5W-01

画素の最適選択とトラッキング強化による顔映像からの心拍数測定精度の向上

小池　実輝哉†　　藤田　悟†

法政大学　情報科学部†

1. はじめに

新型コロナウイルス感染症(COVID-19)の流行に伴い，日常の健康管理の重要性が再認識されている．そこで，本研究では健康管理指標の1つである心拍数に着目し，顔動画から「非接触型」で心拍数を測定するシステムを提案する．

2. 心拍数測定

遠隔フォトプレチスモグラフィー(rPPG)は，外部の光源を必要とせずに，皮膚表皮下の色の小さな変化を解析し，非接触的に心拍数を測定する技術である．先行研究では，人間の顔の動画から毎フレームごとにRGB値を測定し，測定されたRGBを独立成分分析(ICA)[1]やCHROM法[2]などに適用して心拍に関する時系列信号を抽出する手法が研究されてきた．

また，人間の頸動脈の血流変化により，頭に微細な動きが生じる．その上下方向の頭の位置の周期変動を解析し，心拍数を推定する手法も提案されている[3]．

これらの「非接触型」の心拍数測定システムは，身体的影響(呼吸，姿勢の調整)や環境的影響(照明，影)によるノイズの影響を受けるため，測定精度に課題を残している．

3. ノイズに耐性のある心拍推定法

3.1 特徴点トラッキング

本研究では，顔の特徴点トラッキングを行うことで，RGB値の測定揺らぎを低減し，高い精度の心拍数推定の実現を目指す．まず，動画の1フレーム目を取り出し，haar-like特徴量を用いて顔の検出を行う．検出された顔領域には，顔だけでなく背景も含まれる場合が多いため，抽出領域を，縦幅は中央90%，横幅は中央50%に縮小して使用する．さらに口や目の領域付近は，まばたきなどの動きがノイズとなるため，領域の上端から高さ20%の領域と上端55%の位置から高さ20%の領域に分割する．その領域内において，肌と認識されたピクセルをランダムにN個選択し，Lucas-kanade法を用いてトラッキングし，各特徴点のRGB値の時間的変化を測定する．

3.2 特徴点の判別

肌表面の物理形状の影響により，顔と光源との角度が変化し，RGB値の変化量が大きい特徴点が含まれてしまうことがある．そこで，短いウィンドウを設定し，そのウィンドウ内でRGB値の変化が小さい特徴点の選別を行う．

i番目の特徴点において，フレームtから$t+1$の位置の移動を$\overrightarrow{d_i}(dx_i, dy_i)$と置いたとき，RGB値の変化量を正規化して式(1)に示す．

$$dC_i^{t\to t+1} = \begin{pmatrix} \dfrac{R_i^{t+1}(x+dx_i, y+dy_i) - R_i^t(x,y)}{R_i^{t+1}(x+dx_i, y+dy_i) + R_i^t(x,y)} \\ \dfrac{G_i^{t+1}(x+dx_i, y+dy_i) - G_i^t(x,y)}{G_i^{t+1}(x+dx_i, y+dy_i) + G_i^t(x,y)} \\ \dfrac{B_i^{t+1}(x+dx_i, y+dy_i) - B_i^t(x,y)}{B_i^{t+1}(x+dx_i, y+dy_i) + B_i^t(x,y)} \end{pmatrix} = \begin{pmatrix} dR_i^{t\to t+1} \\ dG_i^{t\to t+1} \\ dB_i^{t\to t+1} \end{pmatrix}. \quad (1)$$

$dR_i^{t\to t+1}$，$dG_i^{t\to t+1}$，$dB_i^{t\to t+1}$の値は，人種の違いに対して，[-0.02, 0.02]の範囲内に収まることが分かっている[4]．よって，閾値を設定し，それを超える範囲の変化量を持つ特徴点は，環境の影響を受けた特徴点として取り除きたい．しかし，環境に依存してこの閾値は変化するため，測定対象の特徴点数を確保するために，変化量の小さい特徴点から順にSN個選択し，他の特徴点を取り除くこととする．まず，フレーム数$a\cdot\frac{l}{2}$から$a\cdot\frac{l}{2}+l(a=0,1,2,\cdots)$において，変化量が最大となる値を$max\left(dR_i^{a\cdot l\to(a+1)\cdot l}\right)$，$max\left(dG_i^{a\cdot l\to(a+1)\cdot l}\right)$，$max\left(dB_i^{a\cdot l\to(a+1)\cdot l}\right)$とすると，これらの値の平均値を$max\left(dC_i^{a\cdot l\to(a+1)\cdot l}\right)$と示す．ここで，$max\left(dC_i^{a\cdot l\to(a+1)\cdot l}\right)$が小さい特徴点から$SN$個選択する．その選択された$SN$個の特徴点のRGBの平均をそれぞれ$\bar{R}^t$，$\bar{G}^t$，$\bar{B}^t$と表す．

Improvement of Heart Rate Monitoring from Face Video Using Effective Pixel Selection and Tracking

† Mikiya Koike, Satoru Fujita, Faculty of Computer and Information Sciences, Hosei University

Copyright ©2021 Information Processing Society of Japan.
All Rights Reserved.

3.3 正規化

3.2節で得られた信号$\bar{R}^t$を式(2)により正規化する．$ave()$は，その信号の平均値を表す．また，同様に$\bar{G}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}$，$\bar{B}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}$と定義する．

$$\bar{R}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}=\frac{\bar{R}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}}{ave\left(\bar{R}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}\right)}-1 \quad (a=0,1,2,\cdots). \tag{2}$$

3.4 CHROM法

RGB値の時系列信号から，血流が強く表れるクロミナンス信号を取り出す．クロミナンス信号は式(3)(4)(5)で定義される．

$$\bar{X}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}=3\bar{R}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}-2\bar{G}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}$$
$$\bar{Y}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}=1.5\bar{R}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}+\bar{G}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}-1.5\bar{B}_{normalize}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}. \tag{3}$$

次に，短い時間ウィンドウに対して，心拍に関する時系列信号を式(4)で求める．

$$\bar{P}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}=\bar{X}_f{}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}-s\bar{Y}_f{}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}. \tag{4}$$

$$s=\frac{\sigma\left(\bar{X}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}\right)}{\sigma\left(\bar{Y}^{a\cdot\frac{l}{2}\to a\cdot\frac{l}{2}+l}\right)}. \tag{5}$$

lとσは，スライディングウィンドウに対する長さと標準偏差を示す．また$\bar{X}_f$と$\bar{Y}_f$は，$\bar{X}$と$\bar{Y}$にバンドパスフィルタを適用したものである．本研究では，lを48と定義して行う．さらに，スライディングウィンドウから推定された$\bar{P}$トレースは，オーバーラップされ，ハニング窓を用いて追加していく．以上のCHROM法で推定された時系列信号をP_{chrom}とする．

3.5 最終処理

CHROM法は，白色灯の下で高い精度の血流測定を実現しているが，白色灯や太陽光以外の環境では，むしろ単純に緑の輝度を利用する方が高い精度を示す．そこで，心拍測定環境を考慮した最終的な心拍信号Pを式(6)により決定する．

$$P=\begin{cases}P_{chrom}, & \text{if 白色灯 or 太陽光}\\ \bar{G}_f, & \text{otherwise}\end{cases}. \tag{6}$$

式(6)において，$\bar{G}_f$は，$\bar{G}$にバンドパスフィルタを適用したものである．

以上の手順により求めた心拍信号PにFFTを適用して，心拍数を推定する．FFTを適用したときに，スペクトルの最大値における周波数をf_Pとすると，心拍数は式(7)で推定できる．

$$HR_{estm}=f_P\times 60. \tag{7}$$

表 1　実験結果(MAE)

	G	ICA	CHROM[2]	Head[3]	提案手法
顔領域を固定選択	25.44	26.79	16.56		
Lucas-kanade法	7.89	10.52	6.38	15.75	3.30

4. 実験

iPhone SE 第2世代のカメラで撮影された動画(解像度:4k，フレームレート:30fps)を用いる．光源は太陽光と室内の照明のいずれかの環境で，複数の男女で実験を行う．カメラと被験者の距離は，0.3メートルから0.5メートルの間で，被験者はカメラに顔の正面を向けて静かに座る状態にした．記録されるビデオの長さは，30秒から45秒とする．被験者は，シチズン電子血圧計(CH-602B)装着状態で撮影を行う．

トラッキングする特徴点の数は，2000個(N=2000)用意する．また，選択する変化量の小さい特徴点を200個(SN=200)に設定する．バンドパスフィルタは，バターワース型の設計で[0.8, 4.0]Hz([48, 240]bpm)の範囲とする．また，分解能を1bpm単位で推定するため，測定される信号すべてにゼロパディングを行う．

他の手法との精度の比較を表1に示す．本研究の提案手法は，MAEが3.30bpmで推定することができた．

5. まとめ

本研究では，顔の特徴点トラッキングや特徴点の選別を行うことで，身体的影響と環境的影響に対応したシステムを実現することができた．

謝辞

本研究はJSPS科研費JP20K11777の助成を受けたものです．

参考文献

[1] Christian Wiede, Julia Richter, Andre Apitzsch, Fajer KhairAldin and Gangolf Hirtz, “Remote Heart Rate Determination in RGB Data.”, ICPRAM(2016).
[2] G. de Haan and V. Jeanne, “Robust pulse rate from chrominance-based RPPG.”, IEEE Trans. Biomed. Eng(2013).
[3] Guha Balakrishnan, Fredo Durand, John Guttag, “Detecting Pulse from Head Motions in video.”, IEEE CVPR(2013).
[4] Wenjin Wang, Sander Stuijk, and Gerard de Haan, “Exploiting Spatial Redundancy of Image Sensor for Motion Robust rPPG.”, IEEE Trans. Biomed. Eng(2015).

Copyright ©2021 Information Processing Society of Japan.
All Rights Reserved.