



以毫米波雷達與虛實轉移深度強化 學習訓練之即時避障系統

1) Abstract & Motivation

自走車具有毫米波雷達感測器，在以往的例子中，多是使用Point Cloud之方式進行深度資料處理。然而在這樣的處理方式下，受限於毫米波雷達在環境中障礙體的複雜度造成許多雜訊，多個毫米波雷達可能互相干擾的狀況下，造成偵測錯誤或是無法偵測的情況，特別是在未知環境以及員工活動較頻繁的環境。

然而，毫米波雷達之波段，可於多霧、多懸浮微粒、粉塵及無任何光線之環境蒐集深度資料，非常適合戶外、草地、海上、地下之環境進行無人載具避障任務。

專題使用深度強化學習方式進行深度資料處理，透過在虛擬及現實環境下雙重收集資料，不僅使得資料之多樣性、完整性提升，也因為此技術不需使用相機等感測器進行影像處理，更能使儀器能在不受光線強弱之影響下，閃避障礙物。

指導教授：王學誠、楊春美

專題學生：張博凱

深度資料預處理之方法：

以毫米波雷達蒐集之原始資料，以深度學習方法設置限制條件，例如範圍、密度，濾除原始資料之雜訊。

將一或多個毫米波雷達所收集之深度資料合併且整理出三層高度（0~50、50~150、150~250公分）之360度環景深度資訊，其中每度內有一筆深度資訊，代表與環境障礙物之距離。已處理過之點雲資料，放入強化學習模型，得知避障需進行之移動。

深度資訊處理之方法：

以深度強化學習方法進行 obstacle avoidance (閃避障礙物)

利用兩組類神經網絡(actor net, critic net)，critic net 負責根據當下環境輸入（深度資料）來評估當下以及未來可能的狀態價值，actor net 負責根據當下環境輸入（深度資料）來輸出載具之動作值。

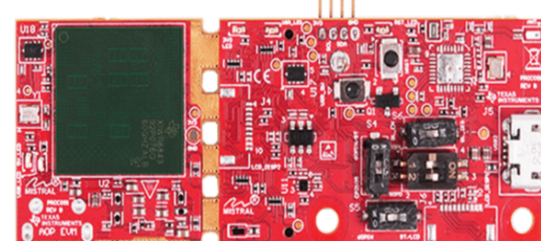
actor net 根據critic net所輸出之價值進行深度學習之梯度反向傳波。actor net 之梯度critic net則利用人為設計之獎勵方程式來計算當下以及未來之價值誤差，進行梯度反向傳波。

$$L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i | \theta^Q))^2 \text{ 價值誤差。}$$

獎勵方程式是獎勵載具直走之行為，直走之行為會有最大之獎勵，隨著轉彎程度增加，獎勵值便指數型下降，再依據深度資訊，越靠近障礙物會有線性的懲罰值，最後將計算出之獎勵值正規化。

2) Environment & Software

- ROS
- Point Cloud
- TI IWR6843 AOP
- TensorFlow



3) Functions & Approach

深度資料蒐集方法：

以一或多個毫米波雷達蒐集深度資料，裝設於無人載具，或是同時使用以上兩組毫米波雷達蒐集深度資訊，主要蒐集資訊為與側邊牆面之距離及速度、前方障礙物之相對位置。

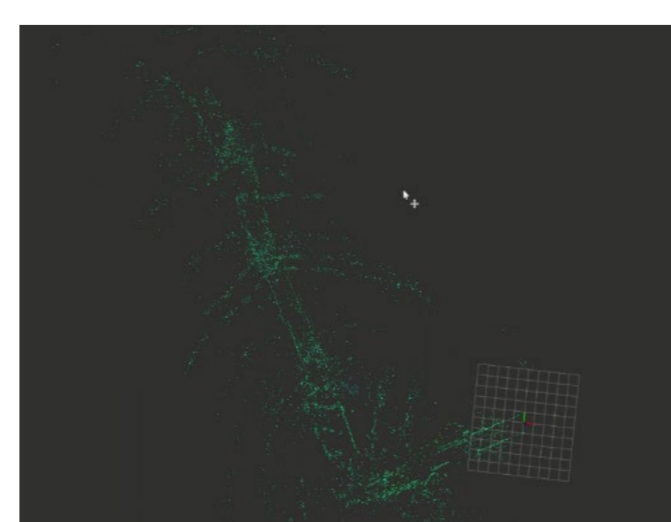
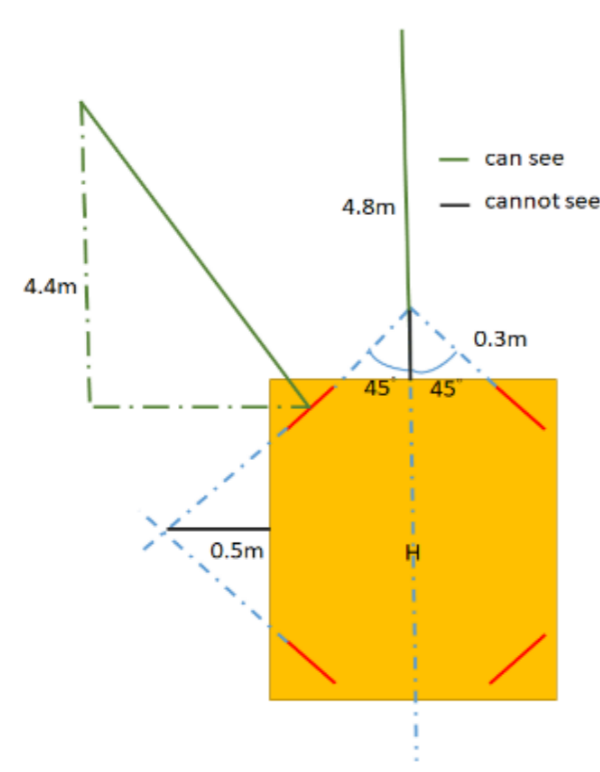
以多個毫米波雷達蒐集深度資訊，進行360度全景偵測，蒐集載具周圍360度之深度資訊。

4) Conclusion

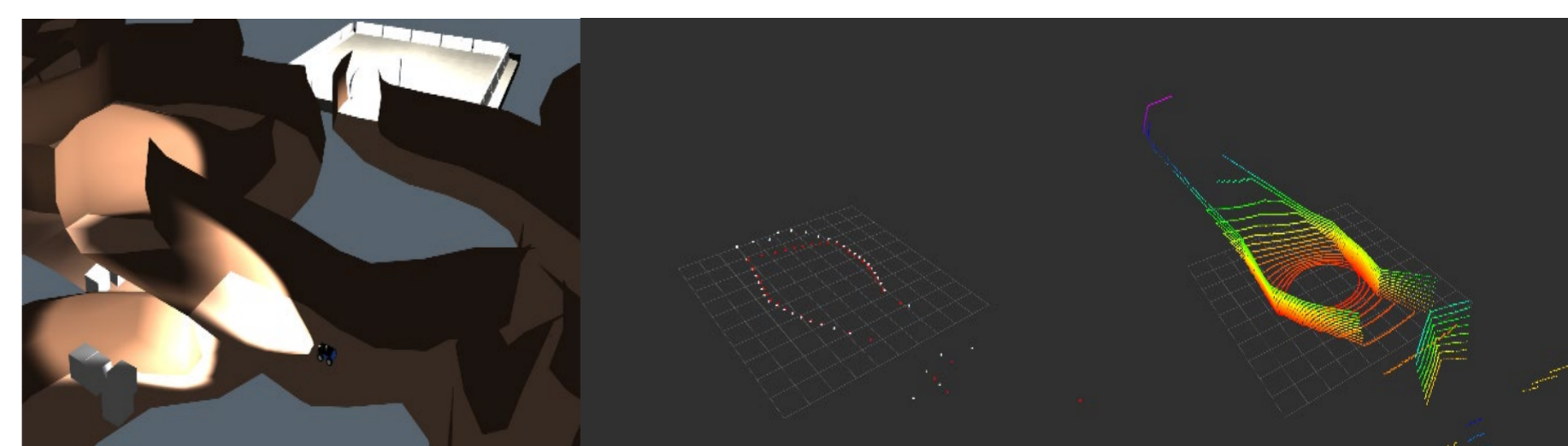
1. 以一個或多個毫米波雷達蒐集深度資訊。
2. 多個毫米波雷達進行全景360度蒐集資訊
3. 透過深度強化學習進行障礙物閃避
4. 使用現實及虛擬的方式進行資料蒐集與訓練
5. 可以即時學習之系統，包含毫米波雷達及深度強化學習模型。



(a) 於載具45°角位置裝設毫米波雷達



(b) 工五走廊 Point Cloud Map



(c) 於虛擬環境進行深度強化學習模型訓練