

基于相关峰细化时延估计的舱内服务机器人发话人定位研究

王啸臻¹, 王兆魁², 张育林^{1,2}

(1. 国防科技大学 空天科学学院, 湖南 长沙 410073; 2. 清华大学 航天航空学院, 北京 100084)

摘要:发话人定位是舱内服务机器人有效区分航天员与环境并获取与航天员相对位置关系的重要手段。针对时延估计的机器人声定位算法精度受采样频率和噪声限制的问题,提出了一种基于相关峰精确插值的空间六元阵列发话人定位方法。该方法基于机器人球形结构设计,利用信号预处理和二次相关降低噪声干扰,通过线性调频 Z 变换(MCZT)取代快速傅里叶变换(FFT)计算细化频谱,突破已有时域采样率的限制,能有效弱化 FFT 带来的栅栏效应,提高相关函数分辨率、时延估计精度及发话人定位精度。实验结果表明:基于相关峰精确插值算法的发话人定位方法,其时延估计性能有明显提升,能较好地对声源目标进行定位,且算法精度优于基于广义互相关的发话人定位方法,能满足舱内服务机器人的定位需求。

关键词:服务机器人;发话人定位;时延估计;相关峰精确插值

中图分类号:TP242

文献标志码:A

DOI:10.19328/j.cnki.1006-1630.2018.05.002

Study on Speaker Localization for Assistant Robots in Spacecraft Based on Fine Interpolation of Correlation Peak

WANG Xiaozhen¹, WANG Zhaokui², ZHANG Yulin^{1,2}

(1. College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, Hunan, China;

2. School of Aerospace Engineering, Tsinghua University, Beijing 100084, China)

Abstract: Speaker localization is an important means for assistant robot in spacecraft to effectively distinguish the astronauts from the environment and acquire the relative position relationship with the astronauts. To solve the problem that the accuracy of robot sound localization algorithm based on time delay estimation (TDE) is limited by sampling frequency and noise, a speaker localization algorithm based on six-element cone-shaped array via fine interpolation of correlation peak (FICP) is proposed in this paper. The improved method based on the spherical structure uses signal pre-processing and secondary correlation to reduce the noise, and uses the modified chirp-Z transform (MCZT) rather than the fast Fourier transform (FFT) to calculate the refined spectrum for breaking through the limitation of the existing time domain sampling rate. The improved method can effectively weaken the fence effect caused by FFT. It can raise the resolution of correlation function and improve the accuracy of TDE and speaker localization. Experimental results show that the speaker localization method based on FICP algorithm can locate the sound source target efficiently, and the accuracy of the improved method is better than that of the method based on generalized cross-correlation. Thus, the method can satisfy the localization requirement of the assistant robot.

Keywords: assistant robot; speaker localization; time delay estimation; fine interpolation of correlation peak

0 引言

当前空间站的稳定运营主要依赖于航天员的长期在轨工作。航天员需要独立完成飞船操控、在轨科学实验操作、科学演示等多项任务,在轨工作效能有限。随着机器人技术的快速发展,利用机器人辅助完成在轨任务成为可能。国际上已先后提出多个舱内辅助机器人项目,如 PSA^[1]、Astrobee^[2],以及曾在国际空间站服役的 Smart Spheres^[3]和 Int-Ball 等。这些机器人多采用遥控方式运行,其自主运行、人机智能交互能力有限,需要航天员在轨照料。为发展更自主的舱内服务机器人,提升机器人的任务辅助能力,需要解决机器人对航天员的信息感知问题。惯性与视觉导航主要用于获取舱内的相对位置和姿态,而舱内机器人无法有效实时地获取航天员的相对位置。利用麦克风阵列进行声源定位,感知航天员的语音信息,可有效区分航天员与环境,实时获取航天员角度位置关系,这一方法计算速度快,结构简单。

目前使用麦克风阵列进行声源定位的方法一般可分为 3 类:基于高分辨率谱估计、基于可控功率响应和基于时延估计。而机器人声源定位系统一般要求麦克风阵列数量少且易搭载,机器人运动特性则要求定位方法高效迅速且具备一定的抗混响和抗噪声能力^[4]。由于基于时延估计的声源定位方法计算复杂度低,实时性高,且硬件成本较低,因此在机器人系统中得到了广泛应用。

基于时延估计的定位方法是一种双步定位方法,它首先估计声源到达不同麦克风的时间差,再利用得到的时延结合麦克风阵列结构,确定声源的位置^[5]。自 1976 年以来相关学者提出了许多时延估计的算法,如基于声学传递函数比的算法^[6]、基于基音加权的算法^[7]、基于自适应特征值分解的算法^[8]、基于广义互相关函数的算法^[9]等。广义互相关函数法由于其计算速度快且符合实时解算要求,因此应用最广泛,它通过计算两路信号的互相关函数求取峰值,相关函数峰值对应的时间就是两路信号的时延差。这类方法原理简单,但受噪声和时域采样频率的限制,所以时延估计精度不高。

为提高时延估计精度,针对广义互相关时延估计方法使用快速傅里叶变换(FFT)会在频域上产生栅栏效应,从而影响时延估计精度的问题,提出了基于相关峰精确插值(FICP)的空间六元阵列发话人

定位方法。该方法基于机器人球形结构设计,利用谱减法 and 二次相关降低噪声干扰,通过对声音信号进行频谱细化提高相关函数的时间分辨率,能有效改善 FFT 带来的栅栏效应,提高时延估计及定位精度。改进算法既继承了相关峰插值算法的优点,又增强了语音处理的抗噪声能力。本文对传统广义互相关算法和改进算法在时延估计准确性方面进行了比较分析,实验结果表明:改进算法可有效改善发话人定位性能。

1 基于球形结构的空六元阵列定位原理

考虑到舱内服务机器人为球形结构设计(直径 $D=230\text{ mm}$),因此系统中的麦克风阵列结构采用空间六元阵。空间六元阵可进行全空域声源定位,相比平面四元阵在保证分维特性的同时还拥有更优的定位精度^[10]。舱内服务机器人结构如图 1 所示。

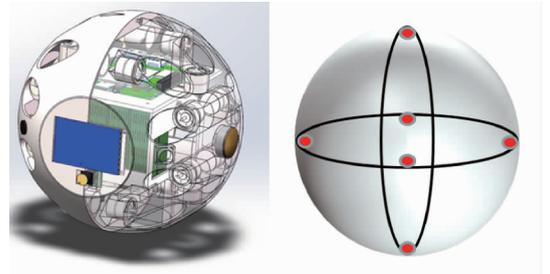


图 1 舱内服务机器人结构

Fig. 1 Layout of assistant robots

在直角坐标系下对空间六元麦克风阵列进行描述,如图 2 所示。6 个阵元的坐标分别为 $M_1(d/2, 0, 0)$ 、 $M_2(0, d/2, 0)$ 、 $M_3(-d/2, 0, 0)$ 、 $M_4(0, -d/2, 0)$ 、 $M_5(0, 0, d/2)$ 、 $M_6(0, 0, -d/2)$,其中 d 为阵列孔径; r 为目标(声源) $S(x, y, z)$ 到阵列中心(坐标原点)的距离; φ 为目标方位角; θ 为目标俯仰角; C 为空气中声音传播速率。假设目标为点声源,到达各阵元的传播时间为 $\tau_i (i = 1, 2, \dots, 6)$,则目标到各阵元的距离为 $r_i = C \cdot \tau_i (i = 1, 2, \dots, 6)$ 。若到达两阵元的时延差为 τ_{ij} ,则声源到 M_1 和 M_3 的距离差 $d_1 = C \cdot \tau_{31}$,到 M_2 和 M_4 的距离差 $d_2 = C \cdot \tau_{42}$,到 M_6 和 M_5 的距离差 $d_3 = C \cdot \tau_{65}$ 。

由麦克风阵列和目标的几何位置关系可联立方程组

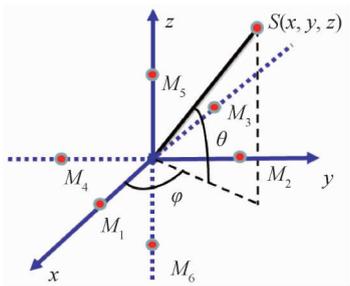


图 2 空间六元阵列结构

Fig. 2 Schematic diagram of six-element cone-shaped microphone array

$$\begin{cases} x^2 + y^2 + z^2 = r^2 \\ (x - d/2)^2 + y^2 + z^2 = r_1^2 \\ (y - d/2)^2 + x^2 + z^2 = r_2^2 \\ (x + d/2)^2 + y^2 + z^2 = r_3^2 \\ (y + d/2)^2 + x^2 + z^2 = r_4^2 \\ (z - d/2)^2 + y^2 + x^2 = r_5^2 \\ (z + d/2)^2 + y^2 + x^2 = r_6^2 \\ r_3 - r_1 = d_1 \\ r_4 - r_2 = d_2 \\ r_6 - r_5 = d_3 \end{cases} \quad (1)$$

由直角坐标与球坐标的转换关系可得目标方位信息 r, φ, θ 。因 $r, r_i \gg d_j$, 故可得

$$\sin \varphi = d_2 / \sqrt{d_1^2 + d_2^2} \quad (2)$$

$$\cos \theta = d_3 / \sqrt{d_1^2 + d_2^2 + d_3^2} \quad (3)$$

对式(2)、(3)进行简化和反三角变换可得方位角 φ 和俯仰角 θ 的解算表达式为

$$\begin{cases} \varphi = \arcsin\left(\frac{\tau_{42}}{\sqrt{\tau_{31}^2 + \tau_{42}^2}}\right), & \varphi \in [0^\circ, 360^\circ) \\ \theta = \arccos\left(\frac{\tau_{65}}{\sqrt{\tau_{31}^2 + \tau_{42}^2 + \tau_{65}^2}}\right), & \theta \in [0^\circ, 180^\circ) \end{cases} \quad (4)$$

平面四元阵列的方位角和俯仰角的解算表达式^[11]为

$$\begin{cases} \varphi = \arcsin\left(\frac{\tau_{41} - \tau_{21}}{\sqrt{\tau_{31}^2 + (\tau_{41} - \tau_{21})^2}}\right), & \varphi \in [0^\circ, 360^\circ) \\ \theta = \arcsin\left(\frac{c \sqrt{\tau_{31}^2 + (\tau_{41} - \tau_{21})^2}}{d}\right), & \theta \in [0^\circ, 180^\circ) \end{cases} \quad (5)$$

根据实际测量经验,一般得到的方位角和俯仰角误差较小,而距离的测量误差较大。考虑到机器人的实际需要,仅得到高精度的角度测量信息即可

准确找到人员位置,从而能为视觉导航与人机交互提供辅助,因此本文仅给出角度测量公式。式(5)中测角精度受声速估计误差和阵元安装误差影响。由式(4)可知,空间六元阵测量方位角 φ 和俯仰角 θ 仅与时延 τ_{ij} 有关,与声速 C 和阵列孔径 d 无关,定位精度受时延 τ_{ij} 和声速 C 的估计误差影响。对比平面四元阵,空间六元阵的测角受干扰更小,定位精度更高。由此可见,得到高精度的时延估计值即可求得较准确的角度位置信息。

2 基于相关峰精确插值的改进时延估计算法

2.1 相关峰精确插值算法

为提高相关函数的分辨率,对相关峰进行插值计算,一般采用基于信号内插法的时延估计算法,其估计精度受限于内插倍数,复杂度较高。如果要达到 0.01 采样间隔估计精度,需要计算 100 倍插值,不仅计算量大,还会带来新的误差源,因此需要计算相关峰的精确算法。相关峰精确插值算法的基本思想是针对两路长度均为 N 点的语音信号,计算其互功率密度谱并进行频域补零,构建完整的频谱序列以提高频谱采样率及时域内相关函数分辨率。FICP 算法是在线性调频 Z 变换(MCZT)的基础上,利用互谱的前后两端计算相关函数的峰值波形,并对相关峰进行细化。FICP 算法的具体实现过程如下:

1) 利用 MCZT 分别计算两路长度为 N 的信号 $x_1(n)$ 和 $x_2(n)$, 得到频谱 $X_1(k)$ 和 $X_2(k)$ (频谱间隔由 N_1 决定), 再计算两路信号的互功率谱 $R(k) = X_1(k)X_2^*(k)$, 频谱细化 N_1/N 倍。由谱的共轭对称性得到 N_1 点的完整互谱

$$R_1(k) = \begin{cases} R(k), & k = 0, 1, \dots, N-1 \\ 0, & k = N, N+1, \dots, N_1 - N - 1 \\ R^*(N_1 - k), & k = N_1 - N, N_1 - N + 1, \dots, N_1 - 1 \end{cases} \quad (6)$$

2) 对 $R_1(k)$ 补零, 将互谱拉长至 N_2 ($N_2 > N_1$) 点, 得到新的互谱

$$R_2(k) = \begin{cases} R(k), & k = 0, 1, \dots, N-1 \\ 0, & k = N, N+1, \dots, N_2 - N - 1 \\ R^*(N_2 - k), & k = N_2 - N, N_2 - N + 1, \dots, N_2 - 1 \end{cases} \quad (7)$$

3) 对 $R_2(k)$ 进行 IMCZT 逆变换得到相关函数 $R(n)$, 由此可突破时域采样率限制, 提高相关函数分辨率, 其分辨率可提高 N_2/N_1 倍。因阵列信号的时延估计值在有限范围内, 故仅计算相关函数主峰左右各 N 点, 组成长度为 $2N$ 的新相关函数, 降低运算量。

对于 $n = 0, 1, \dots, N-1$ 的时域范围, 相关函数为

$$R_{n_1}(n) = \frac{1}{N_2} \sum_{k=0}^{N-1} R(k) \exp\left(j \frac{2\pi}{N_2} kn\right) + \frac{1}{N_2} \sum_{k=N_2-N}^{N_2-1} R(k) \exp\left(j \frac{2\pi}{N_2} kn\right) \quad (8)$$

对于 $n = N_2 - N, N_2 - N + 1, \dots, N_2 - 1$ 的时域范围, 取 $n = N_2 - N + n_1$, 对应 $n_1 = 0, 1, \dots, N-1$ 的相关函数为

$$R_{n_2}(N_2 - N + n_1) = \frac{1}{N_2} \sum_{k=0}^{N-1} R(k) \exp\left(j \frac{2\pi}{N_2} k(N_2 - N + n_1)\right) + \frac{1}{N_2} \sum_{k=N_2-N}^{N_2-1} R(k) \exp\left(j \frac{2\pi}{N_2} k(N_2 - N + n_1)\right) \quad (9)$$

由 $R_{n_1}(n)$ 和 $R_{n_2}(n)$ 组成的相关函数为

$$R_n(n) = \begin{cases} R_{n_2}(N + n), & n = -N, -N + 1, \dots, -1 \\ R_{n_1}(n), & n = 0, 1, \dots, N-1 \end{cases} \quad (10)$$

经过复杂公式推导^[12], 可得化简后的相关函数表达式

$$R_{n_1}(n_1) = \frac{1}{N_2} \left\{ D_{\text{IMCZT}}[R(k)] + \exp\left(j \frac{2\pi}{N_2} kn\right) \times D_{\text{IMCZT}}[R^*(N-k)] \right\} \quad (11)$$

$$R_{n_2}(n_2) = \frac{1}{N_2} \left\{ D_{\text{IMCZT}}\left[R(k) \cdot \exp\left(j \frac{2\pi}{N_2} kN\right)\right] + \exp\left[j \frac{2\pi}{N_2} (N^2 - n_1 N)\right] \times D_{\text{IMCZT}}\left[R^*(N-k) \cdot \exp\left(j \frac{2\pi}{N_2} kN\right)\right] \right\} \quad (12)$$

线性调频 Z 变换为计算细化频谱的有效方法, 计算精度高且运算速度快。 D_{MCZT} 和 D_{IMCZT} 分别代表线性调频 Z 变换及其逆变换, 其定义式分别为

$$X(k) = D_{\text{MCZT}}[x(n)] = \sum_{n=0}^{N-1} x(n) \exp\left(-j \frac{2\pi}{N_1} kn\right) \quad (13)$$

$$x'(n) = D_{\text{IMCZT}}[X(k)] = \sum_{k=0}^{N-1} X(k) \exp\left(j \frac{2\pi}{N_1} kn\right) \quad (14)$$

2.2 算法分析与比较

选取一段时长 0.3 s 的语音信号进行分析, 采样频率 $f_s = 44\ 100$ Hz, FICP 计算参数 $N_1 = 44\ 100$, $N_2 = 500\ 000$ 。分别用 FICP 和 FFT 计算频域频谱, 图 3 选取了包含谱峰的一段频谱图(频率为 400~600 Hz), 可以看出用 FICP 得到的频谱信息比 FFT 更丰富, 谱峰高度也更精确。在谱峰附近, 用 FFT 计算得到的频率依次为 470.1、478.2、486.4、502.6、510.7 Hz, 用 FICP 计算得到的频率依次为 484、485、486、487、488 Hz, FICP 的频谱泄漏更少。通过互功率谱计算能提高时延估计值精度, 频谱分辨率从 8 Hz 提高到 1 Hz。用 FICP 和 FFT 计算得到的相关函数波形比较如图 4 所示。可以看出改进算法相当于对峰值附近的一段相关函数进行插值, 使峰值附近的相关函数波形更精细, 局部分辨率明显提高, 时延估计值更逼近真实值, 精度更高。

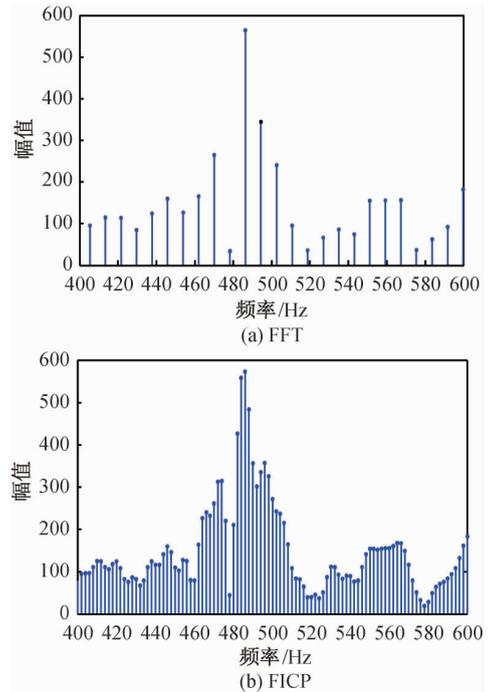


图 3 FICP 和 FFT 计算频谱比较

Fig. 3 Calculated spectra of FICP and FFT algorithms

2.3 时延估计算法的改进

FICP 算法虽然能提高相关函数的分辨率, 但并未改进抗噪声性能^[13-14]。因此, 通过语音信号预处理

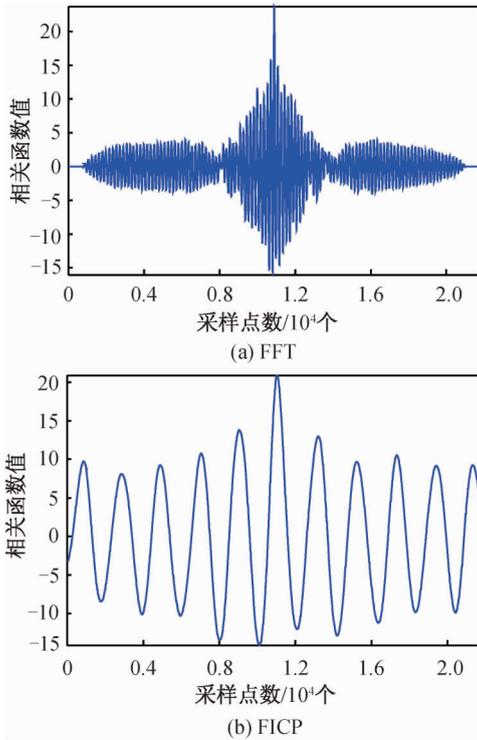


图 4 FICP 和 FFT 计算相关函数波形图比较
Fig. 4 Signal cross-correlation waveforms calculated by FICP and FFT algorithms

理和二次相关进一步提高算法时延估计的鲁棒性。

首先利用谱减法对语音信号进行降噪处理,设定前导无话段时长估计环境噪声能量,将谱减算法与原信号“作差”得到降噪信号,并利用基于倒谱距离的端点检测筛选有话段,减小互相关函数计算长度,从而提高计算效率。

采用二次相关的目的在于抑制噪声,每进行一次相关计算,都能提高信号的信噪比^[15]。对信号 1 作自相关,对信号 1 和 2 作互相关,将信号 1 的自相关函数和互相关函数作相关计算,进一步削弱噪声影响。时延估计改进算法流程如图 5 所示。

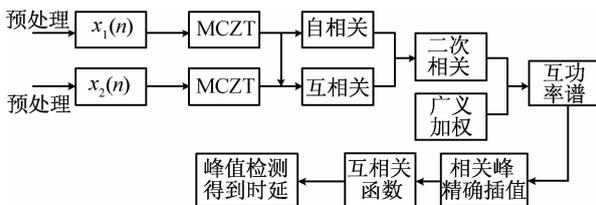


图 5 时延估计改进算法流程

Fig. 5 Flowchart of improved time delay estimation algorithm

完整的改进时延估计算法如下:1)利用 MCZT 取代 FFT 计算两路预处理后的语音信号 1 和 2 的

细化频谱来提升频谱精度;2)计算信号 1 的自功率谱及信号 1 和 2 的互功率谱,在此基础上进行二次相关,提高抗噪性能,得到二次互功率密度谱;3)对上述互功率密度谱进行相关峰精确插值,提高二次相关函数的分辨率,通过峰值检测计算得到两路信号的时延差值。

3 算法仿真与实验测试

3.1 FICP 算法仿真

实验中分别选取频率 f_s 为 20 000、40 000、80 000 Hz 的单频正弦模拟信号,预设两接收信号时延差 $T_d=1.7 \times 10^{-4}$ s,采样点数 $N=5 000$,FICP 算法参数 $N_1=10 000$, $N_2=40 000$,预设信噪比 $R_{SN}=8$ dB,仿真次数为 10 次,如图 6 所示。

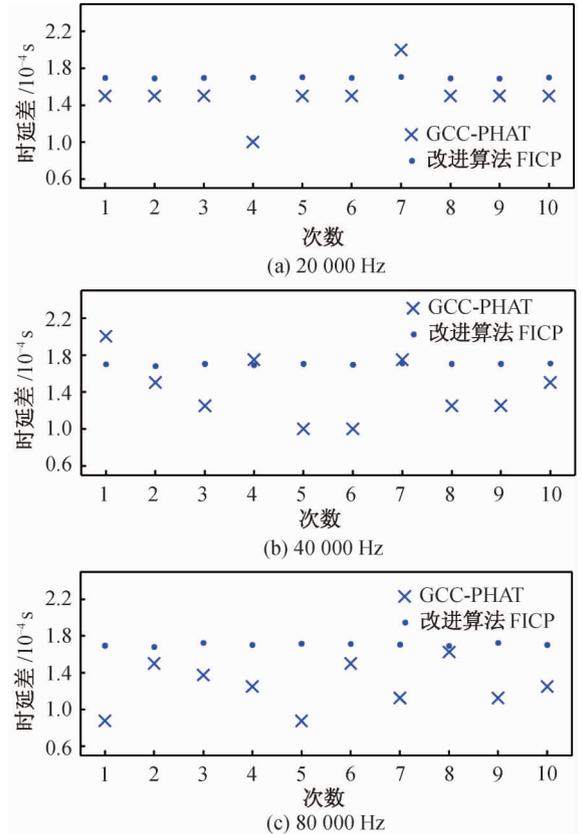


图 6 10 次计算结果比较

Fig. 6 Comparison of calculation results for ten times

由图 6 可知,当真实时延值为采样间隔的整数倍时,基于 FFT 计算的 GCC 法计算的时延值始终解算为采样间隔的整数倍,而 FICP 法计算的时延值则可准确求取真值,有

$$T_s = (1.7 \pm 0.02) \times 10^{-4} \text{ s} \quad (15)$$

式中: T_s 为时延估计值。

3.2 真实语音信号处理测试

实验中选取采样频率 $f_s = 16\ 000$ Hz, 采样点数

$N = 16\ 347$ 的两路真实语音信号, FICP 算法参数 $N_1 = 16\ 000$, $N_2 = 100\ 000$ 。经过谱减法降噪处理、倒谱距离端点检测筛选有话段后计算时延差。端点检测后得到有话段为 58~137 帧, 计 6 321 个采样点。语音信号处理如图 7 所示。

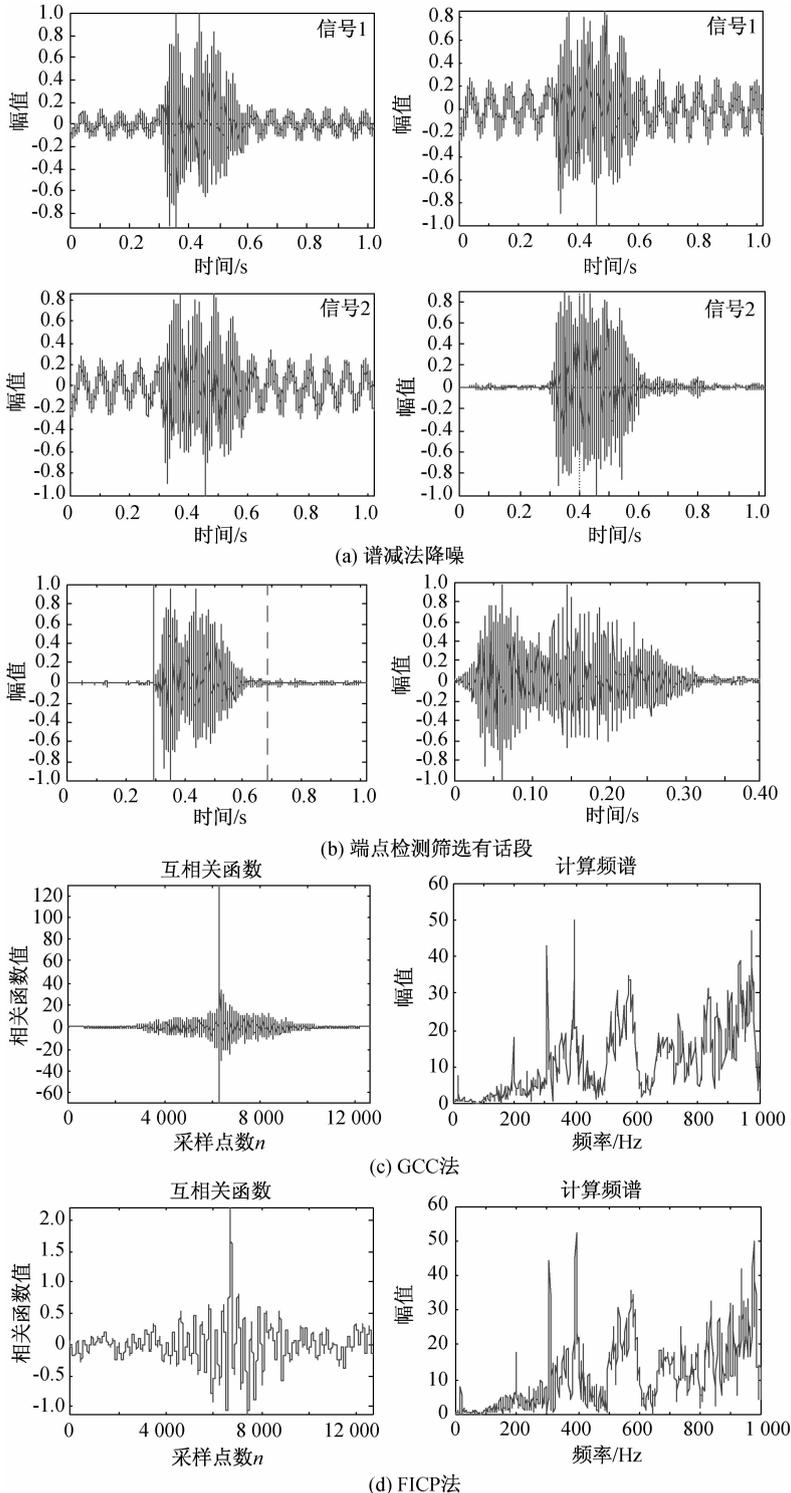


图 7 语音信号处理

Fig. 7 Voice signal processing

从图 7 可看出,用 FICP 法得到的计算频谱比 GCC 法更精细,频谱分辨率得到提升,峰值附近的相关函数波形更精细,局部分辨率明显提高。用 GCC 法计算的时延估计值为 3.75×10^{-4} s,即 6 个采样时长;用 FICP 法得到的时延估计值为 3.94×10^{-4} s,位于 3.75×10^{-4} s(6 个采样时长)和 4.375×10^{-4} s(7 个采样时长)之间。实验结果表明:用 FICP 法计算的时延值更接近时延差的真实值。

3.3 麦克风阵列定位测试实验

为验证空间六元麦克风阵列的定位效果,进行了静止目标的定位实验,如图 8 所示。实验条件如下:六元阵列及信号处理模块预设于简易结构的球形模型中,模型直径 23 cm,放置在高约 20 cm 的空心基座上。采用全指向性的 PDM 数字 MEMS 麦克风,频率响应 20 ~ 20 kHz,信号采样频率为 16 000 Hz。实际测量中,声源(音响)固定在距离麦克风阵列的一定范围内,通过改变目标声源的相对位置,完成对方位角和俯仰角的实时测量。



图 8 空间六元麦克风阵列

Fig. 8 Six-element cone-shaped microphone array

第 1 组实验中,声源相对阵列距离分别为 50、60、70 cm,保持同一水平高度,方位角 φ 间隔 30° 进行音频采集,定位求解结果如图 9 所示。

第 2 组实验中,声源保持同一水平高度,方位角 φ 设为 40° 和 70° ,声源相对距离为 40、50、60、70、80、90 cm 的 6 组,采集 12 次音频,定位求解结果如图 10 所示。

图 9、10 对比分析了基于广义互相关与相关峰精确插值这 2 种声源定位方法的定位结果。发话人定位旨在为机器人视觉导航提供与人员的相对位置关系,由于机器人服务工作在极其有限的舱内空间,

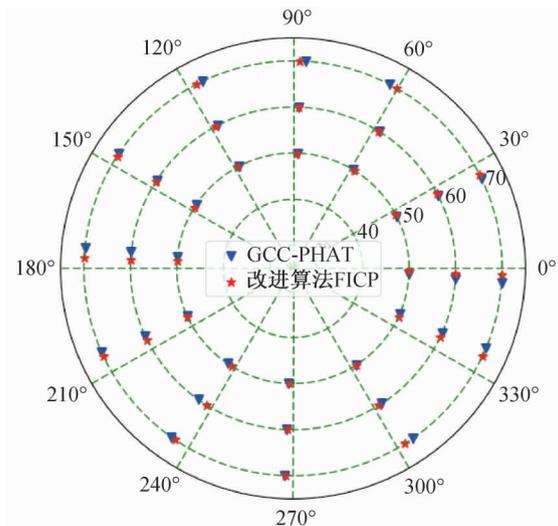


图 9 方位角测定值

Fig. 9 Results of azimuth angle measurement

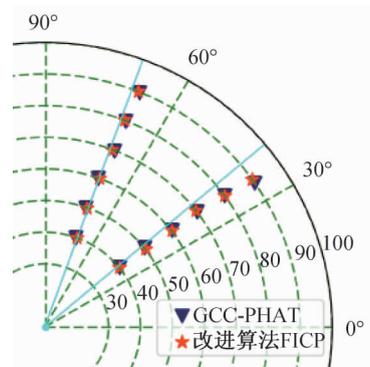


图 10 方位角测定值

Fig. 10 Results of azimuth angle measurement

优于 5° 的稳定定位精度即可满足机器人对人员的定位需求。随着相对距离的增大,基于 GCC 法的方位角测量误差逐渐增大,而基于改进算法的定位准确率则能保持在 90% 以上。可以看到,基于改进算法的声源定位结果明显优于基于 GCC 法的声源定位结果。

第 3 组实验中,声源保持同一方位角和水平距离,方位角 φ 设为 70° ,俯仰角 θ 为 70° 、 80° 、 90° 、 100° 、 110° 的 5 组,采集 10 次音频,定位求解结果如图 11 所示。

图 11 对比分析了空间六元阵列与平面四元阵列的俯仰角测量结果。可以看到,平面四元阵列的俯仰角测量误差很大,而空间六元阵列的俯仰角测量精度相对较高。综上所述,基于改进时延估计算法的空间六元阵列定位方法可有效改善发话人定位性能。

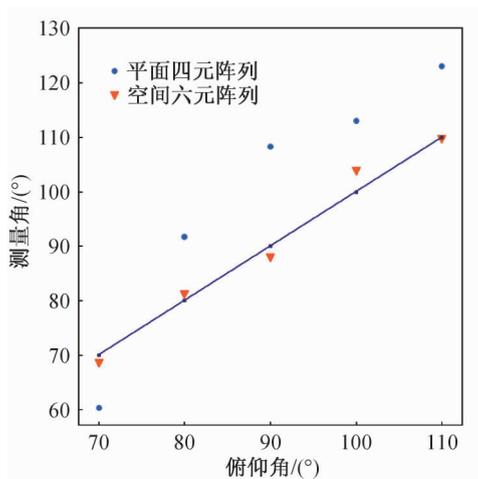


图 11 俯仰角测定值

Fig. 11 Results of pitching angle measurement

4 结束语

发话人定位是空间站舱内服务机器人实现人员相对位置关系确定及人机交互任务辅助的重要前提。本文提出了基于 FICP 和二次相关的发话人定位算法,并对算法的流程进行了推导说明。该方法基于机器人球形结构设计,利用信号预处理和二次相关降低噪声干扰,通过对声音信号进行频谱细化以提高相关函数的时间分辨率,能有效改善 FFT 变换带来的栅栏效应,提高时延估计及发话人定位精度。实验结果表明:相对基于广义互相关的定位方法,基于 FICP 的空间六元阵列发话人定位方法能得到更高的时延估计精度,从而对发话人进行更准确的定位。未来将围绕发话人定位与视觉导航方法的融合问题展开进一步研究,实现舱内辅助机器人的稳定跟踪。

参考文献

[1] DORAIS G A, GAWDIK Y. The personal satellite assistant: an internal spacecraft autonomous mobile monitor[C]//2003 IEEE Aerospace Conference Proceedings. Big Sky, MT:IEEE, 2003: 333-348.

[2] SMITH T, BARLOW J, BUALAT M, et al. Astrobe: a new platform for free-flying robotics on the international space station [C]//13th International Symposium on Artificial Intelligence, Robotics, and Automation in Space (i-SAIRAS). San Diego: NASA

Ames Research Center, 2016: 83-86.

[3] FONG T, MICIRE M J, MORSE T, et al. Smart spheres: a telerobotic free-flyer for intravehicular activities in space[C]//AIAA Space 2013 Conference and Exposition. San Diego: NASA Ames Research Center, 2013: 5338-5352.

[4] 李晓飞, 刘宏. 机器人听觉声源定位研究综述[J]. 智能系统学报, 2012, 7(1): 9-20.

[5] 崔玮玮, 曹志刚, 魏建强. 声源定位中的时延估计技术[J]. 数据采集与处理, 2007, 22(1): 90-99.

[6] BENESTY J. Adaptive eigenvalue decomposition algorithm for passive acoustic source localization[J]. Journal of the Acoustical Society of America, 2000, 107(1): 384-391.

[7] BRANDSTEIN M S. A pitch-based approach to time-delay estimation of reverberant speech[C]// Applications of Signal Processing to Audio and Acoustics. New Paltz:IEEE, 1997: 625-637.

[8] YEGNANARAYANA B, PRASANNA S R M, DURAI SWAMI R, et al. Processing of reverberant speech for time-delay estimation[J]. IEEE Transactions on Speech & Audio Processing, 2005, 13(6): 1110-1118.

[9] KNAPP C, CARTER G. The generalized correlation method for estimation of time delay[J]. IEEE Transactions on Acoustics Speech & Signal Processing, 2003, 24(4): 320-327.

[10] 王学青, 时银水, 朱岩. 空间六元麦克风阵列及其定位精度[J]. 哈尔滨工业大学学报, 2006, 38(8): 1392-1393.

[11] 马少春, 刘庆华, 黄灵鹭. 基于相关峰插值的五元十字阵被动声定位算法[J]. 探测与控制学报, 2014(5): 94-98.

[12] 杨亦春, 马驰州. 相关峰细化的精确时延估计快速算法研究[J]. 声学学报, 2003(2): 159-166.

[13] ZHANG Q, ZHANG L. An improved delay algorithm based on generalized cross correlation[C]// Information Technology and Mechatronics Engineering Conference. [S.l.]:IEEE, 2017: 395-399.

[14] 茅惠达, 张玲华. 声源定位中广义互相关时延估计算法的研究[J]. 计算机工程与应用, 2016, 52(22): 138-142.

[15] 唐娟, 行鸿彦. 基于二次相关的时延估计方法[J]. 计算机工程, 2007, 33(21): 265-267.

(本文编辑:姚麒伟)