摘要

一首音樂的主旋律通常很容易就能被人類辨識，尤其是歌唱類型的音樂，在這類的音樂中歌唱音高通常就是主旋律。但是要利用電腦來直接辨識音樂中的歌唱音高是一件相當困難的事情，對於電腦來說背景音樂就像是干擾人聲的雜訊。在本論文中我們使用了一個基本方法，此方法是以一連串的頻譜分析演算法所構成，大部分的演算法目的都在於提高人聲並且降低音樂，藉此使動態規劃方法的音高追蹤準確度提高。但此方法在人聲端點與頻率快速變化區域容易得到錯誤的音高，所以我們利用反轉訊號的時間軸得到相異的結果，並使用疊合找出不穩定的音高，再輔以隱藏式馬可夫模型訓練的音高抽取方法，使用投票法來對不穩定的音高進行修正。在本論文的方法裡，我們改進了基本方法在弱點區域的準確度，使得整體辨識率得到明顯的提升。

Abstract

Human can easily recognize the main melody of a piece of music, especially of a song with singing voice, because pitch of singing voice usually represents the main melody of a song. However, it is not as easy for a computer to automatically detect the singing pitch from a song, because the background music acts as an interfering signal to the singing voice. In this thesis we propose a method that is composed of a series of spectrum analysis algorithms. Most of the existing algorithms focus on enhancing singing voice while reducing the background music so that the accuracy of singing pitch extraction can be improved. But these methods tend to yield incorrect pitch values near both endpoints of singing voice and sound segments with fast-varying frequencies. We therefore reverse the time axis of the song signal and overlap it with the original signal to find the segments that have unreliable pitch values. This method is assisted with another singing pitch extraction method incorporating hidden Markov models. A voting mechanism is adopted to justify and correct the unreliable pitch values. Experimental results show that the proposed method yields a better performance than the original baseline system in terms of raw pitch accuracy.