

摘要

本論文中，我們提出了一種有效改善哼唱選歌（query by singing/humming, QBSH）的整合架構。其中包含了三種不同的改進方法。第一種方法，是利用黃金比例搜尋法（golden section search）減少傳統線性伸縮（linear scaling）的比對耗時。第二種方法，是針對音高向量（包括使用者的哼唱以及資料庫歌曲）中的休止符加入不同的權重，以減少休止符對距離計算的影響。第三種方法，則是在比對音高向量時，利用序列誤差向量（sorted error vector）的概念，忽略一部分差異過大的距離值，而改使用剩餘的距離值作為比對距離。這是為了減少因使用者哼唱技巧不足或是音高追蹤錯誤，導致的短暫音高偏差所造成的影響。

我們提出的整合方案，不僅能夠縮短辨識所需的時間（方法一），同時也提升了辨識的正確率（方法二、方法三）。根據我們在 MIR-QBSH 資料庫與測試語料的實驗中，我們獲得了 21.4% 的誤差縮減比例（error reduction rate）並減少了 49.3% 的比對耗時。

關鍵字：音樂檢索、哼唱選歌、線性伸縮、黃金比例搜尋法、序列誤差向量

Abstract

This thesis proposes an improved framework for improving both the efficiency and the effectiveness of a query by singing/humming (QBSH) system. The proposed framework is based on three methods. Method 1 uses golden section search to reduce the computation time in traditional linear scaling (LS) algorithm. Method 2 assigns different weights for rests (in both database songs and in queries) so that these rests now have less effect on computing the weighted distance. Method 3 utilizes a sorted error vector to ignore the LS distances that are overly large and only considers the rest of the LS distances in the computation. This reduces the effect of pitch deviation in a short time span, probably due to the singer being out of tune or errors in pitch tracking.

The proposed framework improves the baseline system in both the computation time reduction (via scheme 1) and recognition accuracy (via schemes 2 and 3) of LS-based QBSH. Our experiment shows an error reduction rate of 21.4% in accuracy and 49.3% decrease in computation time on the MIR-QBSH dataset.

Keywords: music retrieval, query-by-singing/humming, linear scaling, golden section search, sorted error vector