



Trường Đại học Công nghệ Thông tin - ĐHQG TP.HCM

MÔ HÌNH SINH GIỌNG HÁT VỚI CÁC THUẬT TOÁN CHỈNH SỬA CAO ĐỘ

NGÔ CAO LỘC & ĐẶNG PHƯỚC SANG

{21521088, 21521377} @gm.uit.edu.vn

GIỚI THIỆU



Trong ngành công nghiệp âm nhạc hiện đại, việc chỉnh sửa âm thanh cũng như giọng hát đang ngày càng quan trọng. Như vậy làm thế nào với một người bình thường có thể có cho mình một bản audio hay mà không cần phải học nhiều về chuyên môn. Bài viết này sẽ đề xuất các phương pháp xử lý cao độ để có thể giúp tạo ra một giọng hát hay hơn và vẫn giữ được sự tự nhiên của giọng hát ban đầu

WHAT?

Chúng tôi đem đến một framework tạo ra những giọng hát hay đến người nghe, bao gồm:

- Mô hình sinh làm cốt lõi và học từ các bộ dữ liệu giọng hát.
- Đề xuất một thuật toán chỉnh sửa cao độ cho giọng hát.
- Đánh giá thực nghiệm khách quan và chủ quan.



WHY?

- Chỉnh sửa giọng hát ảnh hưởng rất lớn đối với ngành âm nhạc. Được tạo ra để nâng cao trải nghiệm khán giả, đổi mới trong ngành, cải thiện quá trình tạo ra một bản audio không mất nhiều công sức, tích hợp cộng đồng âm nhạc, tận dụng tiềm năng thương mại, và định hình tương lai của âm nhạc.
- Sự kết hợp giữa mô hình sinh giọng hát và thuật toán chỉnh sửa cao độ không chỉ là một bước tiến trong công nghệ âm nhạc mà còn là nguồn động viên cho sự sáng tạo và phát triển trong ngành.

XÂY DỰNG VÀ HUẤN LUYỆN

Từ các đặc trưng đã trích xuất cũng như cao độ đã chỉnh sửa, ta sẽ sử dụng kiến trúc mô hình CVAE kết hợp với thuật toán SADTW để có thể tạo ra được một giọng hát mới với cao độ đã chỉnh sửa và mang những đặc trưng riêng của giọng hát ban đầu



MÔ TẢ PHƯƠNG PHÁP

TRÍCH XUẤT ĐẶC TRƯNG

- Tìm hiểu các kiến thức biểu diễn âm thanh, đặc trưng của giọng hát
- Sử dụng các Encoder giải mã các đặc trưng như âm sắc, cao độ,...

CHỈNH SỬA CAO ĐỘ

- Áp dụng các thuật toán chuyên dụng (SADTW) bằng cách sử dụng thông tin hình dạng các biến đổi cao độ.

