

基于耦合多隐马尔可夫模型和深度图像数据的人体动作识别

张全贵, 蔡 丰*, 李志强

(辽宁工程技术大学 电子与信息工程学院, 辽宁 葫芦岛 125105)

(*通信作者电子邮箱 545130868@qq.com)

摘 要:为解决使用 RGB 图像进行特征提取时容易受外界因素干扰,且计算复杂度高等问题,采用一种更加有效的解决方案,即使用深度数据进行人体动作识别。利用 Kinect 采集的关节点数据,首先将人体关节划分成五个区域,对每个区域的向量夹角离散化从而描述不同的状态,再通过 Baum-Welch 算法学习出各区域的多隐马尔可夫模型(multi-HMM),并使用前向算法建立生成区域与动作类别概率矩阵。在此基础上,对区域及动作类别进行内耦合和间耦合分析,从而表达各关节点之间的交互关系。最后使用基于耦合的 K 最邻近(KNN)算法完成整体的动作识别。通过实验测试对五种动作的识别率均达到 90% 以上,并与 3D Trajectories 等方法进行对比,实验得到的综合识别率高于对比方法,具有明显的优势。

关键词:Kinect; 人体动作识别; 划分区域; 多隐马尔可夫模型; 耦合 K 最邻近

中图分类号: TP391.41; TP18 **文献标志码:** A

Human action recognition based on coupled multi-Hidden Markov model and depth image data

ZHANG Quangui, CAI Feng*, LI Zhiqiang

(School of Electronics and Information Engineering, Liaoning Technical University, Huludao Liaoning 125105, China)

Abstract: In order to solve the problem that the feature extraction is easy to be affected by external factors and the computational complexity is high, the depth data was used for human action recognition, which is a more effective solution scheme. Using the joint data collected by Kinect, the human joint was divided into five regions. The vector angle of each region was discretized to describe different states, and then Baum-Welch algorithm was used to study multi-Hidden Markov Model (multi-HMM), meanwhile, forward algorithm was used to establish the generation region and action class probability matrix. On this basis, the region and action categories were intra-coupled and inter-coupled to analyze, thus expressing the interaction between the joints. Finally, the K-Nearest Neighbors (KNN) algorithm based on coupling was used to complete the action recognition. The experimental results show that the recognition rates of the five actions reach above 90%, and the comprehensive recognition rate is higher than that of the contrast methods such as 3D Trajectories, which means that the proposed algorithm has obvious advantages.

Key words: Kinect; human action recognition; divide area; multi-HMM (multi-Hidden Markov Model); coupled K-Nearest Neighbors (KNN)

0 引言

人体动作识别是为了让计算机通过一些方法判别出捕捉到的人体动作类别,在行为科学^[1]、社交媒体等方面得到了广泛的应用及发展,如智能视频监控、家庭服务机器人^[2]等。在微软公司发布 Kinect 之后,使得获取的图像数据受较少外部因素干扰而更为准确,因此近年来基于深度图像数据的人体动作识别成为一个研究热点。例如,申小霞等^[3]利用 Kinect 获取图像信息及金字塔特征来描述行为信息,并且通过使用支持向量机(Support Vector Machine, SVM)算法进行行为分类;Shotton 等^[4]提出一种新的方法从一个单一的深度图像预测人类姿势,该方法基于当前目标识别策略,通过设计中表示部分把复杂姿态问题化为简单问题;张毅等^[5]将行为

识别应用到医疗康复训练方面,同样采用 Kinect 采集数据,同时使用了朴素贝叶斯模型,并通过调整模型阈值提高识别率。

以上这些方法都具有一定的优势,并且能达到较好的识别效果。但现有的基于人体关节点进行人体动作识别的研究一般假设各关节之间是相互独立的,即没有考虑关节之间的相互关系。根据人体运动解剖学,人体在完成某个动作的时候各个关节之间是相互制约和协同工作的,即关节之间具有一定的耦合关系^[6]。例如,从桌子上拿起水杯这个动作,胳膊的肘部活动一定会影响到手腕的活动。如果忽略各部位之间的交互,这样分析可能会导致大量信息的损失,因此本文将考虑人体各部分之间的相互影响关系,即耦合关系。建立人体动作的多隐马尔可夫模型(multi-Hidden Markov Model, Multi-HMM),并运用耦合相似度分析与 K 最邻近

收稿日期:2017-08-09;修回日期:2017-09-09。 基金项目:辽宁省自然科学基金面上项目(2015020100)。

作者简介:张全贵(1978—),男,辽宁葫芦岛人,副教授,博士,CCF 会员,主要研究方向:计算机视觉、机器学习、数据挖掘; 蔡丰(1992—),女,辽宁大连人,硕士研究生,主要研究方向:计算机视觉; 李志强(1993—),男,辽宁锦州人,硕士研究生,主要研究方向:计算机视觉、数据挖掘。

(K-Nearest Neighbors, KNN)算法^[7]进行基于耦合的人体动作识别,以便充分利用人体各个关节点之间的协同交互关系。

1 特征值提取

本文将 Kinect 提取的人体 20 个节点分成 5 个区域,从上到下、从左到右依次为左上肢、躯干、右上肢、左下肢和右下肢,分别用 R_1, R_2, R_3, R_4 和 R_5 来表示。在每个区域中选取合适的向量及夹角,通过对向量夹角的划分进行特征值的提取,完成人体动作特征的描述。

人体各关节点名称如图 1(a) 所示;为方便描述将每个节点名称用英文字母代替全称,如图 1(b) 所示。例如,在 R_1 区域内,节点 c 与节点 d 形成向量 V_{cd} ,节点 c 与节点 b 形成向量 V_{cb} ,向量 V_{cd} 与向量 V_{cb} 所形成夹角用 A_{cd-cb} 表示。在其他四个区域以同样的方式选取向量及夹角。具体所选的向量如表 1 所示。

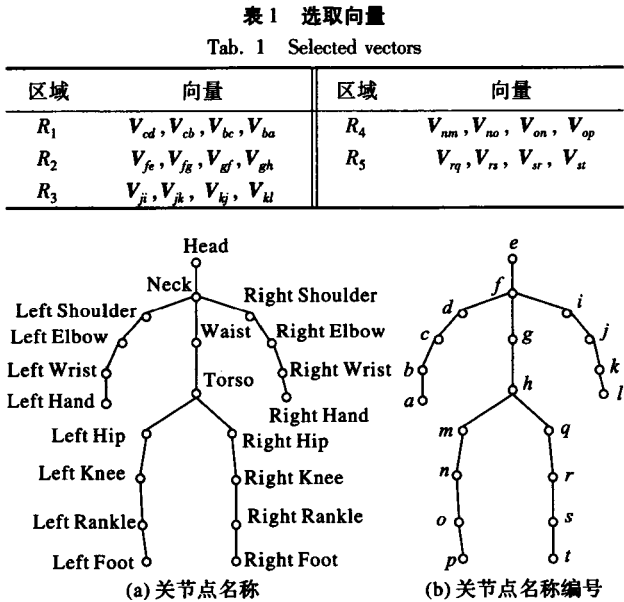


图 1 关节点名称和编号示意图

Fig. 1 Schematic diagram for name and label of joint nodes

在空间坐标系中左上肢与右上肢相对应,左下肢与右下肢对应,结合实际情况活动的范围大致相同,所以会有相同的角度划分。例如在区域 R_1 中,左上臂与左前臂活动角度的范围比较大,所以设置的角度范围是最小 0° 到最大 180° ,以间隔为 20° 将此区间划分成 9 个区间,依次编号为 0~8。左前臂与左手动作范围较小,所以设置为 0° 到 90° ,间隔区间为 10° ,同样将此范围划分成 9 个区间,编号为 0~8。右上肢将有与左上肢相同的角度划分。本文将五个区域中的向量夹角均划分成 9 个区间,编号为 0~8。经过实验分析得到每个角度的最佳变化范围及划分区间如表 2 所示。

表 2 角度划分

Tab. 2 Angle division

向量夹角	角度范围	向量夹角	角度范围
A_{cd-cb}	$0^\circ \sim 180^\circ$	A_{gf-gh}	$0^\circ \sim 180^\circ$
A_{bc-ba}	$0^\circ \sim 90^\circ$	A_{nm-no}	$0^\circ \sim 180^\circ$
A_{fe-fg}	$0^\circ \sim 180^\circ$	A_{on-op}	$0^\circ \sim 180^\circ$

当求两个空间向量之间的夹角时,需要使用余弦函数,例如有向量 $a = (x_1, y_1, z_1)$ 和向量 $b = (x_2, y_2, z_2)$,则向量 a, b

之间夹角可以由式(1) 计算得到:

$$\langle a, b \rangle = \arccos \frac{a \cdot b}{|a| \cdot |b|}$$

(1)

其中 $|a| \neq 0, |b| \neq 0$ 。如果 $|a| \neq 0$ 且 $|b| \neq 0$,而 $a \cdot b = x_1x_2 + y_1y_2 + z_1z_2, |a|^2 = x_1^2 + y_1^2 + z_1^2$,则 $\langle a, b \rangle = 0$,即 a 和 b 为正交向量。例如,通过式(1) 求得某个区域的一个夹角大小为 23° 时,根据角度及编号的划分可以得到:当此向量角处于 $0^\circ \sim 180^\circ$ 范围时,它对应的编号为 1;当此向量角处于 $0^\circ \sim 90^\circ$ 范围时到此角度对应标号为 2。

本文将取 T 帧视频序列,每个分区将形成一个观察序列 $(o_{n1}, o_{n2}, \dots, o_{nT})$,其中 n 表示分区,取值为 $\{1, 2, \dots, 5\}$ 。

2 动作识别

2.1 基于隐马尔可夫模型的局部分类

最初的隐马尔可夫模型(Hidden Markov Model, HMM)^[8]是在统计学中提出的。一个 HMM 包含两组状态集合和三组概率集合,可以用五个元素来描述;记为 $\lambda = (S, O, \pi, A, B)$ 。其中: S 表示隐含状态,所有隐含状态之间满足马尔可夫性; O 表示可观测状态; π 表示初始状态概率矩阵; A 表示隐含状态转移概率矩阵; B 表示观测状态转移概率矩阵。也可以将 HMM 用一个三元组 (π, A, B) 来表示。

在第 1 章中已经获取了每个分区的观测序列,接下来将使用 Baum-Welch 算法训练得到各个分区的多个 HMM,并使用前向算法计算观察序列的概率。具体流程如下:

算法 1 分区训练 HMM 并计算观察序列概率。

输入 分区观察序列 $O = \{o_n\}$,其中 $n = \{1, 2, \dots, 5\}$, $1 \leq t \leq T$ 。

过程:

- 1) 初始化参数模型 λ_0 。
- 2) 根据式(2) 计算 $t+1$ 时刻状态为 j 的概率。

$$\varepsilon_t(i, j) = P(q_t = i, q_{t+1} = j | O, \lambda) = \frac{P(q_t = i, q_{t+1} = j, O | \lambda)}{P(O | \lambda)} = \frac{a_i(i) a_{ij} b_{j, t+1} \beta_{t+1}(j)}{P(O | \lambda)} = \frac{a_i(i) a_{ij} b_{j, t+1} \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N a_i(i) a_{ij} b_{j, t+1} \beta_{t+1}(j)}$$

(2)

其中: $1 \leq i \leq N, 1 \leq j \leq N, 1 \leq k \leq M, N$ 表示隐含状态的个数, M 表示观测状态的个数; $\beta_{t+1}(j)$ 表示后向概率。

- 3) 根据式(3) 计算 t 时刻状态 i 的概率。

$$\gamma_t(i) = \sum_{j=1}^N \varepsilon_t(i, j)$$

(3)

其中: $1 \leq i \leq N, 1 \leq t \leq T$ 。

- 4) 根据式(4)、(5)、(6) 计算分别得到模型的三个参数值,即能获得模型 λ_{i+1} 。

$$\pi_i(i) = \gamma_t(i)$$

(4)

$$a_{ij} = \sum_{t=1}^{T-1} \varepsilon_t(i, j) / \sum_{t=1}^{T-1} \gamma_t(i)$$

(5)

其中: $1 \leq i \leq N, 1 \leq j \leq N, 1 \leq t \leq T-1$ 。

$$b_{jk} = \sum_{t=1, o_t=k}^T \gamma_t(i) / \sum_{t=1}^T \gamma_t(i)$$

(6)

其中: $1 \leq j \leq N, 1 \leq k \leq M$ 。

- 5) 重复步骤 2)、3)、4),循环迭代直到 π_i, a_{ij}, b_{jk} 三个参

数收敛,为这次训练的最终模型。

6) 模型训练结束后,取新的观察序列。使用前向算法:当 $t = 1$ 时,根据式(7)计算局部概率;当 $t > 1$ 时,采用递归的方法,根据式(8)计算当前 HMM 下一个观察序列的概率。

$$a_1(j) = \pi(j)b_{j_1} \quad (7)$$

$$a_{t+1}(j) = \left[\sum_{i=1}^N a_t(i)a_{ij} \right] b_{j_{t+1}} \quad (8)$$

7) 根据式(9)计算给定观察序列的概率等于 T 时刻所有局部概率之和。

$$P(O/\lambda) = \sum_{j=1}^N a_T(j) \quad (9)$$

输出 结果为当前的观察序列在 HMM 下的概率值。

2.2 基于 HMM 的局部分类

观察序列输入到模型后会计算得到概率矩阵如下: $P = \{p_{xy}\}$, p_{xy} 表示第 x 分区中第 y 个 HMM 对观察序列的概率,其中 x, y 取值均为 $\{1, 2, \dots, 5\}$ 。

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{15} \\ p_{21} & p_{22} & \dots & p_{25} \\ \vdots & \vdots & & \vdots \\ p_{51} & p_{52} & \dots & p_{55} \end{pmatrix}$$

当某个动作发生时,通过矩阵 P 可以知道是身体哪个区域发出这个动作,但是并不能知道此动作是由身体的单一区域还是多个区域共同完成。基于以上工作,通过进一步分析各个概率值之间耦合关系,并使用基于耦合的 KNN 算法得到最终的动作分类。因此需要分析处理:将得到的矩阵 P 变换成 25 维的向量,即 $(p_{11}, p_{12}, \dots, p_{15}, p_{21}, p_{22}, \dots, p_{25}, \dots, p_{51}, p_{52}, \dots, p_{55})$, 并且这个向量在训练数据中有相应的分类标签。每一个视频序列都形成一个向量,那么取 m 个视频序列 $\{s_1, s_2, \dots,$

$$R_{-le}(a_f | \{a_g\}_{g \neq f}) = \begin{pmatrix} \beta_{11}(f|g_1) & \dots & \beta_{1L}(f|g_1) & \dots & \beta_{11}(f|g_{n-1}) & \dots & \beta_{1L}(f|g_{n-1}) \\ \beta_{21}(f|g_1) & \dots & \beta_{2L}(f|g_1) & \dots & \beta_{21}(f|g_{n-1}) & \dots & \beta_{2L}(f|g_{n-1}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ \beta_{L1}(f|g_1) & \dots & \beta_{LL}(f|g_1) & \dots & \beta_{L1}(f|g_{n-1}) & \dots & \beta_{LL}(f|g_{n-1}) \end{pmatrix}$$

2.2.3 对象的耦合表示及 KNN 分类

每一个属性值扩展后的集合为 $\bar{A} = \{\langle a_1 \rangle^1, \langle a_1 \rangle^2, \dots, \langle a_1 \rangle^L, \langle a_2 \rangle^1, \langle a_2 \rangle^2, \dots, \langle a_2 \rangle^L, \dots, \langle a_n \rangle^1, \langle a_n \rangle^2, \dots, \langle a_n \rangle^L\}$, 设更新信息函数 $\tilde{f}_f^v(u)$ 表示对象 u 在属性 $\langle a_f \rangle^v$ 上对应值,对象 u 在属性值 a_f 上的向量为:

$$\tilde{u}(a_f) = [\tilde{f}_f^1(u), \tilde{f}_f^2(u), \dots, \tilde{f}_f^L(u)]$$

对象 u 在其他的属性值上的向量为:

$$\tilde{u}(\{a_g\}_{g \neq f}) = [\tilde{f}_{g_1}^1(u), \tilde{f}_{g_1}^2(u), \dots, \tilde{f}_{g_1}^L(u), \tilde{f}_{g_2}^1(u), \tilde{f}_{g_2}^2(u), \dots, \tilde{f}_{g_2}^L(u), \dots, \tilde{f}_{g_{n-1}}^1(u), \tilde{f}_{g_{n-1}}^2(u), \dots, \tilde{f}_{g_{n-1}}^L(u)]$$

由公式可以得到对象 u 在属性 a_f 上的 $1 \times L$ 向量:

$$u^c(a_f | \bar{A}, L) = \tilde{u}(a_f) \odot \mathbf{w} \otimes [R_{-le}(a_f | \{a_g\}_{g \neq f})]^T + \tilde{u}(\{a_g\}_{g \neq f}) \odot [\underbrace{\mathbf{w}, \mathbf{w}, \dots, \mathbf{w}}_{n-1}] \otimes [R_{-le}(a_f | \{a_g\}_{g \neq f})]^T$$

其中: \mathbf{w} 是一个 $1 \times L$ 的常量向量, $\mathbf{w} = \left[\frac{1}{(1!)}, \frac{1}{(2!)}, \dots, \frac{1}{(L!)} \right]$; \odot 表示 Hadamard 乘积; \otimes 表示矩阵

s_m 便可以得到 m 个 25 维向量。本文共取 100 个视频序列。每一帧看作一个属性,计算各属性的内耦合及间耦合关系,通过属性的内耦合及间耦合来分析各个动作序列之间的隐含关系。本文借鉴文献[6]的方法计算内耦合、间耦合。

2.2.1 内耦合分析

本节将属性值 a_f 扩展出 $L-1$ 个属性值,本文的 L 取值为 2,分别为 $\langle a_f \rangle^2, \langle a_f \rangle^3, \dots, \langle a_f \rangle^L$, 那么属性值 a_f 的内耦合关系可以表示成 $L \times L$ 的矩阵 $R_{-la}(a_f)$, 其中 $\langle a_f \rangle^v$ 表示属性 a_f 的 v 次幂。

$$R_{-la}(a_f) = \begin{pmatrix} \varphi_{11}(f) & \varphi_{12}(f) & \dots & \varphi_{1L}(f) \\ \varphi_{21}(f) & \varphi_{22}(f) & \dots & \varphi_{2L}(f) \\ \vdots & \vdots & & \vdots \\ \varphi_{L1}(f) & \varphi_{L2}(f) & \dots & \varphi_{LL}(f) \end{pmatrix} \quad (10)$$

其中: $\varphi_{pq}(f) = \text{Cor}(\langle a_f \rangle^p, \langle a_f \rangle^q)$ 是属性 $\langle a_f \rangle^p$ 和属性 $\langle a_f \rangle^q$ 的皮尔逊相关系数。属性 a_f 和 a_g 之间的皮尔逊相关系数为:

$$\text{Cor}(a_f, a_g) = \frac{\sum_{u \in U} (f_f(u) - \mu_f)(f_g(u) - \mu_g)}{\sqrt{\sum_{u \in U} (f_f(u) - \mu_f)^2} \sqrt{\sum_{u \in U} (f_g(u) - \mu_g)^2}}$$

其中 μ_f, μ_g 分别表示属性值 a_f, a_g 的均值。

2.2.2 间耦合分析

假设有 n 个属性值(这里 $n = 25$), 将每一个属性值 a_f 都扩展出 $L-1$ 个属性值: $\langle a_f \rangle^2, \langle a_f \rangle^3, \dots, \langle a_f \rangle^L, \langle a_f \rangle^v$ 表示属性 a_f 的 v 次幂。那么数值型属性 a_f 与属性 a_g 及属性 $\langle a_f \rangle^v (f \neq g)$ 的间耦合关系可以表示为大小为 $L \times (L \times (n-1))$ 的矩阵 $R_{-le}(a_f | \{a_g\}_{g \neq f})$ 。计算公式如下:

相乘。

到此,已完成了多个 HMM 形成的概率矩阵的耦合相似性分析。基于此,使用 KNN 算法完成最后的人体动作分类。在使用 KNN 算法时,使用以上介绍的对象耦合相似度计算方法代替传统的欧氏距离方法。

3 实验分析

3.1 实验数据集及实验过程

实验在公开数据集 MSR-Daily-Activity 3D (MSR Daily Activity 3D Dataset)^[9] 上进行。MSR-Daily-Activity 3D 数据集包含 16 种人体动作,每种动作分别由 10 个人来做,每类动作做 2 次,一共有 320 个样本。该数据集是在客厅中录制的,基本上涵盖了人类的日常活动。行为者是站在沙发附近完成的各种动作,使得骨骼追踪非常复杂,因此使得这个数据集更具有挑战性。该数据集同时提供了三种数据信息:深度图像、骨骼关节位置和 RGB 视频,本文选择骨骼关节位置中的 drinking、cheer up、walk、sit down 和 stand up 共 5 种动作的数据进行实验。

实验时取出上述 5 种动作序列的数据集后分成训练数据

集与测试数据集两组,其中第一个到第八个人的动作数据集用作训练,第九、第十这两个人的动作数据集用作测试。为避免过拟合现象,本文从这些数据集中按照等间距标准从原视频序列选取关键帧组成新的序列。根据第2章中特征提取方法把每种视频序列用矩阵表示,进行HMM训练得到参数 λ_i ,再利用前向算法计算观察序列的概率即完成分区动作识别,最后根据基于耦合关系的KNN算法完成整体的动作识别。

3.2 实验结果与分析

本文动作识别算法的动作识别率结果如表1所示,均在90%以上,其中walk的识别率最高为96.86%,这说明本文算法的识别率水平较高。

表3 动作识别率
Tab. 3 Recognition rates of actions

动作	识别率/%	动作	识别率/%
drinking	90.48	sit down	92.16
cheer up	91.00	stand up	93.14
walk	96.86		

为了进一步验证本文算法的有效性,将其与同样使用此数据集进行实验的文献[10-12]的方法对所有动作的综合识别率进行比较。其中,文献[10]的实验中将动作分为了两类,一类是低运动动作,另一类是高运动动作,两种类型的识别率分别为60%和85%。由于本文所选取的动作属于高运动动作,所以只考虑与文献[10]中高运动动作的识别率比较。文献[11]的数据获取较为复杂,且整体计算量较大,其设置了多组实验,因此这里取平均值与本文方法进行比较。文献[12]使用的是骨骼形状轨迹的方法,而MSR-Daily-Activity 3D数据集本身所处的环境较复杂,所以在用骨骼追踪时比较困难,以至于识别率较低。本文取与文献[12]相同的实验设置,即取前5个人作为训练,后5个人作为测试。

经过实验,本文方法的综合识别率为87.16%,3D Trajectories方法^[10]识别率为85%,Actionlet Ensemble Model方法^[11]识别率为85.75%,Skeletal Shape Trajectories方法^[12]识别率为70.00%。可以看出,本文所提出的识别方法与其他三种方法相比具有较高的识别率,并且达到了较好的预期识别效果,由于能够在复杂的环境下表现良好,所以体现出了较强的鲁棒性。

4 结语

本文使用关节角度值描述特征,通过Baum-Welch算法建立多HMM得到相应参数,根据前向算法对人体分区进行动作识别,然后利用基于耦合关系的KNN算法完成整体的动作识别,充分利用了人体的各个部分之间的隐含交互关系。实验结果表明本文提出的方法具有一定的优越性。

本文方法只识别了五种动作,没有对更复杂背景下的复杂动作进行识别验证,这将是我们要继续研究的内容。

参考文献:

- [1] ZACHARIAS G L, MacMILLAN J, VAN HEMEL S B, et al. Behavioral modeling and simulation: from individuals to societies [J]. Journal of Artificial Societies & Social Simulation, 2008, 12(3): 291-304.
- [2] 田国会,尹建芹,韩旭,等.一种基于关节点信息的人体行为识别

- 新方法[J]. 机器人, 2014, 36(3): 285-292. (TIAN G H, YIN J Q, HAN X, et al. A new method of human behavior recognition based on joint information [J]. Robot, 2014, 36(3): 285-292.)
- [3] 申小霞,张桦,高赞.基于Kinect和金字塔特征的行为识别算法[J]. 光电子·激光, 2014(2): 357-363. (SHEN X X, ZHANG H, GAO Z. Behavior recognition algorithm based on Kinect and pyramid feature [J]. Journal of Optoelectronics · Laser, 2014(2): 357-363.)
- [4] SHOTTON J, FITZGIBBON A, COOK M, et al. Real-time human pose recognition in parts from single depth images [C]// CVPR 2011: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, SCI 411. Washington, DC: IEEE Computer Society, 2011: 1297-1304.
- [5] 张毅,黄聪,罗元.基于改进朴素贝叶斯分类器的康复训练行为识别方法[J]. 计算机应用, 2013, 33(11): 3187-3189. (ZHANG Y, HUANG C, LUO Y. Rehabilitation training behavior recognition method based on improved naive Bayesian classifier [J]. Journal of Computer Applications, 2013, 33(11): 3187-3189.)
- [6] WANG C, SHE Z, CAO L. Coupled attribute analysis on numerical data [C]// IJCAI '13: Proceedings of the 2013 International Joint Conference on Artificial Intelligence. Menlo Park, CA: AAAI Press, 2013: 1736-1742.
- [7] FIGUEIREDO F, ROCHA L, COUTO T, et al. Word co-occurrence features for text classification [J]. Information Systems, 2011, 36(5): 843-858.
- [8] MOGHADDAM Z, PICCARDI M. Training initialization of hidden markov models in human action recognition [J]. IEEE Transactions on Automation Science & Engineering, 2014, 11(2): 394-408.
- [9] WANG J, LIU Z, WU Y, et al. Mining actionlet ensemble for action recognition with depth cameras [C]// CVPR 2012: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2012: 1290-1297.
- [10] KOPERSKI M, BILINSKI P, BREMOND F. 3 D trajectories for action recognition [C]// ICIP 2014: Proceedings of the 2014 IEEE International Conference on Image Processing. Piscataway, NJ: IEEE, 2014: 4176-4180.
- [11] LI W, ZHANG Z, LIU Z. Action recognition based on a bag of 3D points [C]// CVPRW 2010: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Washington, DC: IEEE Computer Society, 2010: 9-14.
- [12] AMOR B B, SU J, SRIVASTAVA A. Action recognition using rate-invariant analysis of skeletal shape trajectories [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(1): 1-13.

This work is partially supported by Liaoning Provincial Natural Science Foundation Project (2015020100).

ZHANG Quanguai, born in 1978, Ph. D., associate professor. His research interests include computer vision, machine learning, data mining.

CAI Feng, born in 1992, M. S. candidate. Her research interests include computer vision.

LI Zhiqiang, born in 1993, M. S. candidate. His research interests include computer vision, data mining.