

doi: 10.12068/j.issn.1005-3026.2018.01.006

基于声源定位的头部姿态检测系统研究与实现

康 雁, 宁国琛

(东北大学 中荷生物医学与信息工程学院, 辽宁 沈阳 110169)

摘 要: 研究了声源定位技术在鼾声检测上的应用,主要研究鼾声的声学特性、声源定位技术原理,并利用声源定位技术对人体打鼾时头部姿态进行识别.采用广义互相关函数法对多路声音采集系统采集到的鼾声信号进行延迟估计计算,根据得到的相对时延信息进行声音定位,结合定位信息和朝向信息推断出当前受试者打鼾时头部姿态.采用数据采集卡进行高速信号采集,使用虚拟仪器 Labview 进行信号处理和算法实现,实验中受试者以不同头部姿态在该系统监测下模拟打鼾,通过该系统检测头部姿态,综合判断正确率达到80%.

关 键 词: 鼾声检测;声源定位;头部姿态识别;信号处理;麦克风阵列

中图分类号: TP 274.1 **文献标志码:** A **文章编号:** 1005-3026(2018)01-0026-06

Research and Implementation of Snoring Head Posture Detecting System Based on Sound Source Localization Technology

KANG Yan, NING Guo-chen

(School of Sino-Dutch Biomedical and Information Engineering, Northeastern University, Shenyang 110169, China. Corresponding author: KANG Yan, E-mail: kangyan@bmie.neu.edu.cn)

Abstract: Sound source localization technology in the application of snoring detecting was studied, mainly including the aspects of acoustic characteristics of snoring as well as sound source localization technology principle and was used to detect the head posture of snoring human. The generalized cross-correlation (GCC) function algorithm was used to calculate relative time-delaying estimation of snoring signals from mulita-channel sound acquisition system and sound position was obtained based on the relative delay information. Face orientation and head location information were combined to deduce the current head posture. The head posture detecting system contained high speed signal acquisition which was compiled by using DAQ. Besides, algorithm and signal processing was compiled by Labview. In the experiment, the subject simulated snoring under the detecting of system and the accuracy of the results is up to 80%.

Key words: snoring detecting; sound source localization; head posture recognition; signal processing; microphone array

声源定位技术有着广泛的应用场景,早在 20 世纪七八十年代,就已经开始将麦克风阵列应用于语音信号处理的研究,进入 21 世纪后,基于麦克风阵列的声音信号处理已经成为众多领域研究热点^[1-2].打鼾是人体睡眠中常见的慢性症状之一,国外临床统计显示,全球每天约有 3 000 人的死亡与阻塞性睡眠呼吸暂停低通气综合征 (obstructive sleep apnea-hypopnea syndrome, OSAHS)有关.近 10 年来,应用计算流体动力学 (CFD)方法对呼吸道流场进行数值模拟的研究与实验有效结合,定量分析呼吸道解剖结构与功能的交互关系,其中以鼻腔、气管-支气管单个器官的报道多见^[3-5].从声学角度来讲,鼾声是由软腭、咽腔壁、会厌、舌等咽部组织结构的振动产生的,气流受限一般是产生振动的前提,而气流受限又与各种原因引起的上气道截面积减小密切相

关^[6]. 本研究中的头部姿态检测系统利用基于麦克风阵列的声源定位技术进行非接触检测. 通过结合受试者身体和环境特征建立鼾声传播模型, 利用广义互相关 (generalized cross - correlation, GCC) 函数算法得到麦克风阵列接收到鼾声信号的相对时延并计算得到发声坐标, 结合通过鼾声幅值特征得到的打鼾时头部朝向判断头部姿态. 该方法实现了长时间的非接触式头部姿态识别, 创新性将声源定位技术应用在人体信号检测上, 对未来研究鼾声与睡姿关系提供了模型与实验基础.

1 打鼾时头部姿态检测系统介绍

本研究利用声源定位技术, 结合人体与环境模型对鼾声进行检测, 并通过计算得到打鼾时受试者的头部姿态. 系统以 4 路相同的麦克风声音采集电路作为声音传感器, 以 1 m × 1 m 矩阵方式放置于受试者头部四周采集鼾声. 信号采集通过使用 NI - DAQ 数据采集卡进行多路、高速、同步的 AD 转换, 转换后的数据传入上位机中. 上位机使用 Labview 对采集到的数据进行信号处理和算法分析, 使用广义互相关 (GCC) 函数法计算多路信号相对时延, 通过将时间延迟参数代入鼾声传播模型并结合头部朝向得到头部姿态数据. 图 1 为本系统流程图.

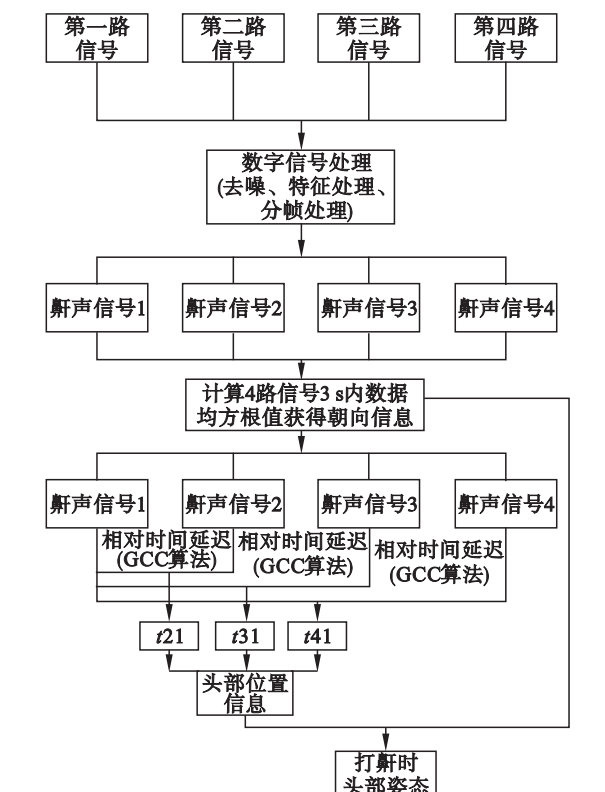


图 1 头部姿态检测系统流程图
Fig. 1 Flow chart of head posture detecting system

图 2 为信号采集示意图, 4 路麦克风传感器以阵列形式放置于受试者枕头四周, 枕头的柔软特性可以防止声音通过震动传播带来误差.

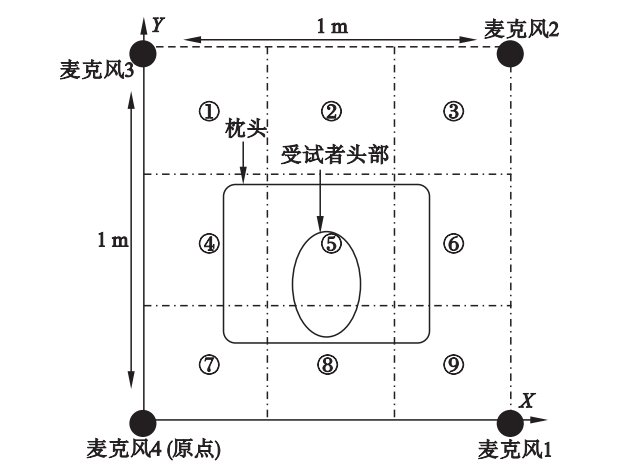


图 2 鼾声监测示意图
Fig. 2 Diagram of snoring monitoring

2 原理与实现

2.1 鼾声声源模型的构建

声源信号的空间基本模型是麦克风阵列系统处理中最主要的部分, 该空间模型携带声源信号距麦克风阵列的距离和相位信息^[7]. 因为鼾声频率在 350 ~ 400 Hz 之间, 所以在该麦克风阵列中认为鼾声信号为近场模型, 且鼾声主要由于上气道存在多面狭窄^[4]造成, 人体头部本身对鼾声的传播起到了很大的阻碍和反射作用. 因此通过鼾声对头部姿态进行识别时需要考虑麦克风接收到的鼾声时间和幅值均不同, 需要综合考虑幅值衰减与相位差两个因素, 并同时作为头部朝向信息和头部位置信息的参数.

由麦克风采集到的鼾声信号理想模型是对真实物理环境的简化, 在该模型中第 i 个麦克风所接受到的信号表示为

$$x_i(t) = a_i m_i(t) s(t - \tau_i) + n_i(t) \quad (1)$$

其中: $m_i(t)$ 表示第 i 个麦克风接收到的混响信号; $s(t - \tau_i)$ 为声源信号, 选用 IMAGE 模型^[7]产生该单位冲激响应; a_i 为鼾声声源到第 i 个麦克风之间的衰减系数; $n_i(t)$ 为环境及系统噪声.

鼾声在时域上的特征与语音相似, 在一段时间内, 单纯打鼾者的鼾声振幅基本一致, 间隔均匀, 每一次鼾声由多个重复出现的复合波组成, 波形似锯齿样, 振幅、间隔大致相仿^[4]. 根据鼾声声源模型对采集到的原始信号进行信号处理, 包括去噪、分帧处理、加重、分频处理等. 将 1 s 长度且采样频率为 44 kHz 的鼾声按照鼾声模型仿真得

到的鼾声信号如图 3 所示。

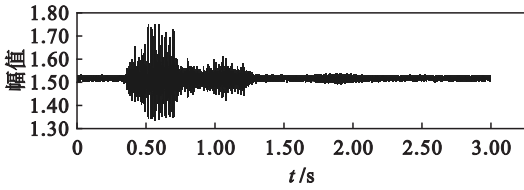


图 3 仿真鼾声信号

Fig. 3 Emulation of snoring signal

2.2 麦克风阵列结构和定位算法概述

麦克风阵列结构是由一定数量的麦克风组成特性拓扑结构的麦克风组^[8]。其中麦克风阵列中各个麦克风阵元的摆放方式和结构决定了声源定位的特性。根据鼾声传播模型和已有麦克风几何结构,鼾声定位算法采用基于到达时间差的定位算法(TDOA)^[9],该算法运算量小,计算方法简单,满足实时定位需求。其主要定位过程分为两个步骤:首先根据各个麦克风采集到的鼾声信号计算得到声音到达时间差,然后将得到的相对时间差代入麦克风几何模型得到声源位置。根据模型得到距离方程:

$$\left. \begin{aligned} (x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2 &= (r+c\tau_1)^2, \\ (x-x_2)^2 + (y-y_2)^2 + (z-z_2)^2 &= (r+c\tau_2)^2, \\ (x-x_3)^2 + (y-y_3)^2 + (z-z_3)^2 &= (r+c\tau_3)^2, \\ (x-x_4)^2 + (y-y_4)^2 + (z-z_4)^2 &= (r+c\tau_4)^2. \end{aligned} \right\} \quad (2)$$

由于人体睡眠时头部高度位置不会有大幅度变化,为了化简计算过程,只对二维平面进行定位计算,以 4 号麦克风为原点建立声源与各麦克风之间的距离方程:

$$\left. \begin{aligned} (x-x_1)^2 + (y-y_1)^2 &= (r+c\tau_1)^2, \\ (x-x_2)^2 + (y-y_2)^2 &= (r+c\tau_2)^2, \\ (x-x_3)^2 + (y-y_3)^2 &= (r+c\tau_3)^2, \\ x^2 + y^2 &= r^2. \end{aligned} \right\} \quad (3)$$

其中:\$(x,y,z)\$为鼾声声源位置坐标;\$(x_i,y_i,z_i)\$,\$i=2,3,4\$为各个麦克风的坐标位置;\$\tau_i\$, \$i=2,3,4\$为声源到达第 1 号麦克风和第 \$i\$ 号麦克风之间的时间差;\$r\$为声源至原点的距离。根据本文麦克风阵列模型确定 \$x_i\$ 值,化简方程组得到声源坐标公式:

$$\left. \begin{aligned} x &= \frac{1-2rc\tau_1-c^2\tau_1^2}{2}, \\ y &= \frac{1-2rc\tau_3-c^2\tau_3^2}{2}, \\ r &= c \frac{\tau_2^2-\tau_3^2-\tau_1^2}{2(\tau_1-\tau_2+\tau_3)}. \end{aligned} \right\} \quad (4)$$

根据鼾声发声模型的声学特性,将录好的鼾声通过加入混响、噪声以及头部阻碍造成的衰减系数进行信号仿真。

2.3 鼾声时延估计算法与仿真

影响实验估计精度和稳定性的关键因素主要包含估计的精度和稳定性。好的时延估计算法应该在信号中存在混响和较低信噪比的情况下,能较精确地估计出信号间相对时延,并且有较快的运算速度和较少的运算量。

由于人体打鼾时头部姿态并不会在短时间内发生较大变化,较长的时间窗可以在相同的采样频率和其他条件下得到更高的精度,所以本文采用互功率谱相位(CSP)法^[10]计算信号相对时延。根据定义,两路信号互相关函数的傅里叶变换为互功率谱函数。广义互相关函数原理如下:根据麦克风矩阵的信号模型,任意两个传声器 \$m_i, m_j\$ 采集信号 \$x_i(t), x_j(t)\$ 之间的相关函数为

$$R_{x_{p_j}}(\tau) = E\{x_i(t)x_j(t+\tau)\} = \alpha_i\alpha_jR_{ss}(t-\tau_{ij}) + R_{n_{p_j}}(\tau). \quad (5)$$

\$R_{ss}(t-\tau_{ij})\$ 为信号的自相关序列。因为基本互相关函数对噪声、声音混响及观测数据长度影响较大,需要通过加权函数对噪声大的频带进行抑制,锐化信号峰值。之后在信号频域内对计算得到的互功率谱函数 \$G_{ln}(\omega)\$ 加权,以抑制噪声和混响。处理后的信号通过傅里叶反变换回到时域,得到的互相关函数就是广义互相关函数:

$$R_{ln}(t) = \int_0^\pi \Psi_{ln}(\omega) X_1(\omega) X_n^*(\omega) e^{-j\omega t} d\omega. \quad (6)$$

其中,\$X_1(\omega), X_n(\omega)\$ 为两个传声器信号 \$x_1(t), x_n(t)\$ 的傅里叶变换,则 \$X_1(\omega) X_n^*(\omega)\$ 为互功率谱。为了进一步锐化相关函数的峰值,广义互相关函数法中对互功率谱函数进行加权,加权函数为

$$\Psi_{ln}(\omega) = \left| \frac{1}{G_{x_1x_j}(\omega)} \right|. \quad (7)$$

由式(5)知,CSP 法对功率谱函数进行白化(whitening)处理^[11-12],抑制了睡眠环境带来的混响以及头部本身的反射和衰减干扰,通过对功率谱的归一化,将信号的幅度信息排除,保留信号相位特性,最大程度抑制混响造成的干扰^[13]。将 2.1 节中的仿真鼾声信号以 4 号麦克风为原点坐标, \$(0.3, 0.5)\$ 为点声源,以此坐标计算出理论相对时延。将理论相对时延分别对应加入 4 路麦克风信号中构成仿真鼾声信号,进行算法仿真验证。通过广义互相关函数法计算出仿真信号的结果如图 4 所示。

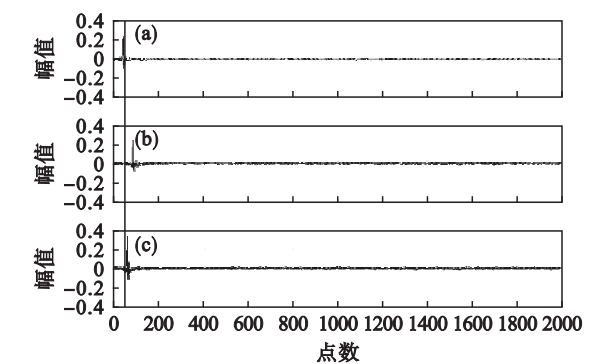


图 4 仿真鼾声信号 GCC 计算结果
Fig. 4 GCC calculation result of emulational snoring signal
(a)—2 号麦克风; (b)—3 号麦克风; (c)—4 号麦克风.

根据结果中最大计算值点位置计算相对时延 $\tau_i = p/f, i=2,3,4$, 其中 p 为最大值点所在位置, f 为信号采样频率. 将 τ_i 的计算值代入式(3)中计算得到仿真信号的声源位置为 $(0.35, 0.58)$, 验证了定位算法的可行性.

2.4 打鼾时头部姿态判断方法概述

单点的近场声源模型与头部姿态模型的差别在于人体头部自身对声音传播产生很大的影响, 鼾声在近场模型中由于发声方式和头部自身对信号传输的阻碍, 所以鼾声的传播方向性较强, 会有因为面部朝向与距离较近的麦克风相反而导致距离头部较近麦克风幅值反而较小的现象产生. 因此单纯根据麦克风接收到的信号能量和定位位置不能判断出头部的姿态. 考虑到存在打鼾者面部朝下的特殊情况, 打鼾者的面部将与系统中的枕头充分接触而导致信号能量被大量吸收. 试验中分析了打鼾者面部向下时采集到的声音信号.

试验中受试者以面部向下发出声音, 在图 2 中⑥区域发声, 采集到的波形图如图 5 所示.

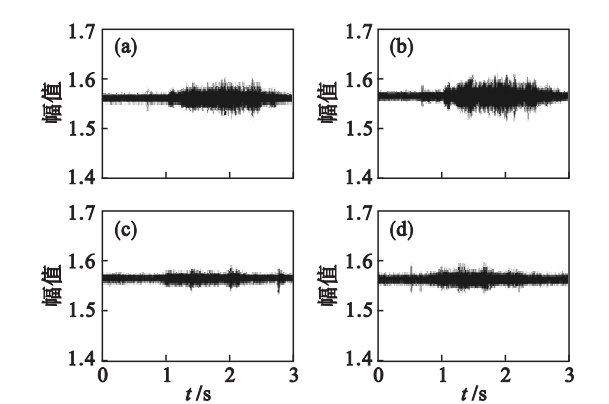


图 5 面部朝下时波形图
Fig. 5 Snoring signal oscillograms with head down
(a)—1 号麦克风; (b)—2 号麦克风;
(c)—3 号麦克风; (d)—4 号麦克风.

计算面部朝下时采集到的信号的广义互相关函数结果如图 6 所示.

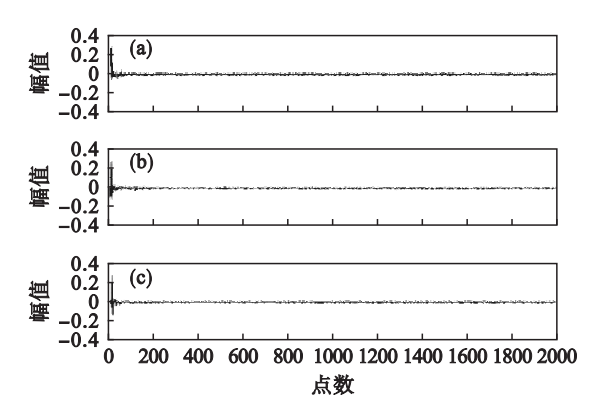


图 6 面部朝下时广义互相关函数结果图
Fig. 6 GCC calculation result of emulational snoring signal with head down
(a)—2 号麦克风; (b)—3 号麦克风; (c)—4 号麦克风.

经过多组重复实验, 面部朝下时头部在不同位置的情况下各麦克风之间仍然存在幅值差异, 但是广义互相关函数计算结果基本不变, 即面部贴合于枕头上时, 整个枕头取代人体成为声源, 因此通过广义互相关函数计算得到位置为固定不变值. 面部朝上时头部位置变换必然会导致各路信号幅值相应变化, 这与面部朝下时相比为明显差异. 因此本文根据此特性作为判断面部朝下这种特殊情况的判断依据, 并且通过各路信号幅值差异判断面部朝下时头部位置.

根据位置和信号能量参数判断头部姿态的流程如图 7 所示.

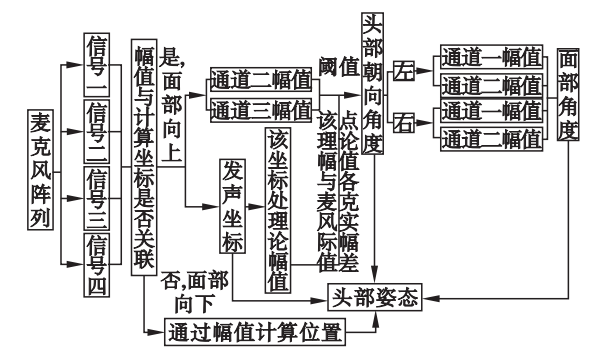


图 7 头部姿态判断流程
Fig. 7 Flow chart of head posture detecting

根据图中流程, 通过广义互相关函数排除幅值对信号定位影响后, 4 路信号计算出发生坐标. 之后对比该点坐标在信号模型中的理论信号能量参数与 2, 3 号麦克风采集到的实际信号能量参数以得到头部朝向信息. 其中采用 2, 3 号麦克风而不使用 1, 4 号麦克风信号是为了排除身体阻碍对信号产生的影响. 在得到头部朝向信息后, 以头部朝向信息为根据对比该方向麦克风组信号能量参数与该坐标点理论信号能量参数以得到面部朝向

角度.面部朝下时依次通过对比 1,2 号麦克风能量参数与 3,4 号麦克风能量参数判断头部位置.本文中信号能量参数采用信号功率,以排除信号采集过程中通过公式 $\frac{1}{N} \sum_{n=0}^{N-1} |x(n)|^2$ 计算相同长度信号的信号功率作为判断多路信号强弱的参数.

3 实验结果与分析

由于鼾声较强的方向性和人体头部姿态的复杂性,为了更简明地描述头部姿态,本文对头部姿态做了定性的区分.面部朝向分为向左、向上、向右. β 为面部与身体平面夹角,角度范围为 $190^\circ \sim 90^\circ$.如图 8 所示,综合头部中心位置作为发声坐标,得到打鼾时头部姿态的 3 个参数:面部朝向、头部夹角与头部位置.

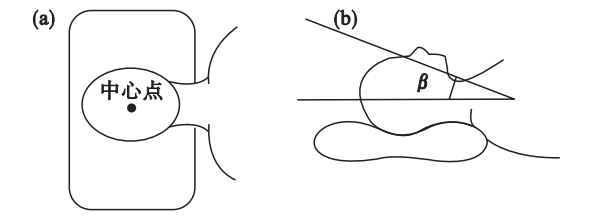


图 8 头部姿态参数示意图

Fig. 8 Schematic diagram of head posture

(a)—俯视;(b)—侧视.

试验中受试者根据以上判定方法以不同姿态躺在测试系统中的枕头上进行模拟打鼾,模拟打鼾者按照典型鼾声的发声方式进行模拟.本文通过声源定位技术对鼾声进行定位,根据声源定位原理,信号本身的声音特性对定位不会产生直接影响,因此可将模拟鼾声作为发声时判断头部姿态的实验采集对象.模拟打鼾者根据 2.4 节中头部姿态判断方法进行计算,测试结果见表 1.

从表 1 中可以看出,本文声源定位算法有效判断出了受试者打鼾时头部在枕头上的姿态.当受试者以仰卧姿态模拟打鼾时判断结果精确度达 90%,侧卧时面部朝向准确度达到 60%,位置精度达到 80%,俯卧时由于只能通过对比幅值判断头部位置,而幅值又受到身体本身等多因素影响,判断误差在 30% 左右.本文定位算法在面部向上时有效排除了幅值对定位的影响,侧卧姿态人体本身对声音传播影响较大,且不同种类的鼾症的发声部位不同^[6],受试者模拟打鼾时为尽量接近真实打鼾,模拟不同的打鼾方式,导致朝向角度判定产生一定误差.

表 1 模拟打鼾时头部姿态识别结果
Table 1 Head posture identification results of emulational snoring

编号	理论			实际		
	$\beta/(^\circ)$	朝向	位置	$\beta/(^\circ)$	朝向	位置
1	0	上	(50,75)	0	上	(48,72)
2	0	上	(50,50)	0	上	(49,51)
3	0	上	(50,25)	0	上	(54,23)
4	0	上	(25,25)	0	上	(22,21)
5	0	上	(25,50)	0	上	(27,55)
6	0	上	(75,25)	0	上	(79,27)
7	0	上	(75,50)	0	上	(77,56)
8	-25	左	(50,75)	-19	左	(77,56)
9	0	左	(50,50)	3	左	(77,56)
10	45	左	(50,25)	33	左	(71,41)
11	0	左	(25,25)	2	左	(22,71)
12	45	左	(25,50)	56	左	(28,44)
13	0	左	(75,25)	-11	左	(65,23)
14	-25	左	(75,50)	-31	左	(81,47)
15	-25	右	(50,75)	-36	右	(41,69)
16	0	右	(50,50)	-5	右	(57,48)
17	45	右	(50,25)	56	右	(58,21)
18	0	右	(25,25)	17	右	(28,22)
19	45	右	(25,50)	32	右	(29,43)
20	0	右	(75,25)	15	右	(80,20)
21	-25	右	(75,50)	-31	右	(78,46)
22	—	下	(75,50)	—	下	(86,33)
23	—	下	(50,50)	—	下	(61,35)
24	—	下	(25,50)	—	下	(39,37)
25	—	下	(25,75)	—	下	(36,88)

4 结 论

本文利用声源定位技术,根据鼾声的声学特性和传播特性,在定位出发声位置的基础上通过对比发声位置理论信号幅值与实际信号幅值判断出受试者打鼾时头部朝向,结合头部朝向与发声位置得到打鼾时头部姿态.实验结果表明,在受试者仰卧时计算判断出的睡眠姿态与实际睡眠姿态吻合率达 90%,俯卧时由于声源的发声方式特殊,判断误差在 30% 左右,主要由人体自身影响与该姿态声音传播模型特殊造成,侧卧时计算判断出的打鼾时头部姿态与实际姿态吻合率达 80%.