Heart Rate Monitoring using Sparse Spectral Curve Tracing

Menglian Zhou, and Nandakumar Selvaraj*

Abstract— Heart rate (HR) monitoring under real-world activities of daily living conditions is challenging, particularly, using peripheral wearable devices integrated with simple optical and acceleration sensors. The study presents a novel technique, named as CurToSS: CURve Tracing On Sparse Spectrum, for continuous HR estimation in daily living activity conditions using simultaneous photoplethysmogram (PPG) and triaxial-acceleration signals. The performance validation of HR estimation using the CurToSS algorithm is conducted in four public databases with distinctive study groups, sensor types, and protocols involving intense physical and emotional exertions. The HR performance of this time-frequency curve tracing method is also compared to that of contemporary algorithms. The results suggest that the CurToSS method offers the best performance with significantly (P<0.01) lowest HR error compared to spectral filtering and multi-channel PPG correlation methods. The current HR performances are also consistently better than a deep learning approach in diverse datasets. The proposed algorithm is powerful for reliable long-term HR monitoring under ambulatory daily life conditions using wearable biosensor devices.

Index Terms— Mobile Health, Photoplethysmography, Heart Rate Monitoring, Time-frequency Spectrum, Curve Tracing, Motion Artifacts Removal

I. Introduction

Long term heart rate (HR) monitoring is essential for tracking health and fitness. Photoplethysmography (PPG) is a ubiquitous optical technology that measures the variations of pulsatile blood volume at a skin surface. Due to its non-invasive, simple and inexpensive nature, PPG is integrated in most wearable devices and applied for widespread HR monitoring as a substitute for complex electrocardiogram (ECG). However, obtaining a reliable and accurate PPG-based HR under ambulatory real world conditions remains a challenge. Anatomical factors (such as body composition and pigmentation) and sensor location (arm, wrist, forehead) affect the signal quality substantially, and furthermore, PPG signals are sensitive to physical activity and contact surface related motion artifacts (MA).

In recent years, various PPG based HR monitoring approaches have been evolving with MA detection and reduction capabilities. The first category of methods without requiring a simultaneously acquired accelerometer (ACC) signal include empirical mode decomposition (EMD) [1] and independent component analysis (ICA) [2]. The second category, in contrast, employs the simultaneous motion data captured by ACC that lead to robust algorithms including adaptive filtering such as normalized least mean squares (LMS) [3]; spectral filtering method such as SpaMA [4]; and

Authors are with Biofourmis Inc., Boston, MA 02110 USA. (*correspondence e-mail: `kumar@biofourmis.com`)

sparse spectrum estimation based methods such as TROIKA [5] and JOSS [6]. A tremendous research interest continues to contributing other noteworthy algorithms including Wiener Filtering and Phase Vocoder (WFPV) [7] and multiple initialization spectral peak tracking (MISPT) [8]. A detailed literature review can be found in [9].

These HR estimation methods are mostly validated with in-house laboratory testing or public databases [10] consisting of short duration recordings capturing limited human motions (like walking, arm swinging) under controlled laboratory conditions. Thus, validation testing of such HR algorithms may have limited accuracy and robustness for day-to-day real life activities. To address these shortcomings, two datasets [11], [12] were introduced recently with relatively long recording duration and various activities of daily living ranging from low to high intensities. With these new datasets, Reiss et al. [11] developed a deep learning approach that outperforms contemporary algorithms like SpaMA [4]. Novel algorithms with superior performance and dependability are still warranted to meet the clinical standards and needs.

In this paper, a novel algorithm named as CurToSS (CURve Tracing On Sparse Spectrum) for HR estimation using simultaneous PPG and ACC signals has been proposed. The CurToSS algorithm framework is detailed in Section II. The performance validation results on four public datasets are compared with other contemporary methods in section III. The significance and limitations of the proposed method, and future directions are discussed in section IV.

II. Materials and Methods

A. Datasets

The four datasets used in this paper are listed below:

1) IEEE_Training: This dataset [10] contains 12 records comprised of a wristband device embedded with PPG and ACC sensors, and a chest ECG sensor for HR ground truth. During a 5-min protocol, 12 participants walked and ran on a treadmill at three different speeds.

2) IEEE_Testing: This dataset [10] contains 10 records with the same hardware but a different 5-min protocol that involved 8 subjects performing intensive arm movements (such as boxing, running, jump and push-up) or various forearm and upper arm exercises (such as stretch and push).

3) DaLiA (PPG dataset for motion compensation and HR estimation in Daily Life Activities): This dataset [11] contains 15 records with a commerial wristband (Empatica E4) for PPG and ACC signals, as well as a chest ECG (RespiBAN) for HR ground truth. The protocol takes 2−3 hours, and the 15 subjects performed eight physical activities below in sequence with flexible transition time in between:
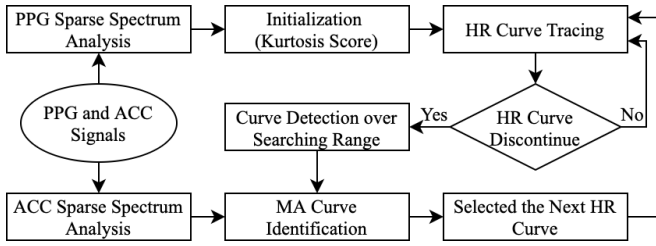
Fig. 1. A block diagram representation of CurToSS Framework

Fig. 2. An example of sparse spectra of (a) PPG signal and (b) x-axis accelerometer signal (ACC$_x$), from subject S2 of the IEEE_Training dataset. The green tracing corresponds to the HR curves whereas MA curves are highlighted with red tracing. The panel (c) shows a very close correspondence between the CurToSS based HR values to that of the ground truth HR values.

sitting still (10 mins), ascending/descending stairs (5 mins), table soccer (5 mins), cycling (8 mins), driving (15 mins), lunch (30 mins), walking (10 mins) and working (20 mins).

4) WESAD (Wearable Stress and Affect Detection): This dataset [12] contains 15 records using the same hardware as DaLiA but with a different 2 hour protocol, in which 15 subjects performed four sedentary activities sequentially with flexible breaks: baseline (sitting/standing for 20 mins at a table, reading magazines), amusement (sitting/standing for about 7 mins, watching funny video clips), stress (sitting/standing for 10 mins, with public speaking and mental arithmetic), and meditation (carrying out controlled breathing exercises for 7 mins while sitting in a comfortable position).

B. CurToSS framework

The proposed CurToSS framework presented in Fig. 1 involves unique spectral peak tracing (SPT) and MA reduction strategies using PPG and ACC derived sparse spectra, as compared to prior sparse spectrum estimation-based methods [5], [6]. The SPT accounts the HR variations manifested as smooth frequency curves or ridges on sparse spectra, restrained by time-varying physiological boundaries and restricted to a limited transient jumps. Thus, the SPT results in tracing of frequency curves pertinent to HR on the PPG spectra.The HR curves are practically not always continuous particularly in the presence of MA and other noise/interference. Once the HR curve found to be discontinued, a curve detection will be performed, MA curves and other interference curves will be excluded, and the curve with highest curve strength will be established as the next HR trace to resume the curve tracing. The major steps of CurToSS are described in detail below.

1) Sparse spectrum reconstruction (SSR): Sparse spectrum reconstruction provides a high resolution, low variance and robust time-frequency spectrum estimation [5]. The raw PPG and ACC signals are first downsampled to 8 Hz to reduce computational load; the time-series PPG and ACC signals are segmented with 8s sliding window (window shift: 2s); and the PPG and ACC segments are processed to estimate their respective sparse spectra. FOcal Underdetermined System Solver (FOCUSS) algorithm [13] is used to estimate sparse spectrum and more details can be found in [5].

2) Initialization: SPT begins with an initialization that normally requires a motion-free baseline condition. Such a baseline period can be identified by higher kurtosis scores calculated from the sparse spectra indicating the presence of motion-free dominant HR curves. To establish a starting
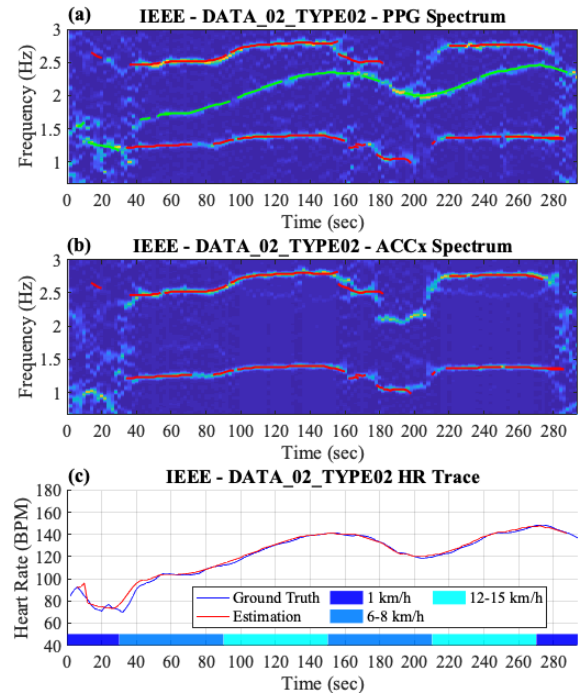
point for SPT, the kurtosis scores of the starting PPG segments within the HR range are tracked until three consecutive segments have kurtosis scores above a preset threshold. The dominant frequency corresponding to highest peak location on the SSR spectrum for the last of the three segments is taken as the initial HR, and curve tracing commences.

3) Curve tracing: The curve tracing is mainly comprised of two steps adapted from [14] for HR estimation.

(i) Identify pixel curve points from the spectral image:
A pixel of an image is identified as a curve point if its 1st directional derivative along the direction perpendicular to the curve (same direction as the eigenvector corresponding to the maximum absolute eigenvalue of the Hessian matrix H) vanishes within a unit square around this pixel [15].

$$H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix} \quad (1)$$

The 2nd partial derivatives, I_{xx}, I_{xy} and I_{yy} are computed using partial differences after convolving the image with a Gaussian smoothing kernels, which makes the derivatives estimation of a noisy spectrum image well posed. Once confirmed as curve point, the orientation θ, sub-pixel location (p_x, p_y) and eigenvalues of Hessian matrix on that pixel can be derived with its 1st and 2nd partial derivatives.

(ii) Track the subsequent curve pixel points:
Based on the orientation of the current curve point, three candidate pixels are identified. If there is only one curve point within the three pixels, that pixel is linked to the current point as the next curve point; if there is more than one

curves point within the three pixels, subpixel distance and orientation difference to the previous curve point are ranked to chose the curve point with minimum differences d as the next curve point as shown in (2);

$$d = \|\mathbf{p_2} - \mathbf{p_1}\|_2 + min\{|\theta_2 - \theta_1|, \pi - |\theta_2 - \theta_1|\} \quad (2)$$

where $\mathbf{p_1}$ and θ_1 are the sub-pixel location and orientation of previous curve pixel point, $\mathbf{p_2}$ and θ_2 are the sub-pixel location and orientation of the new pixel point. If there is no curve point within the three pixels, then the current HR curve is faded (*i.e.*, discontinued). Ideally, the HR curve can be continuously tracked as long as there is no loss of sensor contact or other perturbations. However, in practice, the HR curve fades out from time to time hence a re-establishing step is required to decide the next HR curve.

4) Curve detection (once HR curve is discontinued): Once the HR trace fades out, curve detection will be performed to find all non-vertical curves (including HR, MA and other curves) in the next predetermined number of consecutive segments. To exclude other interference curves such as the respiratory trace and higher-order harmonics curves, this detection step is limited to be within a searching range based on the previous HR, as described by (3):

$$HR_{cur} \geq HR_{prev} - max\{H_{min}, th_{dec} * D * \frac{HR_{prev}}{f_{scaling}}\}$$
$$\leq HR_{prev} + max\{H_{min}, th_{inc} * D * \frac{HR_{prev}}{f_{scaling}}\}$$
$$(3)$$

where HR_{prev} is the last HR extracted from the previous HR curve, HR_{cur} is the first HR from the candidate HR curve, D is the gap (measured as number of segments) between the previous and candidate HR curves, H_{min} is the minimum searching range between two consecutive curves, $f_{scaling}$ is the HR scaling factor, th_{inc} is the HR increasing threshold and th_{dec} is the HR decreasing threshold. For the j_{th} detected curve with length L and its curve points' subpixel locations as $(p_x(l), p_y(l)), l = 1, 2, ..., L$, the correlation with the PPG spectrum $Corr_{PPG}(j)$ and with the ACC spectra $Corr_{ACC,i}(j), i = x, y, z$ are calculated as (4) and (5).

$$Corr_{PPG}(j) = \frac{1}{L}\sum_{l=1}^{L} Eig_{PPG}(p_x(l), p_y(l)) \quad (4)$$

$$Corr_{ACC,i}(j) = \frac{1}{L}\sum_{l=1}^{L} Eig_{ACC,i}(p_x(l), p_y(l)) \quad (5)$$

where Eig_{PPG} and $Eig_{ACC,i}$ are the pixel eigenvalues of the Hessian matrix of their corresponding sparse spectrum. Note that $(p_x(l), p_y(l))$ are curve points' subpixel locations hence are non-integer, so $Eig_{PPG}(p_x(l), p_y(l))$ and $Eig_{ACC,i}(p_x(l), p_y(l)$ are interpolation results using the nearby pixels' eigenvalues. The eigenvalue images Eig_{PPG} and $Eig_{ACC,i}$ are currently used to measure the curves' strength relatively better than using the original sparse spectrum images I_{PPG} and $I_{ACC,i}$.

5) Motion artifact curve identification using ACC sparse spectrum: During the HR curve re-initialization step, the curve detection procedure finds all curves including the MA curves. The MA curves in the PPG spectrum normally have similar, corresponding curves in the ACC spectrum hence the MA curves have high correlation with one or more ACC spectra. On the other hand, the HR curves in PPG spectrum normally have low correlation with the ACC spectrum since the blood volume change can not be easily captured by the ACC sensor at the wrist, arm or forehead. As a result, the MA curves can be effectively identified due to their high correlations with the ACC spectrum using criteria (6) & (7).

$$Corr_{ACC,i} \geq th_{ACC}, i = x, y, z \quad (6)$$

$$\frac{1}{3}\sum_{i=x,y,z} Corr_{ACC,i} \geq Corr_{PPG} * th_{Ratio} \quad (7)$$

where th_{ACC} is the ACC spectrum correlation threshold and th_{Ratio} is the ratio threshold between ACC and PPG spectrum correlations.

6) New HR curve selection based on curve strength: After MA curves are identified and excluded, the remaining candidate curves are ranked based on their curve strength S, defined as (8),

$$S = Corr_{PPG} - \frac{1}{3}\sum_{i=x,y,z} Corr_{ACC,i} - \frac{|HR_{prev} - HR_{cur}|}{f_{scaling}}$$
$$(8)$$

where, the curve with the maximum curve strength is selected as the next HR curve to resume tracking. If no candidate curves are found, then the curve searching range will be expanded to additional consecutive segments until one or more candidate HR curves are found.

Fig. 2 demonstrates the CurToSS algorithm using a sample record from IEEE_Training dataset. In the PPG sparse spectrum shown in (a), the potential HR and MA curves are highlighted in green and red traces, respectively. The red MA curves on the PPG spectrum have high correlation with the ACC_x spectrum shown in (b), while the green HR curves on the PPG spectrum has low correlation with the ACC_x spectrum, as anticipated due to the fact that ACC signal from peripheral site might contain predominantly motion components. Thus, the subsequent HR curve tracing step allows effective rejection of the MA curves based on their high correlation with the three concurrent ACC sparse spectra. A very close correspondence between the predicted HR values of the CurToSS method to that of the ground truth HR values is shown in (c) for comparison from this record.

C. CurToSS Parameters

The CurToss parameters are common for DaLiA and WE-SAD except th_{inc} that can be customized for low, medium or high physical exertion levels, and the parameter values are listed as below: the curve search range is set to 30 segments; $f_{scaling}$ is the HR scaling factor as 100; H_{min} is the minimum search range between two consecutive curves to account for abrupt HR changes and is set to 15 beats per

minute (bpm); th_{dec} is the HR decreasing threshold and is set to 1; th_{inc} is the HR increasing threshold and is set to 1 for WESAD and 3 for DaLiA, since DaLiA involves high intensity activities that boost up HR rapidly; th_{ACC} is the ACC correlation threshold and is set as 0.24; th_{Ratio} is the ratio threshold between averaged ACC and PPG correlation and is set as 0.6.

D. Performance evaluation

The performance of the proposed CurToss algorithm is evaluated by the mean absolute error (*MAE*), the bias and limits of agreement (*LoA*), and the Pearson coefficient.

The *MAE* is defined as

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |HR_{est}(i) - HR_{ref}(i)|, \qquad (9)$$

where N is the total number of segments, and $HR_{ref}(i)$ and $HR_{est}(i)$ are the ground truth and the estimated HR in the ith segment, respectively. The similar error analyses were conducted for contemporary algorithms [4], [11], [16] for direct comparisons to the proposed approach.

The limits of agreement (*LoA*) per Bland-Altman analysis are defined as $[\mu - 1.96\sigma, \mu + 1.96\sigma]$, where μ is the bias (i.e., the average) & σ is the standard deviation of the signed error, $(HR_{est} - HR_{truth})$. The estimation and ground truth HR values for each segment are both rounded to integers prior to performance analyses. The statistical dispersion of the error is evaluated for each database based on their subject-wise *MAE*. The agreement analysis is carried out considering the entire study protocol as well as individual activity types. Paired sample t-tests are also conducted to validate the statistical significance of performance improvement for CurToSS over other contemporary algorithms.

III. RESULTS

The performances of CurToSS algorithm are compared to that of other contemporary algorithms (SpaMaPlus: spectral filtering based method [4]; Schaeck2017: multi-channel PPG correlation-based approach [16]; CNN ensemble: deep learning-based approach [11]) on aforementioned databases (Fig. 3). The CurToSS algorithm provides the best *MAE* performance (specifically with lowest mean error, highest consistency and smallest outliers) across all four datasets, indicating its high accuracy and effectiveness in reliably tracking HR over various intense activities. Among the other three algorithms, CNN ensemble algorithm show reasonable MAE for DaLiA and WESAD, but it is relatively poor for IEEE_Testing; Schaeck2017 has overall poor MAE performance, and the MAE distribution of SpaMaPlus is found between these two methods (Fig. 3).

A representative example using the CurToSS algorithm during the real-world daily activities is shown in Fig. 4. Comparing to IEEE_Training, the DaLiA spectrum is much more complex and contains many interference within the HR frequency range, as shown in (a). However, the proposed CurToSS method correctly identifies and traces the HR curves well, as illustrated. Two incorrectly picked curves are
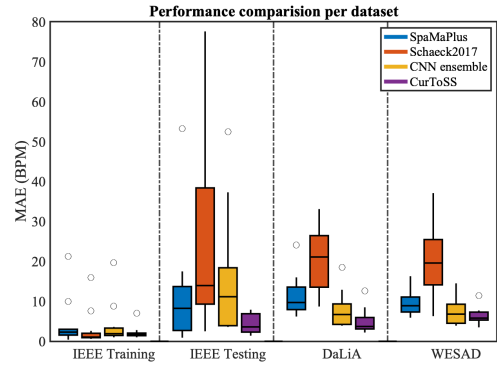

Fig. 3. Performances of HR estimation in mean absolute error (*MAE*) for CurToSS and other contemporary methods on four unique databases.
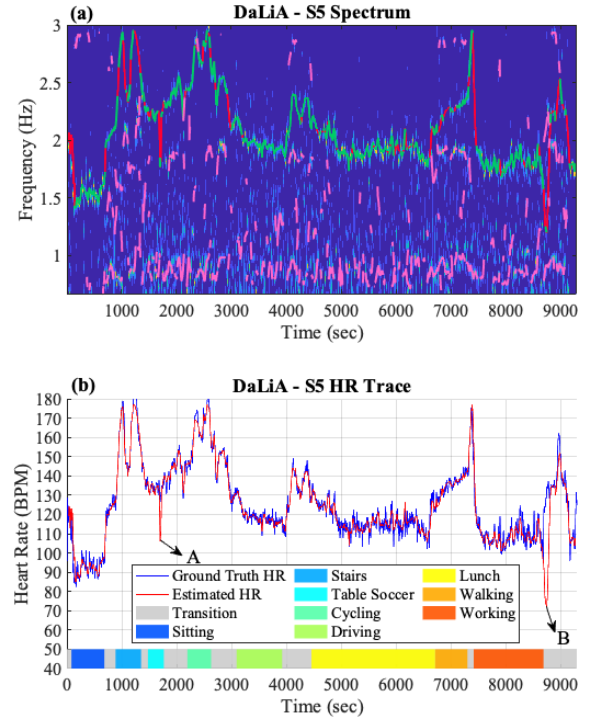

Fig. 4. An example of (a) PPG derived sparse spectrum and (b) comparison of predicted versus the ground truth HR values from DaLiA dataset, Subject S5. In the PPG sparse spectrum, the ground truth HR curves are marked in green and overlaid with the CurToSS estimated HR curves traced in red; the other unrelated curves including motion artifacts are marked in pink.

marked as A and B in (b). These are caused by overestimated curve searching ranges during HR curve re-initialization, which will be discussed further in Section IV-C.

Overall, the CurToSS algorithm achieved the best *MAE* of 2.2, 4.5, 5.0 and 6.4 bpm for IEEE_Training, IEEE_Testing, DaLiA and WESAD, respectively. Subject-wise *MAE*s for DaLiA and WESAD are listed in Table I , along with p values from paired sample t-tests to validate the statistical significance of the CurToSS against each contemporary method. The CurToSS algorithm *MAE* performances are significantly (p<0.01) better than spectral filtering (SpaMaPlus) and multi-channel PPG correlation (Schaeck2017) based approaches for both DaLiA and WESAD databases. Meanwhile, the MAE of CurToss are relatively lower than deep

TABLE I

DaLiA	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	Mean ± STD	p-value[a]
SpaMaPlus	8.9	9.7	6.4	14.1	24.1	11.3	6.3	11.3	16.0	6.2	15.2	12.0	8.5	7.8	8.3	11.1±4.8	<.001
Schaeck2017	33.1	27.8	18.5	28.8	12.6	8.7	20.7	21.8	22.3	12.6	21.1	22.7	27.7	12.1	16.4	20.5±7.1	<.001
CNN ensemble	7.7	6.7	4.0	5.9	18.5	12.9	3.9	10.9	8.8	4.0	9.2	9.4	4.3	4.4	4.2	7.7±4.2	0.062
CurToSS	5.4	4.3	3.0	8.0	2.2	2.8	3.3	8.5	12.6	3.6	3.6	6.1	3.0	5.5	3.7	5.0 ± 2.8	
WESAD	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S13	S14	S15	S16	S17	Mean ± STD	p-value
SpaMaPlus	7.3	12.6	5.9	7.4	6.4	9.2	8.9	6.3	8.2	16.3	13.0	10.2	8.3	10.7	11.2	9.5 ± 2.9	<.01
Schaeck2017	23.3	37.1	26.2	30.5	18.6	21.3	19.6	13.3	20.7	16.7	9.8	6.3	16.6	12.6	27.1	20 ± 8.1	<.001
CNN ensemble	5.1	14.5	7.8	7.7	3.9	6.8	4.3	4.0	8.9	11.1	6.5	5.3	4.2	12.8	9.4	7.5 ± 3.3	0.21
CurToSS	5.5	6.7	7.6	5.8	4.8	7.7	7.4	5.3	5.5	6.9	7.0	3.5	5.3	4.9	11.4	6.4 ± 1.8	

[a] P values from paired sample t-tests against the CurToSS method.

TABLE II

HR Estimation Performance per Activity

	DaLiA[a]										WESAD[b]					
Activity	Act 1	Act 2	Act 3	Act 4	Act 5	Act 6	Act 7	Act 8	Act 9	Overall	Act 1	Act 2	Act 3	Act 4	Act 5	Overall
Duration (mins)	152	109	78	116	228	452	157	284	566	2142	293	167	93	195	666	1415
Mean (bpm)	61.0	118.7	90.4	123.2	84.7	83.6	99.3	76.4	95.1	89.5	72.6	98.3	72.7	69.9	77.3	77.5
Bias (bpm)	-0.5	-2.6	-3.4	-1.1	-0.2	-1.2	-2.6	-0.8	-1.8	-1.4	-1.8	-6.8	-1.8	-2.2	-4.4	-3.7
Lower limit (bpm)	-7.7	-33.1	-24.3	-16.0	-15.0	-15.2	-30.9	-13.6	-26.4	-21.0	-12.5	-36.7	-11.7	-16.0	-25.9	-23.4
Upper limit (bpm)	6.8	27.8	17.5	13.8	14.7	12.9	25.8	11.9	22.9	18.3	8.8	23.1	8.1	11.6	17.1	16.1
Correlation	0.96	0.82	0.83	0.95	0.89	0.89	0.76	0.88	0.85	0.91	0.89	0.74	0.90	0.70	0.72	0.80

[a] Activity types for DaLiA: 1- Sitting, 2- Stairs, 3- Table Soccer, 4- Cycling, 5- Driving, 6- Lunch, 7- Walking, 8- Working, 9- Transition.
[b] Activity types for WESAD: 1- Baseline, 2- Stress, 3- Amusement, 4- Meditation, 5- Break.

learning-based CNN ensemble approach, but the differences are not statistically significant (p>0.05).

The impact of different activities on HR estimation using the CurToSS algorithm is further analyzed: the DaLiA and WESAD datasets are split according to activity types and HR performances are assessed for each activity. Specifically, the correlation, bias and 95% *LoA* between estimated and ground truth HR for each activity type are listed in Table II.

IV. Discussion

Accurate HR estimation using convenient wearable devices is highly valuable for continuous health monitoring, but the obvious human motion including intermittent physical activities and intense exercises during daily life present algorithmic challenges to obtain reliable HR in such real-world conditions. The novel CurToSS algorithm is designed to accurately track HR in ambulatory conditions by analyzing the sparse spectra curves from simultaneous PPG and ACC signals and effectively rejecting the motion artifacts. The study validates the CurToSS algorithm and this method is shown to outperform the contemporary methods for reliable HR estimation during wide range of activities.

A. Advantages of CurToSS algorithm

The proposed CurToSS algorithm solves the problem of estimating HR from the time-frequency spectra by extracting "ridge curves" (i.e. time sequences of certain component's frequency location in spectrum) and tracking spectral peaks using curve tracing and curve detection techniques from image processing and computer vision domains. Conventional SPT searches for the dominant peak within a certain frequency range around the last segment's HR peak from the MA-removed PPG spectrum. This scheme works well when the motions are mostly uncomplicated and well-defined, hence after spectrum subtraction the HR peak will be the dominant peak in the cleansed PPG spectrum.

However, the natural human motions exhibit complex signal dynamics and characteristics. For example, the MA are composed of multiple dominant frequency peaks which are distributed widely over the spectrum and can overlap the HR peaks, making MA subtraction/exclusion ineffective; hence, the dominant peak may not be the HR peak. The HR peak might temporarily disappear from the PPG spectrum due to various reasons like sensor contact loss; interference peaks will be chosen as HR peaks for these segments and the error will also propagate through later segments.

Our curve tracing method could remedy this situation with: 1. Instead of picking the maximum peak's location as HR, the curve tracing decides the searching range adaptively based on both current HR and the HR curve continuity (curve direction and sub-pixel distance) of the previous segments, hence reduces the chance to select a nearby MA peak as HR. 2. When the HR trace discontinues, a curve detection is performed to decide a new HR curve based on the curve strength. As a result, the re-established HR has higher confidence than simply taking the maximum peak of one segment and also avoids the propagation error associated with an incorrect starting point of peak tracking.

B. Impact of activity types to CurToSS performance

Table II presents the impact of different activities to HR estimation using CurToSS algorithm. The relatively poor

performance during stairs, table soccer and walking, might be due to the high amplitude arm movements causing sensor movements on the skin surface and potential ambient light leaking, and eventually leading to substantial degradation of the signal corrupted with motion artifacts, distortions and increased in-band noise. In contrast, cycling requires participants' hands to remain on a handlebar, limiting arm movements. Hence, the performance for cycling across 15 subjects is similar to that of driving and having lunch, even though the activity intensity and mean HR for cycling are much higher than the other two activities. Across all activities including transition, the CurToSS algorithm achieved a correlation of 0.91, bias of -1.4 bpm and 95% *LoA* of [-21.0, 18.3] bpm for DaLiA dataset. Barrios et al. [17] evaluated the same wearable device (Empatica, E4) in 14 subjects while performing similar fitness activities (1-hour protocol involving low/high power biking, walking, jogging and running) and found the HR to have a correlation as 0.41, bias as 17.8 bpm and 95% *LoA* as [-26.6, 62.2] bpm. The current estimation of HR is comparably far better in ambulatory conditions.

C. Limitations and future directions

There are a few limitations of the CurToSS algorithm: Firstly, when the HR trace is discontinued, the HR output will be on hold for certain segments to re-initialize the HR trace. However, with regards to the short-term uncertainty, we established a solid starting point for later HR tracing while other SPT methods would normally establish the new starting point by taking the dominant spectrum peak (which may not be the HR component) from one segment. Secondly, certain thresholds or system parameters are chosen empirically, but such parameters generalize globally. Even though the competitive performance (lowest average *MAE* and standard deviation comparing to other contemporary algorithms) across the independent datasets and recordings indicates that the system parameters are applicable for different devices, subjects and motions. Further improvement can be achieved by customizing these parameters adaptively. For example, th_{inc} and th_{dec} (i.e. the curve searching range) can be updated based on the simultaneous activity level (one example of activity level estimation is the squared derivative of raw ACC signals [18]). Thirdly, a common concern for the SSR-based algorithm is the computational complexity. As we have downsampled the PPG and ACC signals to 8 Hz, the computation load is greatly decreased. Currently, on a computer with Intel Core i7 @2.6GHz and 16 GB RAM, the Matlab run time for a DaLiA dataset with 4722 segments (8 sec per segment) is 58.2 ± 3.4 sec (13.6 ± 0.8 ms per segment), which means this algorithm is real-time realizable and comparable to other contemporary algorithms such as MISPT [8]. However, further improvement is necessitated to accommodate the limited resources (computation capability and energy consumption) of wearable devices.

In conclusion, a novel HR monitoring algorithm is presented using PPG and simultaneous ACC signals from smart wearable devices that uniquely traces HR ridge curves from a PPG sparse spectrum while rejecting the MA curves based on their correlation with ACC sparse spectrum. The validation results suggest that this framework is powerful and suitable for long-term HR monitoring using wearable devices in ambulatory daily-life conditions.

REFERENCES

[1] Q. Wang, P. Yang, and Y. Zhang, "Artifact reduction based on Empirical Mode Decomposition (EMD) in photoplethysmography for pulse rate detection," *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, pp. 959–962, 2010.

[2] J. Yao and S. Warren, "A short study to assess the potential of independent component analysis for motion artifact separation in wearable pulse oximeter signals," *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, vol. 7 VOLS, pp. 3585–3588, 2005.

[3] K. Chan and Y. Zhang, "Adaptive reduction of motion artifact from photoplethysmographic recordings using a variable step-size lms filter," in *SENSORS, 2002 IEEE*, vol. 2.   IEEE, 2002, pp. 1343–1346.

[4] S. M. Salehizadeh, D. Dao, J. Bolkhovsky, C. Cho, Y. Mendelson, and K. H. Chon, "A novel time-varying spectral filtering algorithm for reconstruction of motion artifact corrupted heart rate signals during intense physical activities using a wearable photoplethysmogram sensor," *Sensors (Switzerland)*, vol. 16, no. 1, 2015.

[5] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 522–531, 2015.

[6] Z. Zhang, "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 8, pp. 1902–1910, 2015.

[7] A. Temko, "Accurate Heart Rate Monitoring during Physical Exercises Using PPG," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2016–2024, 2017.

[8] P. S. Navaneet K. Lakshminarasimha Murthy, Pavan C. Madhusudana, Vijitha Periyasamy, and P. K. Ghosh, "Multiple Spectral Peak Tracking for Heart Rate Monitoring from Photoplethysmography Signal During Intensive Physical Exercise," 2015.

[9] D. Pollreisz and N. TaheriNejad, "Detection and Removal of Motion Artifacts in PPG Signals," *Mobile Networks and Applications*, 2019.

[10] Z. Zhang, "IEEE Signal Processing Cup 2015: Heart Rate Monitoring During Physical Exercise Using Wrist-Type Photoplethysmographic (PPG) Signals," 2015.

[11] A. Reiss, I. Indlekofer, P. Schmidt, and K. Van Laerhoven, "Deep PPG: Large-Scale Heart Rate Estimation with Convolutional Neural Networks," *Sensors*, vol. 19, no. 14, p. 3079, jul 2019.

[12] P. Schmidt, A. Reiss, R. Duerichen, and K. Van Laerhoven, "Introducing WeSAD, a multimodal dataset for wearable stress and affect detection," in *ICMI 2018 - Proceedings of the 2018 International Conference on Multimodal Interaction*.   New York, New York, USA: ACM Press, 2018, pp. 400–408.

[13] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm," *IEEE Transactions on signal processing*, vol. 45, no. 3, pp. 600–616, 1997.

[14] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 20, no. 2, pp. 113–125, 1998.

[15] K. Raghupathy, "Curve tracing and curve detection in images," *Master's Thesis, Cornell University*, no. August, pp. 1–70, 2004.

[16] T. Schäck, M. Muma, and A. M. Zoubir, "Computationally efficient heart rate estimation during physical exercise using photoplethysmographic signals," in *2017 25th European Signal Processing Conference (EUSIPCO)*, Aug 2017, pp. 2478–2481.

[17] L. Barrios, P. Oldrati, S. Santini, and A. Lutterotti, "Evaluating the accuracy of heart rate sensors based on photoplethysmography for in-the-wild analysis," in *ACM International Conference Proceeding Series*, 2019.

[18] J. Y. A. Foo and S. J. Wilson, "A computational system to optimise noise rejection in photoplethysmography signals during motion or poor perfusion states," *Medical and Biological Engineering and Computing*, vol. 44, no. 1-2, pp. 140–145, 2006.