

DOI: 10.11992/tis.201708018

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20180417.0926.005.html>

# 音频感知哈希闭环检测的 无人机仿生声呐 SLAM 算法研究

胡文超, 孙新柱, 陈孟元

(安徽工程大学 安徽省电气传动与控制重点实验室, 安徽 芜湖 241000)

**摘 要:** 针对基于 SLAM 技术无人机在特定高度下构建二维经历图的优化问题, 在 RatSLAM 的基础上, 采用仿生声呐系统代替视觉传感器的 BatSLAM 模型和音频感知哈希闭环检测, 实现在暗光条件下的二维经历图优化。BatSLAM 模型通过绝对差值和 (SAD) 图像处理方法来更新仿声纳模板, 此方法仅仅判断二幅耳蜗图外观是否一致, 不存在几何处理和特征提取。由于耳蜗图在获取和传输过程中会产生各类噪音, 相同位置获得的耳蜗图具有一定的差异, 会导致构建的经历图失真。本文在 BatSLAM 的基础上, 使用音频感知哈希算法对耳蜗图进行特征提取, 并进行闭环检测。改进后的算法不仅考虑到外观, 而且考虑到相邻频带间的能量差异, 通过提高闭环检测准确率, 来改善经历图的失真问题。仿真实验表明: 采用基于音频感知哈希闭环检测的 BatSLAM 模型, 不仅实现了无人机特定高度和暗光条件下二维经历图的构建, 而且提高了闭环检测准确率, 从而改善经历图的失真问题, 实现经历图的优化。

**关键词:** 同步定位与地图构建; 音频感知哈希; 无人机; 闭环检测; 仿生声呐系统; 经历图; 绝对差值和; 耳蜗图  
**中图分类号:** Q811.211; TP751 **文献标志码:** A **文章编号:** 1673-4785(2019)02-0338-08

中文引用格式: 胡文超, 孙新柱, 陈孟元. 音频感知哈希闭环检测的无人机仿生声呐 SLAM 算法研究 [J]. 智能系统学报, 2019, 14(2): 338-345.

英文引用格式: HU Wenchao, SUN Xinzhu, CHEN Mengyuan. Research on BATSLAM algorithm for UAV based on audio perceptual hash closed-loop detection[J]. CAAI transactions on intelligent systems, 2019, 14(2): 338-345.

## Research on BATSLAM algorithm for UAV based on audio perceptual hash closed-loop detection

HU Wenchao, SUN Xinzhu, CHEN Mengyuan

(Anhui Key Laboratory of Electric Drive and Control, Anhui Polytechnic University, Wuhu 241000, China)

**Abstract:** Aiming at the problem of constructing two-dimensional empirical graph based on SLAM technology at certain heights, on the basis of RatSLAM, using the bionic sonar system to replace the visual sensor's BatSLAM mode and audio perceptual hash closed-loop detection, realizing two-dimensional empirical graph optimization under dark conditions. The BatSLAM model uses Sum of Absolute Difference (SAD) image processing methods to update the bionic sonar template. This method only judges whether the appearance of the two cochlear images is consistent, and does not have geometric processing and feature extraction. Because the cochlear images produce various noises during the acquisition and transmission, There are some differences in cochlear maps obtained at the same position. which can lead to the distortion of the constructed empirical map. This article is based on BatSLAM, The audio perceptual hash algorithm is used to extract features of cochlea, and according to closed-loop detection. The improved algorithm not only considers the appearance, but also considers the energy difference between adjacent bands. By improving the accuracy of closed-loop detection, to improve the experience map of the distortion problem. Simulation experiment shows, the BatSLAM model based on audio perceptual hash closed-loop detection is adopted, it not only realizes the construction of two-dimensional experience map under certain height and dark conditions of UAV, but also improves the accuracy of the closed-loop detection, improving the experience map of the distortion problem and implementing the optimization of experience graph.

**Keywords:** simultaneous localization and mapping; audio perceptual hash; unmanned aerial vehicle; closed-loop detection; Experience diagram; bionic sonar; sum of absolute difference; Cochlear map

的传感器获得周围环境的地图,同时确定自己在地图中的位置<sup>[1]</sup>。机器人SLAM可以实现二维空间的经历图构建,但是对于空中某一高度的经历图构建,例如利用SLAM技术用经历图来描述建筑物特定高度的外围轮廓,机器人SLAM已经不再适用。无人机可以在特定高度下飞行,利用无人机SLAM技术可以实现特定高度下二维经历图的构建。无人机SLAM问题应用最成功的是概率的方法,如卡尔曼滤波算法<sup>[2]</sup>、扩展卡尔曼滤波算法<sup>[3]</sup>、粒子滤波算法<sup>[4]</sup>等,文献<sup>[5]</sup>将小型无人机的SLAM问题分解成路径估计与环境地标估计两部分,分别采用粒子滤波器和扩展卡尔曼滤波器进行估计,提出了一种小型无人机的FAST SLAM算法,运算的复杂度大幅度降低,且能更好地表示非线性、非高斯运动模型,这些概率算法的优势是能够处理传感器与环境的模糊性,有很好的SLAM性能,但由于概率方法存在计算量大、搜索时间长、易陷于局部最优解等问题,故不能够完全解决全局地图构建和导航问题。惯性导航设备是无人机最常用的导航与地图构建方案,其主要使用惯性/卫星/磁传感器/气压高度计组合导航,但是无人机在楼群之间、大型建筑物内部或者山区等GPS信号较弱的地区飞行时,由于缺乏卫星导航系统,惯性导航系统的误差会不断发散,很难实现准确的导航与地图构建<sup>[6]</sup>。而视觉SLAM方法仅使用机载摄像头作为外部传感器,目前常用的是单目视觉SLAM<sup>[7]</sup>、双目视觉SLAM<sup>[8]</sup>、RGB\_D SLAM<sup>[9]</sup>、文献<sup>[10]</sup>使用四旋翼飞行器搭载RGB\_D传感器并进行SLAM系统研究,使用稀疏的RGB\_D SLAM方法,提高无人机SLAM系统的实时性的同时也达到了高精度的效果。然而视觉SLAM算法对于光照依赖性较强,必须在适宜的光照强度下,在黑暗或较暗的光强下,很难实现准确的同步定位与地图构建;为了解决此类问题,出现了声呐SLAM算法,声呐SLAM与视觉SLAM相比,声呐具有处理速度快、不受光线条件限制等优点,文献<sup>[11]</sup>利用AUV上的声呐传感器持续探测非合作目标并估计目标航迹的同时,利用探测到的AUV与目标间的相对信息修正其自身航位推算带来的累积估计误差,从而提高AUV的自定位精度。但是声呐SLAM的缺点在于获取的信息量较少,空间分布分散,其感知信息存在较大的不确定性,因此生物声呐逐渐作为研究的热点。生物声呐,能够实现在复杂的环境下进行导航,因为它从回声中提取更多的信息。文献<sup>[12]</sup>提出

可以模仿蝙蝠声呐的发射与接收过程,利用仿生声呐系统实现复杂环境下的定位与导航。

概率方法广泛应用于SLAM的同时,澳大利亚昆士兰大学Michael等<sup>[13]</sup>提出了一种基于复合位姿表征的啮齿动物海马区扩展模型(RatSLAM)。RatSLAM模型提出了一种位姿细胞网络,即连续吸引子网络,作为一种手段来进行路径积分。RatSLAM的路径积分是通过里程计获取速度与角度信息,同时使用视觉传感器对里程计得到的信息进行修正与更新,绘制精确的经历图。文献<sup>[14]</sup>已经证明在RatSLAM基础上,可以使用仿生声呐模板来代替RatSLAM系统中的视觉模板,用双耳平滑耳蜗图来代替视觉图像,可以实现在暗光或黑暗的条件下,使用BatSLAM模型来构建二维经历图。

BatSLAM模型通过双耳平滑耳蜗图作为仿生声呐模板构建二维经历图,使用绝对差值和(SAD)模块进行处理产生局部场景,即通过计算所有模板中的耳蜗图的像素亮度与当前耳蜗图的像素亮度之差并求和后与阈值相比较来进行仿生声呐模板的更新,但是此方法是基于外观的低分辨率、并且只需判断图像外观前后是否一致,不存在几何处理和特征提取。耳蜗图在获取和传输过程中会产生各类噪音,相同位置获得的回波信号构建的耳蜗图具有一定的差异,单独使用SAD模块处理会由于识别出错而导致二维经历图的失真。

文献<sup>[15]</sup>提出对视觉图像进行闭环检测,通常利用当前帧和历史帧的特征点匹配对数,通过设定的阈值来判断是否达到闭环状态。对于音频图像,出现了可以进行声音识别的音频感知哈希技术。音频感知哈希是指可以代表一段音频重要声学特征的基于内容的紧致数字签名,其主要目的是建立一种有效机制来比较两个音频文件的感知听觉质量<sup>[16]</sup>。本文以音频感知哈希检索时的误码率作为标准,将当前耳蜗图构建的音频感知哈希与模板库中耳蜗图构建的音频感知哈希相比较,当最小误码率小于某一阈值时,识别成功,完成闭环检测,从而实现经历图的优化。

## 1 算法原理与设计

采用一种可以在特定高度下进行绘制二维经历图的BatSLAM模型,使用仿生声呐模板来代替RatSLAM系统中的视觉模板,用双耳平滑耳蜗图代替视觉图像来构建二维经历图。使用仿生声呐系统采集回波信号,经过一系列处理后将二次

采样的双耳平滑耳蜗图作为仿生声呐模板,同时生成局部场景细胞。用位姿感知细胞来表示无人机在特定高度下的位置与方向。使用陀螺仪与加速度计信息来更新位姿感知细胞,使用仿生声呐模板对应的局部场景细胞来实现位姿感知细胞的修正,当使用仿生声呐模板识别失败时,则使用音频感知哈希模板进行位姿细胞的修正。考虑到单独使用音频感知哈希进行修正会降低算法的实时性,所以选择优先使用仿生声呐模板修正,辅助使用音频感知哈希修正,在提高闭环检测准确率的同时,实时性也不会有很大的影响。最后由位姿感知细胞、局部场景细胞与音频感知哈希的相互联系来实现二维经历图的构建与校正。改进后的 BatSLAM 算法的结构框图如图 1。

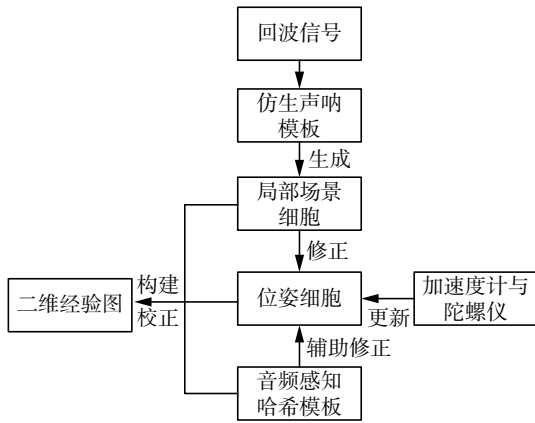


图 1 改进后的 BatSLAM 算法的结构框图

Fig. 1 Structure block diagram of the improved BatSLAM algorithm

### 1.1 仿生声呐模型

蝙蝠发出高频信号,并对反射的回波信号进行分析并进行导航。在头部相关传递函数(head function related transfer, HRTF)的基础上,加入发射器与接收器的位置信息,采用 echolocation related transfer function(ERTF)进行声源定位。左耳与右耳 ERTF 滤波器如式(1)、式(2)所示。

$$H_{\text{ERTF}}^L(f, \theta) = H_e(f, \theta) \cdot H_R^L(f, \theta) \quad (1)$$

$$H_{\text{ERTF}}^R(f, \theta) = H_e(f, \theta) \cdot H_R^R(f, \theta) \quad (2)$$

式中:  $H_e(f, \theta)$  表示发射器方位,  $H_R^L(f, \theta)$ 、 $H_R^R(f, \theta)$  分别表示左右接收器方位。频率为  $f$ , 角度  $\theta$  表示仿生声呐系统中处于特定高度的反射器的方位角。

文献[11]指出双耳 ERTF 差是综合了回声谱、双耳强度差(IID)和双耳时间差(ITD)的声源定位模型,包含了从目标物到反射器位置的详细方位信息,可以实现蝙蝠的回声定位。

在多个反射器的作用下,左右耳蜗接收到的信号如式(3)、式(4)所示。

$$x^L(t) = \sum_{i=1}^{ne} h_e^L(t; \theta_i) * s(t - \delta_i) \quad (3)$$

$$x^R(t) = \sum_{i=1}^{ne} h_e^R(t; \theta_i) * s(t - \delta_i) \quad (4)$$

式中:  $h_e^L(t; \theta_i)$ 、 $h_e^R(t; \theta_i)$  分别为左右耳 ERTF 滤波器的时域脉冲响应;  $s(t - \delta_i)$  为发射器发出的信号;  $\theta_i$  表示第  $i$  个反射器的方向;  $\delta_i$  表示由于第  $i$  个反射器所带来的时间延迟;  $ne$  表示反射器的数量。

为了分析这些双耳回波信号的生理学时频特性,使用 Gammatone 听觉滤波器组来模拟蝙蝠耳廓听觉频率响应,其时域表达式为

$$h^{gt}(t; f_c(n)) = t^3 \cdot e^{-2\pi b \cdot t} \cos(2\pi f_c(n) \cdot t) \quad (5)$$

式中:  $f_c(n)$  为中心频率;  $b$  为带宽。对于接收到的回波信号经过 Gammatone 滤波器组处理后,使用得到的 Gammatone 耳蜗能量谱来表示回波信号的时频特性,接下来进行一阶低通滤波,分别得到左耳与右耳耳蜗图,计算公式如式(6)、式(7)所示。

$$X_L(t) = x_L(t) * h^{gt}(t; f_c(n)) * u^{lp}(t) \quad (6)$$

$$X_R(t) = x_R(t) * h^{gt}(t; f_c(n)) * u^{lp}(t) \quad (7)$$

式中:  $X_L(t)$ 、 $X_R(t)$  分别为左耳与右耳耳蜗图,  $u^{lp}(t)$  为一阶低通滤波,  $x_L(t)$ 、 $x_R(t)$  为左耳与右耳接收的信号。

在 BatSLAM 模型中,将构建的耳蜗图模型作为仿生声呐模板,在使用耳蜗图作为仿生声呐模板时,必须注意对无人机位置细微的变化进行处理,及时进行二次采样的单声道耳蜗图使系统更加具有鲁棒性。这里使用高斯滤波器进行平滑处理,进一步排除细小位置改变带来的影响。二个平滑和二次采样的单声道耳蜗图连接起来,形成一个双耳平滑耳蜗图如式(8)所示。

$$\mathbf{X}^{gt,SS} = \begin{bmatrix} X_L^{gt} * G^S & X_R^{gt} * G^S \end{bmatrix} \quad (8)$$

式中:  $X_L^{gt}$  为左侧二次采样耳蜗图;  $X_R^{gt}$  为右侧二次采样耳蜗图;  $G^S$  为高斯平滑滤波器。

使用标准化双耳平滑耳蜗图,计算公式为

$$\bar{\mathbf{X}}^{gt,SS} = \frac{\mathbf{X}^{gt,SS}}{\sqrt{\sum_t \sum_f |\mathbf{X}^{gt,SS}|^2}} \quad (9)$$

将标准化双耳平滑耳蜗图作为当前仿生声呐模板,计算公式为

$$G_{cur} = \bar{\mathbf{X}}^{gt,SS} \quad (10)$$

### 1.2 位姿感知细胞

定高飞行的无人机在特定的位姿状态时,处于空间位姿感知细胞网络( $X, Y, \theta$ )中,通过连续吸引子网络(CAN)控制无人机位姿感知网络内部的活动。路径积分是通过转换每个位姿感知细胞基于无人机加速度计速度、陀螺仪角速度和

$(X, Y, \theta)$  坐标的活动来更新姿态细胞的活动, 同时使用仿生声呐模板的比较来保持和修正位姿信息, 从而实现位姿感知细胞的重定位。

### 1) CAN 模型

位姿细胞内部动态过程主要包括兴奋度更新, 对所有细胞的全局抑制。使用一个三维离散高斯分布来创建兴奋度权重矩阵  $\varepsilon$ , 如式 (11) 所示:

$$\varepsilon_{a,b,c} = e^{-(a^2+b^2)/k_m} \cdot e^{-(c^2)/k_n} \quad (11)$$

式中:  $k_m, k_n$  分别表示位置与方向的常数;  $a, b, c$  为  $x', y', \theta'$  的分布系数。

由于兴奋性造成的位姿感知细胞的活动变化如式 (12) 所示:

$$\Delta P_{x'y'\theta'} = \sum_{a=0}^{N_x} \sum_{b=0}^{N_y} \sum_{c=0}^{N_\theta} \varepsilon_{(a-x')(b-y')(c-\theta')} P_{abc} \quad (12)$$

式中:  $N_x, N_y, N_\theta$  为  $(x', y', \theta')$  空间中的位姿感知细胞三维矩阵元素。  $P_{abc}$  为位姿感知细胞的活动系数, 对于位姿感知细胞的活动变化的全局抑制如式 (13) 所示:

$$P_{x'y'\theta'}^{t+1} = \max[P_{x'y'\theta'}^t + \varphi(P_{x'y'\theta'}^t - \max(P)), 0] \quad (13)$$

式中: 抑制常数  $\varphi$  控制全局抑制水平且兴奋度水平被限制在非负数范围。

### 2) 直观的路径积分

该模型的路径积分方法是转换现有位姿感知细胞的活性而不是立即将当前活性复制传递。该过程使其性能独立于感知变化更新速率和加速度计速度与陀螺仪角速度, 不仅能生成更加准确的无人机飞行轨迹, 并且不需要参数调整。经过路径积分后, 被更新的位姿感知细胞活性为如式 (14) 所示:

$$P_{x'y'\theta'}^{t+1} = \sum_{a=\delta x'_0}^{\delta x'_0+1} \sum_{b=\delta y'_0}^{\delta y'_0+1} \sum_{c=\delta \theta'_0}^{\delta \theta'_0+1} P^t(x'+a)(y'+b)(\theta'+c) \quad (14)$$

式中:  $\delta x'_0, \delta y'_0, \delta \theta'_0$  为在  $x', y', \theta'$  方向被四舍五入后的整数偏差,  $\delta x'_0, \delta y'_0, \delta \theta'_0$  由式 (15) 计算得到:

$$\begin{bmatrix} \delta x'_0 \\ \delta y'_0 \\ \delta \theta'_0 \end{bmatrix} = \begin{bmatrix} k_x v \cos \theta \\ k_y v \sin \theta \\ k_\theta w \end{bmatrix} \quad (15)$$

式中:  $k_x, k_y, k_\theta$  是路径积分常数;  $v$  表示加速度计速度;  $w$  表示陀螺仪得到的角速度。

### 3) 仿生声呐模板修正位姿感知细胞

采用双耳平滑耳蜗图作为仿生声呐模板, 局部场景细胞由一个一维细胞阵列组成, 每个细胞单元对应一个特定的仿生声呐模板。在路径积分时, 由于位置和方向上的累计误差, 使用仿生声呐模板来进行修正。对于每个新的场景, 算法会建立新的仿生声呐模板, 同时建立局部场景细胞与位姿感知细胞间的联系。进行更新的连接强度

如式 (16) 所示:

$$\beta_{ix'y'\theta'}^{t+1} = \max(\beta_{ix'y'\theta'}^t, \lambda G_i P_{x'y'\theta'}) \quad (16)$$

式中:  $G_i$  为局部场景细胞的活动水平;  $P_{x'y'\theta'}$  为位姿感知细胞的活动水平;  $\lambda$  表示加权值。

对于每一个双耳平滑耳蜗图使用绝对差值和 SAD 模块进行处理产生局部场景。SAD 模块将当前耳蜗图与模板库中的耳蜗图进行比较, 通过计算所有模板中耳蜗图的像素亮度与当前耳蜗图的像素亮度之差, 如式 (17) 所示:

$$d_i = \sqrt{\sum (U_{\text{cur}} - U_i)} \quad (17)$$

式中:  $U_{\text{cur}}$  表示当前耳蜗图的像素亮度;  $U_i$  表示数据库中的第  $i$  个仿生声呐模板的像素亮度。如果像素亮度之差低于某一阈值  $\tau$ , 当前耳蜗图被识别出来, 对位姿感知细胞进行校正:

$$\Delta P_{x'y'\theta'} = \sum_i \beta_{ix'y'\theta'} G_i \quad (18)$$

如果像素亮度之差高于某一阈值  $\tau$ , 当前耳蜗图成为新的仿生声呐模板并加入到模板库中。阈值  $\tau$  为

$$\tau = \alpha_t \sqrt{\frac{\sum_{i=1}^{n_G} d_i^2}{n_G}} \quad (19)$$

式中:  $n_G$  为数据库中的仿生声呐模板的数量;  $\alpha_t$  为调整系统的比例因子,  $\alpha_t$  设置为 0.5。

## 1.3 音频感知哈希算法修正位姿细胞

### 1) 音频感知哈希

音频感知哈希是音频数据集到音频感知摘要集的单向映射, 将具有相同感知内容的音频数字表示唯一地映射为一段数字摘要。映射关系为 PH, 如式 (20) 所示, ph 感知哈希 (perceptual hash):

$$\text{ph} = \text{PH}(n), n \in N \quad (20)$$

生成的数字摘要称为感知哈希 (perceptual hash)。在感知哈希映射作用下, 音频信息间主观测度下的感知距离转化为感知哈希间的客观可计算的数学距离  $n \in N$ 。

### 2) 音频感知哈希的生成

仿生声呐系统将收集到的音频信号  $B$  作为感知哈希函数的输入, 使用一组具有 8 通道的 Gammatone 滤波器进行滤波, 将滤波后的数据根据频带高低进行排序, 由低到高分别为  $S_1, S_2, \dots, S_8$ 。使用 Gammatone 耳蜗能量谱来对回波信号的时频特性进行表示。将每一个频带的数  $S_n^{(r+1)}, S_n^{(r+2)}, \dots, S_n^{(r+m)}$  实行 90% 重叠后进行分帧, 每帧 20 ms (即 80 个采样点), 矩阵  $W$  作为能量谱矩阵, 对于能量谱矩阵里每一个元素  $W_{(n,L)}$ , 如式 (21) 所示:

$$W_{(n,L)} = \sigma \sum_{i=1}^{80} \sqrt{(S_n^{(r+m)})^2} \quad (21)$$

$$r = (L-1) \times 40; n = 1, 2, \dots, 8$$

式中:  $\{S_n^{(r+1)}, S_n^{(r+2)}, \dots, S_n^{(r+m)}\}$  为第  $n$  个频带的第  $L$  帧数据,  $m$  为长度; 每一帧都从  $r$  开始, 能量谱矩阵表示出每个频带的不同时间的能量分布情况。为了消除耳蜗能量谱存在的几何失真, 这里使用非负矩阵分解方法对能量矩阵进行无重叠分块。这里选择分块大小为  $8 \times 8$  的音频感知哈希。将其分解后, 每块为  $Q_k = \{Q(i) | i = 1, 2, \dots, 8\}, k = L/8$ 。信号经过二次采样以及低通滤波、噪声等处理后, 虽然 Gammatone 耳蜗能量谱的能量值发生了一定的变化, 但是相邻频带间能量差异依然没有较大改变。为了保持音频感知哈希的鲁棒性和区别性, 使用差分的方法不仅可以去除耳蜗能量谱值改变时所产生的误差, 还能缩减不同频带之间的相关性。这里使用二阶差分, 如式 (22) 所示:

$$R_k = \Delta^2 Q_k = Q(i) - 2Q(i+1) + Q(i+2) \quad (22)$$

$$i = 1, 2, \dots, 6$$

式中:  $\Delta^2 Q_k$  是对  $Q_k$  做二阶差分,  $R_k$  长度为 6。最后对得到的结果实行二值化处理, 音频感知哈希  $Y$  如式 (23) 所示。

$$Y(t, k) = \begin{cases} 0, & R_k(t) > 0 \\ 1, & R_k(t) \leq 0 \end{cases} \quad (23)$$

$$t = 1, 2, \dots, 6; k = 1, 2, \dots, L/8。$$

式中: 音频感知哈希生成过程如图 2 所示。

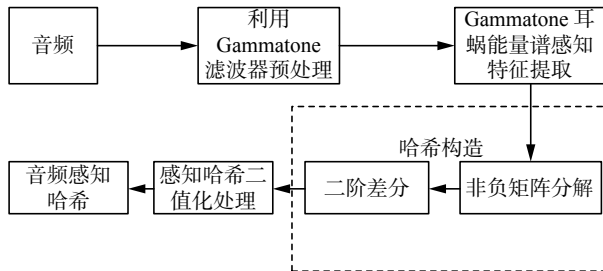


图 2 音频感知哈希生成

Fig. 2 Audio perception Hash generation

### 3) 音频感知哈希模板修正

对于每个新的场景, 算法会建立新的音频感知哈希模板。同时建立音频感知哈希与位姿感知细胞间的联系。进行更新的连接强度如式 (24) 所示:

$$\beta_{ix'y'\theta'}^{t+1} = \max(\beta_{ix'y'\theta'}^t, \lambda Y_i P_{x'y'\theta'}) \quad (24)$$

式中:  $Y_i$  指的是第  $i$  个音频感知哈希模板。这里采用误码率  $P = V_i/V$  来衡量任意二个音频感知哈希之间的相似水平,  $V$  表示待识别音频感知哈希码元个数,  $V_i$  表示模板库中的音频感知哈希进行对比时出现错误的个数。将当前音频感知哈希和数据库中音频感知哈希进行对比, 若低于阈值  $\gamma$  则

对位姿感知细胞进行修正:

$$\Delta P_{x'y'\theta'} = \sum_i \beta_{x'y'\theta'} Y_i \quad (25)$$

如果高于阈值  $\gamma$ , 当前音频感知哈希成为新的音频感知哈希模板并加入到音频感知哈希模板库中。

### 1.4 经历图的构建与校正

每个经验  $e_{x_i, y_i, \theta_i} = \langle E_i, P_{x_i, y_i, \theta_i}, Y_i, G_i \rangle$  关联经验能级  $E_i$ , 位姿细胞  $P_{x_i, y_i, \theta_i}$  和局部场景细胞  $G_i$ , 与音频感知哈希模板  $Y_i$ , 第  $i$  个经验如式 (26) 所示:

$$e_{x_i, y_i, \theta_i} = \langle E_i, P_{x_i, y_i, \theta_i}, Y_i, G_i \rangle \quad (26)$$

当前位姿感知细胞的活性或当前的音频感知哈希决定经验能级  $E_i$  的各组成要素, 如式 (27) 所示。

$$E_i = \begin{cases} 0, & \text{其他} \\ 2 - \frac{\sqrt{(x'_{pc} - x'_i)^2 + (y'_{pc} - y'_i)^2}}{r_a} - \frac{|\theta'_{pc} - \theta'_i|}{\theta_a}, & G_{\text{cur}} = G_i \text{ 或 } Y_{\text{cur}} = Y_i \end{cases} \quad (27)$$

式中:  $x'_{pc}, y'_{pc}, \theta'_{pc}$  为最大活性姿态细胞坐标;  $x'_i, y'_i, \theta'_i$  为与该经验相关的位姿感知细胞坐标;  $r_a$  为  $(x', y')$  平面的区域常数;  $\theta_a$  为  $\theta$  维上区域常数;  $G_{\text{cur}}$  为当前局部场景细胞;  $G_i$  为与经验  $i$  相关的局部场景细胞,  $Y_{\text{cur}}$  为当前的音频感知哈希,  $Y_i$  与经验  $i$  相关的音频感知哈希。

使用当前位姿信息、音频感知哈希和仿生声呐模板与经历图进行匹配时, 当到达经验过的地方时, 进行一次闭环检测, 并对经历图进行校正, 经验位姿改变如式 (28) 所示:

$$\Delta p^i = \alpha \left( \sum_{j=1}^{N_f} (p^j - p^i - \Delta p^{ij}) + \sum_{k=1}^{N_i} p^k - p^i - \Delta p^{ki} \right) \quad (28)$$

式中:  $\alpha$  为校正常数,  $N_f$  为经验  $i$  到其他经验的连接数,  $N_i$  为从其他经验到经验  $i$  的连接数。

## 2 实验与分析

### 2.1 硬件平台数据处理

本文是通过 I2C 协议来读取 MPU6050 数据, 获取的传感器数据经过预处理后, 经过互补滤波算法进行数据融合, 通过姿态解算获取四旋翼飞行器的姿态角与位置, 由 SPI 总线将数据发送到定高飞行的无人机控制器。使用气压计 MES5611 来获取气压数据。因为气压值与海拔高度成固定的比例, 通过气压计数据转换为海拔高度:

$$\text{Alt} = 44\,330 \cdot \left( 1 - \left( \frac{P}{P_0} \right)^{\frac{1}{5.255}} \right)$$

式中: Alt 是海拔高度, m;  $P_0$  为 1 013.25 hPa。

通过 3 个超声波换能器作为蝙蝠的耳朵和

嘴,3个超声波换能器按照“一字型”进行排列。“二耳”分别将采集的信息通过信号接收电路,经过AD芯片转换为数字信号传递给四旋翼飞行器的STM32主控芯片;STM32主控芯片发出信号,经过信号处理电路后使用“嘴”发出声波,仿生声呐系统采集音频信号并传输给上位机获得左右耳蜗图。

## 2.2 音频感知哈希构建基础

图3表示为2组不同回波信号的Gammatone耳蜗能量谱。分别含有了3段回波信号的最初版本、带噪声版本、高斯平滑处理版本。耳蜗能量谱的纵轴表示不同的频带,横轴表示不同时间,能量值越高对应的颜色也相应较深。由图3可以发现,两组音频信号获得的能量谱具有很大的不同,而相同音频虽然经过不同的处理,但是相邻频带间能量差异依然没有较大改变。所以可以通过耳蜗能量谱进行特征提取构建音频感知哈希。

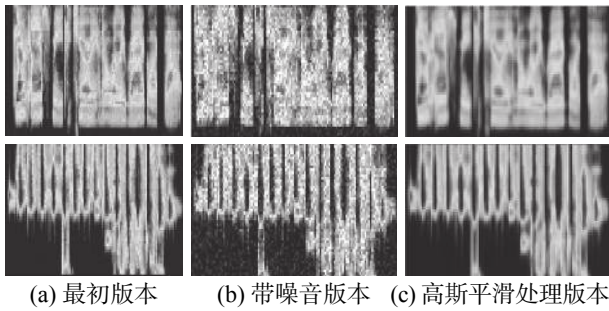


图3 最初音频和处理后的音频耳蜗能量谱比较

Fig. 3 Comparison of the energy spectrum between the initial and the processed audio cochlea.

## 2.3 不同噪音条件下改进算法与原算法的准确率-召回率对比试验

为了分析不同噪音条件下改进算法与原算法的闭环检测的准确率,这里选取典型的两种噪音,即高斯噪音与椒盐噪音。设计了外加高斯噪音、椒盐噪音和未添加噪音下的准确率-召回率对比试验。选取其中一个耳蜗图模板,得到不同噪音版本与原始版本的耳蜗图对比如图4所示。

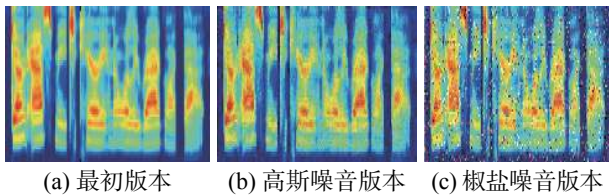


图4 不同噪音版本与原始版本耳蜗图对比

Fig. 4 Comparison of different noise versions and original versions of cochlea

为了分析不同噪音对于闭环检测准确性的影响,本文使用准确率 $P$ 、召回率 $R$ 来分析不同条件下2种算法的性能。其表达式如式(30)、式(31)所示。

$$p = TP/(TP + FP) \quad (29)$$

$$R = TP/(TP + FN) \quad (30)$$

式中:TP表示正确识别的闭环经验单元,FP表示错误识别的闭环经验单元,FN表示未识别的实际闭环经验单元。BatSLAM算法和基于音频感知哈希闭环检测的BatSLAM算法在不同条件下的准确率、召回率如表1、2所示。

表1 BatSLAM在不同条件下的准确率、召回率

Table 1 The accuracy and recall rate of BatSLAM under different conditions

不同条件	准确率/%	召回率/%
无噪音	99.06	87.50
高斯噪音	94.10	84.15
椒盐噪音	82.01	76.05

表2 基于音频感知哈希的BatSLAM在不同条件下的准确率、召回率

Table 2 The accuracy and recall rate of BatSLAM based on audio perception Hashi under different conditions

不同条件	准确率/%	召回率/%
无噪音	100	93.26
高斯噪音	98.23	90.45
椒盐噪音	97.15	86.30

通过表1、2可得,在无噪音条件下,两种算法的准确率大体相同分别为99.06%和100%,但是对于召回率,改进后的算法具有明显提高,未识别的实际闭环经验单元数量明显减少。当噪声干扰较小时,即高斯噪音,准确率提高程度不太明显,当噪音干扰较大时,即椒盐噪音,其准确率具有显著提高,因为改进后的算法不仅考虑到外观,而且考虑到相邻频带间的能量差异,从而提高了正确识别的闭环经验单元数量。当存在多种噪音时,基于音频感知哈希的BatSLAM算法具有明显优势,提高了闭环检测的准确率。

## 2.4 改进算法与原算法经历图对比实验

本文首先设计了硬件平台数据处理方案,实现了对MPU6050姿态数据、气压计MES5611的数据以及音频数据的采集与处理;然后将无人机定高飞行过程中获取的姿态数据与仿生声呐系统获取的音频数据上传到上位机进行了MATLAB仿真实验分析,来验证改进算法的可行性。

选取室外高度为1m的椭圆形方桌构建的环形空间作为实验场地,使用GPS获取位置信息作为基准。无人机分别使用BatSLAM算法和基于音频感知哈希的BatSLAM算法在暗光条件下以

1 m 的特定高度进行绕圈实验,起始点为  $A$ , 终点为  $B$ , 图 5 和图 6 中  $A$  与  $B$  点代表相同的实际位置。如图 5 所示, BatSLAM 算法构建的经历图在第一圈时, 由于声呐模板修正, 当到达终点  $B$  时, 自动定位到起点  $A$ , 继续导航, 当到达  $C$  点时, 由于耳蜗图在获取和传输过程中会产生各类噪音, 仅仅使用 SAD 处理方法会导致识别出错, 同时因为加速度计与陀螺仪漂移出现了  $C$ - $D$  间的误差偏移, 当到达  $D$  点时, 再次识别成功, 自动定位到  $E$  点, 继续导航。如图 6 所示, 将音频感知哈希算法融合进入的 BatSLAM 算法, 经历图在第一圈时, 由于声呐模板-音频感知哈希修正, 当到达终点  $B$  时, 自动定位到起点  $A$ , 继续导航, 虽然在导航过程会存在各种噪音干扰, 但是当使用仿生声呐模板识别失败时, 自动调用音频感知哈希进行修正, 因此第二圈轨迹与第一圈基本重合。因为声呐模板-音频感知哈希修正不仅考虑到外观, 而且考虑到相邻频带间的能量差异, 提高了闭环检测的准确性。基于音频感知哈希闭环检测的 BatSLAM 算法不仅完成了无人机特定高度和暗光条件下的二维经历图绘制, 而且提高了闭环检测准确率, 减小了轨迹偏差, 改善了经历图的失真, 实现了经历图优化。

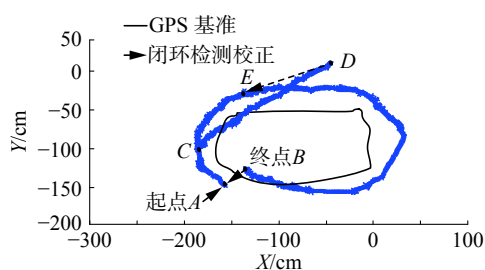


图 5 BatSLAM 经历图

Fig. 5 BatSLAM experience diagram

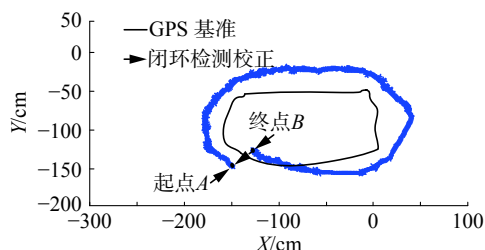


图 6 基于音频感知哈希闭环检测的 BatSLAM 经历图

Fig. 6 BatSLAM experience graph based on audio perception Hashi closed loop detection

### 3 结束语

本文提出的基于音频感知哈希闭环检测的 BatSLAM 模型不仅实现了无人机特定高度和暗光条件下的二维经历图绘制, 同时改善单独使用

SAD 模块处理所导致二维经历图失真问题, 而且提高了闭环检测的准确率, 减小了轨迹偏差, 从而实现经历图的优化。基于本算法的无人机 SLAM 技术具有适应性强的优势, 可以摆脱对于 GPS 的依赖, 不受光照强度的限制等, 可以适用于各种复杂的外界环境下, 如矿洞、楼群间、山区等。同时具有广泛的实际应用前景, 可以实现无人机自主导航的同时并在固定高度下对农作物进行灌溉; 模仿蝙蝠在洞穴中躲避空中钟乳石的原理, 实现在室内或室外复杂环境下躲避空中障碍物; 实现建筑物特定高度外围轮廓的二维经历图构建等。本文仅仅进行了简单地形环境下的 SLAM 二维经历图的构建并证明了算法的合理性与可行性, 复杂环境下的 SLAM 二维经历图的构建将是接下来的工作重点, 正在深入研究中。

### 参考文献:

- [1] MUR-ARTAL R, MONTIEL J M M, TARDÓ S J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. *IEEE transactions on robotics*, 2015, 31(5): 1147–1163.
- [2] HUANG G P, MOURIKIS A I, ROUMELIOTIS S I. A quadratic-complexity observability-constrained unscented Kalman filter for SLAM[J]. *IEEE transactions on robotics*, 2013, 29(5): 1226–1243.
- [3] 张凤, 黄陆君, 袁帅, 等. NLOS 环境下基于 EKF 的机器人定位研究[J]. *控制工程*, 2015, 22(1): 14–19.  
ZHANG Feng, HUANG Lujun, YUAN Shuai, et al. Reach on robot localization based on EKF in NLOS environment[J]. *Control engineering of China*, 2015, 22(1): 14–19.
- [4] 刘卫华, 王希彬, 周强. 基于容积粒子滤波的无人机 FastSLAM 算法研究[J]. *系统仿真技术*, 2014, 10(3): 234–238.  
LIU Weihua, WANG Xibin, ZHOU Qiang. FastSLAM algorithm based on cubature particle filtering algorithm for UAV[J]. *System simulation technology*, 2014, 10(3): 234–238.
- [5] 丛楚滢, 王从庆, 丁臻极, 等. 一种小型无人机的 FastSLAM 算法[J]. *华中科技大学学报(自然科学版)*, 2015, 43(S1): 420–423.  
CONG Chuying, WANG Congqing, DING Zhenji, et al. A FastSLAM algorithm for small unmanned aerial vehicle[J]. *Journal of Huazhong university of science and technology (nature science edition)*, 2015, 43(S1): 420–423.
- [6] 吕品, 赖际舟, 杨天雨, 等. 基于气动模型辅助的四旋翼飞行器室内自主导航方法[J]. *航空学报*, 2015, 36(4): 1275–1284.  
LYU Pin, LAI Jizhou, YANG Tianyu, et al. Autonomous

- navigation method aided by aerodynamics model for an indoor quadrotor[J]. *Acta aeronautica et astronautica sinica*, 2015, 36(4): 1275–1284.
- [7] 蒙山, 唐文名. 单目 SLAM 直线匹配增强平面发现方法[J]. *北京航空航天大学学报*, 2017, 43(4): 660–666.  
MENG Shan, TANG Wenming. Monocular SLAM plane discovery method enhanced by line segments matching[J]. *Journal of Beijing university of aeronautics and astronautics*, 2017, 43(4): 660–666.
- [8] 余小欢, 韩波, 张宇, 等. 基于双目视觉的微型无人机室内三维地图构建[J]. *信息与控制*, 2014, 43(4): 392–397.  
YU Xiaohuan, HAN Bo, ZHANG Yu, et al. Binocular stereo vision based 3D mapping for micro aerial vehicles in an indoor environment[J]. *Information and control*, 2014, 43(4): 392–397.
- [9] 吕科, 施泽南, 李一鹏. 微型无人机视觉定位与环境建模研究[J]. *电子科技大学学报*, 2017, 46(3): 543–548.  
Lü Ke, SHI Zenan, LI Yipeng. Visual localization and environment mapping for micro aerial vehicles[J]. *Journal of university of electronic science and technology of China*, 2017, 46(3): 543–548.
- [10] 王国胜, 郭峰, 刘峰. 无人飞行器视觉 SLAM 特征检测和匹配算法研究[J]. *计算机测量与控制*, 2015, 23(7): 2453–2455.  
WANG Guosheng, GUO Feng, LIU Feng. Research on feature detection and matching algorithm of unmanned aerial vehicles visual SLAM[J]. *Computer measurement & control*, 2015, 23(7): 2453–2455.
- [11] 王闯, 卢健, 黄杰. AUV 同时定位与跟踪算法研究[J]. *西安工程大学学报*, 2014, 28(5): 556–562.  
WANG Chuang, LU Jian, HUANG Jie. Algorithms research of simultaneous localization and tracking for AUVs[J]. *Journal of Xi'an polytechnic university*, 2014, 28(5): 556–562.
- [12] STECKEL J, PEREMANS H. Spatial sampling strategy for a 3D sonar sensor supporting BatSLAM[C]//*Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Hamburg, Germany, 2015: 723–728.
- [13] TUBMAN R, POTGIETER J, ARIF K M. Efficient robotic SLAM by fusion of RatSLAM and RGBD-SLAM[C]//*Proceedings of the 2016 23rd International Conference on Mechatronics and Machine Vision in Practice*. Nanjing, China, 2016: 1–6.
- [14] STECKEL J, PEREMANS H. BatSLAM: simultaneous localization and mapping using biomimetic sonar[J]. *PLoS One*, 2013, 8(1): e54076.
- [15] YU Hongshan, ZHANG Yuan, WANG Yaonan, et al. Fast and robust frontier line segment extracting method based on FCM for robot exploration[C]//*Proceedings of 2013 Chinese Automation Congress*. Changsha, China, 2013: 685–690.
- [16] 王伟, 陈志高, 孟宪凯, 等. 基于熵的音频指纹检索技术研究及实现[J]. *计算机科学*, 2017, 44(S1): 551–556.  
WANG Wei, CHEN Zhigao, MENG Xiankai, et al. Research and implementation of identifying music through performances using entropy based audio-fingerprint[J]. *Computer science*, 2017, 44(S1): 551–556.

#### 作者简介:



胡文超, 男, 1992 年生, 硕士研究生, 主要研究方向为四旋翼飞行器和仿生导航。



孙新柱, 男, 1974 年生, 副教授, 主要研究方向为运动控制系统分析与设计、智能机器人运动控制。发表学术论文 14 篇, 被 EI 检索 4 篇, 获安徽省第六届自然科学优秀学术论文三等奖 1 项。



陈孟元, 男, 1984 年生, 副教授, 主要研究方向为移动机器人地图构建及同步定位。主持安徽省高等学校自然科学研究项目等 10 余项。发表学术论文 30 余篇。