

一、指導老師：吳世弘 老師

二、組員：蘇仕翔（11027013）、顏崇哲（11027097）

陳志中（11027099）、陳憲彥（11027103）

三、系統環境：

軟體：Visual Studio Code、Anaconda、RealVNC Viewer、Python

硬體：人型機器人、樹莓派第四版（Raspberry Pi 4B）、喇叭、語音模組、無線網卡

通訊設備/協定：Wi-Fi/TCP

四、系統功能與特色：

(一)功能

1. 透過 GPT-SoVITS 做語音克隆：

透過已經訓練好的語音克隆模型對文字進行聲音克隆，生成出所需要的虛擬聲音

2. 透過 transformer (GPT-4o-mini) + MLP 做文字生成動作關節數值：

透過訓練好的模型進行文字轉關節數值，也就是當機器人說出某段文字時，會自動根據這段文字生成 16 個對應的關節數值

3. 透過 Socket 傳遞設定讓機器人進行多機互動：

(1)查詢機器人的 IP

(2)根據這個 IP 可以讓 Windows 系統針對不同的機器人進行控制

(3)然後把聲音克隆的音檔以及生成出來的關節數值分別傳遞給不同的機器人，就可以進行多機互動

(二)特色

透過文字或語音指令，機器人能生成對應的動作，多台機器人還能協同合作，發出自然的虛擬聲音，擺脫生硬的機械感。更有趣的是，它們能每次表演不同的相聲劇本，為互動增添更多生動有趣的體驗。



圖 1：兩台機器人說相聲



圖 2：機器人根據文字生成動作