



深度学习下盲人避撞路径导航方法研究

摘要

传统导航方法只能检测出路径上存在的静止障碍物,无法检测出运动障碍物,为此提出基于深度学习的盲人避撞路径导航方法.收集语音信号,利用语音识别模型获取语音特征参数,根据语音特征参数识别出盲人输入的语音序列内容,确定盲人所要到达的目的地.构建障碍物检测模型,检测盲人所在位置与其目的地路径上障碍物的形状特征及其运动方向、速度,并计算初始位置与到达位置的距离.利用深度学习中的卷积神经网络规划出最优避撞路径,实现盲人避撞路径导航.实验结果显示,该方法检测出来的障碍物在 x 轴和 y 轴上的径向运动速度相差无几,可以实时跟踪监测障碍物的移动方向及运动速度.本研究方法所得速度与障碍物的实际运动速度基本一致,误差在 $0.2 \sim 0.4$ cm/s 之间,且在测试时间为 50 min 时,避障精准度达到 96.5%,能够实现最优避撞路径规划及导航.

关键词

深度学习;盲人导航;避撞;路径导航;语音特征;卷积神经网络

中图分类号 TP273

文献标志码 A

收稿日期 2021-03-27

资助项目 安徽省自然科学基金(2008085MF201);安徽省高校自然科学基金重点研究项目(KJ2019A1295);安徽省高校优秀青年人才支持计划(gxyq2020107)

作者简介

张海民,男,硕士,讲师,研究方向为图像处理、深度学习.hmzhang6@iflytek.com

¹ 安徽信息工程学院 计算机与软件工程学院,芜湖,241000

0 引言

目前,导盲手机、导盲鞋、盲人眼镜等盲人导航设备已经在实际中得到了广泛应用,但在导航过程中普遍存在障碍信息采集不完整、路径测量出现盲区等问题.为了解决这些问题,文献[1]提出了基于传感器信息融合的障碍检测及躲避方法,文献[2]提出了基于威胁度评估的机器人神经动态避撞方法,但是这些方法只能检测出路径上存在的静止障碍物,无法检测出运动障碍物,导致盲人避撞效果较差.本文针对上述方法中存在的问题,重点研究盲人避撞路径导航方法,利用深度学习中的卷积神经网络和语音识别模型,让盲人使用的导航工具具有障碍物识别和路径规划等多种功能,且在规划路径的过程中,可以免受其他外界因素的干扰,更好地帮助盲人躲避路径上存在的障碍物.

1 盲人避撞路径导航方法设计

1.1 语音信号处理

在导航语音识别过程中,首先需要收集语音信号,然后对收集到的语音信号进行处理,并提取语音信号特征,使得导航具有语音识别功能^[3].针对语音信号预加重处理,使盲人说话时可以产生 800 Hz 高频端,使信号的频谱变得平坦,可以提高导航过程中盲人语音识别的效率.设语音的频域为 z ,盲人语音输入序列为 H ,频域常数值为 μ ,即 $0.9 \leq \mu \leq 1.0$,通常 μ 取值为 0.98 或 1,则有:

$$H(z) = 1 - \mu z^{-1}, \quad (1)$$

式中, $H(z)$ 为 z 频域时的语音序列,即语音预加重处理结果.由于语音输入时间比较短,且输入过程中的频域特征和某些物理特征参数等几乎可忽略,因此将语音序列导入导航中时,需要在预加重的基础上,分帧识别语音序列特征^[4].用 n 表示帧数,其前后两帧交叠部分称为帧移,用 m 表示,在处理语音序列的过程中所产生的变换用符号 T 表示,则在 n 帧时的语音序列 Q_n 为

$$Q_n = T[H(m)n(z - m)], \quad (2)$$

式中, $n(z - m)$ 表示语音序列分帧时,乘以幅度为 1 的移动窗,即移动窗一定的取值函数.此时将式(1)代入,对语音序列进行二次预加重处理后,采用汉明窗作为语音信号处理的窗函数^[5].设汉明窗函数长度为 L ,则有:

$$w(z) = \begin{cases} 0.54 - 0.46\cos\left[\frac{2z\pi}{(L-1)}\right], & 0 \leq z \leq L-1, \\ 0, & \text{其他,} \end{cases} \quad (3)$$

式中, $w(z)$ 即为 z 频域时的 w 语音序列, 是已经处理过的 $H(z)$ 语音序列. 此时的 $w(z)$ 已经变得平滑, 且更接近于短时语音的频谱, 可以有效避免因截断效应造成的频谱泄露, 屏蔽了语音序列在录入的过程中存在的干扰信息, 能够有效解决语音频域间的相互干扰问题^[6]. 因此在式(3)处理好的语音频谱基础上, 提取语音频谱特征, 以此为后续的语音识别做准备.

1.2 基于语音识别提取语音特征参数

语音特征提取是指在语音识别模型的有限频域内, 获取一个语音输入序列的估计值, 将估计值与实际值进行误差分析, 取分析结果的平方和最小值. 其计算过程是将式(3)处理好的语音输入序列输入语音识别模型, 以此获取语音序列的特征参数. 但是在提取语音参数的过程中, 语音模型内会存在一定的误差^[7]. 因此设语音特征参数提取模型的允许误差值为 $\bar{H}(z)$, 语音序列 $w(z)$ 存在 P 个样点值, 语音识别模型预测值为 a , 系数为 $i, i = 1, 2, 3, \dots, n$, 则有:

$$\bar{w}(z) = \sum_{i=1}^P a_i w(z-i), \quad (4)$$

式(4)计算结果即语音序列 w 的样点值. 频域 z 是一个常数, 将 z 代入式(4), 即可得到语音识别模型预测误差值 e :

$$e(z) = w(z) - \sum_{i=1}^P a_i w(z-i), \quad (5)$$

式中, $e(z)$ 表示 z 频域时的语音识别模型预测误差值 e . 将式(4)和(5)联立, 可得语音识别模型预测均方误差 E :

$$E_z = \sum_z e^2(z) = \sum_z [w(z) - \bar{w}(z)]^2 = \sum_z \left[w(z) - \sum_{i=1}^P a_i w(z-i) \right]^2. \quad (6)$$

将式(4)、(5)和(6)联立, 可得语音特征参数提取方程:

$$\sum_z w(z)w(z-P) = \sum_{i=1}^P a_i \sum_z [w(z-P)w(z-i)]^2. \quad (7)$$

求解式(7), 即可得到语音特征参数, 该参数能够反映出语音序列的特征. 此时的语音序列平滑清晰、特征明显, 因此导航可以精准识别盲人输入的语

音序列^[8].

1.3 语音识别

经过式(3)加窗处理过的语音序列为 $w(z)$, 其在第 n 帧得到的语音序列为 $w_n(z)$, 在语音识别模型 λ 中, 通常会由一组参数分析语言序列内容:

$$\lambda = \{H, z, A, L, w\}. \quad (8)$$

在语音识别模型 λ 中, 汉明窗函数长度为 L , 即语音序列的长度, A 为式(7)的计算结果, 即语音序列所具有的全部特征. 根据语音识别模型中的语言序列识别参数, 有导航语音识别过程 B :

$$B_{Hw}(O) = P(O | H, w) = \frac{1}{(2H)^{P/2} |E_{Hw}|^{1/2}} \exp\left\{-\frac{1}{2}(O - e_{Hw})E_{Hw}^{-1}(O - e_{Hw})\right\}, \quad (9)$$

式中, O 为导航输出的语言序列, e_{Hw} 为式(5)计算出的输入语音序列和输出语音序列之间存在的误差值, E_{Hw} 为式(6)计算出的输入语音序列和输出语音序列之间存在的均方误差值^[9]. 此时导航通过语识别过程, 即可精准识别出盲人输入的目的地. 在此基础上只需确定障碍物特征、盲人位置、障碍物位置、最终目的地位置, 导航即可给出最优避撞路径.

1.4 建立障碍物检测模型

为了让导航可以精准检测出盲人位置到其目的地路径上存在的障碍物位置, 需建立障碍物检测模型^[10].

在此过程中, 利用建立的障碍物检测模型搜索障碍物位置等信息, 可为后续的盲人避撞路径导航奠定坚实的基础.

假设利用机器视觉技术获取的障碍物的长为 l , 宽为 h , 障碍物位置用坐标 (x, y) 表示, 则有障碍物运动过程 W :

$$W = (l, h, x, y). \quad (10)$$

根据障碍物形状, 在整体上可以划分为点和框两种, 因此在障碍物运动过程 W 的基础上, 建立点模型 r 和框模型 s :

$$\begin{cases} r = (x, y, \dot{x}, \dot{y}, \ddot{x}, \ddot{y}), \\ s = (x, y, v, j, \theta, \dot{\theta}), \end{cases} \quad (11)$$

式中, (\dot{x}, \dot{y}) 和 (\ddot{x}, \ddot{y}) 分别表示点模型 r 的速度和加速度, v 和 j 分别为框模型 s 的速度和加速度, θ 和 $\dot{\theta}$ 分别为框模型 s 的运动角度和运动角度的加速度. 利用障碍物检测模型, 可以检测障碍物拐角和边沿特征. 盲人在使用导航时, 导航会搜索最近障碍物点, 由点

模型 r 和框模型 s 同时检测障碍物是否具有拐角点. 当该模型具有拐角点时, 通过框模型 s 计算各边线性度, 并判断其是否属于线段, 属于线段则具有障碍物的拐角特征, 不属于线段时, 则不具有该特征. 当模型不具有拐角点时, 通过点模型 r 计算该障碍物的线性度, 并判断该线性度是否属于线段, 当检测结果为线段时, 具有障碍物的边沿特征, 不属于线段时, 则不具有该特征^[11]. 因此可以通过式(11)确定障碍物形状特征, 并根据式(10)检测障碍物运动速度和运动方向, 此时导航即可根据障碍物运动状态检测结果, 确定障碍物位置.

1.5 路径定位

根据 1.3 节识别出的盲人输入目的地语音, 确定盲人所要去的目的地位置和盲人位置以及两者之间的距离, 还有在盲人前往目的地路径上遇到的障碍物位置. 设这个过程中产生的周边环境信号强度集合为 $R = \{R_1, R_2, \dots, R_n\}$, 盲人位置的平面坐标集合为 $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 此时盲人距离其所要去的目标距离集合为 $\{d_1, d_2, \dots, d_n\}$, 上述集合中 n 为常数, $n = 1, 2, 3, \dots, N$, 则有障碍物位置函数 $f(x, y)$ ^[12]:

$$f(x, y) = \begin{cases} (x - x_1)^2 + (y - y_1)^2 = d_1^2, \\ (x - x_2)^2 + (y - y_2)^2 = d_2^2, \\ \vdots \\ (x - x_n)^2 + (y - y_n)^2 = d_n^2. \end{cases} \quad (12)$$

在求解式(12)过程中, 会受到周边环境信号强弱影响, 因此在接收盲人位置信号时会受到周边环境信号影响. 假设导航在接收盲人位置信号时, 会受到 C, G, V 这 3 个环境信号影响, 并且这 3 个信号圈两两相交, 则会出现 6 个交点, 如图 1 所示. 此时导航要从这 6 个交点中计算出盲人的位置.

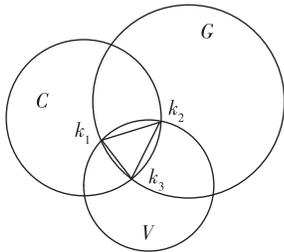


图 1 导航接收信号图

Fig. 1 Navigation receiving signal diagram

图 1 中, k_1, k_2, k_3 为环境信号干扰的 3 个交叉点, 并形成一个小三角形, 三角形的质心即为盲人位

置. 因此有盲人位置函数 $f(x_1, y_1)$:

$$f(x_1, y_1) = \begin{cases} (x_1 - x_C)^2 + (y_1 - y_C)^2 = d_C^2, \\ (x_1 - x_V)^2 + (y_1 - y_V)^2 = d_V^2, \\ (x_1 - x_C)^2 + (y_1 - y_C)^2 \leq d_C^2. \end{cases} \quad (13)$$

求解式(13), 即可得到障碍物位置 (x, y) :

$$\begin{cases} x = \frac{x_1 + x_2 + x_3}{3}, \\ y = \frac{y_1 + y_2 + y_3}{3}. \end{cases} \quad (14)$$

对式(12)降幂, 则有:

$$\begin{cases} 2(x_2 - x_1)x + 2(y_2 - y_1)y = \\ d_1^2 - d_2^2 - (x_1^2 + y_1^2) + (x_2^2 + y_2^2), \\ 2(x_3 - x_2)x + 2(y_3 - y_2)y = \\ d_2^2 - d_3^2 - (x_2^2 + y_2^2) + (x_3^2 + y_3^2), \\ \vdots \\ 2(x_n - x_{n-1})x + 2(y_n - y_{n-1})y = \\ d_{n-1}^2 - d_n^2 - (x_{n-1}^2 + y_{n-1}^2) + (x_n^2 + y_n^2). \end{cases} \quad (15)$$

将或(13)和(15)联立, 则存在:

$$C = \begin{cases} 2(x_2 - x_1)x + 2(y_2 - y_1)y, \\ 2(x_3 - x_2)x + 2(y_3 - y_2)y, \\ \vdots \\ 2(x_n - x_{n-1})x + 2(y_n - y_{n-1})y, \end{cases} \quad (16)$$

$$G = \begin{cases} d_1^2 - d_2^2 - (x_1^2 + y_1^2) + (x_2^2 + y_2^2), \\ d_2^2 - d_3^2 - (x_2^2 + y_2^2) + (x_3^2 + y_3^2), \\ \vdots \\ d_{n-1}^2 - d_n^2 - (x_{n-1}^2 + y_{n-1}^2) + (x_n^2 + y_n^2). \end{cases} \quad (17)$$

可以将式(15)简化成 $Cx - G$, 其中 x 为位置向量, 设 $\epsilon = Cx - G$ 为方程组误差向量, 且令 $F = |\epsilon|^2 = \epsilon^p \epsilon = (Cx - G)^p (Cx - G)$, p 为偏差系数, 当误差向量 ϵ 最小时, F 取最小向量 ζ, ψ , 求导 ζ, ψ , 并令其等于零, 则有:

$$\begin{cases} \frac{dF}{dx} = 2C^p Cx - 2C^p G = 0, \\ \zeta = (C^p C)^{-1} C^p G, \\ \psi = (C^p G)^{-1} C^p V, \end{cases} \quad (18)$$

式中, ζ 为盲人位置坐标向量^[13], X 值为盲人位置距离其所要去位置的距离. 根据 ζ 值可确定盲人所处位置, ψ 值可确定盲人所要前往的位置. 此时只需根据式(11)检测盲人所经过路径上的障碍物运动情况,

再根据式(14) 计算结果,就可确定盲人避撞路径。

2 基于深度学习的路径规划

深度学习是机器学习领域中一个新的研究方向,可以使机器模仿视听和思考等人类的活动,解决了很多复杂的模式识别难题,使得人工智能相关技术取得了很大进步。深度学习是一类模式分析方法的统称,就具体研究内容而言,主要涉及3类方法:卷积神经网络、自编码神经网络以及深度置信网络。

因卷积神经网络采用局部感受野、权重共享和降采样3种策略,降低了网络模型的复杂度,同时对于平移、旋转、尺度缩放等具有尺度不变性以及复杂度低等多种优势,因此本文采用深度学习中的卷积神经网络规划避撞路径。在此过程中需要使路径的线路长度达到最短^[14-15],进而构建由路径长度和障碍物模型两部分组成的等效代价函数^[16]。卷积神经网络结构如图2所示。

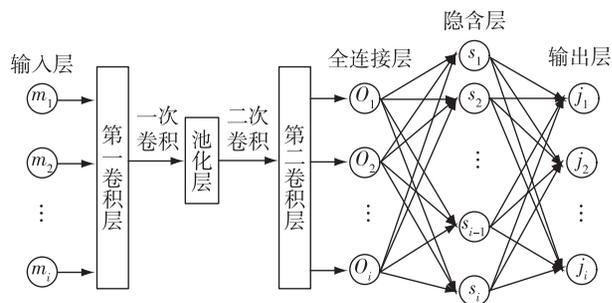


图2 卷积神经网络

Fig. 2 Structure of convolutional neural network

在卷积神经网络中,输入层为 M 、隐含层为 S 、输出层为 J 、全连接层为 O 。

本文所采用的卷积神经网络均来自于TensorFlow,代码可在<https://github.com/TimeIvace/MNIST-TensorFlow.git>中下载,程序名为train.py。卷积神经网络训练过程如图3所示。

初始化卷积神经网络权值,在此基础上输入训练数据,使得这些数据通过卷积层、下采样层、全连接层的向前传播,获取输出值。此时需要计算出卷积神经网络输出值与理想值之间的误差,当误差高于理想值时,需要将误差传回卷积神经网络中,按顺序获取全连接层、下采样层以及卷积层的误差,并将所有误差相加,得到卷积神经网络总误差;当误差等于或小于理想值时,立即停止训练。根据误差计算结果对卷积神经网络权值进行更新,并重新获取隐含层以及输出层各单元的输出。

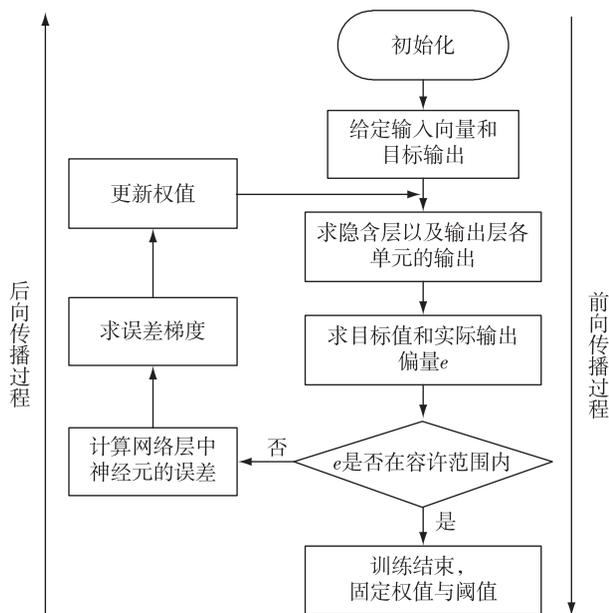


图3 卷积神经网络训练过程

Fig. 3 Training process of convolutional neural network

利用深度学习中的卷积神经网络进行最优路径规划时,会在输入层输入 f ,即障碍物位置,通过卷积层所含有的卷积核进行卷积作用。此时,会在输入层 M 至卷积层 W 之间,产生非线性映射:

$$q_i = g(\| \mathbf{u} - c_w \|), 1 \leq w \leq L, \quad (19)$$

式中, $\| \cdot \|$ 表示距离测量度, $g(\cdot)$ 表示卷积层向输出层传递的函数,即径向基函数, \mathbf{u} 表示输入的障碍物向量, c 表示径向基函数的中心函数, w 表示径向基函数的个数, c_w 表示第 w 个基函数的中心函数。此时设径向基的宽度为 δ ,则有深度学习从卷积层 W 至输出层 J 之间的线性映射函数 $g(x)$:

$$g(x) = \exp\left(-\frac{\mathbf{u}^2}{2\delta^2}\right). \quad (20)$$

设卷积层 S 传递信息至输出层 J 之间产生的连接权值为 ω_{SJ} ,实际输出路径为 D ,则有:

$$J_{\text{best}} = \omega_{SJ}r + \omega_{SJ}s + \omega_{SJ}\zeta + \psi. \quad (21)$$

在输入层至卷积层产生的非线性映射,以及卷积层 W 至输出层 J 之间的线性映射函数都可以看作激发函数,通过式(19)可以检测路径上的障碍物是否满足卷积层限制输入层作用条件。满足条件时,卷积层会输出路径;不满足条件时,卷积层会输出0。式(20)所得结果的值小于式(19)和式(21)所得的路径 D 为最优避撞路径。

3 实验与分析

在Android 5.0版本平台上编写导航程序,并根

据实验地点情况实时调试导航的性能、网络通信、语音播报、地图导航等部分功能.在 Android 5.0 版本平台中,导航的网络通信部分使用的内核由 Google 提供,且确定导航在 3G 网络下数据交互可以稳定进行.

将本文提出的盲人避撞路径导航方法记为实验 A 组,将上文提到的基于传感器信息融合的障碍检测及躲避方法^[1]以及基于威胁度评估的机器人神经动态避撞方法^[2]分别记为实验 B 组和实验 C 组,从障碍物运动速度检测和路径规划 2 个方面,分析 3 种盲人避撞路径导航方法检测障碍物运动速度、障碍物检测性能以及避撞路径规划的精度.为了保证实验测试结果的可靠性和真实性,进行 50 次实验,并将所得障碍物运动速度和语音识别结果按实验次序制成图表,直观上比较不同盲人避撞路径导航方法的差异.

3.1 避撞路径测试

在 Android 5.0 版本平台上,进行避撞路径规划实验,建立以 Android 5.0 版本平台位置为原点的平面直角坐标系,即 Android 5.0 版本平台位置为(0, 0),y 轴方向为平台位置的正北方向,x 轴方向为平台位置的正东方向.Android 5.0 版本平台的大小忽略不计.在平台上输入测试地点,其初始位置为 M,目的地位置为 W,在该段路上存在障碍物和移动障碍物 1,移动障碍物以 5 cm/s 的速度从上至下运动,分别应用 3 种方法规划最优避撞路线,最优路线是 Android 5.0 版本平台生成的,3 种导航方法规划出的最优路径如图 4 所示.

从图 4 中可以看出:实验 C 组会与静止障碍物出现碰撞情况规划,不属于正确路线;实验 B 组规划的避撞路径虽然躲避了静止障碍物,但却出现在移动障碍物前进的道路上;实验 A 组与最优路线最为接近,既躲避了静止障碍物,还绕过了移动障碍物可能出现的位置.由此可见,实验 A 组可以有效实现最优避撞路径的规划.

3.2 障碍物移动速度

选取长、宽、高均为 50 cm 的正方体积木作为障碍物,采用 3 种方法分别检测该积木的移动速度.积木由人为推动,做匀速直线运动向平台的坐标原点方向滑动,且用秒表记录积木的运动时间,并用 Android 5.0 版本平台所具有的数据记录处理功能记录处理数据.积木中心的起始位置为(400, 400),终止位置为(200, 200),即积木的中心坐标移动路径为(400, 400)至(200, 200).如图 5 所示.

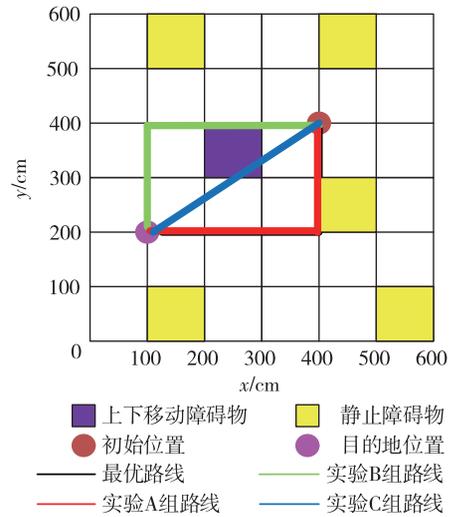


图 4 导航规划避撞路径

Fig. 4 Planning of navigation paths with collision avoidance

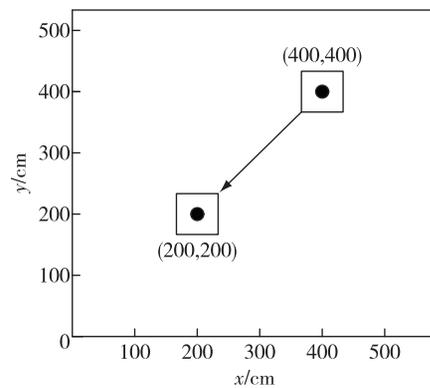


图 5 积木起始和终止位置

Fig. 5 Start and end positions of the moving building blocks

用秒表记录图 5 中积木的运动时间为 9.5 s,其实际运动速度为 33.6 cm/s.实验数据如表 1 所示.

表 1 积木障碍物测试数据

Table 1 Test data sheet of the moving building blocks

实验组	cm/s		
	x 轴速度	y 轴速度	合成速度
A	24	24	33.9
	24	23	33.4
	23	24	33.2
B	19	17	27.6
	18	18	27.5
	18	19	28.2
C	21	16	28.4
	20	18	28.6
	14	19	26.4

从表 1 中可以看出:实验 A 组检测障碍物在 x 轴和 y 轴上的径向运动速度相差无几,可以实时跟踪监测障碍物的移动方向及其在移动方向上的运动速度,且实验 A 组的合成速度与障碍物的实际运动速度基本一致,相差 0.2~0.4 cm/s;实验 B 组检测障碍物在 x 轴和 y 轴的径向运动速度虽然相差无几,可以实时跟踪监测障碍物的移动方向,但是合成速度与障碍物的实际运动速度相差较大,对障碍物的移动速度检测结果不准确;实验 C 组检测障碍物在 x 轴和 y 轴的径向运动速度存在较大差距,且其合成速度与障碍物的实际运动速度相差较大。由此可见,实验 A 组相较实验 B 组和实验 C 组,可以实时跟踪检测障碍物的运动,能够较为准确地测量移动障碍物的运动速度,因此实验 A 组可以满足盲人对避障路径的需求。所以,本研究的导航方法可以精准测量出移动中的障碍物,检测移动速度,正确识别语音输入,给出盲人避障的最优路径。

3.3 避障精准度

在上述实验的基础上对 3 组实验方法的避障精准度进行测试,精准度计算结果为准确避障次数与测试总数的比值,结果如表 2 所示。

表 2 避障精准度测试结果

Table 2 Test results of obstacle avoidance accuracy

测试时间/ min	避障精准度/%		
	实验 A 组	实验 B 组	实验 C 组
10	97.8	84.3	75.6
20	97.2	83.2	73.2
30	96.9	82.6	70.2
40	96.8	82.1	69.1
50	96.5	81.3	66.3

分析表 2 可知,与实验 B 组与实验 C 组相比,实验 A 组的避障精准度最高,说明采用本文方法进行盲人避障路径导航效果最好,可以有效避免盲人与障碍物发生碰撞。

4 结束语

本文研究的盲人避障路径导航方法,充分发挥了深度学习路径规划的优势,降低了传统导航方法对避障路径规划的难度,提高了语音识别精度,并将路径规划和语音识别的精度误差控制在最小范围内,且可以根据盲人位置实时规划出最优避障路径。本方法并未考虑语音识别的速度问题,以及在极端复杂的环境下对避障路径的检测问题。在今后的研

究中,应加强语音识别反映效率研究,以及极端复杂环境下的避障导航方法研究。

参考文献

References

- [1] 赵晓艳,王嘉铭,董燕丽,等.四足机器人的障碍检测及躲避系统设计[J].南方农机,2020,51(8):140-142
ZHAO Xiaoyan, WANG Jiaming, DONG Yanli, et al. Design of obstacle detection and avoidance system for quadruped robot [J]. China Southern Agricultural Machinery, 2020, 51(8): 140-142
- [2] 倪天,魏瑞轩,赵晓林,等.基于威胁度评估的机器人神经动态避障策略[J].机器人,2017,39(6):853-859
NI Tian, WEI Ruixuan, ZHAO Xiaolin, et al. Neural dynamic collision-avoidance strategy for robots based on evaluation of threat degree [J]. Robot, 2017, 39(6): 853-859
- [3] 霍彦明,姜峰,冯怡林,等.基于波束赋形技术与 V290pub 语音模块的新型降噪语音识别系统[J].现代电子技术,2020,43(10):46-50
HUO Yanming, JIANG Feng, FENG Yilin, et al. Novel denoising speech recognition system based on beam forming technology and V290pub speech module [J]. Modern Electronics Technique, 2020, 43(10): 46-50
- [4] 张瑞珍,韩跃平,张晓通.基于深度 LSTM 的端到端的语音识别[J].中北大学学报(自然科学版),2020,41(3):244-248
ZHANG Ruizhen, HAN Yueping, ZHANG Xiaotong. End-to-end speech recognition based on depth-gated LSTM [J]. Journal of North University of China (Natural Science Edition), 2020, 41(3): 244-248
- [5] 孙伟,宋如意,王宇航.视觉/惯性组合导航中的 SWF 与 MSCKF 对比研究[J].中国矿业大学学报,2020,49(1):198-204
SUN Wei, SONG Ruyi, WANG Yuhang. A comparative research of SWF and MSCKF on visual/inertial integrated navigation [J]. Journal of China University of Mining & Technology, 2020, 49(1): 198-204
- [6] 刘娟宏,胡彧,黄鹤宇.端到端的深度卷积神经网络语音识别[J].计算机应用与软件,2020,37(4):192-196
LIU Juanhong, HU Yu, HUANG Heyu. End-to-end speech recognition based on deep convolution neural network [J]. Computer Applications and Software, 2020, 37(4): 192-196
- [7] 卢洵波,李昕.特征融合的 VAD 方法在语音识别系统中的应用[J].电子测量技术,2020,43(7):129-136
LU Xunbo, LI Xin. Speech recognition system with combined features of voice activity detection [J]. Electronic Measurement Technology, 2020, 43(7): 129-136
- [8] 王建平,马兰兰,孙伟.基于声纹识别的智能照明语音识别算法研究[J].传感器与微系统,2020,39(6):37-40,44
WANG Jianping, MA Lanlan, SUN Wei. Research on intelligent lighting speech recognition algorithm based on voiceprint recognition [J]. Transducer and Microsystem Technologies, 2020, 39(6): 37-40, 44

- [9] 李云.基于改进 K 中心点的语音分组识别算法[J].电子设计工程,2020,28(10):152-155
LI Yun. Grouping voice recognition algorithm based on improved K center point[J]. Electronic Design Engineering, 2020, 28(10): 152-155
- [10] 姜正伟,李鹏旭,张斌,等.基于车速依赖静态输出反馈的自主汽车路径跟踪控制[J].南京信息工程大学学报(自然科学版),2021,13(1):41-50
JIANG Zhengwei, LI Pengxu, ZHANG Bin, et al. Path-following control of autonomous vehicles based on velocity-dependent static output-feedback strategy[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(1): 41-50
- [11] 宋广虎,冯全,海洋,等.采用深度学习法优化的葡萄园行间路径检测[J].林业机械与木工设备,2019,47(7):23-27
SONG Guanghu, FENG Quan, HAI Yang, et al. Vineyard inter-row path detection based on deep learning[J]. Forestry Machinery & Woodworking Equipment, 2019, 47(7): 23-27
- [12] Polvara R, Sharma S, Wan J, et al. Obstacle avoidance approaches for autonomous navigation of unmanned surface vehicles[J]. Journal of Navigation, 2018, 71(1): 241-256
- [13] Phan D, Yang J X, Grosu R, et al. Collision avoidance for mobile robots with limited sensing and limited information about moving obstacles[J]. Formal Methods in System Design, 2017, 51(1): 62-86
- [14] Fraga-Lamas P, Ramos L, Mondéjar-Guerra V, et al. A review on IoT deep learning UAV systems for autonomous obstacle detection and collision avoidance[J]. Remote Sensing, 2019, 11(18): 2144
- [15] Sangiovanni B, Incremona G P, Piastra M, et al. Self-configuring robot path planning with obstacle avoidance via deep reinforcement learning[J]. IEEE Control Systems Letters, 2021, 5(2): 397-402
- [16] 王毅,刘波,熊龙焯,等.基于深度学习的果园道路导航线生成算法研究[J].湖南农业大学学报(自然科学版),2019,45(6):674-678
WANG Yi, LIU Bo, XIONG Longye, et al. Research on generating algorithm of orchard road navigation line based on deep learning[J]. Journal of Hunan Agricultural University (Natural Sciences), 2019, 45(6): 674-678

Deep learning-based navigation path planning with collision avoidance for the blind

ZHANG Haimin¹ CHENG Fei¹

¹ School of Computer and Software Engineering, Anhui Institute of Information Technology, Wuhu 241000

Abstract Moving obstacles, unlike stationary ones, cannot be located or avoided by traditional navigation technologies. To address this, a collision avoidance navigation path planning strategy for blind people based on deep learning is proposed. First, a speech recognition system is used to collect speech signal and sort out the speech feature parameters, which is then analyzed to obtain the speech sequence input thus recognize the destination. Second, an obstacle detection model is constructed to detect the edge features as well as moving directions and velocities of the obstacles on the path to destination. Then a convolutional neural network of deep learning is exploited to plan the optimal path with collision avoidance. Finally, experiments are conducted and the results show that the radial velocity of the moving obstacles detected by this model is consistent with actual conditions, specifically, when the actual speed is 33.6 cm/s, the detected speed error is in the range of 0.2-0.4 cm/s, and the accuracy of obstacle avoidance reaches 96.5% when the test time lasts 50 min. It can be concluded that the proposed strategy can realize the optimal path planning and navigation with collision avoidance for the blind people.

Key words deep learning; navigation for the blind; collision avoidance; path navigation; speech feature; convolutional neural network