RGB カメラにより撮影された動画像に基づく 非接触心拍計測による情動工学とその応用

Non-contact Heart Rate Measurement Based on Facial Video Taken by RGB Cameras and Its Applications

津村 徳道

千葉大学 大学院工学研究院
イメージング科学コース

〒263-8522
千葉市稲毛区弥生町 1-33

Norimichi TSUMURA
Department of Imaging Sciences, Graduate
School of Engineering, Chiba University
1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba,
263-8522 Japan

論文要旨：現代人は日々様々なストレスにさらされながら生活している。ストレスが慢性化すると体や心に悪影響が及ぶ危険があるため、それが大きな病気や事故につながる前に対処しなければならない。そこで、本稿では、生体光学に基づく解析により通常の RGB カメラで実現する。肌の RGB 画像からメラニン、ヘモグロビン色素と陰影成分を分離し、分離されたヘモグロビン色素成分画像を利用している。また、陰影成分も分離していることから、照明変動（陰影成分の変化）にロバストな計測を実現しているため、屋内や野外における実用的なシーンにおいて今回紹介する手法は活用可能である。

Abstract: We live under various stresses every day. Serious stress can have adverse effects on the body and mind and must be addressed before it leads to major illness or accidents. Therefore, in this paper, we realized by a normal RGB camera to extract hemoglobin pigment and shadow component from the RGB image of the skin, and the separated hemoglobin pigment component video is used as the pulse wave. Since the shadow component is also separated, robust measurement of lighting fluctuations (change in shadow component) is realized. This method can be used in practical scenes indoors and outdoors.

Key words: stress measurement, non-contact method, RGB video, pulse wave

1 はじめに

現代人は日々様々なストレスにさらされながら生活している。ストレスが慢性化すると体や心に悪影響が及ぶ危険があるため、それが大きな病気や事故につながる前に対処しなければならない。ストレスを評価する際に最もよく用いられる方法は質問紙による自己報告法である。ストレスの原因となるストレスorヤストレス反応を測定する質問紙が多く開発されている。しかし、これらに正確に回答するためには、回答者自身の正確な自己分析が必要になってくる上、回答者自身の意図で得点を

コントロールすることが可能なため、困難な面もある。さらに、その労力自体が回答者へ余計なストレスを加える原因になることもある。

そこで、心拍変動の周期から、交感神経系の活性のバランスを判断し、ストレス負荷の様子をみる事が可能であることから、カメラを用いた非接触の心拍変動計測システムが注目されている。この手法では非接触で行うため被験者に不快感を全く与えることなく計測が可能である。MIT のグループの研究では、通常の RGB カメラにシアン (C) とオレンジ (O) の2つの感度を追加した、5 バンドカメラを用いた心拍変動計測手法によるストレス計測が提案されている¹⁾。しかし、特殊な5 バンドカメラは現状実用性に欠けており、一般的な RGB カメラ

連絡者：津村 徳道
E-mail : tsumura@faculty.chiba-u.jp

を用いた手法の精度の向上が求められている。

そこで、本節では、5バンドカメラと同様の心拍変動計測に関する精度を、生体光学に基づく解析により通常のRGBカメラで実現する。肌のRGB画像からメラニン、ヘモグロビン色素と陰影成分を分離し、分離されたヘモグロビン色素成分画像を利用している。また、陰影成分も分離していることから、照明変動（陰影成分の変化）にロバストな計測を実現しているため、屋内や野外における実用的なシーンにおいて今回紹介する手法は活用可能である。

2 5バンドカメラを用いた非接触心拍計測法（従来法）

まず、先行研究における特殊な5バンドカメラを用いた非接触心拍計測手法¹⁾について述べる。

2・1 撮影環境

通常のRGBカメラにシアンとオレンジの2つのセンサを追加した5バンドのデジタル一眼レフ（DSLR）カメラを用いて、30 fpsで撮影を行う。距離は3.0 m、解像度は720×960に設定し、顔全体が含まれるように位置を調整する。5バンドカメラのRGBの感度は一般的なDSLRカメラと同じである。2分間撮影を行い、その間被験者にはカメラ正面を向いて静止して頂く。照明に関する制限は設けず、一般光源下で撮影を行う。

2・2 関心領域の決定

顔画像の連続写真（動画）から脈波を抽出する際に使用する関心領域（Region of interest : ROI）を決定する。撮影して得られたRAW画像をデモザイキング²⁾してRGBOC5つの画像と、カラー画像を取得。得られたカラー画像を用いて顔画像の特徴点抽出を行う³⁾。これにより、1枚の顔画像からFig. 1のように20個の特徴点が抽出できる。右目の外端から左目の外端までのx座標の距離をWとし、顎の特徴点から縦に2W、横にWの辺をもつ長方形を設定し、瞬きによる影響を受けないように目の周辺領域25%は除外する。この処理によって決定したROIをFig. 2に示す。

2・3 信号取得と前処理

2・2で決定したROI領域の画素の平均値をRGBOCの5バンドそれぞれについて算出し、時間軸に並べることによって2分間の信号を5つ得る。この5つの信号から脈波を抽出する。得られた信号のうちRバンドの信号をFig. 3に示す。

5つの信号は後にピーク検出を行いやすくするために傾き除去を行い⁴⁾、各信号をハミング窓を用いてバンドパスフィルタにかける。通過させる周波数は0.75 Hzから3 Hzまでとする。これは、心拍数にして45から180までの情報に制限することによって脈波以外の情報を可

Fig. 1 特徴点抽出結果

Fig. 2 ROIの決定

Fig. 3 領域の平均画素値の時間変化（R）

Fig. 4 傾き除去処理後の信号（R）

Fig. 5 バンドパスフィルタ適用後の信号（R）

能な限り除去するために行う。Fig. 3の信号に対して傾き除去を行った後の信号と、バンドパスフィルタを適用した後の信号をそれぞれFig. 4, Fig. 5に示す。

2・4 時間軸の独立成分分析

バンドパスフィルタにかけられた5つの信号に対して、時間軸に独立成分分析を適用して信号の分離を行う。独立成分分析のアルゴリズムはJADEを使用する⁶⁾。今回の場合は、ヘモグロビン量の変化による色変化や、顔の動きの変化、照明の変化などの原信号が混合し、RGBOCにおいてそれぞれ別の重み付けをされて記録されると仮定する。Fig. 6に独立成分分析によって分離された5つの信号のうち脈波と思われる信号を示す。

独立成分分析では、分離された5つの信号のどの信号に脈波が現れるかが決まっていない。Fig. 6に示すように、目視で脈波を判断することは可能だが、自動化する必要がある。そのため、得られた5つの信号のフーリエパワースペクトル解析を行い、0.75 Hz から 3 Hz の間に最も大きなパワースペクトルが現れている信号を脈拍だと判断することで自動化が可能である。

2・5 ピーク検出

ピーク検出には、山登り法を用いる。また、誤検出を防ぐために、ピークとピークの間隔が0.4秒以内の場合、より振幅が大きい方をピークとして記録するように実装した。また、独立成分分析後の信号は脈拍が反転して現れることがあるため、それを補正するプログラムとして、山のピークと谷のピークの平均値の絶対値を比較し、より大きい方を脈拍のピークとして記録する処理を行う。先行研究でこの手法による精度の向上が確認されている¹⁾。Fig. 7に、脈波ピーク検出の結果の例を示す。

Fig. 6 独立成分分析結果（脈波と思われる信号）

Fig. 7 脈波ピーク検出

2・6 精度検証

カメラの撮影と同時に、接触型の心拍計を用いて脈波を計測する。本研究ではポリグラフ（日本光電：RMT-1000）を用いて、心電図の計測を行い、得られた心電図の波形に対して、非接触式で利用した手法と同じ手法でピーク検出を行った。また、誤検出や検出漏れがないように、目視で波形とピークを確認し、閾値などの調整の上で補正を行った。得られた脈波波形と、ピーク検出の結果を示す。

カメラで計測される脈波は、心臓が拍動してから顔に血流が届くまでにタイムラグがあるため、波形のピークが記録される時間の誤差を精度検証に用いることができない。そのため、ピークとピークの間隔（脈波間隔）から求められる指標を、精度検証に利用する。検証の指標には、心拍数（HR）と心拍変動スペクトログラムを用いる。心拍数は、脈波間隔の平均値を60から割ることによって求めることができる。

2・7 心拍変動スペクトログラム

心拍変動スペクトログラムは、心拍変動周波数の時間軸の変化を可視化したものである。60秒間を1つの区間としてその間の変動周波数を解析し、その区間を1秒ずつずらして計60回解析を行うことによって、変動周波数の時間変化を捉えることが可能となる。変動周波数の解析には、Lomb-ピリオドグラム⁷⁾を使用する。変動周波数に影響を与える要因は大きく分けて2つある。1つは呼吸変動、もう1つは血圧変動である。呼吸変動は、リラックスしている際に0.15~0.4 Hzの間に現れる。血圧変動はストレスを感じている際もリラックスしているときにも0.05~0.15 Hzの間に現れる。Fig. 8に一般的な心拍変動スペクトログラムと神経活性の関係を示す。Fig. 8の右上の心拍変動スペクトログラムのように、呼吸変動が大きく観測される場合には、副交感神経が活性化し、リラックスしている状態だと判断できる。逆に、Fig. 8の右下の心拍変動スペクトログラムのように呼吸変動が観測されず、血圧変動のみが現れている場合は、交感神経が活性化しており、ストレス負荷がかかっている状態だと判断できる。このように、呼吸変動と血圧変動のバランスによって、心拍変動スペクトログラムからストレス状態を観測することができる。

3 RGB カメラを用いた色素成分分離に基づく非接触心拍計測法（提案法）

ここで紹介する提案方法は、従来手法によってカメラセンサが認識している脈波が、顔表面のヘモグロビン濃度の変化によるものと仮定し、ヘモグロビン成分を効率的に計測する。津村らの手法⁸⁾によって顔画像の色

Fig. 8 一般的な心拍変動スペクトログラムと神経活性の関係

Fig. 9 皮膚モデルおよび皮膚に入射する光の動き

素成分分離を行い、ヘモグロビン成分画像を取得することができる。そのヘモグロビン成分画像の濃度時間変化を脈波として検出する。

3・1 色素成分分離手法

顔画像に対して独立成分分析の技術を適用するために、観測信号である画像のRGB値と皮膚の色素濃度との関係を明らかにする⁸⁾。

人間の皮膚のモデルを Fig. 9 に示す。人間の皮膚は多層構造をとっており、大きく表皮、真皮、皮下組織に分けられる。皮膚にはメラニン、ヘモグロビン、ビリルビンなど様々な色素が含まれ、皮膚の色調はこれらの色素により左右される。これらの色素の中でもメラニン、ヘモグロビンによる変化が皮膚の色調に大きく影響している。メラニン色素は表皮に、ヘモグロビン色素は真皮に毛細血管が張り巡らされているため真皮に多く存在している。従って、表皮をメラニン層、真皮をヘモグロビン

層と仮定することができる。さらに、メラニン色素とヘモグロビン色素は、空間的に独立の分布をしていることが知られている。

皮膚に入射する光は、皮膚表面で反射される表面反射光と、皮膚内部に入射した上で内部で散乱を繰り返した後に皮膚外に射出した内部反射光に分かれる。表面反射光は皮膚の色素に影響されず、光沢等の光源の色を表現する反射光である。一方、内部反射光は皮膚内部の色素により吸収や散乱を繰り返し、皮膚外に射出した光であるため、皮膚の色を表現する反射光である。

以上の仮定を基に観測信号である画像のRGB値と皮膚モデルの関係式が作られる。本研究では、皮膚の色を示す内部反射光のみを観測信号として扱うため、Fig. 9 に示すように、小島らが提案したアルゴリズムに従い、カメラと光源の前部分に偏光板を設置し、偏光板がそれぞれ直交になるように設置することで表面反射光を除去した、内部反射光のみの画像を撮影する⁹⁾。

3・2 空間軸の独立成分分析

メラニン色素は表皮、ヘモグロビン色素は真皮に個別に存在していることから、2つの色素は空間的に独立に分布していると仮定することができる。さらに、皮膚画像の各チャンネルの画素値において Modified Lambert-Beer 法則が成立すると仮定すると、画素値の対数を取り、画像空間から濃度空間へ変換することで、線形の関係式を得ることができる。以上より、メラニン色素とヘモグロビン色素を独立信号とみなすことにより、Fig. 10 に示すように、独立成分分析により RGB 画像（画素値

Fig. 10 皮膚モデルに対する独立成分分析の流れ

Fig. 11 内部反射画像

v_1, v_2, v_3) からヘモグロビン画像, メラニン画像を分離することができる。

3・3 顔画像に対する色素成分分離

独立成分分析による皮膚の色素成分分離を, 撮影した顔画像に対して実行した結果を示す。撮影した顔画像を Fig. 11 に示す。画像は, 表面反射光が除去された内部反射画像である。メラニン色素成分色ベクトル, ヘモグロビン色素成分色ベクトルは顔画像において皮膚の凹凸による陰影のない小領域から推定する。そして, 推定し

Fig. 13 ヘモグロビン画像平均画素値の時間変化

Fig. 14 フィルタ処理後の信号

た色素成分色ベクトルを用いて顔画像に色素成分分離を適用する。顔画像の画素値を色素成分色ベクトルに射影し, 色素成分分離を実行した結果を Fig. 12 に示す。

3・4 脈波検出方法

上で述べた色素成分の分離処理を, 撮影した RGB 画像全てに対して実行する。分離されたヘモグロビン成分画像における, ROI の画素平均値の時間変化を取得する。2・2 で述べた手法と同様に, RGB カラー画像から特徴点抽出を行い, 額と口回りの領域を決定する。そし

(a) ヘモグロビン色素画像 (b) メラニン色素画像 (c) 陰影画像

Fig. 12 独立成分分析による色素成分分離の結果

Fig. 15 ヘモグロビン色素分離により取得した心拍変動スペクトログラム

て、ヘモグロビン成分画像のROIに対応する平均画素値の時間軸の変化を信号として記録する。Fig. 13に取得した信号を示す。さらに、先に述べたバンドパスフィルタと同様の処理を適用し、得られた信号をFig. 14に示す。この処理によって取得した信号に対してピーク検出を行い、得られた脈波間隔から心拍数と心拍変動スペクトログラムを計算する。

4 実験 (RGB カメラを用いた提案法の有効性の検証)

4・1 実験手法

20代の男子学生4名に対しストレス負荷実験を行った。5バンドカメラとポリグラフを用いて同時に計測を行い、得られた心拍変動スペクトログラムを比較して精度を検証する。実験は各被験者に対して安静状態とストレス状態の計2回行う。撮影は3mの距離から5バンドのデジタル一眼レフカメラを用いて行った。色素成分分離の結果とも比較を行うため、人工太陽灯を使用し、暗室下で撮影を行った。表面反射を除去するため、カメラと人工太陽灯の前に偏光板を設置している。

1回目に安静状態での撮影を行った後、2回目のストレス状態の撮影では被験者に暗算を行って頂いた。手や顔は動かさずに、4000から7を連続して減算し続ける、という作業を全員に共通して行って頂いた。4桁を毎回頭の中で復唱する人、下2桁のみで暗算を行う人などいたが、どの被験者もこの作業にストレスを感じたと実験後に話した。

4・2 実験結果

4・2・1 色素成分分離による脈波の抽出

先に紹介した色素成分分離を、5バンドカメラから取得したRGBカラー画像(bmp形式)に対して適用し、脈波を取得した。拍動間隔から心拍変動スペクトログラムを算出した。取得した心拍変動スペクトログラムの一部をFig. 15に示す。

4・3 考察

本研究を行うにあたって、先行研究のMcDuffらの手法に倣い5バンドカメラを使用した。McDuffらの手法と大きく異なる点がある。McDuffらが5バンドカメラの5つの信号のうち、グリーン(G)、シアン(C)、オレンジ(O)の3つの信号を用いて時間軸の独立成分分析を行うことによって心拍変動を精度良く取得することができていた。しかし、RGBの信号のみでは心拍変動スペクトログラムに関して、ストレス状態と安静状態を判定するのに十分な精度を得ることができていなかった。

しかし、今回提案する手法では、Fig. 15に示すように実用的なRGB画像から、ストレス状態と安静状態を

Fig. 16 情動喚起映像を視聴する被験者の顔画像の撮影環境

判定するのに十分な精度で、心拍変動スペクトログラムを得ることができた。また、提案する手法では、陰影成分を除去していることから、屋内や屋外における照明変動にロバストに計測することができるため様々な分野への応用が可能である。下の紹介する情動(感情)のモニタリングもその応用の1つの例である。

5 情動(感情)のモニタリング

より良い製品やサービスを生み出していくためには、消費者がそのどのように感じているかを調査し、改善していくことが非常に重要である。調査の際に消費者が本来感じたものとは異なる回答を行う可能性があるが、消費者に評価されていることを感じさせずに感情を推定可能であれば、正確な評価を得ることができる。生理的反応を伴う感情は情動と呼ばれ、レーザー計測機器を用いて生理的反応を計測することで感情の推定は可能である¹¹⁾。しかし、レーザー計測機器は一般的に普及していないため実用的ではない。そこで本研究では、上で紹介したように、RGBカメラを用いて撮影した顔の動画画像に色素成分分離を適用し生体情報を計測することで、実用的で正確な感情の推定を行う。

5・1 実験

20代の学生7名に対して、amusement, anger, disgust, fear, sadness, surpriseの6つの情動を喚起する映像¹⁰⁾を視聴させた際の顔の動画画像をRGBカメラにより撮影する。撮影環境をFig. 16に示す。撮影した顔の動画画像に対し色素成分分離を行うことで、ヘモグロビン色素画像を作成する。Fig. 17に示すヘモグロビン色素画像の平均画素値の時間変化は、顔面皮膚血流の変化量を表す信号と考えられるため、信号から脈拍間隔を検出することができる。

検出した脈拍間隔を時間領域解析、周波数領域解析及

Fig. 17 計測した顔面皮膚血流の変化量

Fig. 18 全特徴量を用いた際の分類精度

び非線形解析することで25種類の特徴量を取得する。解析区間は平常状態として映像視聴前に30秒間、及び情動状態として映像視聴中に最も情動が喚起された場面から30秒間とする。また、関心領域を額及び頬に設定し、それぞれの領域の皮膚血流量も特徴量として平常状態を10秒間撮影した動画像群及び情動状態を10秒間撮影した動画像群から取得する。

情動状態の特徴量から平常状態の特徴量を減算して得た計27種類の特徴量の変化量を用いて、 k 近傍法による感情分類を行う。教師データは全特徴量から7割を10回ランダム選択し、残りの3割のデータを分類した際の分類成功率の平均を分類精度とする。また、分類精度向上のため、個別最適化法によって全27種類の特徴量から9種類の特徴量を選択し分類を行う。

5・2 実験結果

全特徴量を用いて6情動を最も分類精度が高い $k=4$ で分類した際の分類精度をFig. 18に示す。fearの分類精度が著しく低い、これは映像刺激によってfearを正しく喚起できなかったためと思われる。分類精度を上げるため、fearを除外し個別最適化法によって選択された、平均脈拍間隔、心拍数の標準偏差、フーリエ解析から得たHF、AR法から得たLF及びLF/HF、頬の血流変化量、pNN50、TINN、SD2の9種類の特徴量を用いて $k=4$ とした際の分類精度をFig. 19に示す。94%の精度で5情動を分類できていることがわかる。

Fig. 19 特徴選択後の分類精度

5・3 まとめと今後の課題

被験者に情動喚起映像を視聴させた前後の顔動画像に対し、色素成分分離を適用することで生体情報を取得し、それらを解析して得た特徴量を用いて k 近傍法による感情分類を行った。さらに、個別最適化法による特徴選択を行うことで分類精度が向上した。今後は、異なる情動喚起刺激の使用によるfearの分類精度向上及びプログラムの高速化によるリアルタイム処理を目指す。

6 おわりに

本研究では、実用的なRGB画像から、生体光学に基づく色素成分分離の手法を用いることによって、ヘモグロビン色素を分離した顔画像を作成し、顔表面の血流変化を的確に捉えた。捉えた顔表面の血流変化から、高い精度で心拍変動を計測することに成功した。また、計測された心拍変動波形などから有効な特徴量を抽出することにより、5種の情動(感情)を平均94%の精度で分類することができた。

今後の課題として、リアルタイム化を目指しプログラムの高速化を行うことがあげられる。また、ストレスや感情モニタリングの他に、異性に魅力を感じたときに心拍変動がどう変化するかなど、魅力と生体情報の関連を調べる研究など、多方面に応用研究を展開していく予定である。

文 献

- 1) McDuff, D.; Gontarek, S.; Picard, R.W. Improvements in remote cardiopulmonary measurement using a five band digital camera. *IEEE Trans. Biomed. Eng.* **61**, 2593-2601 (2014).
- 2) 紋野雄介, 田中正行, 奥富正敏. 単板撮像素子を用いたワンショット撮影によるマルチスペクトル画像生成. 画像の認識・理解シンポジウム (MIRU2011), 1042-1049 (2011). (Monno, Y. et al. Meeting on Image

- Recognition and Understanding (MIRU2011), 1042-1049 (2011))
- 3) Martinez, B.; Valstar, M.F.; Binefa, X.; Pantic, M. Local evidence aggregation for regression-based facial point detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 1149-1163 (2013).
- 4) Tarvainen, M.P.; Ranta-aho, P.O.; Karjalainen, P.A. An advanced detrending method with application to hrv analysis. Biomedical Engineering, *IEEE Transactions*, **49**, 172-175 (2002).
- 5) 赤穂昭太郎. カーネル多変量解析. 岩波書店 (2008). (Ako, S. *Kernel Multivariate Analysis: New Trends in Nonlinear Data Analysis*. Iwanami Shoten, Publishers (2008))
- 6) Cardoso, J.F.; Souloumiac, A. Blind beamforming for non-gaussian signals. *IEEE Proceedings F (Radar and Signal Processing)*, **140**, 362-370 (1993).
- 7) 正規化 Lomb ピリオドグラム
<http://www.mathworks.com/matlabcentral/fileexchange/22215-lomb-normalized-periodogram> (Lomb normalized periodogram)
- 8) Tsumura, N.; Ojima, N.; Sato, K. *et al.* Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Transactions on Graphics*, **22**, 770-779 (2003).
- 9) Ojima, N.; Minami, T.; Kawai, M. Transmittance measurement of cosmetic layer applied on skin by using processing. *Proceeding of The 3rd Scientific Conference of the Asian Societies of Cosmetic Scientists*, 114 (1997).
- 10) Sato, W.; Noguchi, M.; Yoshikawa, S. Emotion elicitation effect of films in a Japanese sample. *Soc. Behav. Personal.*, **35**, 863-874 (2007).
- 11) Kashima, H.; Hayashi, N. Basic taste stimuli elicit unique responses in facial skin blood flow. *PLoS ONE* **6**, e28236 (2011).