

성문진동 패턴의 정량적인 해석을 위한 새로운 시스템 설계와 음성분석

김종찬*·이재천**·김덕원***·오명환****·윤대희*****·차일환*****

*산업기술시험원, **아주대학교 시스템공학과,
연세대학교 의용공학과, *한국과학기술연구원, *****연세대학교 전자공학과
(1999년 4월 29일 접수, 1999년 5월 25일 채택)

A New EGG System Design and Speech Analysis for Quantitative Analysis of Human Glottal Vibration Patterns

J.C. Kim*, J.C. Lee**, D.W. Kim***, M.H. Oh****, D.H. Youn*****, I.W. Cha*****

KTL*, Ajou Univ. Department of System Engineering**, Yonsei Univ. Department of Medical Engineering***,
Korea Institute of Science and Technology****, Yonsei Univ. Department of Electronic Engineering*****

(Received April 29, 1999, Accepted May 25, 1999)

요 약 : 본 논문에서는 고음질 음성부호화 및 압축, 음성인식 및 합성 등의 성능개선에 있어서 중요한 파라미터인 피치정보를 실시간으로 검출하기 위해 연구를 수행하였다. 이를 위하여 변복조 방식의 스폿(spot) 전극을 적용한 새로운 EGG (Electroglottograph) 측정 시스템을 개발하여 이용한 안정된 피치검출 알고리즘을 연구하였다. 구체적으로 EGG신호에 의한 성문 닫힘시점을 실시간으로 결정하여 EGG기반의 피치계측 알고리즘을 개발하였고, 음성신호만의 피치계측 추적기와 성능비교를 수행하여 우월한 성능을 가짐을 보였다. 또한, EGG신호를 이용한 음성분석을 수행하여 한국인에 있어서 성문의 다양한 진동신호 패턴의 측정 및 분석을 통해 한국인 음원의 모델과 성문신호 패턴에 대한 정량적인 해석을 하였다.

Abstract : The purpose of the study is to develop an improved pitch extraction method that can be used in a variety of speech applications such as high-quality compression and vocoding, and recognition and synthesis of speech. To do so, we develop a new electroglottograph (EGG) measurement system that is based on the four modulation-demodulation type spot electrodes for detecting the EGG signals. Then, the glottal closure instant(GCI) is determined from the EGG signals on a real-time basis. We can obtain the pitch contour using the information on the GCI. It turns out that the new pitch contour algorithm (PCA) operates more reliably as compared to the conventional speech-only-based algorithm. In addition, we study the speech source models and glottal vibratory patterns for Koreans by measuring and analyzing the diversified vibration patterns of the vocal folds from the EGG signals.

Key words : EGG, Glottal vibration pattern, Pitch detection, Glottal closure instant, Pitch contour

서 론

EGG시스템은 음성을 발생할 때 성대의 진동을 추출하기 위해 성대의 떨림을 임피던스 변화로 바꾸어 시변신호로 출력시키는 시스템이다.[1] 출력되는 EGG신호는 성문과 관련한 후두의 상태를 정량적으로 파악할 수 있는 방법 중의 한가지이다.[2-4] EGG신호는 인체에 고주파 정전류원을 인가할 때 성대의 진동

운동으로 생기는 피치에 의해서 변조된 전류를 다시 복조해서 얻은 신호이다. EGG신호를 사용한 이유는 첫째, 음성발생을 위한 여기신호가 완전한 주기성을 갖지 못하고, 준주기성을 가진다. 특히, 음성신호 분석에서 중요시되는 유성음의 시작의 첫부분과 끝부분에서는 주기성이 약해지고, 주기내에서도 불규칙한 신호의 변화를 나타낸다. 둘째, 음성신호는 성도와 여기신호의 상호작용에 의한 시변신호이다. 따라서 성도의 저역 포탄트가 여기신호를 변화시킬 수 있으므로 실질적인 주기를 찾는 것이 어려울 수 있다. 셋째, 음성신호를 분석하기 위해 녹음하거나 디지털화 할 때 주변잡음에 의해 유성음의 섬세한 구조가 왜곡되

표 1. 한국인의 성문진동 패턴
Table 1. Korean glottal vibration patterns

sample	sex	vowel	F ₀ (Hz)	Closing time (ms)		Opening time (ms)		Open time (ms)	
				mean	sd	mean	sd	mean	sd
kcs	M	a	128	0.68	0.04	5.14	0.09	2.35	0.11
		e	119	0.69	0.03	5.44	0.13	2.15	0.17
		o	124	0.71	0.05	5.10	0.32	2.25	0.31
		u	129	0.93	0.04	5.30	0.26	2.95	0.41
		i	123	0.66	0.04	5.14	0.23	2.33	0.34
lbd	M	a	115	0.58	0.04	5.17	0.22	3.31	0.19
		e	122	0.63	0.02	5.15	0.32	3.45	0.31
		o	120	0.61	0.03	5.16	0.15	3.46	0.34
		u	115	0.64	0.04	5.09	0.35	3.25	0.41
		i	110	0.62	0.15	5.16	0.33	3.37	0.56
pyr	F	a	251	0.62	0.04	2.12	0.03	1.77	0.12
		e	249	0.71	0.03	2.32	0.06	1.66	0.15
		o	256	0.68	0.04	2.21	0.09	1.69	0.09
		u	247	0.78	0.02	2.17	0.05	1.57	0.11
		i	240	0.59	0.03	2.55	0.12	1.68	0.21
lms	F	a	271	0.45	0.04	1.87	0.13	1.55	0.27
		e	276	0.59	0.05	1.54	0.17	1.74	0.14
		o	281	0.53	0.06	1.67	0.21	1.53	0.09
		u	267	0.48	0.03	1.62	0.15	1.68	0.05
		i	288	0.58	0.04	1.89	0.12	1.58	0.13

는 경향이 있다. 넷째, 진폭의 크기가 작은 유성음 및 무성음의 구분이 어렵다. 다섯째, 유성음 구간에서 피치의 시작과 끝을 찾기가 어렵다.

EGG신호에 의한 피치 검출방법은 임피던스 방식으로는 최선의 방법이고, 임피던스 방식을 적용한 것은 다음과 같은 EGG 신호만의 장점이 있기 때문이다.

EGG신호는 여기신호와 성도의 포만트 사이의 상호작용에 영향을 받지 않으면서 발생시의 주변잡음에도 성대의 떨림에 관한 정확한 정보를 제공하고, 성대의 떨림에 관한 정보를 정확하게 진단하므로 피치를 검출할 때의 문제점들을 해결해 줄 수 있으며, 피치검출을 위한 기준신호로 적합하다고 볼 수 있고, 성문의 주기적인 진동패턴에 준한 후두의 공기흐름에 따른 성문의 개폐 시점을 나타낼 수 있다.

새로운 EGG시스템 설계

EGG신호는 성대 안에 있는 두 겹의 막이 붙어서 공기의 흐름이 적어질 때와 떨어져서 공기의 흐름이 많아질 때 공기 임피던스의 변화를 전압의 파동형태로 얻어진다.[5] EGG신호가 나

타내는 것은 성도의 임피던스가 크면 성분이 열려 있는 것으로 전압이 높다는 것을 의미하며, 성도의 임피던스가 작으면 성분이 닫혀 있는 것으로 전압이 낮아지는 것을 의미한다. 따라서 EGG신호가 피치를 검출하는데 안정적이고, 효과적인 것을 알 수 있다.

EGG시스템은 성대가 위치한 후두좌우에 4개의 고감도 전극을 배치하여 일정한 크기의 전류와 주파수를 갖는 고주파 전류를 인가하면 성대의 진동, 성문의 열림과 닫힘에 따라 두 전극 사이의 임피던스가 변화하고, 진폭 변조된다. 이러한 전기적 변화를 전기적 신호로 검출하는 방식이고, 후두부위의 공기저항을 검출하기 위해 필요한 방법으로 음성신호를 이용하여 피치를 검출하기 위해 후처리(post-processing)할 필요없이 간단한 회로 구성으로 실시간적인 피치를 검출하는 것이 가능하다. 본 연구에서는 발생시 상부 후두계(supra-glottal system)의 작용이 배제된 성문막(vocal folds)의 농작특성을 분석하고, 실시간으로 피치를 추출하기 위해 개발하였다. 그림 1은 변복조기를 이용하여 설계한 EGG시스템의 기능적인 회로도이다. 인가되는 전류에 대한 인체의 전기적 안전[6-7]을 고려하여 자극전류의 크기는 1mA_{p-p}로 결정하였고, 고주파 주파수는 차단 주파수(cut-off frequency)에 따라 다르고, 최종 감지신호가 1kHz이므로 샘플

표 2. 피치검출 알고리즘

Table 2. Pitch detection algorithm

- 음성과 DEGG신호에서 영교차(zero-crossing)를 기준으로 각 프레임 사이의 30ms 길이로 20ms 중복하여 프레임을 나누고, 각 프레임의 주기를 산출하고, 최대, 최소주기를 갖는 프레임을 결정한다.
- 처음 30ms 주기동안 음성신호(*speech avg*), DEGG신호(*DEGG avg*), 음성 영교차율(*speech zcr*), EGG신호 영교차율(*EGG_zcr*)의 절대 평균값을 구하고, 각 프레임의 최대진폭과 최소진폭을 찾고, 판별의 문턱치를 결정한다.
- 문턱치 = $0.3 * (\text{최대진폭} - \text{최소진폭}) + \text{최소진폭}$
- $\text{THR1} = \text{speech avg} \times 2.5$
 $\text{THR2} = \text{DEGG avg} \times 5$
 $\text{THR3} = \text{speech zcr} \times 3$
- $i = 1, \dots, n$ 'th 각 프레임에 대하여 만약 감지된 음성신호의 절대 평균값 $< \text{THR1}$ 이면, 그 때의 프레임은 무성음으로 판별한다.
만약 DEGG신호의 절대 음의 피크값 $< \text{THR2}$ 이고, $\text{speech_zcr} > \text{THR3}$ 이면, 그 때의 프레임은 무성음으로 판별한다.
- EGG신호 프레임에서 평균값과 최대값으로 $E_{\text{mean}}(i)$, $E_{\text{max}}(i)$ 을 구한다. $\text{THR4} = E_{\text{max}}(i)$.
- 한 프레임에서 예측된 영교차율을 계산하여 최대주기와 최소 주기로 Max_zcr 와 Min_zcr 을 찾는다.
만약 $\text{EGG_zcr}(i)$ 가 최대주기와 최소주기인 Max_zcr 와 Min_zcr 범위 안에 있지 않으면, 그 때의 프레임은 무성음으로 판별한다.
만약 EGG신호 프레임에서 최대값인 $E_{\text{max}}(i) > \text{THR4}$ 이고, $\text{EGG_zcr}(i)$ 가 Max_zcr , Min_zcr 범위 안에 있으면, 그 때의 프레임은 유성음으로 판별한다.
만약 $E_{\text{max}}(i) < \text{THR4}$ 이고, $\text{EGG_zcr}(i)$ 가 Max_zcr 와 Min_zcr 범위 안에 있으면서 현재의 i 'th 프레임의 $(i-1)$ 'th와 $(i+1)$ 'th 프레임이 유성음이면 유성음이고, 그렇지 않으면 그 때의 프레임은 무성음으로 판별한다.

링 주파수 때문에 최소 20kHz 이상의 신호를 변조시켜야 하는데 인체의 전기적 안전(let go current)을 고려하고, 정확한 신호를 추출하기 위해 주파수를 70kHz로 하였다. 복조기는 인가한 정전류원에 의해 후두의 공기흐름에 대한 공기 임피던스 신호가 진폭 변조되기 때문에 원하는 신호를 얻으려면 복조를 해야 한다. 0.5Hz 고역 통과필터는 후두부위의 자체적인 고유의 임피던스(bulk impedance)가 있으므로 이를 제거하고 성대의 공기 저항치만을 추출하기 위한 것이다. 증폭기는 복조된 신호가 상당히 작으므로 EGG신호 측정을 위해 1000배 증폭시키는 것이고, EGG신호의 왜곡을 방지하기 위해 1kHz 저역 통과필터를 사용하고, 성대의 진동주파수가 800Hz를 넘지 않으므로 성대신호만을 얻기 위해 불필요한 고주파 잡음을 제거하는

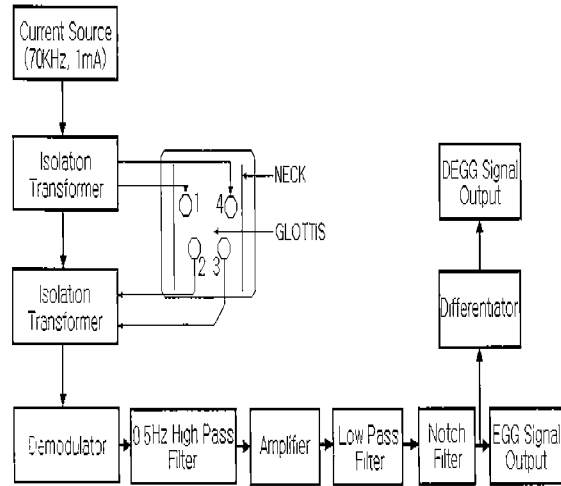


그림 1. 제안한 EGG시스템의 블록선도
Fig. 1. Block diagram of proposed EGG system

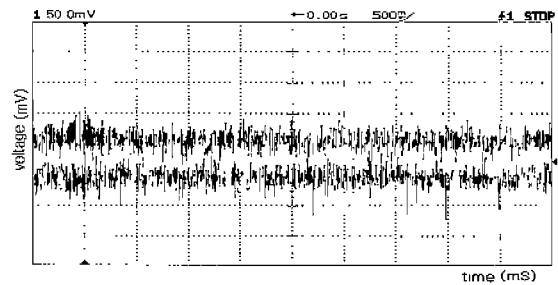


그림 2. 오실로스코프로 변조신호를 측정된 결과
Fig. 2. Measured result of modulation signal using oscilloscope

것이다. 60Hz 전원잡음을 제거하기 위해 노치필터(notch filter)를 사용하고, 최종적으로 EGG신호와 미분기를 통한 DEGG (Differential EGG) 신호의 출력은 0.5V가 되도록 설계하였다. 다른 방식에 비해 4-전극 방식의 장점은 다음과 같다. 후두에서 성문이 열렸다가 닫히는 진동을 이용하는 진동센서 방식도 있으나, 움직임에 의한 잡음(motion artifact)으로 인해 정확한 신호의 검출이 어렵고, 2-전극 임피던스 방식의 시스템도 고려해 보았으나, 2-전극 시스템에서는 전극 2개로 발진신호와 검출신호를 송수신하므로 정밀한 변복조기를 사용하더라도 미세한 생체 신호 임피던스(0.1Ω 이하) 변화를 감지하기에는 문제가 발생하여 후두 좌우의 4-전극 방식을 사용하였으며, 회로가 가장 간단해지고 휴대하기도 용이하다. 변조신호를 측정된 결과 진폭변조 신호에 잡음이 포함되어 있음을 알 수 있었고, 복조신호와 EGG 출력신호의 파형이 같음을 알 수 있었다. 그림 2는 오실로스코프로 변조신호를 측정된 결과이고, 그림 3은 복조신호와 EGG 출력신호의 파형을 실험한 결과이다.

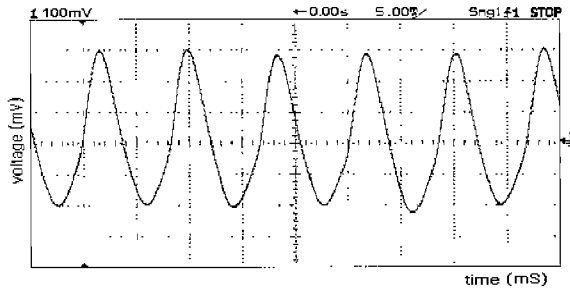


그림 3. 복조신호와 EGG 출력신호의 파형을 실험한 결과
 Fig. 3. Experiment result of demodulation and EGG output signal

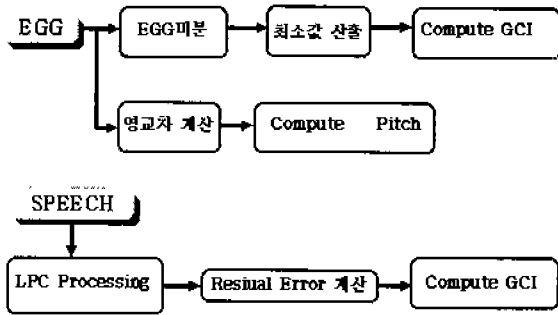


그림 4. 제안한 음성분석의 블록선도
 Fig. 4. Block diagram of proposed speech analysis

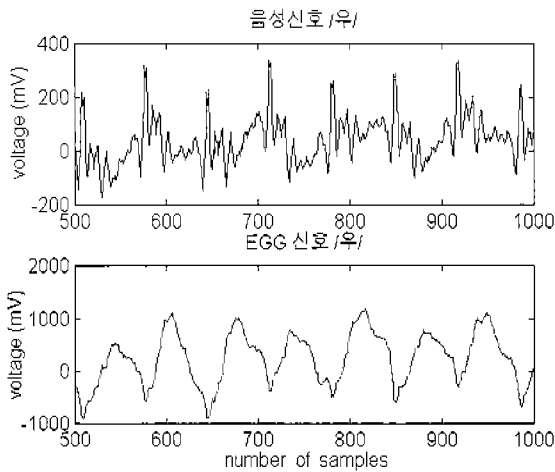


그림 5. 성인남성의 음성신호와 EGG신호
 Fig. 5. Speech signal and EGG signal of adult man

실험 방법

본 연구에서는 음성데이터 수집과 음성신호 분석을 위한 데이터의 디지털처리와 데이터처리를 위해 윈도우 프로그램을 구성하고, 실시간적으로 피치를 검출하게 되며, 디지털 변환은 GUI

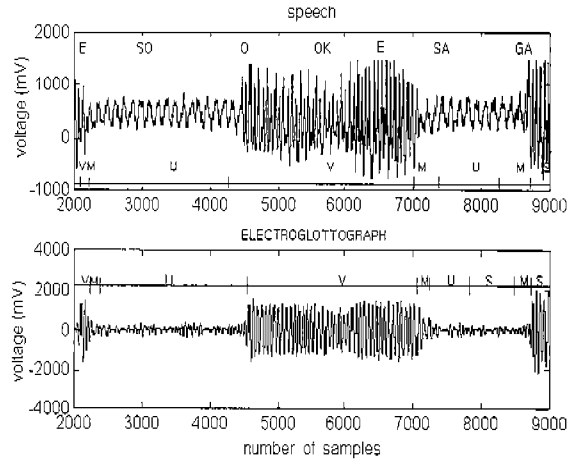


그림 6. 유성음/무성음/묵음의 분류
 Fig. 6. Classification of Voiced/Unvoiced/Silent

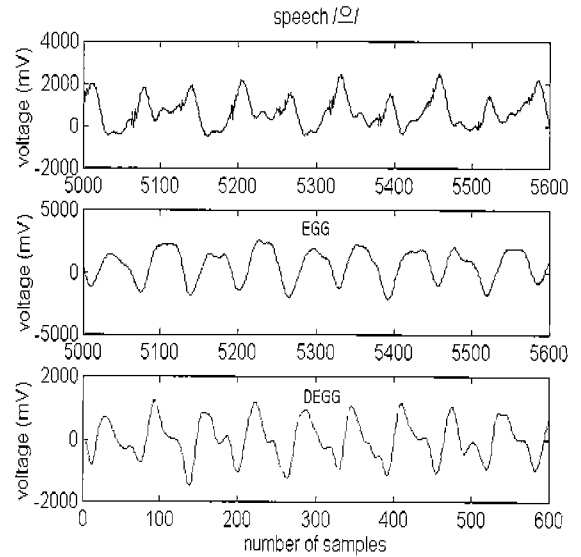


그림 7. 음성, EGG 및 DEGG신호
 Fig. 7. Signals of Speech, EGG, and DEGG

프로그램인 DT VBB[8]를 활용하여 디지털처리를 설계하고 피치분석을 하였다.

한국인 성문신호의 패턴을 분석하고 성분신호를 기반으로 하는 실시간 피치검출을 구현하기 위한 표준적인 어음재료로 단모음 5개(/아/, /에/, /오/, /우/, /이/)와 문장 '성문파를 실험하고 있습니다'를 방음실에서 고음질 마이크로폰[9]을 이용하여 2초 동안 발성하도록 하였다. 발성시간을 2초로 한 것은 음성신호가 발성 후 안정할 때까지 발생된 구간을 처리하고, 10kHz 16비트로 표본화하였다. 표본화한 후 4.7kHz의 저역 통과필터를 통과시켜 미리 강세(pre-emphasis)를 두었다. 데이터를 디지털

필 처리한 후 첫째, EGG신호에서 산출된 성문 닫힘시간동안 LPC처리를 수행하고, 얻어진 잔차신호로부터 포먼트와 대역폭을 산출하기 위해 평활화 과정을 실행하였다. 둘째, 음성신호만을 이용한 LPC처리와 잔차신호를 가지고 피치신호를 추정하여 EGG신호 입력과의 비교를 통해 피치 검출의 성능평가를 수행하였다. 셋째, EGG신호를 기준으로 하여 문장 '성문파를 실험하고 있습니다'에 대한 유성음/무성음/목음 분류를 하였다. 그림 4에 제안한 음성신호 분석 알고리즘의 전체적인 흐름을 도시하였다. 성문신호 패턴의 비교를 위해 30명을 대상으로 데이터 수집을 하였고, 본문의 표 1에는 지면상 대표적 샘플만을 표현하였으며, 데이터의 통계처리는 t-test를 이용하여 평균값과 표준편차를 얻었다.

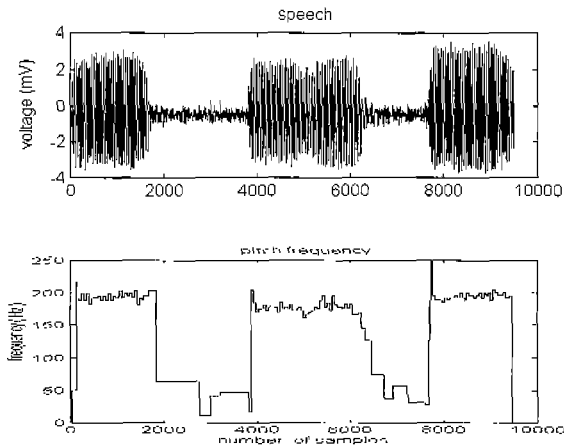


그림 8. 한국어 '성문파를 실험하고 있습니다'에 대한 피치계적
 Fig. 8. Pitch contour of Korean 'secung mun pa reul sil heom ha go it seum ni da'

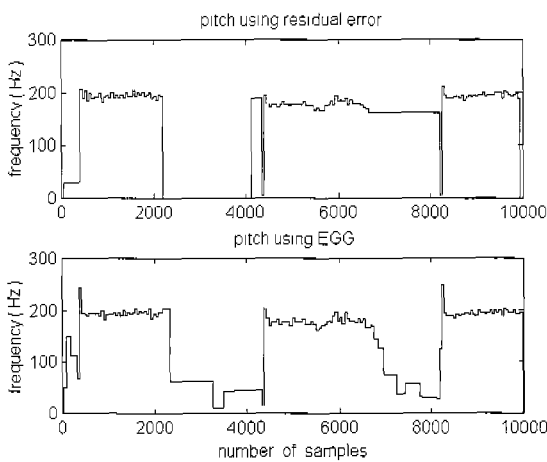


그림 9. 음성신호에 의한 LPC방법의 피치계적과 EGG신호에 의한 피치계적의 성능비교
 Fig. 9. Performance comparison between the pitch contour of only the LPC method by the speech signal and that of the EGG signal.

EGG신호에 의한 음성분석 및 유성음/무성음/목음의 판별

EGG신호를 이용하여 피치검출 알고리즘을 제안하고 실험하였으며, EGG신호는 음성신호의 각 피치구간마다 정확하게 동기되고, 성도에 의한 포먼트의 영향을 거의 받지 않으므로 그 주기성을 검출하여 피치를 구하는 것이 실시간적으로 가능하다. 본 연구에서는 EGG신호의 기저선 변동이 최소가 되도록 설계하여 그 출력신호가 영교차하는 점을 순차적으로 검출하여 피치를 구하였다. 그림 5는 정상인 25세 남성이 단모음 /우/를 발성할 때 후두의 열림과 닫힘의 순간변계를 나타내고 있으며, 상단의 음성신호의 최대, 최소피크에 EGG신호가 정확히 동기되는 것을 알 수 있다. 그림 5의 EGG신호의 양의 피크는 성문막이 열리는 순간을 나타내고, 음(negative)의 피크는 성문막이 닫히는 순간을 나타내는데, 이 때 음성신호의 최대피크와 정확히 동기되어 정확한 피치와 피치주기를 산출할 수 있다. 그러나, 화자의 움직임에 의한 잡음이나 유성자음과 같은 혼합음에서는 음성신호 또는 EGG신호만으로는 정확한 피치검출이 어렵게 된다. 본 연구에서는 영교차하는 EGG신호와 이에 대응하는 음성신호의 분석으로 피치를 검출하였다. 제안한 피치검출 알고리즘은 EGG신호 피크의 절대치 평균값을 구하고, EGG신호의 영교차점을 구하고, 최소 주기구간과 최대 주기구간을 제외한 나머지 주기들을 가지고, 평균주기를 구한다. 유성음을 발성할 때의 성대의 진동운동에서 매 주기마다 성대가 서로 최대로 접촉하는 시점에 음의 피크값을 가진다는 성질을 이용하고, 음성신호의 왜곡 및 불규칙한 변화는 EGG신호의 분석을 통해 상호보완해 줄으로써 시간영역에서 피치동기된 순시적인 피치주기를 정확히 검출할 수 있는 알고리즘을 개발하였고, 표 2는 본 연구에서 적용한 개발된 알고리즘을 나타낸다. 그림 6은 음성신호의 유성음/무성음/목음을 효과적으로 분류하고, 판별하기 위하여 개발된 알고리즘을 적용하여 분류한 특성을 나타낸 결과로서 음성과 EGG신호 분석에 의해 상호 보완하여 유성음의 주기성이 불규칙한 시삭의 첫 부분과 끝 부분, 주기내의 신호의 변화, 녹음하거나 디지털화할 때 왜곡이 있는 부분, 진폭의 크기가 작은 유성음과 무성음 구간을 정확히 판별한다. 그림 6에서 V는 유성음, U는 무성음, M은 혼합음, S는 목음을 나타낸다.

성문신호 패턴의 정량적인 분석

표 1은 각각의 발성된 보음으로부터 20 사이클을 평균하여 얻어진 평균값과 표준편차 값들로서 닫힘시간(closing time)이 가장 표준편차가 적었고, 평균 피치주파수는 120Hz의 변화(130~250Hz)가 있음을 알 수 있다. 닫힘시간과 피치주파수와는 어떠한 상관관계도 보이지 않았으며, 피치주파수의 약간의 변동은 주로 열림시간(opening time)과 열린시간(open time)에 의해

생겨났고, 이들의 값이 적으면 적을수록 피치주파수는 더 크게 나타나게 된다. 한국인 남성의 경우 닫힘시간은 각 사이클에서 피치주기의 3~10%, 여성은 10~21%를 나타낸다. 즉, 여성이 남성보다 EGG신호의 한 주기가 짧다는 것을 의미한다.

EGG신호에 의한 성문 닫힘시점 검출

성문이 닫히는 시점을 GCI(glottal closure instant)라 정의하며, 성문 닫힘시점이 음성신호의 피치와 일치하므로 GCI는 정확한 피치를 추출하기 위한 방법이고, 고음질 합성, 음성압축, 음성코딩이나 인식율의 개선을 위한 중요한 파라미터로서 연구가 활발히 진행되고 있다. 성문 닫힘시점 검출을 위한 방법은 크게 음성신호 자체만의 잔차신호에 의한 방법과 EGG 신호에 의한 방법이 있다. 성문이 여닫히는 순간 기류의 급격한 변화가 생길 것이라는 가정으로 시도된 음성신호만의 잔차신호를 이용한 방법은 원순음과 비음에서 큰 오차를 발생시키는 까닭에 GCI검출에서 오류를 유발하는 원인이 되고 있다. 그림 7은 정상인 25세 남성이 단모음 /오/를 발성할 때 후두의 열림과 닫힘의 순간관계를 나타내고 있으며, 상단의 음성신호의 최대, 최소피크에 EGG신호와 DEGG신호가 정확히 동기되는 것으로 성문 닫힘시점을 정확히 찾을 수 있다. 본 연구에서는 이러한 오차를 발생하지 않는 방법으로서 EGG신호를 이용한 알고리즘을 구현하기 위해 시스템의 출력신호인 EGG신호와 DEGG신호로부터 각 구간에서의 최대 음의 피크를 찾는 방식을 이용하여 정확한 값들을 구할 수 있다. 음성의 여기형태(type of excitation)에 따른 성문의 패턴분류를 EGG신호에 적용하여 성문 닫힘구간을 찾아 유성음을 해석하는 것이다. 기존의 방법들이 음성의 신호크기가 적은 초기구간과 자음과 모음의 천이구간에서는 문제점들을 보이며, 비음과 원순모음의 경우는 LPC를 이용하여 잔차신호를 구하고, 이를 힐버트(Hilbert) 변환하여 처리하는 방식인 SGCI(sequential glottal closure instant detection)는 성문 닫힘구간이 일정한 구간만큼 편이되는 현상이 발생되기 때문에 정확한 위치검출을 위해 이를 보상해 주어야 하는 문제가 있다. 즉, 이러한 방법은 피치의 간격결정은 가능하나, 성문 닫힘구간의 결정은 어려운데, 본 연구에서 제시한 EGG신호를 이용한 성문 닫힘시점의 검출방법은 비음과 원순음에 대해서도 안정되고, 정확한 성문의 닫힘시점을 찾아 주었다. 또한, 기존 방법에서의 예측오차를 이용한 방법을 적용할 때에 나타나는 음성 세그먼트 중간에서의 잘못된 피치검출에 의한 오차를 고려하지 않아도 된다.

피치계적 추적에 의한 성능평가

개발된 EGG시스템의 피치검출 성능을 평가하기 위한 방법으로 한국어 문장 '성문과를 실험하고 있습니다'를 어음자료로 한 결과로서 음성신호만을 이용한 방법과 음성 및 EGG신호를 이

용한 분석 프레임 내에서의 평균적인 피치계적을 비교하였다. 음성신호만을 이용한 방법에서는 음성신호의 LPC예측 오차신호를 구한 후 AMDF를 적용하여 피치를 구하였다. 그림 8은 EGG신호에 동기된 피치계적을 나타내며, 시간에 대한 피치주파수를 연속적으로 나타낸 궤적의 패턴이다. 그림 9에서 실선은 EGG방식을 적용한 궤적이고, 점선은 음성신호만의 궤적이다. 본 연구에서 제안한 EGG신호와 음성신호를 이용하여 각 프레임 피치를 적용한 궤적이 더 연속적이며, 정확한 피치계적이 얻어짐을 알 수 있다. 이러한 특징은 합성음의 음질향상에 EGG기반의 피치적용이 기존의 성분신호 모델링보다 우수함을 나타내는 것이다. 따라서 음성신호만을 이용하여 새롭게 개발되는 피치검출 알고리즘의 객관적인 비교 및 평가기준으로 사용될 수 있고, 정확한 음성분석 과정에 유용하게 이용될 수 있다.

결론

본 논문은 음성신호에 의한 피치검출의 여러 문제점을 해결하고, 한국인의 성분신호 패턴을 정량적으로 분석하기 위한 EGG 시스템을 설계하였으며, 객관적인 성능비교 및 평가기준이 될 수 있는 EGG기반의 피치검출 알고리즘을 연구하였고, 후두의 상태를 판단할 수 있는 정량적인 방법들을 분석하는 임상 의뢰 진단, 음성합성, 음성인식, 암호해독 등의 국방문제와 범죄수사 등에 응용할 수 있을 것으로 사료되며, 얻어진 결과는 다음과 같다.

1. EGG신호의 각 사이클에서 닫힘시간이 가장 표준편차가 적었고, 평균 피치주파수는 130~250Hz 사이에서 나타남을 알 수 있다. 그리고, 닫힘시간과 피치주파수와는 어떠한 상관관계도 보이지 않았으며, 피치주파수의 약간의 변동은 주로 열림시간과 열림시간에 의해 생겨났고, 이들의 값이 작으면 작을수록 피치주파수는 더 크게 나타나게 된다.
2. 한국인 남성의 경우 닫힘시간은 각 사이클에서 피치주기의 3~10%, 여성은 10~21%를 나타낸다.
3. 기존의 방법들이 음성신호의 크기가 적은 초기구간과 자음과 모음의 천이구간에서 문제점들을 보이며, 비음과 원순모음의 경우는 성문 닫힘구간이 일정한 구간만큼 편이되는 현상이 발생되기 때문에 정확한 위치검출을 위해 이를 보상해 주어야 한다. 즉, 이 방법은 피치의 간격결정은 가능하나 성문 닫힘구간의 결정은 어려운데, 본 연구에서 제시한 EGG신호를 이용한 성문 닫힘시점의 검출방법은 비음과 원순음에 대해서도 안정되고, 정확한 성문의 닫힘시점을 찾아 주었다.
4. 기존 방법에서의 예측오차를 이용한 방법을 적용할 때에 나타나는 음성 세그먼트 중간에서의 잘못된 피치검출에 의한 오차를 고려하지 않아도 된다.
5. 본 연구에서는 피치검출에 의한 오차를 발생하지 않는 방법으로서 EGG신호를 이용한 알고리즘을 구현하기 위해 시스템의 출력신호인 EGG신호를 이용하였고, 변복조 방식의 스크린

극 방식의 EGG시스템을 개발함으로써 기존 성문신호 모델링에 의한 성분신호 해석을 정량적이고 체계적으로 분석하였다.

6. 유성음을 발생할 때의 성대의 진동운동에서 매 주기마다 성대가 서로 최대로 접촉하는 시점에 DEGG신호가 음의 피크값을 갖는다는 성질을 이용하고, 음성신호만의 불규칙적이고 왜곡이 있는 부분을 상관성이 있는 EGG 프레임을 이용하여 시간영역에서 실시간적으로 피치동기된 순차적인 피치검출 알고리즘을 개발하였다.

7. EGG신호에 의한 피치계측을 활용하여 음성신호의 유성음/무성음/착음 분류에도 효율적으로 적용할 수 있다. 본 연구에서 적용한 피치검출 알고리즘은 음성신호만을 이용한 기존의 피치검출 알고리즘에 비해 훨씬 정확한 피치계측을 나타냈다. 따라서 음성신호만을 이용한 피치검출 알고리즘의 객관적인 성능 비교 및 평가기준으로 사용할 수 있고, 한국어의 음성분석에 많은 기여를 할 수 있을 것으로 기대한다.

참 고 문 헌

1. S. Smith, "Research on the principle of electroglottography," *Folia Phoniatr.* vol. 33, pp. 105-114, 1981.
2. T. Baer, A. Lofqvist, and N. S. McGarr, "Laryngeal vibration: a comparison between high-speed filming and glottographic techniques," *J. Acoust. Soc. Am.* vol. 73, pp. 1304-1308, 1983.
3. D. E. Metz, R. Whitehead and D. Peterson, "An optical illumination system for high-speed laryngeal cinematography," *J. Acoust. Soc. Am.* vol. 67, pp. 719-720, 1980.
4. M. B. Higgins, R. Netsell, and L. Schulte, "Aerodynamic and electroglottographic measures of normal voice production: intrasubject variability within and across session," *J. Speech Hearing Research.* vol. 37, pp. 38-45, 1994.
5. D. G. Childers, D. M. Hicks, G. P. Moore, and Y. A. Alsaka, "A model for vocal fold vibratory motion, contact area, and the electroglottogram," *J. Acoust. Soc. Am.* vol. 80, pp. 1309-1320, Nov. 1986.
6. I. Chatterjee, D. Wu, and O. P. Gandhi, "Human body impedance and threshold currents for perception and pain for contact hazard analysis in the VLF MF band," *IEEE Trans. BME* vol. 33, pp. 486-494, May 1986.
7. J. G. Webster, *Medical Instrumentation: Application and Design*, Houghton Mifflin Co., 1978.
8. *Data translation product handbook*, 1998.
9. *Modular precision sound level meter type 2231 instruction manual*, Bruel & Kjaer.