ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

A Standards-First Adaptive Quality-Gated Fusion Method for Multi-Source Physiological Signals in Wearable Devices

Shuowei Cao [1,a], Yaqi Gen [2,b], Shun Liu [1,c], ,Lin Wang [1,d,*], and Geng Liu [2,e]

[1] School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China.

[2] Weiqiao-UCAS (Binzhou) Science and Technology Co., Ltd., Binzhou, Shandong, 256606, China;

[a] 17686941828@163.com, [b] gengyq485@wqmail.cn , [c] 15835024985@163.com.

[d] linwang@nwpu.edu.cn, [e] npuliug@nwpu.edu.cn

Abstract. Wearable multi-source physiological monitoring is vulnerable to time-varying motion artifacts, poor contact, and optical interference. Fixed-weight fusion may therefore become less accurate than a reliable single-channel estimate. This paper proposes a standards-first adaptive quality-gated fusion method in which heterogeneous ECG, PPG, temperature, and acceleration observations are first mapped to a common task-estimation layer. Each sensor registers a quality specification bound to a contract hash and configuration epoch. A fast motion pre-check and sliding-window signal quality indices (SQIs) generate quality states that are attached to data frames and drive hard gating, reliability weighting, and four-state switching. Mechanism-level simulations using MIT-BIH/NSTDB data and rule-based stress scenarios show that heart-rate MAE decreases from 31.61 to 17.47 bpm in the running stress case (44.7% improvement). Hardware motion pre-check latency is 1.8 ms, whereas software SQI processing requires 6.7-7.8 ms after a window is ready; the update interval is governed by the 2.5-s window step. Timeline reconstruction reduces the clock-drift contribution to PTT timing error from 29.7 to 0.1 ms at ±200 ppm drift. A 10,000-frame software consistency test detects all injected contract, epoch, sequence, and CRC violations without false alarms.

Keywords: wearable devices; multi-source physiological signals; signal quality index; adaptive fusion; quality gating

1. Introduction

Wearable health-monitoring systems increasingly integrate electrocardiography (ECG), photoplethysmography (PPG), temperature sensing, and inertial measurements. Although these modalities are complementary, signal quality changes markedly during motion, perspiration, strap loosening, and illumination variation. Motion artifacts, poor contact, saturation, and optical interference can therefore make a nominally redundant sensor harmful to fusion rather than helpful.

Signal-quality research provides the basis for quality-aware fusion. Multi-index ECG quality assessment has been studied in [1,2], and wearable motion-artifact handling is reviewed in [3]. Recent work on wearable multi-sensor fusion, embedded architecture, and remote physiological monitoring emphasizes dynamic reliability and synchronization [4-6]. PPG quality and physiological measurement are addressed in [7-10]; wearable digital-biomarker and activity-monitoring systems are reported in [11,12]; motion-robust wrist PPG processing is developed in [13,14]; and nonlinear multi-sensor estimation is supported by unscented filtering theory [15].

Most existing quality-assessment studies focus on determining whether a single signal segment is acceptable for interpretation, while fusion studies usually assume that all available inputs can be combined once basic preprocessing has been completed. These two problems are related but not identical. A signal can remain partially informative without being sufficiently reliable to receive the same fusion weight as a clean channel. In addition, different wearable modalities fail in different ways: ECG is strongly affected by electrode motion and contact impedance, PPG is sensitive to motion and optical coupling, temperature changes slowly and may lag rapid physiological events, and inertial sensing often acts as contextual evidence rather than as a direct source of the physiological parameter. Consequently, a practical multi-sensor system needs more than a binary good/bad label.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39
It needs a quality representation that can be interpreted consistently across acquisition, transport, fusion, storage, and replay. The proposed framework addresses this requirement by separating modality-specific quality assessment from task-level fusion. This separation also makes it possible to retain useful low-quality data for audit while preventing them from influencing an online estimate when their quality falls below the task threshold.

A degraded channel can retain enough weight to make the fused estimate worse than the best available single-channel estimate. We refer to this failure mode as catastrophic fusion in this paper. The broader wearable-fusion literature likewise emphasizes the need to account for time-varying sensor reliability [4]. Many existing pipelines assess quality only after acquisition or absorb reliability into a downstream model. Our approach moves quality information upstream so that each frame carries an explicit, traceable state that can directly control fusion.

The motivation for moving quality information upstream is also architectural. If quality is calculated only inside a host-side fusion algorithm, the meaning of a recorded sample depends on software state that may not be preserved with the data. Replaying the same stream after thresholds, firmware, or sensor settings have changed can then produce a different interpretation of the same measurements. By contrast, attaching a versioned quality contract and configuration epoch to the acquisition stream makes the assumptions used for quality evaluation explicit. The fusion layer can still evolve independently, but it receives a traceable statement of how each channel was assessed at acquisition time. This design is particularly useful for wearable prototypes in which sampling rates, gain settings, optical drive levels, or contact conditions may change frequently during development. The standards-first concept is therefore intended as an interface discipline between sensing and fusion rather than as a new physiological standard.

Here, “standards-first” means that a predefined, machine-verifiable sensor quality specification is established before fusion; the term does not imply compliance with a particular regulatory standard. This study makes four contributions: (1) a quality contract built around the sensor quality specification, contract_hash, and cfg_epoch; (2) a task-level fusion model that avoids direct arithmetic fusion of physically heterogeneous raw signals; (3) a composite SQI/motion score for hard gating, reliability weighting, confidence output, and four-state control; and (4) an STM32F103/Python reference implementation evaluated through quality trends, threshold transfer, runtime, fusion accuracy, ablation, timeline reconstruction, state behavior, confidence, and protocol-consistency tests.

2. System Architecture and Standards-First Quality Description

2.1 Embedded and Host Architecture

The system contains acquisition, transport, and processing layers. An STM32F103 (72 MHz) manages sensor drivers, a fast motion pre-check, window-level SQI computation, quality-state generation, and frame packing. The transport layer separates control commands from streaming data, while the Python/PyQt5 host parses frames, maintains task-relevant channels, performs quality-gated fusion, displays signals, and records data for replay. SQIs are updated per window; status_flags carry the latest quality state on subsequent frames.

Table 1. Sensor configuration and major degradation modes

Signal	Device	Sampling rate	Main degradation / role
ECG	AD8232	250 Hz	Motion artifact; poor contact
PPG	MAX30102	100 Hz	Optical interference; poor contact
Temperature	DS18B20	1 Hz	Loose wearing / slow response
Acceleration	MPU6050	50 Hz	Motion-quality auxiliary input

Sensor drivers follow a plug-in interface. Each outgoing frame carries the sample together with MOTION, BAD_CONTACT, SATURATED, and DEGRADED flags. The online host consumes these states directly; repeated SQI computation is reserved for offline audit.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

The division of responsibilities between the embedded node and the host is chosen to keep the real-time path deterministic while avoiding unnecessary duplication. Fast decisions that can prevent an obviously degraded sample from receiving normal fusion weight are placed close to the sensor, where transport delay cannot postpone the response. More computationally involved processing is organized as periodic window operations so that its execution time is bounded and easy to schedule. The host receives both the measurement and its latest quality state, allowing visualization and fusion to use the same interpretation without recalculating the online decision. This arrangement also supports heterogeneous update rates. A 250-Hz ECG stream and a 1-Hz temperature stream do not need identical processing schedules; each channel can maintain its own descriptor, quality update interval, and task role while sharing the same frame semantics. The architecture is therefore modular at two levels: drivers hide device-specific acquisition details, whereas descriptors and status flags expose a common interface to downstream processing.

2.2 Data Contract and Configuration Traceability

Each variable-length binary frame contains seq, t_master or block_start_t_master, sensor/channel identifiers, contract_hash, cfg_epoch, status_flags, payload length, and CRC. STREAM, PARAM, and EVENT frames carry waveforms, low-rate parameters, and system events; GET_DESC, SET_CFG, COMMIT_CFG, START, STOP, and SYNC_TIME form the control plane.

At registration, each sensor provides a quality specification containing supported states, decision rules, thresholds, normalization parameters, and a version. The normalized specification is hashed as

$$\text{contract_hash} = H(C_{norm}) \tag{1}$$

where H denotes SHA-256. When a configuration is committed, the active configuration epoch is bound to the current quality contract:

$$\text{cfg_epoch} \leftarrow \text{bind}(\text{cfg_epoch}, \text{contract_hash}) \tag{2}$$

The contract_hash/cfg_epoch pair supports semantic consistency, configuration-range isolation, and reproducible replay. The complete chain is still being integrated into the hardware data path, so Section 4 reports software-level consistency verification only and does not treat it as end-to-end hardware validation.

The two identifiers serve different purposes. The contract hash identifies the semantic definition of quality: the set of states, thresholds, normalization parameters, and decision rules that determine how a quality score should be interpreted. The configuration epoch identifies the interval over which a particular committed acquisition configuration is active. Keeping both fields avoids a common ambiguity in experimental logs. Two frames may have the same signal name and sampling rate but still have been evaluated under different quality thresholds, while two configurations may use the same quality contract but different device gains or sampling parameters. During replay, a change in cfg_epoch can be used to segment the stream, and a change in contract_hash can indicate that quality values before and after the change should not be compared without checking the corresponding specification. This mechanism does not improve signal quality by itself; its role is to make the provenance of the quality decision explicit and machine-verifiable. The software consistency test in Section 4 therefore evaluates detection of semantic and configuration mismatches rather than physiological accuracy.

3. Standards-First Adaptive Quality-Gated Fusion

3.1 Task-Level Observation and Catastrophic-Fusion Criterion

Let M heterogeneous sensing modalities observe a latent physiological state z(t):

$$x_i(t) = h_i\big(z(t)\big) + n_i(t) \tag{3}$$

where h_i denotes the modality-specific observation mapping and $n_i(t)$ represents motion, contact, and environmental interference. Because ECG, PPG, temperature, and acceleration have different

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39
units and physical meanings, their raw values are not directly summed. Each modality capable of estimating target y instead produces

$$\hat{y}_i(t) = g_i[x_i(t)] = y(t) + \varepsilon_i(t),\ i \in \Omega_y \tag{4}$$

where Ω_y is the task-relevant channel set. For heart rate, Ω_{HR} contains ECG and PPG; acceleration supplies motion-quality information and temperature does not enter HR numerical fusion. The fused estimate and weight constraints are

$$\hat{y}(t) = \sum_{i\in\Omega_y} w_i(t) \cdot \hat{y}_i(t) \tag{5}$$

$$w_i(t) \geq 0,\ \sum_{i\in\Omega_y} w_i(t) = 1 \tag{6}$$

With approximately uncorrelated task-estimation errors, the fused error variance is

$$\mathrm{Var}[e(t)] = \sum_{i\in\Omega_y} w_i^2(t) \cdot \sigma_i^2(t) \tag{7}$$

A fixed weight cannot immediately react to a sharp increase in $\sigma_i^2(t)$. We define catastrophic fusion as the condition in which fused MSE exceeds that of the best currently available task channel:

$$\mathrm{MSE}_{fusion}(t) > \min_{i\in\Omega_y} \mathrm{MSE}_i(t) \tag{8}$$

Quality gating excludes channels below T_{bad} and requires their weight to vanish as quality approaches zero:

$$w_i(t) = 0,\ \text{if } Q_i(t) < T_{bad} \tag{9}$$

$$\lim_{Q_i(t)\to 0} w_i(t) = 0 \tag{10}$$

Equations (8)-(10) define a testable failure condition and the corresponding gating requirement. Because the current logs do not include a window-by-window best-single baseline, the experiments quantify improvement over fixed weighting rather than claiming that Eq. (8) is avoided in every window.

The task-level formulation is important because it separates sensing diversity from estimation compatibility. ECG and PPG are both related to cardiac activity, but their raw amplitudes are not commensurate and should not be averaged directly. After each modality has been converted to an estimate of the same target, such as heart rate, the fusion problem becomes well defined: the estimator combines quantities with a common physical meaning while their modality-specific quality scores regulate trust. Under the uncorrelated-error approximation in Eq. (7), a channel with rapidly increasing error variance can dominate the fused variance if its weight is not reduced. Hard gating provides a conservative boundary by setting the contribution of a BAD channel to zero, whereas quality-dependent weighting handles intermediate degradation without forcing an abrupt transition for every small fluctuation. The method therefore combines a continuous mechanism, through Q and the reliability term, with a discrete safety mechanism, through Tbad and the state controller. This interpretation also clarifies the role of auxiliary sensors. Acceleration can influence the quality assigned to an ECG or PPG estimate without being treated as a heart-rate estimate itself.

3.2 Signal Quality Indices and Composite Quality Score

For a window $X = \{x_1, x_2, \ldots x_N\}$, kurtosis is used to detect impulsive artifacts:

$$\mathrm{kSQI} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^4}{\sigma^4} \tag{11}$$

Because kurtosis is not monotonic with quality, its normalized component is based on distance from a sensor-specific acceptable interval. Spectral concentration is measured by

$$\mathrm{pSQI} = \frac{\int_5^{20} P(f)df}{\int_0^{fs/2} P(f)df} \tag{12}$$

The present experiments use 5–20 Hz for ECG and 0.5–5 Hz for PPG. Baseline wander is quantified as

$$\mathrm{basSQI} = 1 - \frac{\int_0^1 P(f)df}{\int_0^{fs/2} P(f)df} \tag{13}$$

With 0–1 Hz as the low-frequency drift band, motion intensity from three-axis acceleration is

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

$$M = \frac{1}{N}\sum_{i=1}^{N} \left|\sqrt{a_{x,i}^2 + a_{y,i}^2 + a_{z,i}^2} - \bar{a}\right| \tag{14}$$

The motion term is converted to a normalized motion-quality valueM_i^q that decreases with motion intensity. The composite quality score is

$$Q_i = 0.30K_i + 0.30P_i + 0.20B_i + 0.20M_i^q \tag{15}$$

K_i, P_i, B_i and M_i^q are mapped to [0,1] using parameters stored in the quality specification. Monotonic indicators use clipped bad-to-good normalization, while kurtosis uses distance from an acceptable interval. Operating thresholds are Q≥0.80(VALID), 0.50≤Q<0.80(DEGRADED), and Q<0.50(BAD); they are configurable rather than universal physiological constants.

Using several SQIs is intended to reduce dependence on any single artifact signature. Kurtosis responds strongly to impulsive disturbances and waveform-shape changes, but an unusually high or low value is not universally equivalent to poor quality. Spectral concentration measures whether energy remains in the frequency region expected for the target waveform and is therefore sensitive to broadband contamination or loss of physiological periodicity. Baseline-wander power captures a different failure mechanism: slow variation caused by motion, electrode displacement, or coupling changes can preserve local peaks while still distorting the signal baseline. The acceleration-derived motion term supplies contextual evidence that is available before severe distortion is fully expressed in the physiological waveform. Because these indicators have different units and numerical ranges, their raw values are not combined directly. The quality specification stores the mapping from each indicator to a normalized score, so the composite Q has a consistent direction and range. The fixed coefficients in Eq. (15) should be understood as engineering weights used in the present prototype rather than population-level physiological constants. Their purpose in this study is to test the gating mechanism under a reproducible scoring rule; sensor-specific calibration remains necessary for deployment.

3.3 Reliability Weighting, Gating, and State Control

For simulations with known noise levels, dB SNR is converted to linear reliability $\rho_i = 10^{\mathrm{SNR}_{i,dB}/10}$:

$$w_i^{base} = \frac{\rho_i}{\sum_{j\in\Omega_y}\rho_j},\ \rho_i = 10^{\mathrm{SNR}_{i,dB}/10} \tag{16}$$

The proposed gated weight within Ω_y is

$$w_i = \frac{\mathbb{I}[Q_i \geq T_{bad}]\cdot Q_i\cdot\rho_i}{\sum_{j\in\Omega_y}\mathbb{I}[Q_j \geq T_{bad}]\cdot Q_j\cdot\rho_j} \tag{17}$$

In deployment, known SNR is unavailable, so ρ_i should be replaced by an online estimate such as inverse recent residual variance, $\rho_i = 1/(\sigma^{i^2} + \varepsilon)$. The task output is

$$\hat{y}(t) = \sum_{i\in\Omega_y} w_i(t)\cdot\hat{y}_i(t) \tag{18}$$

The number of valid task channels is

$$S(t) = \begin{cases} S_1\ (\text{full} - \text{quality}), N_{valid} = |\Omega_y| \\ S_2(\text{degraded}), 1 < N_{valid} < |\Omega_y| \\ S_3\ (\text{single} - \text{channel}), N\text{valid} \quad (t) = 1 \\ S_4(\text{unavailable}), N\text{valid} \quad (t) = 0 \end{cases} \tag{19}$$

Here, Nvalid denotes the number of channels in the task-relevant set Ωy whose quality score is at least Tvalid at time t. The controller uses S1 when all task channels are valid, S2 when more than one but not all are valid, S3 when exactly one remains, and S4 when none is valid. ACC is not counted as an HR output channel. State transitions are recorded as EVENT frames for replay and audit. Fusion confidence is defined by Eq. (20).

$$C(t) = \sum_{i\in\Omega_y} w_i(t)\cdot Q_i(t) \tag{20}$$

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

A larger C(t) indicates that a greater share of the fusion weight is assigned to high-quality channels.

The four controller states provide an explicit interpretation of what the numerical weights mean at each instant. S1 represents normal fusion when all task-relevant channels pass the validity threshold. S2 preserves redundancy but restricts the normalization to the channels that remain usable, preventing a rejected channel from receiving weight through renormalization. S3 represents a deliberate fall-back to a single estimate rather than forcing a nominal multi-sensor output when only one source remains credible. S4 marks the result as unavailable when no task channel meets the requirement. This distinction is useful for downstream applications because a numerical output alone cannot indicate whether it came from two agreeing sensors or from a single surviving channel. The confidence measure C(t) complements the discrete state by providing a continuous summary of the quality carried by the current weight distribution. In a practical implementation, hysteresis or a minimum dwell time can be added around the thresholds to suppress rapid state toggling; the present study leaves that extension for future work so that the effect of the basic gating logic remains observable.

3.4 Timing and Multi-Rate Timeline Reconstruction

The 50-Hz motion pre-check provides a rapid prior; its event-to-flag latency is

$$T_{hw} = t_{flag} - t_{motion} \tag{21}$$

Software SQIs are evaluated on 5-s windows with a 2.5-s step. Processing time is defined after a window becomes available:

$$T_{comp} = t_{Q,out} - t_{window,ready} \tag{22}$$

Thus, 6.7-7.8 ms denotes computation only, not end-to-end degradation-detection latency; the first decision requires approximately 5 s and steady-state waiting is dominated by the 2.5 s step.

ECG, PPG, temperature, and acceleration are sampled at 250, 100, 1, and 50 Hz. A common master clock and block_start_t_master reconstruct sample times. For PTT, the experiment isolates only clock-drift error; 100 Hz PPG still carries approximately ±5 ms nearest-neighbor quantization uncertainty, so a 0.1 ms reconstructed clock error is not total PTT accuracy.

This timing decomposition prevents two different notions of latency from being conflated. The motion pre-check reacts to a newly observed acceleration condition and can therefore provide a millisecond-scale warning. The SQI path, in contrast, evaluates statistics over a completed window; its computational cost may be only several milliseconds, but the quality state can change only when the window is updated. For this reason, the reported software timing is a processing benchmark rather than an end-to-end detection delay. Multi-rate alignment introduces a separate timing issue. Even when every channel has a correct local sampling interval, independent clocks can accumulate relative phase error over a long record. Recording a common block start time and reconstructing sample timestamps from the master time base limits this drift contribution and allows derived cross-signal quantities to be computed on a consistent axis. The PTT experiment isolates this synchronization effect so that it does not imply sub-millisecond physiological feature localization from a 100-Hz PPG waveform.

4. Experimental Validation

4.1 Data and Protocol

The experiments use ECG records 100, 101, 103, 105, and 106 from the MIT-BIH Arrhythmia Database together with baseline-wander (bw), muscle-artifact (ma), and electrode-motion (em) noise from the MIT-BIH Noise Stress Test Database (NSTDB). Noise is added at 0, 6, 12, and 18 dB. Rest, Walking, Running, Loose fit, and Optical interference are rule-based stress labels rather than participant activity labels. For joint-fusion simulation, a proxy PPG is synchronized to the reference heart-rate sequence and acceleration provides the motion-quality component. The experiments therefore test the fusion mechanism rather than clinical performance.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

ECG is high-pass filtered at 0.5 Hz before SQI computation; Welch PSD uses 1024-sample segments with 50% overlap. SQIs use 5-s windows and a 2.5-s step. For heart-rate estimation, ECG is band-pass filtered at 8-20 Hz and R peaks are detected with scipy.signal.find_peaks. Fixed reliability weighting is compared with the proposed quality-gated method.

The evaluation is structured as a mechanism test rather than as a clinical performance study. Controlled noise injection makes it possible to vary degradation strength while preserving a known clean ECG reference, which is useful for checking whether the selected SQIs respond in the expected direction. The rule-based stress labels then group combinations of signal-quality conditions into scenarios that exercise different parts of the controller. They should not be interpreted as measured biomechanics of a participant walking or running. Similarly, the proxy PPG is used only to create a second task-relevant heart-rate estimate under synchronized conditions; it does not substitute for a validation data set containing real simultaneous ECG and PPG waveforms. These design choices allow the study to examine quality gating, timing, state switching, and protocol traceability independently of subject variability. They also define the boundary of the conclusions: the reported MAE differences support the behavior of the proposed fusion mechanism under the constructed stress conditions, whereas clinical accuracy and generalization across devices or populations require additional synchronized wearable recordings.

4.2 Quality Response and Threshold Transfer

Table 2. SQIs and composite quality score versus SNR (Record 100 + muscle artifact)

SNR (dB)	kSQI	pSQI	basSQI	Q	State
18 (clean)	30.87	0.5945	0.9896	0.8763	VALID
12	28.97	0.5830	0.9752	0.8699	VALID
6	23.90	0.5456	0.9291	0.8490	VALID
0	15.60	0.4574	0.8228	0.7833	DEGRADED

In the Record 100 example, kSQI, pSQI, and basSQI decrease by 49.5%, 23.1%, and 16.9%, respectively, as SNR is reduced, confirming complementary sensitivity to degradation [2].

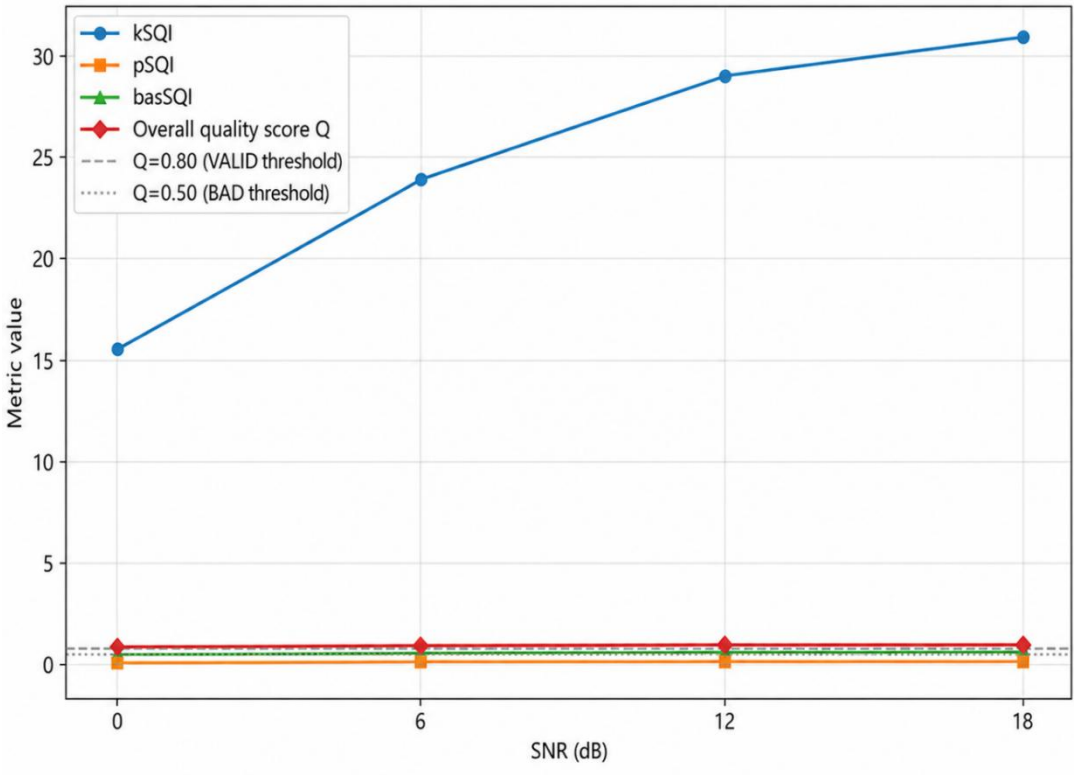


Fig. 1. SQI metrics and composite quality score versus SNR.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

Table 3. Fixed-threshold transfer results under rule-based stress scenarios

Scenario	Windows	Accuracy	Recall	F1
Rest	3605	89.29%	29.76%	31.45%
Walking	10815	48.67%	16.22%	21.83%
Running	10815	4.83%	1.61%	3.07%
Loose fit	14420	26.72%	34.60%	20.66%
Optical interference	14420	27.40%	35.13%	21.94%

Threshold transfer is weak across the rule-based stress domains, especially in the running case (F1=3.07%). The result indicates that fixed quality thresholds are sensitive to domain shift and should be calibrated for the specific sensor, placement, and wearing condition rather than reused as universal values.

The contrast between Table 2 and Table 3 is instructive. Within the controlled Record 100 noise experiment, the individual indices and the composite Q change systematically as SNR is reduced. However, applying one fixed set of thresholds across the broader stress labels produces substantially weaker classification results. This difference suggests that monotonic response to one degradation axis does not guarantee domain-independent decision boundaries. Sensor placement, artifact morphology, and the relative contribution of motion and optical disturbance can shift the distributions of the same SQI. The standards-first framework addresses this issue by treating threshold values and normalization rules as versioned sensor configuration rather than embedding them as universal constants in the fusion code. The low transfer F1 values therefore identify a calibration requirement rather than contradicting the gating principle itself. In a deployment study, threshold selection should be performed on device- and placement-specific data and then frozen or versioned through the quality specification used by the acquisition system.

4.3 Runtime and Fusion Performance

Table 4. Timing statistics and computational cost

Operation	Timing / update definition	Mean / cost	Std. dev.	Range	Avg. CPU
Motion pre-check	Event-to-flag; 50 Hz	1.8 ms latency; 0.05 ms/iteration	0.5 ms	1.1-3.2 ms	0.2500%
Contact/saturation SQI	Window ready → output	6.7 ms	1.8 ms	4.2-11.3 ms	—
Optical-interference SQI	Window ready → output	7.8 ms	2.1 ms	4.8-13.7 ms	—
kSQI	Every 2.5 s	0.29 ms	—	—	0.0116%
pSQI	Every 2.5 s	0.66 ms	—	—	0.0264%
basSQI	Every 2.5 s	0.22 ms	—	—	0.0088%
Composite score/weight	Every 2.5 s	0.02 ms	—	—	0.0008%

Software timing values refer to processing after a complete 5 s window is available. The three SQIs plus composite scoring require about 1.19 ms per 2.5 s update (≈0.048% average CPU), and the 50 Hz motion pre-check contributes ≈0.25%, yielding an estimated total average computation load of ≈0.30% on the 72 MHz STM32F103.

Fusion accuracy.

Table 5. Heart-rate MAE: fixed weighting versus adaptive quality gating

Scenario	Fixed weighting	Adaptive gating	Change
Rest	0.66	0.80	-21.4%
Walking	23.36	20.14	13.8%
Running	31.61	17.47	44.7%
Loose fit	2.80	2.58	7.8%
Optical interference	15.50	10.52	32.1%

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

In the high-quality resting condition, adaptive gating increases MAE slightly from 0.66 to 0.80 bpm. Its benefit appears mainly when signal quality degrades: MAE falls by 44.7% in the running stress case and by 32.1% under optical interference.

The fusion results also show why a quality gate should not be evaluated only by average improvement. In the resting condition, where both task channels are already high quality, the additional quality logic does not improve the estimate and produces a small MAE increase. Under stronger degradation, however, the difference between fixed weighting and adaptive gating becomes much larger because the fixed scheme continues to assign weight to an unreliable channel. The running result is the clearest example in the present data. The ablation study further indicates that the fast motion cue and the window-based SQIs serve different roles rather than duplicating one another. Motion-only gating approaches the full-method result in the running stress condition, as expected for a scenario dominated by motion, whereas removing the motion pre-check produces a larger error. The small remaining gap between motion-only and the full method suggests that waveform-based quality information still contributes beyond the fast contextual cue. These observations support a layered design in which rapid context detection protects the short-term response and slower signal analysis provides modality-specific confirmation.

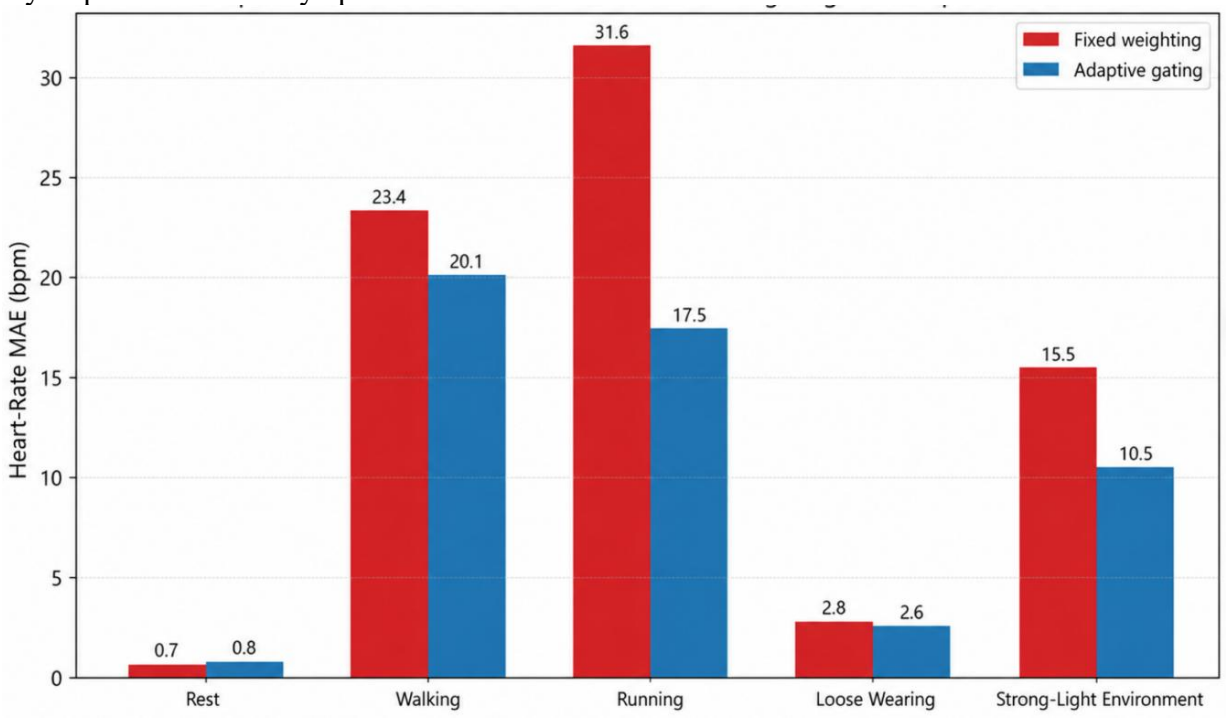


Fig. 2. Heart-rate error comparison between fixed weighting and adaptive quality gating.

Table 6. Ablation study in the running stress scenario

Configuration	Description	MAE (bpm)	Δ vs. full
Full method	Motion pre-check + SQI + four-state gating	17.47	—
SQI only	No fast motion pre-check	23.45	+5.98
Motion only	No software SQIs	17.87	+0.40
Fixed weighting	No quality decision or hard gate	31.61	+14.14

The ablation results separate the contributions of the two quality mechanisms. Removing the motion pre-check increases MAE to 23.45 bpm, while motion-only gating gives 17.87 bpm. The full method achieves 17.47 bpm, compared with 31.61 bpm for fixed weighting.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

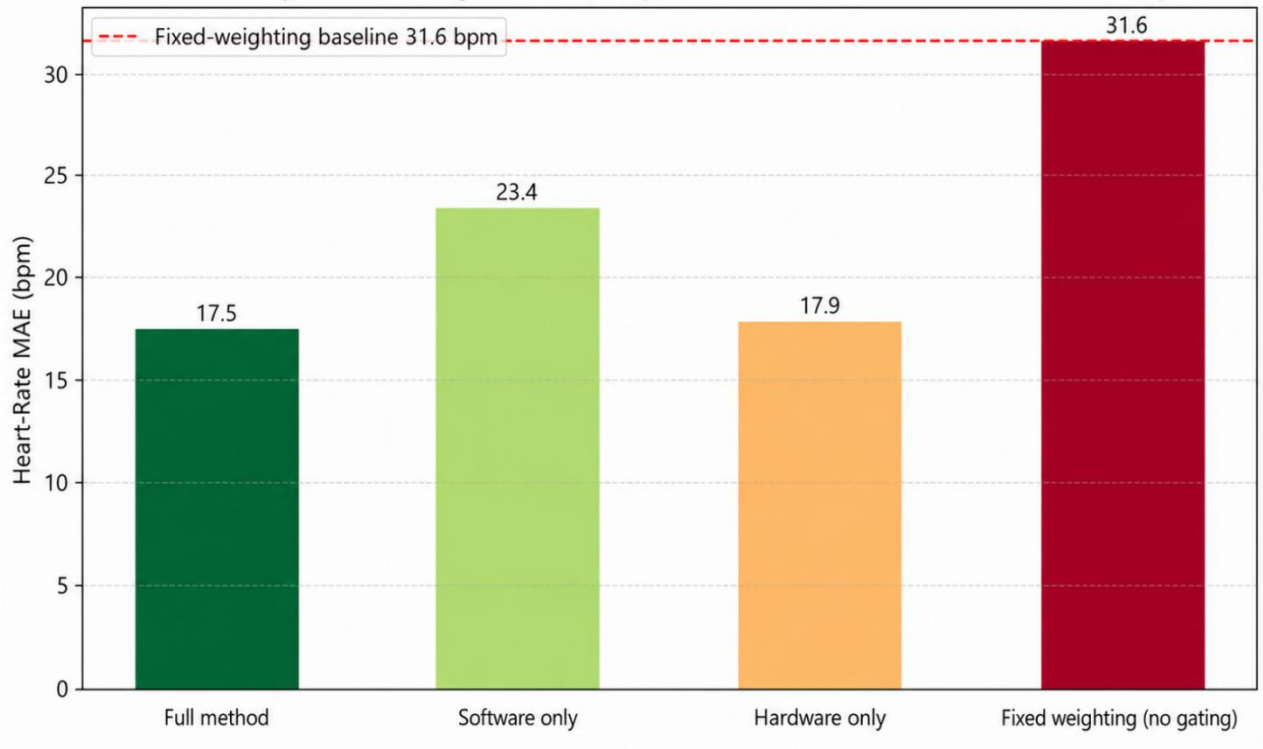


Fig. 3. Ablation results under the running stress scenario.

4.4 Temporal and System-Level Validation

Table 7. Clock-drift contribution to PTT timing error

Clock condition	Without reconstruction	With reconstruction	Reduction
No drift	0.1 ± 0.0 ms	0.1 ± 0.0 ms	0.0%
±50 ppm	7.4 ± 0.2 ms	0.1 ± 0.0 ms	98.2%
±200 ppm	29.7 ± 0.7 ms	0.1 ± 0.1 ms	99.6%

Timeline reconstruction removes most of the clock-drift contribution to the PTT time-base error. At ±200 ppm, this component decreases from 29.7 ms to approximately 0.1 ms. PPG sampling quantization and fiducial-point localization remain separate error sources and are not included in this number.

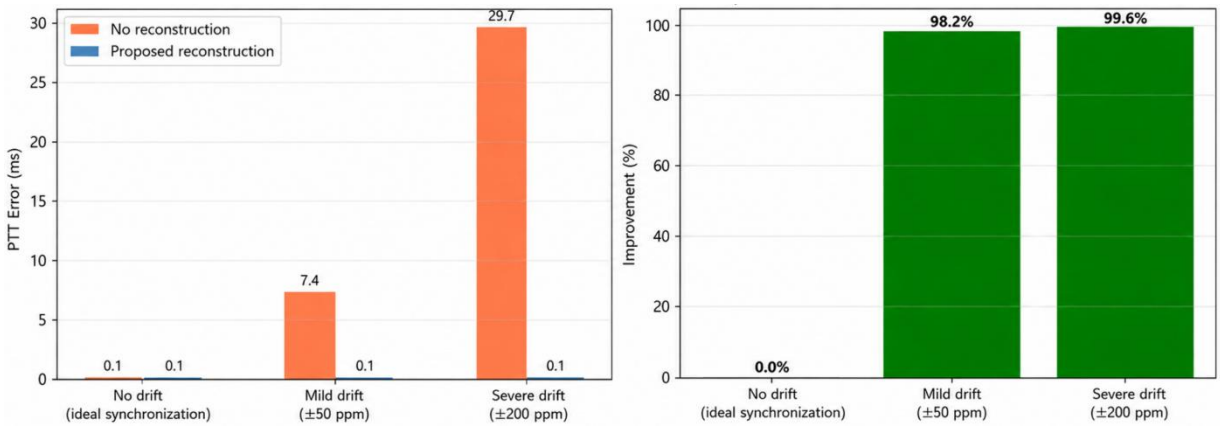


Fig. 4. Clock-drift contribution to PTT error before and after timeline reconstruction.

Figure 5 shows a representative 30-s segment in which all four controller states occur. Table 8 summarizes state occupancy and dwell-time statistics over the full 60-s running stress sequence.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

Table 8. Four-state switching statistics for the 60-s running stress sequence

Statistic	S1 full-quality	S2 degraded	S3 single-channel	S4 unavailable
Time occupancy	62.7%	17.6%	12.8%	6.9%
Mean dwell time	3.8 s	2.1 s	1.5 s	0.9 s
Longest continuous segment	12.3 s	8.7 s	5.2 s	2.8 s

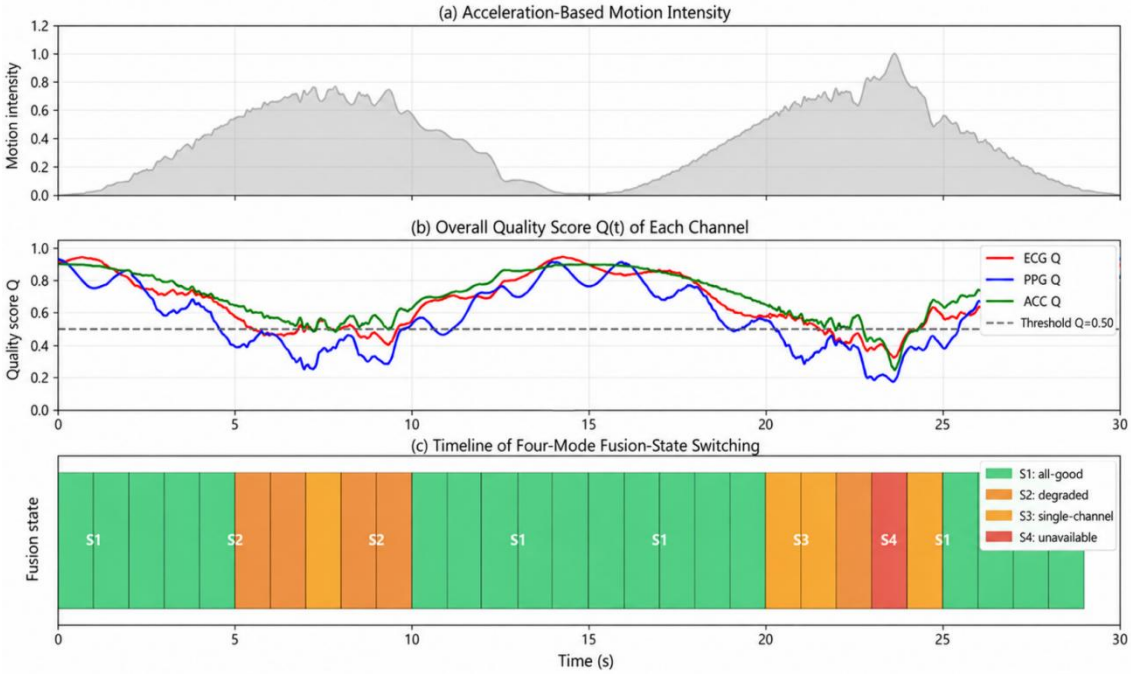


Fig. 5. Four-mode state-transition timeline.

Across the analyzed windows, fusion confidence is negatively correlated with heart-rate estimation error (Pearson r=-0.73). Because adjacent 5-s windows overlap by 50%, the correlation is reported descriptively without an independence-based significance p-value.

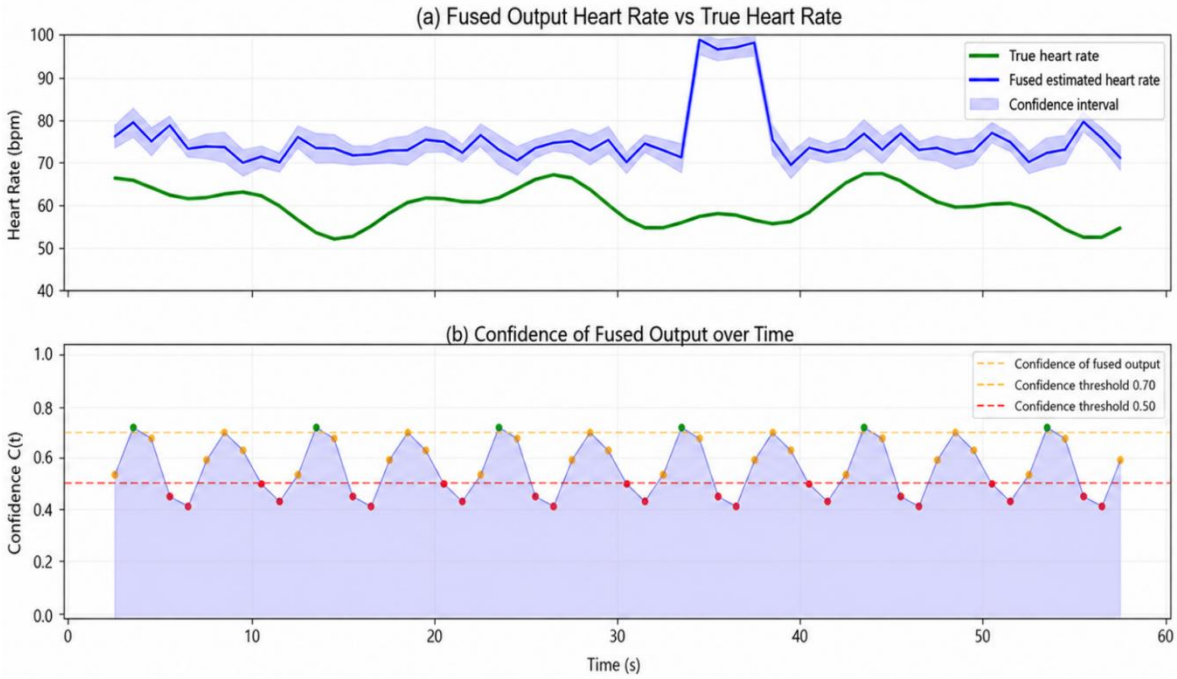


Fig. 6. Fused heart-rate output and confidence over time.

Table 9. Software-level consistency test of the standards-first protocol

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39

Injected violation	Injected	Detected	Detection rate
contract_hash mismatch	100	100	100%
cfg_epoch mismatch	100	100	100%
CRC error	100	100	100%
Sequence error	100	100	100%

The software-level test generates 10,000 frames across four configuration epochs and injects 100 non-overlapping violations of each type. All injected violations are detected, and none of the remaining 9,600 clean frames produces a false alarm. Because these checks are deterministic, the result verifies the protocol logic in software; it does not establish packet-loss, clock, or hardware robustness on a physical link.

Taken together, the temporal and protocol experiments evaluate aspects that are not captured by heart-rate MAE alone. Timeline reconstruction tests whether asynchronous acquisition can be placed on a common time base without attributing clock drift to the physiological estimator. State occupancy and dwell-time statistics show that the controller does not remain locked in a single operating mode under the running stress sequence; instead, it traverses full-quality, degraded, single-channel, and unavailable conditions. The negative association between confidence and heart-rate error is consistent with the intended interpretation of C(t), although the overlapping windows mean that the reported correlation should remain descriptive. Finally, the protocol-consistency test verifies that data carrying an unexpected contract, epoch, sequence, or CRC value can be identified before replay or downstream analysis. These checks are deliberately separated from physiological validation. They demonstrate that the proposed architecture can expose quality, timing, and configuration context to the fusion layer, while the limitations stated in Section 5 identify the additional evidence needed before making claims about clinical or cross-device performance.

5. Conclusions

This paper presents a standards-first adaptive quality-gated fusion framework for wearable multi-source physiological monitoring. Heterogeneous observations are first mapped to a common task-estimation layer, while the sensor quality specification, contract_hash, and cfg_epoch make quality semantics and configuration scope explicit. A fast motion pre-check complements the slower window-based SQIs, and hard gating prevents severely degraded task channels from retaining fusion weight. In mechanism-level experiments, the running-scenario heart-rate MAE is 44.7% lower than with fixed weighting. The timing analysis separates millisecond-scale computation from second-scale window updates and gives an estimated average CPU load of about 0.30%. Timeline reconstruction removes the clock-drift component of PTT timing error, and the state, confidence, and protocol tests show that the proposed control and traceability mechanisms behave as intended.

The present study has several limitations. The quality-score weights are empirical, real synchronized ECG+PPG+ACC wearable data have not yet been used for validation, and the logs do not contain a window-by-window best-single baseline. The PTT experiment isolates clock error rather than total PTT error. End-to-end STM32 protocol validation and hysteresis for state transitions will be addressed in future work.

Acknowledgements

Binzhou Institute of Technology, Weiqiao-UCAS Science and Technology Project (GYY-KT-2025-ZY-002).

References

[1] Z. Zhao and Y. Zhang. SQI quality evaluation mechanism of single-lead ECG signal based on simple heuristic fusion and fuzzy comprehensive evaluation. Frontiers in Physiology, 2018, 9: 727.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P39
[2] G. D. Clifford, J. Behar, Q. Li, and I. Rezek. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. Physiological Measurement, 2012, 33(9): 1419-1433.

[3] D. Seok, S. Lee, M. Kim, J. Cho, and C. Kim. Motion artifact removal techniques for wearable EEG and PPG sensor systems. Frontiers in Electronics, 2021, 2: 685513.

[4] A. John, A. P. James, B. Cardiff, and D. John. A review on multi-sensor data fusion for wearable health monitoring. Information Fusion, 2026, 133: 104319.

[5] M. Toptsis, N. Karkanis, A. Giannakoulas, and T. Kaifas. A review of embedded software architectures for multi-sensor wearable devices: sensor fusion techniques and future research directions. Electronics, 2026, 15(2): 295.

[6] D. Tegegne, S. R. Niakan Kalhori, P. Haas, V. M. G. Sobotta, J. Warnecke, and T. M. Deserno. Wearable devices for remote monitoring of chronic diseases: systematic review. JMIR mHealth and uHealth, 2026, 14: e74071.

[7] M. Elgendi. Optimal signal quality index for photoplethysmogram signals. Bioengineering, 2016, 3(4): 21.

[8] J. Allen. Photoplethysmography and its application in clinical physiological measurement. Physiological Measurement, 2007, 28(3): R1-R39.

[9] Q. Li and G. D. Clifford. Dynamic time warping and machine learning for signal quality assessment of pulsatile signals. Physiological Measurement, 2012, 33(9): 1491-1501.

[10] Y. Sun and N. V. Thakor. Photoplethysmography revisited: from contact to noncontact, from point to imaging. IEEE Transactions on Biomedical Engineering, 2016, 63(3): 463-477.

[11] B. Bent, P. J. Cho, M. Henriquez, A. Wittmann, C. Thacker, M. Feinglos, M. J. Crowley, and J. P. Dunn. Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches. npj Digital Medicine, 2021, 4: 89.

[12] A. Reiss and D. Stricker. Introducing a modular activity monitoring system. Proc. IEEE EMBC, 2011: 5621-5624.

[13] Z. Zhang. Heart rate monitoring from wrist-type photoplethysmographic signals during intensive physical exercise. Proc. IEEE GlobalSIP, 2014: 698-702.

[14] Y. Ye, Y. Cheng, W. He, M. Hou, and Z. Zhang. Combining nonlinear adaptive filtering and signal decomposition for motion artifact removal in wearable photoplethysmography. IEEE Sensors Journal, 2016, 16(19): 7133-7141.

[15] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. Proceedings of the IEEE, 2004, 92(3): 401-422.