

디지털비디오캠코더 소음 저감 알고리즘 구현

박재하, 오윤학, 이혁재

삼성전자 디지털미디어연구소 Audio Lab.

Implementation of Noise Reduction for Digital Video Camcorder.

Jaeha Park, Yoonhak Oh, Hyuckjae Lee

Jaeha, yoonharkoh, hyuckjae.lee@samsung.com

요약

본 논문에서는 TeakLite DSP 프로세서를 이용하여 캠코더에서 레코딩을 할 때 모터 소음과 주변 잡음이 입력되어 오디오 신호의 명료도가 떨어지는 문제점을 해결하기 위한 잡음 제거 기법의 실시간 구현에 대해서 기술하고자 한다. 잡음 제거를 위해서는 일반적으로 많이 사용되고 있는 Spectral Subtraction 기법을 사용하였다. 알고리즘 구현시 MIPS 강소에 효과적이었던 최적화 기법들을 적용하여 TeakLite DSP 프로세서에서 최적화되어 동작하도록 하였다. 최적화된 Spectral Subtraction 어셈블리 코드는 TeakLite DSP 프로세서에서 32 KHz, 16 bit 입력에 대해 40 MIPS에서 동작하였다.

1. 서론

캠코더를 사용하여 카메라 촬영 시 줌 모터(zoom motor)나 드럼 모터(drum motor)등의 회전에 의해서 소음이 발생한다. 이러한 소음 또는 잡음은 마이크 등의 입력단자를 통하여 오디오 신호와 함께 테이프에 녹음되므로 캠코더 오디오를 재생했을 때 기기의 움직임을 저하시키는 요인이 된다. 따라서 이러한 주변 환경에 의한 소음을 제거하기 위한 잡음 제거 기술이 필요하게 된다. 일반적으로 잡음 제거 장치는 배경 잡음을 제거하기 위해 스펙트럼 차감법(spectral subtraction)[1][2][3]을 이용하고 있다. 본 논문에서는 이 Spectral Subtraction 기법을 이용한 잡음제거 기술을

캠코더 제품에 적용한 내용과 이 기술을 TeakLite DSP 를 이용하여 실시간 동작을 구현한 내용에 대해 설명하고자 한다.

2. 본론

2.1 Spectral Subtraction

일반적인 스펙트럼 차감법은 그림 1 에서 도시한 흐름도와 같이 다음과 같이 동작한다.

- 입력되는 아날로그 신호를 디지털 신호로 변환한다.
- 디지털 신호는 시간영역에서 프레임 단위화한다.
- 프레임 단위의 신호는 프레임간의 정보 단절과 왜곡을 줄이기 위해 윈도우를 취한다.
- 윈도우를 적용한 신호는 푸리에 변환(Fast Fourier Transform: FFT)을 통해 주파수 스펙트럼으로 변환한다.
- 스펙트럼 정보는 크기정보와 위상정보로 이루어진다. 이때 스펙트럼의 크기 정보는 스펙트럼 차감에 이용되고, 위상 스펙트럼 정보는 Inverse Fast Fourier Transform 에 이용된다.
- 스펙트럼 차감은 음성과 잡음이 섞인 크기 스펙트럼에서 추정된 잡음스펙트럼을 빼는 연산이다.
- 이때 일반적으로 잡음 스펙트럼은 잡음 구간의 크기 스펙트럼을 평균하여 음성 구간의 잡음 스펙트럼으로 추정한다. 잡음 특성이 정상적일 경우 추정된 잡음 스펙트럼은 실제 잡음 성분의 스펙트럼과 유사하다. 따라서 스펙트럼 차감에 의해 얻어지는 크기 스펙트럼은

근사적으로 잡음이 제거된 음성 신호만의 크기 스펙트럼이 된다.

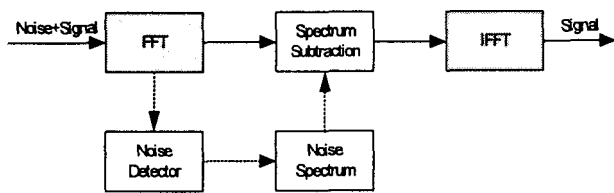


그림 1. Spectral Subtraction 구조

2.2 개선된 스펙트럼 차감 알고리즘 개선

TeakLite DSP 에 구현한 노이즈 저감 알고리즘은 앞서 소개한 일반적인 스펙트럼 차감법을 개선한 방법을 적용하였다. 하지만, 기존의 스펙트럼 차감법은 일정한 주파수 형태를 갖는 잡음 신호가 오디오 신호에 인가되는 경우 잡음 신호만의 주파수 형태를 미리 파악하였다가 잡음과 오디오 신호가 혼합된 신호가 입력되면 해당하는 잡음 주파수대역의 성분만 빼내는 것이다. 그러므로 만약 잡음이 아닌 오디오 신호의 주파수 성분을 잡음 신호라고 판단하면 잡음 이외의 오디오 신호도 제거되며, 잡음 신호의 주파수 성분인데도 불구하고 잡음 신호가 아니라고 판단하면 잡음 신호를 제대로 제거하지 못하는 문제점을 가진다. 따라서 노이즈 스펙트럼을 갱신하는 방법을 변경할 필요가 있다. 그림 2 에서와 같이 Noise Detector 부는 FFT 처리된 현재 프레임의 신호가 잡음만 있는 프레임인지 또는 잡음이 섞인 오디오 신호가 있는 프레임인지를 세 가지의 임계값을 이용하여 판별하며, 노이즈 프레임으로 판별된 현재 프레임 신호 $\{X_n[w]+N_n[n]\}$ 를 Noise Spectrum Update 부로 전달한다. 즉, Noise Detector 부는 현재 프레임의 에너지와 제 1 임계값, 현재 프레임 스펙트럼과 잡음 스펙트럼 간의 에너지 차와 제 2 임계값, 현재 프레임의 스펙트럼과 잡음 스펙트럼 간의 차와 제 3 임계값을 동시에 만족하는 입력 프레임을 노이즈 프레임으로 판단한다. 이때 제 1, 제 2, 제 3 임계값은 잡음 스펙트럼 갱신 구간별로 다르게 설정된다. Noise Spectrum Update 부는 Noise Detector 부에서 현재 프레임이 잡음만 있는 프레임으로

판별되면 그 프레임의 스펙트럼 $\{X_n[w]+N_n[n]\}$ 과 이전의 잡음 스펙트럼을 이용하여 잡음 스펙트럼을 갱신한다. Noise Spectrum Update 부에서 갱신되는 잡음 스펙트럼을 수식으로 표현하면,

$$(N_n[w])=N_{n-1}[w]*(1-\alpha)+\{X_n[w]+N_n[n]\}*\alpha \quad (1)$$

으로 표현할 수 있으며, 여기서 $N_{n-1}[w]$ 는 이전 프레임의 잡음 스펙트럼이며, $X_n[w]$ 는 현재 프레임의 스펙트럼으로 이상적인 경우 0 의 값을 가지며, $N_n[n]$ 는 현재 프레임의 노이즈 스펙트럼이며, α 는 노이즈 스펙트럼 갱신 계수이며 실험적으로 0.2 의 값에서 효과적이었다.

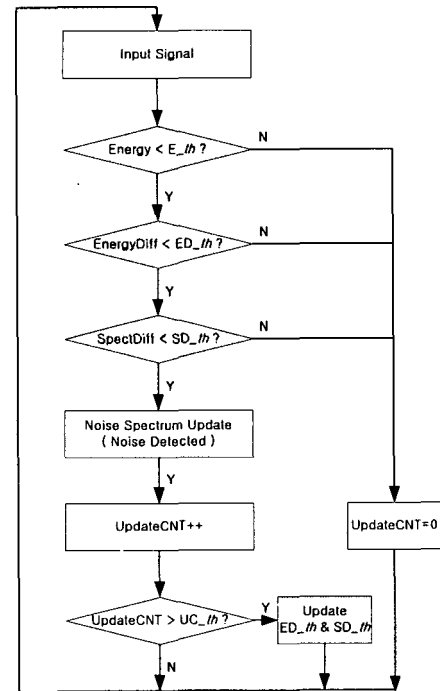


그림 2. 잡음 스펙트럼 갱신 기법의 흐름도

그림 2 는 정교한 방식의 잡음 스펙트럼 갱신 기법 흐름도이다. 입력되는 신호는 매 프레임 마다 주파수 스펙트럼으로 변환된다. 이어서, 입력 신호의 에너지 레벨을 모니터링한다. 여기서 현재 프레임의 에너지 레벨과 제 1 임계값(E_{th})을 비교한다. 이때 잡음만

존재하는 프레임은 잡음과 혼합된 신호가 존재하는 프레임보다 에너지 레벨이 적으므로 현재 프레임의 에너지 레벨이 제 1 임계값과 비교하여 작아야 한다. 그리고 현재 프레임의 에너지가 제 1 임계값보다 작으면 다시 입력 신호의 매 프레임별 에너지 레벨의 변화량을 모니터링한다. 여기서 현재 프레임의 에너지 레벨과 이전 프레임에서 갱신된 노이즈 스펙트럼의 에너지 레벨간의 차와 제 2 임계값(ED_th)을 비교한다. 이때 현재 프레임이 잡음 프레임이라면 그 잡음 프레임은 잡음과 혼합된 신호 프레임보다 에너지의 변화량이 적으므로 그 에너지 레벨 차가 제 2 임계값과 비교하여 적어야 한다. 반면에 현재 프레임의 에너지가 제 1 임계값보다 크면 업데이트 카운트(UpdateCNT)를 초기화시킨다. 이어서, 현재 프레임의 에너지 레벨과 이전 프레임에서 갱신된 노이즈 스펙트럼의 에너지 레벨간의 차가 제 2 임계값보다 적으면 다시 입력 신호의 스펙트럼의 변화량을 모니터링한다. 여기서 현재 프레임의 스펙트럼과 이전 프레임에서 갱신된 노이즈 스펙트럼간의 차와 제 3 임계값(SD_th)을 비교한다. 이때 현재 프레임 구간이 잡음 구간이라면 스펙트럼의 변화량이 적으므로 그 스펙트럼 차가 제 3 임계값과 비교하여 작아야 한다. 반면에 현재 프레임의 에너지와 이전 프레임에서 갱신된 노이즈 스펙트럼의 에너지 레벨간의 차가 제 2 임계값보다 크면 업데이트 카운트(UpdateCNT)를 초기화시킨다. 이어서, 현재 프레임의 스펙트럼과 이전 프레임에서 갱신된 노이즈 스펙트럼간의 차가 제 3 임계값보다 작으면 노이즈 스펙트럼 업데이트를 수행한다. 반면에 현재 프레임의 스펙트럼과 이전 프레임에서 갱신된 노이즈 스펙트럼간의 차가 제 3 임계값보다 크면 업데이트 카운트(UpdateCNT)를 초기화시킨다. 따라서 에너지(energy), 에너지 차(Energydiff), 스펙트럼 차(Spectdiff)들이 각각 제 1, 제 2, 제 3 임계값을 모두 만족하면 업데이트 카운트(UpdateCNT)를 1 증가 시킨다. 이어서, 업데이트 카운트가 임계값(UC_th)보다 크면

제 2 임계값(ED_th) 및 제 3 임계값(SD_th)을 업데이트시킨다.

2.3 청각적 특성을 이용한 유지컬 노이즈의 감쇠

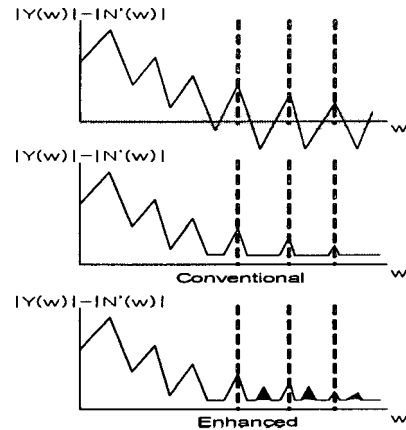


그림 3. 청각적인 특성을 이용한 유지컬 감쇠 기법

유지컬 노이즈를 감쇠시키기 위해 그림 3 과 같은 기법을 사용하였다. 절차는 다음과 같다.

- 입력되는 신호는 잡음과 오디오가 더해진 신호이다. 이 신호의 스펙트럼에서 잡음 스펙트럼을 차감시킨다.
- 발생하는 유지컬 노이즈를 줄이기 위해 사람의 청각적인 특성을 이용하여 comfort 노이즈를 인접 대역에 삽입하여 유지컬 노이즈를 masking 시킴으로써 사람이 유지컬 노이즈를 인지하지 못하도록 한다.

이 과정을 수식으로 나타내면 다음과 같다.

$$\text{IF } (Y(w) - N'(w)) \geq \text{threshold} \\ Y(w) - N'(w) = (Y(w) - N'(w)) \quad (2)$$

$$\text{IF } 0 < |Y(w) - N'(w)| < \text{threshold} \\ Y(w) - N'(w) = \text{threshold} \quad (3)$$

$$\text{IF } (Y(w) - N'(w)) \leq -\text{threshold} \\ Y(w) - N'(w) = -0.5 * (Y(w) - N'(w)) \quad (4)$$

그림 3 은 magnitude 가 일정한 임계값보다 작은 경우 사람의 청각적인 특성을 이용하여 comfort 노이즈를 인접 대역에 삽입하여 유지컬 노이즈를 masking 시킨 스펙트럼을 나타내고 있다.

2.4 TeakLite DSP 개요

TeakLite DSP는 범용 16비트 DSP이며, 개략적인 특징을 살펴보면 다음과 같다.

- TeakLite 16-bit core
- 8 K words of data RAM on the X memory bus
- 8 K words of data RAM on the Y memory bus
- Downloadable 10 K words program RAM
- On chip Interrupt Control Unit

노이즈 저감 알고리즘을 구현한 프로그램은 TeakLite 를 제어하는 마이크로(micro-controller)이 ROM 에 있는 프로그램을 프로그램 버스를 통해서 Program-RAM 에 loading 을 한다. 이후 내부에서 일어나는 연산은 X-RAM, Y-RAM 을 이용하여 노이즈 저감 알고리즘이 실행되도록 설계하였다.

2.5. TeakLite DSP 에서의 연산량 절감 방안

- 효과적인 FFT 알고리즘 적용하는 것이 중요하다. 노이즈 저감 알고리즘에서 가장 많은 비중을 차지하는 것은 FFT 함수 이다. 그래서 FFT 함수 구현을 어떻게 하느냐에 따라 DSP 효율적으로 사용할 수 있는지 없는지가 결정된다. 본 알고리즘에서는 비교적 효율적으로 구성된 Real-Point FFT 를 사용하게 되었으며, 이 방법을 적용했으며, 이 방법은 일반적인 FFT 방법보다 30%정도의 연산량을 감소시킬 수 있었다.
- FFT 함수에서 내부 연산이 아닌 index 연산이 FFT 함수 연산의 1/3 정도를 차지하게 되어, 연산량의 부담을 가져온다. index 연산을 테이블로 가져가면 FFT 함수 연산에서 30%이상 연산량 감소효과를 볼 수 있다.
- 포인터 연산 최적화로 부가적인 레지스터 및 명령어 수를 줄여야 하며, 이것이 FFT 와 같이 많은 loop 를 사용하는 곳에서는 상당한 MIPS 개선 효과를 갖는다.

2.6. 구현결과

스테레오 마이크가 장착된 디지털 비디오 캠코더에 DSP 를 장착하여, 노이즈 저감 알고리즘을 구현하였다. 노이즈가 포함된 신호와 기존의 노이즈 저감 알고리즘을

개선한 알고리즘을 구현한 결과를 그림 4 에 나타내었다. 검게 표현된 선택구간에 대하여, 녹음된 신호(a)의 잡음은 -43dB, 노이즈 저감 알고리즘을 구현한 결과는 -48dB 의 평균 RMS 값을 나타내어 약 5~6dB 의 잡음감쇄효과를 확인하였다.

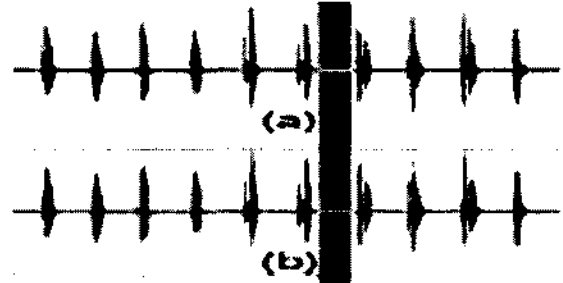


그림 4. 노이즈 저감 알고리즘 적용 전(a)과 후(b)

3. 결론

TeakLite DSP 프로세서를 이용하여 캠코더에서 레코딩을 할 때 모터 소음과 주변 잡음이 입력되어 오디오 신호의 명료도가 떨어지는 문제점을 해결하기 위한 잡음 제거 기법의 실시간 구현을 하였고, 잡음 제거를 위해서 Spectral Subtraction 기법을 사용하였다. 알고리즘은 어셈블리 코딩시 MIPS 감소에 효과적인 기법을 적용하였고, TeakLite DSP 프로세서에서 최적화되어 동작하도록 하였다. 최적화된 Spectral Subtraction 어셈블리 코드는 TeakLite DSP 프로세서에서 32 kHz, 16 bit 입력에 대해 40 MIPS 에서 동작하였으며 5~6dB 의 노이즈 저감 효과를 보였다.

참고문헌

1. Saeed V. Vaseghi, "Advanced Signal Processing and Digital Noise Reduction", Wiley Teubner, 1996
2. Brouti M., Schwartz R., Makhoul J., "Enhancement of Speech Corrupted by Acoustic Noise", ICASSP-79, p. 208-211, 1979
3. Juang B.H., Rabiner L.R., "Signal Restoration by Spectral Mapping", Proc. IEEE, Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP-87 Texas, 1987