**BỘ GIÁO DỤC VÀ ĐÀO TẠO BỘ QUỐC PHÒNG**

**HỌC VIỆN KỸ THUẬT QUÂN SỰ**

**Nguyễn Thị Phương Mai**

**NGHIÊN CỨU CÁC KỸ THUẬT VÀ PHƯƠNG PHÁP XÁC ĐỊNH**

**NGƯỜI NÓI TRONG ÂM THANH**

**CHUYÊN NGÀNH: HỆ THỐNG THÔNG TIN**

**ĐỀ CƯƠNG ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC**

**Hà Nội - Năm 2018**

BỘ GIÁO DỤC VÀ ĐÀO TẠO BỘ QUỐC PHÒNG

**HỌC VIỆN KỸ THUẬT QUÂN SỰ**

**ĐỀ CƯƠNG ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC**

Chuyên ngành: **Hệ thống thông tin**

Mã số:

Ngày giao đồ án: tháng … năm 2018

Ngày nộp đồ văn: tháng … năm 2019

Tên đề tài: **NGHIÊN CỨU CÁC KỸ THUẬT VÀ PHƯƠNG PHÁP XÁC ĐỊNH**

**NGƯỜI NÓI TRONG ÂM THANH**

Học viện thực hiện: Nguyễn Thị Phương Mai

Lớp: HTTT13

NGƯỜI HƯỚNG DẪN KHOA HỌC:

Hướng dẫn:

Họ và tên :

Học hàm học vị :

Đơn vị : Khoa CNTT Học viện KTQS

**Hà Nội - Năm 2018**

**ĐỀ CƯƠNG ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC**

**Tên đề tài:** **NGHIÊN CỨU CÁC KỸ THUẬT VÀ PHƯƠNG PHÁP XÁC ĐỊNH**

**NGƯỜI NÓI TRONG ÂM THANH**

**Chuyên ngành**: Hệ thống thông tin

**Thời gian thực hiện**:

* **Cơ sở khoa học và tính thực tiễn của đề tài**:
* *Cơ sở khoa học*: Đề tài có liên quan đến nhiều nội dung kiến thức như xử lý tín hiệu số, các kiến thức toán học, trích rút đặc trưng, các kỹ thuật học máy cho phân lớp. Hơn nữa, vấn đề xác định người nói trong âm thanh chỉ mới được quan tâm nghiên cứu gần đây và các kết quả đạt được còn nhiều hạn chế, đặc biệt là với ngôn ngữ Tiếng Việt.
* *Tính thực tiễn*: Đây là lĩnh vực có tiềm năng ứng dụng lớn, như xây dựng các chương trình tự động tổng hợp văn bản, ý kiến từ giọng nói từ các hội nghị, hội thảo và ứng dụng mở rộng của chúng .Do sự khác biệt về mặt ngữ âm giữa các ngôn ngữ việc áp dụng các kỹ thuật xử lý tiếng nói đã cho các ngôn ngữ như tiếng Anh, tiếng Nhật, sang tiếng Việt đòi hỏi sự cải tiến. Một hệ thống xác định người nói trong âm thanh phải được xây dựng trên nền tảng của tiếng nói Tiếng Việt. Do đó, đề tài “Nghiên cứu các kỹ thuật và phương pháp xác định người nói trong âm thanh” được thực hiện nhằm đáp ứng nhu cầu thực tế hiện nay là giao tiếp người – máy trở nên hiệu quả hơn.
* **Mục tiêu của đề tài**:
  + *Mục tiêu tổng quát*:
* Nghiên cứu các phương pháp phân đoạn tiếng nói;
* Nghiên cứu các phương pháp trích chọn đặc trưng tiếng nói;
* Tìm hiểu các kỹ thuật phân biệt người nói qua các đặc trưng tiếng nói;
* Xây dựng chương trình thử nghiệm phân biệt người nói trong dữ liệu âm thanh.
* *Nội dung chi tiết:*

Đề tài được trình bày trong 3 chương, phần mở đầu, phần kết luận và kiến nghị và phần phụ lục.

Phần mở đầu nêu tổng quan về mục đích, nội dung, lý do chọn đề tài và ý nghĩa khoa học của đề tài.

Chương 1. Tổng quan về tiếng nói và bài toán xử lý tiếng nói. Trong chương 1 tập trung tìm hiểu:

* Tổng quan về tiếng nói và một số đặc trưng vật lý của tiếng nói
* Tổng quan về xử lý tiếng nói và các lĩnh vực xử lý tiếng nói;
* Trình bày cơ sở lý thuyết của bài toán xác định người nói;
* Một số nghiên cứu và ứng dụng của bài toán xác định người nói;

Chương 2 Tập trung tìm hiểu các phương pháp và kỹ thuật được áp dụng để phân biệt được người nói trong dữ liệu âm thanh. Các nội dung nghiên cứu chính bao gồm:

* Âm thanh và vấn đề số hóa và các định lưu trữ âm thanh;
* Một số kiến thức chung về xử lý dữ liệu âm thanh;
* Mô hình các bước của hệ thống phân biệt người nói trong dữ liệu âm thanh;
* Đặc trưng LPC, MFCC và một số phương pháp trích chọn đặc trưng tiếng nói;
* Các kỹ thuật phân đoạn, xác định vùng có người nói và không có người nói;
* Một số mô hình và phương pháp học phân biệt người nói.

Chương 3 Nghiên cứu đề xuất chương trình cài đặt các mô hình xác định người nói không phụ thuộc vào từ khóa, các nội dung chính trong chương 3 bao gồm:

* Xây dựng bộ dữ liệu thử nghiệm cho bài toán xác định người nói;
* Mô hình hóa người nói: trích chọn đặc trưng sử dụng các thuật toán theo Mô hình lượng tử hóa vector (Vector Quantization) và mô hình hỗn hợp Gauss (Gaussian Mixture Model) để mô hình hóa người nói;
* Xây dựng kịch bản thử nghiệm các thuật toán và mô hình phân biệt giọng nói theo VQ hay GMM;
* So khớp mẫu: đưa ra phương pháp so khớp mẫu trong giai đoạn kiểm thử ứng với từng mô hình hóa người nói VQ hay GMM;
* Đề xuất cấu trúc chương trình thử nghiệm;
* Tiến hành cá thực nghiệm và đánh giá kết quả thực nghiệm với nhiều tiêu chí khác nhau.

Phần kết luận và kiến nghị chỉ ra những vấn đề đã giải quyết được, những tồn tại và định hướng nghiên cứu tiếp theo của đề tài.

* **Phương pháp nghiên cứu:**
* *Về lý thuyết:* Nghiên cứu trên các trang mạng và các đồ án có cùng đề tài liên quan và vận dụng các kỹ năng và kiến thức để học để áp dụng
* *Về thực nghiệm*: ……
* **Nội dung nghiên cứu**:

Nội dung đề tài gồm một số vấn đề chính sau:

**Đặt vấn đề**

**Chương 1: Tổng quan về tiếng nói và bài toán xử lý tiếng nói.**

1.1. Tổng quan về tiếng nói.

1.1.1 Đặc tính của âm thanh.

1.1.2. Đặc tính âm học của tiếng nói

1.2. Các lĩnh vực xử lý tiếng nói.

1.3. Bài toán xác định người nói.

1.3.1. Cơ sở lý thuyết về xác định người nói

1.3.2. Phân loại bài toán xác định người nói

1.3.3. Xác định người nói không phụ thuộc vào từ khóa

1.4. Kết luận chương 1.

**Chương 2: Các phương pháp và kỹ thuật được áp dụng để phân biệt được người nói trong dữ liệu âm thanh.**

2.1. Các định dạng lưu trữ âm thanh

2.2. Một số kiến thức chung về xử lý tín hiệu tiếng nói

2.2.1. Chuyển từ tín hiệu tương tự sang tín hiệu số

2.2.2. Chuẩn hóa biên độ

2.2.3.Lọc nhiễu và làm rõ tín hiệu

2.3. Đặc trưng và trích rút đặc trưng tiếng nói

2.3.1. Mục đích của trích chọn đặc trưng tiếng nói.

2.3.2. Phân loại trích chọn đặc trưng

2.3.3. Trích chọn đặc trưng bằng Mel – frequency cepstral coefficients

2.4. Các phương pháp phân đoạn tiếng nói

2.5. Kết luận chương 2.

**Chương 3: Thực nghiệm đánh giá kết quả**

3.1. Dữ liệu thực nghiệm

3.2. Xây dựng kịch bản thử nghiệm

3.3. Mô hình phân biệt người nói

3.3.1. Xây dựng mô hình người nói

3.3.2. So khớp tiếng nói

3.3.3. Ra quyết định phân lớp

3.4. Kết quả thực nghiệm

**Kết luận và kiên nghị**

**5. Tài liệu tham khảo:**

1. - Voice Recognition Using MFCC Algorithm ( Trích International Journal of Innovative Research in Advanced Engineering), Koustav, Chakraborty, Asmita Talele Prof. Savitha Upadhya
2. Speech Feature Extraction Techniques: A Review , Shreya Narang , Ms. Divya Gupta , 2015
3. Automatic Speaker Recognition Using Gaussian Mixture Speaker Models, Douglas A. Reynolds
4. Speaker Recognition Systems: A Tutorial (African Journal of Information and Communication Technology, Vol. 3, No. 2, June 2007), Abimbola A. Fisusi Thomas K. Yesufu
5. Nhận dạng phương ngữ tiếng việt sử dụng MFCC và tần số cơ bản, Phạm Ngọc Hưng, Trịnh Văn Loan, Nguyễn Hồng Quang

**6. Dự kiến kế hoạch thực hiện:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TT | Nội dung công việc | Thời gian  thực hiện | Địa điểm  thực hiện | Ghi chú |
| 1 | Xác định tên đề tài nghiên cứu | 15/08/2018-20/08/2018 |  |  |
| 2 | Nghiên cứu các khái niệm và lý thuyết và tìm hiểu các nghiên cứu trước đây. | 20/08/2018- 30/08/2018 |  |  |
| 3 | Xây dựng giả thuyết | 03/09/2018- 14/09/2018 |  |  |
| 4 | Xây dựng đề cương | 15/09/2018- 25/09/2018 |  |  |
| 5 | Thu thập và phân tích dữ liệu, thông tin | 26/09/2018- 15/10/2018 |  |  |
| 6 | Viết báo cáo | 16/10/2018 – 10/11/2018 |  |  |
| 7 | Xây dựng chương trình |  |  |  |

7. Các cơ quan, đơn vị cần liên hệ:

* Khoa CNTT - HVKTQS

8. Kinh phí thực hiện đề tài, điều kiện đảm bảo:

* Tự thực hiện.

|  |  |
| --- | --- |
| **CHỦ NHIỆM BỘ MÔN** | *Ngày 24 tháng 09 năm 2018*  **NGƯỜI LẬP ĐỀ CƯƠNG**  Nguyễn Thị Phương Mai |
| **CHỦ NHIỆM KHOA** | **GIÁO VIÊN HƯỚNG DẪN** |