

采用支持向量机评估老年人步态对称性的研究

吴建宁, 王珏

(西安交通大学生物医学信息工程教育部重点实验室, 710049, 西安)

摘要: 采用支持向量机(SVM)评估老年人步态的对称性. 将鉴别老年人下肢左、右两侧的步态模式相似性问题转化为二分类问题, 通过识别老年人下肢左、右两侧的步态模式, 确定其两侧步态模式相似性的差异, 判断其步态的对称性. 采集 24 名健康老年人下肢左、右两侧的步态数据, 采用交叉验证方法评估 SVM 泛化能力, 测试了多项式核、径向基核、线性核. 结果表明, 多项式核、径向基核的泛化能力优于线性核, 基于多项式核的 SVM 识别左、右两侧老年人步态模式的分类正确率较高(88%), 可有效地提取步态模式的非线性信息. SVM 有望成为评估老年人步态对称性的一个有效的工具, 有助于及早预防老年人跌倒和老年性疾病发生, 提高老年人生活质量.

关键词: 步态分析; 对称性; 支持向量机; 老年人

中图分类号: R318.01 **文献标识码:** A **文章编号:** 0253-987X(2007)08-0995-05

Novel Approach for Evaluation of Elderly Gait Symmetry

Wu Jianning, Wang Jue

(Key Laboratory of Biomedical Information Engineering of Ministry of Education, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The application of support vector machine (SVM) for evaluation of the symmetry of elderly gait is investigated following the assumption that the determination of similarity between the lower limbs can be considered as a binary classification task. Adopting SVM to classify the right and left sides gait patterns of the elderly lower limbs during normal walking, the difference between the corresponding gait parameter curves can be determined effectively to assess the gait pattern symmetry. The bilateral kinetic gait data of 24 elderly participants are recorded by a force platform during normal walking. The cross-validation method is employed to evaluate the generalization ability of SVM, and three types of kernel functions (linear, polynomial and radial basis function) are validated. The experimental results demonstrate that SVM with polynomial kernel outperforms SVM with linear kernel and radial basis function in classification performance to identify the right and left sides gait patterns with an accuracy up to 88%. It is suggested that SVM serves as an efficient tool for evaluation of the symmetry of elderly gait to identify at-risk gaits in older population.

Keywords: gait analysis; symmetry; support vector machine; elderly

目前,随着我国人口老龄化程度不断提高,老年人的健康问题越来越受到全社会的关注. 医学研究表明,在老年人的衰老过程中,由于老年人的神经系统功能和生理机能衰退,导致了老年人行走能力的下降,使其左、右两侧步态不对称,造成老年人在站

立、步行时容易跌倒,从而引发各种老年性疾病^[1-4]. 在临床上,医生仅依靠视觉和经验评估老年人的步态对称性,其评估结果缺乏客观性,严重地影响了老年人异常步态的早期诊断,这就需要客观地量化评估其步态的对称性^[5].

量化评估步态对称性是步态分析研究中的一个热点问题^[6]. 早期的研究主要是利用人体下肢左、右两侧的步态参数(时间-距离参数、运动学参数、动力学参数、肌电参数等)计算绝对对称性指标(A_{SI})判定步态的对称性^[7]. 近年来,一些学者采用传统统计分析方法计算人体下肢左、右两侧的步态参数在统计意义上有无显著差异,判断下肢左、右两侧的步态是否具有相似性,从而判定步态的对称性^[8]. 这些方法虽然简便实用,但是算法简单,常常丢失一些有效的步态信息,增大了量化评估误差^[6]. 目前,为了更好地鉴别人体下肢左、右两侧在行进时的步态相似性,有些学者采用基于多元统计模型的线性主成分分析(PCA)方法,分析人体下肢左、右两侧的动力学步态参数提取步态数据结构特征,比较主成分曲线的相似性,确定步态的对称性^[6,9]. 虽然主成分分析能够通过提取步态数据结构特征进行,但是它仅能获取一些步态数据结构中的线性信息,使得步态数据结构中大量的非线性信息丢失,难以反映步态参数在对称性中的作用,降低了鉴别步态参数曲线相似性的敏感性,评价结果误差较大^[10]. 因此,需要更有效的方法来解决这些不足. 目前,基于 Vapnik-Chervonenkis 理论和结构风险最小化原则的支持向量机(SVM)由于能够解决传统学习算法中的小样本、非线性、高维性、局部极小点等问题,在小样本步态数据的二分类任务中,表现出了较好的泛化能力^[11-12]. 为了能够较准确地鉴别左、右两侧的步态相似性的差异,本文尝试将鉴别下肢左、右两侧步态相似性差异的问题转化为二分类问题. 采集了 24 名健康老年人左、右两足的步态动力学数据,选用垂直方向上的足-地面间反作用力作为步态参数,采用 SVM 识别老年人人体下肢左、右两侧的步态模式,并使用交叉验证方法评估 SVM 的泛化性能,评估老年人步态的对称性.

1 步态数据采集

挑选 24 名健康、神经和运动系统无任何损伤的老年人作为受试者,他们的平均年龄为 68.4 ± 4.3 岁,身高为 165 ± 4.8 cm. 数据采集系统采用加拿大 Bertec 公司开发的标准三维测力台(FP4060-10),该系统由三维力传感器、模拟放大器、模/数转换器、计算机组成. 测力平台尺寸为 $600 \text{ mm} \times 400 \text{ mm} \times 100 \text{ mm}$,嵌放在长 10 m、宽 2 m 的实验水平步道的中间,顶部与水平步道等高. 为避免鞋对采集数据的影响,每个受试者被要求光脚行走. 在采集过程中,要

求每个受试者在实验步道上按自己习惯的步速行走^[3],采样频率为 200 Hz. 为了消除受试者之间的个体差异,对采集的步态参数进行归一化处理. 就是通过对每个受试者的足-地面间接接触力除以其相应的体重,消除体重的差异;通过将每个受试者的足-地面间接接触时间长度归一化为步态周期中支撑期的百分比,消除不同受试者的速度所造成的误差^[3]. 图 1 为一名受试者左、右两足在垂直方向上的足-地面间接接触力(F_z)随时间变化的曲线图.

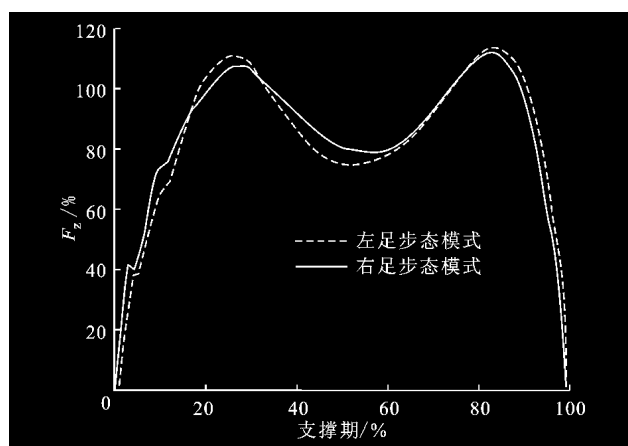


图1 足-地面间接接触力随时间变化的曲线图

2 步态数据分析方法

鉴别左、右两侧的步态模式的相似性,本质上是一个二分类问题,本研究通过采用 SVM 识别左、右两侧的步态模式,判断左、右两侧步态相似性是否存在差异. 假设步态样本训练集 $D = \{(x_i, y_i)\}_{i=1}^l$, 每个原始输入量 $x_i \in R^d$, d 为原始输入量的特征维数, l 为样本数. 首先,通过非线性函数 $\phi: R^d \rightarrow F$ 将原始输入量 x_i 映射到高维特征空间 F , 我们的目的就是在 F 中找到一个最优线性分类面 f , 将左、右两侧的步态模式映射到其相应的类标志空间 $y_i \in \{-1, +1\}$, 在本文中, -1 表示左侧的步态模式, $+1$ 表示右侧的步态模式

$$f: R^d \rightarrow \{-1, +1\} \quad (1)$$

$$x_i \mapsto y_i$$

最优线性分类面 f 可表示为

$$f(x) = \text{sgn}\left(\sum_{i=1}^l \beta_i y_i K(x_i, x) + b\right) \quad (2)$$

式中:核函数 $K(x_i, x) = \phi(x_i) \phi(x)$ 必须满足 Mercer 定理^[11]; b 为训练集的估计偏置; β_i 为最优分类超平面的系数,通过求解下列二次规划问题可得到

$$\left. \begin{aligned} \min W(\beta) &= -\beta^T I + \frac{1}{2} \beta^T D \beta \\ \text{s. t. } \beta^T y &= 0; \beta_i \in [0, C] \end{aligned} \right\} \quad (3)$$

对于线性不可分数据来说,惩罚因子 C 可以作为一个变化范围较大的调节参数,调节 SVM 算法中的最大间隔与训练样本数据的最小训练误差之间的平衡,详细的求解过程见文献[11]。

在支持向量机的学习过程中,不同的核函数得到不同的支持向量机算法,因此核函数的类型及其参数的选择影响支持向量机的泛化性能。目前,如何选择合适的核函数及其参数,还没有统一的标准。对小样本步态数据而言,交叉验证方法是估计预测误差最常用的方法[12-13]。本文把采集的 24 名受试者左、右两侧的步态数据作为样本集(样本总数为 48),每个步态模式为一个 101 维的特征向量。我们采用六折交叉验证方法评估分类器的泛化性能和选取最优的核函数及参数、惩罚因子 C ,把 48 个样本数据平均分成 6 个部分,其中 5 个部分作为样本训练集,构造支持向量机的最优分类超平面,样本训练集包含 20 个左侧的步态模式和 20 个右侧的步态模式。剩余的一个部分作为样本测试集,测试分类器。样本测试集包含 4 个左侧的步态模式和 4 个右侧的步态模式。重复上述过程 6 次,使每一个部分都作为样本测试集,测试分类器,将所获得的 6 次结果平均就可得到最终的分类结果。

为了更准确地评估分类器性能,采用分类正确率(A_c)、敏感率(S_e)和特效率(S_p)作为分类器性能评估指标[12]。正确率用来评估 SVM 正确识别左、右两侧的步态模式的能力,敏感率反映了分类器正确识别左侧的步态模式的能力,特效率反映了分类器正确识别右侧的步态模式的能力。其定义如下

$$A_c = \frac{T_1 + T_2}{T_1 + N_1 + T_2 + N_2} \times 100\% \quad (4)$$

$$S_p = \frac{T_2}{T_2 + N_1} \times 100\% \quad (5)$$

$$S_e = \frac{T_1}{T_1 + N_2} \times 100\% \quad (6)$$

式中: T_1 是分类器正确识别左侧的步态模式的个数; T_2 是分类器正确识别右侧的步态模式的个数; N_1 是错误识别左侧的步态模式的个数; N_2 是错误识别右侧的步态模式的个数。

本研究采用序列极小优化算法(SMO)和下列 3 种核函数测试支持向量机分类器:

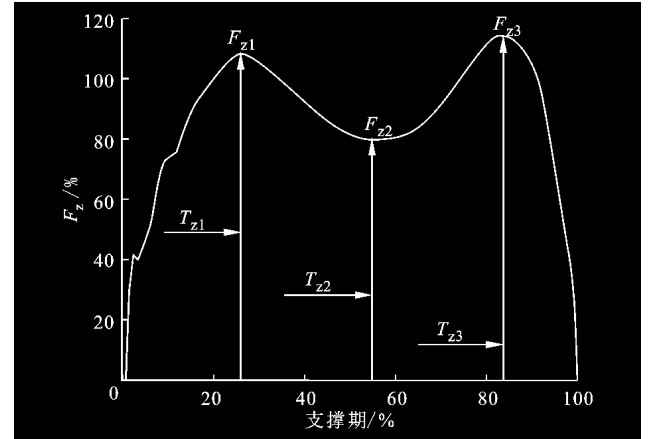
(1) 线性核函数 $K(x_i, x) = x_i x$;

(2) 高斯径向基核函数 $K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right)$, σ 是径向基函数的宽度;

(3) 多项式核函数 $K(x_i, x) = (x_i x + 1)^d$, d 是多项式的阶。

3 实验结果

首先,分别从左、右两足的步态模式中选取步态分析中常用的 6 个步态参数特征(见图 2)。由于这 6 个步态参数特征含有大量与人体运动有关的信息[3,6,14],考虑到人体步态差异性的特点,采用变异系数(C_v)评估所采集的动力学步态数据的变异性[7-8,14]。其定义为数据的标准差与数据的平均值的比值,若 $C_v < 12.5\%$,数据的变异性为可接受的水平[7-8,14]。另外,采用绝对对称指标 A_{sl} 的绝对值判定左、右足不对称性,其定义为 $A_{sl} = \left| \frac{2(Y_R - Y_L)}{Y_R + Y_L} \right| \times 100\%$,其中 Y_R 为右足步态参数的平均值, Y_L 为左足步态参数的平均值。若 $A_{sl} < 10\%$,认为步态对称是在一个正常范围内,确定步态为对称性。本实验采用 T-检验法分析左、右两侧步态的差异性,若检验显著性水平值 $P < 0.05$,则表明左、右两侧步态显著性差异[7-8,14],计算结果见表 1。



F_{z1} : 足跟着地时的最大承受力; T_{z1} : F_{z1} 相应的时刻; F_{z2} : 全足着地时的最小承受力; T_{z2} : F_{z2} 相应的时刻; F_{z3} : 足尖离地时的最大承受力; T_{z3} : F_{z3} 相应的时刻

图 2 选取步态模式中的步态参数特征

从表 1 可以看到,左、右足步态参数的变异系数都小于 12.5%,表明采集的实验数据的变异性在可接受水平范围内。尽管 A_{sl} 的绝对值都小于 10%,左、右两侧步态初步被认为是对称的,但是基于 T-检验法的分析结果,我们可以看到步态参数 F_{z1} 、 F_{z2}

表1 左、右足的步态参数统计分析结果

步态参数	左足		右足		$A_{SI}/\%$	P
	$C_V/\%$	步态参数平均值±标准差/%	$C_V/\%$	步态参数平均值±标准差/%		
F_{z1}	8.77	109.98±9.48	9.23	111.37±8.12	2.23	0.034
T_{z1}	10.23	24.43±2.94	8.84	23.50±1.77	4.74	0.079
F_{z2}	6.78	75.60±6.88	8.21	76.05±7.32	4.09	0.025
T_{z2}	5.76	48.92±5.38	5.35	51.90±3.50	3.27	0.039
F_{z3}	9.71	117.57±6.57	8.39	117.08±7.58	2.57	0.087
T_{z3}	4.98	75.92±2.25	3.58	78.03±2.65	0.81	0.063

和 T_{z2} 的检验显著性水平值 $P < 0.05$, 表明左、右两侧步态存在着一定的差异. 为了更准确地鉴别左、右两足之间的步态差异, 我们采用 SVM 算法鉴别左、右足的步态模式的差异. 为了避免步态模式中有效信息的丢失, 我们直接分析了步态模式中的全部步态参数(步态模式为一个 101 维的特征向量). 表 2 详细地给出了整个样本集 48 个左、右两侧步态模式在 3 种不同核函数的 SVM 上的最优分类结果. 从表 2 可以看到, 3 种不同核函数的 SVM 分类正确率、敏感率、特效率都大于 81%, 能够较好地识别左、右两侧步态模式. 相比较而言, 多项式核的分类性能最优(分类正确率、敏感率、特效率均为 88%), 线性核较差. 因此, SVM 算法通过核函数能够将老年人左、右两侧步态模式映射到高维特征空间, 有效地分离左、右两侧不同的步态模式. 此外, 选取最优的核函数参数、惩罚因子 C , 见表 2.

表2 3种核函数的支持向量机分类测试结果

核函数类型	C	d	σ	$A_c/\%$	$S_e/\%$	$S_p/\%$
多项式核	0.75	3		88	88	88
高斯径向基核	0.55		850	85	88	83
线性核	0.3			81	83	83

4 讨论与结论

众所周知, 步态的对称性反映了人体步态整体特征, 是评估人体步态功能的一个重要指标^[15]. 特别是在人衰老的过程中, 大脑神经中枢系统和生理机能的衰退, 影响着人体下肢左、右两侧的步态, 使其产生变异. 由于人体运动是一个复杂的动态系统, 其过程是非平稳的, 包含着大量步态变异的非线性信息^[14]. 因此, 如何从人体下肢左、右两侧的步态模

式中提取有效信息是评估其相似性差异和对称性的关键所在. 从表 1 的结果可以看到, 一些传统的步态对称性指标仅反映了步态周期中个别时刻的步态参数在对称性中的作用, 忽视了步态周期中任一时刻的步态参数在评估步态对称性中的敏感性^[15], 因此难以全面反映步态参数在对称性中的作用, 限制了及早发现步态的不对称性. 采用 SVM 算法识别左、右两侧的步态模式, 能够较好地解决上述问题. 从表 2 的结果可以看到, 非线性核(多项式核、高斯径向基核)的 SVM 分类性能优于线性核, 说明基于非线性核的 SVM 在映射的高维特征空间中可有效地提取左、右两侧步态模式中的非线性信息, 增加了信息量, 克服了基于线性变换算法造成非线性信息丢失的问题, 提高了步态参数在评估步态对称性中的敏感性, 能有效地鉴别左、右两侧步态模式相似性的差异, 有利于及早发现步态的不对称性.

在 SVM 的学习过程中, 核函数参数、惩罚因子 C 对分类性能影响较大. 目前, 选取核函数参数也没有统一的标准, 由于这些参数在学习过程中相互影响, 因此必须通过反复实验仔细地挑选. 另外, 数据结构特征也影响着 SVM 的泛化能力^[11], 若数据结构特征包含更多的两类分离信息, 则能极大地提高 SVM 的泛化能力. 为了进一步提高 SVM 识别左、右两侧步态模式的泛化能力, 下一步需要寻找有效的数据结构特征提取方法, 提取左、右两侧步态模式更具有代表性的步态数据特征.

总之, SVM 克服了以往传统算法在评估步态对称性时存在的不足, 能够较准确地识别老年人左、右两侧步态模式, 及鉴别人体下肢左、右两侧的步态相似性的差异, 有利于评估其步态的不对称性, 有助于预防老年人跌到和老年疾病的发生, 提高老年人生活质量, 并且 SVM 也可成为临床上早期诊断步态疾病和评估治疗效果的有效工具.

参考文献:

- [1] Francois P, Helene C, Rejean H, et al. Gait in the elderly [J]. *Gait and Posture*, 1997, 5(2):128-135.
- [2] 赵芳,周兴龙. 老年人站立及行走稳定性的生物力学研究[J]. *北京体育大学学报*, 2003, 26(2):188-191.
Zhao Fang, Zhou Xinglong. Biomechanical study on the standing and walking stability of the aged: a review [J]. *Journal of Beijing University of Physical Education*, 2003, 26(2):188-191.
- [3] 朱晓兰,赵芳. 老年人步态特征的分析及其评价系统的初步建立[D]. 北京:北京体育大学运动人体科学学院, 2002.
- [4] 肖美痕,赵芳. 老年人常速行走与快速行走特征的生物力学分析[D]. 北京:北京体育大学运动人体科学学院, 2004.
- [5] Simon S R. Quantification of human motion:gait analysis—benefits and limitations to its application to clinical problems[J]. *Journal of Biomechanics*, 2004, 37(2): 1869-1880.
- [6] Heydar S, Paul A, Francois P, et al. Symmetry and limb dominance in able-bodied gait: a review [J]. *Gait and Posture*, 2000, 12(1): 34-45.
- [7] Giannis G, Vasilios B. Time and frequency domain analysis of ground reaction forces during walking: an investigation of variability and symmetry [J]. *Gait and Posture*, 1997, 5(3): 189-197.
- [8] Sadeghi H, Allard P, Duhaime M. Functional gait asymmetry in able-bodied subjects [J]. *Human Movement Science*, 1997, 16(2):243-258.
- [9] Sadeghi H. Local or global asymmetry in gait of people without impairments [J]. *Gait and Posture*, 2003, 17(3): 197-204.
- [10] Chau T. A review of analytical techniques for gait data: part1 [J]. *Gait and Posture*, 2001, 13(1): 49-66.
- [11] Vapnik V N. Statistical learning theory [M]. New York: John Wiley and Sons Inc. , 1998.
- [12] Begg R, Kamruzzaman J. A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data [J]. *Journal of Biomechanics*, 2005, 38(3): 401-408.
- [13] Ding C H Q, Dubchak I. Multi-class protein fold recognition using support vector machines and neural networks[J]. *Bioinformatics*, 2001, 17(4):349-358.
- [14] 耿春亚,马军,郭忠武,等. 关于正常青年人步态垂直方向力的检测与统计分析[J]. *航天医学与医学工程*, 2003, 16(5):364-367.
Geng Chunya, Ma Jun, Guo Zhongwu, et al. Measurement and statistical analysis of the vertical force in normal young people's gait [J]. *Space Medicine & Medical Engineering*, 2003, 16(5):364-367.
- [15] 杨年峰,金德闻,黄昌华. 人体运动协调规律及其参数描述[D]. 北京:清华大学机械工程学院, 2001.

(编辑 杜秀杰)