

# 음성 분석 요소의 적용을 통한 아나운서 음성의 특징 추출 및 분석

\*김봉현, \*이세환, \*\*가민경, \*\*조동욱, \*\*\*배영래

\*한밭대학교 컴퓨터공학과, \*\*충북도립대학 정보통신학과, \*\*\*충북도립대학 전자상거래과

\*bhkim@hanbat.ac.kr, \*sianlee@nate.com, \*\*kplus@cbpc.ac.kr, \*\*ducho@cbpc.ac.kr, \*\*\*yljb@cbpc.ac.kr

## Feature Extraction and Analysis of Announcers Voice through Application of Voice Analysis Parameter

\*Bong-Hyun Kim, \*Se-Hwan Lee, \*\*Min-Kyoung Ka, \*\*Dong-Uk Cho, \*\*\*Young-Lae J.Bae

\*Hanbat National Univ., \*\*Chungbuk Provincial Univ.

### 요 약

방송 산업의 발전과 더불어 사회적 관심이 증대되면서 방송 분야는 현대 사회에서 문화적 혜택을 누릴 수 있는 기회를 제공하는 분야로 급속히 성장하고 있다. 이와 같은 방송 산업에서 시청자들의 관심과 호응을 유발할 수 있는 중요한 요소가 방송 진행자의 역할이라 할 수 있다. 따라서 본 논문에서는 아나운서의 다양한 능력 요소 중에서 음성에 대한 특징을 추출하여 분석하는 실험을 수행하였다. 이를 위해 남·녀 아나운서들을 피실험자로 선정하고 이들의 음성을 각각 수집하여 비교, 분석을 행하였다. 특히, 음성을 통해 전달할 수 있는 영향력을 분석하기 위해 피치, 지터, 짐머, 안정도, 스펙트럼 및 스펙트로그램 등의 다양한 음성 분석 요소를 적용하여 아나운서의 음성에서 도출할 수 있는 안정성, 전달성 및 친밀성 등의 음성 정보 특징을 분석하여 상호간의 유의적 관계를 추출하였다.

### I. 서 론

방송 매체의 기술적 발전, 시청자의 수준 향상 및 문화 산업의 변화로 인해 현대 사회에서 방송 분야는 대중화를 넘어 필수화가 되어 가고 있다[1]. 특히, 다양한 장르에 대한 방송 분야가 시청자들에게 접근하면서 다수의 시청자들에게 호응을 얻을 수 있는 요소들의 반영이 필수적이며 이러한 세부 요소들에 의해 특정 방송의 성장에 막대한 영향을 미치게 되었다. 이와 같이 방송에 대한 일상 생활속의 관심이 증대되면서 방송의 체계적 구성과 흐름이 중요시되고 있으며 이러한 시스템에서 방송을 진행해나가는 아나운서의 역할은 매우 중요한 요소로 자리잡고 있다. 아나운서는 방송 시간, 주요 시청자, 패널 및 게스트의 상황 및 방송 내용 등의 다양한 변수에 대비하면서 매끄러운 진행과 깔끔한 말투로 방송을 이끌고 나가야 하는 어려움을 수행해야 한다[2].

따라서 본 논문에서는 아나운서에 대한 음성을 수집하여 다양한 음성 분석 요소들을 적용하고 이에 따른 결과값을 기반으로 아나운서의 음성 특성에 대한 특징적 정보를 추출하여 일반인의 음성과의 비교, 분석을 행하는 실험을 수행하였다. 이를 위해 아나운서들을 대상으로 남녀 성별을 구분하여 음성을 수집하였으며 이와 동일한 음성을 기반으로 일반인의 음성 자료를 수집하였다.

### II. 연구 과정 및 방법

#### 2.1 음성 분석 요소의 적용

첫째로, 음성의 안정성에 대한 분석은 진행자들이 해당 멘트를 낭독하는 과정에서 숨을 고르는 시간을 전체 음성 낭독 시간에 적용하여 듣

는 사람으로 하여금 편안함을 줄 수 있는 분석 요소를 의미한다. 즉, 낭독 시간에 비해 쉬는 시간이 짧을 경우 안정도가 낮아지며 반대로 쉬는 시간이 길어질 경우 안정성이 높아진다.

둘째로, 성대 진동율은 말하는 사람의 감정과 정서의 변화에 따라 달라지며 성대의 진동이 1초에 몇 번 있는가를 나타내는 것으로 성대의 크기와 길이, 질량 등에 영향을 받는다[3]. 음성 파형의 정보는 일정 구간에서 유사한 모양이 계속적으로 반복되는 형태를 나타낸다. 이러한 반복 주기는 음성 파형을 특징화하는데 가장 중요한 정보를 제공하며 이를 음성 정보의 기본 주기인 피치(pitch)라 한다. 또한, 단위시간 안의 발음에서 성대의 진동인 피치의 변화가 얼마나 많은지를 나타내주는 것이 지터이다. 지터는 음성파형 1회 진동에 대한 음성파형 3회 진동길이의 비율이다.

마지막으로 스펙트로그램은 음성에서 각각의 성분음의 위상은 청각적으로 느끼지 못하기 때문에 생략하고 주파수와 진폭의 크기만을 나타낸 것을 스펙트럼이라 하며 스펙트럼 정보를 시간점마다 나타낸 것을 스펙트로그램이라 한다. 스펙트럼 정보에서 진폭 값이 높은 언덕 부분은 진하게 표시하고 낮은 골 부분은 연하게 각 시간점마다 연속하여 표시하면 스펙트로그램이 된다. 이 부분에서 진폭 값의 꼭지점에 해당하는 부분은 진하게 나타내고 골에 해당하는 부분은 연하게 표시한 것이 스펙트로그램이다[4].

#### 2.2 실험 자료 수집

본 논문에서는 다양한 음성 분석 요소의 적용에 따른 아나운서의 음성적 정보의 특징 요소를 추출하기 위한 실험을 수행하였다. 이를 위해

아나운서의 음성을 수집하여 실험 집단으로 구성하고 이와 동일한 문장을 낭독한 일반인의 음성을 수집하여 비교 집단으로 각각 구성하였다.

아래 표 1은 실험에 사용된 피실험자 집단을 나타낸 것이다.

표 1. 피실험자 집단의 분류표

구분	성별	소속	구분	성별	소속
AM-01	남성	아나운서	NM-01	남성	일반인
AM-02			NM-02		
AM-03			NM-03		
AM-04			NM-04		
AM-05			NM-05		
AF-01	여성		NF-01	여성	
AF-02			NF-02		
AF-03			NF-03		
AF-04			NF-04		
AF-05			NF-05		

### III. 시플레이션 결과 및 분석

본 논문에서는 수집된 음성 자료를 기반으로 음성 분석 프로그램인 프라트(Praat) 5.0.23을 사용하여 음성 분석 요소들을 적용시켰다. 아래 표 2와 표 3은 아나운서 실험 집단의 음성 분석 결과를 나타낸 것이며 표 4와 표 5는 일반인 실험 집단의 음성 분석 결과를 나타낸 것이다.

표 2. 남성 아나운서 집단의 음성 분석 결과

구분	안정도(%)	집머(dB)	지터(%)	피치(Hz)
AM-01	28.425	1.132	1.853	146.520
AM-02	24.053	0.850	2.042	163.854
AM-03	31.624	0.925	1.935	143.732
AM-04	26.830	1.205	1.733	150.256
AM-05	22.394	0.982	2.147	134.957
평균값	26.67	1.02	1.94	147.86

표 3. 여성 아나운서 집단의 음성 분석 결과

구분	안정도(%)	집머(dB)	지터(%)	피치(Hz)
AF-01	25.820	1.125	1.628	254.268
AF-02	23.614	0.858	1.285	238.651
AF-03	26.418	1.254	2.099	261.052
AF-04	24.935	0.982	2.124	245.933
AF-05	22.862	0.887	1.835	248.602
평균값	24.73	1.02	1.79	249.70

표 4. 남성 일반인 집단의 음성 분석 결과

구분	안정도(%)	집머(dB)	지터(%)	피치(Hz)
NM-01	18.520	1.052	2.412	152.528
NM-02	14.285	0.988	1.892	142.850
NM-03	24.384	1.250	3.402	171.417
NM-04	21.387	1.173	2.572	139.584
NM-05	17.640	0.922	2.394	150.382
평균값	19.24	1.08	2.53	151.35

표 5. 여성 일반인 집단의 음성 분석 결과

구분	안정도(%)	집머(dB)	지터(%)	피치(Hz)
NF-01	20.538	0.988	2.820	248.523
NF-02	18.965	1.024	2.304	270.856
NF-03	14.520	1.217	1.968	252.880
NF-04	22.385	0.876	3.482	239.346
NF-05	16.574	0.975	2.574	264.204
평균값	18.60	1.02	2.63	255.16

또한 아래 그림 1은 남·녀 아나운서의 스펙트로그램 파형이며 그림 2는 남·녀 일반인의 스펙트로그램 파형을 나타낸 것이다.

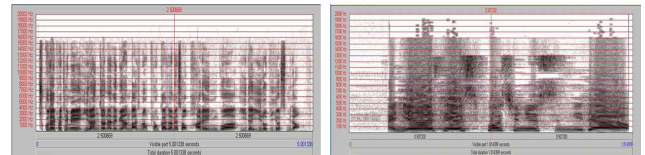


그림 1. 남·녀 아나운서의 스펙트로그램 분석 파형



그림 2. 남·녀 일반인의 스펙트로그램 분석 파형

실험 결과에서 알 수 있듯이 아나운서 실험 집단이 일반인 집단에 비해 음성의 안정도가 20% 중반대로 추출되었으며 이를 통해 아나운서의 음성에서 안정성이 높게 형성되는 것을 알 수 있다. 또한, 지터값이 일반인 집단보다 낮은 1% 후반으로 추출되어 비교적 발음이 정확하게 표현되어 전달성이 높은 것을 알 수 있다. 마지막으로 아나운서 실험 집단의 스펙트로그램 파형이 주파수 대역폭은 높으며 에너지의 세기가 균등하게 분포되어 있는 것으로 저주파 대역폭에 집중된 일반인 집단의 음성보다 친밀성이 낮은 음성임을 알 수 있다.

### IV. 결론

방송 산업의 시대적 변화속에서 지속적으로 관심의 대상이 되고 있는 것이 시청자들의 수준 및 변화의 초점이며 이를 파악하여 원활한 방송의 진행을 주도해야 하는 것이 아나운서의 역할이다.

따라서 본 논문에서는 아나운서에 대한 음성을 수집하여 다양한 음성 분석 요소들을 적용하고 이에 따른 결과값을 기반으로 아나운서의 음성 특성에 대한 특징적 정보를 추출하는 실험을 수행하였다. 실험 결과에서 나타나듯이 아나운서 실험 집단이 일반인 실험 집단보다 안정성 및 전달성은 높게 분석되었으나 친밀성은 낮게 분석되었다. 이는 정확한 전달을 위한 아나운서 집단의 음성적 특징임을 알 수 있다. 이와 같은 결과를 기반으로 방송 장르에 적합한 음성을 표현하는 것이 방송의 효과를 증대시킬 것으로 생각된다.

### 참 고 문 헌

- [1] 백소연 외 2, “통합적 뉴미디어로 변모한 TV에 대한 시청자의 인식과 능동성, 만족도에 관한 연구,” 한국전자거래학회지, 2008.
- [2] 강인수, 방송 산업 발전을 위한 정책과제, 정보통신정책연구원, 1998.
- [3] 정옥란, “음성 총괄 평가,” 대한음성언어의학회지, 1994.
- [4] 이형석 외 5명, “Dr. Speech Science를 이용한 정상 및 후두질환 환자의 음향 분석,” 대한음성언어의학회지, 1997.