

고전 영화의 디지털 음원 복원을 위한 강인한 노이즈 검출 기법

*유수정 **조남익

서울대학교 전기컴퓨터공학부

*sjyou21@gmail.com **nicho@snu.ac.kr

Robust Noise Detection for Digital Audio Restoration in Old Films

*You, Su-Jeong **Cho, Nam-Ik

Dept. of Electrical Engineering and Computer Science, Seoul National University

요약

본 논문에서는 단일 채널 디지털 오디오 신호에서 스펙트로그램과 영상 처리 기법을 이용하여 크래클 잡음을 검출하는 알고리즘을 제안한다. 오디오 신호의 주파수 특성을 효율적으로 분석하기 위해 스펙트로그램을 특정 컬러맵을 이용하여 컬러 영상으로 변환한 후 영상 처리 기법을 적용하여 크래클 잡음이 존재하는 구간을 검출하여 디지털 오디오 복원에 이용한다. 특히 고전영화에 나타나는 크래클 잡음은 에너지와 신호 길이가 음성이나 음악 신호와 유사하여 기존의 스펙트럴 음성 검출 기법으로는 검출에 어려움이 있다. 이에 비해 스펙트로그램 영상에서는 크래클 잡음이 다른 신호들과 구분되는 특성을 나타내므로 영상 처리 기법을 적용하여 경계 검출과 Hough 변환에 의한 선 검출을 이용하여 크래클 잡음을 검출한다. 제안된 알고리즘은 고전 영화 디지털 오디오 복원에 적용하였으며 크래클 잡음 검출에 우수한 성능을 보여준다.

1. 서론

일반적으로 오디오 신호에서 음성과 잡음을 분리하기 위해서 음성 검출기를 이용하여 왔다. 기본적인 음성 검출은 주어진 시간 구간에서 신호의 에너지, 즉 신호대잡음비(이하 SNR, signal to noise ratio)가 일정한 임계값보다 크면 이 구간을 음성 구간, 작으면 잡음 구간이라고 결정한다. 그러나 오디오 신호에 존재하는 잡음들은 다양한 특성을 가지고 있기 때문에 이러한 방법으로 잡음과 음성을 분리하는 데에는 한계가 있다. 특히 고전 영화의 오디오 신호에 나타나는 크래클 잡음은 크기나 신호 길이 등이 음성이나 음악 구간과 비슷하여 기존의 음성 검출기에 의해서는 잡음으로 검출되지 못한다. 최근까지 음성 검출기의 성능을 향상시키기 위해 SNR을 추정하는 방법[1], soft decision method[2]와 같은 많은 연구가 있었지만 고전 영화의 크래클 잡음 특성상 그대로 적용하기에 어려움이 있다.

음성 스펙트로그램은 1차원 신호를 주파수와 시간, 즉 2차원 신호로 분석할 수 있기 때문에, 음성 신호 분석에 있어서 매우 강력한 도구이다. 한편 영상처리는 2차원 데이터를 처리하는 유용한 방법이며 이를 위한 방대한 연구가 이루어져왔다. 따라서 영상 처리 알고리즘을 스펙트로그램 영상에 적용한다면 세밀한 분석이 가능할 것이다. 이러한 연구는 음성과 음악 신호를 분리하는 과정에서 시도되기도 하였다[3].

본 논문에서는 스펙트로그램과 영상 처리 알고리즘을 사용하여 고전 영화의 단일 채널 디지털 오디오 음원에 존재하는 크래클 잡음을 검출한다. 우선, 입력 신호에서 스펙트로그램을 이용하여 주파수-시간적 특성을 추출하고 이를 칼라 영상과 그레이 영상으로 변환하여 2차원 특징들로 변환시킨다. 그리고 경계 검출과 Hough 변환과 같은 영상 처리 알고리즘으로 크래클 잡음 검출을 위한 선 검출을 수행한다. 일련의 과정에 의해 검출된 선의 위치는 크래클 잡음이 존재할 수 있는 후보 위치로 선정되고 이를 기준으로 스펙트럼 크기를 비교하여 최종적으로 크래클 잡음의 위치를 검출한다. 실험 결과 기존의 음성 검출기에서는 음성 구간으로 구분되는 크래클 잡음이 음성과 분리되어 검

출되어 고전 영화의 디지털 음원 복원 성능이 향상되었다. 제 2 장에서는 스펙트로그램에 나타나는 크래클 잡음의 특성을 분석하고 이를 검출하기 위해 적용되는 Hough 변환에 대해 설명한다. 제 3 장에서는 제안된 알고리즘을 설명하고 제 4 장에서는 실험결과를 제시한 후 제 5 장에서는 결론으로 마무리한다.

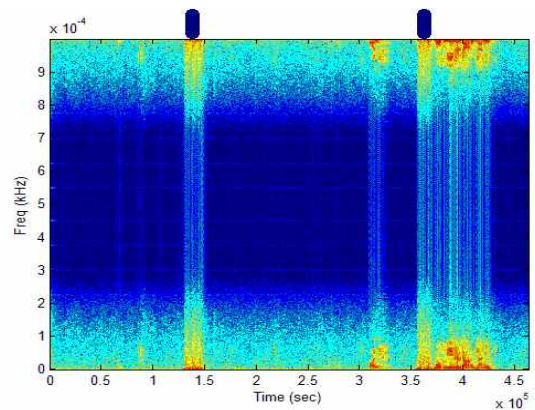


그림 1 Noise+음성 신호의 spectrogram
(위에 표시된 부분이 크래클 노이즈)

2. 기본 개념

가. 잡음 특성

그림 1에서는 크래클 잡음과 음성이 섞여 있는 신호의 스펙트로그램을 보여준다. 그림 상단의 표시되어 있는 부분이 크래클 잡음의 위치를 나타낸 것이다. 크래클 잡음으로 표시된 부분은 수직 기둥 형태를 이루며 영상의 컬러를 분석하면 다른 구간에 비해 서로 비슷한 값을 가진 특징들로 구성되어 있음을 알 수 있다. 한편 오른쪽에 표시된 크래클 잡음은 음성구간에 섞여 있기 때문에 검출에 어려움이 있다. 그림

1에서는 음성의 시작 구간에 존재하기 때문에 비교적 용이하지만 음성이나 음악 구간의 중간에 위치할 경우에는 검출이 매우 어렵다.

나. Hough Transform

스펙트로그램에서 정확한 크래클 잡음의 위치를 검출하기 위해 잡음에 강인하고 안정한 Hough 변환을 적용한다. Hough 변환에서 직선 위의 화소, (x,y) 는 다음과 같이 정의된 시그모이드 함수상의 영역으로 매핑된다.

$$H: (x,y) \rightarrow \rho = x \cos \theta + y \sin \theta \quad (1)$$

여기서 θ 와 ρ 는 각각 선의 기울기와 원점으로부터의 거리를 나타낸다. 경계 검출시 각 경계상의 화소는 (ρ, θ) 공간 영역에서 정현곡선으로 나타난다. 이러한 공간에서 같은 직선상의 모든 화소들은 같은 점에서 만나는 곡선들의 군집을 이룬다. 즉 이 곡선들의 군집을 이용하여 영상에서 존재하는 선을 검출하게 된다.

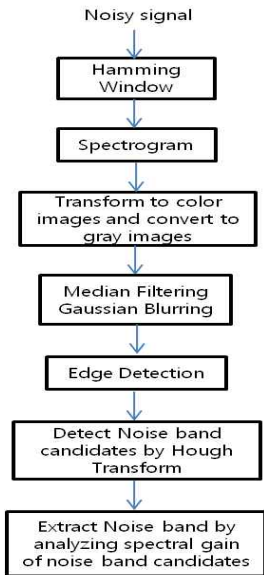


그림 2 제안 알고리즘의 블록도

3. 제안 알고리즘

그림 2는 본 논문에서 제안하는 알고리즘의 블록도를 보여준다. 디지털 오디오 음원을 25ms 크기의 Hamming 윈도우를 사용하여 세그먼트 단위로 나누어서 DFT를 계산한다. 이렇게 스펙트로그램을 구한 후에 주파수-시간영역으로 표시된 값을 특정 컬러맵을 이용하여 컬러 영상으로 변환한다. 이 컬러 영상은 그레이 영상으로 변환된 후 median filtering과 Gaussian blurring 과정을 통해서 잡음을 제거하고 경계 검출을 용이하게 한다. 전처리된 영상에 Canny edge 검출 알고리즘을 적용하여 경계를 검출한 후 Hough 변환에 의한 선 검출을 수행한다. 크래클 잡음은 수직 기둥을 이루고 있기 때문에 Hough 변환시 이를 고려한 선들만 추출한다. 추출된 선들의 위치는 크래클 잡음의 후보 영역이 되고 각 위치에서 일정 영역의 스펙트럼 크기 평균을 구하고 이를 사용자가 지정한 크래클 잡음 영역과 비교하여 크래클 잡음

의 위치를 검출한다.

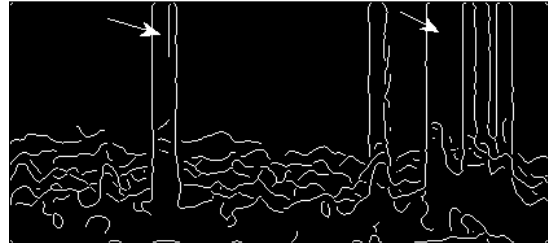


그림 3 Hough transform이용한 크래클 잡음 검출

4. 실험 결과

실험 데이터로 고전 영화 '두만강'이 잘 있거라'의 디지털 음원을 사용하였으며 이는 한국영상자료원에서 제공한 것이다. 그림 3은 제안된 방법으로 검출된 크래클 잡음의 위치를 나타낸 것이다. 이 결과는 그림 1에서 크래클 잡음으로 표시된 부분과 일치하는 것이다.

그림 1에서 오른쪽 음성구간에도 크래클 잡음 이외에 수직 기둥들이 보이는데 Hough 변환으로 선 검출시 이러한 선들도 검출되어 크래클 잡음의 후보군이 된다. 검출된 각 선의 위치에서 사용자가 지정한 참조 크래클 잡음과 스펙트럼 크기 평균값을 비교하면 주어진 음원에서는 쉽게 크래클 잡음을 검출할 수 있다.

5. 결론

본 논문에서 고전 영화 디지털 오디오 신호에서 스펙트로그램과 영상처리 알고리즘을 이용하여 음원에 존재하는 크래클 잡음의 위치를 검출하는 방법을 제안하였다. 고전 영화에 존재하는 크래클 잡음은 크거나 신호의 길이가 음성이나 음악 구간과 비슷하여 기존의 음성구간 검출기로는 검출이 용이하지 않다. 이에 비해 제안된 알고리즘은 스펙트로그램에 나타난 크래클 잡음의 특성을 경계 검출과 Hough 변환 등의 영상처리 알고리즘을 적용하여 그 위치를 정확하게 찾을 수 있다. 제안된 방법은 크래클 잡음 이외에도 비슷한 특성을 갖는 광대역 잡음의 검출에도 적용할 수 있다.

감사의 글

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2010년도 콘텐츠 산업기술지원사업의 연구결과로 수행되었습니다.

참고문헌

- [1] Boll, "Suppression of Acoustic Noise in Speech using Spectral Subtraction," *IEEE Trans. on ASSP*, vol. 29, April 1979.
- [2] McAulay, Malpass, "Speech Enhancement Using a Soft-Decision Noise Suppression Filter," *IEEE Trans. on ASSP*, vol. 28, no. 2, April 1980.
- [3] Jin, Wang, "Speech Separation from Background of Music Based on Single-channel Recording," in *the Proc. of ICPR IEEE 2006*.