**TIÊU ĐỀ BÁO CÁO**

**Phan Văn Vũ, Võ Minh Đức, Lê Đức Minh, Nguyễn Sỹ Tuấn Thành**

**Nhóm 11, lớp HP: 17Nh11**

|  |  |  |  |
| --- | --- | --- | --- |
| **Điểm**  (dành cho  GV ghi) | **Bảng phân công nhiệm vụ**  (SV ghi càng cụ thể thì GV càng dễ đặt câu hỏi và cho điểm mỗi SV) | | **Chữ ký của SV**  (mỗi SV ký xác nhận trước khi nộp báo cáo) |
|  | Lê Đức Minh | Cài đặt thuật toán phân đoạn giọng nói và khoảng lặng, kết quả thực nghiệm, tìm các yếu tố ảnh hưởng đến độ chính xác thuật toán. |  |
|  | Nguyễn Sỹ Tuấn Thành | Đọc tài liệu, cài đặt và viết báo cáo về thuật toán Zero-crossings, viết báo cáo phần đặt vấn đề, sơ đồ khối thuật toán. |  |
|  | Phan Văn Vũ (Nhóm Trưởng) | Đọc tài liệu và viết báo cáo phần kết luận, tìm các yếu tố ảnh hưởng đến độ chính xác thuật toán, kết quả thực nghiệm, làm slide thuyết trình PowerPoint. |  |
|  | Võ Minh Đức | Đọc tài liệu, cài đặt và viết báo cáo về thuật toán Năng lượng , viết báo cáo phần kết quả thực nghiệm, làm slide thuyết trình PowerPoint |  |

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 ghi rõ trong phần VII. 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— Tìm các đặc trưng tần số của tín hiệu là bài toán cần thiết trong xử lý tín hiệu âm thanh, đặc biệt là tín hiệu tiếng nói. Bài báo cáo này thực hiện việc tìm các đặc trưng tần số của tín hiệu tiếng nói trên miền thời gian và tần số thông qua các biên độ của ảnh phổ. Các thử nghiệm với tín hiệu của 5 nguyên âm (/a/, /e/, /i/, /o/ và /u/) cho thấy bảng thống kê 3 tần số formant. Kết quả thực nghiệm cũng cho thấy ta có thể sử dụng các đặc trưng tần số trên miền thời gian và miền tần số để phân biệt tín hiệu tín hiệu tiếng nói.

Từ khóa— Formant, STFT(Short time Fourier Transform), DFT, FFT.

Mục lục

[I. ĐẶT VẤN ĐỀ 3](#_Toc26101323)

[II. LÝ THUYẾT XỬ LÝ TÍN HIỆU TIẾNG NÓI VÀ CÁC THUẬT TOÁN 3](#_Toc26101324)

[A. Lý thuyết 3](#_Toc26101325)

[*1.* Sơ đồ khối thuật toán 3](#_Toc26101326)

[*2.* STFT ( Short-time Fourier Transform ) 3](#_Toc26101327)

[*3.* The Spectrograph & Spectral Analysis 4](#_Toc26101328)

[*4.* FFFFFFFF. 4](#_Toc26101329)

[B. Kích thước khác 4](#_Toc26101330)

[*1.* Phần tiêu đề 4](#_Toc26101331)

[*2.* Tác giả 4](#_Toc26101332)

[III. MÃ CHƯƠNG TRÌNH CÀI ĐẶT CÁC THUẬT TOÁN 4](#_Toc26101333)

[IV. KẾT QUẢ THỰC NGHIỆM 4](#_Toc26101334)

[A. Hình vẽ 4](#_Toc26101335)

[B. Bảng biểu 5](#_Toc26101336)

[V. KẾT LUẬN 5](#_Toc26101337)

[VI. TÀI LIỆU THAM KHẢO 5](#_Toc26101338)

# ĐẶT VẤN ĐỀ

Xử lý tiếng nói từ khi xuất hiện đã có một vai trò rất quan trọng trong cuộc sống của chúng ta. Cùng với sự phát triển của khoa học kỹ thuật, nhu cầu xử lý tiếng nói của con người ngày càng tăng cao. Xử lý tiếng nói có ứng dụng về nhiều mặt, về cơ bản có ứng dụng như nhận dạng tiếng nói, người nói, tăng chất lượng giọng nói và tổng hợp tiếng nói. Để làm được điều đó, việc nghiên cứu biên phổ của tín hiệu là rất quan trọng. Nhiều nghiên cứu về xử lý và nhận dạng tiếng nói đã chỉ ra rằng các tham số formant là ứng cử viên tốt nhất cho việc biểu diễn phổ. Các phương pháp xác định formant liên quan đến việc tìm kiếm các đỉnh trong các biểu diễn phổ, thường là từ kết quả phân tích phổ theo phương pháp STFT hoặc mã hóa dự đoán tuyến tính (LPC).

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

Phần này trình bày các lý thuyết có liên quan đến vấn đề cần giải quyết, cơ sở lý thuyết của các thuật toán, sơ đồ khối và các tham số quan trọng của mỗi thuật toán, phân tích các vấn đề của mỗi thuật toán và đề ra giải pháp khắc phục (nếu có). Nên dùng hình vẽ để minh hoạ ý tưởng.

## Lý thuyết

### Sơ đồ khối thuật toán

Màu sắc (Biên độ phổ)

Thời gian

Phân khung tín hiệu

Tín hiệu vào x(n)

Lấy DTFT trên từng khung

STFT của x(n)

Tần số

### STFT ( Short-time Fourier Transform )

#### Định nghĩa

* Biến đổi fourier thời gian ngắn là sự phân chia chuỗi thời gian các khối chồng nhau (overlaping blocks) có chiều dài bằng nhau và áp dụng biến đổi Fourier nhanh (FFT) cho mỗi khối một cách tuần tự.
* Đầu tiên tín hiệu được nhân với một hàm cửa sổ và sau đó thực hiện biến đổi Fourier, kết quả sẽ cho một biến đổi hai chiều (two indexed)

STFT():

STFT()=

#### Tính chất

Trong biến đổi Fourier thời gian ngắn (STFT) các hàm sử dụng trong mở rộng thu được bằng cách làm trễ và điều chỉnh ham cơ sở (t)

=

từ đó dẫn đến một dạng mở rộng :  
 STFT()= .

Hàm f(t) có thể khôi phục lại được theo công thức sau:

.d

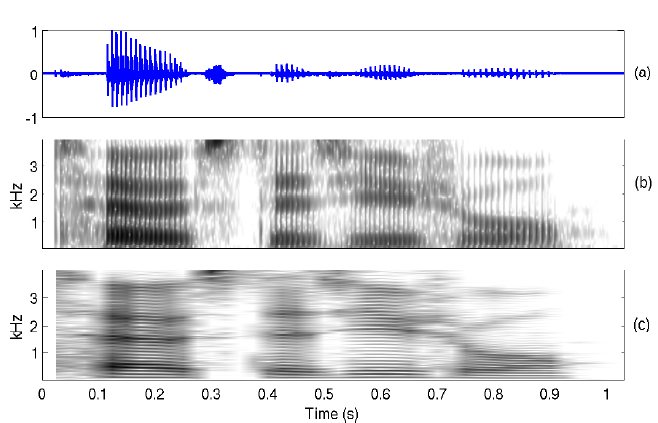
STFT không có tính chất bảo toàn năng lượng

Để thực hiện phương pháp này một cách tốt nhất thì yêu cầu phải chọn khoảng thời gian của các đoạn để phân chia sao cho tín hiệu ở mỗi khoảng thời gian đó có thể coi là tĩnh. Vì STFT chỉ xử lí số liệu tĩnh trên mỗi đoạn nên nó chỉ tính một cặp giá trị biên độ và pha

STFT là một phương pháp phổ biến và tính toán hiệu quả. Nhược điểm lớn nhất của phương pháp này là khi tín hiệu có một dải động lớn thì cụm tần số thấp. Trong trường hợp đó hướng tạp âm tần số cao có thể che cấu trúc tín hiệu tần số cao

### Narrowband và wideband:

Băng thông là dãy tần số tập trung hầu hết năng lượng (công suất) của tín hiệu. Nếu năng lượng tín hiệu tập trung vào một dãy tần số nào đó giữa tần số thấp và tần số cao thì đó là tin hiệu thông tải. Lúc này khái niệm băng hẹp (narrowband) được dùng để chỉ tín hiệu có băng thông F2 - F1 rất nhỏ (khoảng 10% hoặc nhỏ hơn) so với tần số trung tâm (F1+F2)/2. Ngược lại gọi là băng thông rộng.



Hình 1: Narrowband và wideband

### Spectrogram:

Spectrogram là một trong những công cụ cơ bản của phân tích phổ tín hiệu tiếng nói, trong đó nó chuyển đổi dạng sóng tín hiệu tiếng nói hai chiều thanh cấu trúc ba chiều (biên độ/tần số/thời gian). Trong đồ thị spectrogram, thời gian và tần số tương ứng là các trục ngang và dọc, còn biên độ được biểu diễn bởi độ đậm nhạt của màu biểu diễn. Các đỉnh của phổ tín hiệu xuất hiện là các dải nằm ngang màu đậm. Tần số trung tâm của các dải thường được coi là tần số formant. Các âm hữu thanh tạo ra các mảng dọc trong biểu đồ spectrogram bởi vì có một sự tăng cường biên độ tín hiệu tiếng nói mỗi khi thanh quản đóng lại. Nhiễu trong các âm vô thanh tạo ra các cấu trúc đậm hình chữ nhật và kết thúc ngẫu nhiên với nhiều đốm nhạt do sự thay đổi tức thì của năng lượng tín hiệu. Lược đồ spectrogram chỉ diễn tả biên độ phổ của tín hiệu mà bỏ qua các thông tin về pha bởi vì các thông tin về pha được cho rằng không có vai trò quan trọng trong hầu hết các ứng dụng liên quan đến tiếng nói.

## Thuật toán

### Phần tiêu đề

Tên bài báo dùng chữ in hoa như Template file này (Font chữ Arial 14pt, in đậm)

### Tác giả

Tên tác giả bao gồm cả cơ quan, địa chỉ email như Template file này

# MÃ CHƯƠNG TRÌNH CÀI ĐẶT CÁC THUẬT TOÁN

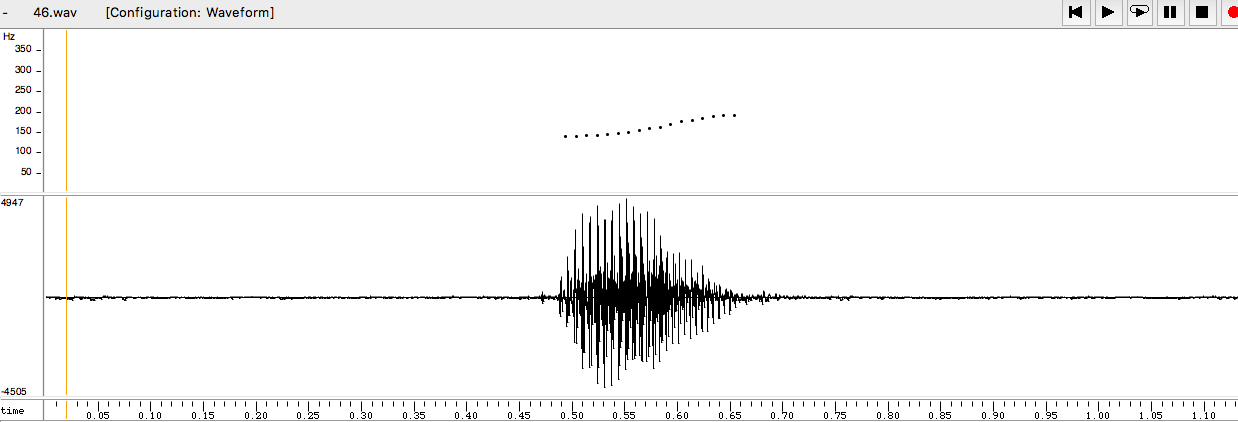
Trình bày mã nguồn cài đặt các thuật toán (copy & paste mã nguồn từ Editor của IDE) kèm theo chú thích (comment) từng khối code theo các sơ đồ khối mô tả trong phần II.

# KẾT QUẢ THỰC NGHIỆM

Mô tả dữ liệu dùng để đánh giá độ chính xác của các thuật toán, đưa ra các đánh giá định tính và định lượng, so sánh các thuật toán đã cài đặt với nhau và với các cài đặt (hoặc thuật toán) khác.

## Hình vẽ

Hình vẽ trong bài viết được đánh số bắt đầu từ 1, được canh lề Justified, lời chú thích được viết dưới hình vẽ với kích thước font chữ là 9pt như Hình 1.



1. Kết quả tính F0 trong trường hợp tốt nhất (chụp màn hình nên bị mờ)

Các hình vẽ tiếp theo được đánh số như Hình 2.



1. Đây là một hình khác xuất từ Matlab figure theo định dạng .eps (rõ nét)

## Bảng biểu

Bảng biểu cũng tương tự như hình vẽ; tuy nhiên dòng chú thích được viết ở phía trên như Bảng 1.

1. Bảng biểu hướng dẫn

|  |  |
| --- | --- |
| Font chữ toàn văn | Times New Roman |
| Kích thước toàn văn | 10pt |
| Kích thước tiêu đề | 14pt |

# KẾT LUẬN

Tóm lại các kết quả đã đạt được và đề xuất các hướng phát triển/hướng cải thiện trong tương lai.

# TÀI LIỆU THAM KHẢO

1. Author1\_Name, Author2\_Name,“Paper Title ACASH: An Adaptive Web Caching method based on the Heterogeneity of Reference Characteristics”, Journal of AICIT, AICIT(Publication\_Name), vol. 10, no. 4, pp.169-711, 2015.
2. Author1\_Name, Author2\_Name, Web Caching and Replication, Addison-Wesley(Publication\_ Name), USA, 2014
3. Link: https://en.wikipedia.org/wiki/Window\_function

**(chú ý chỉ đưa vào các tài liệu có trích dẫn [1], [2], … trong báo cáo)**