

音視訊處理實驗室

Video/Audio Processing Laboratory

指導教授 | 張寶基 Pao-Chi Chang 教授
E-mail | pcchang@ce.ncu.edu.tw

研究領域

音視訊處理實驗室屬國立中央大學 通訊工程學系。

研究範圍包括語音音訊、影像視訊之編碼與處理，尤其重視在寬頻網路與無線通訊上之應用。近年來亦投入機器學習與深度學習於多媒體上之應用，如影像檢索、音樂辨識、特徵比對等研究。

Research areas include speech/audio coding, image/video compression, and video/audio applications on Internet and Wireless communications.

Recent studies have involved machine learning and deep learning in multimedia applications such as image retrieval, music recognition, and feature matching.

近期研究成果

- ◆深度學習應用在H.266編碼單元之快速演算法:
利用CNN網路學習之編碼模式預測演算法。
Intra Mode Prediction for H.266/FVC Video Coding based on CNNs
- ◆結合心理特徵與情緒標籤訓練之語音情感辨識技術:
使用了迴歸模型以及分類模型，迴歸模型用來進行心理狀態程度的預測。
Speech Emotion Recognition Based on Joint Training by Self-Assessment Manikins and Emotion Labels
- ◆基於平行膠囊神經網路之聲音事件偵測:
利用深度學習中的膠囊神經網路作為方法，提出應用於聲音事件偵測的系統。
Parallel Capsule Networks for Sound Event Detection
- ◆基於遞迴神經網路之聲學回聲消除技術利用 (Long Short Term Memory Network, LSTM) 及(Gated Recurrent Unit, GRU) 的遞迴神經網路。估計理想比例掩蔽 (Ideal Ratio Mask, IRM)來分離信號，從而達到去除回聲的目的。
Acoustic echo cancellation based on recurrent neural network
- ◆基於遞迴神經網路於多重裝置下之硬舉動作辨識及應用針對重訓類動作進行辨識
Deadlift Recognition and Application based on Multiple Modalities using Recurrent Neural Network