

## 음성 분석을 통한 최근 보이스피싱의 음성 특징 규명

이 범 주\*, 조 동 욱°, 정 연 만\*

Identification of Voice Features for Recently Voice Fishing  
by Voice Analysis

Bum Joo Lee\*, Dong Uk Cho°, Yeon Man Jeong\*

## 요 약

국가적 그리고 사회적 피해 방지 노력에도 불구하고 보이스피싱으로 인한 재산 피해의 규모가 좀처럼 줄지 않고 있는 실정이다. 이는 최근의 보이스피싱이 세련된 말투와 전문용어를 사용함으로써 범죄자를 인지하기 어렵게 만들기 때문이다. 더 나아가 개인정보 취득과 실제 관공서등에서 근무하는 종사자들이 실제 현장에서 사용하는 전문용어를 구사하여 젊은 층을 집중적으로 속이고 있는 실정이다. 이는 결과적으로 노년층보다 판단력이 있는 20~30대의 젊은 층의 피해가 급증하고 있는 상황과 직결된다. 이를 위해 본 논문에서는 실제 보이스피싱에 사용된 보이스피싱의 음원과 보이스 피싱 범죄자와 동년배의 일반 젊은 층을 대상으로 같은 문장을 읽고 이에 대한 음성을 비교 분석을 행하였다. 실험은 2011년 이후 최근까지 보이스 피싱을 행한 범죄자 목소리의 음성적 특징이 무엇인가를 음성의 음높이와 그 대역폭, 음성에 실리는 에너지 및 발화속도, 음색 등을 기반으로 수행하였다. 실험 결과 음성에 실리는 에너지와 발화속도에 있어 의미 있는 차이가 있음을 규명해 낼 수 있었다.

**Key Words** : voice analysis, voice fishing, voice color, crime prevention, identification

## ABSTRACT

The scale of financial damages on voice fishing has not been decreased despite of national and social efforts to reduce the amounts of voice fishing damage. One of these reasons is a sophisticated and vernacular speech style that makes it difficult to recognize the offenders. Furthermore, nowadays, young men have intensively been deceived by not only sophisticated and vernacular speech style which is used the employer of real public offices but also obtained personal information. As a result, this lead directly to the financial damages of younger people who has a stronger judgement than older. For this, we investigated the comparison and analysis between the criminals of voice fishing and the same generation younger people for identifying voice features. The experiment was carried out based on the pitch, bandwidth of pitch, energy, speech speed and voice color for searching the difference of voice characteristics between the criminals of voice fishing and the same generation younger people since 2011. The experimental result shows that there is a significant difference in energy and speech speed between the criminals of voice fishing and the same generation younger people.

\* First Author : Youngdong Fire Stations, stegnography@korea.kr, 종신회원

° Corresponding Author : Chungbuk Provincial University, ducho@cpu.ac.kr, 종신회원,

\* Gangneung-Wonju National University, ymjeong@gwu.ac.kr, 종신회원

논문번호 : KICS2016-08-179, Received August 1, 2016; Revised September 29, 2016; Accepted September 29, 2016

## I. 서론

최근 진화하고 있는 보이스피싱으로 인하여 이에 따른 경제적 재산 피해 규모가 줄어들지 않고 있는 실정이다. 사실 초기의 보이스 피싱은 특정 지역의 어눌한 사투리와 억양 및 한국 실정에 맞지 않는 내용으로 말을 하다 보니 보다 용이하게 보이스피싱이라는 판단 및 판별이 가능하였으나 최근의 보이스 피싱 조직은 한국인이 보이스 피싱 전문 범죄 조직에 전문적으로 함께 참여 하여 보다 지능화하여 한국인의 정서와 말투 및 실정에 맞도록 콘텐츠를 개선함과 더불어 유출된 개인정보를 바탕으로 이를 숙지하여 범행을 저지르는 것으로 나타나고 있다<sup>1)</sup>. 실제로 아래 표 1의 보이스 피싱 통계에서 알 수 있듯이 2012년 ~ 2015년 7월까지 발생한 총 보이스피싱 범죄건수는 23,500건이며 여기에 따라 검거된 인원은 29,026명, 그리고 그 피해액은 2,095억원에 이르고 있는 실정이다. 더욱 문제가 심각한 것은 이에 대한 국가적, 사회적 대책 마련과 법 집행이 강하게 이루어지고 있음에도 불구하고 최근까지도 보이스피싱에 의한 범죄 피해가 급증하고 하고 있는 것으로 나타나 그 심각성이 더욱 커지고 있다. 한 예로 2015년 1월 ~ 7월 사이에 발생한 보이스 피싱 범죄건수 5,391건 7,378명이며, 재산상 피해액이 784억으로 나타나 그 문제점이 시간이 경과될수록 심화되고 있는 현상을 보이고 있다. 따라서 이 같은 문제를 해결하기 위해서는 보이스피싱의 음성 특징을 규명해내는 작업이 이루어져야 한다. 그러나 이를 위한 기존의 연구는 전무한 실정이다. 따라서 이를 위해 본 논문에서는 최근의 보이스피싱의 음성적 특징은 어떤 특징이 있는지에 대한 규명 작업을 행하고자 한다. 이 같은 작업을 달성하고자 본 논문에서는 다음과 같은 두 가지 작업을 수행하고자 한다. 첫째, 보이스피싱을 행한 범죄자가 말한 문장을 동년배의 일반인이 똑 같이 해당 문장을 읽게 하여 그 실험 음성군의 음성 특징 차이가 어떤 것이 있는지에 대한 비교 분석을 행하고자 한다. 둘째, 유튜브 등에서 평범하지만 일반적인 보이스피싱의 음성이라고 여겨질 수 있는 보이스 피싱 음원을 대상으로 보이스피싱을 행하는 범죄자들의 공통적인 음성 특징은 무엇인가에 대한 규명 작업을 행하고자 한다. 이 같은 두 가지 작업을 통해 일반인과 보이스 피싱 범죄자들 간의 음성 특징의 차이 그리고 보이스피싱 범죄자들이 가지고 있는 차별성 있는 음성 특징은 어떤 것인가에 대한 규명작업을 행하고자 하며 차후 이 같은 보이스피싱 범죄자들의 음성 특징 규명을 통해 보이스피싱 예방 앱을 개발하고자 한다.

표 1. 보이스 피싱 피해에 대한 통계자료<sup>2)</sup>  
Table 1. Statistics data of voice fishing<sup>2)</sup>

	2012 Year		2013 Year	
	the number of Voice fishing	financial damage (billion won)	the number of voice fishing	financial damage (billion won)
Total	5,709	59.5	4,765	55.2
Seoul	1,934	21.6	1,647	18.5
Busan	571	7.3	315	4.3
Daegu	358	3.9	351	4.4
Incheon	347	3.5	217	2.3
Gwangju	52	0.6	48	0.6
Daejeon	260	2.7	130	1.7
Ulsan	193	1.6	131	1.6
Gyeonggi	521	5.2	833	9.0
Gangwon	84	1.0	90	1.2
Chungbuk	196	1.8	138	1.8
Chungnam	330	2.5	148	1.5
Jeonbuk	76	0.6	149	1.6
Jeonnam	145	1.0	118	1.2
Gyeongbuk	280	2.6	222	2.6
Gyengnam	262	2.4	168	2.1
Jeju Island	100	1.2	60	0.8

	2014 Year		2015 Year (January to July)	
	the number of voice fishing	financial damage (billion won)	the number of voice fishing	financial damage (billion won)
Total	7,635	97.4	5,391	78.4
Seoul	2,822	34.2	1,959	30.1
Busan	517	7.0	338	4.4
Daegu	470	6.4	269	3.9
Incheon	388	5.2	285	4.2
Gwangju	230	2.5	88	1.3
Daejeon	136	1.7	178	2.6
Ulsan	248	4.0	164	2.0
Gyeonggi	1,479	17.7	1,131	16.4
Gangwon	140	1.9	121	1.2
Chungbuk	140	1.3	107	2.1
Chungnam	141	4.0	188	1.7
Jeonbuk	238	2.8	105	1.6
Jeonnam	231	3.1	132	2.8
Gyeongbuk	198	2.2	144	1.9
Gyengnam	18	2.5	108	1.3
Jeju Island	74	0.9	74	0.9

## II. 실험 음원의 선정 및 실험에 사용한 음성 분석요소

우선 실험에 필요한 실험 음원의 선정 및 획득은 유튜브에서 2011년 ~ 2015년 보이스피싱에 대한 음원을 찾아 추출하여 실험을 수행하였다. 또한 유튜브에서 획득. 선정된 음원을 바탕으로 보이스피싱 범죄

자의 목소리와 동년배의 일반인들 간의 음성 차이를 분석하기 위해 음성의 전달력이 좋은 20대 본 대학 재학생의 음성을 녹취하여 실험 음원으로 선정하였다. 아울러 일반인의 음성을 20대 대학 재학 중인 사람들로 선정한 것은 보이스피싱에 동원되는 범죄자들이 20대가 많은 관계로 동년배간의 음성 차이를 알아보기 위함이며, 20대 대학 재학 중인 사람들의 경우 음성의 전달력 등이 좋은 나이대인 관계로 실험 대상자들로 선정하였다. 이때 음성 분석 프로그램은 프라트<sup>[8]</sup>를 이용하였으며 음성 분석 도구로는 음 높이, 음 높이의 편차, 강도, 지터, 짐머, NHR값 및 발화 속도 등을 선정하였다. 따라서 보이스피싱 범죄자의 음성과 일반인들의 음성 간에 음 높이와 음 높이 편차의 차이는 어느 정도인지 그리고 음성의 에너지 및 공신력과 연계된 음색 음성 분석 요소들 간의 편차는 어느 정도 발생하는지 등에 대한 분석을 행하였다. 최종적으로 음성 분석 요소 중 가장 중요한 요소인 발화속도의 차이는 어느 정도 편차를 보이는 지에 대한 분석을 행하였다.

### 2.1 음 높이(pitch)

음 높이란 일상적인 개념과 같이 청각적으로 느끼는 음의 높이를 말한다<sup>[4]</sup>. 공학적으로는 1초 동안 성대의 진동이 몇 회 있었는가를 의미하는데 성대의 크기와 길이, 질량 등에 영향을 받는다. 예로서 여성의 경우 남성 보다 성대가 짧고 그 크기가 작은 관계로 이 수치가 보다 높다. 또한 음의 높이인 사람의 감정과 정서의 변화에 따라 그 높이에 변화가 발생하며 통상 평균값을 음 높이를 대표하는 값으로 사용하고 있지만 카메라 셔터 소리나 특정한 소리의 음원으로 인해 이 값이 영향을 받는 경우가 있으므로 평균값(Mean)과 중간값(Median)을 비교해 볼 필요가 있다.

### 2.2 음 높이의 편차

음 높이의 편차란 음 높이의 최대값과 최소값의 편차를 나타낸다<sup>[5]</sup>.

### 2.3 음에 실리는 에너지

음에 실리는 힘 즉, 에너지는 음성파형에서 소리의 크기를 나타내는 척도를 의미한다. 일반적으로 여성의 경우 60[dB]대이며, 남성의 경우는 70[dB]대이다. 강조하고 싶은 경우나 음성에 에너지를 느끼게 할 경우 이 값을 증가시키며, 역으로 자신감이 떨어지거나 부드러움 등을 나타내고자 할 때는 이 값을 적게 한다<sup>[6]</sup>.

### 2.4 발화속도

발화속도란 1분 당 몇 음절을 말하느냐를 의미한다. 일반적으로 1분에 300음절대의 발화 속도를 보일 경우 보통 빠르기로 말하는 경우로 해석한다<sup>[7]</sup>.

### 2.5 음색

음색은 말하는 사람의 공신력과 가장 밀접한 관계를 갖는다<sup>[8,9]</sup>. 다시 말해 음색이 풍부할 경우 말하는 사람의 공신력이 높게 여겨지며, 역으로 음색의 거친 경우 발화자의 공신력이 떨어지는 소리로 들리게 된다. 일반적으로 음색을 구성하는 요소는 크게 지터<sup>[10]</sup>와 짐머<sup>[11]</sup> 그리고 NHR<sup>[12]</sup> 등으로 구성되며 이 중 지터(Jitter)는 단위시간 안의 발음에서 성대의 진동인 피치의 변화가 얼마나 많은지를 나타내준다. 이에 비해 짐머(Shimmer) 분석 요소는 음성파형에서 각 지점의 진폭 값의 변화가 얼마나 규칙적인지를 나타내 주는 것으로, 상대진폭의 변이 양상을 보여주어 진폭의 정확도를 의미한다. 또한 NHR은 소음 대 배음비(Noise-to-Harmonics Ratio)로 단위는 [%]를 사용한다.

## III. 보이스피싱 범죄자들과 동년배의 일반인들 간의 음성 차이에 대한 실험 결과

보이스피싱 범죄자들의 음성 특징을 규명하기 위해서는 일단 이들의 음성과 연령대가 같고 성별이 같은 일반인들을 선정하여 보이스피싱 범죄자들의 음성 특징 차이가 있는지 그리고 만일 있다면 어떤 요소에서 있는지를 규명해 보는 것이 가장 중요한 일이 된다. 이를 위해 유튜브에서 획득한 보이스피싱 범죄자들의 음성 특징을 추출해 내고 범죄에 사용된 똑같은 문장을 동년배 및 같은 성별의 일반인들이 읽게 하여 음성 특징을 추출한 후 두 집단 간의 음성 특징 차이를 파악해내고자 한다. 첫 번째 실험 음원은 보이스 피싱 범죄자가 “최근 사회적으로 많은 문제가 되고 있는 개인 정보 유출에 해당 사항이 있으십니까?”라는 음성이고, 두 번째 음원은 “아, 저희가 이걸 여쭙보는 이유는 저희가 2014년 4월 25일 불법도박금융사기단 김형태 일당을 검거 했는데요”라는 음성이다. 마지막으로 세 번째 음원의 경우 보이스 피싱 범죄자가 “아, 지금 본인명의로 도움이 되셔가지고요, 지금 대포통장 금융사기 업체에 연루가 되셨습니다”라는 음성이다. 이에 따른 실험 결과로 아래 표 2에 음원 1에 대한 범인과 일반인들의 음 높이, 음 높이의 편차 그리고 강도에 대한 실험 결과 수치를 나타내었으며, 표 3에 음원 1에 대해 음색과 연계된 지터, 짐머, NHR 값 및 발화

표 2. 음원 1에 대한 평균 음높이, 대역폭 및 음성에 실리는 힘  
Table 2. Average Pitch, Pitch bandwidth & energy of source 2

	Average Pitch[Hz]	Pitch Bandwidth[Hz]	Energy[dB]
criminal 1	141	95	64
ordinary people 11	118	140	76
ordinary people 12	141	160	73
ordinary people 13	142	101	73
ordinary people 14	155	75	69
ordinary people 15	137	74	74
Average of ordinary people	138.6	110	73

표 3. 음원 1에 대한 음색 및 발화속도  
Table 3. Results of related to voice color and speech speed for source 1

	Jitter[%]	Shimmer [dB]	NHR[%]	Speech Speed
criminal 1	2.1	1.319	0.224	494
ordinary people 11	1.5	1.279	0.183	470
ordinary people 12	1.9	1.363	0.183	509
ordinary people 13	1.9	1.263	0.168	505
ordinary people 14	1.6	1.186	0.13	530
ordinary people 15	1.5	1.018	0.135	447
Average of ordinary people	1.68	1.2218	0.1598	492.2

표 4. 음원 2에 대한 평균 음높이, 대역폭 및 음성에 실리는 힘  
Table 4. Average Pitch, Pitch bandwidth & enrgy of source 2

	Average Pitch[Hz]	Pitch Bandwidth[Hz]	Energy[dB]
criminal 1	127	109	63
ordinary people 11	112	171	63
ordinary people 12	128	136	71
ordinary people 13	137	111	70
ordinary people 14	149	133	71
ordinary people 15	114	52	67
Average of ordinary people	128	120.6	68.4

속도에 대해 실험한 결과수치를 표하였다. 마찬가지로 표 4와 표 5는 음원 2에 대한 비교 실험 결과값을, 그리고 표 6과 표 7은 음원 3에 대한 실험 결과 수치를 나타내었다.

표 5. 음원 2에 대한 음색 및 발화속도  
Table 5. Results of related to voice color and speech speed for source 2

	Jitter[%]	Shimmer [dB]	NHR[%]	Speech Speed
criminal 1	1.882	1.269	0.167	632
ordinary people 11	1.951	1.244	0.215	494
ordinary people 12	1.791	1.098	0.139	452
ordinary people 13	1.666	1.071	0.152	423
ordinary people 14	1.758	1.013	0.105	445
ordinary people 15	1.697	1.063	0.148	435
Average of ordinary people	1.772	1.097	0.1518	450

표 6. 음원 3에 대한 평균 음높이, 대역폭 및 음성에 실리는 힘  
Table 6. Average pitch, pitch bandwidth & energy of source 3

	Average Pitch[Hz]	Pitch Bandwidth[Hz]	Energy[dB]
criminal 1	145	167	65
ordinary people 11	123	57	76
ordinary people 12	145	167	69
ordinary people 13	139	70	72
ordinary people 14	151	205	66
ordinary people 15	134	87	71
Average of ordinary people	138.4	117.2	70.8

표 7. 음원 3에 대한 음색 및 발화속도  
Table 7. Average pitch, pitch bandwidth & energy of source 3

	Jitter[%]	Shimmer [dB]	NHR[%]	Speech Speed
criminal 1	2.186	0.990	0.123	419
ordinary people 11	1.541	1.188	0.224	286
ordinary people 12	2.186	0.990	0.123	419
ordinary people 13	1.513	0.971	0.133	291
ordinary people 14	1.439	1.124	0.095	361
ordinary people 15	1.751	1.041	0.129	358
Average of ordinary people	1.686	1.062	0.140	343

위의 실험 결과를 정리. 요약하면 아래 표 8 및 표 9와 같다.

위의 표 8과 표 9에서 알 수 있듯이 남성 범인들과 일반인들의 경우 음 높이, 음 높이가 편차, 음색 관련 수치들의 경우 그 차이가 크게 발생하지 않음을 알 수 있었다. 이 이유 중 하나는 남성 보이스피싱 범죄자와 같은 동년배의 음성이기 때문이다. 그러나 같은 동년배 남성의 음성이라 할지라도 유의미하게 확실한 차이를 보이는 부분이 있는데 그것이 바로 음성에 실리는 에너지와 발화속도 부분이다. 이 중 음성에 실리는 에너지 값이 작은 것은 부드럽게 말을 해서 마치 관공서 사람처럼 여기게 하려는 의도가 있는 것으로 여겨지며, 또 다른 이유는 사실상 거짓말을 하고 있는 것이어서 심적으로 이에 대한 부담감으로 말미암아 음성에 에너지가 제대로 실리지 않는 것으로 해석된다. 마지막으로 발화속도는 일반인들에 비해 빠른 속도여서 실제 해당 분야에 종사하는 사람임을 느끼게 하려는 것과 같은 의도가 내포되어 있거나 또는 같은 문장의 말을 상당히 여러 번 반복함으로써 자연스럽게 해당 문장이 입에 배어 발화속도가 빨라진 것으로 여겨진다.

표 8. 평균 음높이, 대역폭 및 음성에 실리는 힘에 대한 실험 결과의 요약정리  
Table 8. Summary of experimental results for average pitch, pitch bandwidth & energy

	Average Pitch[Hz]	Pitch Bandwidth[Hz]	Energy[dB]
Sound source1	2.4	-15	-9.0
Sound source2	-1.0	-11.6	-5.4
Sound source3	6.6	49.8	-5.8

표 9. 음색 및 발화속도에 대한 실험 결과의 요약정리  
Table 9. Summary of experimental results for related to voice color and speech speed

	Jitter[%]	Shimmer[dB]	NHR[%]	Speech Speed
Sound source1	0.4200	0.0972	0.0642	1.8
Sound source2	0.1094	0.1712	0.0152	182
Sound source3	0.5000	-0.0728	-0.025	76

#### IV. 보이스피싱 범죄자들만의 음성에 대한 특징 규명 실험 결과

보이스피싱 범죄자들의 음성 특징을 규명하기 위해

유튜브에서 2011년부터 2015년까지 보이스피싱 범죄자들의 음원을 녹취하여 이에 대한 실험을 행하였다. 이 실험의 목적은 일반인을 배제하고 보이스피싱 범죄자들만의 음성 특징을 분석해 보고자 함이다. 이를 위해 실험을 수행한 결과 아래 표 10에 2011년부터 2015년까지의 실제 보이스피싱 범인들에 대한 실험 결과가 도출되었다. 또한 아래 그림 1 ~ 그림 5에 각년도 별 보이스피싱 범죄자들에 대한 실험 결과 과정의 예를 나타내었다. 이제 실험 결과가 갖는 의미를 분석해 보면 우선 아래 표 10에서 알 수 있듯이 평균 음높이는 남성의 경우 131[Hz], 여성의 경우 257[Hz]로 나타났다. 이는 일반적인 남성과 여성의 평균값과 유사한 결과로 동년배의 일반인들과 큰 차이가 없는 실험 결과 자료이다. 이에 비해 차별성을 보이며 유의미한 실험 결과 자료가 나온 부분은 바로 음성에 실리는 에너지 즉, 음성의 강도가 남성 보이스피싱의 범죄자의 경우 전체 평균이 64[dB]가 나왔다는 사실이다. 아울러, 말하는 속도인 발화속도 역시 분 당 414 음절로 동년대의 일반인들 평균인 분당 300 음절대를 넘어서는 빠른 발화속도를 보이고 있다. 이는 앞 3장의 결론과 동일한 결과를 보이고 있는 부분이다. 즉, 남성 보이스피싱 범죄자들은 음성의 에너지를 낮게 가져감으로서 그 음성을 듣는 사람들에게 부드러운 느낌을 느끼게 하려는 의도와 더불어 보이스피싱이란 근본적으로 거짓말을 하고 있는 경우이기 때문에, 음성에 실리는 에너지가 심적 요인에 의해 제대로 안 실리고 있는 것으로 해석된다. 발화속도 부분은 말을 빨리함으로써 해당 분야 관공서에서 근무하는 전문 담당자로 느끼게 하려는 의도와 더불어 반복되는 숙지된 말인 관계로 발화속도가 빠른 것으로 여겨진다. 이에 비해 여성 보이스 피싱 범죄자의 경우 음성에 실리는 에너지가 평균 69[dB]로 여성의 평균치 보다 음성에 실리는 힘이 더 높은 수치를 보이고 있다. 또한 말하는 속도도 일반인 평균보다는 높은 수치를 보이고 있다. 결국 여성 보이스피싱 범죄자들은 공공기관이나 관공서에 근무하고 있다는 것을 의도적으로 나타내기 위해 여성 특유의 부드러움보다는 역으로 강함을 선택하여 말을 하고 있는 것으로 분석된다. 아울러 남성 보이스피싱 범죄자나 여성 보이스피싱 범죄자에 관계없이 전체적인 문체의 심각성은 일부음원의 경우는 유출된 개인정보를 기반으로 하여 이를 숙지함과 함께 세련된 말투 및 전문용어를 구사함으로써 피해자로 하여금 실제 해당 분야 종사자의 전화로 느껴지게 할 의도가 있는 것으로 분석되어 그 심각성이 더해지고 있는 관계로 개인 정보 유출에 대한 국가적 대비책

도 강하게 행해져야 할 것으로 여겨진다.

표 10. 2011 ~ 2015년 음원에 대한 실험 결과  
Table 10. Experimental results of 2011~2015 Year Sound Source

	Average Pitch [Hz]	Bandwidth [Hz]	Energy [dB]	Jitter [%]
2011 Year	134	116	64	1.782
2012 Year	253	284	70	2.083
2013 Year	261	200	67	1.933
2014 Year	122	155	64	2.194
2015 Year	137	123	64	2.056
Male average	131	132	64	2.011
Female average	257	242	69	2.008
Whole average	182	176	66	2.010

	Shimmer [dB]	NHR [%]	Speech Speed	Sex
2011 Year	1.314	0.257	393	male
2012 Year	1.031	0.099	384	female
2013 Year	1.172	0.178	389	female
2014 Year	1.191	0.242	388	male
2015 Year	1.193	0.171	515	male
Male average	1.232	0.223	432	male
Female average	1.102	0.138	386	female
Whole average	1.180	0.189	414	whole

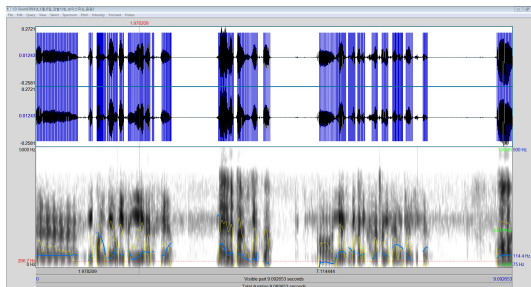


그림 1. 2011년 보이스피싱 음원에 대한 분석 결과 그림  
Fig. 1. Result picture for speech analysis of voice fishing in 2011

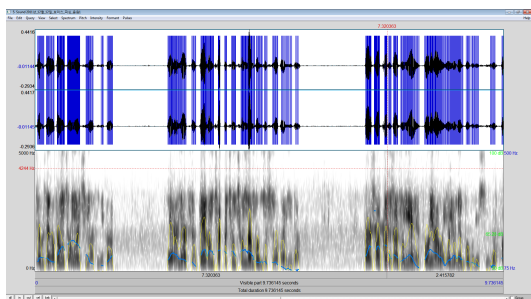


그림 2. 2012년 보이스피싱 음원에 대한 분석 결과 그림  
Fig. 2. Result picture for speech analysis of voice fishing in 2012

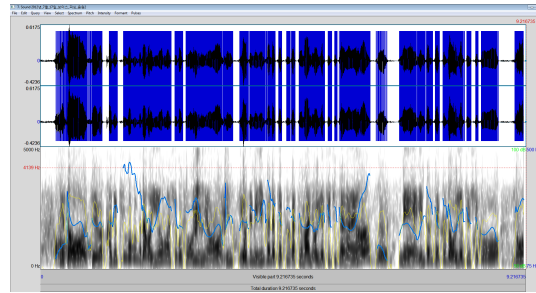


그림 3. 2013년 보이스피싱 음원에 대한 분석 결과 그림  
Fig. 3. Result picture for speech analysis of voice fishing in 2013

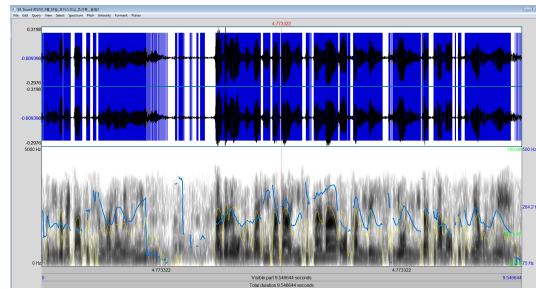


그림 4. 2014년 보이스피싱 음원에 대한 분석 결과 그림  
Fig. 4. Result picture for speech analysis of voice fishing in 2014

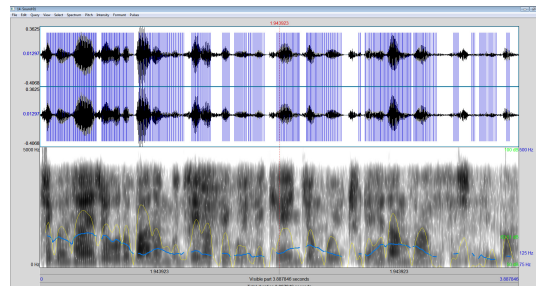


그림 5. 2015년 보이스피싱 음원에 대한 분석 결과 그림  
Fig. 5. Result picture for speech analysis of voice fishing in 2015

## V. 결 론

본 논문에서는 사회적으로 큰 문제가 되고 있는 보이스피싱의 음성 특징은 무엇이고 또 그 음성에는 어떠한 특징과 의미가 있는지를 분석 및 규명을 수행해 보는 연구를 수행하였다. 이를 위해 보이스피싱을 행한 범죄인들 그리고 그들과 동년배 사람들 간의 음성 차이가 무엇이 있는지에 대한 규명 연구 및 2011년부터 2015년까지 보이스 피싱을 행한 사람들의 음성 특징을 추출해 내는 연구를 수행하였다. 연구를 수행한

결과 보이스 피싱을 행한 범죄자들의 경우 일반인들에 비해 상대적으로 음성에 실리는 에너지와 말하는 속도 즉, 발화속도에서 있어 가장 큰 유의미한 차이가 있음을 확인해 낼 수 있었다. 다시 말해, 보이스피싱을 행하는 범죄자들은 일반인들과 달리 음성에 실리는 에너지 값이 작고 말하는 속도는 빨리하는 것으로 나타났다. 이는 첫째로, 관공서에서 일반 민원인들에게 접근하는 방식 및 숙지한 개인정보와 범행에 사용할 전문용어를 숙지 후 범행을 시도함으로써 사람들로 하여금 실제로 해당 분야 관공서에서 근무하는 사람으로 착각하게 하여 전화 통화 내용을 믿게 하고자 하는 의미가 있음을 알 수 있었다. 둘째로는, 훈련을 통해 보이스 피싱 해야 할 해당 문장을 익히고 이를 계속적으로 반복되어 말하게 됨으로서 문장 및 단어에 대한 숙지도로 말미암아 발화속도가 빨라지며 음색이 풍부하다는 분석이 가능하다. 향후는 이 같은 실험 결과를 바탕으로 하여 관련 기관과의 협조를 통해 보다 다양한 보이스피싱에 대한 음원 분석 작업을 수행하여 이 음성이 보이스피싱 음성인 지 아니면 실제 관공서에서 건 음성인 지를 구분해 내는 시스템 개발에 대한 연구를 지속적으로 행해야 하리라 여겨진다. 아울러 예로서 수사기관을 사칭할 때와 대출 업자를 사칭할 때의 음성 등에 차이가 발생할 수밖에 없다. 금 번 논문은 주로 수사 기관 등을 사칭할 때의 음성을 기반으로 한 실험이지만 앞으로 보이스 피싱의 전체 유형을 구분해내고 이에 각 유형별 음성 차이와 특징을 구분해내는 연구도 지속적으로 수행하고자 한다. 마지막으로 보이스 피싱을 행하는 시간 중 보이스 피싱에 걸려들기 전과 걸려든 후의 발화속도를 중심으로 한 음성 특징을 규명해 내는 작업도 행할 예정이다.

## References

- [1] M. H. Yang, et al., "Features identification of 'Voice Fishing' voice by using voice analysis," in *Proc. KICS Conf. 2015(KICS Fall Conf. 2015)*, Seoul National Univ., Korea, Nov. 2015
- [2] S. K. Lim(Member of the National Assembly in Korea), *Voice fishing statistics*, 2015.
- [3] B. G. Yang, *Theory and Reality of Voice Analysis Using Praat*, ManSu Publishing Co., 2003.
- [4] D. U. Cho, et al., "Application of vocal fold vibration analysis parameter for infant congenital heart diseases diagnosis," *J. Korea Acad.-Ind. Cooperation Soc.*, vol. 10, no. 10, pp. 2708-2714, Oct. 2009.
- [5] D. U. Cho and B. H. Kim, "Application of preference for Korean pop music by applying acoustic signal analysis techniques," *J. KIPS*, vol. 19-D, no. 3, pp. 211-219, Jun. 2012.
- [6] D. U. Cho and B. H. Kim, "Pronunciation influence analysis of carbonate drink and eucalyptus fragrance by applying speech signal processing techniques," *J. KICS*, vol. 37-C, no. 5, pp. 420-428, May 2012.
- [7] D. U. Cho, et al., "Voice features extraction of lung diseases based on the analysis of speech rates and intensity," *J. KIPS*, vol. 16-B, no. 6, pp. 471-478, Dec. 2009.
- [8] R. H. Park, "A study on the communication effect of voice analysis elements," *J. Speech & Commun.*, vol. 11, no. 11, pp. 293-327, Dec. 2009.
- [9] D. U. Cho, "The communicability observations of broadcasting programs MC by extracting voice feature," *J. KBS*, vol. 59, no. 6, pp. 36-73, Dec. 2009.
- [10] B. H. Kim, et al., "Variation measurement and analysis of jitter and shimmer parameter value by hemodialysis in diabetic and hypertensive," *J. KICS*, vol. 36-C, no. 7, pp. 834-840, Jul. 2012.
- [11] H. B. Kang, et al., "Identification of voice features changes by era of representative announcer," in *Proc. KICS Conf. Commun.*, Jeju Island, Korea, Jun. 2015.
- [12] Y. M. Jeong, et al., "Voice features analysis of broadcaster by applying ICT technologies," in *Proc. KICS Conf. Commun.*, Jeju Island, Korea, Jun. 2015.

**이 범 주 (Bum Joo Lee)**



2006년 8월 : 충북대학교 컴퓨터공학 학사  
2016년 2월~현재 : 충북대학교 컴퓨터과학과 석박사 통합과정  
2009년~현재 : 영동소방서 화재조사관

1996년~현재 : 화재감식평가기사 외 17개 자격증 취득  
2016년 6월 한국통신학회 우수논문상  
<관심분야> 음성분석, 빅데이터, 국가재난 시스템구축

**정 연 만 (Yeon Man Jeong)**



1983년 2월 : 송실대학교 전자공학  
1985년 2월 : 송실대학교 전자공학 석사  
1991년 8월 : 송실대학교 전자공학 박사  
1993년~현재 : 강릉원주대학교 정교수

<관심분야> 음성신호처리, 통신신호처리, 무선통신 시스템, RF IC 설계

**조 동 욱 (Dong Uk Cho)**



1983년 2월 : 한양대학교 전자공학 학사  
1985년 8월 : 한양대학교 전자공학 석사  
1989년 2월 : 한양대학교 전자통신공학과 박사  
1989년 3월~1990년 2월 : 한양

대학교 박사후과정 연구원  
1989년 9월~1991년 2월 : 동양미래대학교 통신공학과 교수  
1991년 3월~2000년 2월 : 서원대학교 정보통신공학과 교수  
1999년 : 미국 Oregon State University 교환교수  
2000년 3월~현재 : 충북도립대학교 교수  
2007년 9월 : 대통령 표창  
2008년 12월 : 한국정보처리학회 학술대상  
2012년 11월 : 한국통신학회 LG 학술대상  
<관심분야> 음성 분석, 신호처리