**TỰ ĐỘNG PHÂN ĐOẠN TÍN HIỆU THÀNH TIẾNG NÓI VÀ KHOẢNG LẶNG**

**DỰA VÀO NĂNG LƯỢNG NGẮN HẠN CỦA TÍN HIỆU**

**Nguyễn Nhật Tùng, Ninh Hải Hoàng**

Khoa Công nghệ thông tin, Trường Đại học Bách Khoa, Đại học Đà Nẵng

nhattungnguyen.2kgl@gmail.com, ninhhaihoang@gmail.com

**Nhóm X, lớp HP: 18N15**

|  |  |  |  |
| --- | --- | --- | --- |
| **Điểm** | **Bảng phân công nhiệm vụ** | | **Chữ ký của SV** |
|  | Nguyễn Nhật Tùng (nhóm trưởng) | Viết báo cáo (tr. 1, 3), viết báo cáo kết quả thực nghiệm và kết luận (tr. A-b). |  |
|  | Ninh Hải Hoàng | Viết báo cáo (tr. 1, 3), đọc tài liệu, viết báo cáo về cơ sở lý thuyết,mã cài đặt thuật toán và tài liệu tham khảo (tr. X-y) |  |

Lời cam đoan: Chúng tôi, gồm các sinh viên có chữ ký ở trên, cam đoan rằng báo cáo này là do chúng tôi tự viết dựa trên các tài liệu tham khảo liệt kê ở cuối báo cáo. Các số liệu thực nghiệm và mã nguồn chương trình nếu không chỉ dẫn nguồn tham khảo đều do chúng tôi tự làm. Nếu vi phạm thì chúng tôi xin chịu trách nhiệm và tuân theo xử lý của giáo viên hướng dẫn.

TÓM TẮT— Việc tự động phân đoạn tín hiệu thành tiếng nói và khoảng lặng là công việc cơ bản mà quan trọng giúp xác định các khoảng có tiếng nói và khoảng không có tiếng nói từ đó giúp hệ thống khác trích xuất tín hiệu giọng nói và dùng cho việc nhận diện giọng nói sau đó. Thuật toán này chia tín hiệu âm thanh đầu vào thành các đoạn ngắn, sau đó dựa vào năng lượng của tín hiệu trên từng đoạn đem so sánh với giá trị năng lượng làm mốc xác định để phân đoạn tín hiệu đầu vào thành các đoạn tiếng nói và khoảng lặng (kq sẽ bổ sung sau).

Từ khóa— xử lý tín hiệu số, phân biệt tiếng nói và khoảng lặng, năng lượng ngắn hạn.

Mục lục

[I. ĐẶT VẤN ĐỀ 3](#_Toc55096468)

[II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN 3](#_Toc55096469)

[III. MÃ CHƯƠNG TRÌNH CÀI ĐẶT CÁC THUẬT TOÁN 3](#_Toc55096470)

[IV. KẾT QUẢ THỰC NGHIỆM 3](#_Toc55096471)

[V. KẾT LUẬN 3](#_Toc55096472)

[VI. TÀI LIỆU THAM KHẢO 3](#_Toc55096473)

# ĐẶT VẤN ĐỀ

Kể từ khi các kỹ sư có thể dựa vào các đặc trưng âm thanh để phân biệt các âm tiết, người ta đã nảy ra nhiều ý tưởng để nhận biết các loại âm tiết khác nhau, từ các âm tiết cơ bản tới các âm tiết phức tạp, nhằm mục đích nhận biết các từ ngữ dựa vào tín hiệu âm thanh được ghi âm.

Tuy nhiên để nhận biết được chính xác hơn, hệ thống nhận biết từ ngữ đó cần phải sử dụng các đoạn âm thanh chỉ chứa tiếng nói (hay các tín hiệu âm thanh có giá trị). Vì vậy việc tự động phân đoạn tín hiệu thành tiếng nói và khoảng lặng giúp xác định các khoảng có tiếng nói và khoảng không có tiếng nói từ đó giúp hệ thống khác trích xuất tín hiệu giọng nói và dùng cho việc nhận diện giọng nói sau đó.

Thuật toán được trình bày sau đây thực hiện việc tự động phân đoạn tín hiệu thành tiếng nói và khoảng lặng dựa vào năng lượng ngắn hạn (short-time energy) của tín hiệu âm thanh. Cụ thể: Thuật toán này chia tín hiệu âm thanh đầu vào thành các đoạn ngắn, sau đó dựa vào năng lượng của tín hiệu trên từng đoạn đem so sánh với giá trị năng lượng làm mốc xác định để phân đoạn tín hiệu đầu vào thành các đoạn tiếng nói và khoảng lặng.

Bài viết có bố cục như sau: Phần II trình bày tổng quan về cơ sở lý thuyết liên quan tới tín hiệu âm thanh, nguyên lý của các thuật toán, những vấn đề phát sinh trong thuật toán và cách khắc phục. Phần III ghi mã nguồn cách cài đặt thuật toán bằng Matlab. Phần IV trình bày kết quả thu được và các đánh giá kết quả đó khi áp dụng thuật toán lên các dữ liệu mẫu; so sánh thuật toán chuẩn hóa hiện tại với thuật toán chuẩn hóa khác (!). Cuối cùng là kết luận rút ra được trình bày ở phần V.

# LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN

# MÃ CHƯƠNG TRÌNH CÀI ĐẶT CÁC THUẬT TOÁN

# KẾT QUẢ THỰC NGHIỆM

# KẾT LUẬN

# TÀI LIỆU THAM KHẢO