

전통적 신호처리 기반 기침 검출·카운팅: RMS-EMA와 이중 임계치

정찬희^{1*}, 홍혁기¹, 조영창

한국전자기술연구원 메디컬IT융합연구센터¹

Cough Detection and Counting via Classical Signal Processing: RMS-EMA with Dual-Threshold Hysteresis

Chan Hee Jeong^{1*}, Hyuck Ki Hong¹, Young Chang Jo¹

Medical IT Convergence Research Center, Korea Electronics Technology Institute, Korea

*chjeong@keti.re.kr

Abstract

We present a classical signal-processing approach for cough detection and counting that delivers low latency, low computational cost, and high interpretability without large training datasets or specialized hardware. The method tracks the background noise level online via RMS exponential moving average (RMS-EMA), defines adaptive dual thresholds (high/low) relative to the tracked RMS, and confirms events through a two-state hysteresis machine with tolerance, padding, and minimum duration constraints. This design yields scale invariance to input gain changes, robustness against gradual noise shifts, and stable boundary decisions under fluctuating energies. The approach is practical for field deployment, supports reproducible parameterization, and can serve as a reliable baseline or annotation aid prior to integrating more complex learned models.

1. 서론

본 논문은 전통적 신호처리 기법에 근거한 기침 검출·카운팅 알고리즘을 제시한다. 제안 방법은 입력 신호의 에너지 수준을 지수이동평균(RMS-EMA)으로 추적해 배경 잡음에 실시간 적응하고, 추정 RMS를 기준으로 한 상·하한 이중 임계치와 히스테리시스 상태 전이를 통해 이벤트를 안정적으로 확정한다. 종료 판정에는 일정 샘플 수 동안의 하한 유지(*tolerance*), 시작·종료 주변 보정을 위한 *padding*, 최소 지속시간 제약을 함께 사용해 스파이크성 잡음이나 단발 충격음을 배제한다. 이 접근은 대규모 학습 데이터나 고성능 하드웨어 없이 낮은 지연·낮은 계산량·높은 해석 가능성을 달성하며, 입력 스케일 변화에 대한 스케일 불변성과 점진적 배경 변동에 대한 강건성을 보인다. 임상·현장 배치에서의 재현성과 품질관리 측면에서도 실무적 이점을 제공한다.

2. 기침 검출 알고리즘

제안한 검출기는 입력 신호를 프레임 단위로 순차 처리하면서 배경 잡음의 장기 레벨을 RMS 지수이동평균(RMS-EMA)으로 추정하고, 그 추정값을 기준으로 이중 임계치를 설정해 이벤트의 시작과 종료를 안정적으로 결정한다. 먼저 프레임 RMS를 다음과 같이 정의한다.

$$rms(x) = \sqrt{\frac{1}{N} \sum_i x_i^2}$$

샘플링 주파수 f_s , 프레임 길이 N , 시간상수 τ 에 따른 갱신 계수와 RMS-EMA 갱신식은 다음과 같다.

$$\alpha = 1 - \exp\left(-\frac{N}{f_s \tau}\right)$$
$$r\widehat{ms} \leftarrow (1 - \alpha)r\widehat{ms} + \alpha rms(x)$$

상한·하한 임계치는 배수 m_h, m_l 를 곱해 산정한다.

$$th_h = m_h r\widehat{ms}, th_l = m_l r\widehat{ms}$$

실제 비교 연산은 계산량을 줄이기 위해 샘플 에너지(제곱값)와 임계치의 제곱 사이에서 수행한다. 이 방식은 입력 이득이 변해도 임계치가 같은 비율로 따라가므로 스케일 불변성을 확보한다.

이벤트 경계는 두 상태(*Idle*, *InCough*)를 가지는 히스테리시스 상태기계로 결정한다. 대기 상태에서 에너지가 상한 임계치를 넘는 순간을 이벤트 시작으로 간주하고, 시작 인덱스는 *padding* 만큼 과거로 당겨 선행부를 포함한다. 이벤트 내부에서는 에너지가 하한 임계치 아래로 떨어지는지를 감시하되, 일시적 하강에 의한 흔들림을 피하기 위해 *tolerance* 동안 연속적으로 낮은 상태가 유지될 때만 종료 후보로 인정한다. 종료 판정이 성립하면 종료 인덱스는 *padding* 만큼 미래로 늦춰 잔향을 포함한다. 종료 후 유효 지속시간은 다음과 같이 계산하며, 사전 정의된 최소 지속시간 이상이고 스트리밍 초기의 워밍업 시간이 지난 경우에만 이벤트로 확정하여 카운트를 증가시킨다.

$$L_{eff} = (i_{end} - i_{start} + 1) - 2 \cdot padding$$

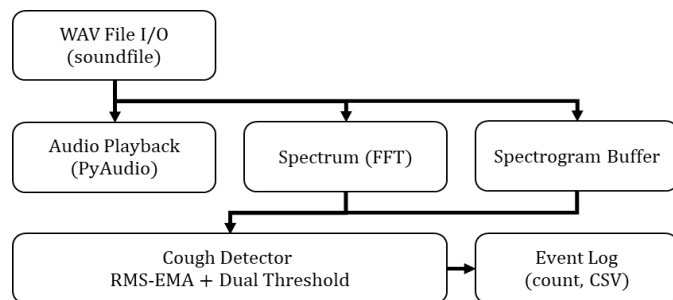


그림 1. 알고리즘 흐름도

알고리즘 지연은 *tolerance*와 *padding*에 의해 주로 결정되며, 다음과 같이 근사할 수 있다.

$$\Delta t_{lat} \approx \frac{tolerance + padding}{f_s}$$

연산은 제곱, 1차 지수평활, 임계치 비교가 주를 이루므로 메모리는 상수 수준이고, 시간 복잡도는 스트리밍 길이에 선형으로 비례한다. 파라미터는 배수(m_h, m_l), 시간(τ , *tolerance*, *padding*, *warmup*), 길이(*min_len*)로 표현되어 민감도와 특이도의 균형을 의도적으로 조정할 수 있다. 실무적으로는 m_h 2~4, m_l 0.1~1.0 범위에서 시작하고, τ 0.3~1.0초로 배경 변동을 추적하게 하며, *tolerance* 수 밀리초, *min_len* 수백 밀리초, *padding* 수백 밀리초로 오탐 억제와 경계 보정을 병행한다. 이 구성은 급상승-완만감쇠라는 기침 파형의 전형을 포착하면서 클릭성 충격음, 짧은 무성 간극, 대화 중 파편적 발성 등 비목표 이벤트를 체계적으로 배제한다.

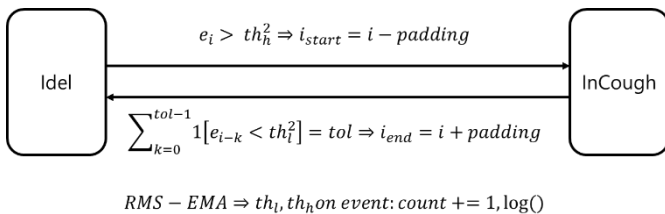


그림 2. 이중 임계치 히스테리시스 상태 머신

3. 연구 방법

평가 체계는 합성 데이터와 실제 녹음을 병행한다. 합성 세트는 백색 잡음과 협대역(narrowband) 톤을 중첩한 배경 위에, 밴드 제한 잡음에 지수 감쇠 포락선(envelope)을 곱한 버스트를 특정 시각에 삽입해 기침 유사 파형을 만든다. 한 버스트의 모델은 다음과 같다.

$$b(t) = (h * n)(t) \cdot e^{-\beta t} 1_{[0,L]}(t)$$

여기서 $n(t)$ 는 가우시안 잡음, h 는 짧은 해닝 윈도우 필터, 베타는 감쇠율, L 은 평균 지속시간이다. 이벤트 시각 집합은 간격과 중첩을 통제할 수 있도록 설계한다. 실제 데이터는 16kHz 단일 채널을 기준으로 실내-복도-야외 등 다양한 환경에서 수집하고, 독립된 주석자 두 명 이상이 시작, 종료, 확정 시점을 표기한 뒤 합의 절차로 기준 라벨을 마련한다.

정확도 평가는 이벤트 단위의 정밀도, 재현율, F1로 산출하며, 매칭은 검출 확정 시점과 라벨 시점의 절대 오차가 허용 범위 안에 있을 때 참양성으로 간주한다. 시점 정합도는 절대 오차의 중앙값과 95% 구간을 보고하여 경계 안정성을 확인한다. 지연 측정은 이벤트 종료 이후 확정까지의 시간으로 정의해, 파라미터가 실효 지연에 미치는 영향을 관찰한다.

강건성 실험에서는 SNR을 0, 5, 10, 20 dB로 변화시키고, 마이크 거리와 방향, 배경 소음 유형(대화, 충격음, 바람 소리 등)을 교차하여 성능 변동을 측정한다. 비교 대상으로는 배경 적응이 없는 고정 임계치 에너지 검출기, 히스테리시스가 없는 단일 임계치 방식, 간단한 수작업 특징(대역 에너지, 제로교차율, 스펙트럴 센트로이드)에 로지스틱 회귀를 적용한 경량 머신러닝 베이스라인을 포함한다. 이를 통해 적응형 RMS와 이중 임계치, 시간 제약의 결합이 각각의 단점을 어떻게 보완하는지 정량적으로 평가한다.

재현성을 위해 모든 파라미터($\tau, m_h, m_l, tolerance, padding, min_len, warmup$)를 실험 카드로 기록하고, 검출 결과는 이벤트 로그 CSV로 저장한다. 논문 도식은 코드로

자동 생성하며, 스펙트로그램 위에 이벤트 시각을 수직선으로 오버레이하여 시간-주파수 맥락에서 분절을 시각화하고, 확정 시점 기준 카운트의 계단형 증가와 플래시 윈도우를 함께 제시해 온라인 동작의 특성을 명확히 전달한다. 이러한 절차는 배치 전 튜닝 단계에서 파라미터 민감도를 빠르게 탐색하고, 데이터셋 간 이전성을 일관된 방법으로 점검하게 해준다.

4. 결론 및 향후 과제

본 논문은 RMS-EMA + 이중 임계치 히스테리시스 + 시간 제약을 결합한 경량 기침 검출-카운팅 절차를 정식화하고, 재현 가능한 평가 설계를 제안하였다. 이 방식은 낮은 지연/낮은 계산량/높은 해석 가능성을 동시에 확보하며, 입력 이득·환경 변화에 스케일 불변적으로 적응한다. 향후에는 (i) 대역별 에너지·스펙트럴 센트로이드-ZCR 등 다특징 결합, (ii) 준지도/소량학습으로 임계 배수-시간상수의 자동 최적화, (iii) 실제 임상 데이터셋에서 정밀도·재현율·지연의 체계적 보고와 파라미터 표준 운영절차(SOP) 정립을 진행할 계획이다. 또한 멀티채널 빔포밍과 간단한 후단 분류기를 가미해 오탐을 추가로 억제하는 경량 하이브리드 구성이 유망하다.

5. Acknowledgements

이 연구는 산업부 [No. 20023916, 하이브리드 근무환경의 감정노동자 감정위험 조기대응과 의사결정을 지원하는 AI 상담어시스트 시스템 개발 및 실증], 과학기술정보통신부-정보통신기획평가원의 핵심전략 R&D사업 지원을 받아 수행된 연구임(IITP-2025-RS-2025-02219190)

6.참고 문헌

- [1] J. D. Hunter. "Matplotlib: A 2D graphics environment." *Computing in Science & Engineering*, Vol. 9, No. 3, p. 90–95, 2007.
- [2] C. R. Harris, K. J. Millman, S. J. van der Walt, et al. "Array programming with NumPy." *Nature*, Vol. 585, p. 357–362, 2020.
- [3] M. R. Schroeder. "Number Theory in Science and Communication." Springer, 2006.
- [4] S. Mallat. "A Wavelet Tour of Signal Processing." Academic Press, 2009.
- [5] PortAudio Project. "PortAudio: Portable Cross-platform Audio I/O." Online documentation.
- [6] PySoundFile Developers. "PySoundFile (SoundFile) Documentation." Online documentation.