**三、研究計畫內容（以中文或英文撰寫）：**

1. 研究計畫所要探索的課題。
2. 研究方法及進行步驟。
3. 預期達成目標及可行性評估指標。
4. 研究計畫所要探索的課題
5. 說明機器人辩識聲音來源方位的重要性
6. 說明機器人辨識聲音來源方位方法上的限制
7. 指出探索課題

何謂複雜環境?

何謂仿生計算?

聽覺週邊圖

腦幹聽覺訊號路徑圖

系統方塊圖

Nengo概念圖

1. 既有的方位辨識方法

4.1 麥克風陣列方法

4.2 雙耳辨識方法

4.3 Simple ASA

Multiple sources

Binaural cues

Onset detection

Harmonic structure

Streaming

4.4 仿生方法

聽覺訊號傳播處理路徑

HRTF

外耳

中耳

內耳

MSO

LSO

IC

ASA

* 1. 神經系統模擬軟體Nengo, SPA, 及 Spaun

Neural Engineering Framework

Nengo

SPA

1. 研究方法及進行步驟
2. HRTF from Brian
3. 聽覺周邊 : IPEM 模型
4. Jeffrey model in Nengo
5. LSD model in Nengo
6. IC model in Nengo
7. 預期達成目標及可行性評估指標

預期目標

本研究預期產生一篇投稿國際會議論文。

所完成的軟體系統，預備於下列四種類型測試場景，求得聲源方位俯仰(elevation)角與水平(azimuthal)角，畫出類似(Liu, Erwin, &Wermter, 2008)之圖9(d)的圖表(複製於本計畫書圖 1)，並探討其準確性及相關物理生理機制。



圖 2.聲音方位辨識結果呈現，圖中方塊大小代表Estimated Angle附近的神經脈衝數與整體脈衝速率的比例大小。本圖取自(Liu et al., 2008)文中的Fig. 9(d)。

測試場景類型1: 無反射(anechoic)環境，單聲源(距離頭部中心1.28公尺，俯仰角-20o至20o，間隔10o; 水平角 -90o至90o， 間隔30o，正前方為90o)，不同訊雜比(Signal to Noise Ratio，SNR)的聲音訊號強度(sound pressure level)及White Gaussian、冷氣聲、與交通背景雜訊。

場景1: 聲音訊號為弦波低頻 (100 Hz、100 x 151/15 Hz、100 x 152/15 Hz、…、1500 Hz)，預期ITD影響較ILD為大。

場景2: 聲音訊號為弦波高頻 (1200 Hz、1200 x 151/15 Hz, 1200 x 152/15 Hz、…、18,000 Hz)，預期ILD影響較ITD為大。

場景3: 聲音訊號為自然語音(五個英語字：”hello,” “look,” “fish,” “coffee,” “tea”，(Liu et al., 2008) 指出這五個單字包含相當大範圍的日常語音型態)。雖然我們人形機器人的頭部形狀與(Liu et al., 2008)的機器人頭部不同，語音特徵也有差異，但是仍可與其結果參考比較。

場景4: 國語數字1到10的發音、老人低沉呼救聲。

場景5: 重物落地聲(模仿老人摔倒發出的聲音)、門窗開啟聲、玻璃破裂聲。

測試場景類型2: 多重反射環境(長方形空房間，高3公尺，長寬依照擬考慮之回聲延遲時間調整，利用影像法求各次反射距離與角度)，單聲源(距離頭部中心1.28公尺，俯仰角-20o至20o，間隔10o; 水平角 -90o至90o， 間隔30o，正前方為90o)，不同訊雜比(Signal to Noise Ratio，SNR)的聲音訊號強度(sound pressure level)及White Gaussian、冷氣聲、與交通背景雜訊。

場景1: 聲音訊號為自然語音(五個英語字：”hello,” “look,” “fish,” “coffee,” “tea”)。

場景2: 國語數字1到10的發音、老人低沉呼救聲。

場景3: 重物落地聲(模仿老人摔倒發出的聲音)、門窗開啟聲、玻璃破裂聲。

測試場景類型3: 無反射環境，相對機器人運動的單聲源(不同速度及不同軌跡，以Robot Operating System 及 其中的機器人模擬軟體Gazebo呈現其3D場景)，不同訊雜比(Signal to Noise Ratio，SNR)的聲音訊號強度(sound pressure level)及White Gaussian、冷氣聲、與交通背景雜訊。以Robot Operating System 及其中的機器人模擬軟體Gazebo呈現3D場景。

場景1: 腳步聲(快、慢)，機器人靜止。

場景2: 聲音訊號為自然語音(五個英語字：”hello,” “look,” “fish,” “coffee,” “tea”)，機器人移動。

場景3: 國語數字1到10的發音、老人連續低沉呼救聲，機器人移動。

場景4: 重物落地聲(模仿老人摔倒發出的聲音)、門窗開啟聲、玻璃破裂聲，機器人移動。

測試場景類型4: 多重反射環境(長方形空房間，高3公尺，長寬依照擬考慮之回聲延遲時間調整)，相對機器人運動的單聲源(不同速度及不同軌跡)，不同訊雜比(Signal to Noise Ratio，SNR)的聲音訊號強度(sound pressure level)及White Gaussian、冷氣聲、與交通背景雜訊。以Robot Operating System 及其中的機器人模擬軟體Gazebo呈現3D場景。

場景1: 腳步聲(快、慢)，機器人靜止。

場景2: 聲音訊號為自然語音(五個英語字：”hello,” “look,” “fish,” “coffee,” “tea”)，機器人移動。

場景3: 國語數字1到10的發音、老人連續低沉呼救聲，機器人移動。

場景4: 重物落地聲(模仿老人摔倒發出的聲音)、門窗開啟聲、玻璃破裂聲，機器人移動。

可行性

評估指標

Liu, J., Erwin, H., &Wermter, S. (2008). Mobile robot broadband sound localisation using a biologically inspired spiking neural network. *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2191–2196. http://doi.org/10.1109/IROS.2008.4650760