A23-028

具有深度估測距技術與智慧喚醒之自主移動跟隨護理工作車

Equipped with Depth Ranging Technology and Intelligent Wake-up Autonomous Mobile Follow-up Nursing Work Vehicle

隊伍名稱｜請跟著我！
Please Follow Me!
隊　　長｜林羿恒 / 中央大學電機工程研究所
隊　　員｜李俊霖 / 中央大學電機工程研究所

指導教授

蔡宗漢｜中央大學電機工程學系

臺灣大學電機工程學士、碩士、博士，現為中央大學電機工程系教授、中央大學資電學院智慧晶片及系統中心主任，並擔任智慧電子國家型科技計畫分項召集人。

研究領域

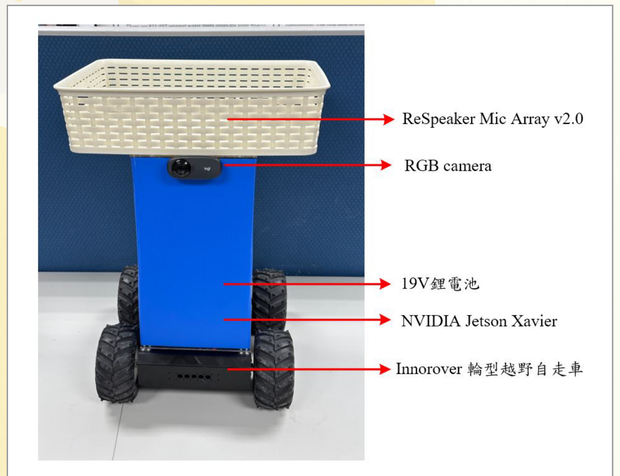
VLSI電路設計、SOC系統設計、多媒體數位訊號處理、視訊與音訊編碼、生醫信號處理、視覺及音訊神經網路

作品摘要

本作品充分扣合了智慧醫療的居家照護與無人載具應用的主題，隨著COVID-19疫情的爆發，導致醫療系統的超載，因為醫護人員人力的不足，醫護人員在工作上的勞務負擔越來越重。本作品從此方向進行設計，設計出一台具有深度估測距技術與智慧喚醒之自主移動跟隨護理工作車，目的用於取代現今各大醫院內傳統需要人力手推的護理工作車，成為護理人員的助手，減緩護理人員負荷和壓力，改善護理人力的問題。且考量在實際應用中，商品量產、實施與後續維護的成本，所形成的導入門檻，本作品以價格低廉的RGB攝影機配合深度估計演算法取代高規格高單價的深度攝影機（RGBD camera），並在Nvidia所推出的嵌入式硬體設備上運行，不需要高效能的硬體設備，這大大降低了實際應用的難易度，在未來可以廣泛的運用在各大醫院或是診所當中。改善護理人力的問題，打造智慧、科技、安全的照護服務，結合最新科技技術提升臺灣醫療照護人員的工作環境。

本作品結合深度估計技術與即時追蹤演算法，再加上智慧語音辨識指令及馬達控制，設計出一台可以自動跟隨並提供深度估計與物件辨識的自主移動跟隨護理工作車，本系統設計工作車共有三大部分：智慧語音辨識、深度估計技術、追蹤演算法、馬達控制。

系統開始執行時，先運行智慧語音辨識，判斷是否為喚醒詞，且為了使自主移動跟隨護理工作車更佳的客製化，我們使用speech to text演算法，若判斷為喚醒詞，系統根據聲源方向傳送控制訊號給為控制器進行轉向，此時自主移動跟隨護理工作車面向喚醒者，根據自主移動跟隨護理工作車上RGB攝影機獲取結果選定跟隨目標（護理人員），接著開始執行跟隨演算法，透過Siamese Networks提取目標Template與整個frame的深度特徵進行特徵向量距離計算，搜尋最近距離的目標，使系統可以準確地在複雜的挑戰下實時地獲得欲追隨目標的位置，並結合深度估計演算法獲取深度資訊。深度估計的任務類型為回歸任務，他不像普通的分類任務是藉由產生類別機率來得到分類結果。若是能夠有效地訓練影像深度估計網路，便能良好的輔助其他感測器，如光達（LiDAR）、毫米波雷達（mmWave Radar）等距離感測器來得到更加精確的測距值。深度估計甚至能夠取代其他價格昂貴的感測器，只需要一個解析度夠高的鏡頭，即可得到精確的空間距離資訊，藉此能使自主移動跟隨護理工作車準確地跟隨目標並保持距離。



圖一 護理工作車實體圖。

Abstract

This work fully integrates the themes of smart home healthcare and unmanned vehicle applications. With the outbreak of the COVID-19 pandemic, the healthcare system has been overloaded, leading to a shortage of medical personnel. Healthcare professionals are facing increasingly heavy workloads. This work is designed to address this issue by creating an autonomously moving nursing cart equipped with depth estimation and intelligent wake-up technologies. The goal is to replace the traditional manually operated nursing carts in hospitals and assist healthcare professionals, thereby reducing their burden and stress, and improving the issue of healthcare workforce shortage.

Considering the cost of mass production, implementation, and subsequent maintenance in practical applications, this work adopts a cost-effective approach by using low-cost RGB cameras combined with depth estimation algorithms instead of high-specification and high-priced depth cameras (RGBD cameras). The system runs on embedded hardware devices provided by Nvidia, eliminating the need for high-performance hardware equipment. This significantly reduces the difficulty of practical application and allows for widespread use in hospitals and clinics in the future. It aims to improve the healthcare workforce shortage and create intelligent, technological, and safe care services by combining the latest technology advancements to enhance the working environment for medical caregivers in Taiwan.

This work combines depth estimation technology with real-time tracking algorithms, along with intelligent voice recognition commands and motor control, to design an autonomously moving nursing cart that can automatically follow and provide depth estimation and object recognition. The system design of the nursing cart consists of three main parts: intelligent voice recognition, depth estimation technology, tracking algorithms, and motor control.

When the system starts running, it first activates the intelligent voice recognition module to determine if the wake-up word is detected. To achieve better customization of the autonomously moving nursing cart, we use speech-to-text algorithms. If the wake-up word is detected, the system sends control signals to the controller based on the direction of the sound source, causing the autonomously moving nursing cart to face the wake-up caller. The RGB camera on the autonomously moving nursing cart captures the results and selects the follow target (healthcare professional) based on the obtained data. Then, the tracking algorithm is executed, utilizing Siamese Networks to extract the target template and the depth features of the entire frame for calculating the feature vector distance. The system searches for the nearest target, allowing it to accurately track the desired target's position in real-time under complex conditions. In conjunction with the depth estimation algorithm, the system obtains depth information to accurately follow and maintain distance from the target.

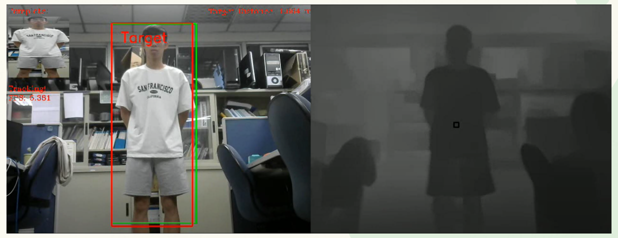


Fig.2 System tracking bounding box and depth estimation map.