

声场可视化系统中声像阵列空间关系标定研究*

温俊杰, 杨殿阁, 苗丰, 张凯, 连小珉

(清华大学 汽车安全与节能国家重点实验室 智能新能源汽车协同创新中心 北京 100084)

摘要:在声场可视化系统中,摄像机与传声器阵列的空间位置及角度的关系不准确将导致声源定位出现误差。而目前对于声场可视化系统中的声像阵列空间关系的研究较少,缺少对声像阵列空间关系的简单有效的标定方法。因此,提出了一种可计算任意位置与姿态的摄像机与传声器阵列之间空间关系的标定方法,可有效消除由传声器阵列与相机的空间关系不匹配导致的声源定位误差。通过对已知声源在多个空间位置的声源定位误差进行最小化处理,利用分步迭代计算的方法,分别计算出摄像机与传声器阵列的空间位置与角度,最终得到声像阵列之间真实的空间位置与角度关系。仿真与实际实验均表明本方法可对任意位置与姿态的摄像机与传声器阵列的空间关系进行准确标定,能有效提高声场可视化系统的声源定位精度。提出的声像阵列空间关系标定方法可应用于各种声场可视化系统中,对提高声场可视化系统的声源定位准确度有着积极的意义。

关键词:声源定位;声全息;空间关系标定

中图分类号: TH89 文献标识码: A 国家标准学科分类代码: 460.4099

Research on spatial relationship calibration technique for microphone array and camera in sound field visualization system

Wen Junjie, Yang Diange, Miao Feng, Zhang Kai, Lian Xiaomin

(State Key Laboratory of Automotive Safety and Energy, Collaborative Innovation Center of Intelligent New Energy Vehicle, Tsinghua University, Beijing 100084, China)

Abstract:In sound field visualization systems, imprecise spatial position and angle relationship between the microphone array and camera could lead to the error of the sound source localization. However, few of the previous studies have considered the spatial relationship between the camera and microphone array in sound field visualization systems. So there lacks a simple and effective spatial relationship calibration method for the spatial relationship between the camera and microphone array. In this paper, the calibration method for calculating any spatial orientation and position relationship between microphone array and camera is proposed, which can effectively eliminate the sound source localization error caused by the mismatched spatial relationship between the microphone array and camera, and then improve the precision of the sound source localization for sound field visualization system. Through minimizing the sound source localization errors of the known sound source in multiple positions, the spatial position and orientation between the camera and microphone array could be calculated respectively with stepped iteration calculation. Simulation and experiment results both show that the proposed method can accurately calculate any spatial orientation and position relationship between the camera and microphone array in a sound field visualization system, and could effectively improve the sound source localization accuracy of sound field visualization systems. The proposed spatial relationship calibration method could be used in various sound field visualization systems, and has an active significance for improving the sound source localization accuracy of sound field visualization systems.

Keywords:sound source localization; acoustic holography; spatial relationship calibration

1 引言

声场可视化系统通常由传声器阵列和摄像机组成。系统中传声器阵列获取声音信号,而摄像机获取声源的实景图像,最后将重建的声场与实景图像融合,获得直观的声场可视化图像^[1]。

在声场可视化系统中,由于重建的声场需要与摄像机所拍摄的图像进行融合,而重建声场与拍摄图像进行准确融合的前提是已知系统中传声器阵列与摄像机的空间关系。若传声器阵列与摄像机的空间关系不准确,将使真实声源投射在声源重建面与相机拍摄面的位置不重合,导致摄像机所拍摄的声源与声场识别的声源位置不匹配,从而产生声源的误识别,不利于噪声的有效识别和控制。

对于大多数声场可视化系统来说,摄像机一般固定于传声器阵列中心,如文献[2-3]中所使用的采集测试系统。然而,由于制造、装配等误差,会导致可视化系统中阵列与摄像机的空间关系不准确。同时,有些可视化系统中的传声器和摄像机的位置并不固定^[4-6],需要根据实验条件调整摄像机与传声器阵列的空间关系。这类系统的优点是可以根据实际情况获取被测物体的最佳拍摄位置及角度,甚至可以利用双目定位的原理对运动声源载体进行准确地定位以及测速。然而,由于其摄像机与传声器阵列的位置并不固定,往往导致在实际的实验中,实验者需要通过手动调整摄像机的位置和角度使得摄像机拍摄声源位置与识别的声源位置重合^[7]。这种方法不仅繁琐且不准确,而且不适用于对数据进行离线处理的声场可视化系统。

目前对于声场可视化系统中的声像阵列的空间关系研究比较少。OptiNav公司的Dougherty R. P.^[7]提到了一种在实时采集系统下手动调整摄像机的空间位置和角度的方法;奥克兰大学的Legg M等人^[8]提出了一种声像阵列空间关系的标定方法。但此种方法需将多个音箱固定于标定板上,标定方法复杂且不适用于户外声学测量。

本文提出了一种对于任意空间状态下摄像机与传声器阵列的空间关系的标定方法,可有效解决声场可视化系统中声像阵列空间关系不精确的问题。经仿真和实验可知此方法可以准确计算声场可视化系统中摄像机与传声器阵列的空间位置和角度关系,能消除由传声器阵列与摄像机空间关系不匹配导致的声源定位误差,提高声源定位精度。

2 声源定位方法

目前,声源定位方法主要包括声全息、波束形成及

声达时差(time difference of arrival, TDOA)等^[9-13]。本文所用的声源定位方法为文献[14-15]中提出的基于基尔霍夫衍射积分的远场声全息方法,其原理如图1所示。

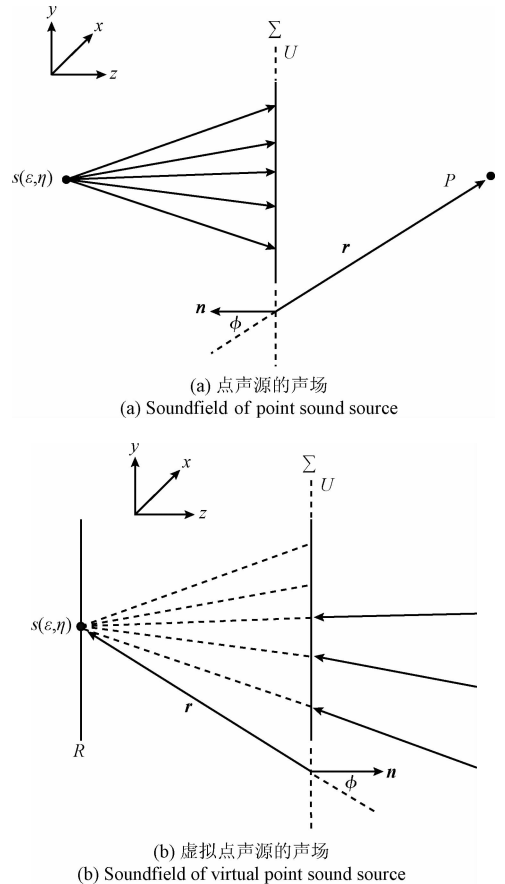


图1 基于基尔霍夫衍射积分的声全息原理

Fig. 1 The schematic diagram of acoustic holography based on Kirchhoff diffraction integration principle

图1中 n 为全息面外法线的方向, r 为矢径, ϕ 为 r 与全息面外法线方向 n 的夹角, Σ 为在测量面位置的无限大空间平面, R 为假定声源点 s 所在的声源面。 U 可以看做点声源 s 所发出的球面波在 Σ 处的空间声场分布。由图1(b)可看做虚拟声源 s_1 所发出的声波汇聚于 s ,则此声波与 s 点实际发出的球面波互为共轭波,则可得到基于基尔霍夫衍射积分的声全息重建公式为:

$$P_s(\varepsilon, \eta, f) = \frac{Ck}{j} \sum_M U^*(i, f) \left[1 + \left(1 - \frac{1}{jkr} \right) \frac{z_0}{r} \right] \frac{e^{-jkr}}{r} \Delta S \quad (1)$$

式中: M 为传声器个数, U^* 为第 i 个传声器的共轭声压, r 为重建面上点 $s(\varepsilon, \eta)$ 与第 i 个传声器的距离, ΔS 为传声器网格的面积。

3 可视化成像原理

图2所示为相机成像的坐标系系统,其中 $O_w X_w Y_w Z_w$ 为世界坐标系, $O_c X_c Y_c Z_c$ 为相机坐标系, OXY 为图像平面坐标系,而 OUV 为图像像素坐标系。假设空间某点 P 的世界坐标系坐标为 $\mathbf{X}_w = (x_w, y_w, z_w)^T$,且其相机坐标系坐标为 $\mathbf{X}_c = (x_c, y_c, z_c)^T$ 。则从世界坐标系转换为相

$$\mathbf{R}_c = \begin{bmatrix} \cos\beta\cos\theta & \cos\alpha\sin\theta + \cos\theta\sin\alpha\sin\beta & \sin\alpha\sin\theta - \cos\alpha\cos\theta\sin\beta \\ -\cos\beta\sin\theta & \cos\alpha\cos\theta - \sin\alpha\sin\beta\sin\theta & \cos\theta\sin\alpha + \cos\alpha\sin\beta\sin\theta \\ \sin\beta & -\cos\beta\sin\alpha & \cos\alpha\cos\beta \end{bmatrix} \quad (3)$$

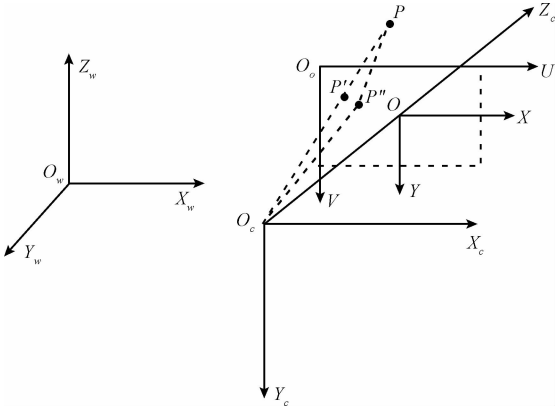


图2 相机成像坐标系

Fig. 2 The camera imaging coordinate system

对于相机的针孔成像模型,将相机坐标系坐标 $\mathbf{X}_c = (x_c, y_c, z_c)^T$ 转换为图像像素坐标 $\mathbf{X}_p = (u, v)^T$ 可以表示为:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_{cu} & 0 & u_0 \\ 0 & f_{cv} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (4)$$

式中: (f_{cu}, f_{cv}) 为相机焦距在图像像素坐标系横纵轴上所表示的像素大小,而 (u_0, v_0) 为图像平面坐标系原点在图像像素坐标系中的坐标,即主点坐标。参数 $(f_{cu}, f_{cv}, u_0, v_0)$ 共同组成了相机的内部参数。由图2可以看出,在理想情况下,若根据针孔成像模型,若根据针孔成像模型,点 P 将投射在图像点 P' ,而在实际情况下,由于镜头的畸变,点 P 将投射在点 P'' 。一种常用的非线性畸变成像模型如下:

$$\begin{cases} x' = x(1 + k_1 r^2 + k_2 r^4) + 2p_1 xy + p_2(r^2 + 2x^2) \\ y' = y(1 + k_1 r^2 + k_2 r^4) + 2p_2 xy + p_1(r^2 + 2y^2) \end{cases} \quad (5)$$

式中: $r^2 = x^2 + y^2$,而 $(x, y)^T = (x_c/z_c, y_c/z_c)^T$ 为针孔模型下的归一化相机坐标系坐标, $(x', y')^T$ 为修正后的归一化坐标,参数 (k_1, k_2, p_1, p_2) 为相机的内部畸变参数。根据式(5)即可对镜头成像的非线性模型进行畸变的校正。

机坐标系可以表示为:

$$\mathbf{X}_c = \mathbf{R}_c \mathbf{X}_w + \mathbf{T}_c \quad (2)$$

式中: \mathbf{R}_c 为 3×3 旋转矩阵,表示相机坐标系与世界坐标系的空间角度关系; \mathbf{T}_c 为 3×1 平移向量,表示了相机坐标系与世界坐标系的空间平移关系; \mathbf{R}_c 与 \mathbf{T}_c 共同组成了相机的外部参数。若将世界坐标系先后绕 $O_w X_w$ 、 $O_w Y_w$ 、 $O_w Z_w$ 轴按右手定则旋转一定角度 α, β, θ ,则旋转矩阵 \mathbf{R}_c 可以表示为式(3)。

4 声像阵列空间标定算法

假设空间某点 P 的世界坐标系坐标为 \mathbf{X}_w ,它的相机坐标系坐标为 \mathbf{X}_c ,而其在传声器阵列坐标系中的坐标为 \mathbf{X}_m 。显然, $\mathbf{X}_c, \mathbf{X}_m$ 与 \mathbf{X}_w 的转换关系如下式:

$$\begin{cases} \mathbf{X}_c = \mathbf{R}_{uc} \mathbf{X}_w + \mathbf{T}_{uc} \\ \mathbf{X}_m = \mathbf{R}_{um} \mathbf{X}_w + \mathbf{T}_{um} \end{cases} \quad (6)$$

式中: $\mathbf{R}_{uc}, \mathbf{R}_{um}$ 分别为世界坐标系相对于相机坐标系和传声器阵列坐标系的旋转矩阵,而 $\mathbf{T}_{uc}, \mathbf{T}_{um}$ 为对应的平移向量。由此可得点 P 的相机坐标与传声器阵列坐标的转换关系为:

$$\mathbf{X}_c = \mathbf{R}_{uc} \mathbf{R}_{um}^{-1} \mathbf{X}_m - \mathbf{R}_{uc} \mathbf{R}_{um}^{-1} \mathbf{T}_{um} + \mathbf{T}_{uc} \quad (7)$$

由式(7)可知,传声器阵列坐标系转换至相机坐标系的空间旋转矩阵 \mathbf{R}_{mc} 和平移向量 \mathbf{T}_{mc} 分别为:

$$\begin{cases} \mathbf{R}_{mc} = \mathbf{R}_{uc} \mathbf{R}_{um}^{-1} \\ \mathbf{T}_{mc} = \mathbf{T}_{uc} - \mathbf{R}_{uc} \mathbf{R}_{um}^{-1} \mathbf{T}_{um} \end{cases} \quad (8)$$

因此,若要标定相机与传声器阵列的空间角度和位置关系,则需要计算世界坐标系相对于相机坐标系和传声器阵列坐标系的空间关系。而将世界坐标系固结在传声器阵列上,则只需计算世界坐标系相对于相机坐标系的空间关系,可减少计算复杂度。

若已知声源距离传声器阵列的距离为 d_m ,声场重建的声源在图像上的广义像素坐标为 $\mathbf{X}_{pm} = (u_m, v_m, 1)^T$,而相机拍摄的声源在图像上的广义像素坐标为 $\mathbf{X}_{pc} = (u_c, v_c, 1)^T$ 。假设相机的内部矩阵为 \mathbf{A} ,则由式(4)可计算得重建声源在传声器阵列坐标系下的坐标为:

$$\mathbf{X}_m = d_m \mathbf{A}^{-1} \mathbf{X}_{pm} \quad (9)$$

假设标定计算的传声器阵列相对相机阵列的旋转矩阵和平移向量分别为 $\tilde{\mathbf{R}}_{mc}$ 和 $\tilde{\mathbf{T}}_{mc}$,则可估计出重建声源在相机坐标系下的坐标为:

$$\tilde{\mathbf{X}}_c = d_m \tilde{\mathbf{R}}_{mc} \mathbf{A}^{-1} \mathbf{X}_{pm} + \tilde{\mathbf{T}}_{mc} \quad (10)$$

由式(10)可得到重建声源距离相机的距离 d_c ,则根据式(4)可由重建声源的广义像素坐标 \mathbf{X}_{pm} 估算出当其

空间旋转矩阵和平移向量为 $\tilde{\mathbf{R}}_{mc}$ 和 $\tilde{\mathbf{T}}_{mc}$ 时所对应的拍摄声源的广义像素坐标为:

$$\tilde{\mathbf{X}}_{pc} = \frac{d_m}{d_c} \mathbf{A} \tilde{\mathbf{R}}_{mc} \mathbf{A}^{-1} \mathbf{X}_{pm} + \frac{1}{d_c} \mathbf{A} \tilde{\mathbf{T}}_{mc} \quad (11)$$

当估算的空间旋转矩阵 $\tilde{\mathbf{R}}_{mc}$ 和平移向量 $\tilde{\mathbf{T}}_{mc}$ 为真实值时, $\tilde{\mathbf{X}}_{pc}$ 与 \mathbf{X}_{pc} 差异最小。 $\tilde{\mathbf{R}}_{mc}$ 可由旋转角度 α, β 和 θ 决定,而 $\tilde{\mathbf{T}}_{mc}$ 则由平移分量 t_x, t_y 和 t_z 决定。因此求解 \mathbf{R}_{mc} 和 \mathbf{T}_{mc} 的问题可以转化为以下最优化问题:

$$\min \{ \mathbf{F}(\alpha, \beta, \theta, t_x, t_y, t_z) = \| \mathbf{X}_{pc} - \frac{d_m}{d_c} \mathbf{A} \tilde{\mathbf{R}}_{mc} \mathbf{A}^{-1} \mathbf{X}_{pm} + \frac{1}{d_c} \mathbf{A} \tilde{\mathbf{T}}_{mc} \|_2 \} \quad (12)$$

此问题是一个多目标的非线性最优化问题,可用分步迭代法求解。求解步骤如下:

- 1) 设定初始的旋转角度 $(\alpha^{(0)}, \beta^{(0)}, \theta^{(0)})$ 和平移分量 $(t_x^{(0)}, t_y^{(0)}, t_z^{(0)})$;
- 2) 将 $(\alpha^{(0)}, \beta^{(0)}, \theta^{(0)})$ 和 $(t_x^{(0)}, t_y^{(0)}, t_z^{(0)})$ 代入式(12),保持平移分量 $(t_x^{(0)}, t_y^{(0)}, t_z^{(0)})$ 不变,求出一次迭代旋转角度 $(\alpha^{(1)}, \beta^{(1)}, \theta^{(1)})$;
- 3) 将 $(\alpha^{(1)}, \beta^{(1)}, \theta^{(1)})$ 和 $(t_x^{(0)}, t_y^{(0)}, t_z^{(0)})$ 代入式(12),保持旋转角度 $(\alpha^{(1)}, \beta^{(1)}, \theta^{(1)})$ 不变,求出一次迭代平移分量 $(t_x^{(1)}, t_y^{(1)}, t_z^{(1)})$;
- 4) 满足迭代终止条件,终止迭代计算,得到第 N 次的迭代旋转角度 $(\alpha^{(N)}, \beta^{(N)}, \theta^{(N)})$ 和平移分量 $(t_x^{(N)}, t_y^{(N)}, t_z^{(N)})$,并将其作为标定结果。

由解得的 $(\alpha^{(N)}, \beta^{(N)}, \theta^{(N)})$ 和 $(t_x^{(N)}, t_y^{(N)}, t_z^{(N)})$ 即可估算出相应的旋转矩阵 \mathbf{R}_{mc} 和 \mathbf{T}_{mc} ,用此结果修正相机坐标系和传声器坐标系的空间角度和位置关系,即可使重建声源位置与相机拍摄声源位置重合,可有效消除由于声像阵列空间关系不精确导致的声源位置重建误差。

5 仿真验证及分析

假设相机图像的分辨率为 640×512 ,相机的等效焦距分别为 $f_{cu} = 600, f_{cv} = 600$,且其主点像素坐标值为 $u_0 = 319.5, v_0 = 255.5$ 。若令相机坐标系相对传声器坐标系的旋转角度 $(\alpha, \beta, \theta) = (0.05, -0.05, 0.06)$,且平移分量为 $(t_x, t_y, t_z) = (0.4, 0.3, -0.3)$ 。若单声源在 $s = 33$ 个不同的空间位置发声,且它们与传声器阵列的距离均为 $d_m = 5m$,如图3所示。初始迭代的旋转角度 $(\alpha^{(0)}, \beta^{(0)}, \theta^{(0)})$ 及平移分量 $(t_x^{(0)}, t_y^{(0)}, t_z^{(0)})$ 均为 $(0, 0, 0)$ 。

图4所示为在未标定计算声像阵列空间关系时重建声源与相机所拍摄声源在图像上的像素位置。显然,此时重建声源的位置与相机拍摄声源位置偏差较大,若不进行声像阵列空间关系的标定计算,则会导致较大的声源识别误差。

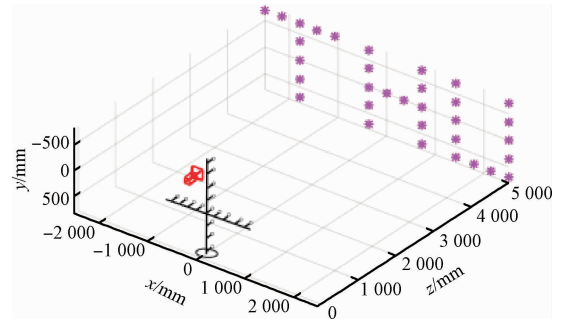


图3 仿真示意

Fig. 3 Sketch diagram of the simulation

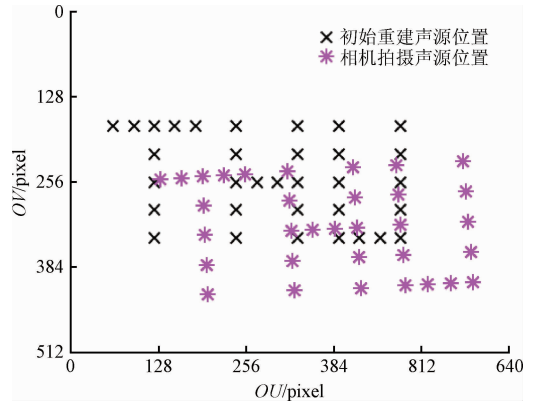


图4 声像阵列标定前重建声源与拍摄声源的位置关系

Fig. 4 Position relationship between the reconstructed and captured sound sources before the spatial relationship calibration of the camera and microphone array

由初始旋转角度 $(\alpha^{(0)}, \beta^{(0)}, \theta^{(0)}) = (0, 0, 0)$ 及初始平移分量 $(t_x^{(0)}, t_y^{(0)}, t_z^{(0)}) = (0, 0, 0)$ 开始迭代 $N = 150$ 步,各步迭代计算的结果如表1所示。随着迭代次数的不断增加,计算结果不断逼近真实值 $(\alpha, \beta, \theta) = (0.05, -0.05, 0.06)$ 和 $(t_x, t_y, t_z) = (0.4, 0.3, -0.3)$,由此可证明此标定方法的准确性。

表1 各步迭代计算结果 (n 为迭代步数)

Table 1 Calculation results of each iteration step (n is the number of iteration steps)

n	$(\alpha^{(n)}, \beta^{(n)}, \theta^{(n)})$	$(t_x^{(n)}, t_y^{(n)}, t_z^{(n)})$
1	$(0.103, -0.103, 0.004)$	$(0.118, 0.028, -0.289)$
10	$(0.094, -0.079, 0.061)$	$(0.243, 0.089, -0.298)$
50	$(0.062, -0.053, 0.060)$	$(0.383, 0.242, -0.302)$
100	$(0.052, -0.050, 0.060)$	$(0.397, 0.289, -0.300)$
150	$(0.050, -0.050, 0.060)$	$(0.400, 0.298, -0.300)$

利用迭代150次的标定结果去重建声源后,所重建声源位置与相机拍摄声源位置的关系如图5所示。与图5中未标定的重建结果相比,标定后的重建声源位置

能够与相机所拍摄的声源位置准确地匹配融合。可以看出,本标定方法可以在传声器阵列与相机的空间角度与位置关系均未知的情况下对它们的空间关系进行准确标定,并且利用此标定方法的结果可以有效地消除由于声像阵列空间关系不匹配所导致的声源重建位置误差。

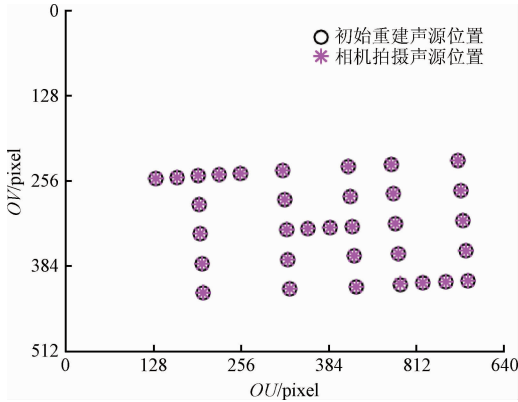


图5 声像阵列标定后重建声源与拍摄声源的位置关系(迭代次数 $N = 150$)

Fig. 5 Position relationship between the reconstructed and captured sound sources after the spatial relationship calibration of the camera and microphone array (number of iterations $N = 150$)

6 实验验证及分析

在实际的声像阵列空间关系标定实验中,需要在距离传声器阵列某个已知的距离 $z_m = d_m$ 平面内缓慢移动发声的有源音箱,如图6所示。摄像机和传声器同时捕捉信号,选取 $s (s > 20)$ 个位置的图像所对应的信号段进行声场重建分析。根据重建结果获取各组重建声源及拍摄声源的像素坐标,并由式(12)进行迭代计算即可标定出真实的旋转角度和平移分量。

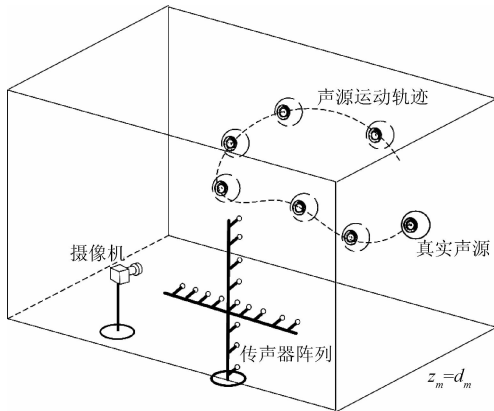


图6 声像阵列空间关系标定示意图

Fig. 6 Sketch diagram of the spatial relationship calibration of the camera and microphone array

标定实验所用相机图像的分辨率为 658×492 ,相机的等效像素焦距为 $f_{cu} = 837.9$ 和 $f_{cv} = 838.5$,其主点的像素坐标值为 $u_0 = 317.6$ 和 $v_0 = 256.3$ 。相机置于传声器阵列中心,有源音箱在距离传声器阵列中心,有源音箱在距离传声器阵列前 $z_m = 4.5$ m 的平面内缓慢移动并发声。选取 20 组不同时刻进行标定计算,其中 2 组未标定时的声场重建结果如图7所示。



(a) $t = 5.937$ s



(b) $t = 7.939$ s

图7 声像阵列标定前不同时刻的两组重建结果(‘+’为拍摄声源位置,‘O’为重建声源位置)

Fig. 7 Reconstruction results at two different time before the spatial relationship calibration of the camera and microphone array (‘+’ is captured sound source, ‘O’ is reconstructed sound source)

由图7可知,在未标定声像阵列的空间关系时,重建声源的位置虽然与拍摄到的声源比较接近,但仍然存在位置上的偏差。在初始迭代旋转角度和平移分量分别为 $(\alpha^{(0)}, \beta^{(0)}, \theta^{(0)}) = (0, 0, 0)$ 和 $(t_x^{(0)}, t_y^{(0)}, t_z^{(0)}) = (0, 0, 0)$ 时,经过 $N = 100$ 次迭代计算,计算得到声像阵列的空间旋转角度和平移分量分别为 $(\alpha^{(100)}, \beta^{(100)}, \theta^{(100)}) = (0.019, 0.028, 0.011)$ 和 $(t_x^{(100)}, t_y^{(100)}, t_z^{(100)}) = (-0.006, 0.005, -0.022)$ 。用此迭代计算结果去修正传声器阵列的空间位置,并对声场进行重建,则可得到与图7对应的标定后声源识别结果,如图8所示。可以看

出,标定后识别到的声源能够与拍摄的声源位置准确地对应,说明了标定计算的结果是可靠的。



(a) $t=5.937$ s



(b) $t=7.939$ s

图8 声像阵列标定后不同时刻的两组重建结果
(‘+’为拍摄声源位置,‘O’为重建声源位置)

Fig.8 Reconstruction results at two different time after the spatial relationship calibration of the camera and microphone array (‘+’ is captured sound source, ‘O’ is reconstructed sound source)

为了证明标定结果的正确性,利用此标定结果对另一组距离传声器阵列 $z_m = 5$ m 的声源进行了声场重建。声像阵列标定前后的声源识别结果如图9所示。



(a) $t=25.807$ s 标定前结果
(a) $t=25.807$ s, the result before calibration



(b) $t=25.807$ s 标定后结果
(b) $t=25.807$ s, the result after calibration

图9 声像阵列空间关系标定前后的声源识别结果
(测量距离 $z_m = 5$ m, ‘+’为拍摄声源位置,‘O’为重建声源位置)

Fig.9 Sound source identification results before and after the spatial relationship calibration of the camera and microphone array (measurement distance $z_m = 5$ m, ‘+’ is captured sound source, ‘O’ is reconstructed sound source)

可以看出,在标定前,声源定位存在一定的偏差,而利用标定结果对声像阵列的空间关系进行修正后,重建声源与拍摄声源的位置能准确地重合,证明了本标定方法的正确与有效性。

7 结 论

在目前没有简单有效的标定声场可视化系统中传声器阵列与摄像机空间关系方法的情况下,本文提出的声像阵列的空间关系标定方法可以准确计算出系统中传声器阵列和任意姿态相机的空间角度和位置关系,解决了传声器阵列与相机的空间对应问题,使重建声平面与拍摄图像空间准确地融合。仿真和实际的实验均验证了本文提出的标定方法能有效消除由于声像阵列空间关系不匹配导致的声源识别误差,可提高各种声场可视化系统的声源定位结果的准确及可靠性。

参考文献

- [1] O'DONOVAN A, DURAISWAMI R, NEUMANN J. Microphone arrays as generalized cameras for integrated audio visual processing [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2007: 1-8.
- [2] STIER J, BEILSCHMIDT M. Generating train side views from video sequences for microphone array pass-by measurements[C]. Proceedings on CD of the 4th Berlin Beamforming Conference, 2012: 676.
- [3] MEYER A, DÖBLER D, GRÄSER M. Efficient methods for the analysis of moving sound sources with high channel microphone arrays [C]. Berlin Beam-forming Conference, 2014: 1-9.

- [4] WAKABAYASHI Y, KURITA T, YAMADA H, et al. Noise measurement results of Shinkansen high-speed test train (FASTECH360S, Z) [C]. *Noise and Vibration Mitigation for Rail Transportation Systems*, 2008:63-70.
- [5] 李兵, 杨殿阁, 邵林, 等. 基于波束形成和双目视觉的行驶汽车噪声源识别[J]. *汽车工程*, 2008, 30(10): 889-892.
LI B, YANG D G, SHAO L, et al. Noise sources measurement and identification for moving automobiles[J]. *Automotive Engineering*, 2008, 30(10):889-892.
- [6] 邵林, 杨殿阁, 李兵, 等. 基于双目视觉和声阵列的声学摄像机的开发[J]. *仪器仪表学报*, 2009, 30(4):823-827.
SHAO L, YANG D G, LI B, et al. Development of acoustic vidicon based on binocular vision and microphone array [J]. *Chinese Journal of Scientific Instrument*, 2009, 30(4):823-827.
- [7] DOUGHERTY R P. Beamforming in acoustic testing[C]. *Aeroacoustic Measurements*, 2002: 62-97.
- [8] LEGG M, BRADLEY S. A combined microphone and camera calibration technique with application to acoustic imaging[J]. *IEEE Transactions on Image Processing*, 2013, 22(10): 4028-4039.
- [9] MAYNARD J D, WILLIAMS E G, LEE Y. Near field acoustic holography: I. Theory of generalized holography and the development of NAH [J]. *Journal of the Acoustical Society of America*, 1985, 78 (4): 1395-1413.
- [10] 周东旺, 李舜酩, 江星星, 等. 基于传递函数估计的近场声全息的噪声源识别[J]. *仪器仪表学报*, 2015, 36(12):2874-2880.
ZHOU D W, LI SH M, JIANG X X, et al. Noise source identification of near-field acoustical holography based on transfer function estimation [J]. *Chinese Journal of Scientific Instrument*, 2015, 36(12), 2874-2880.
- [11] 张志飞, 陈思, 徐中明, 等. 基于反问题的正则化波束形成改进算法[J]. *仪器仪表学报*, 2015, 36(8): 1752-1758.
ZHANG ZH F, CHEN S, XU ZH M, et al. Modified algorithm of the inverse problem based regularized beamforming [J]. *Chinese Journal of Scientific Instrument*, 2015, 36(8): 1572-1758.
- [12] 褚志刚, 段云炆, 沈林邦, 等. 奇异值分解波束形成声源识别方法[J]. *电子测量与仪器学报*, 2014, 28(11):1177-1184.
CHU ZH G, DUAN Y Y, SHEN L B, et al. Singular value decomposition beamforming method for sound source identification [J]. *Journal of Electronic Measurement and Instrumentation*, 2014, 28 (11): 1177-1184.
- [13] 鲍骏, 郭爱煌. 波束形成在水声定位中的应用[J]. *电子测量技术*, 2014, 37(11):42-45.
BAO J, GUO AI H. The application of beamforming in acoustic positioning [J]. *Electronic Measurement Technology*, 2014, 37(11):42-45.
- [14] YANG D, WANG Z, LI B, et al. Quantitative measurement of pass-by noise radiated by vehicles running at high speeds [J]. *Journal of Sound and Vibration*, 2011, 330(7): 1352-5764.
- [15] YANG D, WANG Z, LI B, et al. Development and calibration of acoustic video camera system for moving vehicles [J]. *Journal of Sound & Vibration*, 2011, 330(11):2457-2469.
- [16] HEIKKILA J, SILVÉN O. A four-step camera calibration procedure with implicit image correction [C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997: 1106-1112.
- [17] ZHANG Z. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330-1334.
- [18] WEN J, YANG D, WANG Z, et al. Camera orientation calibration technique for sound visualization system[C]. *22th International Congress of Sound and Vibration*, 2015:381-389.
- [19] MORSE P M, INGARD K U. *Theoretical acoustics*[M]. 2nd. Newyork: McGraw-Hill, 1968: 842-5249.

作者简介



温俊杰, 2013年于大连理工大学获得学士学位, 2016年于清华大学获得硕士学位, 现为广汽研究院 NVH 部声学包测试工程师, 主要研究方向为车辆噪声测控技术。

E-mail: wenjunjie901023@163.com

Wen Junjie received his B. Sc. degree in 2013 from Dalian University of Technology and M. Sc. degree in 2016 from Tsinghua University. Now, he is a sound package test engineer in AEI, GAG co. ltd. His main research interest includes vehicle noise detection and control.



杨殿阁 (通讯作者), 分别在 1996 年和 2001 年于清华大学获得学士学位和博士学位, 现为清华大学汽车工程系教授, 主要研究方向为车辆噪声测控技术、车辆动态测控技术、智能汽车及智能交通技术。

E-mail: ydg@mail. tsinghua. edu. cn

Yang Diange (Corresponding author) received his B. Sc and Ph. D. degrees both from Tsinghua University in 1996 and 2001, respectively. Now he is a professor in Department of Automotive Engineering, Tsinghua University. His main research interest includes vehicle noise detection and control, vehicle dynamic measurement and control, intelligent vehicle and intelligent transportation.