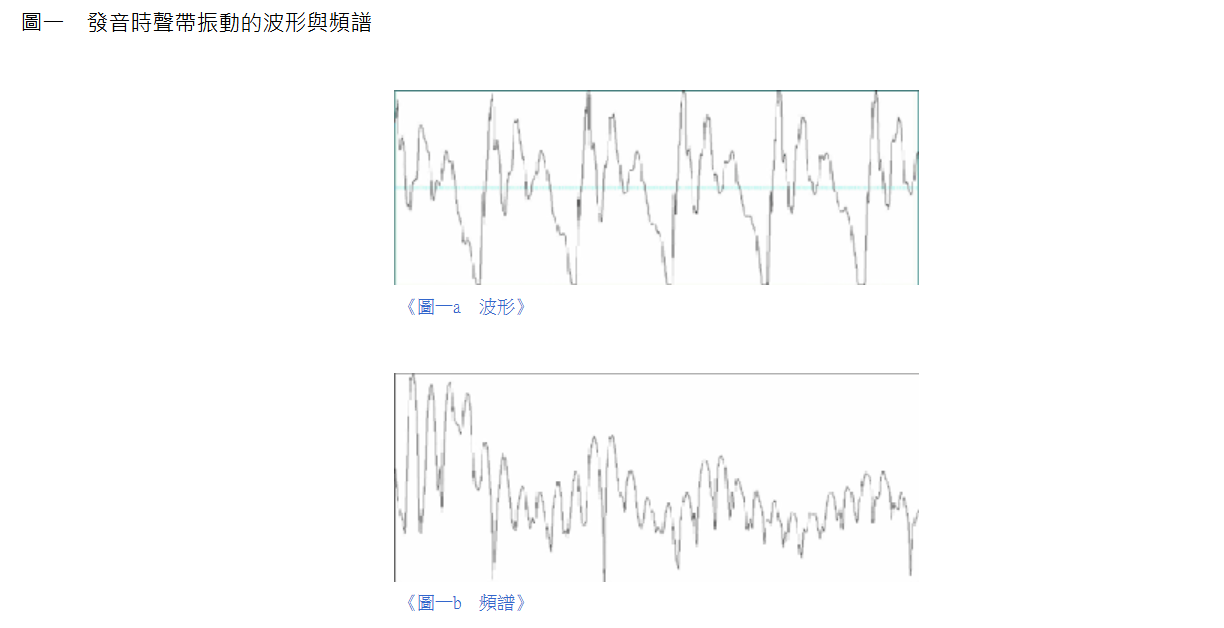
語音是否為真人

聲紋辨識的過程包含兩個階段：

1. 聲紋提取（voiceprint extraction）。

2. 聲紋比對（voiceprint comparison）。

在確認說話者的身分之前，要先有說話者的聲音語料，依說話者提供的聲音語料進行分析，並建立專屬他的聲紋模型。



一般在處理語音訊號時，會將音檔切割成小區段的方式來處理、進行分析，透過聲譜圖上的資訊來分析說話者的聲音頻率、音強、抑揚頓挫等建立專屬他的聲紋模型，並將其聲紋資訊存到系統裡。

當系統裡存有說話者的聲紋資訊後，其實就就能快速的進行一對一的說話者身份驗證，或是進行一對多的說話者辨認，從眾多人找出誰是說話者。

除了辨識身份，聲紋辨識其實也能應用在其他地方。現在也有許多研究團隊開發各種聲紋科技的應用，例如：透過大數據的聲音比對，由電腦判斷出鳳梨的好壞[5]、或是辨識青蛙叫聲的APP

<https://pansci.asia/archives/360983>

說話人確認的判斷準則（decision rule），就是將相似度比值拿來與一個事先設定的門檻值做比較，大於門檻值才認定這個受測語音是屬於宣稱人的，否則就認為說話的人是一個冒充者（impostor）。

上述的說話人辨識方法，基本上是對聲學特徵參數作長時間的統計，可以與文句不相關，但是辨識正確率通常較差。若是限制在某些語音範圍內，例如只講數目字，辨識結果會好一點。如果我們也做語音的辨識與分類，抽取特定的一些語音作比對，甚至於加上韻律參數，如音高軌跡等，作為比對的參數，通常能進一步提升辨識結果的準確率。

<https://www.ctimes.com.tw/DispArt-tw.asp?O=HJO3PA2EO4WAR-STDK>

音頻處理

工具/框架：

Librosa：提取音頻特徵（如 MFCC、Chroma 特徵、零交叉率等）。

<https://librosa.org/doc/latest/index.html>

PyDub：處理音頻文件的格式轉換和基本操作。<https://github.com/jiaaro/pydub>

OpenSMILE：進行聲學特徵提取，檢測音調、頻率等參數。