

음성 주파수 분포 분석을 통한 편집 의심 지점 검출 방법

A Speech Waveform Forgery Detection Algorithm Based on Frequency Distribution Analysis

허희수¹⁾ · 소병민²⁾ · 양일호³⁾ · 유하진⁴⁾

Heo, Hee-Soo · So, Byung-Min · Yang, IL-Ho · Yu, Ha-Jin

ABSTRACT

We propose a speech waveform forgery detection algorithm based on the flatness of frequency distribution. We devise a new measure of flatness which emphasizes the local change of the frequency distribution. Our measure calculates the sum of the differences between the energies of neighboring frequency bands. We compare the proposed measure with conventional flatness measures using a set of a large amount of test sounds. We also compare- the proposed method with conventional detection algorithms based on spectral distances. The results show that the proposed method gives lower equal error rate for the test set compared to the conventional methods.

Keywords: forgery detection, frequency distribution, spectral distance

1. 서론

최근 스마트폰과 같은 디지털 기기들이 보급됨에 따라 어디서나 손쉽게 오디오 파일을 녹음하는 것이 가능해졌다. 이에 따라, 디지털 오디오 파일이 법정 증거자료로서 제출되는 경우도 증가하고 있다. 하지만 디지털 오디오 파일은 컴퓨터로 복사된 뒤, 음성을 편집하는 것이 용이하기 때문에 증거로서 효력을 얻기 위해서는 무결성이 입증되어야 한다. “골드웨이브”나 어도비 사의 “오디션” 등 다양한 오디오 편집 프로그램들이 보급됨에 따라 비전문가도 디지털 오디오 파일을 손쉽게 위변조할 수 있는 반면, 오디오 파일의 위변조 여부를 판별하기 위해서는 음성학, 음성공학 등의 전문지식을 갖춘 전문가의 세밀한 분석이 필요하다. 따라서 증거물로 제출된 오디오 파일의 위변조 여부를 판별하는 데에는 많은 비용이 필요하게 되

므로, 이 과정을 자동화하거나 분석관을 보조할 수 있는 기술이 개발된다면 위변조 여부 판별에 드는 비용을 크게 줄일 수 있을 것이다. 이에 따라 오디오 파일의 위변조 여부 판별에 활용될 수 있는 다양한 연구들이 진행돼 왔다[1]-[9]. 여기에는 발전소에서 보내지는 전기의 주파수를 이용하는 방법, 원본 음성의 채널 특성을 이용하는 방법, 음성이 연결되는 지점의 특성을 파악하는 방법 등이 있다.

Electric network frequency(ENF)는 오디오 파일의 녹음 과정에서 음성과 함께 녹음될 수 있는 고유한 주파수 대역의 신호를 의미한다[1]-[4]. ENF는 전원이 연결된 전자기기에서 발생하며 전원을 공급하는 발전소와 그 발생 시기에 따라서 고유한 패턴을 지니게 된다. 이 신호는 매 순간 변화하는 고유한 패턴을 지니므로 이를 위변조 판별에 활용하는 것이 가능하다. 예를 들어, 매 순간의 ENF 신호와 측정 시각을 함께 기록해 두면 녹음 시각이 알려진 오디오 파일의 진위 여부를 판별하는 데 활용할 수 있다. 특정 시각에 녹음된 것으로 알려진 오디오 파일에 포함된 ENF 신호와 그 시각에 미리 수집해 두었던 ENF 신호에 차이가 있다면 오디오 파일의 시간 정보가 조작됐다는 것을 유추할 수 있기 때문이다. 또한, ENF는 각 시각에서 고유한 패턴을 가지기 때문에 오디오 파일에 포함된 ENF 신호를 미리 기록해 두었던 신호와 비교하여 고유한 패턴이 동일하게 발생하는지 확인 함으로써 오디오 파일의 위변조 여부를 판별할 수도 있다. 하지만 전원이 연결되지 않은 녹음기

1) 서울시립대학교, zhasgone@naver.com

2) 대검찰청, sbm1210@spo.go.kr

3) 서울시립대학교, heisco@hanmail.net

4) 서울시립대학교, hjyu@uos.ac.kr, 교신저자

이 논문은 2015년 대검찰청 연구용원의 지원으로 수행되었습니다. (과제명: 디지털 녹음 파일 위변조 여부 분석 기법 연구)

접수일자: 2015년 12월 3일

수정일자: 2015년 12월 19일

게재결정: 2015년 12월 19일

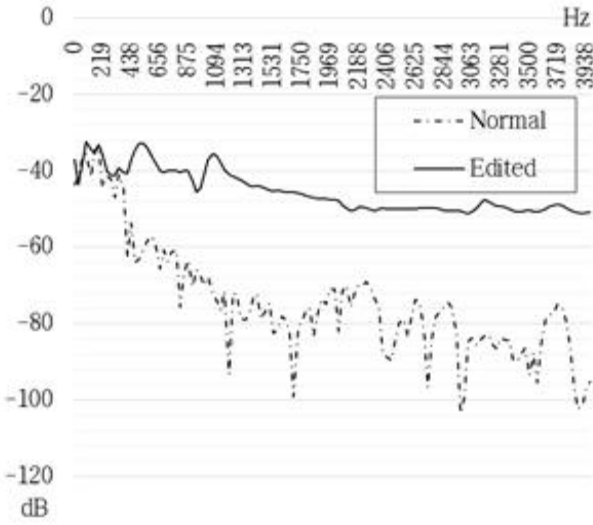


그림 1. 단절 지점을 포함하는 음성 구간에서의 스펙트럼(Edited) 과 일반적인 음성 구간에서의 스펙트럼(Normal)

Figure 1. Spectrum on an edited wave and a normal wave

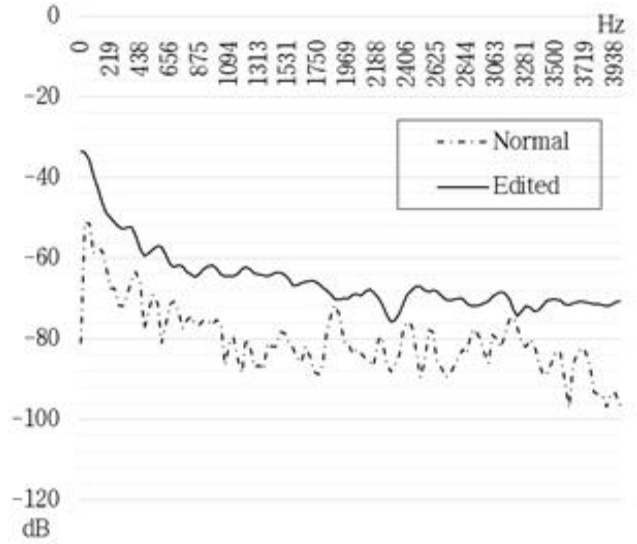


그림 2. 단절 지점을 포함하는 구간에서의 스펙트럼(Edited)과 일반적인 비 음성 구간에서의 스펙트럼(Normal)

Figure 2. Spectrum on an edited and a normal wave (without voice)

기를 사용하는 경우, ENF를 확인하기 어려울 수 있다.

다른 위변조 판별 연구로는 오디오 파일의 채널을 식별해 위변조 여부를 판별하는 연구들이 있다[5]-[7]. 예를 들어, 오디오 파일을 편집할 때, 다른 채널에서 녹음된 다수의 오디오 파일들을 이어 붙여 하나의 오디오 파일을 생성할 수 있다. 정상적인 오디오 파일이라면 모든 구간의 채널이 동일하지만 채널이 다른 여러 원본 오디오 파일을 이어 붙였다면 채널이 다른 구간이 존재하게 된다. 따라서 이러한 편집의 경우, 오디오 파일 각 구간의 채널을 식별하여 편집 여부를 판별할 수 있다. 하지만 동일한 채널에서 녹음한 여러 오디오 파일들을 이어 붙여 편집한 경우, 혹은 하나의 원본 오디오 파일에서 특정 구간을 삭제하여 편집한 경우에는 이 방법으로 편집 여부를 판별할 수 없다.

디오 파일의 편집 의심 지점을 찾는 위변조 판별 방법의 경우, 앞에서 언급했었던 제약 사항들(ENF 녹음 여부, 위변조 파일 내 채널 차이 여부) 없이 위변조를 판별할 수 있다[4], [8]. 이 방법은 오디오 파일을 세밀하게 분석하여 파일 내에 오디오 편집으로 인해 발생한 것으로 추정되는 단절을 포함하는 편집 의심 지점을 찾는다. 본 논문에서는 편집 의심 지점을 찾는 방식의 위변조 여부 판별 연구를 주로 다룬다.

본 논문의 다음 장에서는 오디오 파일 상의 편집 의심 지점을 찾는 기존 연구인 spectral distance 기반의 알고리즘을 소개한다. 3장에서는 편집 의심 지점을 찾기 위하여 제안한 스펙트럼 분석 알고리즘을 소개하고, 4장에서 기존의 알고리즘과 성능을 비교하는 실험 및 결과를 기술한다. 5장에서는 본 연구의 결론을 맺고 향후 연구 계획을 소개한다.

2. Spectral Distance 기반 편집 의심 지점 검출

오디오 파일 내 각 프레임의 spectral distance를 분석하면 음성 스펙트럼의 급격한 변화를 탐지해 편집 의심 지점을 탐색할 수 있다. 음성이 임의로 편집되는 경우, 편집된 지점의 단절 때문에 인접한 구간과 비교할 때 스펙트럼의 급격한 변화가 발견될 수 있다. 스펙트럼의 급격한 변화를 탐지하기 위하여 오디오 파일의 각 프레임으로부터 선형예측계수(linear predictive coding, LPC)를 추출한 뒤, 추출한 LPC 간의 거리(spectral distance)를 활용하는 연구가 제안되었다[4]. 참고문헌 [4]에서는 아래의 수식과 같이 오디오 파일의 t 번째 프레임에서의 spectral distance $D(t)$ 를 정의한다.

$$D(t) = \max D_1(t), D_2(t) \tag{1}$$

$$D_1(t) = \|w(t) - w(t-1)\|^2 \tag{2}$$

$$D_2(t) = \|w_{inv}(t) - w_{inv}(t-1)\|^2 \tag{3}$$

위 식에서 $w(t)$ 는 t 번째 프레임에서 추출한 LPC 벡터를, $w_{inv}(t)$ 는 t 번째 프레임의 샘플들을 역순으로 배열한 뒤, 추출한 LPC 벡터를 의미한다.

해당 연구에서는 spectral distance를 사용하는 오디오 파일의 편집 의심 지점을 검출하는 알고리즘을 제안하였으며, 몇몇 사례 연구를 통해 제안한 알고리즘이 실제 위변조 여부 판별에 효과적으로 사용될 수 있음을 보였다[4].

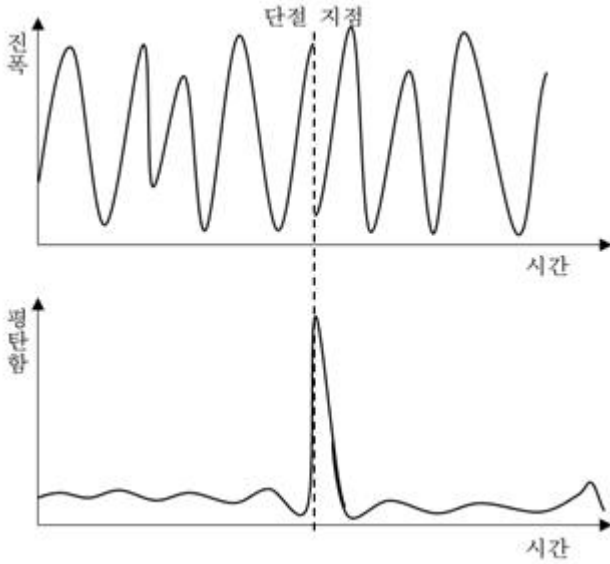


그림 3. 단절 지점 검출 예시
Figure 3. Example of detecting edited point

3. 스펙트럼 분석을 통한 편집 의심 지점 검출

본 논문에서는 오디오 파일 내의 편집 의심 지점을 검출하기 위하여 오디오 스펙트럼을 분석하는 알고리즘을 제안한다. 제안한 알고리즘은 오디오 파일 내에 편집으로 인해 발생하는 단절 지점에서 백색 잡음과 유사한 특성이 발생한다는 점에 착안하였다 (<그림 1>). 또한, 이러한 특성이 음성 구간이 아닌 비 음성 구간에서도 유사하게 발생하는 것을 확인하였다. <그림 2>를 보면 오디오 파일의 비 음성 구간을 임의로 편집하여 발생시킨 단절 지점에서의 스펙트럼(0kHz~4kHz)과 일반적인 비 음성 구간에서의 스펙트럼을 비교할 수 있다. 두 스펙트럼을 비교해 보면 단절 지점에서 발생하는 잡음 성분으로 인하여 스펙트럼에서의 대역별 에너지 변화가 상대적으로 완만한 것을 확인할 수 있다. 따라서 오디오 파일의 각 프레임 별로 스펙트럼의 평탄함을 계산함으로써 편집 의심 지점을 검출하는 알고리즘을 제안한다. 이를 위하여 스펙트럼의 평탄함 f 를 아래의 수식 (5)와 같이 정의한다.

$$\alpha = \sum_{i=1}^{l-1} (freq_i - freq_{i+1})^2 \quad (4)$$

$$f = e^{-\alpha} \quad (5)$$

위 식에서 α 는 평탄함을 계산하기 위해 사용되는 보조 변수, l 은 각 구간에서 추출한 주파수 대역의 개수, $freq_i$ 는 i 번째 대역의 에너지를 의미한다. 따라서 수식 (5)로 계산한 f 값은 스펙트럼이 평탄할수록 큰 값을 가지고 그렇지 않을수록

작은 값을 가지게 된다. 여기서 지수함수를 사용한 이유는 단절된 지점이 강조되어 사람이 눈으로 판별하기 쉽도록 하기 위한 것이다. 수식 (5)로 정의된 스펙트럼의 평탄함을 활용하여 실제로 오디오 파일의 단절 지점을 검출하는 예시는 <그림 3>과 같다. 그림의 위 쪽에서 단절 지점을 포함하는 음성 파형을, 아래 쪽에서 음성의 각 프레임에서 계산한 스펙트럼의 평탄함 값을 확인할 수 있다. 평탄함 값을 보면 일반적인 음성 프레임에서는 유사하게 작은 값들이 계산되지만, 단절 지점을 포함하는 프레임에서는 높은 값이 계산돼 단절 지점을 검출하는 것을 확인할 수 있다. 본 논문에서는 제안한 스펙트럼의 평탄함을 사용한 편집 의심 지점 검출 알고리즘에 다양한 척도를 적용하거나 기존의 위변조 검출 연구와 결합하는 추가 연구를 소개한다.

Spectral flatness measure(SFM)는 스펙트럼의 평탄함을 계산하는데 사용될 수 있는 전통적인 척도이다. SFM 값이 높게 계산된 프레임은 음성의 특성이 백색 잡음과 유사한 프레임이라고 할 수 있으며, 작게 계산된 구간은 주파수 성분의 정점들이 많이 분포해 백색 잡음과 차이가 나는 프레임이라고 할 수 있다[9], [10]. 따라서 이 척도를 편집 의심 지점 검출에 활용하는 것이 가능하다. SFM 값은 아래의 식 (6)과 같이 정의된다.

$$f_{SFM} = \frac{\sqrt[l]{\prod_{i=1}^l freq_i}}{\sum_{i=1}^l freq_i/l} = \frac{\exp(\sum_{i=1}^l \ln(freq_i)/l)}{\sum_{i=1}^l freq_i/l} \quad (6)$$

스펙트럼의 평탄함을 계산하기 위해 정보이론에서 많이 사용되는 엔트로피 개념 또한 사용될 수 있다. 정보이론에서의 엔트로피는 특정 변수에 포함돼 있는 정보량을 측정하는 척도로 사용된다[11]. 따라서 엔트로피를 계산함으로써 확률 값으로 정의된 변수에서 발생할 수 있는 정보량을 계산할 수 있다. 엔트로피 값이 크다면 해당 변수로부터 많은 정보를 얻을 수 있음을, 작다면 적은 정보 밖에 얻을 수 없음을 나타낸다[11]. 이를 스펙트럼에 적용해 생각하면 다음과 같이 해석하는 것이 가능하다. 스펙트럼이 평탄하지 않은 경우, 많은 정보를 얻을 수 있기 때문에 엔트로피 값이 클 것이고, 스펙트럼이 평탄하다면 적은 정보 밖에 얻을 수 없기 때문에 엔트로피 값이 작을 것이다. 따라서 스펙트럼으로부터 엔트로피 값을 계산함으로써 편집 의심 지점을 검출하는 것이 가능할 것이다. 이 때의 엔트로피 값을 통해서 계산한 스펙트럼의 평탄함은 아래의 식 (7)과 같이 정의된다.

$$f_{ent} = -\sum_{i=1}^l p_i \ln(p_i) \quad (7)$$

$$p_i = \frac{freq_i}{\sum_{i=1}^l freq_i} \quad (8)$$

위 식에서 p_i 는 정보이론의 엔트로피를 계산하기 위하여 i 번째 대역의 에너지를 확률 변수로 변환한 것이다.

본 연구에서는 스펙트럼의 평탄함 계산을 위해 다양한 척도들을 적용하는 것 외에도 다른 위변조 검출 관련 연구들을 참고하여 편집 의심 지점 검출 성능을 향상시키고자 하였다. 참고문헌[8]에서는 오디오 샘플에 difference 연산을 반복적으로 적용함으로써 위변조 검출 성능을 향상시키는 것이 가능하다는 연구 결과를 발표하였다. 오디오 샘플의 k 차 difference 연산은 아래의 수식 (9)와 같이 정의 된다.

$$d^k(x_n) = d^{k-1}(x_n) - d^{k-1}(x_{n-1}) \quad (9)$$

위 식에서 x_n 는 n 번째 오디오 샘플을, $d^k(x_n)$ 는 n 번째 오디오 샘플에 k 차 difference 연산을 적용한 결과를 의미한다. 이 때, 0차 difference 연산 결과를 일반적인 오디오 샘플 값으로 정의 한다. 위 식처럼 difference 연산을 반복함으로써 오디오 파일의 편집 의심 지점을 더욱 정교하게 검출하는 것이 가능하기 때문에, 스펙트럼의 평탄함을 활용한 위변조 검출 실험에도 difference 연산을 적용한다.

4. 실험 설계 및 결과

위변조 검출과 관련된 기존의 많은 연구들은 임의로 생성한 소수의 위변조 데이터들만을 사용하여 개발한 알고리즘의 동작을 확인하는 수준의 사례 연구 만을 진행하였다[4], [8]. 본 연구에서는 대량의 데이터를 사용하여 보다 객관적인 성능 평가를 수행하기 위하여 다음과 같이 편집된 오디오 파일 검출 실험을 진행하였다.

4.1 실험 설계

사람의 음성이 이어지는 중간에 다른 음성을 이어 붙이거나 일부 구간을 삭제하는 등의 편집을 가하는 경우에는 파형이나 스펙트럼의 변화가 크므로 상대적으로 위변조 여부 판별이 쉬울 수 있다. 따라서 본 연구에서는 청취 분석만으로는 사람이 판별하기 어려운 수준의 정교한 편집의 경우, 대부분 비 음성 구간에서 이루어질 것이라고 가정하였다. 실험을 위하여 사람의 음성이 발생하지 않는 환경에서 오디오 파일을 약 26분간 녹음하였다. 이렇게 녹음한 오디오 파일은 실제로 사람의 음성이 포함되지 않고 에어컨이나 컴퓨터 팬 소리 등의 생활 잡음을 포함하고 있다. 이렇게 녹음한 오디오 파일의 임의 구간을 편집하더라도 그 단절점이 크게 부각되지 않는다. 따라서

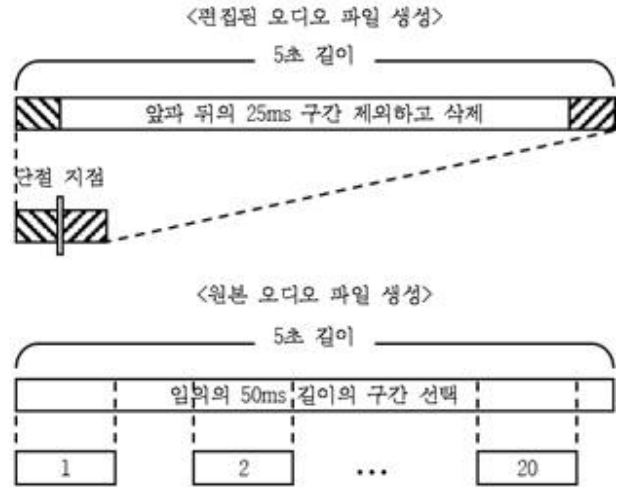


그림 4. 위변조 검출 실험을 위한 위변조 파일 및 원본 파일 생성 과정

Figure 4. Producing process of edited and normal audio for forgery detection experiments

이 비 음성 오디오 파일을 짧은 길이로 분할한 뒤, 임의로 편집하여 위변조 오디오 파일을 생성해 실험을 진행하였다.

먼저 26분 길이의 오디오 파일을 5초 길이의 파일 311개로 분할한다. 분할한 5초 길이의 오디오 파일 앞과 뒤의 25ms 구간을 제외한 중간 부분을 전부 삭제해 50ms 길이의 편집된 오디오 파일을 생성한다. 다시 분할한 5초 길이의 오디오 파일 내에서 50ms 길이의 구간을 임의로 20개 선택하여 원본 오디오 파일을 생성한다. 따라서 약 26분 길이의 오디오 파일에서 311개의 편집된 오디오 파일과 6,220개(311 * 20)의 원본 오디오 파일을 생성한다. <그림 4>는 앞에서 설명한 편집된 오디오 파일과 원본 오디오 파일을 생성하는 과정을 보여준다. 이렇게 생성한 50ms 길이의 오디오 파일들을 편집된 파일과 그렇지 않은 파일로 이진 분류하는 방식으로 실험을 진행하였다. 이 때, 오디오 파일의 길이를 50ms로 설정한 것은 편집 의심 지점 검출시 분석 윈도우의 크기를 25ms로 하는 경우를 가정한 것이다.

분할한 각 음성에서 여러 길이의 window(15ms, 20ms, 25ms, 30ms)를 한 샘플씩 이동시켜 가며 각 window에서의 스펙트럼을 분석하였다. 이 때, 각 window의 오디오 샘플들을 스펙트럼으로 변환하기 위해서 fast Fourier transform(FFT)을 적용하였다. Window의 종류는 Hamming window와 blackman-harris window 두 종류를 사용하였다. FFT를 적용해 계산한 스펙트럼에 앞에서 설명한 척도들을 사용하여 스펙트럼의 평탄함을 계산하였다. 사용한 척도는 제안한 스펙트럼의 평탄함(수식 (5)), SFM(수식 (6)), 엔트로피(수식 (7))이다. 또한, 이전의 위변조 검출 연구를 적용하기 위하여 오디오 샘플에 1차부터 10차까지 difference 연산을 적용시켜가며 편집 의심 지점 검출 성능

을 비교하였다. 마지막으로 기존의 spectral distance를 기반으로 한 편집 의심 지점 검출 알고리즘과 성능 비교를 진행하였다. 모든 성능 평가는 동일 오류율 (equal error rate, EER, %)을 기준으로 진행하였다.

표 1. 주파수 스펙트럼의 평탄함을 사용하였을 때 window 종류와 크기에 따른 위변조 검출 성능

Table 1. Forgery detection performance according to window types

Window type	EER (%)	Window size (ms)				
		10	15	20	25	30
Hamming	20.3	21.7	35.3	45.8	48.2	
Blackman-harris	18.6	18.0	23.2	42.3	49.8	

표 2. 평탄함 척도에 따른 편집 지점 검출 성능

Table 2. Forgery detection performance according to measures for spectral gentleness

EER (%)	계산 척도		
	단순 평탄함	SFM	Entropy
18.0		23.1	32.5

4.2 실험 결과

<표 1>은 분석에 사용한 window의 종류나 크기에 따른 편집 의심 지점 검출 성능을 보여준다. 이 실험에서는 스펙트럼의 평탄함 계산을 통해 편집 의심 지점을 검출하였다. 실험 결과를 보면 Hamming window 보다 blackman-harris window가 편집 지점 검출에 더욱 적합하며, window의 크기가 15ms일 때 가장 좋은 성능을 보이는 것을 알 수 있다.

<표 2>는 스펙트럼의 평탄함 계산 척도에 따른 편집 의심 지점 검출 성능을 보여준다. 이 실험에서는 window 종류를 blackman-harris, 크기를 15ms로 고정한 뒤, 실험을 진행하였다. 실험 결과를 보면 스펙트럼 평탄함의 척도로 전통적으로 사용되었던 SFM이나 정보이론에서의 엔트로피 보다 제안한 평탄함 계산이 더욱 높은 성능을 보이는 것을 알 수 있다. 이는 단절 지점에서의 스펙트럼의 전체적인 평탄함 보다는 인접한 주파수 대역의 성분 진폭의 차이가 단절 지점 검출에 더욱 유효하기 때문이라고 해석할 수 있다.

<표 3>은 스펙트럼의 평탄함을 활용한 편집 의심 지점 검출

알고리즘에 difference 연산을 차수를 다르게 적용해가며 적용한 실험의 결과이다. 이 실험에서는 window 종류를 blackman-harris, 크기를 15ms, 단순한 방식으로 평탄함을 계산하였다. 실험 결과를 보면 difference 연산의 차수에 따라서 적용하지 않은 경우(0차) 보다 편집 지점 검출 성능이 증가하는 경우가 있는 것을 확인할 수 있다. 성능 향상의 폭이 큰 편은 아니지만 간단한 difference 연산을 적용함으로써 성능 향상이 있다는 의미가 있을 수 있다.

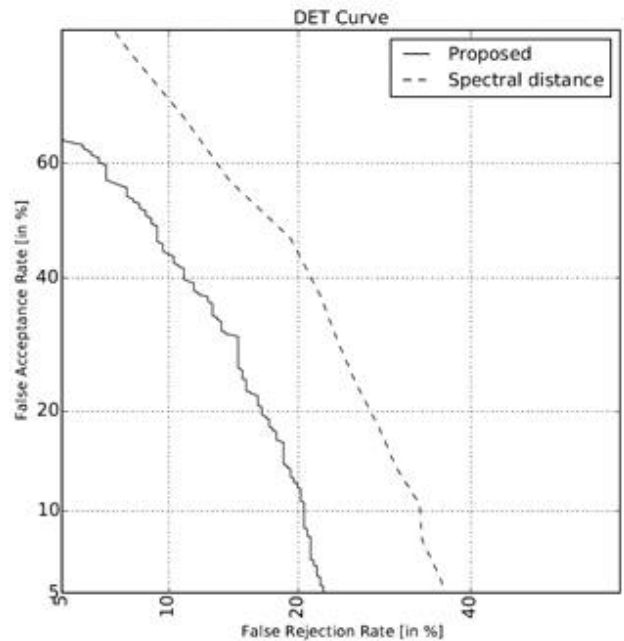


그림 5. 편집 의심 지점 검출 방식에 따른 DET 곡선
Figure 5. DET curves according to forgery detection methods

<그림 5>는 제안한 스펙트럼의 평탄함을 활용한 편집 의심 지점 검출 알고리즘과 spectral distance를 활용한 알고리즘의 성능 비교 결과를 DET 곡선으로 나타낸 것이다. 이 실험에서 spectral distance를 활용한 알고리즘은 참고문헌 [4]에서의 실험 파라미터를 그대로 활용하였다. 결과를 보면 스펙트럼의 평탄함을 기반으로 제안한 방식이 spectral distance를 기반으로 한 방법보다 항상 더 높은 성능을 보이는 것을 알 수 있다. EER을 기준으로 보면 평탄함 기반의 검출은 17.65%, spectral distance 기반의 검출은 25.14%의 EER을 보여서, 평탄함 기반의 편집 의심 지점 검출을 통해서 EER 기준 오류가 약 29% 정도 감소하는 것을 확인할 수 있었다.

표 3. 적용한 difference 연산의 차수에 따른 위변조 검출 성능

Table 3. Forgery detection performance according to orders of difference operation

EER(%)	Difference 차수										
	0	1	2	3	4	5	6	7	8	9	10
18.0	18.9	17.7	17.6	17.9	18.3	18.6	17.9	18.3	19.6	18.6	

5. 결론

본 논문에서는 스펙트럼의 평탄함을 사용한 오디오 파일 내의 편집 의심 지점 검출 알고리즘을 제안하였다. 제안한 방법과 함께 기존의 다양한 평탄함 척도를 적용하고 다른 위변조 검출 연구를 적용하여 성능을 비교한 결과, 제안한 방법이 기존의 여러 방법과 비교하여 우수함을 알 수 있었다. 최종적으로 기존의 spectral distance 기반의 편집 의심 지점 검출 알고리즘과 성능 비교를 진행해 EER을 기준으로 더욱 높은 성능을 보이는 것을 확인하였다.

향후에는 수식적으로 더욱 정교한 평탄함 척도를 적용하는 연구, 대량의 오디오 파일을 사용해 추가적인 성능 비교를 하는 연구 등을 진행할 계획이다. 또한, 기계학습 분야에서 널리 사용되고 있는 deep neural network 구조를 편집 의심 지점 검출에 적용하는 연구도 진행할 계획이다.

참고문헌

- [1] Brixen, E. B. (2007). Techniques for the authentication of digital audio recording, *Audio Engineering Society Convention* 122.
- [2] Ojowu, O. (2012). ENF extraction from digital recordings using adaptive techniques and frequency tracking, *Information Forensics and Security*. Vol. 7, 1330-1338.
- [3] Grigoras, C. (2009). Applications of ENF analysis in forensic authentication of digital audio and video recording, *Journal of the Audio Engineering Society*. Vol. 57, Issue 9, 643-661.
- [4] Nicolade, D. P. & Apolinario, J. A. (2009). Evaluating digital audio authenticity with spectral distances and ENF phase change, in *proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1417 - 1420.
- [5] Hicsonmez, S., Husrev T. S., and Ismail V. (2011). Audio codec identification through payload sampling, *Information Forensics and Security (WIFS)*, 2011 IEEE International Workshop on. IEEE.
- [6] Baek, R. S., Kim, K. W., So, B. M., Yang, I. H., Kim, M. J., Heo, H. S. & Yu, H. J. (2012). The Transmission Channel Identification of Digital Speech Data Using GMM, *Proceedings of the 2012 Fall Conference of the Korean Society of Speech Sciences*.
- (백록선, 김경화, 소병민, 양일호, 김명재, 허희수, 유하진. (2012). GMM을 사용한 디지털 음성 데이터의 전송 채널 식별, 한국 음성학회 가을 학술대회 발표논문집.)
- [7] Kim, I. W., Yang, I. H., Kim, M. J., Heo, H. S., Yoon, S. H. & Yu, H. J. (2015). Audio Codec Identification Based on DNN, *Proceedings of the 2015 Fall Conference of the Acoustic Society of Korean*.
- (김인화, 양일호, 김명재, 허희수, 윤성현, 유하진, (2015). 심층 신경망 기반의 오디오 코덱 식별, 한국음향학회 추계학술대회 발표논문집)
- [8] Cooper, A. J. (2010). Detecting butt-spliced edits in forensic digital audio recordings, *Audio Engineering Society Conference: 39th International Conference: Audio Forensics: Practices and Challenges*. Audio Engineering Society.
- [9] Jayant, N. S. & Noll, P. (1984). *Digital Coding of Waveforms*, Englewood Cliffs, NJ: Prentice-Hall.
- [10] Johnston, J. D., (1988). Transform coding of audio signals using perceptual noise criteria, *IEEE J. Select. Areas Commun.* vol. 6, 314 - 323.
- [11] Balian, R. (2004). Entropy, a Protean concept, In Dalibard, Jean. Poincaré Seminar 2003: Bose-Einstein condensation - entropy. Basel: Birkhäuser, 119 - 144.

• 허희수 (Heo, Hee-Soo)

서울시립대학교 컴퓨터과학부
서울시 동대문구 전농동 90번지
Tel: 02-6490-5697 Fax: 02-6490-2444
Email: zhasgone@naver.com
관심분야: 화자 인식, 음성 인식, 범음성학
현재 컴퓨터과학과 대학원 석사과정 재학 중

• 양일호 (Yang, Il-Ho)

서울시립대학교 컴퓨터과학부
서울시 동대문구 전농동 90번지
Tel: 02-6490-5697 Fax: 02-6490-2444
Email: heisco@hanmail.net
관심분야: 화자 인식, 음성 인식, 범음성학
현재 컴퓨터과학과 대학원 박사과정 재학 중

• 소병민 (So, Byung-Min)

대검찰청
서울시 서초구 반포대로 157
Tel: 02-3480-2593 Fax: 02-3480-2707
Email: sbm1210@spo.go.kr
관심분야: 화자 인식, 음성 인식, 범음성학
현재 대검찰청 NDFC 멀티미디어분석실 수사관

• 유하진 (Yu, Ha-Jin), 교신저자

서울시립대학교 컴퓨터과학부
서울시 동대문구 전농동 90번지
Tel: 02-6490-2448 Fax: 02-6490-2444
Email: hjyu@uos.ac.kr
관심분야: 화자 인식, 음성 인식
현재 컴퓨터과학부 교수