

상체 움직임 기반 공진자 상태공간 모델링을 이용한 비접촉 호흡 추정 기법에 관한 연구

A Study on a Non-Contact Respiration Estimation Method Using Oscillator State-Space Modeling Based on Upper-Body Motion

요약

기존 비접촉 호흡 추정은 얼굴 rPPG 신호에서 호흡 대역(0.08~0.50 Hz) 성분을 분리해 호흡률을 계산한다. 그러나 이 저주파 대역은 조명·자세 변화와 중첩되어 근본적인 잡음 한계를 가진다. 본 연구는 이러한 한계를 관측 대상(얼굴 피부 광학 신호)과 호흡의 생리적 기원(흉곽 운동) 사이의 불일치로 해석하고, 얼굴 대신 흉부 상체 움직임을 직접 관측하는 방식을 제안한다. 흉부 관심영역(ROI)에서 얻은 1차원 움직임 신호를 호흡·비호흡 움직임과 잡음이 섞인 준주기적 시변 신호로 모델링하고, 공진자 상태공간 모델과 두 가지 추적 헤드를 결합해 호흡 주파수와 위상의 시변 궤적을 추정한다. COHFACE 데이터셋에서 5종 움직임 신호와 2종 트래커 조합(총 15개)을 평가한 결과, KF-std는 Baseline 대비 시간 영역 MAE가 DoF에서 4.39, Optical Flow에서 2.28, 1D Profile 계열에서 평균 1.47만큼 감소하여 성능 개선을 보였다. 또한 흉부 상체 움직임 관측과 상태공간 표현을 결합함으로써, 관측 대상과 생리적 기원을 일치시키는 해석 가능한 호흡 신호 표현을 제공한다.

키워드 : 비접촉 생체신호 측정, 호흡 측정 기법, 공진자 상태 공간 모델, 칼만 필터

Abstract

Conventional non-contact respiration estimation extracts the 0.08~0.50 Hz band from facial rPPG signals to compute respiratory rate. However, this low-frequency band fundamentally overlaps with illumination and posture changes, leading to robustness limitations. We interpret this issue as a mismatch between the observation target (skin reflectance) and the physiological origin of respiration (thoracic motion), and thus propose directly observing upper-body chest motion that aligns with respiratory mechanics. One-dimensional motion signals obtained from a chest region of interest (ROI) are modeled as quasi-periodic, time-varying signals where respiratory motion, non-respiratory motion, and noise are mixed. An oscillator state-space model combined with two tracker heads is used to estimate time-varying trajectories of respiratory frequency and phase. Experiments on the COHFACE dataset with 15 combinations of five motion signals and two trackers show that KF-std improves performance by reducing the time-domain MAE compared to the Baselines by 4.39 for DoF, 2.28 for Optical Flow, and an average of 1.47 for the 1D Profile family. By aligning the observation target with respiratory mechanics and employing a state-space representation, the proposed method provides a physiologically interpretable representation of respiration.

Key Words : Non-contact vital sign measurement, Respiratory measurement, Oscillator state-space model, Kalman Filter

Received: Jan. 01, 2015

Revised : Jan. 01, 2015

Accepted: Jan. 01, 2015

†Corresponding author(kim@email.ac.kr)

본 논문은 한국연구재단 연구자지원사업(XXX-XXXX)에서 지원하여 연구하였음.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서 론

카메라 기반 비접촉 호흡 추정에는 얼굴 rPPG 신호에서 호흡 대역(0.08~0.50 Hz)을 분리해 호흡률을 추정하는 방식이 주류였다 [1, 2]. 그러나 이 저주파 대역은 조명·자세·미세 움직임과 중첩되어 구조적인 잡음 한계를 가진다 [2, 3]. 본 연구는 이러한 한계를 알고리즘 성능 문제가 아니라, 관측 대상(피부 광학 신호)과 호흡의 생리적 기원(흉곽 운동) 사이의 불일치에서 비롯된 구조적 문제로 해석한다. 이에 따라 호흡 기전에 직접 정합하는 흉부 상체 움직임을 새로운 관측 대상으로 선택한다 [4]. 상체 ROI에서 프레임 간 광류, 프레임 차분, 1D 프로파일 상관을 이용한 다섯 가지 움직임 신호를 구성하되, 이 신호를 곧바로 호흡으로 간주하지 않는다. 대신 이를 준주기적·시변 공진자로 해석하는 상태공간 표현을 통해 호흡 성분을 복원한다 [5, 6].

이를 위해 관측 계층-공진자 헤드-트래커 헤드로 구성된 3계층 상태 공간 파이프라인을 설계한다. 먼저 관측 계층에서 상체 ROI의 픽셀 변화를 추적해 다섯 종류의 1차원 움직임 신호를 추출한다. 공진자 헤드는 이 신호를 0.08~0.50 Hz 대역의 감쇠·회전 공진자로 표현하여, 위상과 주파수를 포함한 숨은 상태를 안정화한다. 마지막 단계에서는 두 종류의 추적기를 사용한다. KF-std는 공진자 출력의 기본 동역학을 안정적으로 추정하기 위한 기준 추적기로 배치되며 [7], UKF-freq는 비접촉 환경 특유의 비선형적, 비가우시안적 관측 특성을 반영하기 위해 이를 확장한 형태로 설계하였다 [8]. 두 추적기는 서로 다른 관점에서 공진자 출력을 해석함으로써, 호흡 주파수·위상의 시변 특성을 생리적 대역 내에서 추정한다.

본 연구의 기여는 다음과 같다.

1. 관측 대상을 얼굴 피부가 아닌 흉곽 운동으로 전환하여, 호흡의 생리적 기원과 신호 처리 대역을 일치시킨다.

는 비접촉 호흡 추정 파이프라인을 제안한다.

2. (공진자 상태공간 모델과 선형·비선형 기반 두 추적기를 결합하여, 호흡 주파수와 위상을 생리적 대역 내에서 안정적으로 추정하는 방법을 제시한다.
3. COHFACE 데이터셋에서 10개 모션-트래커 조합을 시간·주파수 영역 지표로 체계적으로 비교하여, 얼굴 기반 rPPG와 유사한 호흡률 추정 성능과 함께 상태공간 기반 추정의 해석 가능성을 실증한다.

2. 제안 방법

본 절에서는 상체 움직임 영상으로부터 호흡 주파수와 위상을 추정하기 위한 전체 파이프라인을 기술한다.

그림 1은 제안 방법의 모식도이며, 세 단계로 구성된다.

1. 흉부 ROI 기반 1차원 움직임 관측,
2. 공통 전처리 및 공진자 상태공간 모델 임베딩,
3. 주파수·위상 트래커를 통한 시변 호흡률 복원.

이를 통해 관측 대상(흉곽 운동)과 호흡의 생리적 기원을 일치시키면서, 상태공간 표현을 통해 잡음과 시변성을 일관된 형태로 모델링한다.

2.1 시스템 개요

입력은 상반신이 포함된 RGB 비디오이며, 프레임별로 흉부 영역을 관심 영역(ROI, Ω)으로 정의한다. 연속 프레임 I_t 에서 ROI 내 픽셀 변화를 이용해 다섯 종류의 1차원 움직임 신호 $y(t)$ 를 생성한다(DoF, OF, 1D Profile 3종). 이 신호들은 모두 호흡과 관련된 저주파 성분을 일부 포함하지만, 자세 변화·조명 변화·스파이크 등 비가우시안 잡음이 중첩된 시변 준주기적 신호이다.

생성된 움직임 신호는 공통 전처리를 거쳐 정규화 신호

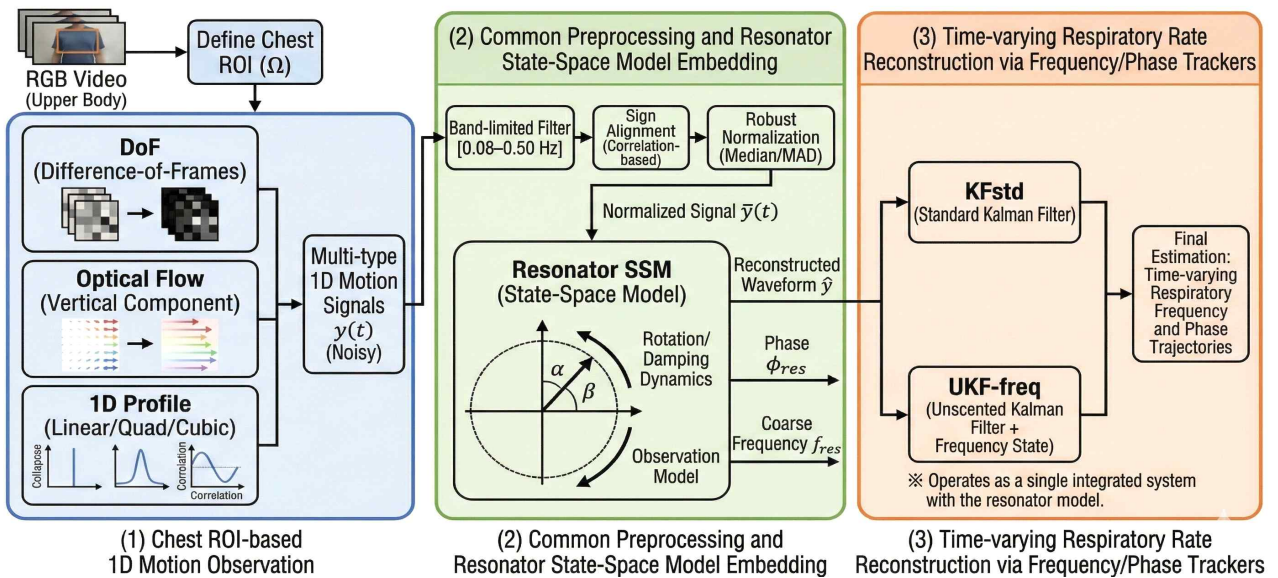


그림 1. 제안하는 기법: 공진자 상태공간 모델을 이용한 상체 움직임 기반 비접촉 호흡 측정

Fig. 1. Proposed method: upper body motion-based respiratory estimation pipeline

$\tilde{y}(t)$ 로 변환되며, 이후 공진자 상태공간 모델에 의해 2차원 공진자 상태 x_t 로 임베딩된다. 이 상태는 호홉 리듬을 협대역에서 회전·감쇠하는 공진자로 해석한 것으로, 위상·진폭·즉시 주파수가 숨은 상태 변수로 자연스럽게 파생된다. 마지막으로 공진자 출력(재구성 파형, 위상, coarse 주파수)을 입력으로 사용하는 두 종류의 트래커 헤드(KF-std, UKF-freq)가 서로 다른 동역학 가정하에서 호홉 주파수 궤적을 추정한다.

2.2 상체 움직임 관측 및 전처리

2.2.1 Difference-of-Frames (DoF)

DoF 신호는 연속 프레임 간 밝기 차이 프레임 간 밝기 차이 $\Delta I_t(x)$ 에서 임계치(τ) 이상 증가한 픽셀을 사건으로 간주하고 그 개수를 합산함으로써 프레임 간 전체 움직임 크기를 정의한다 [3, 4]. 해당 방식은 큰 움직임과 조명 변화에 비교적 민감하다는 특성이 있다.

$$y_{DoF}(t) = \sum_{x \in \Omega} 1[\Delta I_t(x) > \tau] \quad (1)$$

2.2.2 Optical Flow (OF)

OF 기반 관측은 인접 프레임 간 픽셀 이동 벡터 $o_t(x)$ 를 추정하고, 호홉 축에 해당하는 수직 성분 $v_t(x)$ 의 중앙값을 프레임 간 대표 변위로 사용한다 [3, 9]. 이는 반사·주름과 같은 국소 이상치를 억제하면서, 홍곽의 미세한 상하 변위를 비교적 안정적으로 포착한다.

$$o_t(x) = (h_t(x), v_t(x)) \quad (2)$$

$$y_{OF}(t) = \text{median}_{x \in \Omega}(v_t(x)) \quad (3)$$

2.2.3 1D Profile 계열(Linear/Quadratic/Cubic)

1D Profile 계열은 ROI를 행 방향으로 축약한 1차원 신호의 연속적 변화를 이용하여 프레임 간 상대적 움직임을 추정한다 [3, 10]. 각 프레임에서 산출된 프로파일 $p_t(r)$ 은 프레임 간 형태 변화를 민감하게 반영하므로, 인접 프레임에 대해 상호 상관 $c_t(k)$ 를 계산하면 프레임 간 상대적 이동 정도를 정량적으로 평가할 수 있다.

이후, 정수 화소 기반으로 이산적으로 계산된 상호 상관 계수열의 분해능을 향상시키기 위해 m 배 보간(본 연구에서는 $m=100$)을 적용하여 준연속 상호 상관 곡선을 구성한다. 보간된 곡선에서 최댓값을 갖는 인덱스 n^* 를 변위를 나타내는 지표로 사용한다. 최종적으로 상대적 수직 변위 $y_{prof}(t)$ 는 수식 (5)와 같이 정의된다 [10].

$$p_t(r) = \Delta_r(0.5(\mu_t(r) + \sigma_t(r))) \quad (4)$$

$$c_t(k) = \sum_r p_t(r) \cdot p_{t-1}(r-k), \quad y_{prof}(t) = f_s \cdot n_t^* \quad (5)$$

보간 차수에 따라 Linear / Quadratic / Cubic 세 가지 방식으로 구분되며, 차수 변화는 보간된 상호 상관 곡선의 매끄러움과 피크 위치의 세밀도에 영향을 미친다. 그러나 프로파일 생성, 상호 상관 계산, 최대 지표 추출, 변위 산출의 전체 절차는 동일하게 유지되어, 프레임 간 미세한 형태 변화를 안정적으로 반영하는 특성을 갖는다.

2.2.4 전처리

움직임 신호 $y(t)$ 에는 조명 변화, 자동 노출, 상체 자세 이동, 비가우시안 스파이크 등 다양한 잡음이 직접 중첩된다. 공진자 모델과 평가 단계의 주파수 대역을 일치시키고 잡음의 영향을 최소화하기 위해, 다음 세 단계 전처리를 수행한다.

1. 대역 제한: 0.08~0.50 Hz 대역통과 필터를 적용하여 저주파 외란을 억제하고, 신호 에너지를 생리적 호홉 대역에 집중시킨다.
2. 부호 정렬: ROI 구성·카메라 시점에 따라 들숨·날숨 부호가 반전될 수 있으므로, 초기 구간에서 코사인 참조 신호와의 상관을 이용해 필요시 부호를 반전시킨다.
3. 강건 정규화: 중앙값과 중앙 절대편차(MAD)에 기반한 robust z-score 정규화를 적용하고, 필요시 클리핑을 수행하여 이상값이 관측 잡음 공분산 추정에 미치는 영향을 완화한다 [11].

이 과정을 통해 얻은 정규화 신호 $\tilde{y}(t)$ 는 비접촉 호홉 신호의 특성에 맞게 안정화되며, 이후 공진자 상태공간 모델의 입력으로 사용된다.

2.3 공진자 헤드

2.3.1 공진자 상태공간 모델 개요

공진자 상태공간 모델은 협대역 호홉 성분을 회전·감쇠하는 2차원 공진자 상태 $x_t = [x_1(t), x_2(t)]^T$ 로 표현한다 [5, 6]. 이 상태는 실수축과 허수축에 대응하는 공진자 좌표이며, 위상과 진폭은 각각 상태 벡터의 각도와 크기로 해석된다.

2.3.2 상태 전이

상태 전이는 진폭 감쇠 인자 ρ 와 회전 행렬 R 을 따라 수식 (6)과 같은 회전·감쇠 동역학으로 정의된다.

$$x_{t+1} = \rho \cdot R(\omega_t \Delta t) x_t + w_t, \quad w_t \sim N(0, Q) \quad (6)$$

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (7)$$

$$\rho = \exp(-1/(f_s \cdot \tau_{env})) \quad (8)$$

$$Q = q_x \cdot I_x, \quad Q \in \mathbb{R}^{2 \times 2} \quad (9)$$

여기서 ω_t 는 순간 각 주파수이며, f_t 는 Hz 단위 호홉률, $\Delta t = 1/f_s$ 는 샘플 간격을 의미한다. 진폭 감쇠 인자 ρ 는 샘플링 주파수 f_s 와 포락선 시간 상수 τ_{env} 에 의해 결정된다. w_t 는 상태 변화에서 생기는 불규칙성을 반영하는 관측 잡음이며, Q 는 잡음을 흡수하기 위한 등방성 상태 잡음 공분산으로, 잡음의 확산 정도를 조절하는 스칼라 q_x 와 2차 항등 행렬 I 로 구성된다 [5].

구현에서는 $f_s = 64\text{Hz}$ 로 설정하고, 신호 품질 및 모션 강도에 따라 q_x 가 동적으로 스케일되며 UKF-freq에서는 주파수 확산을 위한 별도 항 q_y 를 함께 사용한다.

2.3.3 공통 공진자 헤드

공진자 헤드는 전처리된 연속 신호 $\tilde{y}(t)$ 를 샘플링 주파수 f_s 에서 샘플 인덱스 k 로 이산화하여 얻은 이산 표본 $\tilde{y}_{raw,t}$ 을 구성한다. 해당 이산 표본은 상태의 실수축 성분과의 선형 관계로 모델링되며, 동시에 관측 과정에서 발생하는 불확실성을 잡음항 n_t 로 표현하여, 수식 (10)과 같이 최종 이산 관측 \tilde{y}_t 로 정의된다.

$$\tilde{y}_t = [1 \ 0] x_t + n_t, \quad n_t \sim N(0, R) \quad (10)$$

2차원 공진자 상태는 $x_t = [x_1(t), x_2(t)]^T$ 로 정의된다. 이때, 실수축 $x_1(t)$ 는 이산 관측 \tilde{y}_t 와 선형적으로 대응되며, 허수축 $x_2(t)$ 는 Hilbert 변환 없이 상태 전이 식 (6)을 따라 회전·감쇠 동역학을 내부적으로 생성·갱신한다. 이를 통해 1차원 관측으로부터 직교축이 자동 형성되며, 시변 진폭·위상·주기를 상태공간 내부에서 표현할 수 있다.

$$R = \max((\alpha \cdot \sigma)^2, R_{\min}) \quad (11)$$

관측 잡음 공분산 R 은 식 (11)과 같이 정의하며, 여기서 α 는 관측 스케일 조정 계수, σ 는 전처리 단계에서 추정된 MAD 기반 분산 추정치, R_{\min} 은 관측잡음의 하한을 의미한다. 구현에서는 가드레일로 $\hat{\sigma} \leq 10$, $R \leq 50r_{v, \text{floor}}$ 의 상한을 두고, 스펙트럼, SNR, 움직임 강도 등의 지표를 활용해 비가우시안 스파이크나 국소 밝기 변화를 관측잡음으로 흡수함으로써 필터의 수치적 불안정을 완화한다.

공진자 모델은 다음 세 가지 핵심 출력을 제공한다.

$$\hat{s}(t) = x_1(t) \quad (12)$$

$$\phi(t) = \text{atan2}(x_2(t), x_1(t)) \quad (13)$$

$$\hat{f}[t] = \frac{1}{2\pi \Delta t} \text{wrap}_{(-\pi, \pi]}(\theta[k] - \theta[k-1]) \quad (14)$$

$\hat{s}(t)$ 는 재구성된 호홉 파형이며, $\phi(t)$ 와 $\hat{f}[t]$ 는 후속 트래커 헤드에서 기준 위상과 coarse 주파수로 사용된다. 식 (14)는 인접한 두 샘플($\theta[k], \theta[k-1]$)의 위상 차분을 통해 순간 주파수로 근사하는 방법을 나타낸다. 이때, wrap 함수는 위상 차이가 2π 주기를 넘어갈 때 발생하는 모호성을 제거하기 위해 $(-\pi, \pi]$ 범위로 정규화하며, 주파수 단위로 환산한다.

구현에서는 $\hat{f}[t]$ 를 인접 위상 차분으로 계산한 후, 0.08~0.50 Hz로 클램프하고 지수평활을 적용하여 순간 위상 점프와 고주파 잡음을 완화한다. 초기 주파수 f_0 는 Welch 스펙트럼, 자기상관, Hilbert 기반 후보를 각각의 신뢰도로 가중하여 블렌딩 한 뒤 0.08~0.50 Hz 내에서 선택하며, 유효 후보가 존재하지 않을 경우 대역 중심값을 기본값으로 사용한다. 전 과정에서 주파수는 항상 생리적 대역으로 클램프되며, 상태·관측 공분산은 floor·cap·jitter 설정을 통해 양의정부성 조건을 안정적으로 유지하여 필터 발산을 방지한다.

2.4 주파수·위상 트래커 헤드

상태 공간 모델에서는 확률적 잡음이 포함된 동적 시스템의 관측 불가능한 내부 상태와 확률적 요인을 합리적으로 추정하고 시스템 모델과 실제 관측치 사이의 오차 공분산을 최소화하기 위해, 칼만 필터 기반의 예측-수정 프레임워크가 필수적으로 요구된다.

이에 트래커 헤드에서는 비접촉 환경에서 얻어진 데이터의 불안정성 및 시변하는 호홉 동역학을 정밀하게 추적하는 것을 목표로 한다. 공통 공진자 헤드에서 산출된 이산 출력 $\hat{s}_t, \phi_t, \hat{f}_t$ 를 관측으로 정의하고, 내부 상태 공간 정의와 시스템의 선형성 가정 여부에 따라 KF-std와 UKF-freq라는 두 가지 트래커를 구성하여 최적의 시변 호홉 신호를 추정한다.

2.4.1 KF-std (Standard Kalman Filter)

KF-std는 호홉 신호가 단기적으로 안정적인 주파수를 유지한다는 가정하에, 공진자 상태 $[x_1(t), x_2(t)]^T$ 를 선형 시불변 시스템으로 모델링하는 기준 추적기이다 [7]. 해당 트래커에서는 초기화 단계에서 결정된 기준 주파수를 전 구간 동안 고정 파라미터로 구속함으로써, 시스템의 선형성을 보장하고 계산 복잡도를 최소화하는 동시에 관측잡음에 대한 최적의 선형 추정치를 제공한다.

시스템의 상태 전이 행렬 F 는 수식 (15)와 같이 회전·감쇠 동역학을 따르는 선형 방정식으로 정의된다.

$$F = \rho \begin{bmatrix} \cos(\omega_0 \Delta t) & -\sin(\omega_0 \Delta t) \\ \sin(\omega_0 \Delta t) & \cos(\omega_0 \Delta t) \end{bmatrix}, \quad \omega_0 = 2\pi f_0 \quad (15)$$

이때 상태 전이 행렬 F 는 초기 관측에서 추정된 기준값

ω_t 대신, 초기 관측에서 추정된 기준값 ω_0 를 상수로 고정하여 정의된다. 이는 상태 벡터가 위상 평면 상에서 일정한 속도 ω_0 로 회전하도록 강제함을 의미하며, 상태 잡음 공분산 Q 는 수식 (9)의 등방성 정의를 따름으로써 고정된 회전 궤적 주변에서의 확률적 확산을 허용한다.

관측 단계에서는 공진자 헤드의 이산 출력 \tilde{y}_t 와 예측 상태 간의 오차를 기반으로 상태를 갱신한다. 여기서 주목할 점은, 동역학 모델은 선형 시불변 시스템(LTI)으로 고정되어 선형성을 유지하지만, 관측잡음 공분산 R_t 는 수식 (11)에 정의된 바와 같이 MAD 기반의 동적 스케일링을 유지한다는 것이다. 이를 통해 KF-std는 계산 효율성을 확보하면서도 스파이크나 급격한 진폭 변화와 같은 비정상 잡음 유입에 적응적으로 대응하는 강건함을 갖는다.

필터링을 통해 정제된 공진자 상태 \hat{x}_t 로부터 최종적인 호흡 위상 ϕ_t 와 즉시 주파수 \hat{f}_t 는 각 수식 (16), (17)과 같이 산출한다.

$$\phi_t = \text{atan2}(x_{2,t}, x_{1,t}) \quad (16)$$

$$\hat{f}_t = \text{EMA} \left(\text{clip}_{[f_{\min}, f_{\max}]} \left(\frac{1}{2\pi \Delta t} \Delta(\text{unwrap}(\phi_t)) \right) \right) \quad (17)$$

순간 주파수 \hat{f}_t 는 추출된 위상 ϕ_t 에 대해 불연속성을 제거하는 위상 풀림을 선행한 후, 인접 시점 간 차분을 통해 계산된다. 이후 생리적 호흡 대역 내로 클램핑하고 지수 이동 평균(EMA)을 적용하여 미분 과정에서 증폭된 고주파 잡음을 완화한다.

KF-std는 선형성 가정 덕분에 수렴 속도가 빠르고 실시간 처리에 유리하다. 그러나 실제 호흡수는 생리적 요인에 의해 시변하는 특성을 가지므로, 고정된 기준 주파수 ω_0 를 사용하는 선형 모델은 실제 호흡수와 모델 주파수 간의 괴리가 커질 때 위상 추적에 실패하는 주파수 풀림 현상이 발생할 구조적 한계를 갖는다. 이는 주파수 변화 자체를 확률적 상태 모델링해야 할 필요성을 시사한다.

2.4.2 UKF-freq (Unscented Kalman Filter + Frequency State)

비접촉 측정 환경은 호흡수의 급격한 변화로 인한 시스템 동역학의 비선형성과, 조명·모션 아티팩트에 의한 비가우시안 잡음이 공존하는 특성을 갖는다. KF-std는 이러한 환경에서 선형성 가정을 유지하기 위해 파라미터를 고정하는 전략적 제약을 두었으나, UKF-freq에서는 호흡수 자체를 추정해야 할 확률적 상태로 정의하고 비선형 추적을 도입함으로써 환경적 불안정성을 구조적으로 극복하고자 한다 [8].

이를 위해 기존 2차원 공진자 상태에 로그 주파수 성분(상태) $\log f_t$ 를 추가한 3차원 확장 상태 벡터 z_t 를 수식 (18)과 같이 정의한다.

$$z_t = [x_1(t), x_2(t), \log f_t]^T \quad (18)$$

여기서 주파수에 로그 변환($\log f_t$)을 적용한 이유는, 추정 과정에서 주파수가 항상 양수임을 수학적으로 보장하고, 스케일에 민감하지 않은 안정적인 갱신을 유도하기 위함이다. 확장된 상태 공간에서의 시스템 동역학은 주파수 상태와 공진자 상태의 상호작용에 의한 비선형 전이 특성을 갖는다.

$$\log f_{t+1} = \log f_t + w_t^{(f)} \quad (19)$$

$$\begin{bmatrix} x_{1,t+1} \\ x_{2,t+1} \end{bmatrix} = \rho R(f_t) \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + w_t^{(x)} \quad (20)$$

수식 (19)와 (20)은 각각 주파수 동역학의 유연성과 파형의 형태적 불확실성을 모델링한다. $w_t^{(f)}$ 는 주파수 확산 잡음으로, 랜덤 워크 과정을 통해 호흡 속도의 시변적 변화를 허용한다. $w_t^{(x)}$ 는 공진자 상태 잡음으로, 진폭의 변동이나 비정형적인 찌그러짐과 같은 파형 형태의 불규칙성을 흡수한다. 상기한 두 잡음 성분 ($w_t^{(f)}$, $w_t^{(x)}$)은 상태 잡음 공분산 Q_t 에 블록 대각 형태로 포함되어, 필터가 주파수 변화와 파형 왜곡을 독립적인 확률 변수로 분리하여 추적하게 한다.

핵심은 수식 (20)의 공진자 갱신 과정이다. 전이 행렬 $R(f_t)$ 가 고정된 상수가 아니라 매 시점의 확률적 상태 $f_t = \exp(z_{3,t})$ 에 종속되어 재계산되므로, 전체 시스템은 상태 변수가 시스템 행렬을 결정하는 강한 비선형 구조를 띈다.

이러한 비선형성과 비가우시안 분포와 같은 통계적 특성을 정밀하게 전파하기 위해, UKF-freq는 무항 변환을 적용한다 [12]. 시그마 포인트 생성을 위한 파라미터 (α, β, κ)는 가우시안 분포 가정에 적합한 상수로 설정하여 비선형 함수의 평균과 공분산을 근사하며, 필터 갱신 과정에서의 수치적 안정성은 2.3.3절의 공분산 강건화 기법 (floor-cap:jitter)을 동일하게 적용하여 양의 정부호성을 보장함으로써 확보된다.

관측 단계에서는 UKF의 가우시안 가정하에서도 비가우시안 잡음을 효과적으로 제어하기 위해, 2.3.3절의 동적 관측 모델을 심화 적용한다. 관측잡음 공분산 R_t 는 스칼라 (scalar)로 정의되며, 그 값은 MAD 기반 추정치에 모션 강도, ROI SNR, 스펙트럼 신뢰도를 반영한 스케일링을 거쳐 산출된다. 이는 이상치 구간에서 R_t 를 증폭시켜 칼만 이득(Kalman Gain)을 간접적으로 억제함으로써, 비가우시안 스파이크가 상태 추정에 미치는 영향을 능동적으로 차단한다. 최종 호흡수 \hat{f}_t 는 수식 (21)과 같이 확장 상태 z_t 의 세 번째 성분을 역변환하여 복원된다.

$$\hat{f}_t = \text{EMA} \left(\text{clip}_{[f_{\min}, f_{\max}]} \left(\exp(z_{3,t}) \right) \right) \quad (21)$$

이때 \hat{f}_t 는 별도의 처분 과정 없이 상태 값으로부터 직접 산출된다. 안정적인 출력을 위해 로그 도메인에서 우선적으로 대역 제한을 수행한 후 지수 변환을 거치며, 다시 한번 생리적 대역 $[f_{\min}, f_{\max}]$ 내로 클리핑하고 지수 이동 평균(EMA)을 적용하여 최종적으로 출력된다. 결과적으로 UKF-freq는 비선형 동역학을 처리하는 UKF 알고리즘과 비가우시안 잡음을 억제하는 관측 모델 설계를 결합함으로써, 비접촉 관측 환경의 불확실성을 구조적으로 수용하고 호흡률 변화에 유연하게 적응하는 정합성을 제공한다.

3. 실험 및 결과

본 절에서는 제안한 상체 움직임 기반 비접촉 호흡 정량적으로 평가한다. 기존 비접촉 호흡 연구는 주로 주파수 영역의 호흡률만을 평가하는 경향이 있으나, 이러한 방식은 흡기-호기 피크, 파형의 강도, 시변 궤적과 같은 시간 영역의 유효 정보가 반영되지 않는 한계가 있다. 이에 따라 주파수 영역 지표와 시간 영역 지표를 함께 고려하여, 다섯 종류의 상체 움직임 관측법과 두 종류의 공진자-트래커 조합을 함께 비교 평가한다.

3.1 실험 설정

실험은 COHFACE 데이터셋의 모든 피험자 비디오를 입력으로 사용한다 [13]. 검증용 위한 Ground Truth은 동기화된 센서 기반 호흡 신호를 원시 샘플링률에서 다운샘플링·보간하여 시변 호흡률 형태로

재구성하였다. 상체 움직임 기반 비접촉 호흡 추정을 위해 DoF, Optical Flow, 1D Profile 3종의 다섯 가지 관측 기법을 사용하고, 각 관측 신호에 두 종류의 공진자-트래커(KF-std, UKF-freq)를 결합하여 총 10가지 제안 모델을 구성하였다. 기본 관측 방식 5개를 포함해 총 15개 방식을 비교한다.

모든 방식은 동일한 절차로 처리한다.

- (1) 0.08~0.50 Hz 대역통과 전처리
- (2) 30초 창-1초 스트라이드 기반 시퀀스 평가
- (3) 추정 파형과 GT의 시간축 정렬
- (4) 창 단위 평가지표 산출

산출된 창(window) 단위 결과는 트라이얼 내부에서 중위수로 1차 집계하며, 전체 데이터셋에서도 중위수로 최종 요약하여 이상치 영향과 분산을 최소화한다. 평가는 시간 영역과 주파수 영역에서 병행한다. 시간 영역에서는 정렬된 시변 호흡 파형과 GT 간의 불일치를 RMSE, MAE, PCC, SNR로 측정하여 파형의 시간적 안정성과 위상-진폭 추종 품질을 평가한다. 주파수 영역에서는 Welch 전력 스펙트럼 밀도를 기반으로 최대 에너지 피크를 추출하고 이를 분당 호흡수(RPM)로 변환하여 RMSE, PCC를 산출함으로써 스펙트럼 기반 집중도와 호흡률 추정 정확도를 평가한다 [14].

3.2 실험 결과

[표 1]의 결과에서, 다섯 가지 관측 방식 모두에서 Baseline 대비 KF-std 결합 시 시간 영역 MAE가 DoF에서 4.39, Optical Flow에서 2.28, 1D Profile 계열에서 평균 1.47만큼 감소하며 큰 개선을 보였다. 이는 전 구간에서 안정적인 파형 복원 특성이 확보되었음을 의미한다.

표 1. 호흡추정 기법별 성능 지표 비교.

Table 1. Performance Comparison of Respiration Estimation Methods.

Method	Tracker	RMSE(t)	MAE(t)	PCC(t)	SNR(t)	RMSE(RR)	MAE(RR)
DoF	Baseline	6.410	6.130	-0.100	11.574	4.360	2.930
	KF-std	2.180	1.745	0.190	29.387	2.215	1.760
	UKF-freq	2.630	2.295	0.210	22.898	3.105	2.245
Optical Flow	Baseline	3.640	2.860	0.090	16.841	0.355	0.295
	KF-std	0.725	0.580	0.500	27.060	0.340	0.260
	UKF-freq	1.455	1.230	0.420	21.694	0.430	0.320
1D Profile (Linear)	Baseline	2.780	2.190	0.200	15.363	0.450	0.325
	KF-std	0.720	0.590	0.520	30.238	0.455	0.365
	UKF-freq	0.945	0.760	0.530	24.756	0.535	0.380
1D Profile (Quadratic)	Baseline	2.335	1.920	0.150	18.292	0.260	0.195
	KF-std	0.630	0.530	0.520	30.437	0.330	0.260
	UKF-freq	0.805	0.650	0.370	25.126	0.400	0.305
1D Profile (Cubic)	Base	2.475	1.954	0.140	18.249	0.285	0.195
	KF-std	0.635	0.520	0.500	30.328	0.330	0.250
	UKF-freq	0.805	0.680	0.450	25.224	0.375	0.310

주파수 영역 평가에서도 KF-std가 UKF-freq 대비 전 반적으로 낮은 오차를 기록하며 더 높은 안정성을 보였다. 이는 UKF-freq가 이론적으로는 시변 주파수 추적에 유리한 비선형 구조를 갖추고 있음에도 불구하고, 실제 성능은 상태 잡음 공분산 Q 와 같은 하이퍼파라미터 설정에 매우 민감하게 반응함을 시사한다. 특히 본 실험의 파라미터 최적화 과정이 과형 복원 오차의 최소화와 신호 안정성에 비중을 두었기 때문에, UKF의 주파수 상태가 갖는 가변성(Random Walk)이 과도하게 억제되어 급격한 호흡 변화에 대한 추종 성능이 충분히 발휘되지 못한 것으로 해석된다.

종합하면, 시간 영역과 주파수 영역 모두에서 KF-std가 일관된 성능 우위를 보였으나, 이는 고정 파라미터 모델의 강건함과 비선형 모델의 최적화 난이도 간의 트레이드오프를 보여준다. 경험적으로 UKF-freq가 높은 주파수 추종 잠재력을 보였던 점을 고려할 때, 향후 연구에서는 단순한 과형 복원을 넘어 시간 영역의 안정성과 주파수 영역의 반응성을 동시에 고려하는 복합 목적함수 기반의 최적화가 모델의 잠재력을 온전히 끌어내는 데 필수적임을 확인했다.

3.3 결론 및 향후 연구

본 연구는 상체 움직임 기반 비접촉 호흡 추정에서 관측 기전과 상태공간 구조의 정합성이 성능 안정성에 핵심적으로 작용함을 보였다. 관측 신호 중 1D profile 계열은 흉곽의 상·하 변위를 단일 축 정렬로 직접 추정해 호흡 기전과 가장 일치하며, 측면 흐름·조명 요동 등 비호흡성 교란을 상대적으로 덜 반영해 공진자 기반 추정기의 주파수·위상 추종 안정성을 높였다.

또한 비접촉 호흡 신호가 갖는 비선형성, 비가우시안 잡음, 장기 드리프트는 칼만필터류 상태공간에서 상태와 관측에 대한 합리적 추정을 위해 반드시 고려해야 할 제약임을 확인하였다. 주파수·위상·진폭과 같은 어떤 성분을 상태로 두고 어떤 요인을 관측잡음으로 처리할지에 대한 설계 선택이 모델의 해석성/적응성을 좌우하며, 환경 변화에 대응하는 노이즈 적응과 상태 확장이 요구된다.

향후 연구는 1D profile를 기본 관측 축으로 유지하되, 공진자 상태를 흡기·호기 피크, 진폭, 비대칭성 등 생리적 요인으로 확장해 해석 가능한 호흡 표현을 구축하는 방향으로 전개될 것이다. 이를 통해 단순 호흡률을 넘어 비접촉 잡음 환경에서도 생리적 구조를 보존하는 E2E 호흡 모델로 발전시키고, 다양한 조건의 선형·비선형 특성 및 가우시안·비가우시안 잡음 분석을 바탕으로 적합한 상태공간 설계를 정립하는 것이 목표이다.

Conflict of Interest

저자는 본 논문에 관련된 어떠한 잠재적인 이해상충도 없음을 선언한다.

References

- [1] C. Massaroni, A. Nicolò, D. Lo Presti, M. Sacchetti, S. Silvestri, and E. Schena, "Contactless monitoring of breathing patterns and respiratory rate at the start and end of exercise: A review," *IEEE Sensors Journal*, vol. 19, no. 15, pp. 5898–5907, 2019.
- [2] W. Wang and X. Wang, "Camera-based respiration monitoring: Motion and PPG-based measurement," in *Contactless Vital Signs Monitoring*, W. Wang and X. Wang, Eds. Academic Press (Elsevier), 2021, ch. 12.
- [3] W. Wang and A. C. den Brinker, "Algorithmic in-sights of camera-based respiratory motion extraction," *Physiological Measurement*, vol. 43, no. 7, 07TR01, 2022.
- [4] M. A. Al-Naji, K. Gibson, and J. Chahl, "Remote sensing of physiological signs using a hovering unmanned aerial vehicle," *Bioengineering*, vol. 4, no. 3, p. 55, 2017.
- [5] A. Solin and S. Särkkä, "Explicit link between periodic covariance functions and state space models," *Proceedings of the 17th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Reykjavik, Iceland, vol. 33, pp. 904–912, 2014.
- [6] A. Solin, "Stochastic differential equation methods for spatio-temporal Gaussian process regression," *Doctoral Thesis*, Aalto University, Helsinki, Finland, 2016.
- [7] S. Särkkä, *Bayesian Filtering and Smoothing*, Cambridge University Press, 2013.
- [8] L. Mirmohamadsadeghi and J.-M. Vesin, "Respiratory rate estimation from the ECG using an instantaneous frequency tracking algorithm," *Biomedical Signal Processing and Control*, vol. 14, pp. 66–72, 2014.
- [9] S. M. M. Islam, S. M. Alahi, and S. C. Mukhopadhyay, "Non-contact monitoring of breathing pattern and respiratory rate using optical flow and RGB-D camera," *IEEE Access*, vol. 10, pp. 63889–63901, 2022.
- [10] P. Franchuk and T. Yezerska, "Improved touchless respiratory rate sensing," *Proceedings of the 2022 IEEE 9th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T)*, Kharkov, Ukraine, pp.

235–240, 2022.

- [11] R. Janssen, W. Wang, A. C. den Brinker, and G. de Haan, "Video-based respiration monitoring with automatic region of interest detection," Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, pp. 5062–5065, 2015.
- [12] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," Proceedings of the IEEE, vol. 92, no. 3, pp. 401–422, 2004.
- [13] G. Heusch, A. Anjos, and S. Marcel, "A reproducible study on remote heart rate measurement," arXiv preprint arXiv:1709.00962, 2017.
- [14] J. Lazaro, E. Gil, R. Bailon, A. Mincholé, and P. Laguna, "Deriving respiration from photoplethysmographic pulse width," Medical & Biological Engineering & Computing, vol. 51, no. 1–2, pp. 233–242, 2013.

저 자 소 개
