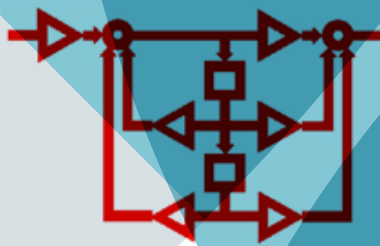


제29회 신호처리합동학술대회 (신호처리로 이해하는 딥러닝)

2019. 9.26(목) ~ 27(금)

KAIST (대전)



- ▶ 주 최 : 대한전자공학회 신호처리소사이어티
- ▶ 주 관 : 대한전자공학회

세션 C – 영상처리

9월 27일

01	딥러닝 기반 압축 영상 기술 분석 CFP 27	39
	정재련, 김동욱, 정승원 (동국대학교)	
02	선형 전역 매핑을 통한 영상 탈색 방법 CFP 57	42
	이혁재, 김창익 (한국과학기술원)	
03	영상의 대비 향상 기술 개발 동향 CFP 54	45
	이승연 ^{1,2} , 김창익 ² (¹ 국방과학연구소 ² 한국과학기술원)	
04	영상 캡셔닝을 위한 효과적인 자가비판 학습법 CFP 78	49
	정승준, 김창익 (한국과학기술원)	

세션 D – 화질향상

9월 27일

01	블라인드 영상 디블러링에서 자기상관을 이용한 커널정보 추출방법 CFP 49	52
	주용관, 김창익 (한국과학기술원)	
02	학습 가능한 선형 보간을 이용한 효율적인 초고해상도화 방법 CFP 44	56
	김종희, 김창익 (한국과학기술원)	
03	합성곱 인공신경망을 활용한 다중 수준 플래시 영상 융합 CFP 32	59
	김재우, 김종욱 (고려대학교)	
04	초고해상도 위성영상 품질 개선을 위한 인공지능 영상처리 기술 CFP 23	61
	임현성, 유수환, 하은재, 백준기 (중앙대학교)	

포스터세션 A

9월 26일

01	단일 이미지를 이용한 패션쇼 애니메이션 생성 기술 CFP 4	64
	안희준, 미나리 마드올 라흐만 (서울과학기술대학교)	
02	GAN을 이용한 얼굴 훈련 영상 데이터의 감정 특성 강화 CFP 5	68
	김준화, 원치선 (동국대학교)	
03	영상 검색을 이용한 CNN 훈련 영상의 오류 라벨 정정 CFP 8	72
	조용현, 팜트린, 원치선 (동국대학교)	
04	Small-scale focused SNIPER for pedestrian detection in drone images CFP 11	76
	Han Xie, Yiran Li, Yizhen Pan, Hyunchul Shin (Hanyang University)	
05	Traffic Light Detection Based on Deep Neural Network by using Spatial Pyramid Pooling and Feature Refinement CFP 12	80
	Ayush Shivali, Han Xie, Hyunchul Shin (Hanyang University)	
06	자폐스펙트럼장애 영유아의 음성 특징에 관한 연구 CFP 14	84
	김성주, 이건우, 김홍국 (광주과학기술원)	

자폐스펙트럼장애 영유아의 음성 특징에 관한 연구

김성주, 이건우, 김홍국
광주과학기술원 전기전자컴퓨터공학부
e-mail : {sjkim3357, goenwoo0801, hongkook}@gist.ac.kr

A Study on Speech Characteristics of Infants with Autism Spectrum Disorder

Seong-Ju Kim, Geon Woo Lee, Hong Kook Kim
School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology

Abstract

본 연구에서는 영유아의 자폐스펙트럼장애 조기진단을 위해 영유아의 음성으로부터 자폐스펙트럼장애 추정을 위한 음성 특징들에 관하여 연구한다. 이를 위해, 자폐스펙트럼장애로 추정되는 영유아들과 그렇지 않다고 추정되는 (비자폐스펙트럼장애) 영유아들의 음성으로부터 pitch, mel-frequency cepstral coefficients (MFCCs) 그리고 spectral centroid 등 9가지의 특징을 각각 추출하여 각 특징간의 유의성 평가를 진행한다. 판별을 위한 분류기로는 가우시안 혼합 모델을 사용하였으며, 부족한 데이터 환경에서의 실험을 위해 cross-validation 방법을 통하여 학습과 성능 평가를 진행하였다. 그 결과, MFCC의 일부 파라미터를 사용하는 경우, 54.47%의 자폐스펙트럼장애의 판별 정확도를 보였다.

I. 서론

자폐스펙트럼장애는 의사소통과 타인과의 사회적 상호작용 능력의 저하를 일으키는 신경발달장애로, 이러한 장애의 진단 지연은 영유아의 전반적 발달에 많은 지장을 주게 된다[1]. 영유아기에 자폐스펙트럼장애를 조기 진단하는 것은 이후에 발생할 수 있는 행동상의 장애를 예방할 수 있는 가능성이 높기 때문에 매우 중요하다[2]. 하지만 자폐스펙트럼장애의 조기진단에는 몇 가지 어려움이 있다. 자폐스펙트럼장애는 앓기,

걷기 등의 신체적 발달은 예정시기에 진행되므로 초기에 장애가 있음을 느끼지 못하다가 언어발달의 장애로 이어질 2세 무렵에 인식하게 되는 경우가 많다[2].

따라서, 본 연구에서는 영유아의 음성으로부터 자폐스펙트럼장애에 대한 정량적인 지표를 제공함으로써 판별에 신뢰성을 줄 수 있는 유용한 음성 특징에 대해 연구하고자 한다. 그 기초적인 단계로, 자폐스펙트럼장애로 추정되는 영유아와 그렇지 않은 영유아의 음성 특징을 추출하여 두 군 간에 각 음성 특징마다 어떤 차이가 있는지 분석한다. 이를 토대로, 유의미한 음성 특징을 선별하여 가우시안 혼합 모델(Gaussian Mixture Model, GMM)을 생성하여 자폐스펙트럼장애의 판별 성능을 평가한다.

II. 본론

2.1 음성 데이터

분석에 사용된 데이터는 자폐스펙트럼장애로 추정되는 영유아의 경우, YouTube[3]에서 “Autism”, “Autism Spectrum Disorder”의 키워드로 검색한 결과에서 게시자가 영상의 영유아가 자폐스펙트럼장애라고 주장하는 영상으로부터 총 5명에 대한 오디오 데이터만을 획득하였다. 반면, 비자폐스펙트럼장애 영유아의 경우, Google Audioset[4]의 “Babbling”(울알거림) 분류에 있는 영상들로부터 자폐스펙트럼장애 영유아 데이터의 수와 동일하게 5명의 영유아의 오디오 데이터를 획득하였다. 해당 데이터의 YouTube 비디오 ID는 다음 표 1과 같다.

표 1. YouTube 비디오 ID.

자폐스펙트럼장애	비자폐스펙트럼장애
DO_sneeTZDQ	-DsIEPiibls
ZdQW7YWltYg	-IxmP19HOUU
4LH6QiyN0I8	1LOZ6z-5HOU
cau0FPiXIKw	1TThHJISLn0
-PeBIWOgiCc	1nC1ddPiDVU

표 2. 분석에 사용한 음성 특징 종류.

음성 특징		
pitch	spectral centroid	MFCC (13 order)
spectral roll-off	spectral flux	spectral entropy
spectral crest factor	spectral flatness	zero crossing rate

표 3. 유의미한 음성 특징의 맨-휘트니 U 검정 결과.

특징	p-value	특징	p-value
pitch	0.067	MFCC 2	0.041
spectral centroid	0.065	MFCC 3	0.051
		MFCC 7	0.002

자폐스펙트럼장애 영유아의 경우 모두 영어권의 영유아이며, 그 중 2명의 나이는 24개월, 27개월이고 나머지 영유아에 대한 연령은 미상이며 성별은 모두 남자이다. 비자폐스펙트럼장애 영유아 데이터의 경우 연령과 성별 모두 미상이다.

2.2 음성 특징

획득한 오디오 파일로부터 표 2와 같이 총 9개의 특징을 추출하였다[5]. 본 논문에서 모든 오디오 파일들의 샘플링레이트는 44.1 kHz로 동일하며, 각 오디오 파일은 50%의 오버랩을 갖는 25 ms 단위의 프레임으로 분할되고, 각 프레임마다 음성 특징을 추출하였다.

추출된 음성 특징들 중 유의미한 차이가 있는 특징만을 분석에 사용하기 위해 맨-휘트니 U 검정[6]을 실시하여 p-value<0.07인 특징들을 선별하였다. 결론적으로, 유의미한 차이를 보이는 특징들은 pitch, 2, 3, 7 번째 차수의 MFCC, 그리고 spectral centroid로 선별되었고, 검정 결과는 표 3과 같다.

2.3 가우시안 혼합 모델

표 3의 음성 특징들로 GMM을 생성하였다. GMM은 MFCC 2, 3, 7차 계수만을 사용한 모델, MFCC 2, 3, 7차 계수와 다른 특징 하나씩을 조합한 모델, 세 특징을 전부 사용한 모델로, 총 4가지 모델을 생성하여 성능을 평가하였다. 데이터 수가 적은 것을 어느 정도 보완하기 위하여, 5 fold cross-validation 방법으로 모델을 학습하고 성능을 평가하였다.

표 4는 4가지 GMM에 대한 판별 성능을 보여 준다. 표에서 보는 바와 같이, MFCC(2, 3, 7)와 pitch를 사용하는 GMM이 가장 높은 판별 정확도를 보였다. 이는 두 영유아 군의 음성 특징으로부터 MFCC의 특정 계

표 4. 음성 특징별 판별 정확도 비교.

모델	정확도(%)
MFCC(2,3,7)	57.21
MFCC(2,3,7) + pitch	57.98
MFCC(2,3,7) + spectral centroid	57.00
MFCC(2,3,7) + pitch + spectral centroid	55.42

수와 pitch에서 차이점이 있다는 것을 의미한다. 이는 [7]에서 언급한 바와 같이, 자폐스펙트럼장애 영유아의 경우, 비자폐스펙트럼장애 영유아에 비해서 느린 언어 발달을 보이는 것에서 기인한 것으로 판단된다.

III. 결론 및 향후 연구 방향

본 논문에서는 획득한 오디오로부터 영유아의 발성 구간에 대해 여러 가지 음성 특징을 추출하고, 추출된 음성 특징 중 유의미한 음성 특징을 선별한 후, 이를 통해 GMM 기반의 판별 실험을 진행하였다. 판별 결과, 이러한 차이는 자폐스펙트럼장애 영유아의 언어 발달의 지연으로부터 기인한 것으로 예상된다. 향후, 자폐스펙트럼장애 조기진단에 있어 정량적인 지표로써 사용될 수 있도록, 본 연구 방법을 더 많은 오디오 데이터에 적용하여 차이점을 분석할 예정이다.

ACKNOWLEDGMENT

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (2019-0-00330, 영유아/아동의 발달장애 조기선별을 위한 행동·반응 심리인지 AI 기술 개발)

참고문헌

- [1] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*, American Psychiatric Pub, 2013.
- [2] 윤현숙, 정보인, “자폐스펙트럼장애의 조기발견을 위한 부모용 행동지표개발,” *연세대학교 박사학위 논문*, 2002.
- [3] <https://www.youtube.com>
- [4] <https://research.google.com/audioset>
- [5] H. Subramanian, *Audio Signal Classification*, M.Tech. Credit Seminar Report, Electronic Systems Group, Dept. of EE. IIT Bombay, 2004.
- [6] P. E. McKnight and J. Najab, “Mann-Whitney U test,” *The Corsini Encyclopedia of Psychology*, 2009.
- [7] E. Patten, *et al.*, “Vocal patterns in infants with autism spectrum disorder: canonical babbling status and vocalization frequency,” *J. Autism Dev. Disord.*, vol. 44, no. 10, pp. 2413-2428, 2014.