

논문 2024-19-20

신호 처리를 통한 비접촉 광용적맥파 기반 얼굴 영역별 산소포화도 추정 성능 비교

(Comparison of rPPG-based Oxygen Saturation Measurement Performance by Face Area with Simple Signal Processing)

조 호 진, 권 남 규*

(Hyojin Jo, Nam Kyu Kwon)

Abstract : This study evaluates the performance of remote photoplethysmography (rPPG)-based oxygen saturation (SpO₂) estimation across different face areas. The information in the red and green channels is extracted through spatial averaging on the detected area, followed by detrending, min-max normalization, and low-pass filtering are performed. Finally, SpO₂ is estimated using the Ratio of Ratios method based on Beer-Lambert law. Using the PURE dataset, SpO₂ estimation accuracy was evaluated with mean absolute error (MAE). Experimental results showed the entire face provided the highest accuracy, followed by the forehead, left cheek, right cheek, and lips.

Keywords : Oxygen saturation, Remote Photoplethysmography, Ratio of ratios, Segmentation, Region of interest

I. 서 론

산소포화도는 환자의 호흡과 혈액 순환 상태를 평가하는 중요한 지표이다. 산소포화도 모니터링은 호흡 문제, 감염, 빈혈 및 기타 건강 상태의 초기 징후를 감지하는 데 필수적이다. 산소포화도를 측정하는 전통적인 방식으로, Photoplethysmography (PPG)의 비침습적 소형 장치인 펄스 옥시미터를 손가락 또는 귀에 착용하는 방법이 존재한다. 또한, 최근 스마트워치와 같은 웨어러블 기기에서 심박수 및 기타 생체 정보를 쉽게 모니터링 할 수 있는데, 이 또한 PPG 센서의 원리를 이용한 것이다. 하지만, 이러한 접촉식 센서나 웨어러블 기기는 신생아나 화상 환자와 같이 피부가 예민한 경우 장시간 사용이 어렵다. 또한 산소포화도 모니터링이 필수적인 폐 또는 심혈관 질환이 있는 환자의 경우, 손가락으로 흐르는 혈류의 양이 부족하여 펄스 옥시미터를 통한 산소포화도 측정이 어렵다. 그 외에도 접촉식 센서는 환자 간 교차 감염의 위험이 증가하는 문제점이 존재한다 [1].

이러한 문제점들을 극복하기 위해 카메라 기반 비접촉식의 remote Photoplethysmography (rPPG) 기반 산소포화도 측정 연구가 활발히 수행되고 있다 [2]. 기존의 방식은 PPG 센서를 직접 착용해야 하지만, rPPG 기반의 방식은 카메라를 통해 혈액량 변화로 인한 피부의 미세한 색 변화를 포착하는 비접촉 방식의 생체 정보 추정 기술이다. 2000년부터 근적외선 조명과 흑백 카메라를 이용한 rPPG 연구가 시작되었으며, 2007년에 주변광과 RGB 카메라를 통해 PPG 신

호를 감지할 수 있음을 발견하였다 [3]. 이를 통해 물리적 접촉으로 인한 한계점들을 보완하고, 피부가 민감한 환자의 경우에도 지속적인 모니터링의 가능성에 대해 연구되었다. rPPG를 통해 얻은 신호로 추정 가능하다고 알려진 건강 지표로는 심박수, 호흡수, 산소포화도, 그리고 혈압이 있다. 특히 심박수와 호흡수에 대한 연구는 활발히 진행되었으며, 얼굴 영역별 및 신체 부위별 성능 비교 분석의 결과가 존재한다 [4, 5]. 하지만 산소포화도는 건강한 사람 기준 그 값이 거의 비슷하여, 심박수 및 호흡수에 비해 다양한 데이터 확보가 쉽지 않아 검증이 어려워 아직도 많은 도전 과제가 남아 있다. 또한, 기존의 연구들은 얼굴, 뺨, 이마, 손목 등의 다양한 부위를 사용하였지만, 이에 대한 얼굴 영역별 성능 비교 분석의 연구는 상대적으로 부족하다 [6]. 특히, 생체 지표 추정을 위해 선행되어야 하는 rPPG 신호 추출에 대한 영역별 성능 비교 분석 연구는 조명 변화와 피험자의 앞머리 및 수염 유무와 같은 변수에 의한 영향을 최소화하였다는 한계점이 존재한다. 이러한 변동성을 보완하기 위해서 본 논문은 rPPG 기반 산소포화도 추정을 위한 공공 데이터셋인 PURE 데이터 셋을 사용하여 얼굴 영역별 산소포화도 추정 성능을 평가한다. 또한 피험자의 특징을 비교하고, 이를 바탕으로 결과를 분석한다.

II. 배 경

1. rPPG 기반 산소포화도 추정

산소포화도는 혈액 내 산소화된 헤모글로빈의 비율을 나타낸다. 이는 식 (1)과 같이 표현한다. HbO_2 는 산소화 헤모글로빈, Hb 는 탈산소화 헤모글로빈을 의미한다.

*Corresponding Author (namkyu@yu.ac.kr)

Received: Jun. 28, 2024, Revised: Jul. 31, 2024, Accepted: Aug. 12, 2024.

N.K. Kwon: Yeungnam University (Prof.)

H.J. Jo: Yeungnam University (B.S. Student)

* 본 논문은 2023년도 정부 (과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NORS-2023-00219725).

$$SpO_2 = \frac{HbO_2}{HbO_2 + Hb} \times 100\%. \quad (1)$$

일반적으로 rPPG 기반 산소포화도 추정은 Ratio of Ratios (RoR) 방식을 사용하며, 이는 두 가지 서로 다른 파장에서 측정된 신호들을 필요로 한다 [6]. 해당 신호들은 카메라를 통해 사람의 신체 피부가 드러나는 부위 일부를 촬영하고, 피부 영역을 검출하는 과정을 거쳐 추출된다.

산소포화도 측정을 위해 사용하는 신체 부위로, 얼굴에 대한 영역을 주로 사용하며, 그 외에 손바닥, 손목 부위를 사용하기도 한다 [7]. 얼굴을 감지한 후 region of interest (ROI)로 가장 많이 사용되는 영역은 얼굴 전체를 포함하는 것이며, 다음으로 뺨과 이마, 마지막으로 입술이 활용된다 [6]. 최근에는 주변 배경 및 얼굴을 가리는 물체 제거에 효과적인 segment anything model (SAM)을 통해 분할한 얼굴 영역을 사용하는 방법이 주목받고 있다 [3, 8].

기존의 접촉식 PPG 센서는 Beer-Lambert 법칙과 광 반사 이론을 기반으로 하며, 두 가지 파장을 사용하여 최대 투과 강도 (I_{DC, λ_1})와 반사 강도 (I_{DC, λ_2})를 측정한다. 이러한 두 파장 대역의 신호를 처리하여 활용하는 RoR 방식은 두 가지 원칙을 따른다. 한 파장 (λ_1)에서 HbO_2 와 Hb 의 흡수 계수 차이가 크게 나야 하고, 다른 파장 (λ_2)에서는 HbO_2 와 Hb 의 흡수 계수가 거의 동일해야 한다는 점이다.

일반적으로 λ_1 은 가시광선의 660 nm 대역을, λ_2 는 적외선의 940 nm 대역을 사용한다 [10]. RGB 카메라를 통해 rPPG 방법으로 산소포화도를 측정하는 경우, λ_1 에는 Red 채널을, λ_2 에는 Green 또는 Blue 채널을 사용한다 [11, 12]. 하지만 Blue 채널의 경우 그림 1에서 확인할 수 있듯이 해당 대역의 변화가 급격하고, 신호 대 잡음비 (signal-to-noise ratio, SNR)가 낮다는 결과를 보인다 [10, 13]. 따라서 본 논문에서는 λ_1 은 Red, λ_2 는 Green 채널을 사용한다.

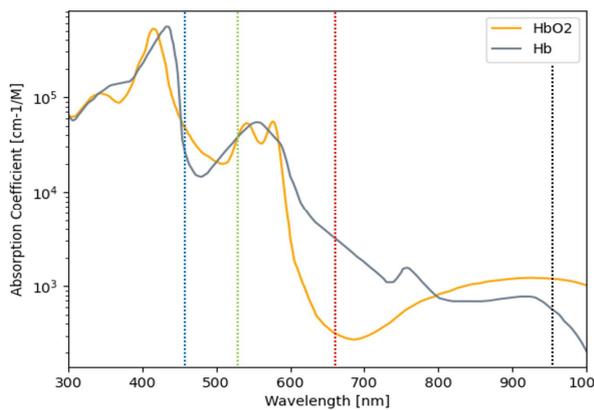


그림 1. 파장에 따른 HbO_2 , Hb 의 흡수 계수

(파란 점선: Blue 채널 대역, 초록 점선: Green 채널 대역, 붉은 점선: Red 채널 대역, 검정 점선: 적외선 대역) [9]

Fig. 1. Absorption Spectra of HbO_2 and Hb (Blue dotted line: Blue channel wavelength band, Green dotted line: Green channel wavelength band, Red dotted line: Red channel wavelength band, Black dotted line: Infrared wavelength band)

$$SpO_2 = A - B \frac{(I_{AC}/I_{DC})_{\lambda_1}}{(I_{AC}/I_{DC})_{\lambda_2}}. \quad (2)$$

최종적으로 두 파장의 신호를 추출하고, 각 신호의 DC 및 AC 성분을 구하여 식 (2)를 통해 산소포화도를 추정한다. DC 성분은 정적인 신호로서 맥동하지 않는 혈액량을 의미하고, AC 성분은 맥박에 의해 유발된 발생하는 혈류량 변화를 반영한다 [13]. 일반적으로 추출한 신호를 전처리한 후 연속적으로 특정 구간을 지정하여 각 구간에 대한 신호의 평균을 DC 성분으로, 표준편차를 AC 성분으로 나타낸다. 이때 A 와 B 는 실험적으로 구하는 보정 인자이다 [10].

산소포화도 값을 산출하기 전에 추출한 신호에 대하여 전처리 과정을 수행한다. 전처리 기술에는 피험자의 움직임과 조명 변화로 인해 발생하는 노이즈 성분인 신호의 추세를 제거하기 위한 Detrending, 서로 다른 피험자들을 대상으로 데이터를 표준 범위로 스케일링하는 Normalization, 생체 정보와 관련된 주파수 성분만 통과시키는 Band-pass Filtering 등이 존재한다 [14]. 이후 산소포화도 계산의 정확도를 높이기 위해 노이즈 아티팩트를 제거하는 Ensemble Empirical Mode Decomposition (EEMD)와 맥박 성분을 분리하는 Independent Component Analysis (ICA) 등의 알고리즘이 사용된다 [12, 15]. 산소포화도 추정을 위해서는 두 가지 파장 대역의 신호가 필요하기 때문에, 기존 심박수 추정에서 사용되는 조명 변화에 강한 색공간 변환을 수행하는 Chrominance (CHROM) 또는 움직임에 강한 색공간 변환을 수행하는 Plane-Orthogonal-to-Skin (POS) 같은 알고리즘을 적용하는 것에는 어려움이 있다 [14, 15]. 최근에는 딥러닝 기술을 적용하여 산소포화도 추정을 하는 연구가 활발히 진행되고 있다 [8, 16].

III. 본 론

1. 사용한 산소포화도 측정 방법

산소포화도 추정을 위해 사용한 방법 및 과정은 그림 2와 같다. 먼저, 데이터 셋의 얼굴 이미지에서 SAM 알고리즘을 통해 얼굴 영역을 분할한 후, 해당 얼굴 영역에서 사전에

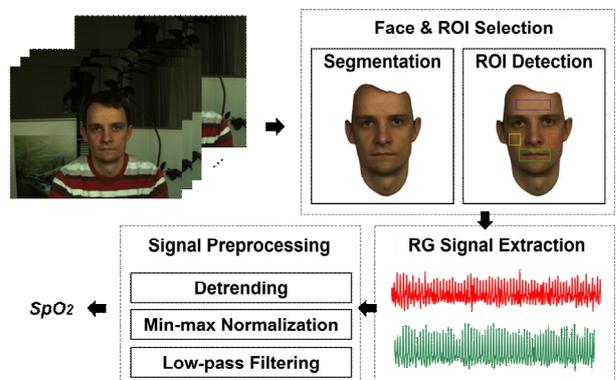


그림 2. 산소포화도 추정 framework

Fig. 2. Oxygen saturation estimation framework

표 1. 피험자 특징 요약

Table 1. Summary of Subject Characteristics

	Sex	Bangs	Beard
subject 1	Male	N	N
subject 2	Male	N	N
subject 3	Male	N	Y
subject 4	Male	N	Y
subject 5	Male	Y	Y
subject 6	Female	Y	N
subject 7	Female	Y	N
subject 8	Male	N	N
subject 9	Male	Y	Y
subject 10	Male	N	N

정 의한 이마, 뺨, 입술 등의 ROI를 검출한다 [17]. 이후 각 영역에 대하여 Red와 Green 신호를 추출하고, Detrending, Min-max normalization, 그리고 Low-pass filtering 과정의 전처리를 수행한다. 최종적으로 얻은 신호에 대하여 Sliding window를 적용하고, 각 Window에 대한 DC와 AC 값을 계산한다. 이를 바탕으로 산소포화도 값을 산출하고, mean absolute error (MAE)를 통해 성능을 비교한다.

2. 데이터 셋

총 10명 (남자 8명, 여자 2명) 의 피험자로 구성되어 있는 PURE 데이터 셋을 사용한다 [18, 19]. 피험자의 상태에 따라 서로 다른 6가지 시나리오가 존재하며, 본 논문에서는 피험자가 움직이지 않는 정적 상태의 시나리오 데이터만을 사용한다. 표 1은 사용한 데이터 셋의 성별, 앞머리 유무, 수염 유무와 같은 피험자의 특징을 요약한 것이다. 앞머리가 이마 ROI 영역을 침범하는 경우 앞머리가 존재한다고 판단하였다.

3. 신호 처리 과정

3.1 얼굴 영역별 검출

얼굴 전체, 이마, 왼쪽 뺨, 오른쪽 뺨, 입술로 총 5가지의 경우를 설정한다. 얼굴 전체는 SAM을 통해 분할한 후, 그 외 영역들은 dlib 라이브러리를 이용하여 그림 3의 68개의 face landmark를 통해 그림 4와 같이 검출하였다. 피험자에 따라 이마, 뺨, 입술의 크기는 다를 수 있으므로, 이를 고려하여 ROI 크기는 각 프레임마다 인식하는 face landmark의 위치에 따라 설정된다.

왼쪽 및 오른쪽 뺨의 ROI의 좌표는 식 (4)-(7), (9)를 통해 구한 값들을 기반으로 식 (8)과 같이 정의한다 [20]. 이때 왼쪽, 오른쪽은 피험자의 시점을 기준으로 한다. 이마는 식 (10), 입술은 식 (12)와 같이 구할 수 있다. 추가로, m_n 은 그림 3의 n 번째 landmark의 좌표를 의미하며, 식 (3)과 같이 표현한다.

$$m_n = [m_{n,x}, m_{n,y}] \in \mathbb{R}^2. \tag{3}$$

수식의 x 축과 y 축의 방향은 그림 4의 좌측 상단에 표기 한 축을 따른다.

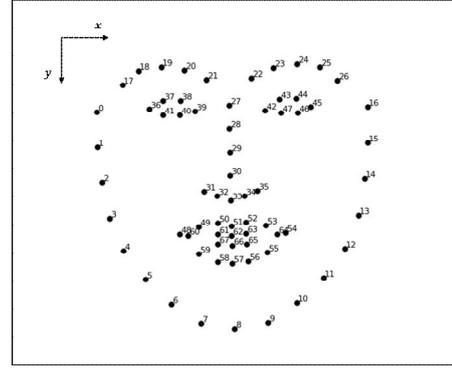


그림 3. 68개의 face landmark [21]

Fig. 3. 68-face landmarks

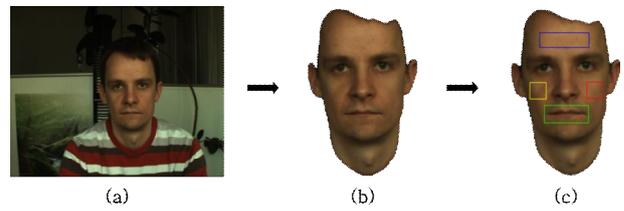


그림 4. 얼굴 영역별 검출 절차. (a) 원본 이미지 (프레임).

(b) SAM을 통해 분할한 얼굴 영역. (c) ROI 검출.

Fig. 4. Detection Procedures by Face Area.

(a) Original Image. (b) Face Area Segmented via SAM.

(c) ROI Detection.

오른쪽 눈의 중앙 좌표는 ER_c 로 나타내며, 다음과 같이 정의한다.

$$ER_c = \frac{m_{36} + m_{39}}{2}. \tag{4}$$

여기서 m_{36} 과 m_{39} 은 오른쪽 눈의 외안각과 내안각의 좌표이다. 이와 비슷하게 왼쪽 눈의 중앙 좌표는 EL_c 로 정의하며, 왼쪽 눈의 내안각 (m_{42}) 및 외안각 (m_{45}) 위치의 평균으로 수식 (5)와 같이 구할 수 있다.

$$EL_c = \frac{m_{42} + m_{45}}{2}. \tag{5}$$

오른쪽 뺨과 왼쪽 뺨의 중앙 좌표는 각각 CR_c 와 CL_c 로 나타내며 다음과 같이 정의한다.

$$\begin{aligned} CR_c &= \left(m_{48,x} - X, \frac{ER_{c,y} + Y}{2} \right), \\ CL_c &= \left(m_{54,x} + X, \frac{EL_{c,y} + Y}{2} \right). \end{aligned} \tag{6}$$

여기서 m_{48} 과 m_{54} 는 오른쪽 입꼬리와 왼쪽 입꼬리의 좌표이다. 또한, X 와 Y 는 뺨 ROI의 중앙점을 찾기 위해 활용되는 인수이며, 식 (7)을 통해 구할 수 있다. 각 뺨의 x 좌표는 각 입꼬리의 x 축의 위치에서 X 만큼 뺨 쪽 방향으로 이동한 위치이며, y 좌표는 눈의 중앙과 입꼬리의 높이에 대한 y 축의 중앙 위치를 의미한다.

$$\begin{aligned} X &= \frac{m_{54,x} - m_{48,x}}{3}, \\ Y &= \frac{m_{54,y} + m_{48,y}}{2}. \end{aligned} \quad (7)$$

X 는 입술의 가로 길이의 1/3 값이며, Y 는 입꼬리의 y 축 좌표의 평균값을 의미한다. 최종적으로 양쪽 뺨의 ROI 좌표는 식 (8)과 같이 나타낸다. ROI의 시작점은 bounding box의 좌측 상단의 꼭짓점을, 도착점은 우측 하단에 위치한 꼭짓점을 말한다. 오른쪽 (왼쪽) 뺨 ROI의 시작점 좌표는 CR_{\min} (or CL_{\min})이며, 오른쪽 (왼쪽) 뺨 ROI의 도착점 좌표는 CR_{\max} (CL_{\max})로 나타낸다.

$$\begin{aligned} CR_{\min} &= (CR_{c,x} - l, CR_{c,y} - l), \\ CR_{\max} &= (CR_{c,x} + l, CR_{c,y} + l), \\ CL_{\min} &= (CL_{c,x} - l, CL_{c,y} - l), \\ CL_{\max} &= (CL_{c,x} + l, CL_{c,y} + l). \end{aligned} \quad (8)$$

여기서 l 은 양쪽 뺨 ROI 크기를 정하기 위해 사용되는 값이며, 식 (9)와 같이 구할 수 있다.

$$l = \frac{EL_{c,x} - ER_{c,x}}{5}. \quad (9)$$

본 논문에서는 왼쪽 눈과 오른쪽 눈 사이의 거리의 1/5 값을 사용하였다. 이마 ROI의 좌표는 식 (10)과 같이 정의하며, 이마 ROI의 시작점과 도착점 좌표는 각각 F_{\min} 과 F_{\max} 로 표현한다.

$$\begin{aligned} F_{\min,x} &= m_{19,x}, \\ F_{\min,y} &= m_{19,y} - h - 10, \\ F_{\max,x} &= m_{24,x}, \\ F_{\max,y} &= m_{24,y} - 10. \end{aligned} \quad (10)$$

이마 ROI의 가로 길이는 오른쪽 눈썹산 좌표 (m_{19})와 왼쪽 눈썹산 좌표 (m_{24}) 사이의 거리를 사용한다. 이때 ROI가 눈썹을 포함하는 것을 방지하도록 y 축에 대해 10의 거리를 감산하여 설정하였다. 이때 h 는 식 (11)과 같이 나타낸다.

$$h = m_{22,x} - m_{21,x}. \quad (11)$$

h 는 미간 사이의 길이이며, 이마 ROI의 세로 길이로 사용된다. 입술 ROI의 시작점과 도착점 좌표는 L_{\min} 과 L_{\max} 로 표현한다.

$$\begin{aligned} L_{\min,x} &= m_{48,x} - M, \\ L_{\min,y} &= m_{51,y} - M, \\ L_{\max,x} &= m_{54,x} + M, \\ L_{\max,y} &= m_{57,y} + M. \end{aligned} \quad (12)$$

입술 ROI는 위쪽 입술의 중앙점 (m_{51}), 아래쪽 입술의 중앙점 (m_{33}), 오른쪽 입꼬리 (m_{48}), 그리고 왼쪽 입꼬리 (m_{54})에 대하여 크기 M 만큼의 간격씩 거리를 둔 bounding box의 크기를 사용한다. 여기서 M 은 식 (13)과 같이 정의한다.

$$M = \frac{m_{51,y} - m_{33,y}}{2}. \quad (13)$$

M 은 위쪽 입술의 중앙점 (m_{51})과 코끝점 (m_{33}) 간격의 1/2 값을 뜻한다.

3.2 신호 추출 및 전처리

분할한 얼굴 영역과 검출한 ROI에 대하여 RG 신호를 추출한다. RG 신호는 시간에 따라 각 프레임의 채널에 대한 평균값을 구하는 Spatial Averaging을 적용하여 추출한다. 이후 신호의 경향성 제거를 위한 Detrending 단계를 거쳐야 하며, 과정은 다음과 같다 [22].

$$D_2 = \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}. \quad (14)$$

Detrending 전의 RG 신호는 $(N-1)$ 의 길이를 가지는 $s(t)$ 이며, 이후의 신호를 $s_d(t)$ 라 정의한다. 이때 smoothing 파라미터 λ 는 10으로 설정하였다. I 는 $(N-1) \times (N-1)$ 크기를 가지는 희소 단위 행렬이다. D_2 는 식 (14)와 같으며, $(N-3) \times (N-1)$ 의 크기를 가진다.

$$s_d(t) = (I - (I + \lambda^2 D_2^T D_2)^{-1})s(t). \quad (15)$$

최종적으로 식 (15)을 통해 신호를 Detrending 한다. 사용한 방식의 Detrending 결과는 그림 5를 통해 확인할 수 있다. 기존의 다항 회귀를 통해 Detrending 하는 경우, 그림 5의 (b)와 같이 신호의 추세를 완벽히 제거하지 못하는 것을 볼 수 있다.

다음으로 식 (16)과 같이 R, G 각각의 신호에 대하여 Min-max Normalization을 수행한다.

$$\frac{s_d(t) - \min(s_d(t))}{\max(s_d(t)) - \min(s_d(t))}. \quad (16)$$

또한 2 Hz 이하의 주파수 대역만을 통과하도록 Low-pass Filtering을 적용하여 혈류량 변화와 관련된 성분만 남도록 한다.

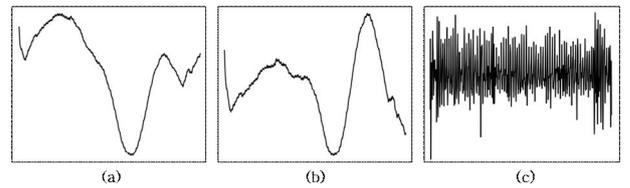


그림 5. Detrending 방식 비교 (a) 추출한 Red 신호.

(b) 다항 회귀 기반 Detrending 결과.

(c) 사용한 Detrending 방식 결과.

Fig. 5. Comparison of Detrending Methods.

(a) Extracted Red Signal.

(b) Detrending Results by Polynomial Regression Method.

(c) Detrending Results by the Employed Method.

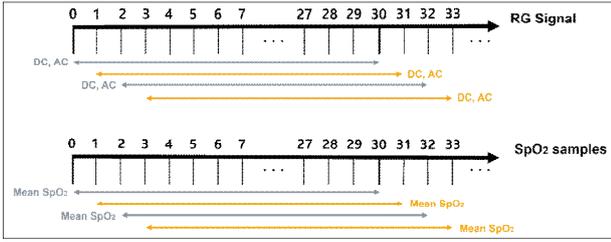


그림 6. Sliding Window 방법 예시 (window size : 30)
Fig. 6. The Sliding Window Method Used

4. 산소포화도 추정

RoR 방식 사용을 위해 Sliding window를 통해 그림 6과 같이 각 구간의 평균 (I_{DC})과 표준편차 (I_{AC})를 구한다. 또한 ground-truth로 사용하는 연속적인 산소포화도 값에도 동일하게 적용한다. PURE 데이터 셋에서 사용한 카메라는 30 fps를 가진다. 따라서 Window의 크기는 1초 동안의 프레임의 정보를 담도록 30으로 설정하였다.

$$SpO_2 = 108 - 15R. \tag{17}$$

반복적인 실험을 통해 식 (17)과 같이 A 는 108, B 는 15로 설정하였다. 또한 산소포화도 출력 범위는 70~100%로 설정하였으며, 이는 실제 호흡기 질환 및 저산소증을 가지고 있는 환자에 대한 적용을 실현하기 위함이다 [23].

IV. 실험 결과

표 2는 얼굴 영역별 산소포화도 추정 성능을 비교한 것이다. 평균 MAE를 기준으로 한 결과, 왼쪽 뺨을 사용했을 때 가장 성능이 높은 것을 확인할 수 있다. 다음으로 오른쪽 뺨, 얼굴 전체, 입술 순으로 좋은 성능을 보였으며, 산소포화도 추정에서 두 번째로 많이 사용하는 영역 중 하나인 이마를 사용했을 때 가장 낮은 성능을 보였다.

표 2. 피험자 및 얼굴 영역별 산소포화도 추정 성능
Table 2. Oxygen Saturation Estimation Performance by Subject and Face Area

	Face	Forehead	Left cheek	Right cheek	Lips
subject 1	3.743	2.696	1.936	1.812	3.986
subject 2	5.856	5.994	2.279	1.556	2.715
subject 3	2.849	3.211	1.696	2.173	2.286
subject 4	2.323	1.783	1.935	2.214	5.921
subject 5	2.705	9.039	3.216	4.995	3.715
subject 6	2.726	4.272	2.900	2.622	3.216
subject 7	2.930	5.080	2.139	3.686	3.412
subject 8	1.954	1.170	4.173	4.775	3.252
subject 9	2.440	4.066	2.092	1.721	2.973
subject 10	4.540	4.728	1.699	2.061	1.956
Average	3.207	4.204	2.407	2.762	3.343

PURE 데이터 셋의 조명 조건은 시간에 따라 바뀌는 조도를 모두 반영한다는 것과 피험자의 움직임을 최대한 제한했다는 것을 비추어 볼 때, 이마는 조명의 변화에 취약한 영역임을 유추할 수 있다. 이마가 가장 좋은 성능을 보였던 subject 4와 8의 경우 거의 일정하게 어두운 밝기가 유지되는 것을 확인하였다.

V. 결론

본 논문에서는 얼굴 영역별 산소포화도 추정 성능을 비교하였다. 산소포화도 추정을 위하여 연속적인 얼굴 이미지에서 Red 및 Green 신호의 정보를 추출하고, Detrending 및 Min-max normalization과 Low-pass filtering의 간단한 전처리를 적용하였다. 이를 통해 얻은 신호에 대하여 Sliding Window를 수행하고, RoR 방식을 사용하여 산소포화도를 추정하였다. 최종적으로 추정된 산소포화도와 실제 접촉식 센서로 측정된 산소포화도를 비교하여 각 피험자에 대한 얼굴 영역별 MAE를 평가 지표로서 사용하였다. 또한, 피험자 및 데이터의 특징을 바탕으로 결과를 분석하고, 원인을 도출하였다. 결론적으로 피험자에 따라 가장 좋은 성능의 영역이 다르게 나타났지만, 평균적으로 뺨을 사용했을 때 가장 좋은 성능을 나타낼 수 있음을 보여준다. 또한 기존에 주로 사용하던 이마는 조명 변화의 영향을 많이 받으며, 산소포화도 추정에 적절하지 않음을 알 수 있다.

기존의 영역별 성능 비교 분석 연구는 조명 변화와 피험자의 앞머리 및 수염 유무와 같은 변수의 영향을 최소화하여 비교하는 것에 중점을 두었다. 또한 심박수 및 호흡수 추정을 목적으로 하는 연구에 비해 산소포화도 추정을 목적으로 하는 연구는 상대적으로 부족하였다. 따라서 본 논문에서는 이러한 한계점들을 인식하고, 다양한 변수를 가진 데이터에서 적용될 수 있는 산소포화도 추정 성능 지표를 제안하였다. 향후 연구 방향은 본 연구를 통해 가장 좋은 성능을 보였던 왼쪽 뺨 영역을 활용하여 rPPG 기반 산소포화도 측정의 정확도를 향상하고자 한다. 또한, 다양한 신호처리 기법과 딥러닝 기술을 활용하여 실생활에 적용이 가능한 산소포화도 추정 방법을 연구할 예정이다.

References

[1] A. d. F. G. Rosa, R. C. Betini, "Noncontact SpO₂ Measurement Using Eulerian Video Magnification," IEEE Transactions on Instrumentation and Measurement, Vol. 69, No. 5, pp. 2120-2130, 2019.

[2] H. Xiao, T. Liu, Y. Sun, Y. Li, S. Zhao, A. Avolio, "Remote Photoplethysmography for Heart Rate Measurement: A Review," Biomedical Signal Processing and Control, Vol. 88, pp. 105608, 2024.

[3] B. Huang, S. Hu, Z. Liu, C. L. Lin, J. Su, C. Zhao, L. Wang, W. Wang, "Challenges and Prospects of Visual Contactless

- Physiological Monitoring in Clinical Study," NPJ Digital Medicine, Vol. 6, No. 1, pp. 231, 2023.
- [4] L. Niu, J. Speth, N. Vance, B. Sporrer, A. Czajka, P. Flynn, "Full-body Cardiovascular Sensing with Remote Photoplethysmography," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5994-6004, 2023.
- [5] S. Kwon, J. Kim, D. Lee, K. Park, "ROI Analysis for Remote Photoplethysmography on Facial Video," in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015: IEEE, pp. 4938-4941.
- [6] A. Gupta, A. G. Ravelo-Garcia, F. M. Dias, "Availability and Performance of Face Based Non-contact Methods for Heart Rate and Oxygen Saturation Estimations: A Systematic Review," Computer Methods and Programs in Biomedicine, Vol. 219, pp. 106771, 2022.
- [7] Q. Fan, K. Li, "Noncontact Imaging Plethysmography for Accurate Estimation of Physiological Parameters," Journal of Medical and Biological Engineering, Vol. 37, pp. 675-685, 2017.
- [8] X. Sun, T. Wen, W. Chen, B. Huang, "CCSpO 2 Net: Camera-Based Contactless Oxygen Saturation Measurement Foundation Model in Clinical Settings," IEEE Transactions on Instrumentation and Measurement, 2024.
- [9] <https://omlc.org/spectra/hemoglobin/>
- [10] L. Kong, Y. Zhao, L. Dong, Y. Jian, X. Jin, B. Li, Y. Feng, M. Liu, X. Liu, H. Wu, "Non-contact Detection of Oxygen Saturation Based on Visible Light Imaging Device Using Ambient Light," Optics Express, Vol. 21, No. 15, pp. 17464-17471, 2013.
- [11] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton, C. Pugh, "Non-contact Video-based Vital Sign Monitoring Using Ambient Light and Auto-regressive Models," Physiological Measurement, Vol. 35, No. 5, pp. 807, 2014.
- [12] A. Al-Naji, G. A. Khalid, J. F. Mahdi, J. Chahl, "Non-contact SpO2 Prediction System Based on a Digital Camera," Applied Sciences, Vol. 11, No. 9, pp. 4255, 2021.
- [13] B. Wei, X. Wu, C. Zhang, Z. Lv, "Analysis and Improvement of Non-contact SpO2 Extraction Using an RGB Webcam," Biomedical Optics Express, Vol. 12, No. 8, pp. 5227-5245, 2021.
- [14] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, J. Dubois, "Unsupervised Skin Tissue Segmentation for Remote Photoplethysmography," Pattern Recognition Letters, Vol. 124, pp. 82-90, 2019.
- [15] A. Lam, Y. Kuno, "Robust Heart Rate Measurement from Video Using Select Random Patches," in Proceedings of the IEEE International Conference on Computer Vision, pp. 3640-3648, 2015.
- [16] X. Ding, D. Nassehi, E. C. Larson, "Measuring Oxygen Saturation with Smartphone Cameras Using Convolutional Neural Networks," IEEE Journal of Biomedical and Health Informatics, Vol. 23, No. 6, pp. 2603-2610, 2018.
- [17] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W. Y. Lo, P. Dollar, R. Girshick, "Segment Anything," in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 4015-4026, 2023.
- [18] R. Stricker, S. Müller, H. M. Gross, "Non-contact Video-based Pulse Rate Measurement on a Mobile Service Robot," in The 23rd IEEE International Symposium on Robot and Human Interactive Communication, 2014: IEEE, pp. 1056-1062.
- [19] <http://www.tu-ilmenau.de/neurob/data-sets/pulse>
- [20] S. R. Lee, S. E. Lee, S. H. Jo, S. H. Kim, "Region-of-Interest Extraction Optimization for Improving the Accuracy of Non-contact Heart Rate Measurement by Using Two-dimensional Face Images," Journal of Institute of Control, Robotics and Systems, Vol. 29, No. 3, pp. 272-278, 2023, doi: 10.5302/J.ICROS.2023.22.0167
- [21] B. Amos, B. Ludwiczuk, M. Satyanarayanan, "Openface: A General-purpose Face Recognition Library with Mobile Applications," CMU School of Computer Science, Vol. 6, No. 2, pp. 20, 2016.
- [22] M. P. Tarvainen, P. O. Ranta-Aho, P. A. Karjalainen, "An Advanced Detrending Method with Application to HRV Analysis," IEEE Transactions on Biomedical Engineering, Vol. 49, No. 2, pp. 172-175, 2002.
- [23] J. S. Hoffman, V. K. Viswanath, C. Tian, X. Ding, M. J. Thompson, E. C. Larson, S. N. Patel, E. J. Wang, "Smartphone Camera Oximetry in an Induced Hypoxemia Study," NPJ Digital Medicine, Vol. 5, No. 1, pp. 146, 2022.

Hyojin Jo (조 효진)

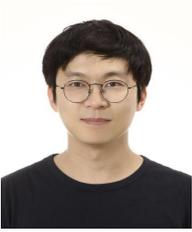


2021~Department of Electronic Engineering
from Yeungnam University (B.S.)

Field of Interests: Biosignal Processing, Artificial Intelligence

Email: hynn0254@gmail.com

Nam Kyu Kwon (권 남 규)



2010 Department of Electrical Engineering from
Pohang University of Science and Technology
(B.S.)

2017 Department of Electrical Engineering from
Pohang University of Science and Technology
(Ph.D.)

Career:

2017~2019 Senior Researcher, Samsung Electronics

2019~ Assistant Professor, Department of Electronic
Engineering, Yeungnam University

Field of Interests: Robust Control, Reinforcement Learning,
Vital Signal Processing

Email: namkyu@yu.ac.kr