

リモート PPG 信号による睡眠／覚醒判別 Sleep/wake classification using remote PPG signals

張 雅文^{*1}
Yawen Zhang

辻川 剛範^{*2}
Masanori Tsujikawa

大西 祥史^{*2}
Yoshifumi Onishi

^{*1} Department of Electronic & Computer Engineering, Hong Kong University of Science and Technology

^{*2} Biometrics Research Laboratories, NEC Corporation

This paper proposes a remote sleep/wake classification method by combining vision-based heart rate (HR) estimation and convolutional neural network (CNN). Instead of directly inputting the estimated HR to CNN, we input remote PPG (Photoplethysmogram) signals filtered by a dynamic HR filter, which can overcome two main problems: low temporal resolution of estimated HR; much noise exists in the estimated remote PPG signals. Evaluation results show that the dynamic HR filter works more effectively compared to the static one, which helps improve AUC (area under the curve) index of the classification to 0.70, as good as the performance (0.71) of HR from a wearable sensor.

1. はじめに

眠気推定や睡眠／覚醒判別など覚醒度低下の定量化は、疲労のリスク管理、運転中、勉強中、仕事時の状態モニタリングや、ヘルスケア等の応用のために広く研究されている。眠気推定手法としては、装着型センサを利用する手法やカメラを利用する手法があるが、中でも顔映像から抽出した顔の動きを用いる手法は、非接触型で使い勝手が良く、また高精度に眠気を推定できる[Oliveira 2018, Tsujikawa 2018]。しかし、その手法は顔を閉じている状況(閉眼時)では常に眠いと判定されてしまうという問題があり、実際に寝ているのか、目を閉じているだけなのか判別できない。

閉眼時でも睡眠／覚醒を判別できる手法として、心電図(Electrocardiogram: ECG)信号や光電式容積脈波記録法(Photoplethysmogram: PPG)信号を用いる手法がある[Aktaruzzaman 2015, Ye 2016, Scherz 2017, Malik 2018]。ECG信号や PPG 信号から、連続した心拍の間隔(Interbeat interval: IBI)の時系列信号を抽出し、IBI の逆数である心拍数(Heart rate: HR)や、IBI の時間変動である心拍変動(Heart rate variability: HRV)を分析して、睡眠／覚醒を判別する。心臓の動きは自律神経により制御されており、覚醒中は交感神経が、睡眠中は副交感神経が支配的になるため、覚醒中は睡眠中より心臓が活発に動く[Somers 1993]。そのため、睡眠／覚醒判別において、心臓の動きを表す HR や HRV は重要な生理学的指標とされている。

近年、HR や HRV を算出するための脈波を含む信号を取得する手法として、離れて配置された(リモート)センサを利用する手法の研究が盛んである。文献[Rahman 2015]は、ドップラーレーダで抽出した HR と呼吸数、体動を用いて、睡眠をセンシングできるシステムを提案している。そのシステムにより、高精度に睡眠／覚醒を判別できると報告されている。しかし、前述の顔映像から眠気を推定するシステムとの統合のためには、リモートセンサとしてカメラを利用することが望ましい。

顔映像から HR を推定する手法は、ますます盛んに研究されている[Wang 2018, Utkarsh 2018]。それは、顔の輝度値に基づく推定手法と顔の動きに基づく推定手法の二種に大別されるが、

前者の研究が盛んである。カメラで捉えた顔の輝度値から PPG 信号(リモート PPG 信号)を抽出した後、高速フーリエ変換(Fast Fourier Transform: FFT)などの周波数分析を用いて、HR を推定する。心臓は周期的に動いているため、パワーが大きい周波数成分を HR 成分とする。欠点は、周波数分析の窓幅に相当する区間の HR しか計算できず、一拍一拍の瞬間的な HR (Instantaneous HR: IHR)の変化を高精度に捉えることが困難な点である。それは、短い周波数分析窓幅では HR 情報が、頭の動きや光の変化等で生じる大きなノイズに埋もれるためである。IHR の変化は HRV の算出において重要であり、その HRV は睡眠／覚醒判別の重要な指標の一つである。したがって、リモートセンサとしてカメラを利用して、睡眠／覚醒を判別する際の課題は、リモート PPG 信号から IHR の情報、すなわち高時間解像度の HR 情報を高精度に抽出することである。

上記の課題に対して、高時間解像度の HR 情報を含むリモート PPG 信号による睡眠／覚醒判別を提案する。頭の動きや表情変化、光の変化により生じる大きなノイズで埋もれた、高時間解像度の HR 情報を抽出するために、低時間解像度の HR の分布情報から動的に設計したフィルタを用いてノイズを除去する。ノイズ除去後のリモート PPG 信号は、手動での特徴抽出を避けるために、畳み込みニューラルネットワーク(Convolutional neural network: CNN)に直接入力し、睡眠／覚醒を判別する。2章で提案法を詳細に説明し、3章で、睡眠／覚醒判別の評価実験において提案法の有効性を示す。

2. 睡眠／覚醒判別法

2.1 従来法

図 1 に文献[Malik 2018]で提案されている従来の睡眠／覚醒判別手法の構成を示す。装着型センサである ECG センサから ECG 信号を得て、連続した心拍の間隔 IBI を算出し、その逆数である IHR を以下の式で算出する。

$$IHR(i) = 60/IBI(i) \quad (1)$$

ただし、 $IHR(i)$ は i 番目の IHR で単位は bpm (beat per minute)、 $IBI(i)$ は i 番目の IBI で単位は秒である。 $IHR(i)$ は不等間隔でサンプリングした時系列信号のため、補間処理により等間隔でサンプリングして HR の時系列信号を得た後、一定の時間窓で区切った HR の時系列信号を CNN に入力する。HR 信号から睡

眠／覚醒を判別できるように学習した CNN により、新たに入力された HR 信号に対して睡眠／覚醒の判別結果を出力する。

文献 [Malik 2018] では、二つの公開データベース ① DREAMS Subject Database¹, ② UCSDSADB²を用いて、学習したモデルの判別精度を評価している。①に対する AUC (Area under the curve) が 0.81, ②に対する AUC が 0.72 で判別可能と報告されている。

上記の従来法における装着型センサから算出した HR (Wearable HR) を、カメラから算出した HR (Camera HR) に置き換えると、次の問題がある。主問題は、Camera HR の時間解像度が低いことである。Camera HR は、一定の推定精度を得るために、通常、IBI より長い時間窓で PPG 信号を周波数分析して算出される。もう一つの問題は、ノイズに起因する推定誤差である。次節で説明する提案法では、高い時間解像度の HR 情報を含む PPG 信号に混在したノイズを動的なフィルタで除去し、それを CNN に入力することで上記の二つの問題を解決する。

2.2 提案法

図 2 は、提案する睡眠／覚醒判別手法の構成を示す。それは、文献 [Utkarsh 2018] の Camera HR 推定法と文献 [Malik 2018] で適用された CNN を組み合わせた手法である。その Camera HR 推定法は他手法と比較して高精度に HR を推定できると報告されている。まず、提案法ではカメラで取得した顔映像から PPG 信号 (Camera PPG) を抽出する。次に、Camera PPG から Camera HR を推定し、Camera HR の値に基づいて動的に設計したフィルタ (Dynamic Camera-HR filter) で Camera PPG に含まれるノイズを除去する。ノイズ除去後の Camera PPG を図 3 に示す CNN に入力し、睡眠／覚醒を判別する。以下、Camera PPG に含まれるノイズの除去手順と文献 [Malik 2018] の CNN とのモデル構造の違いについて説明する。その他に関しては、文献 [Utkarsh 2018] と [Malik 2018] を参照して欲しい。

Dynamic Camera-HR filter は、Camera-PPG に含まれるノイズを除去することで、HR 情報を強調するバンドパスフィルタである。それは Camera HR を用いて設計される。一定時間窓内の HR が正規分布に従うと仮定し、その平均を μ 、標準偏差を σ とすると、 $\mu \pm 3\sigma$ の範囲に 99.7% の値が含まれる。窓番号 I の Dynamic Camera-HR filter のカットオフ周波数の下限 $L_D(I)$ と上限 $H_D(I)$ を $\mu(I)$ と $\sigma(I)$ を用いて次のように算出する。

$$L_D(I) = \frac{\mu(I) - 3\sigma(I)}{60}, H_D(I) = \frac{\mu(I) + 3\sigma(I)}{60} \quad (2)$$

窓番号 I 毎に、推定した HR に応じてバンドパスフィルタの下限と上限を動的に調整する。その下限と上限を用いて、2 次のパワースバンドパスフィルタを設計する。

Dynamic Camera-HR filter の効果を検証するため、次章の評価で用いる学習データセット (DREAMS) の HR 情報から設計した静的なバンドパスフィルタ (Static HR filter) と比較する。Static HR filter においては、下限 $L_S = 0.53\text{Hz}$ 、上限 $H_S = 1.75\text{Hz}$ である。それらの値は次式で算出した。

$$L_S = \min[L_{\text{wake}}, L_{\text{sleep}}], H_S = \max[H_{\text{wake}}, H_{\text{sleep}}] \quad (3)$$

$$L_{\text{wake}} = \frac{\mu_{\text{wake}} - 3\sigma_{\text{wake}}}{60}, L_{\text{sleep}} = \frac{\mu_{\text{sleep}} - 3\sigma_{\text{sleep}}}{60} \quad (4)$$

$$H_{\text{wake}} = \frac{\mu_{\text{wake}} + 3\sigma_{\text{wake}}}{60}, H_{\text{sleep}} = \frac{\mu_{\text{sleep}} + 3\sigma_{\text{sleep}}}{60} \quad (5)$$

ただし、 μ_{wake} と μ_{sleep} は、それぞれ学習データにおいて覚醒状態または睡眠状態とラベル付けされた時刻の HR の平均、同様に σ_{wake} と σ_{sleep} は標準偏差である。それらを用いて、覚醒状態における HR の分布の下限 L_{wake} と上限 H_{wake} 、睡眠状態にお

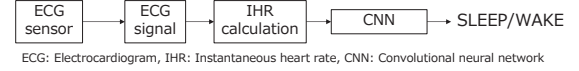


図1 従来の睡眠／覚醒判別手法の構成

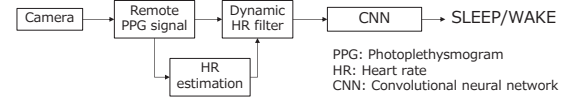


図2 提案する睡眠／覚醒判別手法の構成

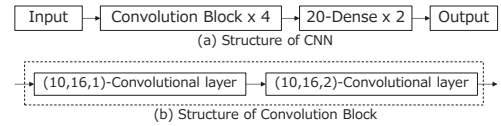


図3 (a) PPG 信号を入力する場合の 1 次元 CNN の構造。 m -Dense は m ノードの全結合層。 (b) 単一の畳込ブロックの構造。 (f, k, s) -convolution はフィルタ数 f 、カーネルサイズ k 、ストライド s の畳込層。

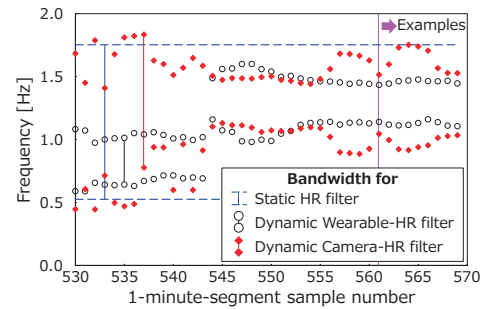


図4 Static HR filter と Dynamic HR filter の周波数帯域。

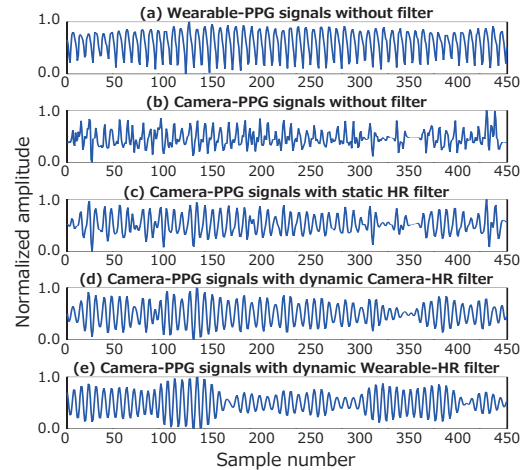


図5 Static HR filter と Dynamic HR filter の適用例。

ける下限 L_{sleep} と上限 H_{sleep} を算出し、最小値演算と最大値演算により、Static HR filter の下限 L_S と上限 H_S を算出した。

図 4 に Static HR filter と Dynamic HR filter の周波数帯域を例示する。比較のため、Dynamic HR filter は、Wearable HR から設計した Dynamic Wearable-HR filter と Dynamic Camera-HR filter の二種を示す。二種の Dynamic HR filter の周波数帯域は

¹ <http://www.tcts.fpms.ac.be/~devuyt/Databases/DatabaseSubjects>

² <https://physionet.org/pn3/ucddb>

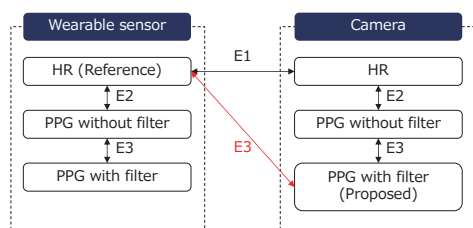


図6 三つの評価の関係. E1:評価1, E2:評価2, E3:評価3.

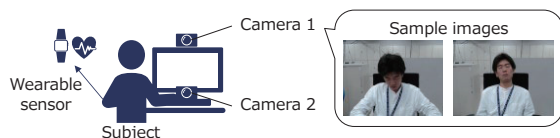


図7 検証用データの収集

表1 評価に用いたデータ

| | Dataset | # of subjects | Recording length | Contents |
|----------|--|---------------------------------|---------------------------------------|--|
| Train | DREAMS Subjects Database+ DREAMS Patients Database | 7 healthy subjects +12 patients | ~8hours/person, 19182 1-min. segments | Wearable sensor: PPG signals, ECG signals. |
| Validate | Data collected by ourselves | 25 healthy subjects | ~30min/person, 1581 1-min. segments | Wearable sensor: PPG signals, IBI Camera: face videos |

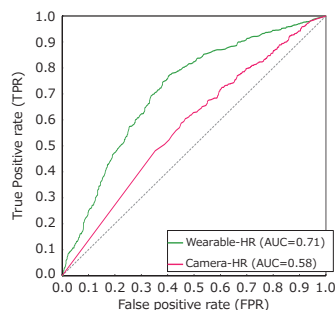


図8 ROCカーブ: Wearable HR vs Camera HR.

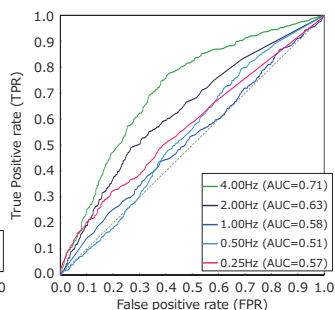


図9 ROCカーブ: Wearable HR における時間解像度 (0.25/0.5/1/2/4 Hz) の違い

Static HR filter と比較して狭い。人や状態(睡眠/覚醒)に応じて HR の分布が異なるため、その分布に応じて動的に算出することで周波数帯域が狭くなる。その周波数帯域の外側に存在するノイズを除去するため、帯域が狭い Dynamic HR filter の方がノイズの除去量が多くなる。さらに、Dynamic Wearable-HR filter の方が Dynamic Camera-HR filter より周波数帯域が狭い。Dynamic Camera-HR filter の周波数帯域は HR の推定誤差が含まれるためである。しかし、多くの場合に Dynamic Camera-HR filter の周波数帯域は Dynamic Wearable-HR filter の帯域を含んでおり、ノイズ除去量は劣るが、HR 情報は殆ど欠落しない。

図5に、図4の窓番号 $I=561$ の Camera PPG に対して、(b)ノイズ除去フィルタ無しの場合、(c) Static HR filter を掛けた場合、

(d) Dynamic Camera-HR filter の場合、(e) Dynamic Wearable-HR filter の場合を例示する。なお、参考のため、(a)には装着型センサで得た PPG 信号(Wearable PPG)を示す。(b)と(c)には多くのノイズが含まれるが、(d)と(e)は Dynamic HR filter により HR 情報を残しつつノイズが除去されている。

ノイズ除去後の Camera PPG を図3に示す CNN に入力して睡眠/覚醒を判別する。図3に示す CNN の構造は文献[Malik 2018]のモデル構造に従う。そのモデル構造は睡眠/覚醒判別に必要な HRV を適応的に抽出できるものである。文献[Malik 2018]との違いは、入力サイズとカーネルサイズを2倍にした点のみである。それは、同等の HR 情報を含むように、HR の時系列信号に対して、PPG 信号のサンプリングレートを2倍に設定したためである。HR に関する情報量は同じであるため、PPG 信号にノイズが含まなければ、文献[Malik 2018]と同等の精度で睡眠/覚醒を判別できると想定される。

3. 睡眠/覚醒判別評価

提案法の有効性を確認するため、CNN への入力を変えて、睡眠/覚醒の判別結果に対する AUC 指標を比較した。具体的には、以下三つの評価を行うことで、文献[Malik 2018]の Wearable HR を入力した場合(従来法)に対する、Camera PPG を入力した場合(提案法)の有効性を確認した。各評価での確認事項は⇒印で示す通りである。

評価1: Wearable HR vs Camera HR

⇒HR の時間解像度の低さが問題であること

評価2: HR vs PPG 信号

⇒PPG 信号が高時間解像度の HR 情報を含むこと

⇒PPG 信号がノイズを含み、それが問題になること

評価3: PPG 信号へのノイズ除去フィルタの適用

⇒Dynamic HR filter によりノイズ除去が可能なこと

⇒Camera PPG でも睡眠/覚醒判別が可能なこと

図6は評価1~3の関係図である。最終的に、図6の赤矢印の比較により、Camera PPG (提案法)でも、Wearable HR (従来法)と同等精度で睡眠/覚醒判別が可能なことを確認する。

3.1 評価に用いたデータ

表1に、評価に用いた学習データと検証データを示す。学習データは、公開データ(DREAMS Subject Database¹とDREAMS Patient Database²)から選択した計19名分である。それらには Wearable HR と Wearable PPG の両方が含まれる。Wearable HR 算出のための ECG 信号と Wearable PPG のオリジナルのサンプリングレートは200Hzであった。1分区切りの19182サンプルのうち覚醒(Positive)が4239サンプル、睡眠(Negative)が14942サンプルであった。

検証データは、我々が収集した。図7に示すように、装着型センサ(Empatica 社 E4)と上下の顔向き変動に対処するための二つのカメラ(640×480画素, 30fps)を用いた。収集は覚醒と睡眠の二つのステージを含む。覚醒ステージで、実験協力者は15分間スマートフォンを利用した。睡眠ステージの準備として単調な計算タスクを10分間行った。その後の40分間が睡眠ステージである。最後に、実験協力者は目覚めた時刻と、睡眠ステージ開始から何分ぐらいで眠りに落ちたかを質問紙に回答した。睡眠の平均時間は12.3分、標準偏差は6.6分であった。1分区切りの1581サンプルのうち覚醒(Positive)が898サンプル、睡眠(Negative)が683サンプルであった。

¹ <http://www.tcts.fpms.ac.be/~devuyt/Databases/DatabaseSubjects>

² <http://www.tcts.fpms.ac.be/~devuyt/Databases/DatabasePatients>

3.2 評価1: Wearable HR vs Camera HR

CNN への入力として, Wearable HR と Camera HR を比較した. 何れも 4Hz でサンプリングされた HR の時系列信号で, 窓幅は 1 分, つまり入力サイズは 240 で固定である. 睡眠/覚醒判別結果として, ROC(Receiver Operating Characteristic)カーブを図 8 に示す. Wearable HR の AUC は 0.71 で, Camera HR の 0.58 より高い. その理由として次の二つが考えられる. ① Camera HR は推定誤差が大きい. ② Camera HR の時間解像度が低い. まず理由①の確認のため, Wearable HR を正解として, Camera HR の推定誤差を算出した. RMSE は 11.1bpm, 相関は 0.49 であり, 比較的誤差が大きい.

理由②の確認のため, CNN の入力に含まれる HR 情報の時間解像度を低くして睡眠/覚醒判別の評価を行った. 具体的には, 4Hz の Wearable HR の算出において, 0.25/0.5/1/2 Hz でサンプリング後に, 4Hz でサンプリングした. 図 9 において, 最も高時間解像度である 4Hz の AUC が 0.71 と最も高く, 時間解像度が低いと AUC は低下する. 低時間解像度の HR 情報では, 睡眠/覚醒判別に有効な, HR と HRV の情報が欠落している.

3.3 評価2: HR vs PPG 信号

Camera HR では高時間解像度の HR 情報を得ることが困難なため, Camera PPG を CNN に入力し, Camera HR と比較した. Camera PPG のサンプリングレート 30Hz は顔映像のフレームレートと同じ 30fps だが, 比較する 4Hz の HR と情報量をそろえるため, 8Hz にダウンサンプリングした. CNN の入力サイズは 8Hz × 60 秒間で, 480 である. Wearable HR と Wearable PPG の比較も参考のために行った.

図 10(a)は装着型センサを, (b)はカメラを用いた場合の睡眠/覚醒判別結果である. (a)では Wearable HR の AUC(0.71)が Wearable PPG の AUC(0.62)より高いが, (b)では Camera PPG の AUC(0.66)が Camera HR の AUC(0.58)より高い. (b)の結果は高時間解像度の HR 情報が睡眠/覚醒判別において重要で, それが含まれる PPG 信号から抽出できることを示す. (a)の結果と, Camera PPG の AUC が Wearable AUC より低い原因としては, PPG 信号に含まれるノイズが問題と想定される. 次節で, ノイズ除去フィルタの効果を示す.

3.4 評価3: PPG 信号へのノイズ除去フィルタの適用

PPG 信号に対してノイズ除去フィルタを適用し, 睡眠/覚醒判別結果を比較した. 図 11 の(a)装着型センサと(b)カメラの両方で, ノイズ除去フィルタの適用による判別精度改善を確認できる. また, Dynamic HR filter は Static HR filter より判別精度の改善効果が高く, (a)と(b)の両方で, Wearable HR の AUC=0.71 とほぼ同じ精度を達成した.

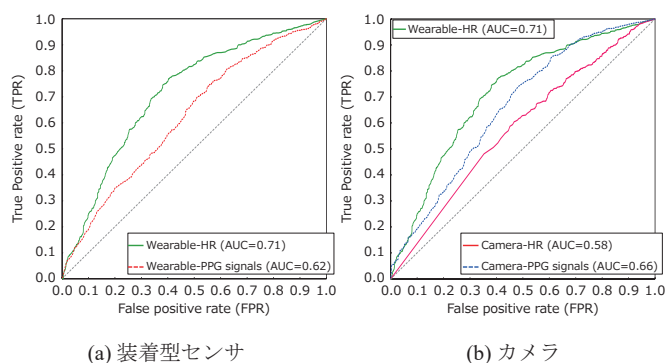


図10 ROC カーブ: HR vs PPG 信号

予想した通り, (b)において, Dynamic Wearble-HR filter の改善効果が Dynamic Camera-HR filter より少し高い. その理由は図 4 で示したように, Dynamic Wearble-HR filter の帯域幅が Dynamic Camera-HR filter より狭く, ノイズ除去効果が高いためである. Dynamic HR filter の改善効果が Static HR filter より高い理由も同じである.

4. まとめ

リモート PPG 信号(Camera PPG)による睡眠/覚醒判別方法を提案した. 提案法は, Camera HR と CNN を組み合わせた手法であり, まぶたの動きに基づく眠気推定方法と統合可能である. Camera HR を直接 CNN に入力する代わりに, Dynamic HR filter で処理した Camera PPG を CNN に入力する. それは, Camera HR の時間解像度の低さ, Camera PPG に含まれるノイズという二つの問題を解決する. 評価において, Dynamic HR filter が Static HR filter より効果的であり, 判別精度指標 AUC を 0.70 に改善できることを示した. それは, Wearable HR の AUC (0.71)と同等である.

参考文献

- [Oliveira 2018] L. Oliveira, et al.: Driver drowsiness detection : a comparison between intrusive and non-intrusive signal acquisition methods, 2018 7th European Workshop on Visual Information Processing (EUVIP), IEEE, 2018.
- [Tsujikawa 2018] M. Tsujikawa, et al.: Drowsiness Estimation from Low-Frame-Rate Facial Videos using Eyelid Variability Features, 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018.
- [Aktaruzzaman 2015] M. Aktaruzzaman et al.: The addition of entropy - based regularity parameters improves sleep stage classification based on heart rate variability, Medical & Biological Engineering & Computing, vol. 53, Issue 5, pp. 415-425, Springer, 2015.
- [Ye 2016] Yanqing Ye et al.: Automatic Sleep and Wake Classifier with Heart Rate and Pulse Oximetry : Derived Dynamic Time Warping Features and Logistic Model, 2016 Annu. IEEE Syst. Conf., IEEE, 2016.
- [Scherz 2017] W. D. Scherz et al.: Heart rate spectrum analysis for sleep quality detection, EURASIP Journal on Embedded Systems, EURASIP, 2017.
- [Malik 2018] J. Malik et al.: Sleep-wake classification via quantifying heart rate variability by convolutional neural network, Physiological Measurement, 29.8, 2018.
- [Somers 1993] V. K. Somers et al.: Sympathetic-Nerve Activity during Sleep in Normal Subjects, New England Journal of Medicine, 328.5, pp. 303-307, 1993.
- [Rahman 2015] T. Rahman et al.: DoppleSleep : A Contactless Unobtrusive Sleep Sensing System Using Short-Range Doppler Radar, Proceedings of the 2015 ACM international Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2015.
- [Wang 2018] C. Wang et al.: A Comparative Survey of Methods for Remote Heart Rate Detection From Frontal Face Videos, Frontiers in Bioengineering and Biotechnology, vol. 6, no. May, pp. 1-16, 2018.
- [Utkarsh 2018] S. Utkarsh et al.: Adaptive Heart rate estimation from face videos, The 8th Symposium on Biometrics, Recognition and Authentication, S2-9, IEICE, 2018.

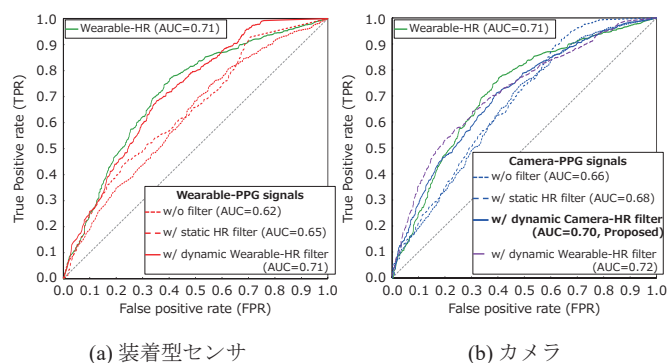


図11 ROC カーブ: PPG 信号に対するノイズ除去フィルタの効果