

無人飲料站：具看、聽、說功能之雙手臂機器人
Unmanned Beverage Station:
Vision-and-Voice-Driven Dual-Arm Robot

Advisor: 連震杰 教授

Students: 謝卓熹 李彥平 薛巧妍

Robotics Lab.

0. Table of contents

1.1 Motivation and Objective

1.2 Scenarios

1.3 System Framework

1.4 System Setup

2.1 人機互動

2.2 物體位姿辨識

2.3 Robot Arm 夾取物體

3.1 Experimental Result

1.1 Introduction: Motivation and Objective

1. Motivation

- 傳統人工製作與調配飲品，流程勞力密集且耗費時間。
- 如何透過自動化流程優化工作效率，已成為自動化領域的重要議題。

2. Objective

- 打造出一套智慧語音飲料機器人，達成「自動化流程」服務。
- 對此，我們整合以下軟硬體技術：
 1. RGB-D相機: RealSense D415
 2. 人機互動 (HRI) + 機械手臂 (Robot Arm):
 - 1) 語音辨識、嘴型同步
 - 2) 杯子、水壺辨識位姿後抓取與倒飲料

1.2 Scenarios (1/3)

1. 喚醒機器人

- 1) 透過喚醒詞 “你好，小美” 來喚醒機器人



2. 透過語音指令送點飲料請求

- 1) 使用者透過語音指令控制機器人



1.2 Scenarios (2/3)

3. 準備飲料

- 1) 偵測與夾取水壺與紙杯，並同時利用大語言模型與使用者進行語音互動



您的咖啡大約需要一分鐘準備，在這過程中，要我幫您查詢天氣或講一個笑話嗎？



幫我查一下天氣

- 2) 倒飲料



今天的氣溫約35度....

1.2 Scenarios (3/3)

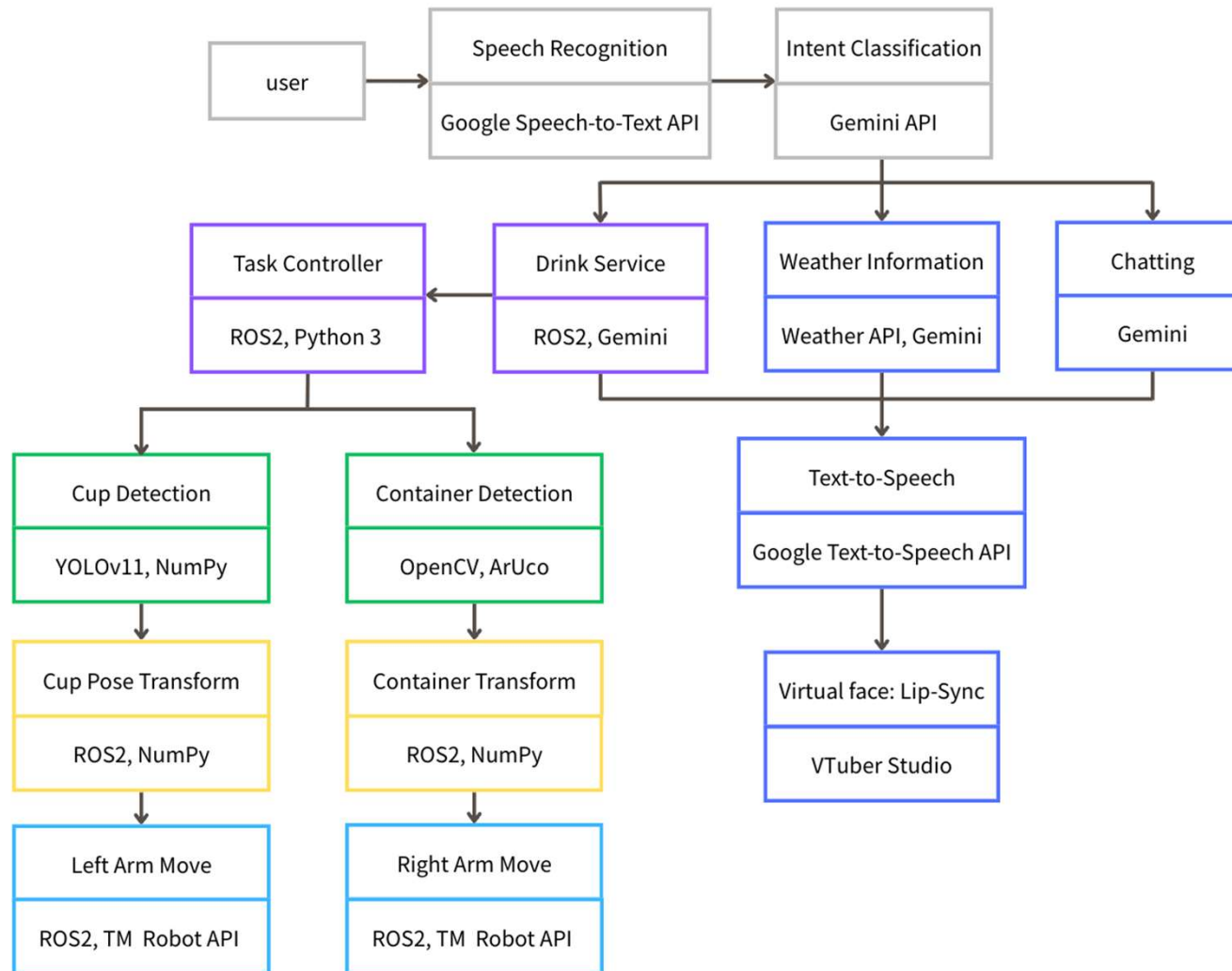
4. 遞交飲料

- 1) 將飲料遞交給使用者後，將水壺放回原處，且手臂復位



您的咖啡已完成，請拿取

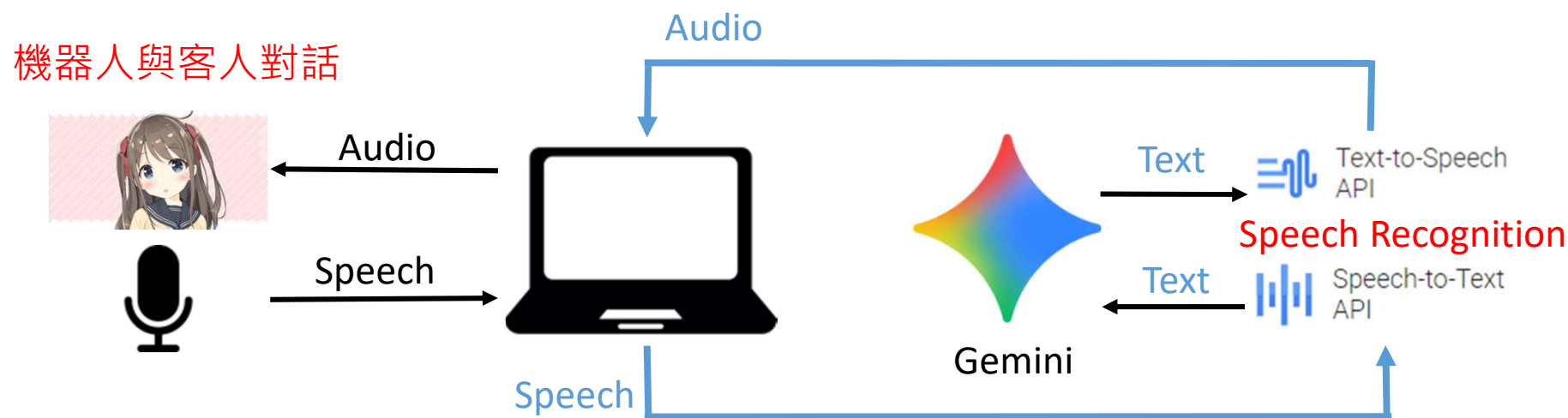
1.3 System Framework



2.1 人機互動

使用 PyAudio 進行錄音，使用 Google Cloud Speech API 進行文字語音間的轉換，並利用VTube Studio產生人物模型與脣形同步以增加互動體驗。

- Speech-to-Text：將接收到的語音轉換成文字辨識，以利後續大語言模型推理。
- Text-to-Speech：將大語言模型回復由文字轉換成語音與使用者互動。
- VTube Studio：人物模型與脣形同步，以增加使用者體驗。



2.2.1 物體位姿辨識-紙杯

1. 有向目標偵測 (YOLO-OBB Inference) :

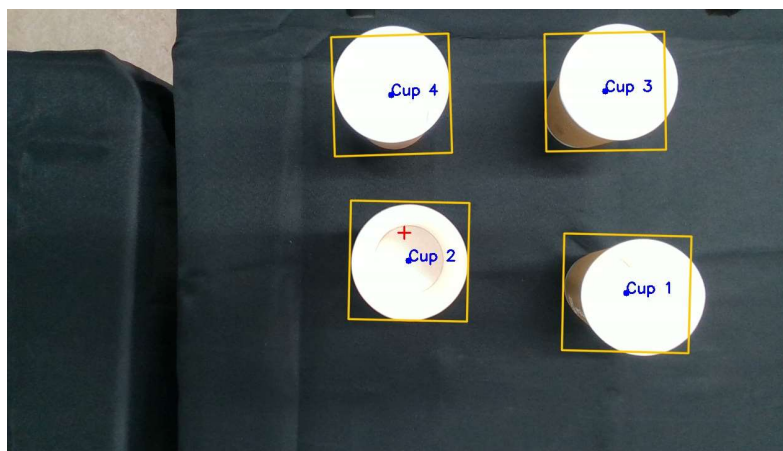
- 部署自行訓練之 YOLOv11-OBB 模型，精確提取紙杯之三維空間質心與二維旋轉角度。

2. 時序訊號濾波與幾何面積過濾 (Spatial-Temporal Filtering) :

- 結合 NumPy 矩陣運算進行鞋帶公式面積篩選，主動剔除背景雜訊與非目標物特徵。
- 導入時間序列多維中位數濾波器，有效抑制深度感測器之時序掉幀抖動。

3. 多物件空間幾何排序 (Spatial Sorting Mechanism) :

- 透過質心投影干涉檢查演算法，自動規劃多紙杯之最優抓取序列，實現雙手臂無碰撞連。



2.2.2 物體位姿辨識-容器

1. 二維標記追蹤 (ArUco Marker Tracking) :

- 利用 OpenCV 的 cv2.aruco 模組，即時定位固定於水壺 (Container) 表面之特定編碼標記。

2. 六自由度姿態解算 (6-DoF Pose Estimation) :

- 結合相機內參矩陣 (Camera Intrinsics) ，調用 OpenCV PnP 求解器 (cv2.solvePnP) 計算標記相對相機之旋轉向量 (rvec) 與平移向量 (tvec) ，鎖定水壺 6-DoF 位姿。

3. 動態路徑校正與容錯 (Dynamic Recalibration) :

- 將解算之實體坐標與手眼校正矩陣進行剛體轉換，提供即時空間參數；當物件遭遇外力產生位移時，強制刷新快取並重啟軌跡追蹤，維持抓取強健度。



2.3.1 Robot Arm 夾取物體

1. 雙臂非對稱任務分配 (Asymmetric Task Allocation)

- 左臂 (Left Arm)：專職夾取「紙杯」。通過得到的三維空間質心與二維旋轉角度，並以特定姿態夾取杯身，隨後移動至「接水預備點」。
- 右臂 (Right Arm)：專職夾取「飲料容器」。利用 ArUco 標籤推算出「水壺把手」的 6-DoF 側向姿態，進行側面夾取。

2. 兩段式防撞軌跡規劃 (Two-Stage Approach Trajectory)

- 預備與進入 (Approach & Grasp)：為避免夾爪碰撞桌面或物體，程式中皆會先計算目標物正上方或側邊的「預備點 (Approach Pose)」，到位後再以「直線運動 (LINE_T)」筆直進入「抓取點 (Grasp Pose)」。

3. 夾爪控制與硬體交握 (Gripper Control & IO Handshake)

- Robotiq 2F-85 夾爪控制：透過 ROS2 Action 異步控制夾爪開合度，針對不同物體（紙杯、把手）給予不同的夾取行程與力度（例如紙杯設定為 0.3 的開合比例）避免捏爆紙杯。



雙手臂到達預備點

$$X_c = \frac{(u - c_x) \cdot \text{depth}}{f_x}$$

$$Y_c = \frac{(\frac{u}{\phi} - c_y) \cdot \text{depth}}{f_y}$$

$$Z_c = \text{depth}$$

(杯子2D->3D 公式)

2.3.2 Robot Arm 倒水

4. 協同倒水與復位規劃 (Pouring Strategy & Homing)

- 多航點平滑軌跡：右臂夾取水壺後，會依序經過多個預先定義好的倒水航點（POUR WAYPOINTS），以直線運動平滑改變傾角完成倒水動作。
- 安全退回機制：任務完成後，採用與倒水完全相反的軌跡退回，再分別將容器與杯子放置回指定位置，確保雙臂在狹小空間內不發生干涉與碰撞。



3.1 Experimental Result

Demo video: <https://youtu.be/BZxMK9pSeqk?si=VZt4aBu18n8PBqlC>