Journal of Engineering and Applied Sciences 15 (2): 452-459, 2020
ISSN: 1816-949X

Comparison of Voice Analysis Programs for Fundamental Frequency Measurement in Elderly Voice Signals Through Gender Analysis

Min Ju Seo and JiYeoun Lee
Department of Biomedical Engineering, Jungwon University, Donbu-ri #5, Goesan-eup, Goesan-gun, 367-805 Chungcheongbuk-do, South Korea

Abstract: Changes in the vocal folds due to aging may change the pitch of the voice. An elderly signal can be automatically distinguished from a normal signal through various analyses. With most smart biomedical devices, elderly voices have been neglected due to optimization that does not take the elderly into account. The objective of this study was to use a symmetric higher-order differential energy function to analyze the elderly signal and extract the fundamental frequency. This study suggests a symmetric higher-order differential energy function based on gender analysis. The elderly voices of 40 Korean subjects (20 females and 20 males) ranging in age from 70-80 years were used. Symmetrical instantaneous frequency estimators with orders 5 and 4 were selected for female and male voices, respectively in this study through gender analysis. The experiments were compared to the F0 extracted by various methods such as manual extraction, WaveSurfer, TF32, Praat and an instantaneous frequency estimator based on before-and-after gender analysis. The F0 value obtained through the instantaneous frequency estimator after gender analysis is the most similar to the results from manual extraction, exhibiting an accuracy of 80%. The results will help to provide ease of access for the elderly by means of speech. Future investigations will incorporate multiple analytical methods to implement more reliable detectors for automated medical diagnostic systems.

Key words: Elderly voice, symmetric higher-order differential energy function, fundamental frequency, instantaneous frequency estimator, diagnostic, disorder voice

INTRODUCTION

Speech signal processing is a field of researching algorithms that computers automatically understand and process natural language uttered by human. It is necessary to understand the contents and emotions of words, recognize meaning and exchange natural conversations (Lee, 2017). According to the statistics of the National Statistical Office, Korea has already entered an aged society and expects to increase to 15.1% by 2020 (Cho *et al*., 2018).

The aging of the body brings morphological changes in the tissues of the vocal cords and larynx structure that are directly relevant to voice. Therefore, elderly voices have to be understood in conjunction with the acoustic properties of the sound as the larynx changes due to aging. In other words, anatomical and physiological changes in the larynx and vocal cords may change the pitch of the voice which is measured by the fundamental frequency (F0) (Kahane, 1981). In conclusion, elderly voice is distinguished from that of such groups as young adults and middle-aged persons.

Computerized smart devices with speech interface are actively produced for medical welfare systems (Lee, 2014). However, as is the case for most medical devices, elderly voices have been excluded from the speech recognition and synthesis systems due to an interface that does not take the elderly into account (Kim and Kim, 2001). A speech interface that supports a medical device currently uses an optimized method based on the average speech patterns of young adults, the middle-aged and the elderly. If the deviation of the pattern from the standard is too great, it may result in a phenomenon that degrades the performance of voice synthesis and recognition (Kim and Kim, 2001; Song, 2012; Lee and Kwon, 2014). Therefore, in order to respond to the changes of aged society, the voice signal analysis technique of the elderly should be utilized in the clinic for the purpose of interpreting the voice in the multidimensional aspect and providing appropriate intervention plan. Analysis of the elderly voice can be the beginning of gender innovation.

Although, many studies related to elderly signals have been published they have been based on the acoustic analysis of voice samples including jitter and shimmer (Sataloff *et al*., 1997; Kim and Ko, 2008; Song, 2012). According to the "Change and characteristics of acoustic indicators of Koreans according to Age" published by Kim *et al*. (2000) the acoustic index for voice analysis showed a high correlation between the age difference and the changes due to aging phenomenon in the 40 and 50 sec, respectively. Therefore, it is considered inappropriate to analyze the acoustic index of the elderly signals by the

Corresponding Author: Ji Yeoun Lee, Department of Biomedical Engineering, Jungwon University, 85 Munmu-ro Goesan-eup Goesan-gun, Chungbuk 28024, Korea, +82-10-6751-1476

uniformized threshold and the threshold value for each acoustic index is applied with the same number of sex samples according to each age group. Since, the acoustic index are also based on fundamental frequency, a highly reliable pitch detection algorithm is necessary to measure voicing irregularities in elderly signals (Lee *et al*., 2010; Lee, 2012). Perturbation analysis has been found to be sensitive to pitch variations in analysis tools such as Multi-Dimensional Voice Profile (MDVP), CSpeech and TF32.

Some papers have utilized Higher-Order Statistics (HOS) for pitch estimation. Many studies have applied HOS to disordered voices, since, Alonso *et al*. (2001) published on the automatic detection of voice pathologies by HOS-based parameters (Lee *et al*., 2008a, b; Lee and Hahn, 2010; Lee *et al*., 2011; Lee, 2012). Furthermore, the combination of HOS analysis and the Linear Predictive Coding (LPC) residual may help to effectively construct important information to distinguish the signal types of disordered voices (Lee and Hahn, 2010; Lee *et al*., 2011; Lee and Choi, 2012).

In this study, we analyze the speech of the elderly by means of gender analysis. Gender analysis provides a basis for robust evaluation of the differences between women and men and removes the possibility of analysis based on incorrect assumptions and stereotypes. An instantaneous frequency estimator using symmetric higher-order differential energy function is presented in this study for the analysis of elderly voices and a comparison of the various programs of fundamental frequency extraction such as TF32, WaveSurfer, Praat, and manual techniques (Iem, 2010, 2011). This approach will enhance the speech recognition performance of existing smart medical systems for the elderly. It is also expected to help provide a means of easy access for the elderly and people with disabilities who were excluded from rapid socialization with speech.

Gender analysis methods: Gendered innovations aim to pursue excellence in the research and development of science and technology by integrating sex and gender analysis. Gendered innovations activities were initiated by Dr. Londa Schiebinger of Stanford University and the concept, methods and case studies are presented in detail on the following website: http://genderedinnovations.stanford.edu.

Sex and gender can influence all stages of research or development processes, ranging from strategic considerations for establishing priorities and building theory to the more routine tasks of formulating questions, designing methodologies and interpreting data as shown in referenced website.

Gender innovation is an approach to applying gender analysis in science and engineering. As the need for "innovation that reflects gender" is raised, Stanford University and the European Commission (EC), led by experts in gender, basic science, engineering, health and medicine, policy and technology development have developed a systematic analysis method that considers gender and sex. Gender innovation methods and case studies were developed through seven international workshops from 2010-2013. The twelve methods for gender and sex analysis in each research field are as follows: rethinking research priorities and outcomes, rethinking concepts and theories, formulating research questions, analyzing sex, analyzing gender, analyzing how sex and gender interact, analyzing factors intersecting with sex and gender, engineering innovation processes, designing health and biomedical research, participatory research and design, rethinking standards and reference models and rethinking language and visual representations. We should consider the above research processes in each step to analyze sex and gender.

In this study, we reported the sex of research subjects and the differences that exist within groups of females and males as sex analysis. We also conducted research on the elderly voice in response to gender needs as gender analysis. Therefore, the experiments were conducted by setting the elderly voice as the target, not using voice from all ages. Finally, we considered standards and reference models to decrease gender and sex bias.

MATERIALS AND METHODS

Elderly voice database: Elderly voice samples were collected in the Speech information Technology and Industry Promotion Center (SiTEC). The database includes the elderly voices of 40 Korean subjects (20 females and 20 males), ranging in age from 70-80 years. Two sentences and ten words were used in this study. The voice samples were sampled at 16 kHz. Detailed database information is shown in Table 1.

Extraction method of fundamental frequency: A symmetric higher-order differential energy function is utilized to extract fundamental frequency. The instantaneous frequency is a single frequency value at a certain instance in time. That is it is a time varying frequency as a function of time. The definition of classical instantaneous frequency can be reviewed in (Iem, 2010, 2011).

For the discrete signal $x[n] = A\cos(\Omega n+\theta)$, the higher order differential energy function is defined as follows:

$$\Gamma_k\{x[n]\} = x[n]x[n+k-2]-x[n-1]x[n+k-1] \qquad (1)$$

Where:

- $x[n]$: Discrete signal at time n
- k : The order of the energy operator

A symmetric higher-order differential energy function can be defined as follows:

Table 1: Elderly voice database

Sex	Age	No.	Korean sentences
Female	70-79	5	Who then came forward to her desk (geuttae nuga geunyeoui chaegsang ap-eulo dagawassda)
		5	Then a stranger approached and asked (geuttae wen nachseon salam-i dagawa mul-eossda)
		10	Blue house (cheong-wadae), vaccinia (udukeoni), kolitical retirement (jeong-gyeeuntoeleul), by-election (bogwolseongeo), out of sight (siyaleul), iron ax (soedokkileul), benefit (hyetaeg-eul), hugging (kkeul-eoango), premium (boheomlyo), factitious (in-wijeog-in)
Male	70-78	5	Same
		5	Same
		10	Same

$$\Xi_k\{x[n]\} = \begin{cases} \dfrac{\Gamma_k\{x[n]\}+\Gamma_k\{x[n-k+2]\}}{2}, & \text{odd k} \\ \Gamma_k\left\{x\left[n-\dfrac{2}{k}+1\right]\right\}, & \text{even k} \end{cases} \quad (2)$$

Instantaneous frequency using a symmetric higher-order differential energy function can be derived as follows:

$$\Omega[n] = \frac{1}{k-1}\cos^{-1}\left(\frac{\Xi_{2k-1}\{x[n]\}}{2\cdot\Xi_{2k-1}\{x[n]\}}\right) \quad (3)$$

This equation manages the time misalignment of the generalized instantaneous frequency. To obtain the frequency prediction performance with order of the instantaneous frequency estimator an AM-PM signal is used as a test signal. The signal is defined as follows:

$$x[n] = \left[1+\kappa\cdot\cos\left(\frac{\pi n}{100}\right)\cdot\cos\left[\frac{\pi n}{5}+20\cdot\lambda\cdot\sin\left(\frac{\pi n}{100}\right)\right]\right], \; n = 1,2,\ldots,400 \quad (4)$$

$$(\kappa, \lambda)\in\{(0.05i, 0.05j), \; i, j = 1,2,3,\ldots,10$$

The κ and λ control an amount of AM and PM, respectively. They increase the values from 5-50% in a stepwise fashion by increments of 5%. Then, instantaneous frequency of the signal is defined as follows:

$$\Omega[n] = \frac{\pi}{5}+\frac{\pi\lambda}{5}\cdot\cos\left(\frac{\pi n}{100}\right) \quad (5)$$

The frequency estimation error of the instantaneous frequency is defined as follows:

$$\text{Error}(\%) = \frac{1}{40000}\sum_{i=1}^{10}\sum_{j=1}^{10}\sum_{n=1}^{400}\left|\frac{\Omega_{ij}[n]-\hat{\Omega}_{ij}[n]}{\Omega_{ij}[n]}\right| \quad (6)$$

Where:

$\Omega_{ij}[n]$: The real values of instantaneous frequency

$\hat{\Omega}_{ij}[n]$: Estimated values of instantaneous frequency in specific (i, j)

Table 2: Error (%) of various instantaneous frequency estimators

Estimation orders	Noise free	SNR 15 dB	SNR 20 dB
Order k = 2	0.1401	84.1846	58.5612
Order k = 3	0.2675	13.4978	7.3614
Order k = 4	0.3480	12.0056	6.3703
Order k = 5	0.2675	7.2245	7.2871

RESULTS AND DISCUSSION

Order decision of instantaneous frequency estimator: Table 2 shows the performance of various instantaneous frequency estimators. In the noise free case, the symmetric higher-order differential energy operator based on the instantaneous frequency estimator with order k = 2 shows the best result. Other orders also produce acceptable results. In noisy situations, such as SNR 15 and 20 dB, the symmetric higher-order differential energy operators based on the instantaneous frequency estimator with order k = 5 and k = 4 yielded the best results. When the order is 4, acceptable results (7.2245 and 7.2871%) are shown in two noise environments. Then, the second order estimator shows the worst results.

Fundamental frequency estimator: Figure 1 shows the structure of the fundamental frequency estimator. Elderly signals such as Fig. 2a are first processed with a low pass filter with a cutoff frequency of 250 Hz. Then, the 250 Hz lowpass filtered signal is shown as in Fig. 2b. The phonated sentence is " Then a stranger approached and asked (geuttae wen nachseon salam-i dagawa mul-eossda)" in Korean. Symmetrical instantaneous frequency estimators with orders 5 and 4 are selected for female and male voices, respectively as shown in Table 2. The higher-order differential energy function used to estimate an instantaneous frequency is utilized as a characteristic parameter to classify unvoiced and voiced sections. The fundamental frequency can be found through the moving average filter with estimated instantaneous frequency values for the obtained voiced sections.

Figure 2 shows the procedure for the extraction of fundamental frequency in elderly female voice waveform. In Fig. 2c, the threshold of the differential energy operator is fixed to 800 to classify the unvoiced and voiced sections in elderly female waveform. Then, instantaneous frequency is estimated in voiced sections as shown in Fig. 2d. To obtain the final fundamental frequency

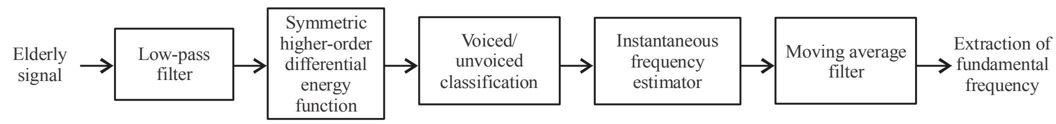

Fig. 1: Structure of a fundamental frequency estimator

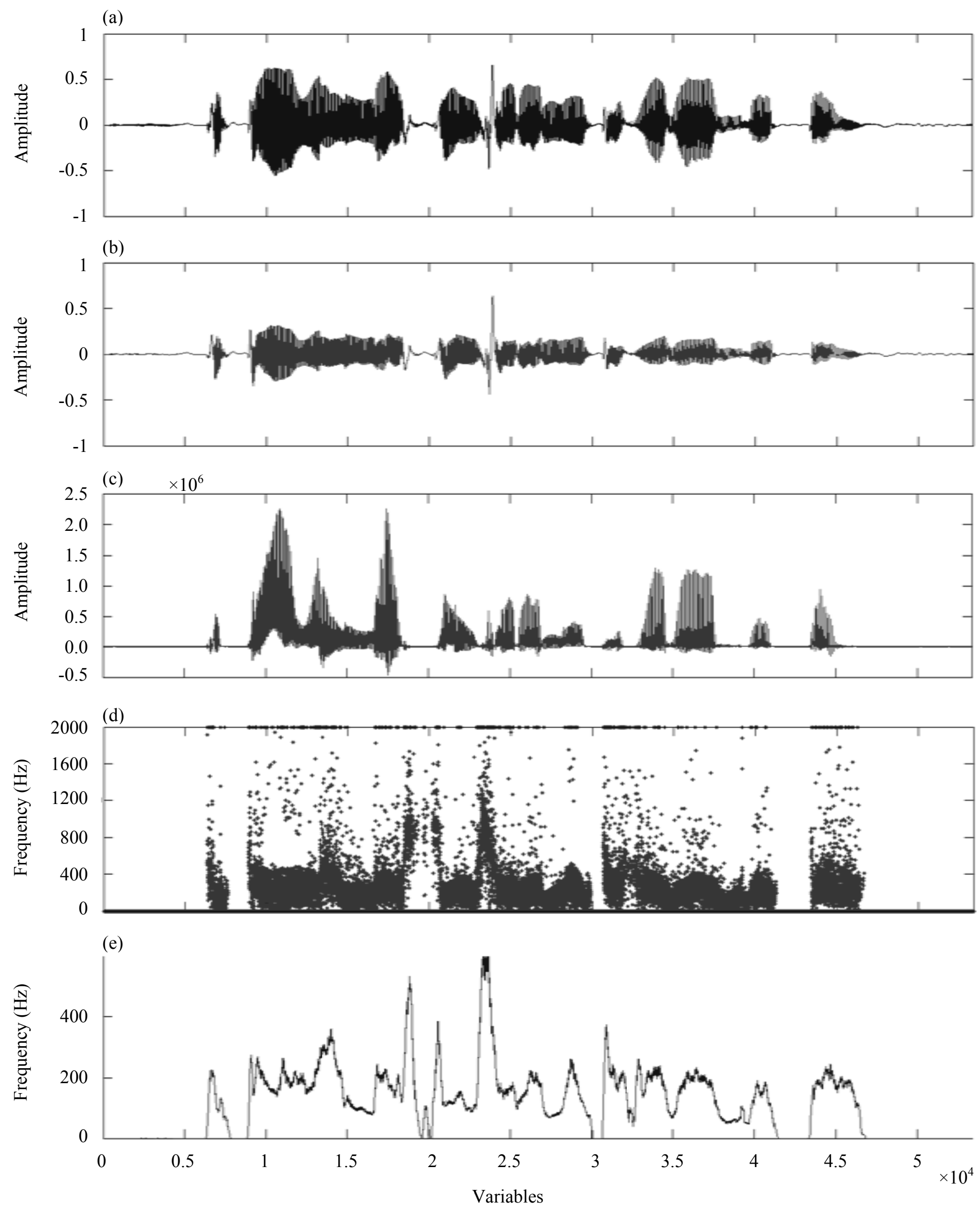

Fig. 2: Higher order differential energy functions extracted from elderly female voice (a) Original waveform of elderly female voice, (b) 250 Hz lowpass filtered waveform, (c) Differential a energy operator, (d) Instantaneous frequency, (e) Extracted fundamental frequency

of Fig. 2e, a moving average filter with a length of 200 is utilized. The same procedures are used for the elderly male waveform in Fig. 3.

Table 3 shows the results of various fundamental frequency extractions. From F_1-F_5 and from M_1-M_5, the subjects pronounced the sentence as

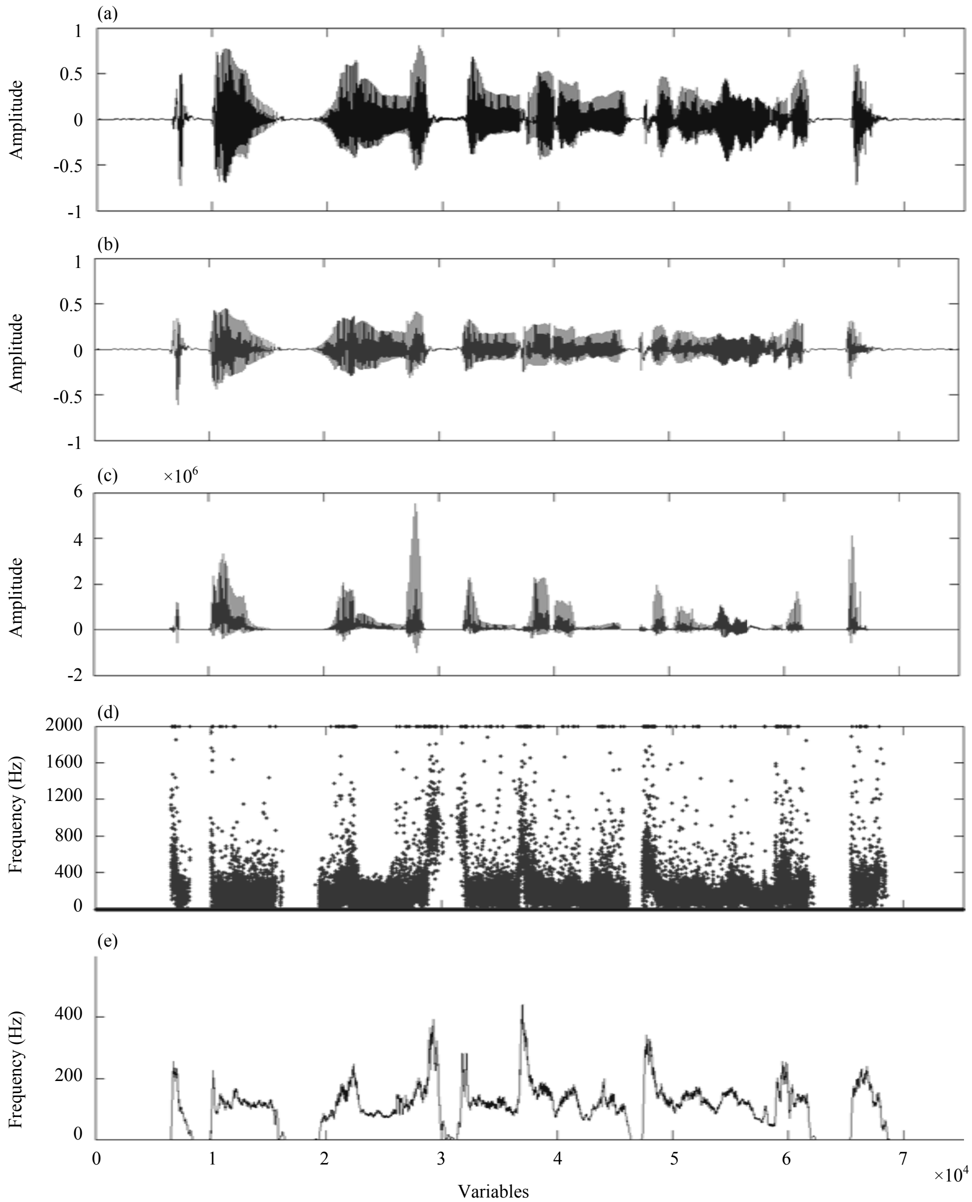

Fig. 3: Higher order differential energy functions extracted from elderly male voice (a) Original waveform of elderly male voice, (b) 250 Hz lowpass filtered waveform, (c) Differential energy operator, (d) Instantaneous frequency, (e) Extracted fundamental frequency

"Who then came forward to her desk (geuttae nuga geunyeoui chaegsang ap-eulo dagawassda.)" in Korean. From F_6-F_10 and from M_6-M_10, the subjects pronounced the sentence as "Then a stranger approached and asked (geuttae wen nachseon salam-i dagawa mul-eossda)" in Korean. From F_11-F_20 and from M_11-M_20, the subjects pronounced the words as "Blue house (cheong-wadae), vaccinia (udukeoni), political retirement (jeong-gyeeuntoeleul), by-election (bogwolseongeo), out of sight (siyaleul), iron ax (soedokkileul), benefit (hyetaeg-eul), hugging (kkeul-eoango), premium (boheomlyo), factitious (in-wijeog-in)" in order.

The first column represents F0 which was manually extracted by our researchers. The second and third columns indicate F0 that were extracted by TF32, WaveSurfer and Praat. Finally, F0 was extracted with the instantaneous frequency estimator with order 5 for female and male signals. It is the last stage before gender analysis that does not consider characteristics of

Table 3: Comparisons of various extraction methods

	Before gender analysis							After gender analysis (Instantaneous frequency estimator (Hz))			
Variables	Manual F0 (Hz)	TF32 F0 (Hz), error rate (%)		Wave surfer F0 (Hz), error rate (%)		Praat F0 (Hz), error rate (%)		Order k = 5 for all		Order k = 5 (females) and 4 (males), error rate (%)	
F_1	162	**162**	**0.00**	166	2.47	179	10.49	**162**	**0.00**	**162**	**0.00**
F_2	173	160	7.51	169	2.31	177	2.31	**174**	**0.57**	**174**	**0.58**
F_3	183	187	2.19	195	6.56	206	12.57	**185**	**1.09**	**185**	**1.09**
F_4	167	161	3.59	**164**	**1.80**	**170**	**1.80**	181	8.38	181	8.38
F_5	210	207	1.43	**210**	**0.00**	212	0.95	179	14.76	179	14.76
F_6	170	166	2.35	174	2.35	174	2.35	**168**	**1.18**	**168**	**1.18**
F_7	178	164	7.87	169	5.06	**182**	**2.25**	**182**	**2.25**	**182**	**2.25**
F_8	203	200	1.48	200	1.48	204	0.49	**203**	**0.00**	**203**	**0.00**
F_9	206	217	5.34	219	6.31	236	14.56	**208**	**0.97**	**208**	**0.97**
F_10	200	203	1.50	206	3.00	208	4.00	**199**	**0.50**	**199**	**0.50**
F_11	154	141	8.44	145	5.84	189	22.73	**156**	**1.30**	**156**	**1.30**
F_12	173	163	5.78	168	2.89	173	0.00	**173**	**0.00**	**173**	**0.00**
F_13	206	**205**	**0.49**	202	1.94	203	1.46	203	1.46	203	1.46
F_14	168	175	4.17	181	7.74	181	7.74	**171**	**1.79**	**171**	**1.79**
F_15	178	175	1.69	181	1.69	197	10.67	**178**	**0.00**	**178**	**0.00**
F_16	187	164	12.30	179	4.28	**186**	**0.53**	180	3.74	180	3.74
F_17	214	205	4.21	210	1.87	**212**	**0.93**	198	7.48	198	7.48
F_18	220	**221**	**0.45**	222	0.91	231	5.00	203	7.73	203	7.73
F_19	208	200	3.85	203	2.40	203	2.40	**209**	**0.48**	**209**	**0.48**
F_20	233	234	0.43	235	0.86	235	0.86	**233**	**0.00**	**233**	**0.00**
M_1	131	124	5.34	133	1.53	143	9.16	**130**	**0.76**	**130**	**0.76**
M_2	116	101	12.93	112	3.45	160	37.93	**114**	**1.72**	**114**	**1.72**
M_3	176	169	3.98	170	3.41	179	1.70	**174**	**1.14**	**174**	**1.14**
M_4	121	110	9.09	121	0.00	130	7.44	**121**	**0.00**	**121**	**0.00**
M_5	125	110	12.00	120	4.00	131	4.80	**127**	**1.60**	**127**	**1.60**
M_6	240	232	3.33	**236**	**1.67**	237	1.25	200	16.67	**237**	**1.25**
M_7	106	100	5.66	**104**	**1.89**	104	1.89	130	22.24	115	8.49
M_8	120	**118**	**1.67**	118	1.67	123	2.50	144	20.00	135	12.50
M_9	133	131	1.50	**132**	**0.75**	136	2.26	152	14.29	**134**	**0.75**
M_10	130	127	2.31	124	4.62	130	0.00	**130**	**0.00**	**130**	**0.00**
M_11	147	146	0.68	149	1.36	169	14.97	161	9.52	**146**	**0.68**
M_12	138	129	6.52	130	5.80	131	5.07	**141**	**2.17**	**141**	**2.17**
M_13	116	105	9.48	108	6.90	110	5.17	**114**	**1.72**	**114**	**1.72**
M_14	122	121	0.82	123	0.82	123	0.82	127	4.10	**122**	**0.00**
M_15	118	115	2.54	114	3.39	115	2.54	**116**	**1.69**	**116**	**1.69**
M_16	142	148	4.23	148	4.23	162	14.08	136	4.23	**140**	**1.41**
M_17	219	203	7.31	214	2.28	217	0.91	**217**	**0.91**	**217**	**0.91**
M_18	120	112	6.67	116	3.33	116	3.33	**118**	**1.67**	**118**	**1.67**
M_19	188	157	16.49	192	2.13	183	2.66	184	2.13	**186**	**1.06**
M_20	127	**126**	**0.79**	123	3.15	124	2.36	**128**	**0.79**	**128**	**0.79**

females and males. After gender analysis, F0 are extracted with the instantaneous frequency estimator with order 5 and 4 for female and male signals, respectively. F0 of female elderly and male elderly was average 190 and 142 Hz. The average F0 value of adult males is about 100-130 Hz and the average F0 value of adult females is about 190-230 Hz. F0 of the older females tend to decrease and F0 of the older males tend to be higher.

The values of the instantaneous frequency estimator before-and-after gender analysis were 75 and 80%, respectively. The accuracies of WaveSurfer, TF32 and Praat were 12.5, 12.5 and 10%, respectively. In addition, comparing the manual and various extraction methods with the F0 error rate (%), the F0 value of the instantaneous frequency estimator after gender analysis was closest to that of manual extraction with an accuracy of 80%. The gray cells represent extracted F0 values closest to the manually extracted F0 values. As shown in Table 3, the best accuracy is obtained with F0 extracted through the instantaneous frequency estimator with order k = 5 and 4 for females and males, respectively, after gender analysis.

CONCLUSION

Speech interfaces that are used in smart medical devices currently uses an optimized method based on the average speech patterns of young and middle-aged adults and the elderly. If the speech pattern has even a slightly larger deviation from the standard model, it may result in a phenomenon that degrades the performance of voice synthesis and recognition. Elderly voices have been neglected from the speech recognition and synthesis system of most smart medical devices due to speech interfaces that do not take the elderly into account.

Gender analysis provides a basis for robust analysis of the differences between women and men and removes the possibility of incorrect assumptions and stereotypes by means of speech analysis. In this study, we reported the sex of research subjects and the differences that exist within groups of females and males as sex analysis. We also conducted research on the elderly voice in response to gender needs as gender analysis. Therefore, the experiments were conducted by setting the elderly voice as the target, not using a voice created from the average speech patterns of all ages. Finally, we confirmed the standards and reference models to decrease gender and sex bias.

The elderly voices of 40 Korean subjects (20 females and 20 males) ranging in age from 70-80 years were used. The symmetrical instantaneous frequency estimators with order 5 and 4 were selected for female and male voices, respectively in this study through gender analysis. The higher-order differential energy function to estimate an instantaneous frequency is utilized as characteristic parameter to classify unvoiced and voiced section. The fundamental frequency can be found through moving average filter with the values of an instantaneous frequency estimated in the obtained voiced section. Performance was estimated as a comparison of F0 extracted by various methods, such as manual techniques, WaveSurfer, TF32, Praat and an instantaneous frequency estimator based on before-and-after gender analysis. The F0 values through the instantaneous frequency estimator after gender analysis are the most similar to the manual extraction results with an accuracy of 80% being observed.

Our research will enhance the speech recognition and synthesis performance of existing smart medical systems for the elderly. This study is also expected to help provide a means of easy access to such enhancements as speech for the elderly and people with disabilities who are excluded from rapid socialization.

ACKNOWLEDGEMENTS

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2017R1A2B4011373).

REFERENCES

Alonso, J.B., J.D. Leon, I. Alonso and M.A. Ferrer, 2001. Automatic detection of pathologies in the voice by HOS based parameters. EURASIP. J. Appl. Signal Process., 4: 275-284.

Cho, B.C., S. Cheon, K.N. Kim and H.S. Yuk, 2018. A policy study for the voice recognition technology based on elderly health care. J. Digital Convergence, 16: 9-17.

Iem, B.G., 2010. An instantaneous frequency estimator based on the symmetric higher order differential energy operator. IEICE. Trans. Fundam. Electron. Commun. Comput. Sci., 93: 227-232.

Iem, B.G., 2011. Estimation of fundamental frequency using an instantaneous frequency based on the symmetric higher order differential energy operator. Trans. Korean Inst. Electr. Eng., 60: 2374-2379.

Kahane, J.C., 1981. Anatomic and Physiologic Changes in the Aging Peripheral Speech Mechanism. In: Aging, Communication Processes and Disorders, Beasley, D.S. and G.A. Davis (Eds.). Grune and Stratton, New York, USA., ISBN:9780808912811, pp: 21-45.

Kim, H.T., S.H. Cho, S.M. Youn, D.I. Sun and M.S. Kim, 2000. The changes and characteristics of acoustic parameters with aging in Korean. Korean J. Otorhinolaryngology Head Neck Surg., 43: 69-74.

Kim, I.C. and D.Y. Kim, 2001. A study on the features for building Korean digit recognition system based on multilayer perceptron. J. Korea Ind. Inf. Syst. Res., 6: 389-391.

Kim, S.H. and D.H. Ko, 2008. Fundamental frequencies in Korean elderly speakers. Speech Sci., 15: 95-102.

Lee, J. and S.H. Choi, 2012. Perturbation analysis using a moving window for disordered voices. Intl. J. Eng. Sci. Innovative Technol., 3: 1-10.

Lee, J., 2014. KHIDI brief. Korea Health Ind. Dev. Inst., 140: 1-2.

Lee, J., S. Jeong, M. Hahn, A.J. Sprecher and J.J. Jiang, 2011. An efficient approach using HOS-based parameters in the LPC residual domain to classify breathy and rough voices. Biomed. Signal Process. Control, 6: 186-196.

Lee, J.Y. and M. Hahn, 2010. Automatic assessment of pathological voice quality using higher-order statistics in the LPC residual domain. EURASIP. J. Adv. Signal Process., 2009: 1-8.

Lee, J.Y., 2012. A two-stage approach using Gaussian mixture models and higher-order statistics for a classification of normal and pathological voices. EURASIP. J. Adv. Signal Process., 2012: 1-8.

Lee, J.Y., S. Jeong and M. Hahn, 2008a. Pathological voice detection using efficient combination of heterogeneous features. IEICE. Trans. Inf. Syst., 91: 367-370.

Lee, J.Y., S. Jeong, H.S. Choi and M. Hahn, 2008b. Objective pathological voice quality assessment based on HOS features. IEICE. Trans. Inf. Syst., 91: 2888-2891.

Lee, J.Y., S.H. Choi, J.J. Jiang, M.S. Hahn and H.S. Choi, 2010. Perturbation and perceptual analysis of pathological sustained vowels according to signal typing. Phonetics Speech Sci., 2: 109-115.

Lee, K.N., 2017. Voice language processing technology, how far have you come. National Korean Lang. Inst. N. Lang. Life, 27: 99-116.

Lee, S.J. and S.I. Kwon, 2014. Elderly speech analysis for improving elderly speech recognition. J. Korean Inst. Inf. Sci. Eng., 32: 16-20.

Sataloff, R.T., D.C. Rogen, M. Hawkshaw and J.R. Spiegel, 1997. The three ages of voice: The aging adult voice. J. Voice, 11: 156-160.

Song, Y.K., 2012. Prevalence of voice disorders and characteristics of Korean voice handicap index in the elderly. Phonetics Speech Sci., 4: 151-159.