

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

TRẦN ÍCH ĐẠI

**NGHIÊN CỨU PHƯƠNG PHÁP BIẾN ĐỔI CẢM XÚC
NGƯỜI NÓI TRONG TIẾNG NÓI DÙNG KỸ THUẬT
PHÂN RÃ MA TRẬN KHÔNG ÂM**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

THÁI NGUYÊN, 2019

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

TRẦN ÍCH ĐẠI

**NGHIÊN CỨU PHƯƠNG PHÁP BIẾN ĐỔI CẢM XÚC
NGƯỜI NÓI TRONG TIẾNG NÓI DỪNG KỸ THUẬT
PHÂN RÃ MA TRẬN KHÔNG ÂM**

Chuyên ngành: Khoa học máy tính

Mã số: 8 48 01 01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Người hướng dẫn khoa học: PGS. TS. PHÙNG TRUNG NGHĨA

THÁI NGUYÊN, 2019

LỜI CẢM ƠN

Lời đầu tiên, em xin chân thành cảm ơn **PGS.TS. Phùng Trung Nghĩa**, người đã trực tiếp hướng dẫn em hoàn thành luận văn. Với những lời chỉ dẫn, những tài liệu, sự tận tình hướng dẫn và những lời động viên của thầy đã giúp em vượt qua nhiều khó khăn trong quá trình thực hiện luận văn này.

Em cũng xin cảm ơn quý thầy cô giảng dạy chương trình cao học "Khoa học máy tính" đã truyền dạy những kiến thức quý báu, những kiến thức này rất hữu ích và giúp em nhiều khi thực hiện nghiên cứu.

Cuối cùng, em xin gửi lời cảm ơn tới gia đình và bạn bè đã luôn ủng hộ động viên giúp đỡ em trong suốt những năm học vừa qua.

Em xin chân thành cảm ơn!

Thái Nguyên, ngày tháng năm 2019

Học viên

Trần Ích Đại

LỜI CAM ĐOAN

Em xin cam đoan: Luận văn này là công trình nghiên cứu thực sự của cá nhân, được thực hiện dưới sự hướng dẫn khoa học của **PGS. TS. Phùng Trung Nghĩa**

Các số liệu, những kết luận nghiên cứu được trình bày trong luận văn này trung thực và chưa từng được công bố dưới bất cứ hình thức nào.

Em xin chịu trách nhiệm về nghiên cứu của mình.

Học viên

Trần Ích Đại

MỤC LỤC

TRANG PHỤ BÌA	
LỜI CẢM ƠN	i
LỜI CAM ĐOAN	ii
MỤC LỤC	iii
DANH MỤC BẢNG	v
DANH MỤC HÌNH	vi
MỞ ĐẦU	1
1. Đặt vấn đề	1
2. Đối tượng và phạm vi nghiên cứu.....	2
3. Hướng nghiên cứu của luận văn	3
4. Những nội dung nghiên cứu chính.....	3
5. Phương pháp nghiên cứu.....	4
6. Ý nghĩa khoa học của luận văn:	4
CHƯƠNG 1	5
TỔNG QUAN VỀ TIẾNG NÓI VÀ VẤN ĐỀ BIẾN ĐỔI THÔNG TIN CẢM XÚC NGƯỜI NÓI TRONG TIẾNG NÓI	5
1.1. Thông tin tiếng nói	5
1.2. Tín hiệu tiếng nói	5
1.3. Quá trình tạo tiếng nói	7
1.4. Cơ quan thính giác	10
1.5. Xử lý tiếng nói.....	12
1.6. Thông tin cảm xúc người nói trong tiếng nói	13
1.7. Biến đổi cảm xúc người nói trong tiếng nói và ứng dụng	16
1.7.1. Phương pháp biến đổi thay đổi tham số trực tiếp	17
1.7.2. Biến đổi cảm xúc người nói dựa vào thay thế khung	20
1.7.3. Biến đổi thông tin người nói bằng TD-GMM	21

1.7.4. Biến đổi thông tin người nói bằng HTD [9]	23
CHƯƠNG 2.....	30
BIẾN ĐỔI CẢM XÚC NGƯỜI NÓI BẰNG KỸ THUẬT PHÂN RÃ MA	
TRẬN KHÔNG ÂM NMF	30
2.1. Kỹ thuật phân rã ma trận không âm NMF	30
2.1.1. Tổng quan.....	30
2.1.2. Lịch sử.....	30
2.1.3. Nền tảng	30
2.1.4. Tính chất phân cụm.....	32
2.1.5. Các kỹ thuật phân rã ma trận không âm	33
2.2. Ứng dụng kỹ thuật phân rã ma trận không âm NMF trong biến đổi	
tiếng nói [10]	39
CHƯƠNG 3.....	41
ĐÁNH GIÁ THỰC NGHIỆM CÁC PHƯƠNG PHÁP BIẾN ĐỔI CẢM	
XÚC NGƯỜI NÓI TRONG TIẾNG NÓI	41
3.1. Ngữ âm tiếng Việt.....	41
3.2. Cơ sở dữ liệu tiếng nói tiếng Việt.....	43
3.3. Tổng hợp tiếng nói tiếng Việt	46
3.4. Lựa chọn cơ sở dữ liệu.....	47
3.5. Thực nghiệm các phương pháp	47
3.6. Đánh giá các phương pháp.....	47
3.6.1. Tiêu chí và kết quả đánh giá khách quan.....	47
3.6.2. Tiêu chí và kết quả đánh giá khách quan.....	48
KẾT LUẬN.....	51
TÀI LIỆU THAM KHẢO	52

DANH MỤC BẢNG

Bảng 3.1: Cấu trúc âm tiết tiếng Việt.....	43
Bảng 3.2: Sáu thanh điệu tiếng Việt	43
Bảng 3.3: Khoảng cách MCD.....	49
Bảng 3.4: Độ sai lệch F0.....	49
Bảng 3.5: Kết quả đánh giá chủ quan ABX giữa các phương pháp TD-GMM (1); HTD (2); và NMF (3).....	49

DANH MỤC HÌNH

Hình 1.1: Dạng sóng tiếng nói một câu tiếng Việt	6
Hình 1.2: Tiếng nói hữu thanh	6
Hình 1.3: Bộ phận cung cấp làn hơi.....	7
Hình 1.4: Dây thanh âm	7
Hình 1.5: Cấu trúc cơ quan phát âm	8
Hình 1.6: Hình dáng cơ quan phát âm thay đổi trong quá trình phát âm.....	9
Hình 1.7: Mô hình hóa cơ quan phát âm.....	9
Hình 1.8: Biểu diễn mô hình hóa cơ quan phát âm đầy đủ bằng máy tính.....	9
Hình 1.9: Mô hình cơ quan thính giác	10
Hình 1.10: Thang tần số Bark.....	11
Hình 1.11: Ngưỡng nghe	11
Hình 1.12: Mặt nạ thời gian (che âm thanh liền trước và liền sau)	11
Hình 1.13: Mặt nạ tần số (che âm thanh có tần số khác nhau phát cùng thời điểm)	12
Hình 1.14: Một số ứng dụng của xử lý tiếng nói	13
Hình 1.15: Hệ thống nhận dạng cảm xúc người nói – một trong các ứng dụng xử lý thông tin người nói.....	13
Hình 1.16: Người nói khác nhau có cơ quan phát âm và cách phát âm khác nhau dẫn tới tiếng nói khác nhau khi nói với các cảm xúc khác nhau	14
Hình 1.17: Đặc trưng phổ formant đặc trưng cho cơ quan phát âm	15
Hình 1.18: Tính toán vector đặc trưng	16
Hình 1.19: Mô hình học máy thống kê GMM được sử dụng để mô hình hóa người nói	17
Hình 1.20: Phân tích phổ, F0 và tái tạo bằng STRAIGHT	18
Hình 1.21: Thuật toán điều chỉnh trực tiếp tham số tiếng nói	19
Hình 1.22: Lựa chọn khung đích phù hợp và thay thế khung nguồn.....	21

Hình 1.23. Phương pháp biến đổi TD-GMM.....	23
Hình 1.24. Mô hình biến đổi giọng người nói HTD	24
Hình 1.25. Ví dụ phân tích / tái tạo tiếng nói bằng MRTD với N khung và K điểm sự kiện	26
Hình 2.1. NMF tuần tự	36
Hình 2.2. Xây dựng từ điển nguồn và đích cho mỗi câu tiếng nói trong pha huấn luyện	40
Hình 2.3. Pha chuyển đổi	40
Hình 3.1: Đường F0 sáu thanh điệu tiếng Việt theo, dấu ? ở thanh ngã chỉ ra rằng đường F0 của thanh ngã không thống nhất giữa các mẫu ở vùng giữa. .	42

DANH MỤC CHỮ VIẾT TẮT VÀ KÝ HIỆU

Ký tự	Ý nghĩa
F0	Tần số giao động cơ bản
AMDF	Hàm hiệu biên độ trung bình
LP	Phương pháp dự đoán tuyến tính
PCM	Kỹ thuật điều chế xung mã
MAV	Dữ liệu âm thanh không nén
LSF	Tham số phổ đường
ABX	Thang điểm đánh giá theo cặp
GMM	Mô hình Gaussian hỗn hợp
P	Chỉ số hiệu năng
MOS	Thang điểm đánh giá chủ quan trung bình
MCD	Khoảng cách phổ Mel
RMSE	Lỗi trung bình phương
HMM	Mô hình thống kê
HTT	Thay thế mẫu và ghép nối