

较小样本动态声发射信号多元统计分析技术*

陈宏志^{1,2}, 芦永明¹, 王丽娜¹, Lik-Kwan Shark², John Goodacre^{2,3}

(1. 冶金自动化研究设计院混合流程工业自动化系统与装备技术国家重点实验室 北京, 100071)

(2. Applied Digital Signal and Image Processing Research Centre, University of
Central Lancashire Preston, United Kingdom, PR1, 2HE)

(3. School of Health and Medicine, Lancaster University Lancaster, United Kingdom, LA1, 4YD)

摘要 利用多传感信息集成系统,以两组平均年龄对应的受试对象往复运动过程中获取的动态声发射信号和角度信号为对象,研究了适用于较小样本的动态声发射信号多元统计分析技术。通过同步记录的角度信号,将往复运动分解为若干个独立运动周期和运动过程;利用累计概率分布,选取具备较显著差异的特征;结合多元统计技术,减小数据量,建立动态声发射信号的可视化模型,证实了使用较小样本声发射信号实现膝盖骨关节诊断的可行性。

关键词 较小样本; 动态; 声发射; 多元统计; 骨关节炎

中图分类号 TN911.72; R331

引言

作为人体中最大的关节和最为复杂的结构之一,膝关节承担着承载数倍于身体重量的任务。随着年龄增长和不同人群的使用情况,膝关节的功能将发生改变^[1]。骨关节炎(OA)是一种慢性关节炎,多发于中老年人群,对他们的生活质量已造成严重影响^[2]。由于目前难以使用准确的物理或数学模型描述其成因和发展过程,在临床界除全膝替换术(TKR)外,还没有更有效的治疗方法^[3]。当前,OA诊断主要使用的是超声成像、X射线以及核磁共振的方式,结合临床经验进行判断。然而,对临床医学经验要求高、潜在放射性危害、高昂的诊断费用以及便携性差已被认为是此类技术的瓶颈^[4]。

以往的研究证实,膝关节在屈伸运动过程中产生振动,基于振动所产生的声信号(包括可闻声和声发射)已被广泛研究,目的在于研发便携、快速、简单、低成本、安全及具备普适性的膝骨关节炎诊断技术^[4-5]。可闻声研究方面,通过研究分析加速度计收集到的频率在20kHz以下的膝关节振动信号,结合统计和智能建模技术,证实了可闻声作为膝骨关节状态生物标识的可能性^[6]。膝关节在振动时,可同时产生频率范围高于20kHz的声发射(AE)信号,连续

振动产生的AE信号结合相关角度信号和角速度信号,进行分解重组后建立的多元统计模型可用于分析和诊断不同年龄层的膝盖骨关节炎^[4,7]以及骨关节炎随年龄的发展和劣化趋势^[8-9]。然而,之前的研究成果在信号采集阶段需要参与者执行多次膝关节的往复屈伸运动(10~15次),以获取尽可能多的AE信号用于分析。这除了增大数据采集的执行难度和建立模型时的工作量外,部分OA患者无法自主进行多次此类运动。

笔者综合项目组之前研究成果的利弊,研究了降低信号数量的同时获取相似膝骨关节炎AE信号分离结果的信号处理方法。

1 参与者

本研究项目共使用了19名志愿者的实验数据,其中包含11名年龄在61岁以上的老年健康志愿者(H组),其平均年龄 $MA=71.27$ 岁,标准差 $SD=6.99$ 岁,全部为来自英国Lancashire地区的志愿者;8名年龄在61岁以上的老年骨关节炎患者(OA组)($MA=69.50$ 岁, $SD=6.39$ 岁),全部为在英国Blackpool Fylde&Wyre医院接受物理治疗的骨关节炎患者。两组参与者在年龄上匹配,实验过程在英国NHS国民健康系统监督下进行。在参与测试前,

* 英国 Arthritis Research Campaign 支持计划资助项目(17542);国家重点基础研究发展计划(“九七三”计划)资助项目(2010CB736005)

收稿日期:2012-05-25;修改稿收到日期:2012-09-20

所有志愿者分别通过临床诊断方式确定关节健康状况。为保证测试结果的精确性,根据临床诊断的结果,AE 信号分别从13条H组志愿者和12条OA组志愿者的膝关节中获取,选取的规则如下:

1) H组中,临床诊断为健康且功能正常的膝关节,共有2名参与者的2条膝盖及另外9名参与者的1条膝盖被选取;

2) OA组中,诊断为骨关节炎的膝盖,其中包括4名患者的双膝及其余患者的其中1条膝关节。

2 数据采集

实验中使用JAAS数据采集系统采集膝关节的AE信号和角度信号^[5]。如图1所示,JAAS系统是一套集成了AE传感装置和电子角度测量装置的多源融合信息采集系统,其AE传感装置包括两个压电陶瓷传感器,在图中分别记做S1和S2。为保证信号传输的强度,为每个传感器配置一个40dB增益的前置放大器。电子角度测量装置包括两个电子角度计和对应信号放大单元。两套系统通过改进过的I/O接口同时连接到数采板上,并用PCI接口实现与计算机之间的通信,通过PCI接口接收到的同步电平信号实现系统间的无缝同步。信号获取实验中,AE传感器固定在最接近于膝盖负重骨面的稳定位置,电子角度计分别固定在大小腿两侧。声发射和角度信号通过参与者进行10~15次坐-立-坐(STS)的方式产生,具体的信号产生和采集方案见文献^[8]。

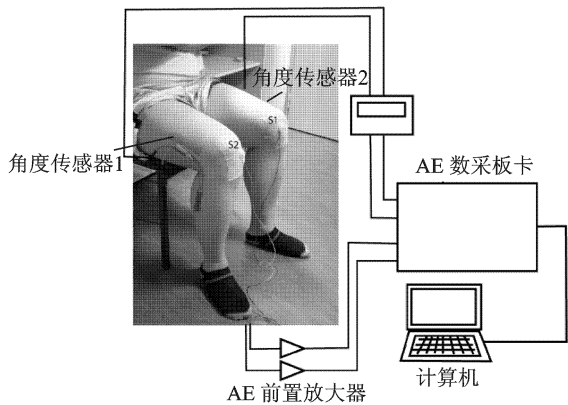
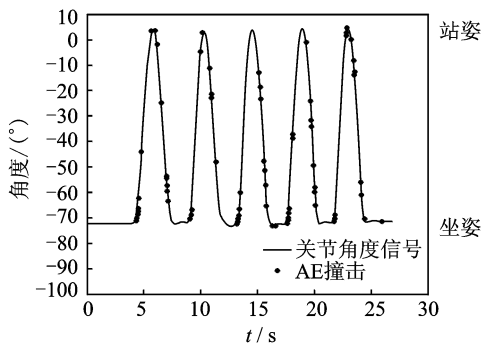


图1 JAAS数据采集系统框图

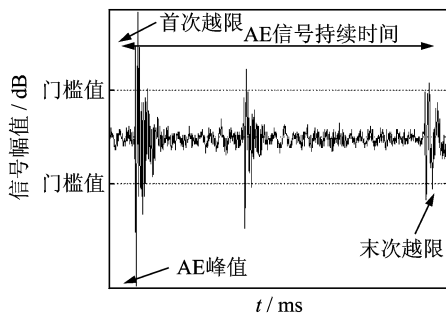
3 信号预处理

由于AE信号是通过往复多组的STS动作来产生,将往复的STS动作分解成若干个独立的运动过程,以尽可能多地从中提取有意义的信息用于后续

分析。每次STS运动的起止时间可通过求解角度信号的1次差分来实现,鉴于参与者的每个动作周期间耗时的高相似度(最大SD值为0.73s)^[9],每个STS周期内的时差未被考虑。图2为一个5次STS动作所产生的角度信号和对应AE信号,图2(a)中的实线表示的是STS过程中采集到的角度信号,嵌入曲线中的每一个点分别表示一个在此过程中获取的如图2(b)所示的AE撞击。由于部分OA患者难以连续进行多次STS运动,为简化数据获取过程和获取难度,提高诊断速度,同时避免由于连续运动造成的膝关节结构疲劳等原因影响结果的精确性,笔者只选用备选参与者执行前5次STS运动所获取的AE信号和对应角度信号为研究对象。



(a) 膝关节角度和对应AE信号示意图



(b) 膝关节AE信号实例

图2 膝关节角度与AE信号示例

图3为一组5个STS动作的角度信号叠加示意图。由图3可见,每个STS运动周期可被再次分解为

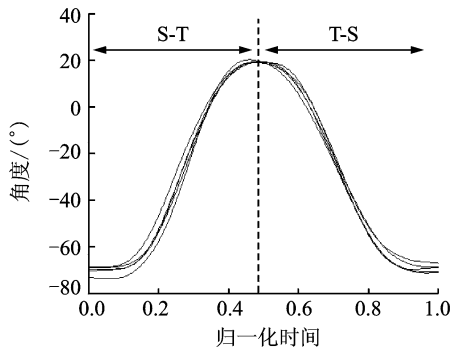
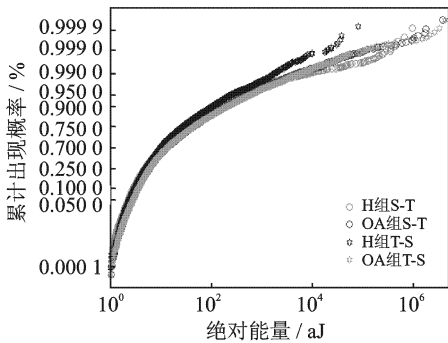


图3 分解后的STS角度信号

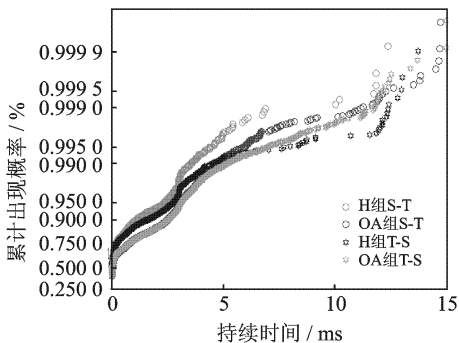
一个从坐到站(S-T)的运动过程和一个从站到坐(T-S)的运动过程,这为建立膝关节AE信号与潜在的生物机械学间的联系提供了纽带^[10]。通过提取和分析动态AE信号的时频域特征,以发掘具有最显著统计差异的AE特征用于区别H组和OA组信号。对5次STS运动所产生的AE信号的特征分析显示,在前5个周期中,信号时、频域具备最显著统计差异的AE信号特征是持续时间、绝对能量和峰值频率。持续时间是AE信号首次越过阈值和最后一次越过阈值间的间隔;绝对能量为在持续时间内信号能量的绝对值(持续时间内信号幅度的积分绝对值);峰值频率是信号峰值所在的FFT频段。

4 信号特征的统计分析

图4分别为绝对能量和持续时间的累计概率分布图,鉴于以上两种特征所呈现的不同统计分布(持续时间接近于指数分布而绝对能量接近于正态分布),图中纵坐标分别采用指数和正态分布尺度来表示绝对能量和持续时间的累计概率分布。此类统计分布表示法中,曲线重合度越低,表明参数的差异性越高。由图5(a)可见,虽然两组受试者的4条绝对能量统计分布曲线具有很高的重合度,但是曲线的尾部显著不同,OA组信号的绝对能量无论在S-T过程还是在T-S过程都比H组信号稍高,超过 10^6 aJ,



(a) AE信号持续时间的累计概率分布图



(b) AE信号绝对能量的累计概率分布图

图4 AE信号特征的累计概率分布图

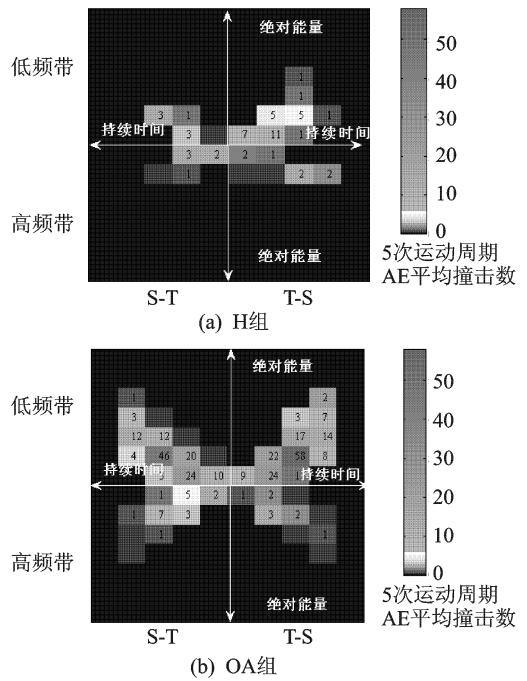


图5 参与者典型AE特征模型的图像表示

而从H组中采集的AE信号能量均不超过 10^6 aJ,特别是在S-T过程,其最大能量仅有约 10^5 aJ。与绝对能量的最大值类似,OA组参与者产生信号的最大持续时间在两个运动过程中也显著高于H组,达到15 ms,而H组在两个运动过程中AE信号的最大持续时间分别为12 ms和13 ms。此外,参与者在连续STS运动过程中所产生AE信号的峰值频率以其传感器中心频率150kHz为界分为低频段和高频段,呈现两种不同的统计分布,其原理及相关具体讨论见文献^[4]。

5 AE信号的多变量统计分析

虽然将连续运动周期缩减为5次,但是由于信号数量仍然处于较大水平(最大516撞击/周期),并且不同参与者所产生的信号总量不一致,直接分析完整波形不仅工作量大,且对获得期望结果的帮助有限^[4]。为了简化分析过程,实现对AE信号的快速精准分析,并获取期望的结果,笔者介绍了一种基于多变量统计的膝关节AE特征模型^[4]。根据上述分析,在前5次STS运动过程中,受试对象的绝对能量和持续时间在S-T和T-S的运动过程中具有最为显著的差异性,此外,H组和OA组AE信号的峰值频率可以中心频率值划分为高、低频带。因此,根据以上3种特征,结合角度信号建立膝关节AE多变量统计模型,模型的表达式如下

$$H = \frac{1}{n} \sum_{i=1}^n (X_i(E_j, D_m, P_{l/h}), \psi) \quad (1)$$

其中: n 为 STS 运动周期数; X_i 为第 i 次 STS 运动过程中特征的不同值域, 用绝对能量、持续时间和峰值频率以 $X_i(E_j \cap D_m \cap P_{l/h})$ 方式进行定义; ψ 为每一次 STS 运动的 S-T 和 T-S 过程; $E_j, D_m, P_{l/h}$ 分别为绝对能量、持续时间和峰值频率。

每个特征值的至于范围定义为

$$10^j \leq E_j \leq 10^{j+1} \quad (j = 0, 1, 2, 3) \quad (2)$$

$$10^m \leq D_m \leq 10^{m+1} \quad (m = 0, 1, \dots, 5) \quad (3)$$

$$\begin{cases} P_l \in \{P_l | P_l \leq 150\} \\ P_h \in \{P_h > 150\} \end{cases} \quad (4)$$

此外, 绝对能量和持续时间的最大值限定条件分别为 $E_j > 10^6 \mu\text{s}$ 和 $D_m > 10^4 \text{aJ}$ 。

根据式(1)~式(4)以及相关前述限定条件, 可将每个参与者的 AE 信号重组成基于信号出现次数的多变量统计的模型。为产生直观视觉效果, 方便诊断和分析, 基于所获取特征模型, 构建了膝关节 AE 多变量统计模型的图像表示, 如图5所示。图像可用高低峰值频率和运动过程分解为4个类似于二维直角坐标系4象限的4个区域, 以原点为中心, 信号持续时间和绝对能量值分别向横、纵向增长, 以颜色表示依据式(1)~式(4)所设定条件返回的 AE 信号数量。比较图5(a)和图5(b), H 组产生的 AE 信号在数量与信号类型上均明显低于 OA 组, 主要为短时、低能量 AE 撞击, 最大撞击数为 11, 出现在 T-S 过程中的低频带; 从 OA 组获取的 AE 信号从种类和信号数量上均显著高于 H 组, 出现了一定数量的大能量、长时 AE 撞击, 最大撞击数为 58, 出现在 T-S 过程中的低频带, 这主要因为 OA 组参与者的膝关节在连贯性 STS 运动过程中由于结构异常造成的较大振动^[8]。

由于 AE 信号的多变量统计模型的维度较高(140 维), 不利于对所有参与者的 AE 模型进行分离和分析, 本研究使用主成分分析法(PCA)^[11]对模型进行适当降维处理, 并在没有先验知识的情况下实现对两组参与者 AE 信号的分离。PCA 是一种较为常用的数据约简和实现盲源分离的技术, 与其相近的同类技术近年来已在复杂结构、设施的特征分析上有较成功的应用^[12-14]。江涌等^[12]通过试验证实了基于两次去相关的盲分离法可用于分离出频带上相互重叠的瞬时振动信号; 崔立林等^[13]通过基于直推式置信机(TCM)与 K 近邻算法(KNN)相结合的改进 TCM-KNN 算法, 实现了水下航行器机械噪声源在较小样本条件下的识别和分类; 艾延廷等^[14]通过

基于独立成分分析(ICA)的最大信噪比分离算法与基于 FFT 的分离算法的结合, 实现了对某双转子航空发动机不同激振源振动信号的提取和分离。

PCA 的中心思想是降维, 通过正交投影变换的方式将研究对象间最显著的特征投影至较低维度的平面上, 以实现研究对象间关系的可视化分析。若分别以 x_i, y_j 表示第 i 个健康组和第 j 个罹患骨关节炎组受试对象的 AE 多元统计模型, 则 H 组和 OA 组受试对象的特征模型可分别通过行层叠的方式组成健康组和患病组特征矩阵, 即 $H = [\dots x_i \dots]^T$ 和 $OA = [\dots y_j \dots]^T$, 通过行层叠方式合并矩阵 H 和 OA , 可得到受试对象的 AE 特征矩阵 $F = \begin{bmatrix} H \\ OA \end{bmatrix}^T$ 。

下面以受试对象的多元统计特征矩阵 F 为例, 介绍 PCA 的实现过程。为保证 F 中的每列值拥有同样的权重, 首先通过式(5)对 F 求列修正矩阵 F_M

$$F_M = (F - \overline{F}_j)^T \quad (5)$$

其中: \overline{F}_j 为矩阵 F 的列平均值。

通过奇异值分解法(SVD)求出 F_M 的主元向量和每个主元向量对应的方差值

$$F_M = USV^T \quad (6)$$

其中: V^T 为 F 的主成分矩阵; V^T 的每一列可构成一个维度的主元平面; S 为一主对角线包含了投影到每一维主元平面的方差值的对角矩阵; u 为荷载矩阵, 满足 $F_M V^T = us$

F 在前 n 个主元面的投影可通过式(7)的线性变换来实现

$$K = F_M V_n^T \quad (7)$$

其中: V_n^T 为 V^T 的前 n 个主元 ($n \ll V_n^T$ 的最大维度); K 为 F 在前 n 个主元平面的投影。

多变量 AE 模型的三维空间聚类如图6所示。通过 PCA 投影, 在前三主元面共保留了 91.13% 的原始信息量。左侧的聚类表示的是 H 组, 右侧的聚类表示的是 OA 组, 每个标记旁的数字代表受试者年龄。由图6可见, H 和 OA 组呈现出两个清晰的不重叠聚

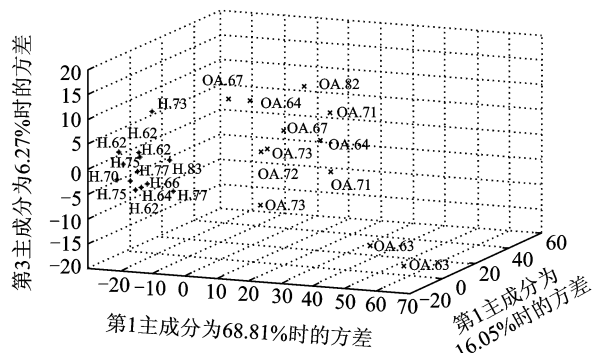


图6 通过PCA产生的参与者膝关节AE特征模型的三维空间聚类图

类分布在原点两侧。此外,和OA组的AE信号相比,由于较低的模型间方差和产生信号的较高相似度,H组模型在PCA三维平面的投影聚类较OA组更为紧凑,而OA组聚类则显示了更高的松散度。这组结果证实了使用笔者所述的多元统计模型,结合PCA,基于较小样本集诊断膝关节炎的可能性。

6 结束语

笔者介绍了一种基于较小样本的膝关节动态AE多元统计分析方法,证实了利用多变量统计方法分析较小样本膝关节AE信号的可行性,这将有助于在后续研究和应用中有效降低参与者的STS运动周期数,减小实验的实施难度,研发快速、精确、便携、适合临床诊断和家庭康复使用的膝关节动态完整性分析与诊断系统。后续研究将加入不同年龄组受试者的AE信号,结合相关统计和智能建模技术,研究最优STS运动周期用于获取AE信号,探讨建立基于较小样本的膝关节AE信号多元统计模型在不同年龄和不同类型病患上的延伸应用,以及基于空间聚类间距离的AE信号模式识别技术。

参 考 文 献

[1] Gray H. Anatomy of the human body [M]. 20 th ed. New York; Bartleby.com,2000:1-32.

[2] Peat G, McCartney R, Croft P. Knee pain and osteoarthritis in older adults; a review of community burden and current use of primary health care [J]. Annals of Rheumatology Disease, 2001, 60: 91-97.

[3] Felson D T, Lawrence R C, Dieppe P A, et al. Osteoarthritis the disease and its prevalence and impact [J]. Annals of International Medicine, 2000,133: 637-639.

[4] Mascaro B, Prior J A, Shark L K, et al. Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints [J]. Medical Engineering and Physics, 2009, 31: 1013-1022.

[5] Prior J A, Mascaro B, Shark L K, et al. Analysis of high frequency acoustic emission signals as a new approach for assessing knee osteoarthritis [J]. Annals of Rheumatology Disease, 2010,69: 929-930.

[6] Mu Tingting, Nandi A K, Rangayyan R M. Screening of knee-joint vibroarthrographic signals using the strict 2-surface proximal classifier and genetic algorithm [J]. Computers in Biology and Medicine, 2008,

38: 1103-1111.

- [7] Shark L K, Chen Hongzhi, Goodacre J. Discovering differences in acoustic emission between healthy and osteoarthritic knees using a four-phase model of sit-stand-sit movements [J]. The Open Medical Informatic Journal, 2010,4: 116-124.
- [8] Shark L K, Chen Hongzhi, Goodacre J. Knee acoustic emission; a potential biomarker for quantitative assessment of knee joint ageing and degeneration [J]. Medical Engineering and Physics, 2011,33: 534-545.
- [9] Chen Hongzhi. Discovery of acoustic emission based biomarker for quantitative assessment of knee joint ageing and degeneration [D]. Preston: University of Central Lancashire, 2011.
- [10] Kerr K M, White J A, Barr D A, et al. Analysis of the sit-stand-sit movement cycle in normal subjects [J]. Clinical Biomechanics, 1997,12:236-245.
- [11] Jolliffe I T. Principal component analysis [M]. New York:Springer-Verlag, 1986:10-61.
- [12] 江涌,章林柯,何琳,等. 两次去相关用于振动信号盲分离[J]. 振动、测试与诊断,2011, 31: 241-245. Jiang Yong, Zhang Linke, He Lin, et al. Blind separation of vibration signals using double decorrelation [J]. Journal of Vibration, Measurement & Diagnosis, 2011,31: 241-245. (in Chinese)
- [13] 崔立林,朱海潮,章林柯,等. 小样本条件下的机械噪声源识别方法[J]. 振动、测试与诊断, 2010,30, 557-560. Cui Lilin, Zhu Haichao, Zhang Linke, et al. Identification of mechanical noise sources on small sample [J]. Journal of Vibration, Measurement & Diagnosis, 2010, 30: 557-560. (in Chinese)
- [14] 艾延廷,费成巍,张凤玲,等. ICA在航空发动机振动信号盲源分离中的应用[J]. 振动、测试与诊断,2010, 30,671-674. Ai Yanting, Fei Chengwei, Zhang Fengling, et al. Blind source separation for aero-engines vibration signal by independent component analysis [J]. Journal of Vibration, Measurement & Diagnosis, 2010,30: 671-674. (in Chinese)



第一作者简介:陈宏志,男,1986年3月生,博士。主要研究方向为数字信号和图像处理技术在智能诊断和视情维护中的应用。曾发表《Knee acoustic emission:a potential biomarker for quantitative assessment of knee joint ageing and degeneration》(《Medical Engineering and Physics》)2011,Vol. 33)等论文。
E-mail:Shadow_c_3186@msn.com