运动特征, 分别是椭圆角度 θ 、椭圆角度变化率 $\Delta\theta$ 、椭圆中心变化 d 、重心高度变化 Δh 、椭圆率 α 和对第 t 帧图像, 则:

$$\theta_t = \theta_t$$

$$\Delta\theta_t = |\theta_t - \theta_{t-1}|$$

$$d_t = \sqrt{(\bar{x}_t - \bar{x}_{t-1})^2 + (\bar{y}_t - \bar{y}_{t-1})^2}$$

$$\Delta h_t = \bar{y}_t - \bar{y}_{t-1}$$

$$\alpha_t = a_t/b_t \quad (14)$$

2.2.2 形态特征

形态特征用于描述人体的形态变化, 在老人摔倒等动作中, 身体的躯干会发生形态上的变化, 而相同的动作类别间其形态变化具有一致性, 不同的动作类别间人体的形态变化区别较大。本文通过提取人体的躯干中心变化来表达运动过程中人体形态的变化。

形态特征的提取主要步骤如下。

1) 根据拟合椭圆的中心坐标和长轴, 提取以椭圆中心坐标为中心点、以椭圆长轴为边长的矩形 ROI 区域。

2) 将 ROI 区域调整为 40×40 大小。

3) 将图像按行平均分成8部分, 提取每一部分的中心点 (\bar{x}_i, \bar{y}_i) 作为该段前景的特征点, 8个特征点可以近似描述躯干的主轴信息, 由于不同的图像8部分中对应部分的 x 坐标近似一致, 故只考虑 y 坐标, 有:

$$\bar{y}_k = \frac{\sum_i 1\{(x^{(i)}, y^{(i)})\} y^{(i)}}{\sum_i 1\{(x^{(i)}, y^{(i)})\}} \quad (15)$$

式中: $1\{(x^{(i)}, y^{(i)})\}$ 为指示函数, 若坐标 $(x^{(i)}, y^{(i)})$ 处为前景像素, $1\{(x^{(i)}, y^{(i)})\} = 1$, 否则 $1\{(x^{(i)}, y^{(i)})\} = 0$ 。为消除运动方向对结果的影响, 取特征点 y_k 对中心的偏移量 Δy_k 为最终的特征数值, 则:

$$\Delta y_k = | \bar{y}_k - 20 | \quad k = 1, 2, \dots, 8 \quad (16)$$

对每一帧图像, 形态特征的特征向量为 $[\Delta y_1 \ \Delta y_2 \ \Delta y_3 \ \dots \ \Delta y_8]$, 结合运动特征的5个特征, 可组成一个13维的向量作为最终的特征。

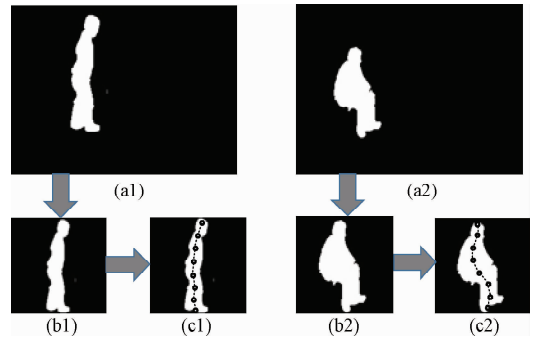


图3 形态特征提取原理

Fig.3 Schematic diagram of extracting morphological features

2.3 动作识别

基于视频的动作识别可以认为是一个时间序列分析问题, 直观上对于每一个动作, 可以认为是经过动作开始、几个动作进行阶段、动作结束的一个过程。实际中观察到的是一个图像序列, 经过特征提取得到的是特征向量序列, 该特征向量序列描述了摔倒等某个特定的动作。HMM 是一种适用于描述动态时间序列的模型, 模型一方面描述了状态之间的转移, 另一方面在描述了状态和观察符号之间的统计对应关系。因此, 本文采用 HMM 实

现对动作的识别。同时考虑到动作过程的特性,假设状态之间的转移具有方向性,且每一个状态对应的观察符号服从于高斯分布。因此,本文采用具有高斯输出的从左到右型的隐马尔科夫模型,其状态转移具有方向性。

隐马尔科夫模型元素构成主要包括以下 5 个部分^[14]。

1) 状态数目为 N , 状态集合可表示为 $S = \{s_1, s_2, \dots, s_N\}$ 。

2) 状态转移矩阵 A 为:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}$$

式中: a_{ij} 表示从状态 s_i 转移到状态 s_j 的概率, 文本从左到右型 HMM, 因此对 A 中, 只有 a_{ii} 和 $a_{i,i+1}$ 不为 0, 其余元素值均为 0。

3) 输出概率密度函数 $B = \{b_j(x) | b_j(x) = p(x | s_j)\}$, 当输出服从单高斯分布时:

$$b_j(x) = p(x | s_j) = N(x; \mu^j, \Sigma^j)$$

4) 序列长度 T 。

5) 特征向量序列 $X, X = \{x_1, x_2, \dots, x_T\}$

隐马尔科夫模型作为一种生成式模型, 其实现动作识别的根据是待测试序列在不同隐马尔科夫模型下产生的概率, 将待测试序列归属为概率最大的那个动作类别。利用 HMM 实现动作分类主要包括两部分内容: 1) 输出概率的计算, 分别计算待测试序列由不同动作所对应的隐马尔科夫模型生成的概率; 2) 模型参数的训练, 主要要求取模型状态转移矩阵参数和每个状态对应的高斯分布的参数。

2.3.1 概率计算与动作分类

要实现对未知类别序列 X 的判断, 需要对每个动作类别建立一个 HMM, 对于一个具有 C 个动作类别的分类器, 选择最能匹配待测试序列的那个 HMM。假设每个 HMM 参数为 $\lambda_i = \{A_i, B_i\}$, 分别计算 $P(X | \lambda_i)$ 按照输出概率最大原则, 将 X 归为使 $P(X | \lambda_i)$ 最大的模型, 从而实现对动作的分类。

概率 $P(X | \lambda)$ 的计算采用前向算法。定义前向变量 $\alpha_t = p(x_1, x_2, \dots, x_t, s(t) = s_j | \lambda)$, 初始化前向变量:

$$\begin{aligned} \alpha_0(s_j) &= 1 \\ \alpha_0(s_j) &= 0, s_j \neq s_1 \end{aligned} \quad (18)$$

利用迭代公式迭代计算前向变量:

$$\alpha_t(s_j) = \sum_{i=1}^N \alpha_{t-1}(s_i) a_{ij} b_j(x_t) \quad (19)$$

则:

$$p(x | \lambda) = \alpha_T(s_E) = \sum_{i=1}^N \alpha_T(s_i) a_{iE} \quad (20)$$

2.3.2 参数模型的学习

当输出概率服从高斯分布时, 所需要确定的参数主要为状态转移矩阵 A 和状态 s_j 所对应的高斯分布的参数 (μ^j, Σ^j) , 记 HMM 模型参数 $\lambda = \{A, \mu, \Sigma\}$ 。本文采用 EM 算法求解模型参数。

利用 EM 算法估计隐马尔科夫模型的参数时, 首先定与前向变量类似的后向变量: $\beta_t(s_j) = p(x_{t+1}, \dots, x_T | s(t) = s_j, \lambda)$ 。

步骤 E: 利用当前 HMM 模型参数, 迭代计算前向变量 $\alpha_t(s_j)$ 和后向变量 $\beta_t(s_j)$; 利用 $\alpha_t(s_j)$ 和 $\beta_t(s_j)$, 计算后验状态概率 $\gamma_t(s_j)$ 和 $\xi_t(s_i, s_j)$ 为:

$$\begin{aligned} \gamma_t(s_j) &= P(S(t) = s_j | X, \lambda) = \\ &= \frac{1}{\alpha_T(s_E)} \alpha_t(j) \beta_t(j) \end{aligned} \quad (21)$$

$$\begin{aligned} \xi_t(s_i, s_j) &= P(S(t) = s_i, S(t+1) = s_j | X, \lambda) = \\ &= \frac{1}{\alpha_T(s_E)} \alpha_t(s_i) a_{ij} b_j(x_{t+1}) \beta_{t+1}(s_j) \end{aligned} \quad (22)$$

步骤 M: 利用 $\gamma_t(s_j)$ 和 $\xi_t(s_i, s_j)$ 重新估计 HMM 模型的参数, 包括均值向量 $\hat{\mu}^j$, 协方差矩阵 $\hat{\Sigma}^j$ 和状态转移概率 \hat{a}_{ij} 。

$$\hat{\mu}^j = \frac{\sum_{t=1}^T \gamma_t(s_j) x_t}{\sum_{t=1}^T \gamma_t(s_j)} \quad (23)$$

$$\hat{\Sigma}^j = \frac{\sum_{t=1}^T \gamma_t(s_j) (x_t - \hat{\mu}^j)(x_t - \hat{\mu}^j)^T}{\sum_{t=1}^T \gamma_t(s_j)} \quad (24)$$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^T \xi_t(s_i, s_j)}{\sum_{k=1}^N \sum_{t=1}^T \xi_t(s_i, s_k)} \quad (25)$$

3 实验结果与分析

本文实现动作识别的算法主要包括人体检测、特征提取和动作识别 3 个主要部分。通过人体检测得到目标前景图像, 利用特征提取得到特征向量, 通过动作识别实现不同动作的判断。本文在模型学习以及测试验证阶段视频数据主要采用文献[15]的视频数据包, 该视频数据包包含了室内摔倒、坐下、蹲下等基本动作, 且拍摄角度与视频监控的要求一致。本文在 PC 端利用 C++ 语言实现算法中的检测和动作的识别, 用 MATLAB 实现 HMM 模型参数的学习。

3.1 人体特征提取

利用本文第 2 节描述的人体检测算法, 并辅以必要的形态学处理, 最终可以得到所需要的人体前景图像。如图 4 所示, 对每一帧原始图像处理都得到了对人体有比较完整描述的二值图像, 可以满足特征提取的要求。



图 4 摔倒过程前景提取结果

Fig. 4 Foreground extraction result of fall

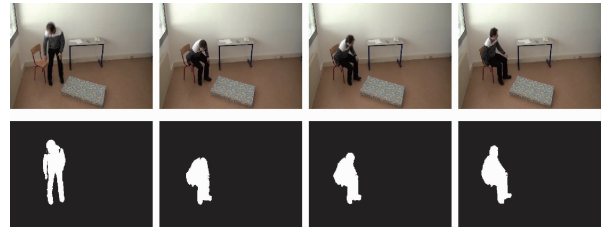


图 5 坐下过程前景提取结果

Fig. 5 Foreground extraction result of siting

人体检测步骤得到了只包含目标前景的二值图像,对每一帧二值图像,通过特征选择步骤可以得到对应的特征向量。实验显示,摔倒、坐下等动作其主要动作变化在 1 s 左右的时间内足够完成,因此,对每一个动作,采用

连续的 30 帧图像进行动作的表达。且每一帧对应的特征为 13 维向量,故每一个动作最终表示为一个 13×30 的时间序列,部分帧特征向量如表 1 所示。

表 1 摔倒动作部分特征向量
Table 1 Partial feature vectors of fall

帧数	1	2	3	4	5	6	7	8	9	10	11	12	13
13	71.13	3.76	6.42	4.87	2.48	7.34	4.54	0.11	0.41	1.27	0.15	1.94	6.00
14	66.77	4.36	7.25	5.77	1.96	8.51	5.21	0.38	0.26	0.94	0.60	1.99	5.95
15	62.02	4.75	7.17	6.24	1.66	9.9	4.9	0.20	0.79	0.81	1.77	3.16	3.37
16	51.65	10.4	9.63	7.67	1.23	10.3	2.82	0.59	0.20	0.49	2.59	4.98	1.93
17	13.93	37.7	9.01	7.81	1.15	12.4	2.62	0.17	0.09	1.44	2.47	3.26	3.80
18	24.34	10.4	12.1	11.3	1.24	11.8	2.06	0.13	1.92	3.70	2.52	1.60	15.0
19	29.17	4.83	9.70	8.87	1.38	20.0	0.03	1.06	2.86	4.54	2.01	10.2	13.0

表 1 是某一个摔倒动作中 13 ~ 19 帧的特征。其中每一帧的 13 个数值分别对应运动特征中的椭圆角度 θ 、椭圆角度变化率 $\Delta\theta$ 、椭圆中心变化 d 、重心高度变化 Δh 、椭圆率 α 和形态特征中的 8 个值。从表中可以看出,摔倒动作发生时,各特征值发生明显的变化;而不同的动作发生变化的趋势不同,故可以利用上述特征通过 HMM 建模实现动作的识别。

3.2 动作识别

在所使用的视频数据包中,主要包括摔倒、坐下、蹲下和行走 4 种动作,本文以分开这 4 类动作为目标,实现动作的识别。经过人体检测和特征提取,已获得每个动作的多个样本。取每个动作的 15 个样本作为训练本来学习模型的参数 $\lambda = \{A, \mu, \Sigma\}$,取每种动作的 10 个样本作为测试样本,来测试模型的准确性。模型参数的训练过程包括以下 4 个步骤。

1) 初始化基本参数:HMM 状态数目 $N = 5$,每个动作序列长度 $T = 30$,每个特征向量长度 $L = 13$ 。

2) 初始化模型参数 $\lambda = \{A, \mu, \Sigma\}$:以序列长度 T 与状态数目 N 之比为每个状态平均分配对应的观测向量

x , 设每个状态 s_i 对应的观测向量的数量为 c_i ,则状态转移变量初始化为 $a_{i,i} = \frac{c_i}{c_i + 1}$, $a_{i,i+1} = \frac{1}{c_i + 1}$ 根据每个状态对应的观测向量初始化 μ^j 和 Σ^j 。

3) 利用 EM 算法学习模型的参数。

4) 保存模型的参数:利用训练出的各动作模型参数,根据第 2 节中的概率计算公式 $P(X | \lambda_i)$ 可以得到待测试动作序列由不同模型产生的概率,比较概率的大小,将待测试动作归类到概率最大的动作类别。

表 2 给出了 5 个摔倒样本在不同动作 HMM 模型下产生的概率的对数值。其中 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 分别代表摔倒、坐下、蹲下和行走的模型参数。

由表 2 可知,摔倒样本在各个动作模型下的输出概率有较大差别,其中明显在摔倒动作模型下有比较大的输出概率,试验中,对其他动作有同样类似的结果。在每个动作分别有 10 个测试动作时,测试结果汇总如表 3 所示,其中 data1 ~ data10 分别代表每个动作的 10 个测试样本,action1 ~ action4 分别表示摔倒、坐下、蹲下和行走 4 种动作,表格中的 1、2、3 和 4 代表每个样本实际判别的动作类别。

表2 摔倒动作由不同动作模型产生的概率

Table 2 The probability of fall sample created by different action models

	$\log P(X \lambda_1)$	$\log P(X \lambda_2)$	$\log P(X \lambda_3)$	$\log P(X \lambda_4)$
摔倒样本1	-724.447	-852.454	-759.367	-920.569
摔倒样本2	-787.056	-2555.03	-1208.46	4570.74
摔倒样本3	-698.186	-1579.92	-861.978	-2752.44
摔倒样本4	-1000.48	-4486.25	-1406.1	-9120.01
摔倒样本5	-1069.84	-2922.07	-1287.8	-4752.33

表3 测试数据动作识别结果

Table 3 Action recognition results of test samples

	data1	data2	data3	data4	data5	data6	data7	data8	data9	data10
action1	1	1	1	1	1	1	1	1	1	1
action2	2	2	2	2	2	2	2	2	2	2
action3	1	1	3	3	3	2	3	3	3	3
action4	4	4	4	4	3	4	4	4	4	4

由表3可知,存在误分类的数据是蹲下动作样本中的3个样本,以及行走动作样本中的1个样本。结合表2和3,摔倒样本和蹲下样本动作过程本身存在一定的相似性,在表2中由摔倒样本在各个动作模型下的输出概率可知这一点。同时,测试结果表明摔倒动作能够全部检测出来,算法的整体准确率达到90%,在实验室实景室内监控测试实验中,也得到了类似的效果。鉴于系统算法本身的目的更侧重于能够及时发现类似摔倒等对老人身体有较大危害的動作的发生。因此本文所提设计方案具有一定的实用价值。

4 结论

本文针对老年公寓中或独居老人安全监护,提出了基于视频监控的动作识别设计方案。在具体实现中,采用基于HSV空间的混合高斯建模方法实现人体的检测,采用运动特征和人体形态特征相结合的特征提取方法实现对动作的描述,采用具有高斯输出的HMM模型实现动作的识别。本文所提出的设计方案,一方面能够适用于光照变化等室内场景,另一方面所采取的特征提取方式也能够适用于动作方向变化等情况,具有较好的鲁棒性。实验结果表明,该设计方案对室内摔倒等动作的识别有比较好的识别结果,具有一定的实际应用价值。

参考文献

[1] CHARFI I, MITERAN J, DUBOIS J, et al. Optimized spatio-temporal descriptors for real-time fall detection: Comparison of support vector machine and Adaboost-

based classification[J]. Journal of Electronic Imaging, 2013,22(4): 041106-041106.

- [2] SENOUCI B, CHARFI I, HEYRMAN B, et al. Fast prototyping of a SoC-based smart-camera: a real-time fall detection case study[J]. Journal of Real-Time Image Processing, 2016,12(4):649-662.
- [3] MUBASHIR M, SHAO L, SEED L. A survey on fall detection: Principles and approaches [J]. Neurocomputing, 2013, 100(2): 144-152.
- [4] 孟明,杨方波,余青山,等. 基于 Kinect 深度图像信息的人体运动检测[J]. 仪器仪表学报,2015,36(2): 386-393.
- MENG M, YANG F B, SHE Q SH, et al. Human motion detection based on the depth image of Kinect [J]. Chinese Journal of Scientific Instrument, 2015, 36(2): 386-393.
- [5] 杜往泽,宋执环,闫文博,等. 单摄像头旋转监控下的快速图像拼接[J]. 中国图象图形学报,2016,21(2): 245-254.
- DU W Z, SONG ZH H, YAN W B, et al. Fast image stitching under single camera rotate monitoring [J]. Journal of Image and Graphics,2016,21(2):245-254.
- [6] JIANG M, CHEN Y, ZHAO Y, et al. A real-time fall detection system based on HMM and RVM[C]. Visual Communications and Image Processing (VCIP), 2013:1-6.
- [7] YU M, RHUMA A, NAQVI SM, et al. A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment[J]. IEEE Transactions on Information Technology in Biomedicine, 2012,16(6):1274-1286.

- [8] RHUMA A, YU M, CHAMBERS J. Posture recognition based fall detection system [J]. Lecture Notes on Software Engineering, 2013, 1(4):350-355.
- [9] FOROUGHHI H, NASERI A, SABERI A, et al. An eigenspace-based approach for human fall detection using integrated time motion image and neural network [C]. 9th International Conference on Signal Processing, IEEE, 2008: 1499-1503.
- [10] NAIT-CHARIF H, MCKENNA S J. Activity summarisation and fall detection in a supportive home environment [C]. Proceedings of the 17th International Conference on Pattern Recognition, 2004: 323-326.
- [11] 张燕平,白云球,赵勇,等.应用改进混合高斯模型的运动目标检测 [J]. 计算机工程与应用, 2010, 46(34): 155-157.
- ZHANG Y P, BAI Y Q, ZHAO Y, et al. Moving object detection based on improved Gaussian mixture models [J]. Computer Engineering and Applications, 2010, 46(34): 155-157.
- [12] ZHAO M, BU J, CHEN C. Robust background subtraction in HSV color space [C]. Convergence of Information Technologies and Communications, 2002: 325-332.
- [13] PRATI A, MIKIC I, TRIVEDI M M, et al. Detecting moving shadows: Algorithms and evaluation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(7): 918-923.
- [14] HAN S, ZHANG M, LI P, et al. SVM-HMM based human behavior recognition [C]. International Conference on Human Centered Computing, 2014: 93-103.
- [15] CHARFI I, MITERAN J, DUBOIS J, et al. Definition and performance evaluation of a robust svm based fall detection solution [C]. 8th International Conference on Signal Image Technology and Internet Based Systems (SITIS), IEEE, 2012: 218-224.

作者简介



潘月浩, 1991年出生, 2014年于合肥工业大学获得学士学位, 现为浙江大学硕士研究生, 主要研究方向为计算机视觉。

E-mail: pyhhfut@163.com

Pan Yuehao was born in 1991, and received B. Sc. from Hefei University of Technology in 2014. Now he is a M. Sc. candidate in Zhejiang University. His main research interest is computer vision.



宋执环, 1962年出生, 1997年于浙江大学获得博士学位, 现为浙江大学教授、博士生导师, 主要研究方向为复杂工程系统安全监测与故障诊断。

E-mail: songzhihuan@zju.edu.cn

Song Zhihuan was born in 1962, and received Ph. D. from Zhejiang University in 1997. Now he is a professor and Ph. D. supervisor in Zhejiang University. His main research interest is safety monitoring and fault diagnosis of complex engineering system.