

마이크로폰의 종류 및 설치거리에 따른 음성인식성능변화의 검토

김연화*, 이광현**, 최대림**, 김봉완**, 이용주*
원광대학교 전기전자 및 정보공학부*, 원광대학교 SiTEC**

The Validation of Speech Recognition Performance Change according to the kind and established distance of the Microphone

Yoen-Whoa Kim*, Kwang-Hyun Lee**, Dae-Lim Choi**, Bong-Wan Kim**, Yong-Ju Lee*
Dept. of Electrical Electronic And Information Engineering, Wonkwang University*,
SiTEC, Wonkwang University**

E-mail : kimyw@wonkwang.ac.kr, {khlee, dlchoi, bwkim, yjlee}@sitec.or.kr

Abstract

Speech recognition performance depends on various factors. One of the factors is the characteristic and established distance of a microphone which is used when speech data is collected. Thus, in the present experiment speech databases for tests are created through the type and established distance of a microphone. Then, acoustic models are built based on these databases, and each of the acoustic models is assessed by the data to determine recognition performance depending on various microphones and established microphone distances.

1. 서론

음성인식시스템의 성능에 영향을 미치는 요인으로는 인식 알고리즘에 의한 내부적인 것 뿐만 아니라 훈련 및 테스트시(사용시) 사용되는 마이크로폰의 종류 및 거리 등도 영향을 미친다. 그러나 아직까지 이들의 영향에 대하여 체계적인 실험 및 분석 결과는 보고 되지 않았고 다만 필자들에 의해 한정된 마이크로폰의 종류에 따른 영향에 관하여 예비 검토한 결과만 있을 뿐이다[1].

본 보고에서는 마이크로폰의 종류에 따른 음성인식 성능의 변화를 분석하기 위한 음성DB와 거리별 영향

의 분석 및 Distance talking 음성인식의 성능평가에 사용가능한 음성DB의 구축에 관하여 보고하고, 이를 이용한 인식성능변화의 개괄적인 경향을 분석한 결과를 보고한다.

2. 시험용 음성 DB의 구축

종류 및 거리별 음성 DB의 구축을 위해서는 기 보고한 바와 같이 준비된 Source Data를 HATS(Head and Torso Simulator)를 통해 재생하면서 종류와 거리를 달리 하여 음성 DB를 구축하였다. 재생된 공간은 압소음 18dB의 방음스튜디오이다.

사용된 Source Data는 총70명(남자:39명, 여자:31명)이 PBW452어절을 2회씩 발성한 음성데이터로서, 방음실에서 HMD224X마이크로폰을 통하여 수집하였으며, 모든 화자의 발성은 어절단위로 분절되어 있다.

종류 및 거리별 음성 DB 구축에 사용된 마이크로폰의 사양은 표 1과 같다.

거리별 음성 DB는 HATS의 입으로부터 5Cm, 10Cm, 20Cm, 30Cm, 50Cm, 100Cm, 200Cm, 500Cm의 위치에 각각 동일한 사양의 마이크로폰을 설치한 후 멀티 채널 레코더를 이용하여 동시에 수집한 것이다.

표 1. 시험용 음성DB구축에 사용한 마이크로폰

Mic. Model	Mount/Unit Type	Directivity	Frequency Response	비 고
HMD 224X (SENNHEISER)	Headset Dynamic	Super directional	50 ~ 12,000Hz	Source Data
280-Pro (SENNHEISER)	Headset Dynamic	Super directional	50 ~ 13500Hz	종류
e825-s (SENNHEISER)	Hand-held Dynamic	Directional	80 ~ 15,000Hz	종류
SM-10A (SHURE)	Headset Dynamic	Directional	50 ~ 15,000Hz	종류
16A (SHURE)	Desktop Condenser	Directional	50 ~ 15,000Hz	종류
C400-BL (AKG)	Boundary Condenser	Directional	150 ~ 15,000Hz	종류, 거리
C414B-ULS (AKG)	Hand-held Condenser	Directional	20 ~ 20,000Hz	종류
ANC-700 (ANDREA)	Headset ECM	Omni directional	100 ~ 10,000Hz	종류
M@B40 (SENNHEISER)	Headset ECM	Directional	40 ~ 12,500Hz	종류
CMT-58 (BSE)	Embedded ECM	Omni directional	50~15,000Hz	종류
FNE-540PB (FONOTEC)	Headset ECM	Omni directional	50~15,000Hz	종류

마이크로폰의 설치거리에 따른 음성데이터 수집시스템은 그림 1에 나타낸다.

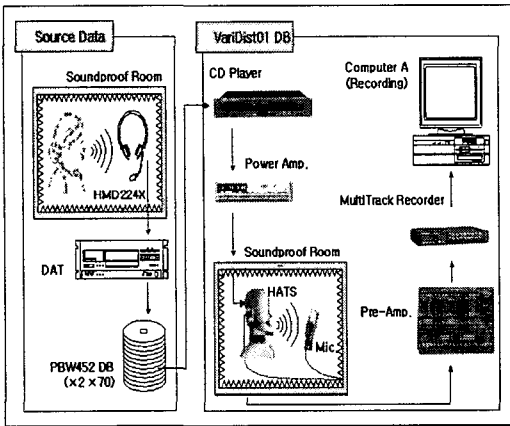


그림 1. 마이크로폰의 설치거리별 음성데이터의 수집시스템 구성도

3. 마이크로폰의 종류별 음성인식의 영향

3.1 실험 방법

마이크로폰의 종류에 따른 음성인식실험에 총 50명(남자:29명,여자:21명)의 마이크로폰별 수록음성데이터를 음향모델의 훈련용음성데이터로 사용하고, 훈련에

참여하지 않은 총 20명(남자:10명,여자:10명)의 음성데이터를 평가용으로 사용하였다.

본 실험에서 사용된 특징파라미터는 MFCC(12), E(1)와 이들의 차분 및 차차분의 39차를 추출하여 사용하였다. 음소모델은 연속 HMM을 이용하여 triphone 모델로 모델링하고, 인식 실험은 HTK3.1.1를 사용하였다.

3.2 결과 검토

마이크로폰의 특성에 따른 음성인식실험의 결과는 표 2에 나타낸다.

표 2. 마이크로폰의 종류별 인식결과(*DM:Dynamic Mic., CM:Condenser Mic., ECM:Electret Condenser Mic.)

평가용 데이터	DM				CM			ECM			평가용 데이터	평균 인식률
	HMD 224X	280 Pro	e825	SM 10A	16A	C400	C414	ANC 700	M@B 40	CMT-58		
HMD 224X	99.16	98.57	98.78	98.23	98.14	98.31	98.51	98.47	97.87	98.05	97.67	98.34
280 Pro	98.69	99.24	98.84	99.21	98.13	98.05	99.12	98.41	98.86	98.62	97.68	98.62
e825	99.02	99.22	99.28	99.21	98.87	98.73	99.19	98.89	98.46	98.81	96.72	98.76
SM 10A	98.63	99.21	98.89	99.20	98.55	98.27	99.14	98.42	99.02	98.45	96.22	98.55
16A	98.38	98.85	98.95	98.73	99.09	98.70	99.04	99.03	97.73	99.16	98.40	98.73
C400	98.80	98.95	98.98	98.87	98.95	99.12	98.85	99.00	98.41	99.03	98.80	98.89
C414	98.22	99.05	98.62	98.84	98.27	97.83	99.15	98.50	98.52	98.86	98.51	98.58
ANC 700	98.86	99.05	99.08	98.94	99.05	98.73	99.08	99.15	98.03	99.15	98.63	98.89
M@B 40	97.16	98.98	96.74	98.84	96.01	97.25	98.85	96.65	99.08	96.88	97.38	97.62
CMT-58	97.22	98.39	97.58	95.32	97.49	96.90	99.12	98.78	96.10	99.23	98.94	97.73
FNE-540PB	87.90	92.23	73.35	70.43	91.27	93.17	98.41	96.80	94.19	98.93	99.11	90.53

• 이로부터 다음과 같은 경향을 알 수 있었다. 즉, 훈련용음성데이터에 사용된 마이크로폰과 평가용음성데이터에서 사용한 마이크로폰의 종류가 동일한 경우 가장 좋은 인식률을 나타낸다.

• 인식 시 사용할 마이크로폰이 정해져 있을 경우는 훈련용DB 구축 시에도 동일 마이크로폰을 사용하는 것이 바람직하고, 유동적일 경우 즉, 사용자에 따라 마이크로폰의 변경이 예상될 경우에는 고급 마이크로폰으로 수록하여 구축된 훈련용DB를 사용하는 것이 좋은 인식결과를 나타내리라 여겨진다.

4. 마이크론의 거리별 음성인식의 영향

4.1 실험 방법

본 실험에서의 훈련용 및 테스트용 DB의 구성방법 및 인식에서 사용된 특징파라미터와 모델구성방법은 3.1과 같다.

4.2 결과 검토

마이크론의 설치거리에 따른 음성인식실험의 결과는 표 3에 나타낸다.

표 3. 마이크론의 설치거리별 인식결과

평가용데이터 훈련용데이터	5cm	10cm	20cm	30cm	50cm	100cm	200cm	500cm	평균 인식률
HMD224X_5	94.08	94.17	93.68	91.77	87.88	62.40	17.16	2.52	67.96
C400BL_5	94.30	94.20	92.68	87.48	73.50	33.01	5.82	1.12	60.26
C400BL_10	94.25	94.36	93.91	92.58	88.93	65.29	15.96	1.11	68.30
C400BL_20	93.43	94.14	94.30	94.21	93.83	91.22	62.59	5.53	78.66
C400BL_30	89.83	93.07	94.29	94.36	94.27	93.73	84.69	16.76	82.63
C400BL_50	78.35	90.54	94.05	94.35	94.34	94.14	90.83	44.12	85.09
C400BL_100	20.58	56.29	91.50	93.91	94.26	94.41	93.60	78.32	77.86
C400BL_200	2.57	11.74	55.09	80.55	88.88	93.76	94.15	91.50	64.78
C400BL_500	0.41	1.14	18.00	45.92	66.43	87.39	93.15	94.14	50.82

• 이 결과에서 훈련 및 평가 시 동일한 거리의 마이크론을 사용하였다고는 하나, 가까운 거리의 마이크론에 의한 것과 원거리 마이크론의 인식결과가 대동소이한 의외의 결과를 보이고 있다. 여기서 예상해 볼 수 있는 점은 원거리에도 가능한 고감도 마이크론인 점과 암소음 19dB 정도의 조용한 환경인 점 등을 감안 하더라도 이에 대해 상세한 분석이 필요하다고 생각된다.

5. 결론

본 논문에서는 음성인식에서의 훈련 및 평가에 사용하는 마이크론의 특성차이 및 마이크론의 설치거리가 인식률에 미치는 영향을 검토하기 위해 시험용 음성DB를 구축하였고, 이를 이용한 인식실험결과를 통하여 그 영향을 개괄적으로 검토하였다.

본 보고는 시험 및 평가용 DB의 구축에 초점을 맞춘 것이고 마이크론의 종류 및 거리별 영향은 개략적

으로 검토한 것이므로 추후 상세한 검토가 필요하며, 여기에서 구축된 음성DB는 알고리즘의 성능평가에 효과적으로 활용이 가능할 것이다.

5. 참고 문헌

- [1] 김연화, 이광현, 정영조, 김봉완, 이용주, "마이크론의 종류에 따른 음성인식 성능의 검토", 대한음성학회 봄 학술대회 발표논문집, pp.183-186, 2003
- [2] S.J. Young, The HTK Book, Cambridge, 1997