The University of Osaka
Institutional Knowledge Archive

Title	IoTのための匿名化人センシング
Author(s)	北島, 利浩
Citation	大阪大学, 2018, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/69608
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

IoTのための匿名化人センシング

平成30年3月

北島　利浩

IoTのための匿名化人センシング

博士（工学）論文提出先
大阪大学大学院基礎工学研究科

平成30年3月

北島　利浩

IoTのための匿名化人センシング

北島　利浩

内容梗概

IoT（Internet of Things）時代の到来により、家電製品に様々なセンサを取り付けて各製品でデータの送受信を行い、サービスを向上させる取り組みが開始されている。家電製品は人の状態に応じて、自動的にサービスを行うことが非常に効果的であるため、様々な場所において人センシングを行う必要がある。本研究におけるセンシングは存在と位置を計測することと定義する。人センシング装置としてはカメラが安価であるため、カメラを搭載する機器が増加している。カメラは得られる情報量が多く、センシング装置として有効な機器であるが、個人を特定する情報も撮影してしまい、人のプライバシが侵害されているという問題点がある。IoT 家電製品は常時インターネットに接続されているため、撮影された使用者の画像が意図せずに流出するという危険性をはらんでいる。

本研究は高い空間分解能と高い搭載性を実現するためにカメラを使用するが、使用者の匿名性を保ったまま、人の存在と位置の検出を行う。カメラの前にすりガラスを置いたり、レンズの位置を移動させ、画像を光学的にぼかす事により画像の鮮明さを低下させて、匿名化画像を作成する。本研究では顔の外見から個人が特定できない画像を匿名化画像と定義する。従来の人検出の画像処理で使用されている画像のエッジ特徴量や明度差特徴量等は、鮮明な画像に対しては有効であるが、不鮮明な画像に対しては効果的ではない。これに対し、本研究ではカメラ画像からの脈拍計測を人検出の要素技術として用いる。本研究の内容を以下の三つに分類し記述する。

(1) 非接触脈拍計測システム

脈拍は心臓が収縮することによって発生する動脈の拍動であり、動脈の血液には酸化ヘモグロビンが多く含まれている。酸化ヘモグロビンは緑色の光を多く吸収する性質がある。この性質をカメラで捉えることにより、非接触に脈拍を計測することが可能となる。顔画像の緑色の変化を観測することにより非接触に脈拍

計測を行う。非接触脈拍計測では原理検証のため、鮮明な画像を用いる。カメラ画像から顔を検出し、顔領域の画素値からRGBのG（緑）だけを抽出し、領域内の平均値を算出する。この平均値の時系列データに対して、平滑化処理、差分処理を行うことにより体動雑音を除去し、脈拍成分だけを抽出する。抽出された信号に対して周波数解析を行うことにより、脈拍数を算出する。従来の研究では、顔が動いた場合に体動雑音が分離できずに誤差が大きくなってしまう問題があったが、提案手法はカメラの前で顔を動かす実験において従来の研究よりも誤差が小さく、高い性能を示した。

(2) 脈拍による顔検出システム

匿名化画像を分割し、各領域に対して非接触脈拍計測の処理を行い、人の脈拍数に相当する値が算出された領域には人の顔が存在すると判定する。静止状態の人を対象として、既存の顔検出手法との性能評価を行った。既存の顔検出手法は匿名化画像では特徴を抽出することができず、低い検出率であったが、提案手法は、高い検出率と低い誤検出率を実現し、高い検出性能を示した。検出された顔位置の周辺を再探索して顔位置を補正することにより顔位置精度が向上し、真値との誤差が1度以下の高精度な位置検出性能を示した。

(3) 人追跡システム

動いている使用者を追跡できるように、パーティクルフィルタの枠組みを導入し、人追跡手法へと拡張を行った。尤度関数に動き成分と脈拍成分を用いることにより、静止状態と動いている状態の両方で追跡が可能となった。性能評価においては、既存の追跡手法である Mean shift 法と既存の顔検出手法を用いて比較し、提案手法の方が高い追跡率を示した。さらに提案手法は実環境で起こるような人が横たわる場面においても、人追跡が可能であり人の姿勢変化に対応できることを確認した。

本研究は匿名化画像から人検出と追跡が可能であり、且つ高精度な位置検出が可能であることを示した。また、提案手法が使用者の匿名性を保ちながら、高い空間分解能と高い搭載性を実現する人センシング手法であることを確認した。

キーワード

IoT、匿名、人検出、脈拍、カメラ画像

Anonymous human sensing for IoT

Toshihiro Kitajima

Abstract

Recent years have witnessed the dawning of the era of the IoT (Internet of Things), and a wide variety of machines and devices connect to the Internet for automated transmission and reception of myriad types of data with no need for human intervention. The electric appliances that we use at home every day are now equipped with various types of sensors and the capacity to send and receive data for the purpose of offering improved services. For household appliances, in order to offer customized services adapted to the particular status of individuals users, it is necessary to perform various types of sensing operations on people in the household. Sensing in this thesis is defined as measuring human presence and position. Cameras have become inexpensive instruments for human sensing and have been increasingly used for this purpose. Because cameras produce large quantities of information, they are powerful tools for sensing. However, because camera images contain information allowing individuals to be personally identified, their use poses risks of personal privacy violations. In addition, because IoT ready home appliances are connected to the Internet, camera captured images of individual users may be unintentionally leaked.

In this research, cameras are used to realize high spatial resolution and high mountability, and detect the existence and position of people while protecting the anonymity of users. Anonymous images are created by placing frosted glass in front of the camera, moving the position of the lens, and optically blurring the image to reduce the sharpness. Images which an individual can not identify from the appearance of a face are defined as anonymous images. Edge features and brightness difference features of images used in conventional human detection method are effective for sharp images, but are not effective for blurred images. In

contrast, the heart rate measurement from camera images is used as an element technology for human detection with this thesis. The contents of this thesis are categorized and described in the following three.

(1) Non-contact heart rate measurement system

Heart rate is a pulsation of the artery caused by contraction and dilation of the heart, and blood of the artery contains a lot of oxygenated hemoglobin. Oxygenated hemoglobin has the property of absorbing the green light. It is possible to measure the pulse without contact by measuring this property with camera images. Non-contact heart rate measurement is performed by capturing changes of green in face images. Clear images are used to verify the principle. A face region is detected from camera images, only G (green) of RGB is extracted from the pixel value of the face region, and the average value in the region is calculated. The body motion is removed and only the hear rate component is extracted by performing smoothing and difference processing on the time series data of the average value. The frequency analysis is performed on the extracted signal to calculate the heart rate. There was the problem that the body motion noise could not be separated and the error increased when the face moved in the conventional research, but the proposed method has smaller errors than the conventional research in the experiment, it showed high performance.

(2) Face detection system using heart rate

The anonymous image is divided, non-contact heart rate measurement is performed for each region, and it is determined that a human face exists in the region where the value corresponding to the heart rate of the human being is calculated. The performance evaluation of comparison with the conventional face detection method was performed for subjects in a stationary state. It was not possible to extract features in anonymous images for the conventional face detection method, and the result was the low detection rate, but the proposed method realized the high detection rate and the low false detection rate. The proposed method showed high detection performance. The face position was corrected by re-search around the detected face position, and the face position accuracy was improved. High accuracy position detection performance with errors of 1 degree or less was shown.

(3) Human tracking system

The particle filter framework was applied to the face detection method so that the moving user can be tracked, and it was expanded to the human tracking method. It became possible to trace in both the stationary state and the moving state by using both the motion component and the heart rate component in the likelihood function. In performance tests involving unclear images, the proposed method yields results superior to those obtained with the existing Mean shift method or with the existing face detection method. The proposed method allows successful tracking even in cases where the posture of the individual changes, such as when the person lies down, a situation that arises in real-world usage environments.

This thesis shows that human can be detected and tracked with anonymous images. Furthermore, human position can be detected accurately. In addition, it was confirmed that the proposed method is a human sensing method that realizes high spatial resolution and high mountability while maintaining the user's anonymity.

Keywords:

IoT, Anonymous, Human detection, Heart rate, Camera images

目 次

1 序論 **1**

1.1 プライバシ保護 1

1.2 IoT とプライバシ問題 2

1.3 プライバシ保護と匿名性 4

1.4 研究の目的 5

1.5 センシングシステム概要 7

1.6 論文の構成 9

2 人検出要素技術 **11**

2.1 人検出装置 11

2.2 プライバシを考慮した人検出 22

2.3 脈拍計測原理 23

2.4 非接触脈拍計測 25

2.5 非接触脈拍計測の意義 26

2.6 まとめ 29

3 非接触脈拍計測システム **31**

3.1 画像からの脈拍計測手法 31

3.2 脈拍計測手法の性能評価 35

3.3 まとめ 42

4 脈拍による顔検出システム **45**

4.1 脈拍を利用した顔検出手法 45

4.2 顔検出率の評価 50

4.3 顔位置検出手法 62

4.4 顔位置精度評価 64

4.5 まとめ 65

5 人追跡システム **67**
5.1 人追跡手法への拡張 67
5.2 人追跡手法の性能評価 75
5.3 匿名性検証 82
5.4 まとめ 88

6 結論 **91**
6.1 匿名化人センシングの実現 91
6.2 社会的展望 92

謝辞 **95**

参考文献 **97**

研究業績 **105**

図 目 次

1.1 人センシングを利用したIoT機器 3
1.2 カメラ付きIoT機器による使用者のプライバシ画像流出 4
1.3 テレビ製品のありたき姿 5
1.4 人検出装置特性分布図 6
1.5 レンズシフト撮影 7
1.6 光学フィルタ使用撮影 8
1.7 鮮明画像 8
1.8 匿名化画像 9

2.1 焦電センサ概要 12
2.2 焦電センサ外観 12
2.3 光電センサ概要 13
2.4 光電センサ外観 13
2.5 熱電対概要 14
2.6 サーモパイルセンサ外観 15
2.7 超音波センサ概要 16
2.8 圧電正効果 16
2.9 圧電逆効果 17
2.10 超音波センサ外観 17
2.11 電波センサ概要 18
2.12 電波センサ外観 18
2.13 サーモカメラ概要 19
2.14 熱画像 20
2.15 サーモカメラ外観 20
2.16 可視光カメラ概要 21
2.17 可視光カメラ外観 22
2.18 撮影後の匿名化加工と撮影時の匿名化 23

2.19 酸化ヘモグロビンの吸光係数 24
2.20 光の波長と色の相関 24
2.21 脈拍計測原理 25
2.22 顔画像からの脈拍計測概要 26
2.23 テレビ視聴者の脈拍計測 27
2.24 運動中の脈拍計測 28

3.1 検出された顔領域 31
3.2 差分画像 33
3.3 顔領域内の緑色の輝度値変化 34
3.4 差分処理により抽出された脈拍成分 35
3.5 DFT による周波数解析 35
3.6 Web カメラ 36
3.7 参照データ用心拍計 37
3.8 距離別性能評価　実験環境 38
3.9 距離別誤差 38
3.10 照度別撮影画像 39
3.11 照度別誤差 39
3.12 運動状態性能評価　実験環境 41
3.13 運動状態での脈拍計測誤差 41

4.1 鮮明画像 46
4.2 匿名化画像 46
4.3 領域分割画像 47
4.4 背景領域 (A) と顔領域 (B) 47
4.5 背景領域 (A) における緑色輝度値の変化 48
4.6 顔領域 (B) における緑色輝度値の変化 48
4.7 背景領域 (A) における DFT パワースペクトラム 49
4.8 顔領域 (B) における DFT パワースペクトラム 50
4.9 各領域における脈拍計測結果 50
4.10 RGB カメラ 51
4.11 焦点がずれた画像 52
4.12 サイクロンフィルタ 52

4.13 サイクロンフィルタ画像 52
4.14 ラジアルフィルタ 53
4.15 ラジアルフィルタ画像 53
4.16 顔検出における ROC 曲線 55
4.17 レンズ位置制御カメラ 56
4.18 焦点の合った画像 58
4.19 レンズ位置 0.3 mm 前方画像 58
4.20 レンズ位置 0.6 mm 前方画像 58
4.21 レンズ位置 0.9 mm 前方画像 59
4.22 0.0 mm 画像における ROC 曲線 60
4.23 0.3 mm 画像における ROC 曲線 60
4.24 0.6 mm 画像における ROC 曲線 61
4.25 0.9 mm 画像における ROC 曲線 61
4.26 顔検出結果 63
4.27 再探索領域 63
4.28 投票平面 64
4.29 顔位置精度計測結果 65

5.1 画像 $I_{t_{n-1}}$ 68
5.2 画像 I_{t_n} 68
5.3 差分画像 69
5.4 パーティクルフィルタのアルゴリズム概要 71
5.5 各パーティクルにおける探索範囲の一例 73
5.6 積分画像 75
5.7 性能評価テストの実験環境 76
5.8 被験者の動作 76
5.9 Mean shift 法概要 77
5.10 追跡位置精度計測結果 79
5.11 人追跡中の水平方向誤差の時間変化 80
5.12 人追跡中の垂直方向誤差の時間変化 80
5.13 人がソファに横たわるシーンにおける追跡結果 82
5.14 ガウシアンフィルタによる鮮明度変化 84
5.15 レンズ位置 0.0 mm 画像 85

5.16 レンズ位置 0.3 mm 画像 86
5.17 レンズ位置 0.0 mm 画像の鮮明度変化 86
5.18 鮮明度変化による類似度 87
5.19 焦点の合わない撮像システム 88

6.1 人検出装置特性分布図における提案手法 92
6.2 増加する IoT 機器 93

表 目 次

2.1 人検出装置特性 29

3.1 PC 仕様 36
3.2 脈拍計測 1 フレーム当たりの処理時間 42

4.1 RGB カメラ仕様 51
4.2 レンズ位置制御カメラ仕様 56

5.1 人追跡成功率 78
5.2 人追跡 1 フレーム当たりの処理時間 81
5.3 個人特定実験結果 84

第1章

序論

“小説なり映画なりがいかに芸術的価値においてみるべきものがあるとしても、そのことが当然にプライバシ侵害の違法性を阻却するものとは考えられない。それはプライバシの価値と芸術的価値の基準とは全く異質のものであり、法はそのいずれが優位に立つものとも決定できないからである。”

これは日本初のプライバシ裁判と言われる三島由紀夫著「宴のあと」をめぐる裁判で、石田哲一裁判長が判決で述べた言葉である [1]。「宴のあと」は、著者が実在するモデルの半生を描いた作品であるが、モデルとなった人物のプライバシを侵害したとして、著者である三島由紀夫が訴えられたのである。

筆者は企業でテレビ等の家電製品を開発する研究員である。上記の判決文は、小説と映画を家電製品に、芸術的価値を商品価値に置き換えても成立する。いくら商品価値が高い製品を開発できたとしても、その製品が使用者のプライバシを侵害してはならないと考える。商品価値とプライバシの価値は別であり、切り分けて考える必要がある。

本章では、1.1 節にて、人々のプライバシ保護行為について述べる。1.2 節ではセンシングの必要性と付随するプライバシ問題について論じ、1.3 節では、本研究で取り扱う匿名性について論じる。1.4 節では、本研究の目的を述べ、既存手法との差異を明確にする。1.5 節では、本研究の提案システムの概要について述べ、1.6 節では、本論文の構成について述べる。

1.1 プライバシ保護

プライバシとは私生活に関する情報を指す言葉であり、人は普段から自身のプライバシを守る行為を行っている。古くは新聞や雑誌の投稿欄にペンネームを使

用したり、ラジオにラジオネームを使用して投稿を行い、匿名性を確保していた。また、外側から内側が見えないプライバシガラスを車で使用したり、すりガラスを家の窓に設置して、外側から内側の私生活が見えない工夫をしている。また、SNS (Social networking service) に自身の写真や他人の写真を投稿する際には、写真を加工して個人を特定できなくする行いも、プライバシを守る行為の一つである。顔写真は一目で個人が特定される情報を持っているため、プライバシ保護の観点からは、厳重に管理することが求められている。しかしその一方で、カメラの小型化、低価格化が進み、様々な機器にカメラが搭載されており、顔画像が撮影される機会が増加している。

1.2 IoT とプライバシ問題

近年、自動車、電車、飛行機等の乗り物、テレビ、冷蔵庫等の家電、さらには医療機器等がネットワークに接続され、人の作業を介さず自動的に多種多様なデータを送受信する IoT（Internet of things）[2, 3] の時代が到来している。IoT はモノのインターネットと呼ばれ、パソコンやスマートフォンだけではなく、すべてのモノがインターネットに繋がる仕組みを指す。我々が普段使用している家電製品に様々なセンサを取り付けて、各製品でデータの送受信を行い、サービスを向上させる取り組みが開始されている [4, 5, 6, 7]。家電製品が常時インターネットに接続されていることにより、遠隔操作が可能になり、外出中に宅内の家電機器を操作することができる。さらに家電製品はプログラムを自動更新し、機器のソフトウェアを最新の状態に保つことができる。また、クラウドシステムを利用することにより、家電機器本体には搭載できない高性能な演算処理装置を用いて、音声認識や画像認識を行うことができる。さらに家電製品は人の状態に応じて、自動的にサービスを行うことにより利便性を向上させている。例えば、人に風を当てないように気流制御を行うエアコン、人の存在範囲にだけ光を照らす照明器具、人を避けて掃除を行うロボット掃除機等がある (図 1.1)。これらの機能を実現するには、屋内の人センシングを行う必要がある。本研究におけるセンシングとは存在と位置を検出することと定義する。

図 1.1 人センシングを利用した IoT 機器

人センシング装置としてはカメラが安価であり、且つ様々な情報をセンシングできるため、多くの場面でカメラが使用されている。カメラは得られる情報量が多く、センシング装置として有効な機器であるが、人の顔、服装、食事等の私生活に関する情報も撮影してしまい、人のプライバシが侵害されるという問題がある。IoT 時代の到来により家電機器がインターネットに接続されているため、撮影された使用者の画像が意図せずに流出してしまう可能性がある (図 1.2)。最新の家電製品はコンピュータと同様に OS (Operating system) を搭載している。コンピュータの OS には脆弱性があることが報告されているが、組込み機器の OS にも脆弱性は存在し、その脆弱性を狙われて、攻撃を受けたり侵入される事がある。外部から OS に侵入されてしまうと、家電製品に接続されているカメラの画像を外部に転送される可能性がある。その点を不安視し、インターネット接続があり、カメラが搭載されている製品の購買を控えている人も存在する。そのため使用者のプライバシを保護した状態で、人センシングが行える装置の開発が期待されている。

図 1.2 カメラ付き IoT 機器による使用者のプライバシ画像流出

1.3 プライバシ保護と匿名性

本研究で取り扱う匿名性について論じる。使用者のプライバシ画像が流出した際に、保護するべき最も大切な情報は何であるだろうかと考える。宅内の画像が流出することを想定すると、部屋にある家具やポスタ、雑誌等も私生活の情報であり、プライバシ情報となる。しかし、これらの情報は個人が特定されなければ、誰の部屋であるか、または誰の所有物であるかが分からないため、個人のプライバシ情報とならない。最も保護するべき情報は個人特定に関する情報である。使用者の服装や体型等は、特定する者が使用者のことを既知であれば、個人特定に繋がる可能性がある。しかし、特定者が使用者を既知でなければ、個人を特定することは困難である。個人特定に繋がる最も強力な情報は顔であると考える。マンションの玄関やスマートフォン等で顔認証システムが採用されており、顔は個人を特定する情報として広く使用されている。また、インターネット上に記録されている使用者の顔画像から、個人が特定されてしまう可能性も考えられる。それ故に顔画像は最も保護するべき対象であると考える。本研究における匿名化は顔画像から個人が特定されないように光学的な加工を行うことを指し、匿名性とは顔画像における匿名性を指す。本研究では、匿名性を保護することが使用者のプライバシを保護することに繋がると考え、顔画像の匿名性を保護することに主眼を置く。

1.4 研究の目的

本研究では人の存在だけではなく、人の位置まで検出することを目標とする。本研究における提案手法は、宅内のテレビで使用されることを想定しており、テレビを視聴している使用者を検出の対象としている。テレビはインターネット接続可能な製品が多数存在し、家電製品の中でプライバシ問題に最も直面している製品であると言える。図 1.3 にテレビのありたき姿を示す。常時インターネットに接続しているため、宅外から操作が可能であり、ソフトウェアの自動更新、クラウドサービスが可能である。また、テレビの将来的な技術として裸眼で 3 D映像が見える技術が開発されており、テレビへの搭載が検討されている。現状の裸眼 3 D技術は、視聴可能な視野角が狭く、視聴位置が限定されてしまう。そのため、テレビの前に居る使用者の位置に応じて、表示する映像を変える方法が検討されている [8]。さらに、指向性のあるスピーカを用いて、テレビを視聴している使用者にのみ音声を提供し、それ以外の人は音が聞こえないシステムも開発されている。加えて、位置検出のためのカメラ画像が流出しても、匿名性が保護されており、使用者に安心感を与える。視聴位置を限定しない裸眼 3 D技術と位置に応じた音声出力の実現に向けて、使用者の位置検出が重要な技術となる。

図 1.3 テレビ製品のありたき姿

図 1.4 は既存の人検出装置を特性毎に表示した分布図である。自動ドアや照明器具等の製品において使用され、世の中に普及している人検出装置を挙げた。縦軸は匿名性、横軸は空間分解能、楕円内の色は搭載性を示している。搭載性は製品適用度合を表しており、小型で安価なセンサは高い搭載性を示す。焦電センサ (Pyroelectric sensor)、光電センサ (Photoelectronic sensor)、超音波センサ (Ultrasonic sensor)、電波センサ (Radio wave sensor) は小型で安価であるため搭載性が高いが、人の存在しか検出できないため空間分解能が低い。サーモパイルアレイセンサ (Thermopile array sensor) は 8 × 8 や 16 × 16 の解像度しかなく、空間分解能が低い。サーモカメラ (Thermal camera) は 320 × 240 や 640 × 480 の解像度を有しているため空間分解能が高いが、高価であるため搭載性が低い。可視光カメラ (Visible light camera) は、高精細な解像度を有しており、空間分解能が高く安価であるが、使用者の顔を撮影してしまうため匿名性が低い。各検出装置の原理と特徴については、2.1 節で後述する。図 1.4 の右上部分を実現する人検出装置は現在存在しない。本研究は図 1.4 の匿名性と空間分解能が高い右上の点線で囲まれた部分の実現を目指す。つまり、使用者の匿名性を保持しながら、高い空間分解能と搭載性に優れた人センシング手法を実現する。

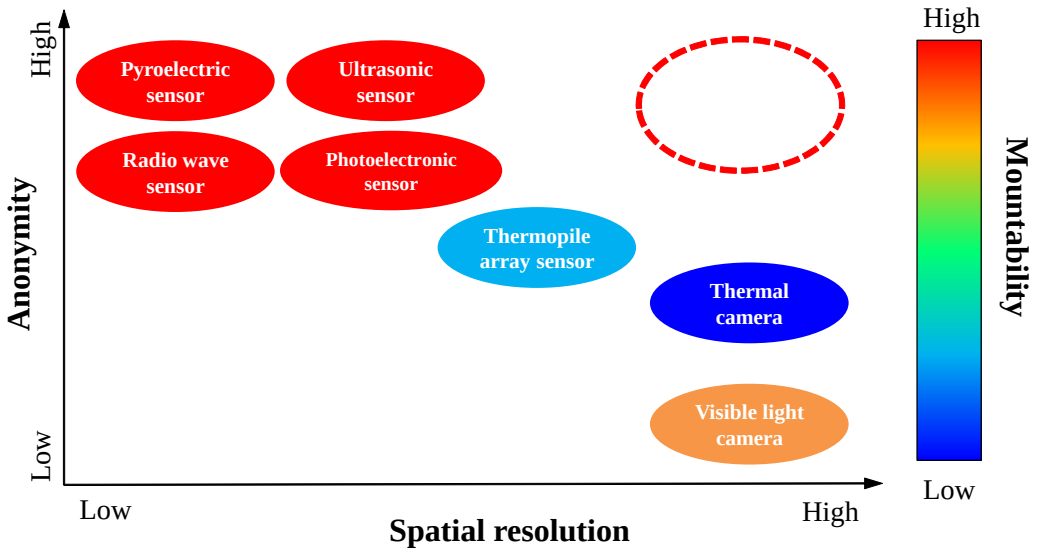


図 1.4 人検出装置特性分布図

1.5 センシングシステム概要

提案システムは高い空間分解能と高い搭載性を実現するために可視光カメラを使用する。しかし、そのまま撮影を行うと使用者の個人が特定されてしまうため、図 1.5、図 1.6 のように、レンズの位置を移動させたり、カメラの前に光学フィルタを設置し、個人の顔が特定できない程度まで画像の鮮明さを低下させる。図 1.7 は、焦点の合った鮮明画像である。被験者のプライバシ保護のため、顔部分に加工を施している。本論文では顔画像部分に目隠し処理やモザイク処理のある画像を記載するが、実験ではモザイク処理のない画像を使用する。図 1.8 はレンズ位置を移動させて不鮮明にした画像である。本研究では、このように個人が特定できない画像を匿名化画像と定義する。この匿名化画像から人検出を行う。従来の画像処理技術で使用されている画像のエッジ特徴量や明度差等は、鮮明な画像に対しては有効であるが、不鮮明な画像に対しては効果的ではない。そこで提案システムでは、生体信号である脈拍の情報を用いて人検出を行う。脈拍は人の肌に現れる生体信号であり、非接触に観測することにより、人検出に応用できると考える。また、脈拍は画像の見た目の特徴を示すエッジ特徴量や明度差等のアピアランス情報を必要としないため、不鮮明な画像においても効果的であると考える。脈拍を用いて画像上での人位置を検出することにより、高い空間分解能による位置検出が可能となる。これにより、使用者の匿名性を保持しながら、高い空間分解能と高い搭載性を実現する人検出システムとなる。

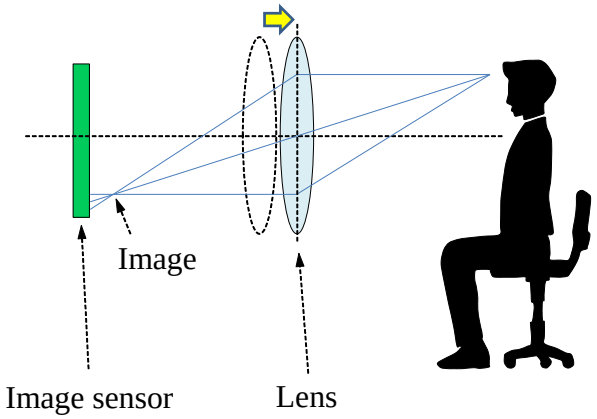


図 1.5 レンズシフト撮影

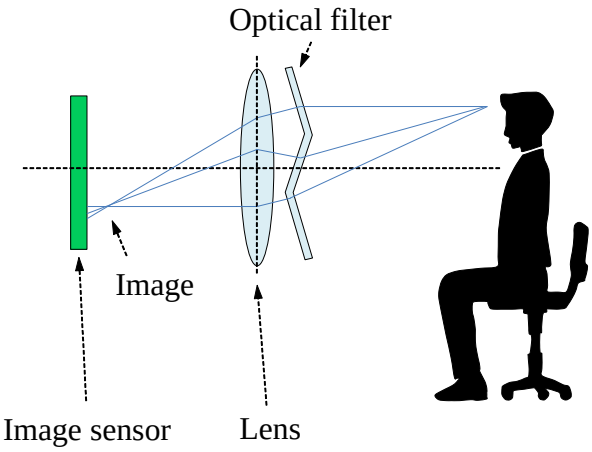


図 1.6 光学フィルタ使用撮影

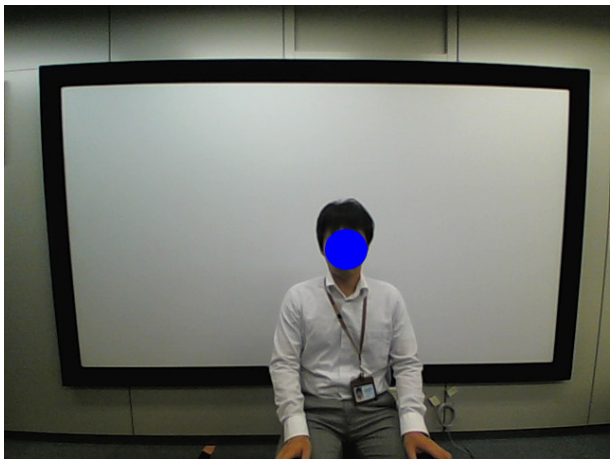


図 1.7 鮮明画像 (被験者のプライバシ保護のため顔部分に加工有り)

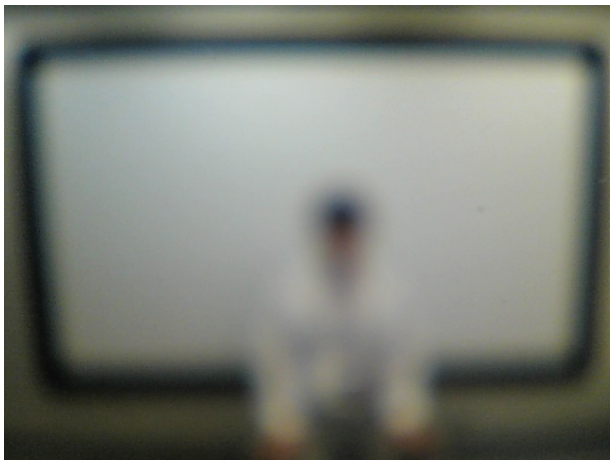


図 1.8 匿名化画像

1.6 論文の構成

本論文では、使用者の匿名性を保持しながら、高い空間分解能と高い搭載性を実現する人検出システムについて述べる。

1 章では、カメラ付き IoT 家電製品の問題点について述べ、プライバシ保護のための匿名化の重要性について論じた。既存の人検出装置の特性を図示し、提案システムの目指す方向とその概要を示した。

2 章では、人検出を行うための要素技術について述べる。既存の人検出装置の原理と特徴について報告し、本研究で使用する脈拍計測技術の原理について解説する。そして、カメラ画像を用いた脈拍計測手法の概要と問題点について述べる。

3 章では、人検出の要素技術となる非接触脈拍計測手法の提案を行う。非接触脈拍計測の意義について議論し、活用場面を示す。また、提案手法の性能評価を行い、その有効性を示す。

4 章では、3 章の脈拍計測手法を用いた顔検出の提案を行う。匿名化画像や鮮明度を変化させた画像における性能評価を行い、有効性を示す。顔位置検出手法についても提案を行い、位置精度の性能評価を行い、有効性を確認する。

5 章では、4 章にて提案した顔検出システムを人追跡システムへと拡張する。人追跡システムの各要素技術について述べ、性能評価によって、追跡率、処理時間において有効性を確認する。

6 章では、本研究の匿名化人センシングの手法と性能について述べ、今後の課題を導出する。さらに、本研究によってもたらされる社会貢献について議論を行う。

第2章

人検出要素技術

本章では、人検出の要素となる技術について述べる。2.1 節で既存の人検出装置とその仕組み、特徴について報告する。2.2 節では可視光カメラによるプライバシを考慮した人検出手法について紹介する。2.3 節で脈拍計測の原理について解説し、2.4 節では非接触に脈拍を計測する方法について説明を行い、非接触脈拍計測の課題について述べる。2.5 節で非接触脈拍計測の意義について述べ、2.6 節では、本章のまとめを行い、本研究に必要な人検出要素技術について述べる。

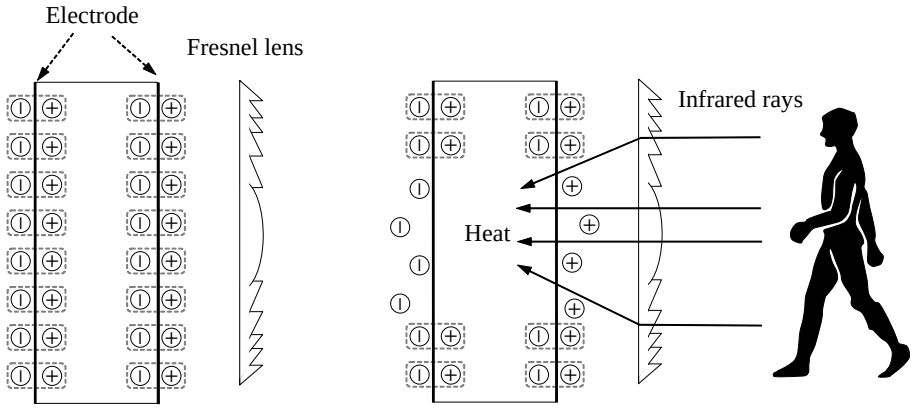
2.1 人検出装置

人検出装置は既に世の中に普及しており、自動ドアや照明器具等の製品に付属して使用されている。人検出装置には様々な種類があり、装置によって検出方式は異なる。以下に人検出装置とその仕組みについて記述する。

- **焦電センサ**

  人体から放射される赤外線を受けて検出を行う。図 2.1 に焦電センサの概要を示す。焦電センサは強誘電体セラミクス、ポリフッ化ビニリデン等の焦電効果を示す物質を電極で挟んだ構造をしている。図 2.1 (a) は熱源である人が計測範囲内に居ない状態を示している。焦電効果を有する物質は分極しており、浮遊電荷と結合しているため、通常は見かけ上の分極は見られない。しかし、この物体に赤外線が入射し、内部が温度上昇すると、分極が減少するにもかかわらず、浮遊電荷は電極表面に一時的に残り、電極間に一過性の電圧変化が現れる。この電圧変化を用いて人検出を行う。実際の焦電センサではフレネルレンズによって赤外線入射の方向性が確保さ

れている [9]。図 2.1 (b) では人が計測範囲内を通過したため、人体から放射される赤外線により、電極間に電圧変化が現れている。焦電センサはセンサの前を通過した人、動いている人のみを検出できる。人の位置は特定できず、存在のみの検出となる。図 2.2 に焦電センサの外観を示す。センサが安価に製造でき、匿名性が高いことが特徴であり、防犯や照明の自動点灯・消灯に頻繁に使用されている [10, 11, 12, 13]。



(a) 赤外線入射無し　　(b) 赤外線入射有り

図 2.1 焦電センサ概要 [9]

図 2.2 焦電センサ外観

- **光電センサ**

発光素子によって投光され、発射された光が人体に反射し、その反射光が受光素子に到達することにより人検出を行う (図 2.3)。光電センサの前に人が居ない場合は、投光した光は受光素子に戻ってこないが、光電センサの

前に人が居る場合は、投光した光が人体に反射し、その反射光が受光素子に入り、受光量が増加する。この受光量の増加を捉えることにより検出を行う。検出物体の表面状態によって光の反射光量、検出安定性が変わるという特徴がある。人がセータ等の凹凸のある衣服を着用している場合には、光が衣服の表面で乱反射し検出されない場合もある。投光に使用される光には可視光と赤外光があるが、大多数は目には見えない赤外光が使用されている。測定範囲内に物体があれば検出する仕組みのため、人以外の物でも検出する。測定範囲内に人が出現するシナリオを利用して人検出装置として使用されている。光電センサはトイレの自動洗浄や手をかざすと水が出てくる自動水栓等に使用されている。図 2.4 に光電センサの外観を示す。安価であり、匿名性は高いが、測定範囲内の物体の存在しか検出できない。

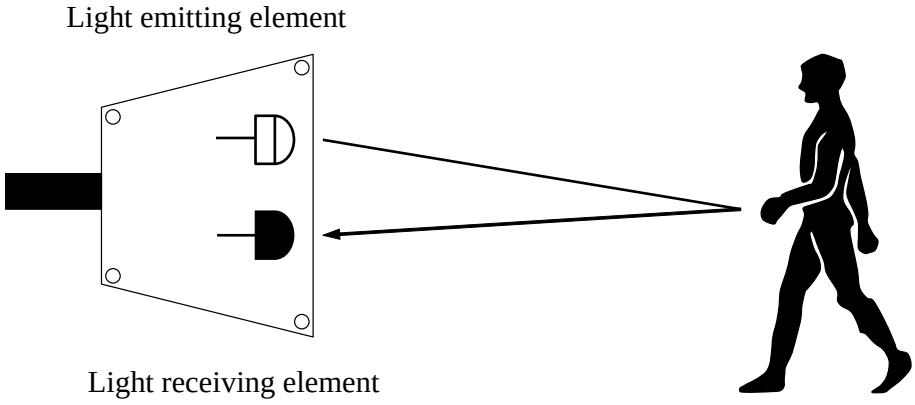


図 2.3 光電センサ概要 [14]

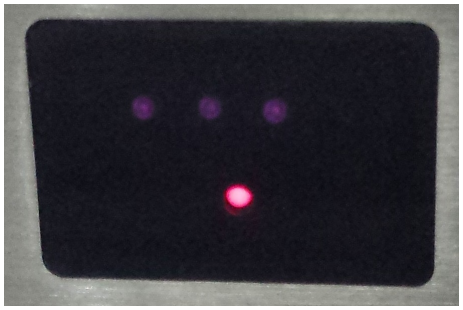


図 2.4 光電センサ外観

- **サーモパイルセンサ**

  赤外線の入射エネルギ量に比例した熱起電力が得られ、その電力を基に人検出を行う。図 2.5 に示すように、二種の異なる金属線 (Ma, Mb) の両端を接続し、その二つの接点間に温度差を与えるとゼーベック効果により、熱起電流が発生する。この原理を利用した温度センサが熱電対である。感度を向上させるため、この熱電対を 100 個程度、ケース内に直列に設置し、赤外線による温接合部の温度上昇を感度よく検知できるようにした装置がサーモパイルセンサである。図 2.6 にサーモパイルセンサの外観を示す。このサーモパイルセンサは動作原理上、特にセンサ自体は電力を必要としないため、小型、軽量、且つ長時間計測に向いているセンサである [9]。焦電センサと異なり、静止した人物も検出可能であることが特徴である。また、サーモパイルセンサ一つでは検出範囲内の人の存在しか分からないため解像度が低いが、複数個を並べてサーモパイルアレイセンサとすることにより、8 × 8 や 16 × 16 の解像度を実現することができる。オフィスや工場等の監視装置として利用されている [15, 16, 17]。現状では、解像度を向上させることが難しく、高い空間分解能が期待できない。匿名性は確保されているが、人検出装置として安価ではない。

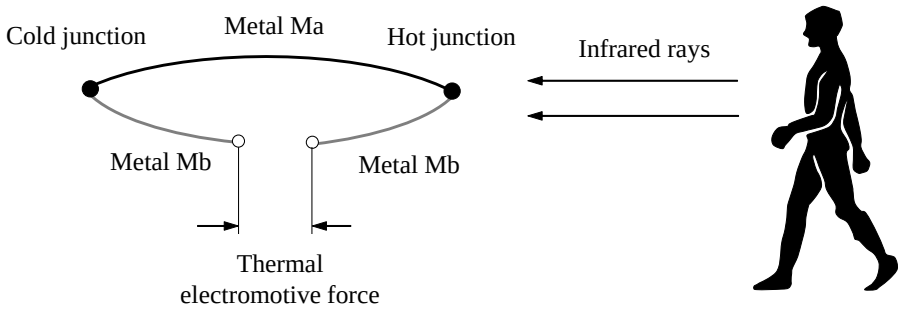


図 2.5 熱電対概要 [9]

図 2.6 サーモパイルセンサ外観

- **超音波センサ**

送波器により超音波を対象物に向け送信し、その反射波を受信することにより、対象物の有無や対象物までの距離を検出する (図 2.7)。超音波センサでは周波数 40 kHz が頻繁に使用され、長い距離を測定するには低い周波数、短い距離を正確に測るには高い周波数が使用される。超音波の送信と受信には、圧電セラミクスを使用している。圧電セラミクスは高純度な粉体（酸化チタン・酸化バリウム等）を高温度で焼き固めた多結晶体セラミクスである。図 2.8 に示すように、圧電セラミクスに外部から力を加えると電圧が発生する。これは圧電正効果と言われ、振動や音を電気信号に変えることができる。また、図 2.9 (a) に示すように、逆に電圧を印加すると極性に応じて圧電セラミクスが伸縮する。これは圧電逆効果と言われ、圧電セラミクスに電気信号を加えると振動や音を出すことができる [18]。図 2.9 (b) は、高周波の交流電圧を加えると、圧電セラミクスは超音波の発振子となることを示している。超音波を送信する際には、圧電逆効果を用いて超音波を発信し、対象物で反射した波を圧電正効果により電気信号に変えて受信を行う。超音波センサは検出対象物体の色や表面状態に影響されないという特長があり、透明体の検出が可能である。光の変化量で判断する光電センサと違い、汚れや埃による影響を受けないという性質もある。超音波の発信から受信までの時間を計測することにより、対象物までの距離を算出することができる。自動車の障害物センサや、ロボット掃除機の障害物センサとして使用されている。図 2.10 に超音波センサの外観を示す。自然界では、コウモリやイルカ、クジラが超音波を出して障害物を検出し、視覚の代わりとしている。検出した物体が人であるか物であるかの判別は困

難であるため、人検出装置としてよりも障害物検出装置として使用されることが多い。安価であり、匿名性は確保されているが、測定範囲内の物体の存在しか検出できず、空間分解能が低い。

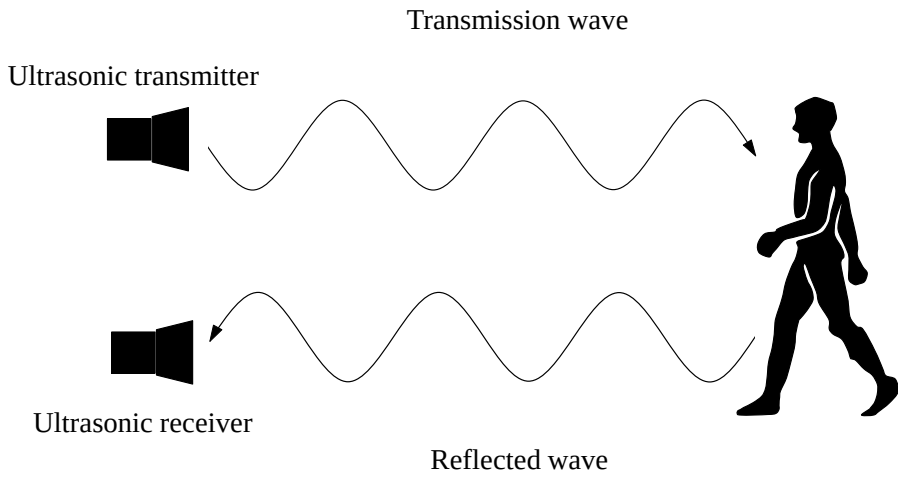


図 2.7 超音波センサ概要

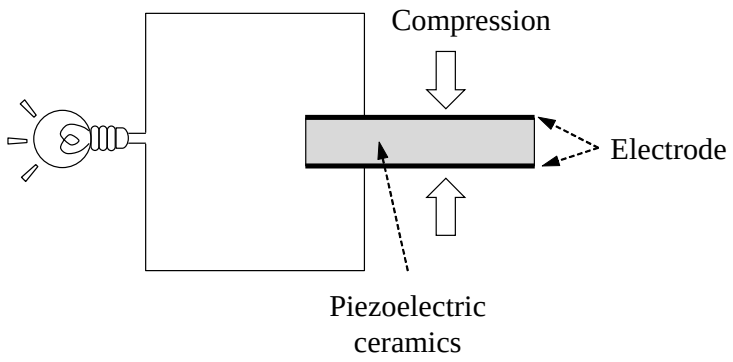
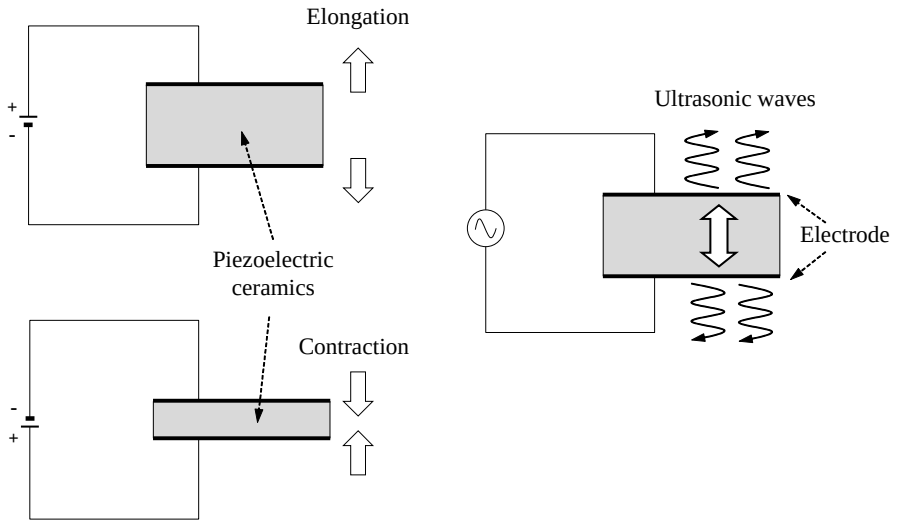


図 2.8 圧電正効果 [18]



(a) 直流回路　　(b) 交流回路

図 2.9 圧電逆効果 [18]

図 2.10 超音波センサ外観

- **電波センサ**

10.5 GHz や 24 GHz のマイクロ波と呼ばれる電波を送信アンテナから発信し、空間内の物体で反射した電波を受信アンテナにて受信する。図 2.11 に電波センサの概要を示す。電波を反射させた物体が移動中の場合には、発信した電波と受信した電波との間に物体の移動速度に応じた周波数の変化が生じる。この現象をドプラ効果といい、電波センサはこのドプラ効果を

利用して検出を行っている。人に反射した電波の周波数と発射した電波の周波数を比較し、変化があった場合には動いている人がいると判断する。電波は各国の電波法によって使用が管理されている。日本の電波法では、10.5 GHz 帯は屋外では使用できないが、24 GHz 帯は屋内と屋外で使用できる。現状では国や場所によって、センサの発する電波の周波数帯が使用できない場合がある。電波センサは検出範囲が広く、金属以外の物体は透過するため、屋内での人検出、見守りセンサとして使用されている [19]。図 2.12 に電波センサの外観を示す。測定範囲が広く、匿名性が確保されており、搭載性が高いという特徴があるが、測定範囲内の移動体の存在しか検出できず、空間分解能が低い。

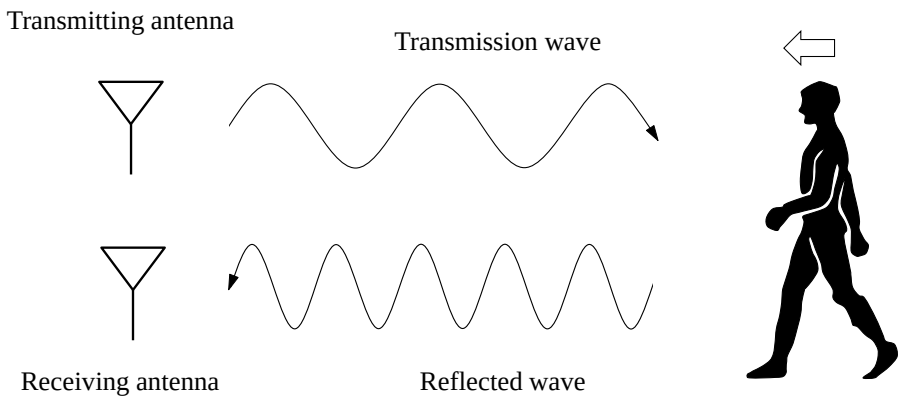


図 2.11 電波センサ概要

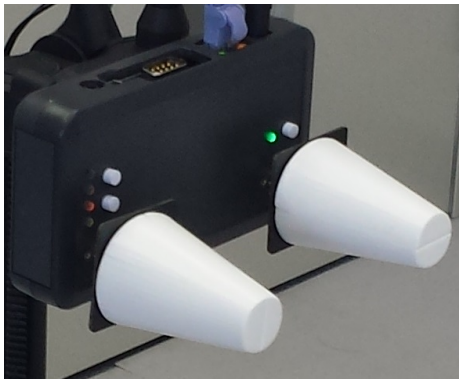


図 2.12 電波センサ外観

- **サーモカメラ**

  対象物から出ている赤外線放射エネルギを電気信号に変換し、可視化する装置である。一般的に受光素子としてはマイクロボロメータが使用される。マイクロボロメータは冷却が不要な小型の赤外線センサで、赤外線が入射するとボロメータ抵抗が熱せられて抵抗値が変化し、この変化を電圧に変換して読み出しを行う。マイクロボロメータの周囲に空気等のガス分子が存在すると、マイクロボロメータで吸収された熱がガス分子に伝わり、センサの感度が低下してしまうため、マイクロボロメータは図 2.13 のように真空パッケージに封止される [20]。また、レンズには、2 μm - 14 μm の波長域で透過率の高いゲルマニウムが材料として使用される。マイクロボロメータ、ゲルマニウムはセンサの材料としては高価であり、製造過程でも真空封止が必要な事から、サーモカメラは装置の価格が高いという欠点がある。しかし、可視光の波長に影響を受けないため、夜間の監視カメラや、自動車の歩行者検出に利用されている [21, 22, 23]。図 2.14 にサーモカメラの画像を示す。被写体の表面温度を画素毎に計測することができ、温度が高い部分は赤色で、背景の温度が低い部分は青色で表示されている。人体と周囲の背景には温度差があるため、この温度差を利用して人検出を行う。図 2.15 にサーモカメラの外観を示す。320 × 240 や 640 × 480 の解像度を有しているため空間分解能は高いが、顔の形状は撮影されるため、匿名性は高くなく、また、高価であるため搭載性が低い。

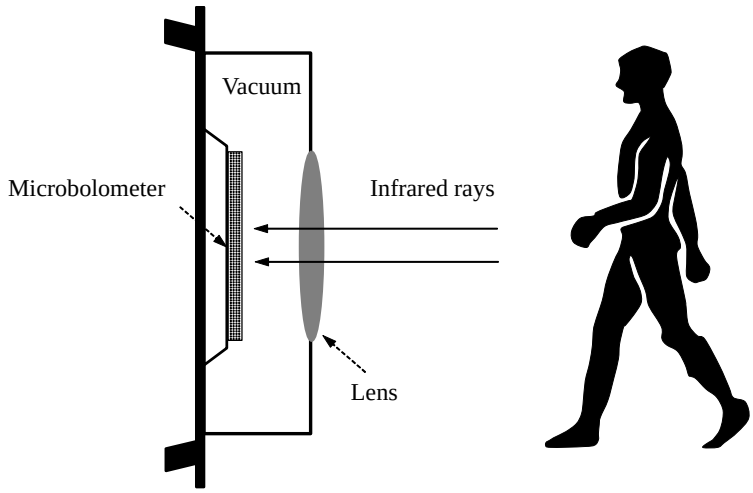


図 2.13 サーモカメラ概要 [20]

図 2.14 熱画像

図 2.15 サーモカメラ外観

- **可視光カメラ**

  人の目に見える光線である可視光が撮影できる装置であり、通常のビデオカメラやデジタルカメラと同様の映像が撮影できる。図 2.16 にカメラが光を受光する仕組みについて示す。光源より発光された光が人体に反射し、その反射光が可視光カメラに入射する。カメラに入射した光はレンズを通過し、カラーフィルタを通過した後にフォトダイオードで受光され電気信号へと変換される。フォトダイオードは光の強弱しか感知できず、色調表現ができないため、カラーフィルタによって特定の波長だけを通過させ、その輝度情報を色データへと変換する。撮像素子としては CCD (Charge coupled devices) や CMOS (Complementary metal oxide semiconductor) があり、この二つは電荷の転送方法が異なり、可視光に対する感度も異なる [24]。近

年では CMOS の方が性能面、価格面において優位であり、広く採用されている。図 2.17 に可視光カメラの外観を示す。可視光カメラの映像だけでは、人を検出することはできないが、画像処理を行うことにより人検出が可能となる。可視光カメラによる既存の画像処理手法には人検出手法が多く存在する。代表的な手法が Haar-like 特徴量と Adaboost を用いた顔検出手法 [25] と HOG (Histograms of oriented gradients) 特徴量を用いた人検出手法 [26] である。これらの手法はデジタルカメラの顔検出によるオートフォーカス機能や自動車の歩行者検出に使用されている。可視光カメラは人の存在だけではなく、位置や向き、移動方向等の様々な情報を抽出できる。さらに高精細な解像度を有しているため空間分解能が高く、安価で小型であるため搭載性も高い。しかし、その一方で使用者の顔や私生活に関する情報も撮影してしまうため、匿名性が低いという特徴がある。

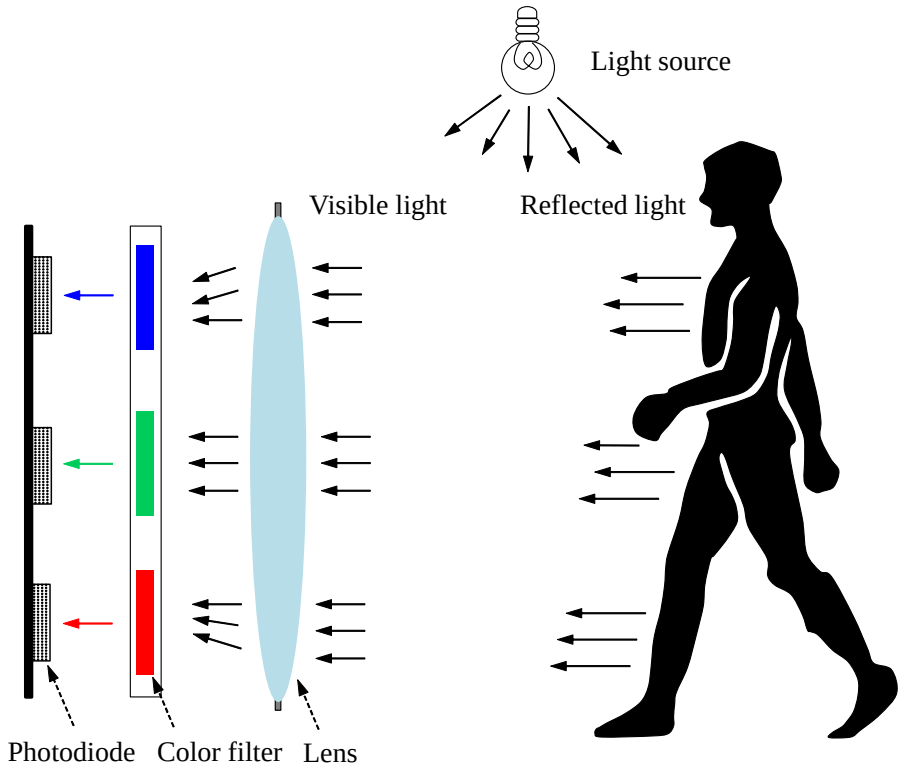
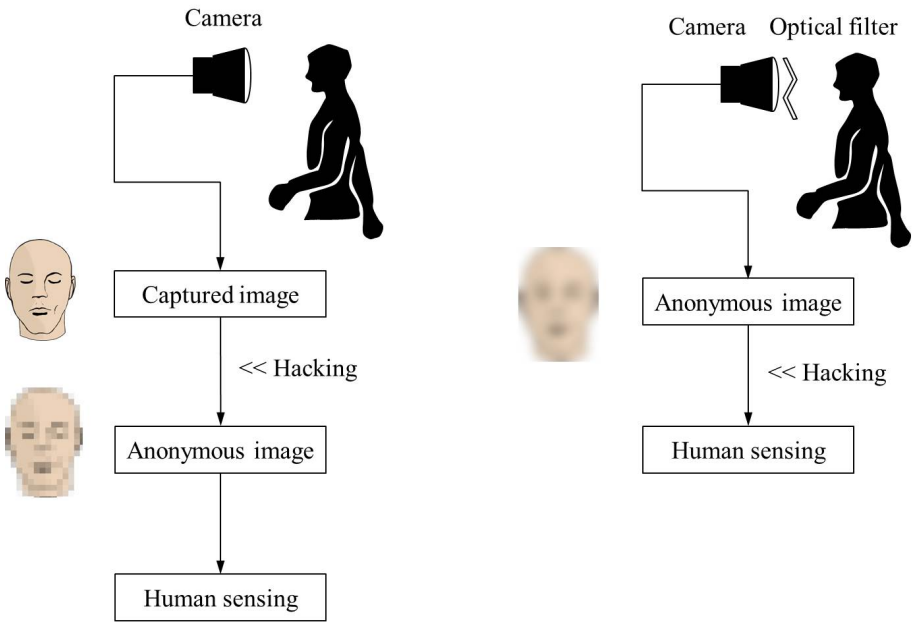


図 2.16 可視光カメラ概要 [24]

図 2.17 可視光カメラ外観

2.2 プライバシを考慮した人検出

可視光カメラを用いた人検出では使用者のプライバシが保護できないという問題があるが、その問題を解決する試みがいくつか報告されている。顔が個人を表す特徴的な情報であるため、撮影した画像の顔領域に対して、ソフトウェアにてモザイク処理や黒く塗りつぶす手法 [27, 28] が報告されている。また、撮影した画像の人の領域を CG (Computer graphics) で置き換えるといった手法 [29] も報告されている。さらに顔や人領域だけでなく、画像全体に処理を行う手法もある。画像の各領域を入れ替えることにより、無秩序な画像に変換する手法 [30] や、認証キーが無いと画像を閲覧できない仕組みにする手法 [31] も報告されている。しかし、図 2.18 (a) に示すように、これらの手法はすべて画像を撮影した後に処理をするため、カメラの元画像が流出すると、使用者のプライバシと匿名性は保護できないという問題がある。より安全な保護を考えると図 2.18 (b) に示すように、画像を取得した段階で匿名性が保護されていることが必要となる。



(a) 撮影後匿名化　　(b) 撮影時匿名化

図 2.18 撮影後の匿名化加工と撮影時の匿名化

2.3 脈拍計測原理

人の肌から脈拍計測を行う仕組みについて解説する。脈拍は心臓が収縮することによって発生する動脈の拍動を表しており、動脈の血液には酸化ヘモグロビンが多く含まれている。図 2.19 [32] は酸化ヘモグロビンの各波長に対する吸収係数を表している。縦軸が吸光係数を、横軸が光の波長を表しており、吸光係数が大きいほど酸化ヘモグロビンは光を多く吸収する。図 2.19 によると、酸化ヘモグロビンは 500 nm - 580 nm の波長の光を多く吸収する性質がある。

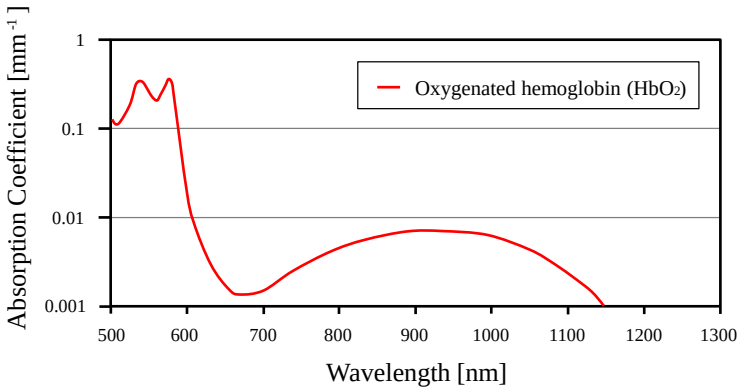


図 2.19 酸化ヘモグロビンの吸光係数 [32]

図 2.20 [33] は波長と色の対応関係を表している。可視光である 400 nm - 700 nm までの波長の色を示している。酸化ヘモグロビンに多く吸収される 500 nm - 580 nm の波長は緑色・黄色に対応している。市販されている脈波センサの指先型 [35]、腕時計型 [36] 等は、緑色が酸化ヘモグロビンに吸収される性質を利用して、緑色の LED (Light emitting diode) を使用している。

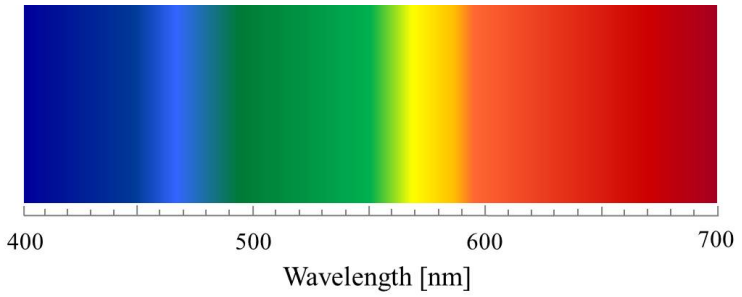


図 2.20 光の波長と色の相関 [33]

図 2.21 [34] に緑色の光を用いた脈拍の計測原理を示す。装置は緑色の LED とフォトディテクタを備えている。LED によって緑色の光を人の肌に照射し、反射した光をフォトディテクタで受光している。拍動により血管の容積が変化するため、フォトディテクタの受光する値は図 2.21 の右側のように心拍に同期した波形信号を得ることができる。この波形信号の一つ一つのピークが拍動を表しているため、ピークを検出することにより、脈拍数の算出が可能となる。

脈拍の計測部位として、指先と装置の固定が容易な耳たぶが利用される。その他にも額や手首、新生児の場合は足指、手のひら、くるぶし等も対象とされる。また、図 2.21 は反射してきた緑色の光量を計測する“反射式”であるが、赤外発光ダイオードを皮膚表面で発光させ、指を透過した赤外光をフォトトランジスタで受光する“透過式”も頻繁に使用される [9]。

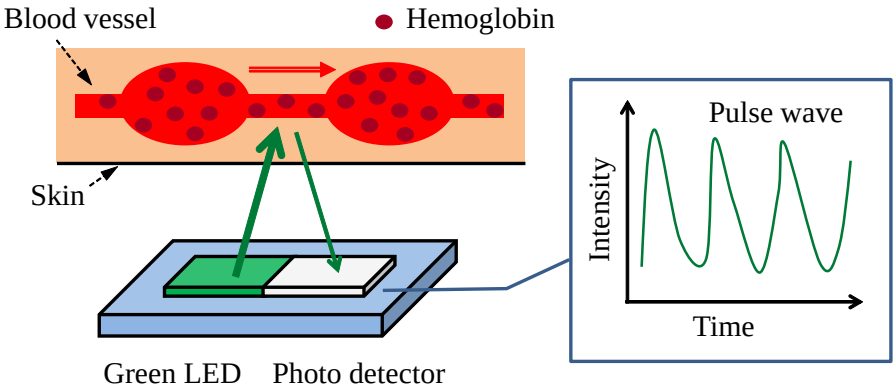


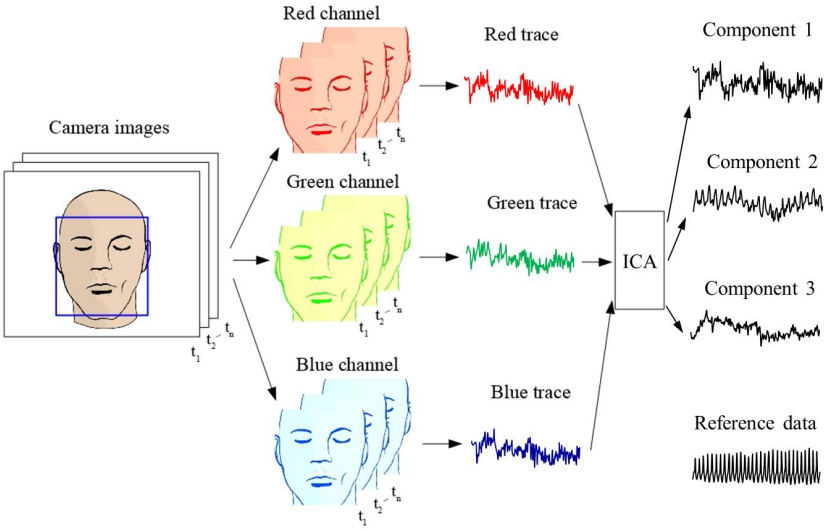
図 2.21 脈拍計測原理 [34]

2.4 非接触脈拍計測

脈拍計測機器は人体の部位と接触して計測を行う接触型が多数存在するが、非接触に脈拍を計測することも可能である。カメラで撮影した顔画像から非接触に脈拍の測定を行う手法が報告されている [37]。この手法では、まず Web カメラから画像を取得し、動画フレームから顔検出技術を用いて顔領域の抽出を行い、ROI (Region of interest) を決定する (図 2.22 (a))。可視光カメラは RGB の 3 チャネル信号を出力するため、図 2.22 (b) では、ROI を RGB の 3 チャネル信号に分割している。次に、図 2.22 (c) では、ROI 内の赤、青、緑の画素の平均値をそれぞれ保持し、時系列に表している。図 2.22 (c) の縦軸は各色の輝度を表しており、横軸は時間を表している。さらに、RGB 信号は ICA（Independent component analysis) によって三つの独立成分に分割される (図 2.22 (d))。ICA は元となる混合信号から独立した信号を抽出するための技術である。図 2.22 (d) の一番下には、参照用データの指の脈拍波形がある。第二成分 (Component 2) がこの参照用の波形に最も似た波形となっており、脈拍のデータを表している。この第二成分の

データのピークを検出することにより、脈拍数の算出が可能となる。この顔画像からの脈拍計測手法 [37] は iPhone の Cardiio [38] という名前でアプリケーションとして配布されており、日常の健康管理を目的として広く利用されている。この手法は ICA を用いているため、以後、ICA 手法と表記する。

顔画像からの脈拍計測は、顔が動いた時には雑音成分が大きくなり、脈拍計測の精度が低下するという問題がある。ICA 手法 [37] では、頭を静止した状態、または少し動かすような動作をした場合には、高精度に脈拍が推定できるが、運動状態では ICA の第二成分に脈拍が表れず、第一成分、第三成分に脈拍が表れ、高精度に脈拍が推定できないという現象が起きる。動いている状態で脈拍を高精度に計測できないことが、非接触脈拍計測の大きな課題となっている。



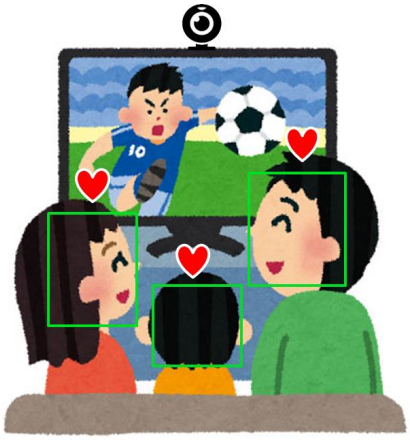
(a) 顔画像　(b) RGB 分割　(c) 時系列値　(d) 独立成分

図 2.22 顔画像からの脈拍計測概要 [37]

2.5 非接触脈拍計測の意義

近年、健康志向の高まりを受け、日常的なヘルスケアや生体情報の取得に対する関心が高まっている。多くの人が毎日視聴するテレビに非接触な計測機能を追加すれば、無意識な計測が行える。また、単なる生体情報の記録だけではなく、

脈拍を用いたテレビ機能の拡張にも有用である。その機能例の一つは、使用者の脈拍数の変化に応じて番組推薦を行うことである。チャネル履歴から使用者が好む番組を推定する手法も存在するが、テレビ画面を表示した状態で退席している場合や複数人で視聴している場合があり、精度が低い。人は興奮すると脈拍数が上昇するため、脈拍数の増減から使用者の関心や嗜好度を推定することができる。さらに、スポーツ番組においては使用者の脈拍数が上昇した場面だけを抽出することにより、自動で要約映像を作成することができる（図 2.23)。また、減量目的の運動を行う際には、運動強度が高すぎると脂肪が燃焼されず効率的な減量効果が得られない。脈拍を計測できると運動強度が算出できるため、運動中の使用者に脈拍数を提示することにより、効果的な減量が実現できる (図 2.24)。



- Automatic digest
- Audience measurement
  (Degree of interest, preference etc.)
- Content recommendation

図 2.23 テレビ視聴者の脈拍計測

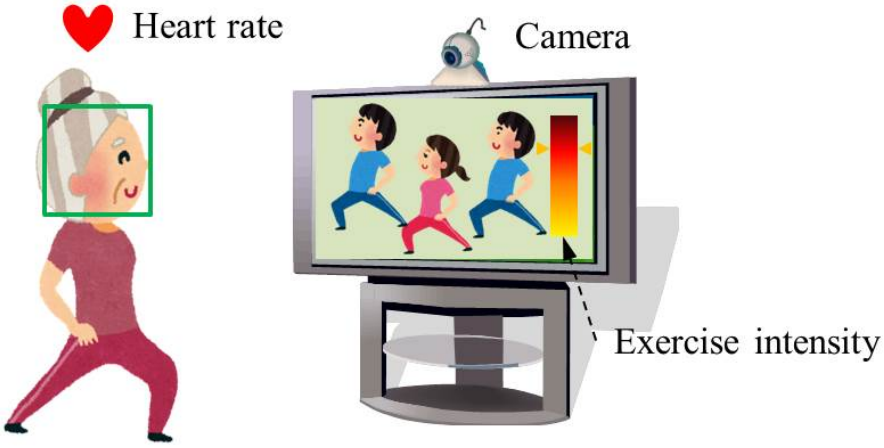


図 2.24 運動中の脈拍計測

現在、市場には様々なタイプの脈拍計が発売されている。指をセンサに入れて測るタイプ [40]、腕時計型 [41] の脈拍計等が発売されているが、いずれも接触する必要があるため、使用者にとって負担が生じる。そのため、使用者が意識せずに非接触で脈拍を計測することが望まれている。

非接触で脈拍・心拍を計測する手法はいくつか提案されている。電波センサを用いた手法 [42]、距離センサを用いた手法 [43]、サーモカメラを用いた手法 [44]、可視光カメラを用いた手法 [37] 等がある。電波センサを用いた手法では、電波が人体を通り抜けることを利用し、電波の反射波から心臓の動きを捉えることにより心拍を計測している。しかし、体動がある場合には心臓の動きと体の動きを切り分けることが難しく、人が動いている時には心拍が計測できないという問題がある。距離センサを用いた手法では、距離精度 1 mm のセンサを用いて胸の表面の動きを計測する。胸の動きは心臓の鼓動による動きであると想定し、胸の動きの変化から心拍を計測する。しかし、この手法においても、人が動いているときには、心臓の鼓動と体の動きを切り分けることができないため、動いている最中には心拍計測ができない。サーモカメラを用いる手法では、血管の拍動を検出することにより、脈拍を測定している。この手法においても動いている最中では、血管の位置を特定することができず動きには対応できない。可視光カメラを用いた手法の代表例としては、2.4 節で記述した ICA 手法がある。顔画像の微小な輝度変化を検出し、その変化が脈拍に相当するという考えに基づいて脈拍計測を行っている。この手法ではスマートフォンのカメラを洗面台に設置して、カメラの近くで静止している人の脈拍を計測している。使用者が静止した状態を想定してい

るため、動いている状態では計測できないという問題がある。

運動中の使用者や宅内で生活している使用者の脈拍を計測するためには、動いている状態でも計測できる必要がある。また、人検出のために非接触脈拍計測を用いる場合においても、動いている使用者を検出する必要があるため、運動状態での脈拍計測が必要となる。さらに人検出では脈拍計測処理を複数領域に対して繰り返し行う必要があるため、計算処理が高速に行える簡素なアルゴリズムが必要である。

2.6 まとめ

本章では、人検出の要素となる技術について述べた。2.1 節では、既存の人検出装置の検出原理と特徴を報告したが、本研究で目標としている匿名性、搭載性、空間分解能を満たす人検出装置は存在しなかった。表 2.1 に各人検出装置の特性を示す。いずれの装置も足りない部分があるが、可視光カメラは搭載性と空間分解能については優れており、匿名性が低い問題を解決すれば、本研究で目指すセンサを実現することになる。そのため、本研究では可視光カメラを使用することを選択した。

表 2.1 人検出装置特性

Human detection sensor	Anonymity	Spatial resolution	Mountability
Pyroelectric sensor	High	Low	High
Photoelectronic sensor	High	Low	High
Thermopile array sensor	Medium	Medium	Low
Ultrasonic sensor	High	Low	High
Radio wave sensor	High	Low	High
Thermal camera	Medium	High	Low
Visible light camera	Low	High	High

2.2 節では、可視光カメラを用いてプライバシと匿名性に関する問題に取り組む先行研究について報告を行ったが、いずれの研究も画像撮影後の加工により、プライバシ保護を行っているため、撮影直後の元画像が流出するとプライバシ保

護ができないという問題がある。そのため、本研究では光学的に不鮮明にした匿名化画像から人検出を行い、さらに人検出の技術要素としてアピアランス情報を必要としない脈拍を使用する方針とした。2.3 節では、脈拍計測の原理について解説し、2.4 節では顔画像から脈拍を計測する先行研究について報告した。2.5 節では非接触脈拍計測の意義について述べた。脈拍計測の先行研究では、顔が動くと計測誤差が増大するという問題がある。本研究で行う人検出は、動いている人も対象としているため、運動状態でも高精度に脈拍計測を行う必要がある。本研究では、3 章において、動いている状態でも高精度に脈拍計測が行える手法を提案する。

第3章

非接触脈拍計測システム

本章では、本研究の重要な技術要素となる非接触脈拍計測について述べる。3.1 節では画像を用いた脈拍数の推定方法について提案を行う。本章では、脈拍計測の原理検証のため、匿名化画像ではなく通常の鮮明な画像を用いて脈拍計測を行う。3.2 節では提案手法の脈拍推定手法の性能評価を行い、3.3 節ではまとめと将来の展望について述べる。

3.1 画像からの脈拍計測手法

血液中に含まれるヘモグロビンが緑色の光を吸収することに着目し、人の顔画像の緑色の変化を捉えることにより脈拍の計測を行う。非接触脈拍計測では原理検証のため、鮮明画像を用いる。まず既存の顔検出手法を用いてカメラ画像から顔を検出する。図 3.1 の枠が顔領域を表している。

図 3.1 検出された顔領域

また、顔全体が動いた場合に対応するために、時系列の連続する 2 枚の画像に対して差分処理を行う。これにより、動きがある部分や照明により変化した部分が抽出できる。

時間差が Δt となる時刻 t_{n-1} と t_n に撮影された画像を $I_{t_{n-1}}$,I_{t_n} とする。$t_{n-1} + \Delta t = t_n$ となる。画像 $I_{t_{n-1}}$,I_{t_n} の画素値は RGB の 3 値から濃淡値に変換した値を使用する。式 (3.1) に変換式を示す。

$$I_{t_n}(x, y) = 0.299 \times r_{t_n}(x, y) + 0.587 \times g_{t_n}(x, y) + 0.114 \times b_{t_n}(x, y) \tag{3.1}$$

x と y は画像 I_{t_n} 中の横方向位置、縦方向位置をそれぞれ表す。$r_{t_n}(x, y), g_{t_n}(x, y), b_{t_n}(x, y)$ は座標 (x, y) の赤、緑、青の輝度値をそれぞれ表し、$I_{t_n}(x, y)$ が座標 (x, y) の濃淡値を表す。二つの画像の差分値である $J_{t_n}(x, y)$ は式（3.2）で定義できる。

$$J_{t_n}(x, y) = |I_{t_n}(x, y) - I_{t_{n-1}}(x, y)| \tag{3.2}$$

この差分値から画素値を二つのクラスに分類することができる。差分の結果から動きの範囲において 2 値化処理を行う。式 (3.3) において T_d は閾値を表し、$D_{t_n}(x, y)$ は 2 値化された画素値を表す。

$$D_{t_n}(x, y) = \begin{cases} 1 & if \quad J_{t_n}(x, y) > T_d \\ 0 & otherwise \end{cases} \tag{3.3}$$

式 (3.3) の $D_{t_n}(x, y) = 1$ となる画素は図 3.2 では白い画素で表されている。鼻のように突起した部分は照明による変化が大きくなってしまうため、周囲の平坦な部分と比較すると動きによる輝度変化が大きくなっている。図 3.1 の枠で指定した領域内で動きによる輝度変化が少ない領域だけを脈拍推定に利用する。

図 3.2 差分画像

図 3.1 の枠で囲まれた領域の画素値から RGB の G (緑) だけを抽出し、式 (3.4) に示すように領域内の空間平均値 G_{t_n} を算出する。

$$G_{t_n} = \frac{1}{N} \sum_{x'=F_{x0}}^{F_{x1}} \sum_{y'=F_{y0}}^{F_{y1}} g_{t_n}(x', y') \tag{3.4}$$

$$D_{t_n}(x', y') = 0 \tag{3.5}$$

F_{x0}, F_{x1} は、顔領域の x 座標の左端、右端を、F_{y0}, F_{y1} は y 座標の上端、下端を表す。x' と y' は式 (3.5) を満たす画素を示す。N は式 (3.4) で使用される画素の総数を示す。領域内の空間平均値を周波数 30 Hz で 30 秒分保存し、空間平均値の時間変化を表すと図 3.3 になる。縦軸は緑成分の輝度値を表し、横軸は時間を表している。図 3.3 の 13 秒未満の部分は顔が動いている状態で、13 秒以降の部分は顔が静止した状態となっている。静止状態の範囲では値が微小な変動を続けているが、この部分が脈による変動を表している。このままでは脈による変動と動きによる変動が判別できないため、図 3.3 の信号に対して平滑化処理を行うことにより、振幅の小さい脈の成分だけを除去する。式 (3.6) に表されるように 5 点の移動平均を用いて平滑化処理を行う。

$$M_{t_n} = \frac{G_{t_{n-2}} + G_{t_{n-1}} + G_{t_n} + G_{t_{n+1}} + G_{t_{n+2}}}{5} \tag{3.6}$$

次に差分処理を式 (3.7) で表す。

$$S_{t_n} = G_{t_n} - M_{t_n} \tag{3.7}$$

M_{t_n} は式 (3.6) で算出された平均値を表し、S_{t_n} は二つの値の差分値を表す。図 3.4 は S_{t_n} の値の変化を表しており、動き成分を除去して脈拍の成分だけを抽出している。この S_{t_n} に対して周波数解析の DFT (Discrete Fourier Transform) を行うと図 3.5 のパワースペクトラムが得られる。DFT のパワースペクトラムを、式 (3.8) に示す。

$$P(k) = \left| \sum_{n=0}^{N-1} S_{t_n} \exp\left(-j\frac{2\pi kn}{N}\right) \right|^2 \tag{3.8}$$

k は周波数を表す。このときの窓関数は矩形窓を使用し、現在時刻から過去に遡り、窓関数を畳み込む。N は窓関数内の S_{t_n} の値の数を表す。図 3.5 の周波数成分においてピークとなる値から脈拍数を計算することができる。図 3.5 ではピーク周波数は 1.03 Hz であるため、60 を積算することで、脈拍数は約 62 bpm (Beats per minute) となる。

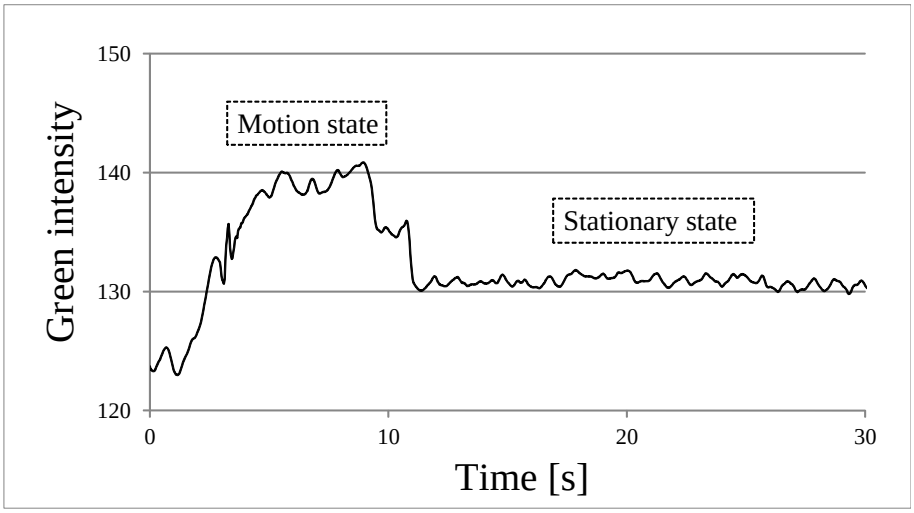


図 3.3 顔領域内の緑色の輝度値変化

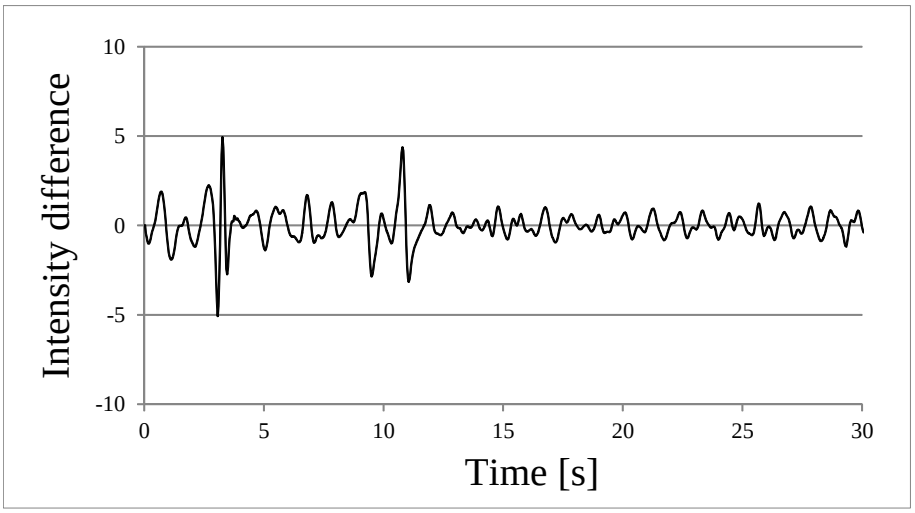


図 3.4 差分処理により抽出された脈拍成分

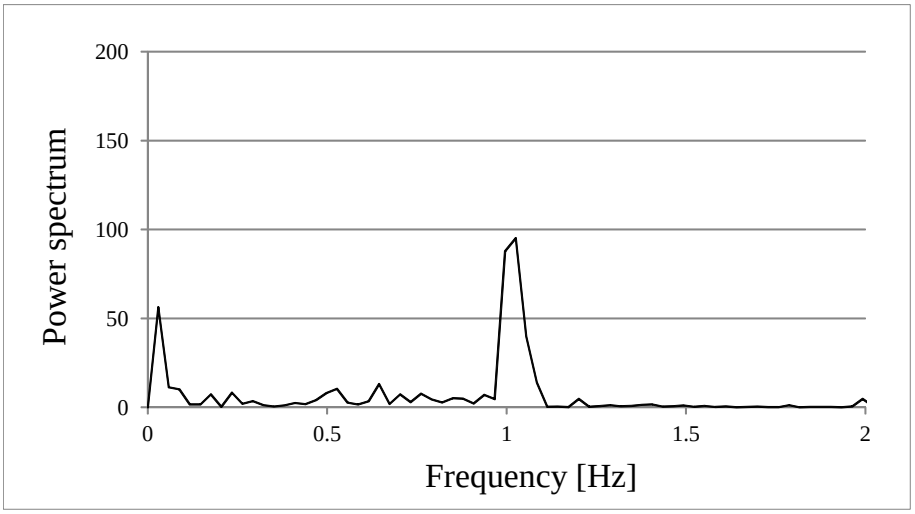


図 3.5 DFT による周波数解析

3.2 脈拍計測手法の性能評価

距離、照度、動き、処理時間の四つのカテゴリーにおいて性能検証を行った。カメラはビデオチャット等で用いられる図 3.6 に示す Web カメラ [45] を使用した。このカメラにて解像度 640 × 480 pixel、 30 fps (frame per second) にて撮影を行い、USB2.0 で PC へ接続し、処理を行った。表 3.1 に実験に用いた PC の仕様を示す。また、脈波計測のソフトウェアは 32 bit アプリケーションとして作成

した。参照データとして図 3.7 の胸に装着するタイプの心拍計 [39] を用いた。心拍計の精度は ± 1 bpm である。誤差は参照データの脈拍数と提案手法による脈拍数の差の絶対値とした。また、照明は天井に設置された蛍光灯のみであった。提案手法は DFT の時間窓を 30 秒間と設定した。画像撮影は 90 秒間行い、最初の 30 秒間は DFT の結果が出力されないため、30 秒間の検出率は評価の対象外とし、残りの 60 秒間の 1800 枚の画像について評価を行った。

表 3.1 PC 仕様

Model	Dell OptiPlex 9010
CPU	Intel (R) Core (TM) i7-3770 3.40 GHz
Memory	32.0 GB
OS	Windows 7 64-bit

図 3.6 Web カメラ (Microsoft LifeCam Studio) [45]

図 3.7 参照データ用心拍計 (Zephyr BioHarness 3)

3.2.1 距離

図 3.8 に示すようにテレビの上部にカメラを設置し、椅子に座っている被験者の顔とカメラまでの距離を計測した。カメラと顔の距離を 2.0 m から 6.0 m までの間で 0.5 m 刻みで移動させ、静止状態の 8 人の被験者に対して 90 秒間の計測を行った。各距離における平均誤差を図 3.9 に示す。エラーバーは標準偏差を表す。図 3.9 の結果から提案手法の誤差は距離に依存していないことが確認できる。すべての距離において脈拍の平均誤差が 3 bpm 以下となっており、十分な精度を示している。

テレビを視聴する場合は、最適視聴距離の目安が定められており、画面の高さの 3 倍が最適であるとされている。1920 × 1080 の解像度においては大型テレビに分類される 52 型で 1.9 m、60 型で 2.2 m が最適視聴距離となる。図 3.9 の結果では 6 m の距離でも脈拍が計測できており、距離性能については十分であると言える。

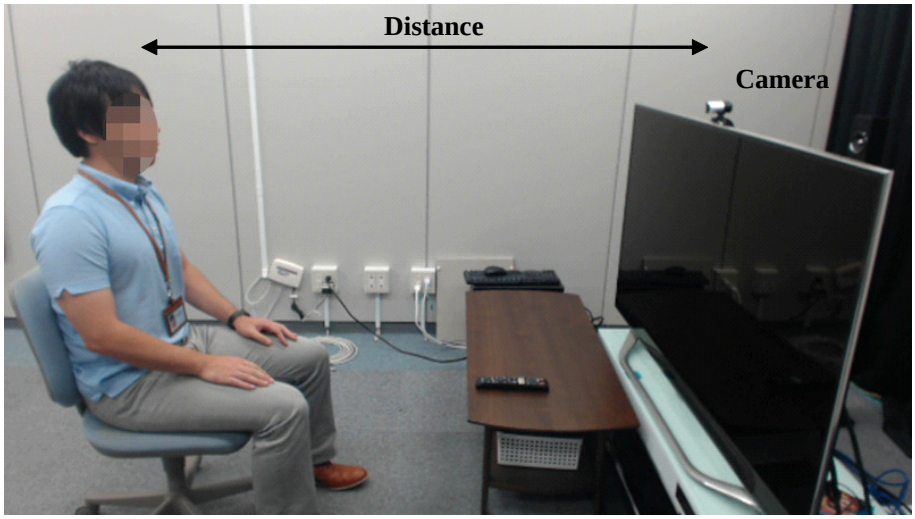


図 3.8 距離別性能評価　実験環境

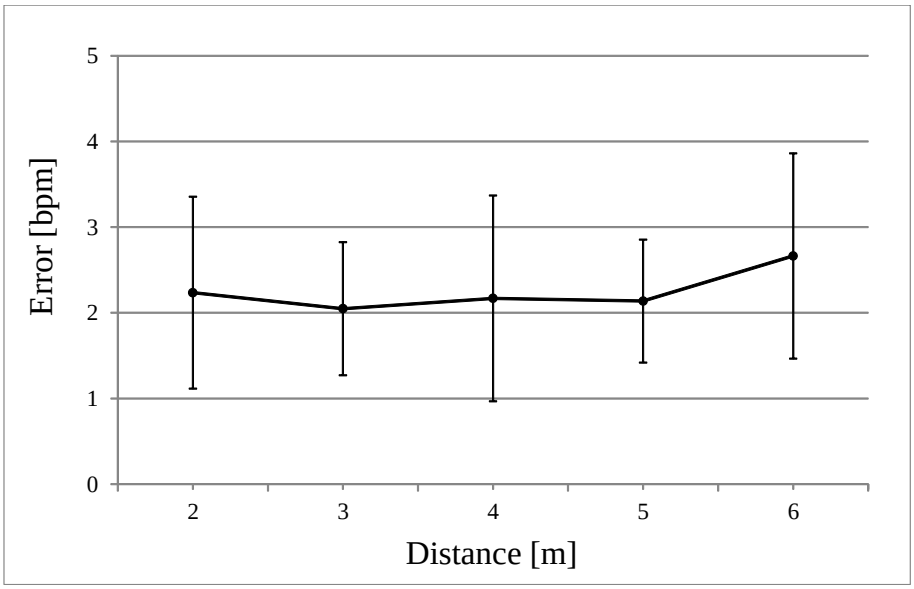


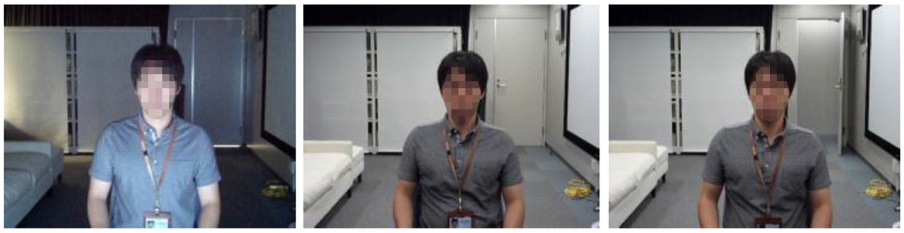
図 3.9 距離別誤差

3.2.2 照度

照度を 30 lx から 2000 lx まで変化させ、図 3.10 に示すように距離 2.0 m の位置に被験者が座り、静止した状態で計測を行った。テレビのバックライトの影響も計測範囲内とするために、テレビ画面には起動時に表示されるトップメニューを表示したまま計測を行った。8 人の被験者に対して、各 90 秒間の計測を行い、

照度は 30 lx, 500 lx, 1000 lx, 1500 lx, 2000 lx にて行った。照度の計測にはデジタル照度計 LX-PRO TM-204 [46] を使用し、被験者の顔と同じ位置で水平に向けて計測を行った。各照度における平均誤差を図 3.11 に示す。

図 3.11 の結果から、30 lx の暗い環境においては性能が低下することがわかる。しかし、500 lx 以上であればすべての計測で平均誤差が 3 bpm 以内に収まっており、通常のリビングルーム環境の照度が 1000 lx 程度であるので、日常環境で使用するには十分な性能を示した。



30 lx　　500 lx　　2000 lx

図 3.10 照度別撮影画像

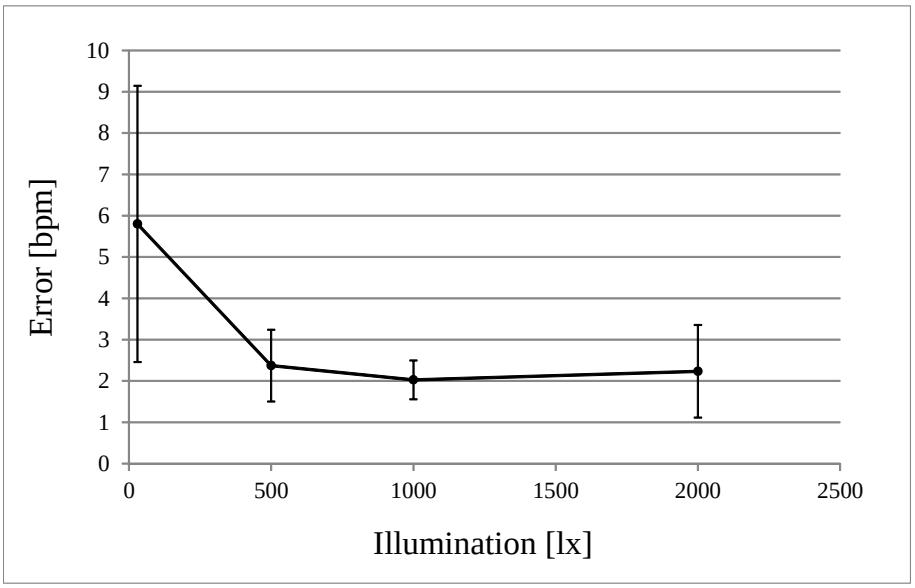


図 3.11 照度別誤差

3.2.3 動き

動きに対する頑健性の性能を計測するために先行研究である ICA 手法 [37] との比較実験を行った。ICA 手法は iPhone の Cardiio [38] というアプリケーションで配布されているため、Cardiio を用いて実験を行った。

図 3.12 のように iPhone と Web カメラを同じ位置に設置し、Cardiio の最適測定距離である 0.4 m で撮影した。Cardiio の脈拍数を画面に表示し、それを Web カメラで撮影することにより、Cardiio と Web カメラの時刻同期を行った。ダンスやジョギング等の動きを想定して顔を角速度 86 度/秒 で往復運動させて計測した。被験者 5 人に対して顔を動かさない場合、垂直方向に動かす場合、水平方向に動かす場合と三つのケースについて各 90 秒間の計測を行った。結果を図 3.13 に示す。

図 3.13 の結果では、静止状態では提案手法と ICA 手法との間に明確な違いはなく、両方とも誤差が 3 bpm に収まる結果となった。顔を横に動かした場合では、提案手法は ICA 手法より小さい誤差となっており、静止状態と同様の良好な結果となった。標準偏差も提案手法の方が小さくなった。顔を縦に動かした場合では、提案手法は ICA 手法より小さい誤差で、標準偏差でも小さい値となった。運動状態の性能評価において、提案手法は先行研究である ICA 手法に対して高い性能を示した。

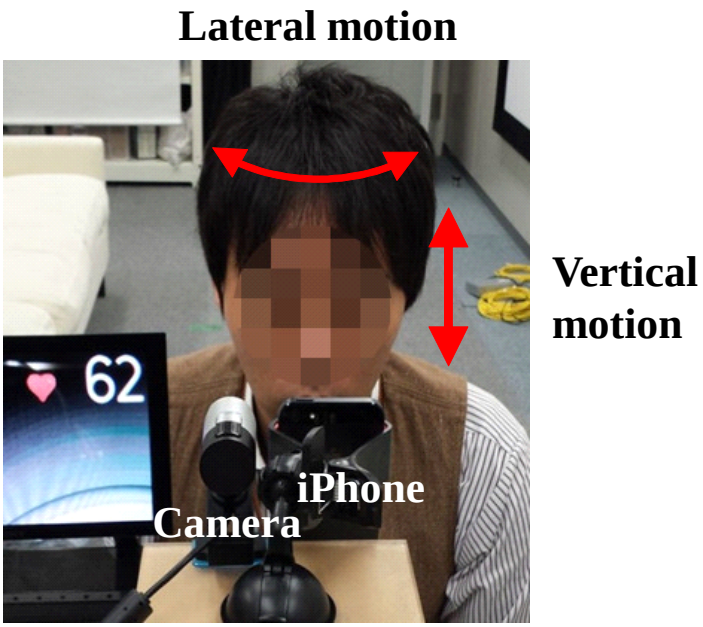


図 3.12 運動状態性能評価 実験環境

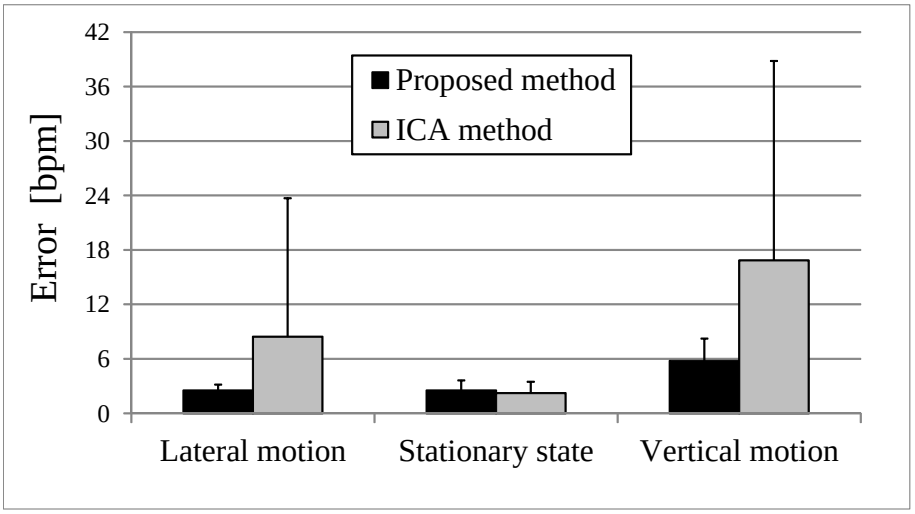


図 3.13 運動状態での脈拍計測誤差

3.2.4 処理時間

提案手法は平滑化と差分処理を用いた簡素なアルゴリズムであるため、高速な処理が期待できる。簡素なアルゴリズムによる効果を確認するために、1 フレーム当たりの処理時間を計測した。計測には 3.2.3 項で撮影した一人の被験者のデー

タ 1800 枚の画像を使用した。平均処理時間と標準偏差を表 3.2 に示す。比較のために独立成分分析を用いた ICA 手法 [37] の処理時間も記載する。表 3.2 の結果から提案手法は ICA 手法よりも 3.2 倍高速になっていることがわかる。また、提案手法はばらつきを表す標準偏差が小さく、常に高速で処理できていることが確認できる。人検出においては複数領域に対して脈拍計測処理を行うため、高速に処理を実行することは非常に重要である。提案手法は ICA 手法と比較して人検出に適していることが確認できた。

表 3.2 脈拍計測 1 フレーム当たりの処理時間

	Average	Standard deviation
Proposed method [ms]	2.6	0.7
ICA method [ms]	8.4	5.8

3.3 まとめ

カメラを用いて非接触に脈拍を計測する手法を提案した。カメラ画像の RGB から G (緑) だけを抽出し、平滑化と差分処理を行うことにより、雑音を除去し脈拍成分の抽出を行った。抽出された信号に対して周波数解析を行うことにより、脈拍数の算出を行った。性能評価においては心拍計を参照データとして、推定された脈拍数の誤差を計測した。距離を変化させた場合には 6.0 m 離れても誤差が 3 bpm 以内となり、照明を変化させた場合には 500 lx 以上であれば 3 bpm 以内になることがわかった。これらの結果から、居間等での日常生活においては、十分に使用できる性能であることがわかった。また、動きに対する頑健性を検証するために、ICA 手法との比較を行った。静止状態では、提案手法と ICA 手法の性能は同等であったが、運動状態では提案手法の方が高い性能を示した。今後はダンスやジョギング等の運動中の脈拍計測への応用が期待できる。人検出を行う場合には、複数領域に対して脈拍計測処理を行うが、人が動いている場合もあるため動きに対する頑健性は必要である。また、人検出では脈拍計測処理を複数回繰り返し行うため、高速な処理が要求される。処理時間の計測において、提案手法は 1 フレーム当たり 2.6 ms の処理時間を示し、ICA 手法の 3.2 倍高速な性能を示した。提案手法は距離変化、照度変化、動き変化に対して頑健であり、高速な実行可能な脈拍計測手法であることが確認できた。さらに、テレビ視聴中の使用

者の脈拍計測に適した脈拍計測手法であり、且つ人検出に適した手法であると言える。

第4章

脈拍による顔検出システム

本章では、脈拍を利用した顔検出システムについて述べ、3 章で記述した非接触脈波計測手法を基に匿名化画像から顔検出を行う。4.1 節では脈拍を用いた顔検出手法について述べ、4.2 節では顔検出手法の検出率の評価を行う。4.3 節では顔位置を高精度に検出する手法について述べ、4.4 節は顔位置精度の性能評価を行う。4.5 節ではまとめを行う。

4.1　脈拍を利用した顔検出手法

従来の可視光カメラ画像を用いた人検出手法では、人の見た目であるアピアランス情報を用いているが、そのためには鮮明な画像が必要である。しかし、使用者の匿名性を保護するためには、人が理解できるような鮮明な画像の撮影はできないため、本研究ではアピアランス情報を用いない手法を提案する。3 章ではカメラ画像から脈拍を計測する手法について述べた。この手法は顔領域を検出し、その領域に対して脈拍計測を行う手法である。この脈拍計測の手法を応用することにより、人検出を行う。カメラ画像の領域を分割して、脈拍計測の処理を各領域に対して行い、人の脈拍数に相当する値が算出された領域には人の顔が存在すると想定する。

まずカメラで画像を撮影する前に、カメラの前にすりガラスを置いたり、レンズの焦点を移動させたりして、個人の顔が特定できない程度まで画像の鮮明さを低下させて、匿名化画像として撮影を行う。通常のカメラで撮影すると図 4.1 の画像が取得できるが、レンズの焦点を移動させてぼかすことにより、図 4.2 の匿名化画像が取得できる。この匿名化画像に対して領域分割を行うと図 4.3 のように画像全体を 30 個の領域に分割することができる。この分割された各領域に対して、脈拍計測の手法を用いて各領域の脈拍数を算出する。

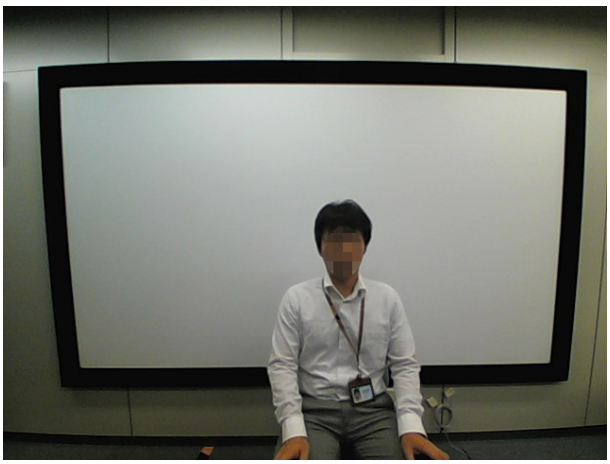


図 4.1 鮮明画像

図 4.2 匿名化画像 (不鮮明画像)

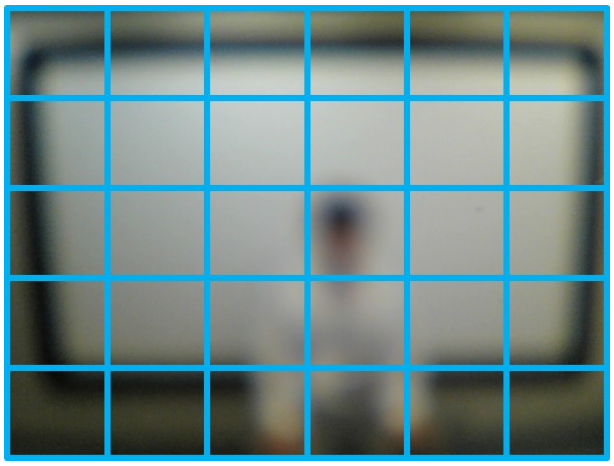


図 4.3 領域分割画像

図 4.4 に示す領域 A は背景領域、領域 B は顔の領域を表している。室内の照明や太陽光には緑色が含まれているため、RGB カメラの G (緑) の輝度値のみを抽出する。二つの各領域内の緑色の平均輝度値の周波数 30 Hz、30 秒間の時間変化を図 4.5 と図 4.6 に表す。図 4.5 は背景領域であり、緑色の輝度値の変化が無い状態であるのに対して、顔の領域である図 4.6 の輝度値には、脈拍を表すような一定間隔の波形が表れていることがわかる。

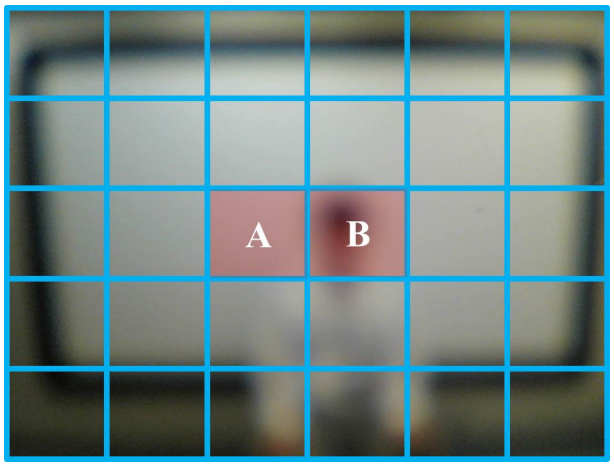


図 4.4 背景領域 (A) と顔領域 (B)

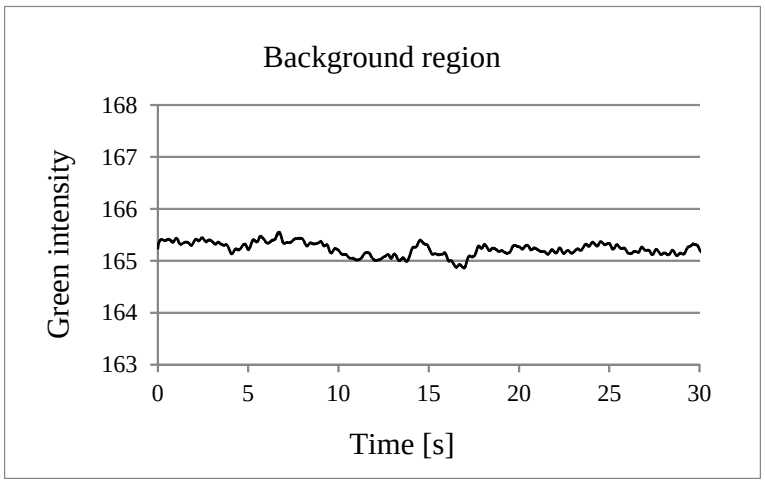


図 4.5 背景領域 (A) における緑色輝度値の変化

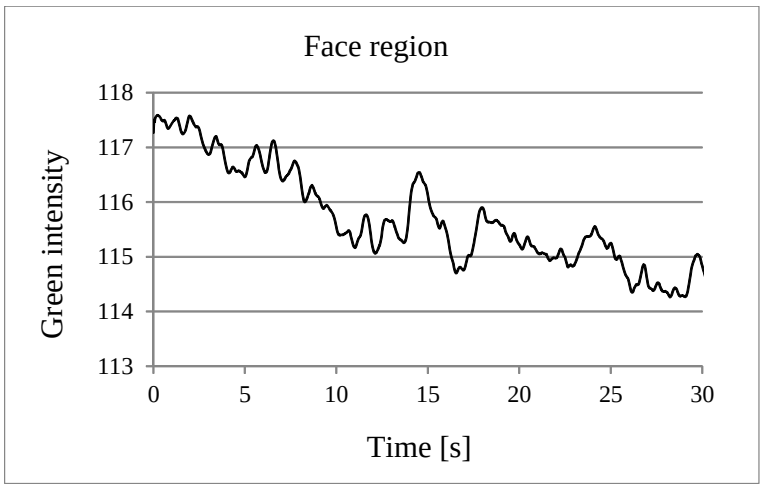


図 4.6 顔領域 (B) における緑色輝度値の変化

各領域の緑色の平均輝度値を入力として、DFT の処理を行う。このときの窓関数は矩形窓を使用し、現在時刻から過去に遡って窓関数を畳み込む。また、時間窓は直前に処理した区間を重複しながら、逐次移動することとする。背景領域と顔領域の各領域に対して、3.1 節に記述した平滑化と差分処理を行い、脈拍成分の抽出を行う。抽出された信号に対して、周波数解析を行うと、図 4.7、 図 4.8 に表されるパワースペクトラムが得られる。このパワースペクトラムの結果から人の脈拍が検知できるかどうかの判定を行う。人の脈拍数として計測されないような値は探索範囲外とするため、脈拍数の範囲設定を行う。この手法では人の脈拍数として 50 bpm - 120 bpm を設定する。この手法においては激しい運動や病気

の場合は対象外としており、宅内の日常生活での行動を対象としている。図 4.7、図 4.8 の灰色の領域は脈拍数が 50 bpm 未満になる領域を表しており、この部分は探索の範囲外とする。また、パワースペクトラムの値が小さい場合には、人間の脈拍では無い可能性が高いため、一定の閾値を設定し、パワースペクトラムの値が閾値以下の場合は顔のある領域ではないと判定する。閾値を超える値が複数ある場合には、パワースペクトラムの最大値の周波数を脈拍数として採用する。図 4.7、 図 4.8 では、経験則から閾値を 20 に設定している。このような判断基準で図 4.7 のパワースペクトラムを判定すると赤点線で表された閾値を超える値が存在しないため、顔が存在しない領域であると判定される。図 4.8 においては周波数が 0.95 Hz 付近で閾値を超えており、60 を積算することで、脈拍数 57 が算出でき、この領域に顔が存在すると判定できる。

この処理をすべての領域に対して行うと図 4.9 のようになる。背景領域で閾値を超えるパワーが存在しなかった領域は脈拍数を 0 として表現している。図 4.9 では人の顔の部分だけが脈拍数 57 となっており、その領域に顔が存在していると考えることができる。このように、画像の領域を分割し各領域に対して脈拍数の算出を行い、人の脈拍数として設定した 50 bpm - 120 bpm の値が検出された領域には顔が存在していると判定することができる。

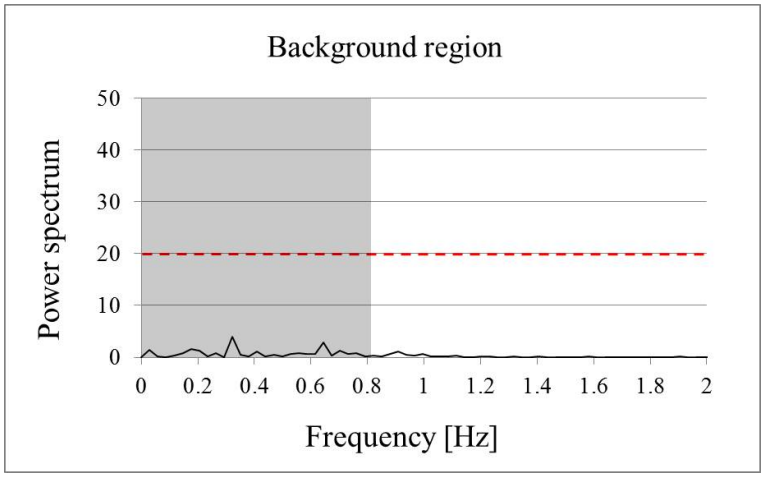


図 4.7 背景領域 (A) における DFT パワースペクトラム

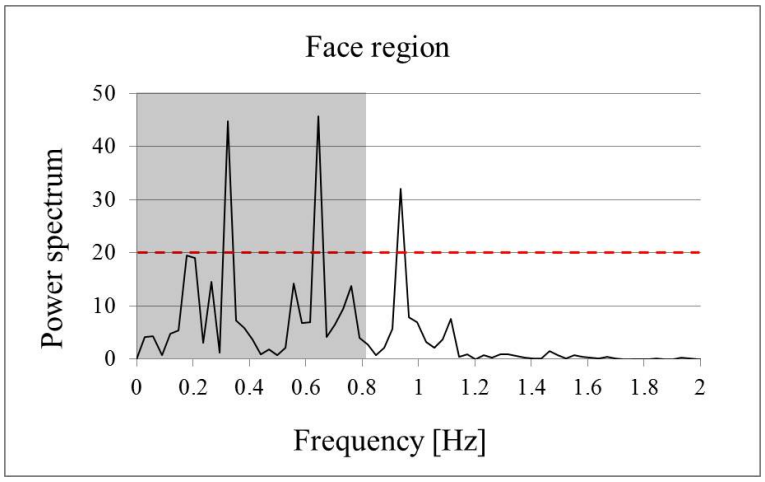


図 4.8 顔領域 (B) における DFT パワースペクトラム

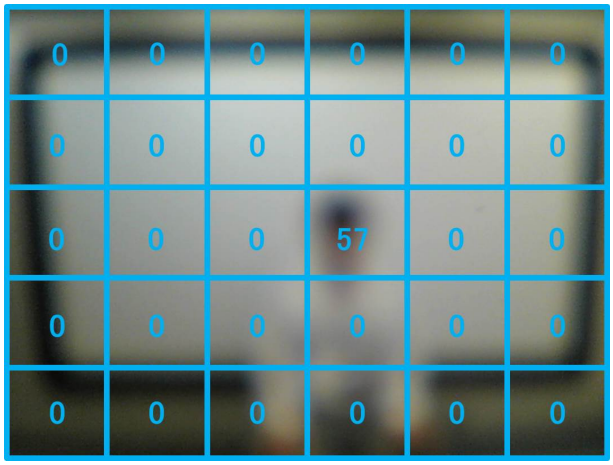


図 4.9 各領域における脈拍計測結果

4.2 顔検出率の評価

提案手法の 3 種類の匿名化画像における性能評価と鮮明度の違いによる性能評価を実施した。

4.2.1 匿名化画像での性能評価

(1) 実験環境

提案手法の顔検出について性能評価を行った。カメラはビデオチャット等で用いられる図 4.10 に示す Web カメラ [47] を使用した。このカメラにて解像度 640 × 480 pixel、30 fps にて撮影を行い、USB2.0 で PC へ接続し、処理を行った。表 4.1 に実験に用いたカメラの仕様を示す。PC は 3.2 節の表 3.1 に示した PC と同じ機材を使用した。また、人検出のソフトウェアは 32 bit アプリケーションとして作成した。

表 4.1 RGB カメラ仕様

Type	RGB
Resolution [pixel]	640 × 480
Frame rate [fps]	30
Gradation [bit]	8

図 4.10 RGB カメラ [47]

カメラから 1 m 離れたところに人が座り、静止した状態で撮影を行った。本実験では静止状態での検出率に主眼を置いているため、人が動いている状態では撮影は行わず、静止状態のみで撮影を行った。また、匿名化画像として 3 種類の画像を用いた。一つ目は図 4.11 のようにレンズの焦点をずらした画像で、二つ目

は図 4.13 のようにサイクロンフィルタで撮影された画像である。図 4.12 に渦巻き状の模様を持つサイクロンフィルタ [48] を示す。三つ目は図 4.15 のようにラジアルフィルタで撮影された画像であり、図 4.14 が放射状の模様を持つラジアルフィルタ [49] である。

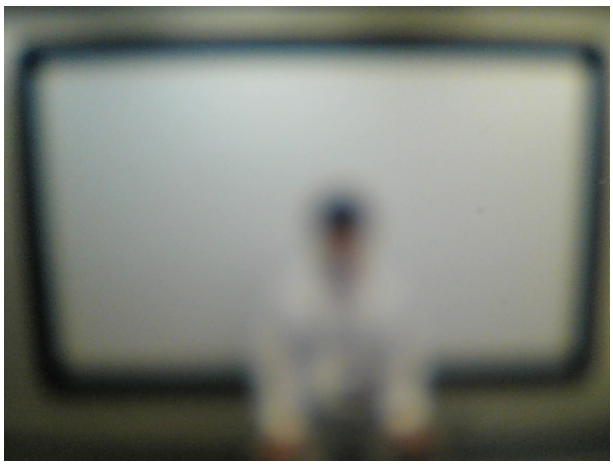


図 4.11 焦点がずれた画像

図 4.12 サイクロンフィルタ [48]

図 4.13 サイクロンフィルタ画像

図 4.14 ラジアルフィルタ [49]

図 4.15 ラジアルフィルタ画像

既存手法との比較のために、代表的な顔検出手法である Haar-like 特徴量と Adaboost を用いた顔検出手法 [25] を比較手法として用いた。この手法を以下、Haar-like 手法と表記する。Haar-like 手法は OpenCV (Ver. 2.4.3) で作成した。Haar-like 手法の最小探索窓は 30 × 30 pixel とし、探索窓を逐次拡大する倍率を 1.1 と設定した。提案手法は 640 × 480 pixel 画像を 40 × 40 pixel の大きさで 192 個の領域に分割した。提案手法は DFT の時間窓の幅を 30 秒と設定した。窓関数は矩形窓を使用し、現在時刻のデータから過去に遡ったデータを時間窓の範囲とした。画像撮影は 90 秒間行い、最初の 30 秒間は DFT の結果が出力されないため、30 秒間の検出率は評価の対象外とし、残りの 60 秒間の 1800 枚の画像についての検出率の評価を行った。比較対象の Haar-like 手法についても同様の 1800 枚について評価を行った。また、今回は 5 人の被験者に対して評価を行った。

検出の判定基準について説明する。まず撮影された画像に対して、人が目視にて顔が存在する領域を設定する。顔検出手法によって算出された顔領域の矩形が人手によって設定された領域と少しでも重複していれば、顔検出は成功したと判定する。お互いの領域が重複していなければ、誤検出として判定する。今回の撮影では 1 フレームに一人しか存在していないため、1 フレームを 1 試行とする。また、1 フレームにおいて顔検出の矩形が複数個所存在し、その矩形がすべて顔検出に成功している場合には、重複した結果であると判断し、一回の成功であると判定している。誤検出の場合も同様に、1 フレーム内に複数個所の誤検出が存在した場合も一回の誤検出であると判定している。

(2) 実験結果

顔検出実験の結果を ROC (Receiver operating characteristic) 曲線 [50, 51] で表す。ROC 曲線は縦軸が検出率（True positives)、横軸が誤検出率 (False positives) を示している。これは各手法の検出の閾値を変化させることにより作成することができる。提案手法の場合は DFT によるピーク値の閾値、Haar-like 手法の場合は、検出領域の重畳回数の閾値を変化させて作成した。ROC 曲線は検出率 100 %、誤検出率 0 % である左上に近ければ近いほど、検出性能が高いとされている。一般的に検出手法において検出率を上げようと調整すれば、誤検出が増えてしまう。逆に誤検出を抑えるようにパラメータを調整すると、今度は検出率が下がってしまう。このため、検出手法の一つのパラメータにおける検出率を比較しても、検出手法全体の性能を比較したことにはならないが、ROC 曲線にて比較することにより、全体の性能比較が可能となる。ROC 曲線は人検出手法 [52, 53] や物体検出手法 [54, 55] の性能評価で使用されている。

この実験では 60 秒間の 1800 枚の画像を三つのフィルタで 5 人の被験者にて行ったため、実験に使用した画像は合計で 27000 枚となる。

図 4.16 に実験結果を示す。ROC 曲線で表された実験結果では、提案手法の曲線の方が Haar-like 手法の曲線よりも左上に近い位置にある。このことから、提案手法の方が高い性能を示していることがわかる。二つの手法の誤検出率が 2.7 % の時の検出率を比較してみると、提案手法は 99.5 % で、Haar-like 手法の方は 4.1 % であり、明らかな違いがあった。Haar-like 手法は、顔検出に必要な明暗差やエッジの情報が不鮮明な画像では検出されないために、全体的に検出率が低い結果となった。反対に提案手法では高い検出率と低い誤検出率が得られる結果となった。この結果から、意図的に不鮮明にした匿名化画像に対しては、既存の顔検出手法では高い検出率が得られなかったが、提案手法では高い検出率を得ることができ、提案手法の有効性が確認された。

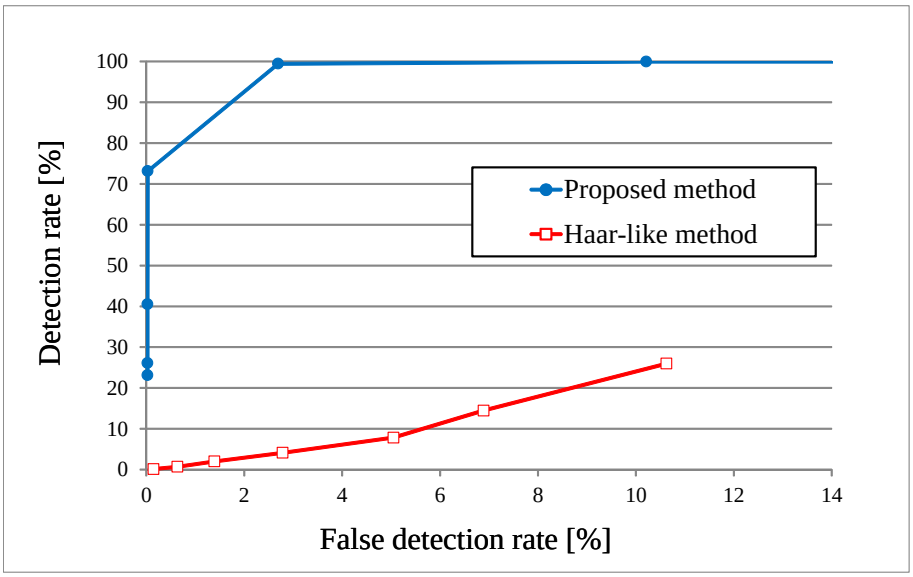


図 4.16 顔検出における ROC 曲線

4.2.2 鮮明度の違いによる顔検出率の評価

(1) 実験環境

画像の鮮明度合を変化させた場合の 4.1 節の顔検出手法について性能評価を行った。カメラはモータを内蔵し、レンズの位置をソフトウェアにて操作した [56]。カメラの外観を図 4.17 に示す。このカメラではレンズの位置を 1000 段階で操作することができる。このカメラにて解像度 640 × 480 pixel、30 fps にて撮影を行い、USB2.0 で PC へ接続し、処理を行った。表 4.2 に実験に用いたカメラの仕様を示す。また、照明は天井に設置された蛍光灯のみであった。

図 4.17 レンズ位置制御カメラ [56]

表 4.2 レンズ位置制御カメラ仕様

Type	RGB
Resolution [pixel]	640 × 480
Frame rate [fps]	30
Gradation [bit]	8
Lens position resolution [μm]	1

カメラから 1 m 離れたところに人が座り、静止した状態で撮影を行った。この実験においては、人が動いている状態では撮影は行わず、静止状態のみで撮影を行った。カメラのレンズ位置制御機能を用いて、レンズ位置を 0.3 mm 刻みで被写体側に移動させて撮影を行った。撮影した画像の一例を図 4.18 - 図 4.21 に示す。図 4.18 はレンズの焦点が合った状態の画像で、図 4.19 は焦点が合った状態からレンズを 0.3 mm 前方に移動させた状態の画像である。図 4.20、図 4.21 はそれぞれ焦点が合った状態からレンズを 0.6 mm、0.9 mm 前方に移動させた状態の画像である。ここでは焦点が合った画像を 0.0 mm と表記し、0.0 mm、0.3 mm、0.6 mm、0.9 mm 前方にレンズを移動させた状態の画像を用いた。提案手法は、640 × 480 pixel 画像を 80 × 80 pixel の大きさで 48 個の領域に分割した。画像撮影は 90 秒間行い、最初の 30 秒間は DFT の結果が出力されないため、30 秒間の検出率は評価の対象外とし、残りの 60 秒間の 1800 枚の画像についての検出率の評価を行った。また、5 人の被験者に対して評価を行った。

既存手法との比較のために、4.2.1 項と同様に顔検出手法の代表的な手法である Haar-like 手法 [25] と提案手法の比較を行った。Haar-like 手法の最小探索窓は 40 × 40 pixel とし、探索窓を逐次拡大する倍率は 1.1 と設定した。Haar-like 手法は MIT の顔データセット [57] を用いて学習を行った。正解画像として顔画像を 2600 枚、不正解画像として背景画像を 5200 枚使用した。また、レンズ位置が 0.3 mm、0.6 mm、0.9 mm の画像は焦点が合っておらず、ぼけた画像であるため、学習画像も擬似的にぼけた画像を使用した。MIT の顔データセット画像に対して、ガウシアンフィルタの処理を行うことにより画像を平滑化し、擬似的に焦点が合っていない画像を生成した。式 (4.1) にガウシアンフィルタを示し、式 (4.2) にガウシアンフィルタを用いた畳み込み処理を示す。

$$U(i,j) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{i^2+j^2}{2\sigma^2}\right) \tag{4.1}$$

$$V_{t_n}(x,y) = \sum_{j=-\omega}^{\omega}\sum_{i=-\omega}^{\omega} I_{t_n}(x+i,y+j)U(i,j) \tag{4.2}$$

式 (4.1) において、i と j はフィルタ内の画素位置を表し、$U(i,j)$ がガウシアンフィルタを表す。σ は標準偏差を表し、σ の値が大きくなると平滑化の効果が大きくなる。式 (4.2) において、I_{t_n} は入力画像を示し、V_{t_n} が出力画像を示す。ω はガウシアンカーネルの大きさを表す係数で、$(2\omega+1)\times(2\omega+1)$ がカーネルの大きさを表す。σ は ω の半分の値とした。ガウシアンカーネルの大きさを変えるこ

とにより、ぼけ具合を変化させ、それぞれのレンズ位置に応じた学習画像を用意した。Haar-like 手法の顔検出器は、学習により合計で四つ作成し、それぞれのレンズ位置に応じた顔検出器を使用した。ガウシアンフィルタは OpenCV (Ver. 2.4.3) で作成した。

図 4.18 焦点の合った画像

図 4.19 レンズ位置 0.3 mm 前方画像

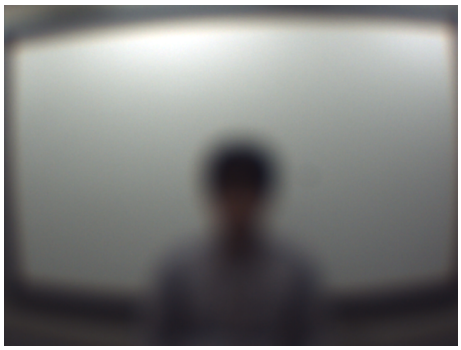


図 4.20 レンズ位置 0.6 mm 前方画像

図 4.21 レンズ位置 0.9 mm 前方画像

(2) 実験結果

鮮明度の違いによる顔検出率実験の結果を ROC 曲線で表す。提案手法による顔検出結果を図 4.22 - 図 4.25 に示す。図 4.22 は焦点が合った画像を使用した結果である。提案手法の ROC 曲線は左上に近い点では検出率 93.0 %、誤検出率 0.1 %と高い性能を示している。また、Haar-like 手法は、左上に近い点で検出率 75.5 %、誤検出率 0.3 %という性能を示している。図 4.22 の結果から提案手法は焦点があった通常の画像において、顔検出が可能であることが分かる。図 4.23 はレンズ移動量 0.3 mm の画像を使用した結果である。提案手法は左上に近い点で検出率 99.8 %、誤検出率 0.1 %という高い性能を示した。Haar-like 手法は非常に低い検出率となり、検出率 7.5 %、誤検出率 7.1 %という性能であった。Haar-like 手法は画像の明度差を利用する手法であるが、鮮明度の低い画像では、顔の特徴がうまく学習できず、顔が抽出できなかったのではないかと考えられる。図 4.24 はレンズ移動量 0.6 mm の画像を使用した結果である。提案手法の ROC 曲線は左上の角の部分が無くなり、高い検出率の場合には誤検出率も高くなってしまう結果となった。左上に近い点では検出率 74.5 %、誤検出率 0.1 %という性能を示した。また、Haar-like 手法は低い検出率となり、検出率 4.8 %、誤検出率 26.3 %という結果であった。図 4.25 はレンズ移動量 0.9 mm の画像を使用した結果である。左上に近い点では検出率 73.9 %、誤検出率 2.6 %という性能を示した。この ROC 曲線を見ると、提案手法は全体的に検出率が増加していないことがわかる。これはレンズ移動量を増やして、画像の鮮明度を過度に低下させると、提案手法の性能が低下することを示している。また、Haar-like 手法は全体的に低い検出率で、検出率が最も高い部分で検出率 6.6 %、誤検出率 36.8 %という結果であった。

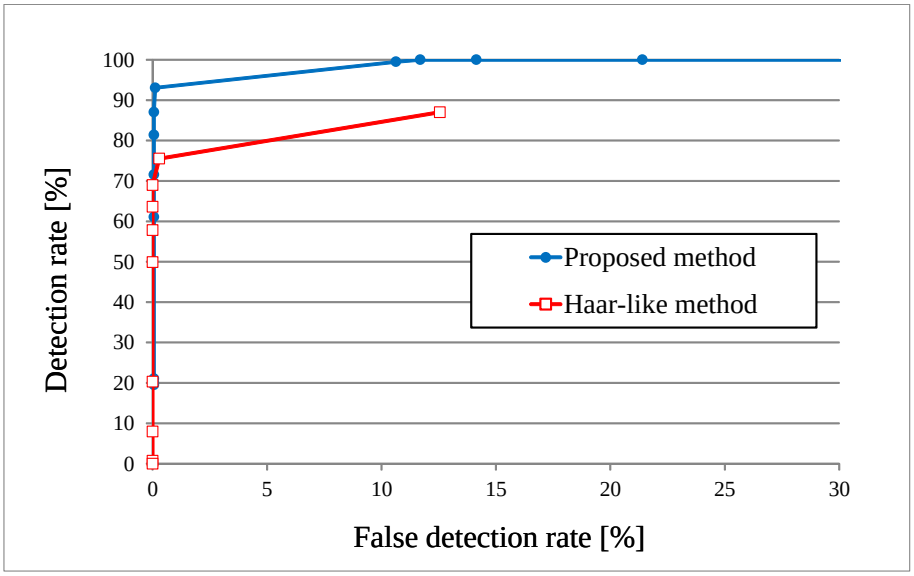


図 4.22　0.0 mm 画像における ROC 曲線

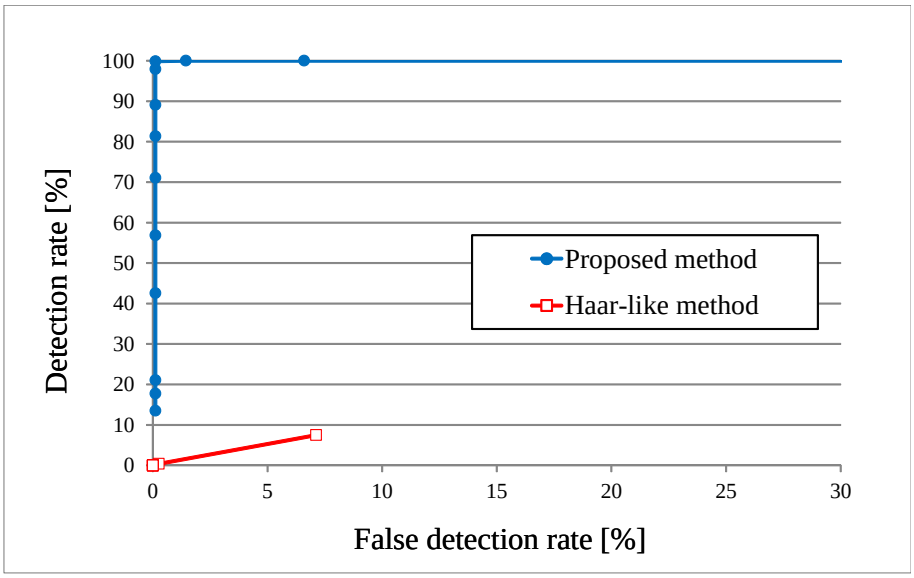


図 4.23　0.3 mm 画像における ROC 曲線

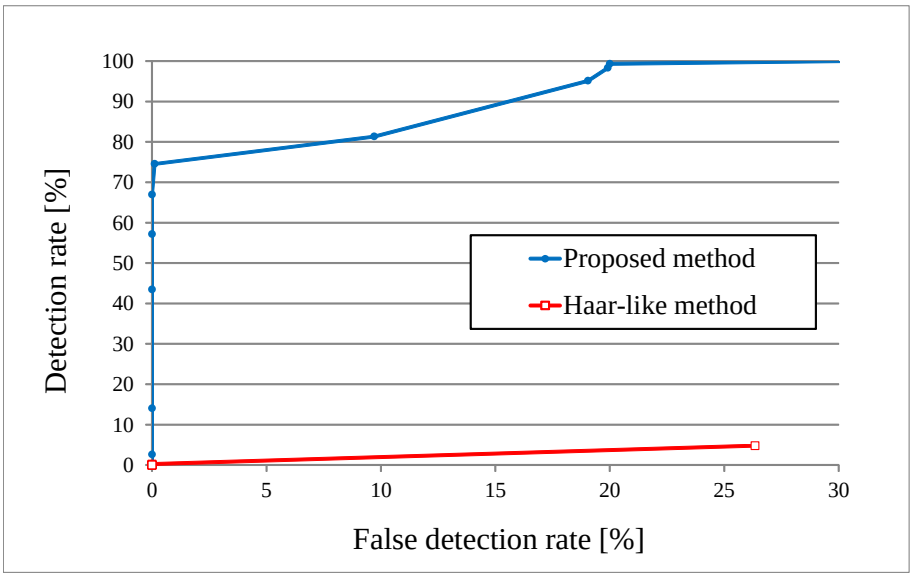


図 4.24 0.6 mm 画像における ROC 曲線

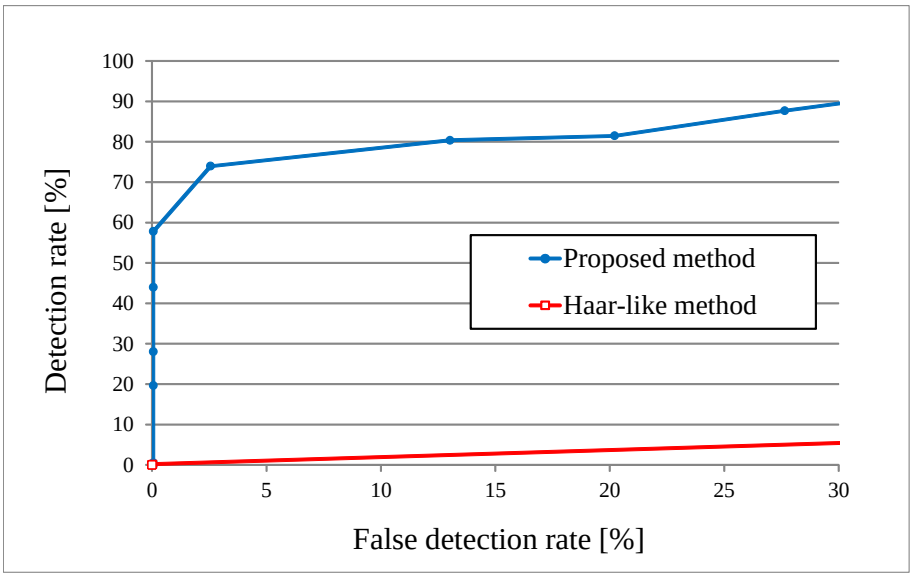


図 4.25 0.9 mm 画像における ROC 曲線

これらの結果から、提案手法はレンズ位置を移動させて意図的に鮮明度合を低下させた画像において顔検出が可能であることを示した。また、鮮明度を落とした画像では、既存手法である Haar-like 手法では検出性能が大きく低下することが確認できた。また、提案手法は焦点が合っている通常の画像においても、顔検出が可能であることを示した。さらにレンズ移動量の増加に伴い、提案手法の

検出性能が低下することが確認できた。これは提案手法は既存手法に比べて不鮮明な画像でも顔検出が可能ではあるが、提案手法にも適切な鮮明度合の範囲があるということを示しており、顔検出システムを開発する上では非常に有用な情報が得られた。

4.3 顔位置検出手法

顔検出手法である 4.1 節の手法だけでは顔の位置検出ができないため、本節は 4.1 節の検出結果を用いて、顔の位置を推定する手法について説明する。まず 4.1 節で出力される結果は人の存在と脈拍が検出された領域である。図 4.26 は 4.1 節の出力であり、矩形は脈拍が検出された領域を示し、中の数字は脈拍数を表している。この段階で検出された矩形の中心を顔の位置とすることもできるが、画像を分割する大きさによって、矩形が顔の左右に均等に出現しない場合や一つだけしか出現しない場合もあるため、矩形の中心位置を補正せずに使用すると位置精度の正確性に欠ける。そこで検出された複数の矩形の中心位置を基準として、その周辺領域を重複しながら探索することにより位置精度の向上を行う。

画像中に複数の矩形が出現することがあるため、隣接している矩形は一つの塊とするラベリング処理を行う。ラベル領域が複数存在する場合には、面積が最大のラベル領域を選択する。そして、そのラベル領域の重心位置を基準位置とする。求められた基準位置の周辺領域について脈拍検出を行う。この時に、脈拍検出のための矩形領域の中心位置 (x, y) と矩形の一辺の大きさ w をパラメータとして変化させて周辺領域の脈拍検出を行う。矩形の中心位置 (x, y) は基準位置の周辺になるように設定し、矩形の一辺の大きさ w は 4.1 節における、画像分割のサイズを基準として大小に変化させる。図 4.27 に周辺領域の脈拍検出の矩形を表す。各矩形を区別するために、緑色の濃淡を用いて矩形を表示している。これらの領域に対して、4.1 節の処理を行い、パワースペクトラムが閾値を超えて、脈拍が検出された領域については、2 次元平面に投票を行う。投票空間は入力画像と同じ解像度の平面を用意する。脈拍が検出された領域内のすべての画素位置において投票を行う。この投票を図 4.27 で示した複数の矩形領域で行うと、図 4.28 となる。領域の明度が投票数を表しており、白い領域が投票が多い画素である。投票数が一定の大きさを超えた領域のみを抽出し、その領域の重心を顔の位置とする。

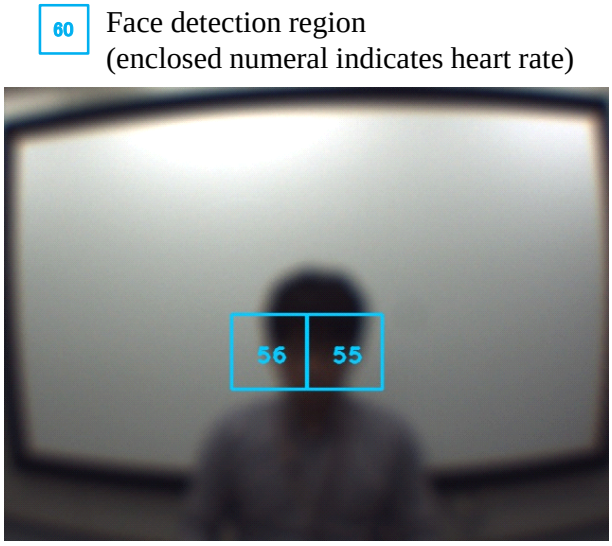


図 4.26 顔検出結果

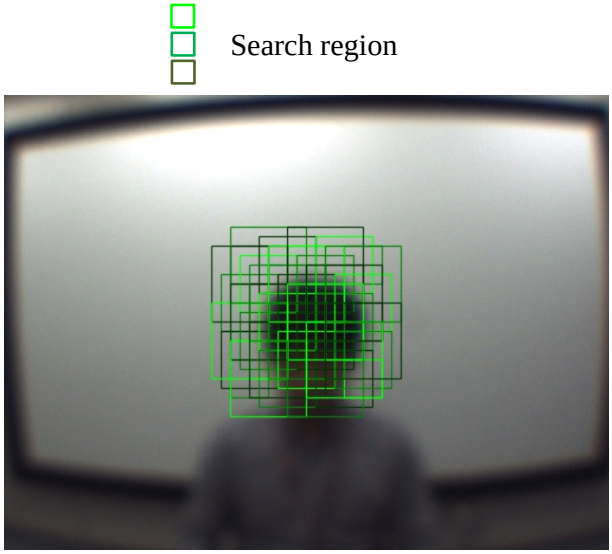


図 4.27 再探索領域

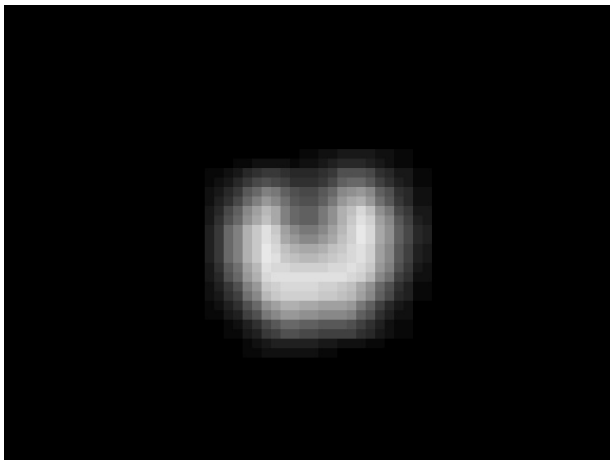


図 4.28 投票平面

4.4 顔位置精度評価

(1) 実験環境

4.3 節で提案した顔位置検出手法の位置精度の計測を行った。実験環境は 4.2.2 項と概ね同じである。図 4.17 のカメラを使用し、カメラから 1 m 離れたところに人が座り、静止した状態で撮影を行った。また、5 人の被験者に対して計測を行った。この実験では位置精度の検証を行うため、被験者はカメラに対して顔を向けた状態で、0 度、10 度、20 度で距離 1 m の位置に座り、撮影を行った。カメラに正対する正面位置を 0 度とする。一つの角度に対して、画像撮影は 90 秒間行い、最初の 30 秒間は評価の対象外とし、残りの 60 秒間の 1800 枚の画像についての位置精度の評価を行った。カメラのレンズの移動量は 0.3 mm で固定し、すべての画像を撮影した。顔位置の真値は、鼻の中心位置を基準として人手により与えた。

(2) 実験結果

提案手法により出力された顔の中心位置と真値の誤差をとることにより誤差を計測した。図 4.29 に位置精度計測の結果を示す。顔の中心位置は画像中の画素位置として出力され、真値との誤差も一旦は画素数にて出力されるが、一般性を持たせるために、画角と解像度の関係を用いて角度に変換している。図 4.29 の

縦軸が真値との誤差であり、灰色の棒グラフが提案手法 4.1 節で検出された矩形の中心をそのまま使用した結果となっている。4.3 節の提案手法にて中心位置を補正した結果を黒い棒グラフで表している。棒グラフの値は平均誤差を表しており、エラーバーは各値の標準偏差を表している。図 4.29 では水平方向、垂直方向ともに 4.3 節の提案手法で位置を補正することにより、ばらつき具合を示す標準偏差が小さくなっていることがわかる。この結果から、提案手法の位置補正の効果が確認できる。また、提案手法で補正された角度は、水平方向で平均誤差 0.26 度、標準偏差 0.48 度、垂直方向で平均誤差 0.51 度、標準偏差 0.62 度という高性能な精度を示した。この結果により、提案手法では匿名化画像から高精度に顔位置を検出できることが確認できた。

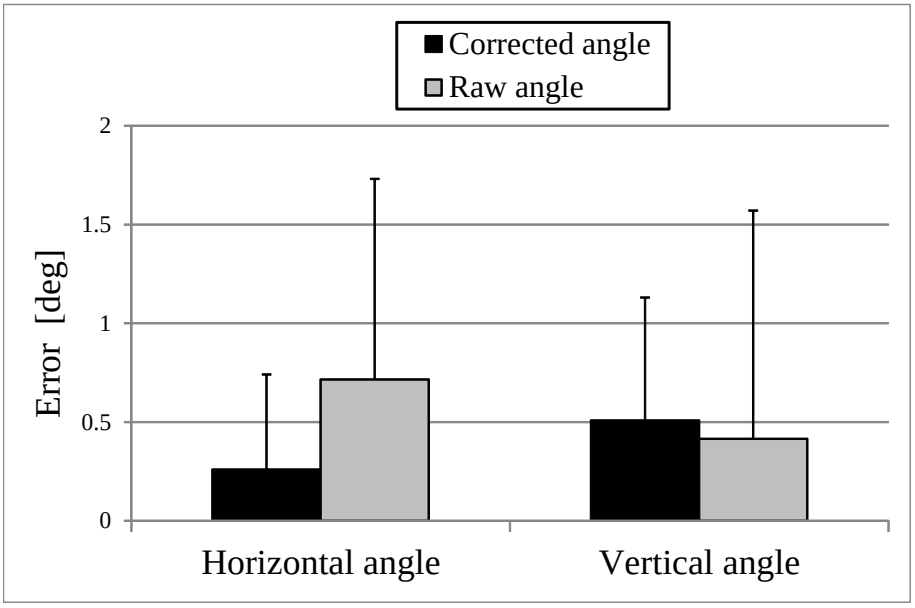


図 4.29 顔位置精度計測結果

4.5 まとめ

使用者の匿名性を保護した顔検出手法を提案した。被撮影者が誰であるか特定できないような不鮮明な匿名化画像に対して、画像を領域分割し、各領域に対して脈拍数を算出することにより、人が存在するかどうかの判定を行った。性能評価では、三つの匿名化画像に対して既存の顔検出手法と比較し、提案手法の方が高い性能を示し、匿名化画像に対して有効な顔検出手法であることを確認した。段階的に鮮明度合を変化させた画像についても既存の顔検出手法と比較し、提案

手法の方が高い性能結果を示した。さらに、画像の鮮明度合が低下するに従い、提案手法の顔検出性能も降下することが確認できた。また、鮮明度合を変化させていない通常画像においても提案手法と既存の顔検出手法と同様に顔検出できることを確認した。また、匿名化画像から人を検出する手法に加えて、位置を高精度に検出する手法を提案した。計測結果では、位置の補正前と補正後では位置精度が大きく向上したことが確認できた。提案手法により、使用者の匿名性を保持しながら顔を検出することができ、さらに顔の位置を高精度に検出できることを確認した。しかしながら、本章の手法では静止状態の顔しか検出できないため、運動状態での検出への対応を行う必要がある。

第5章

人追跡システム

本章では、人追跡システムについて述べる。5.1 節では 4 章で記述した脈拍による顔検出手法を基に、パーティクルフィルタを用いて人追跡を行う方法について述べる。5.2 節では追跡性能の評価を行う。4 章と本章で用いる匿名化画像は本論文で定義した画像であるため、5.3 節では、匿名化画像の匿名性を検証する。5.4 節でまとめを行う。

5.1 人追跡手法への拡張

4 章の顔検出手法は静止状態の人のみを対象としていたが、動いている人を検出するためにパーティクルフィルタの枠組みを用いて追跡手法へと拡張する。静止状態と運動状態の両方を追跡するために、動き検出と脈拍検出を尤度関数に用いる。さらに計算時間を短縮させるために積分画像を用いる。各要素技術について以下に述べる。

5.1.1 動き検出手法

人の動きを検出するためにフレーム間差分を用いる。3.1 節の式 (3.2) と式 (3.3) においてフレーム間差分を行っており、本手法においても同様の処理を行う。3.1 節と重複する部分もあるが、フレーム間差分について記述する。

時間差が Δt となる時刻 t_{n-1} と t_n に撮影された画像を $I_{t_{n-1}}$,I_{t_n} とする。$t_{n-1} + \Delta t = t_n$ となる。画像 $I_{t_{n-1}}$,I_{t_n} の画素値は RGB の 3 値から濃淡値に変換した値を使用する。式 (5.1) に変換式を示す。

$$I_{t_n}(x, y) = 0.299 \times r_{t_n}(x, y) + 0.587 \times g_{t_n}(x, y) + 0.114 \times b_{t_n}(x, y) \tag{5.1}$$

x と y は画像 I_{t_n} 中の横方向位置、縦方向位置をそれぞれ表す。
$r_{t_n}(x,y), g_{t_n}(x,y), b_{t_n}(x,y)$ は座標 (x,y) の赤、緑、青の輝度値をそれぞれ表し、$I_{t_n}(x,y)$ が座標 (x,y) の濃淡値を表す。二つの画像の差分値である $J_{t_n}(x,y)$ は式（5.2）で定義できる。

$$J_{t_n}(x,y) = |I_{t_n}(x,y) - I_{t_{n-1}}(x,y)| \tag{5.2}$$

この差分値から画素値を二つのクラスに分類することができる。差分の結果から動きの範囲において 2 値化処理を行う。式 (5.3) において T_a は閾値を表し、$D_{t_n}(x,y)$ は 2 値化された画素値を表す。

$$D_{t_n}(x,y) = \begin{cases} 1 & if \quad J_{t_n}(x,y) > T_a \\ 0 & otherwise \end{cases} \tag{5.3}$$

図 5.1 は時刻 t_{n-1} に撮影された画像 $I_{t_{n-1}}$ を、図 5.2 は時刻 t_n に撮影された画像 I_{t_n} を表しており、人が左から右へ移動している。図 5.3 は 2 枚の画像の差分画像であり、式 (5.2) の $J_{t_n}(x,y)$ の値を画像で表現している。白い画素は $J_{t_n}(x,y)$ の値が大きいことを意味している。

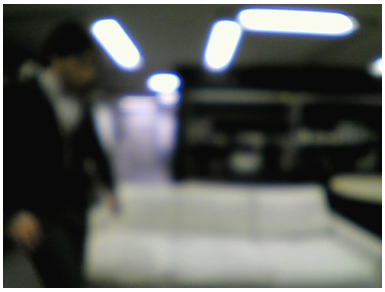


図 5.1 画像 $I_{t_{n-1}}$

図 5.2 画像 I_{t_n}

図 5.3 差分画像

対象領域において $D_{t_n}(x, y)$ の値が 1 となる画素が多ければ、その領域では人の動きがあったと判断できる。

5.1.2 パーティクルフィルタ

本研究では処理時間と追跡の頑健性を考慮して、追跡にはパーティクルフィルタ (粒子フィルタ、モンテカルロ法とも呼ばれる)[58, 59, 60, 61, 62, 63] を導入する。パーティクルフィルタとは、状態推定法の一つであり、状態空間中の多数の粒子により状態の分布を近似し、粒子の値の数値計算により状態分布の時間更新を行う。画像中の対象を追跡する手法として知られており、頑健な追跡を行う場合に使用される [64, 65]。以下のような非線形性および非ガウス性を考慮した状態空間モデルを対象としてパーティクルフィルタによる状態推定を考える。時系列データ $\boldsymbol{B}_{t_n}, n = \{1, 2, ..., N\}$ を所与として非線形非ガウス状態空間モデルは式 (5.4)、式 (5.5) のように定式化される。

$$\boldsymbol{\alpha}_{t_n} = f(\boldsymbol{\alpha}_{t_{n-1}}, \boldsymbol{\nu}_{t_n}) \tag{5.4}$$

$$\boldsymbol{\beta}_{t_n} = h(\boldsymbol{\alpha}_{t_n}, \boldsymbol{\omega}_{t_n}) \tag{5.5}$$

時刻 t_n における対象物体の状態パラメータを状態ベクトル $\boldsymbol{\alpha}_{t_n}$ とし、時刻 t_n において観測された特徴を観測ベクトル $\boldsymbol{\beta}_{t_n}$ とする。$f(\boldsymbol{\alpha}_{t_{n-1}}, \boldsymbol{\nu}_{t_n})$ は、時刻 t_{n-1} と時刻 t_n との状態変化を表す状態遷移関数であり、$\boldsymbol{\nu}_{t_n}$ は或る分布に従って発生

するシステム雑音である。また、$h(\boldsymbol{\alpha}_{t_n}, \boldsymbol{\omega}_{t_n})$ は、状態 $\boldsymbol{\alpha}_{t_n}$ のときに得られる観測ベクトルを表す関数で、$\boldsymbol{\omega}_{t_n}$ は或る分布にしたがって発生する観測雑音である。

時刻 t_1 から t_N までの観測ベクトルの集合を $\boldsymbol{B}_{t_n} = \{\boldsymbol{\beta}_{t_1}, ..., \boldsymbol{\beta}_{t_N}\}$ としたとき、状態推定の問題は $\boldsymbol{\alpha}_{t_n}$ の事後分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_n})$ を推定する問題として定式化できる。観測値 $\boldsymbol{B}_{t_n}$ から状態 $\boldsymbol{\alpha}_{t_n}$ の事後分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_n})$ を直接推定することは困難であるため、パーティクルフィルタではベイズの定理から逐次的な問題に置き換えて推定する。事後分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_n})$ はベイズの定理により式 (5.6) のように表される。

$$p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_n}) = \frac{p(\boldsymbol{\beta}_{t_n}|\boldsymbol{\alpha}_{t_n})p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_{n-1}})}{p(\boldsymbol{\beta}_{t_n}|\boldsymbol{B}_{t_{n-1}})} \tag{5.6}$$

$p(\boldsymbol{\beta}_{t_n}|\boldsymbol{\alpha}_{t_n})$ は、或る状態 $\boldsymbol{\alpha}_{t_n}$ のときに、観測値 $\boldsymbol{\beta}_{t_n}$ を得る尤度であり、式 (5.5) の観測モデルによって与えられる。また、分母の $p(\boldsymbol{\beta}_{t_n}|\boldsymbol{B}_{t_{n-1}})$ は、状態 $\boldsymbol{\alpha}_{t_n}$ とは無関係な項であり、実際の問題では正規化定数 k_{t_n} と置き換えることができる。従って、事後分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_n})$ は事前分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_{n-1}})$ と尤度 $p(\boldsymbol{\beta}_{t_n}|\boldsymbol{\alpha}_{t_n})$ の積で表すことができる。

パーティクルフィルタは各フレームに対して対象物体を追跡する際、繰り返し計算を下記の過程に分けて実行している。

1. 初期化
   対象の初期位置から、パーティクルを生成

2. 予測
   パーティクルを状態遷移モデルに基づき移動

3. 観測
   各パーティクルにおいて尤度を計算

4. 再標本化
   尤度に比例して、パーティクルを再標本化し、[2. 予測] を実行

図 5.4 にアルゴリズム処理手順の概要を示す。図中の白丸は各パーティクルを表している。一時刻前の事後分布である $p(\boldsymbol{\alpha}_{t_{n-1}}|\boldsymbol{B}_{t_{n-1}})$ の各パーティクルを状態遷移モデル（後述 5.1.2) に基づき移動させて、事前分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_{n-1}})$ とする。次に、

そのときの状態 $\boldsymbol{\alpha}_{t_n}$ における尤度 $p(\boldsymbol{\beta}_{t_n}|\boldsymbol{\alpha}_{t_n})$ を尤度関数（後述 5.1.2) に基づき算出する。求めた事前分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_{n-1}})$ と尤度 $p(\boldsymbol{\beta}_{t_n}|\boldsymbol{\alpha}_{t_n})$ から事後分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_n})$ を算出する。事後分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_n})$ に従うように各パーティクルを再標本化して、次の時刻の処理に進む。

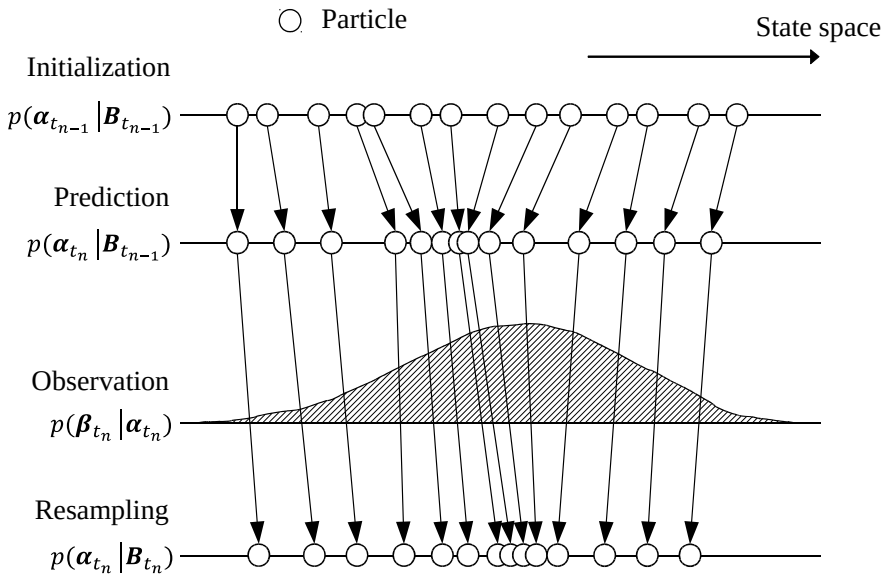


図 5.4 パーティクルフィルタのアルゴリズム概要

状態遷移モデル

式 (5.4) で示した状態ベクトル $\boldsymbol{\alpha}_{t_n}$ の要素を以下のように設定する。

$$\boldsymbol{\alpha}_{t_n} = \{x_{t_n}, y_{t_n}, u_{t_n}, v_{t_n}, w_{t_n}\} \tag{5.7}$$

$$\boldsymbol{\xi}_{t_n} \sim N(0, \Sigma_5) \tag{5.8}$$

$$\boldsymbol{\xi}_{t_n} = \{\xi_{1t_n}, \xi_{2t_n}, \xi_{3t_n}, \xi_{4t_n}, \xi_{5t_n}\} \tag{5.9}$$

$$x_{t_{n+1}} = x_{t_n} + u_{t_n}\Delta t + \xi_{1t_n} \tag{5.10}$$

$$y_{t_{n+1}} = y_{t_n} + v_{t_n}\Delta t + \xi_{2t_n} \tag{5.11}$$

$$u_{t_{n+1}} = u_{t_n} + \xi_{3t_n} \tag{5.12}$$

$$v_{t_{n+1}} = v_{t_n} + \xi_{4t_n} \tag{5.13}$$

$$w_{t_{n+1}} = w_{t_n} + \xi_{5t_n} \tag{5.14}$$

$\boldsymbol{\alpha}_{t_n}$ は式 (5.7) に示すように五つの要素で表される。x_{t_n}、y_{t_n} は時刻 t_n における対象の二次元位置を表しており、画像の横方向が x、縦方向が y である。ここでは $t_n + \Delta t = t_{n+1}$ となり、u_{t_n}、v_{t_n} は対象の x 方向速度、y 方向速度をそれぞれ表しており、対象は等速直線運動を仮定している。w_{t_n} は探索領域を示しており、画像上の領域の大きさを意味している。式 (5.8) は、平均 0 で共分散が Σ_5 の 5 次元正規分布に従って、$\boldsymbol{\xi}_{t_n}$ を無作為に生成することを意味する。この $\boldsymbol{\xi}_{t_n}$ を式 (5.10) から式 (5.14) において各要素に加算している。これはランダムウォークと呼ばれ、無作為性を有しているため複雑な変化がある場合に有効とされている。

図 5.5 に事前分布 $p(\boldsymbol{\alpha}_{t_n}|\boldsymbol{B}_{t_{n-1}})$ の状態ベクトル $\boldsymbol{\alpha}_{t_n}$ を可視化した図を示す。画像中の各矩形領域の中心が各パーティクルの位置である x_{t_n}、y_{t_n} を表しており、矩形の大きさが探索領域の大きさ w_{t_n} を表している。u_{t_n}、v_{t_n} はパーティクルの速度を示すので、画像中には表れていない。図 5.5 では、各矩形の中心位置が一点を中心に散らばり、様々な大きさの矩形があることから x_{t_n}、y_{t_n}、w_{t_n} に無作為性が反映されていることが確認できる。

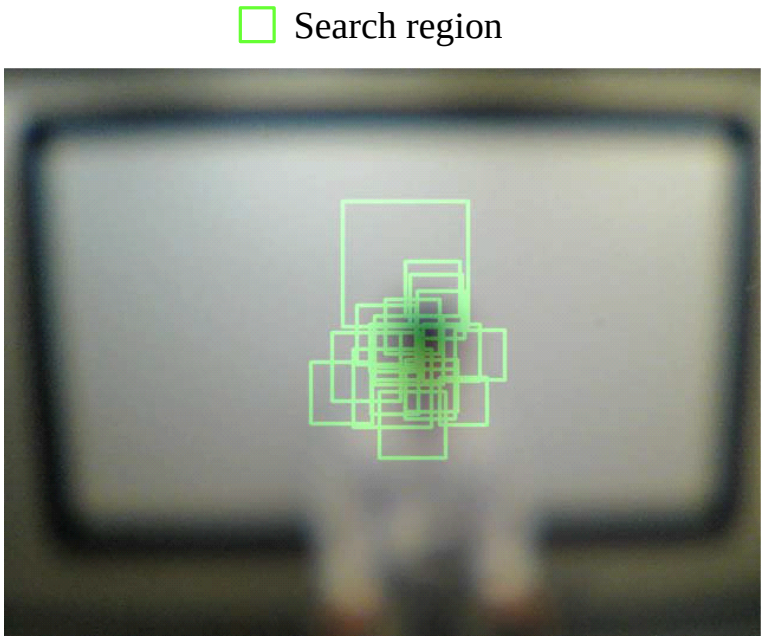


図 5.5 各パーティクルにおける探索範囲の一例

尤度関数

観測時に尤度を得る尤度関数には、人の動きと脈拍の二つの要素を取り入れる。人の動きがある場合には、一定領域内での動きを示す画素数の比率が高い方が尤度が高くなるように設計する。人の動きが無い場合には、領域内の脈拍数が設定した脈拍数に近づくほど、尤度が高くなるように設計する。動きを示す画素の割合は式 (5.15) で示される。

$$r_{t_n} = \frac{Q_{t_n}}{C_{t_n}} \times 100 \tag{5.15}$$

Q_{t_n} は対象領域内において式 (5.3) の $D_{t_n}(x, y) = 1$ となる画素の数を表している。対象領域とは図 5.5 における一つの矩形領域を指す。C_{t_n} は対象領域内のすべての画素数を表しており、r_{t_n} は対象領域内における動きを示す画素の占有面積率を表す。人の動きの有無によって尤度計算に使用する ϵ_{t_n} の値を変化させる。

$$\epsilon_{t_n} = \begin{cases} |100 - r_{t_n}| & if \quad r_{t_n} > T_b \\ |HR_{re} - HR_{t_n}| & otherwise \end{cases} \tag{5.16}$$

式 (5.16) では、式 (5.15) で求めた r_{t_n} が閾値 T_b を超えた場合には、動きの画素の占有面積率に応じて尤度が高くなるように ϵ_{t_n} の値を入力する。r_{t_n} が閾値 T_b を超えない場合には、脈拍数の情報を使用する。HR_{t_n} は対象領域から算出された脈拍数を表しており、HR_{re} は事前に設定される脈拍数である。計測された脈拍数 HR_{t_n} が事前設定された脈拍数 HR_{re} に近ければ、尤度が高くなるように ϵ_{t_n} に値を入力する。ϵ_{t_n} による尤度関数を式 (5.17) に示す。

$$L(\epsilon_{t_n}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\epsilon_{t_n}^2}{2\sigma^2}\right) \tag{5.17}$$

式 (5.17) は、0 を平均、σ を標準偏差として持つ正規分布である。各パーティクルはこの分布に従って追跡を行うこととなる。

積分画像

パーティクルフィルタを使用して追跡を行う場合には、図 5.5 に示すように複数の領域に対して脈拍検出を行う。パーティクルの数に応じて脈拍検出を繰り返し行うと、処理コストが大きくなり、実時間処理が困難となる。処理時間短縮のため、積分画像 (Integral Image) [66, 67] を使用することにより、高速化を行う。積分画像は以下の式で定義される。

$$II_{t_n}(x, y) = \sum_{x'' \leq x, y'' \leq y} I_{t_n}(x'', y'') \tag{5.18}$$

II_{t_n} は積分画像であり、I_{t_n} は元画像である。画像位置 (x, y) の積分画像は左上を原点とした場合に (x, y) より上方と左側の画素値の合計となる。図 5.6 における S_D の領域、つまり矩形 $(x - W, y - H) - (x, y)$ の領域から輝度値の和を求める場合には、式 (5.19) より算出することができる。

$$S_D = II_{t_n}(x, y) - II_{t_n}(x - W, y) - II_{t_n}(x, y - H) + II_{t_n}(x - W, y - H) \tag{5.19}$$

積分画像を事前に計算しておくことで、任意の矩形領域内の画素値合計を数回の加減算で求めることができる。脈拍検出においては RGB 値の内、G (緑) の値のみについて事前に積分画像を計算しておくことにより、領域内の平均値の計算が高速に行える。

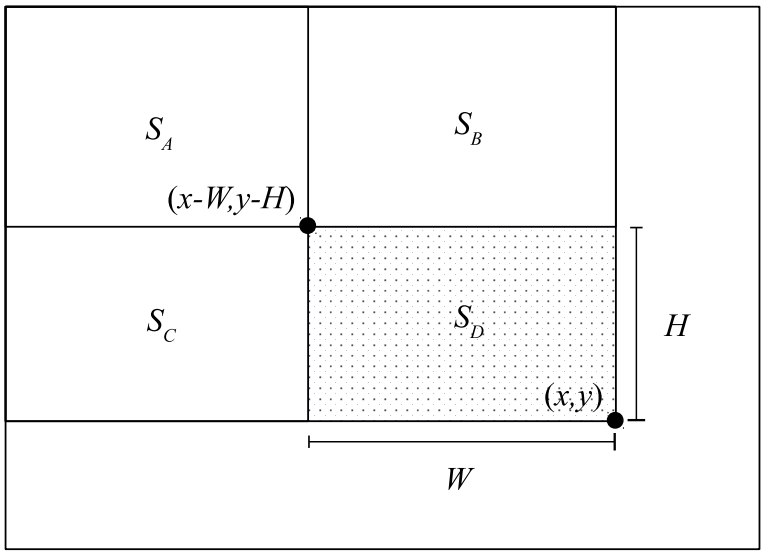


図 5.6 積分画像

5.2 人追跡手法の性能評価

5.2.1 実験環境

提案手法の人追跡について性能評価を行った。カメラは図 4.17 に示すカメラを用いた。モータを内蔵し、レンズの位置をソフトウェアにて操作することができる。このカメラにて解像度 640 × 480 pixel、30 fps にて撮影を行い、USB2.0 で PC へ接続し、画像を取得した。PC はこれまでの実験と同様に表 3.1 の機材を用いた。

居間でテレビを見ている人をカメラで追跡することを想定して、実験を行った。また、匿名化画像とするために、レンズを 0.3 mm 前方に移動させて個人が特定できない程度まで画像をぼかして撮影を行った。図 5.7 に実験の様子を示す。室内の照明は天井に設置された蛍光灯のみであり、窓から日差しが入る環境の室内にて撮影を行った。また、10 人の被験者に対して評価を行った。人が動いている状態と静止している状態の両方を追跡対象とするために、カメラから 1 m 離れたところに被験者が座り、被験者には 30 秒に一回座る位置を変えるように指示を行った。図 5.8 に被験者の動作の様子を示す。

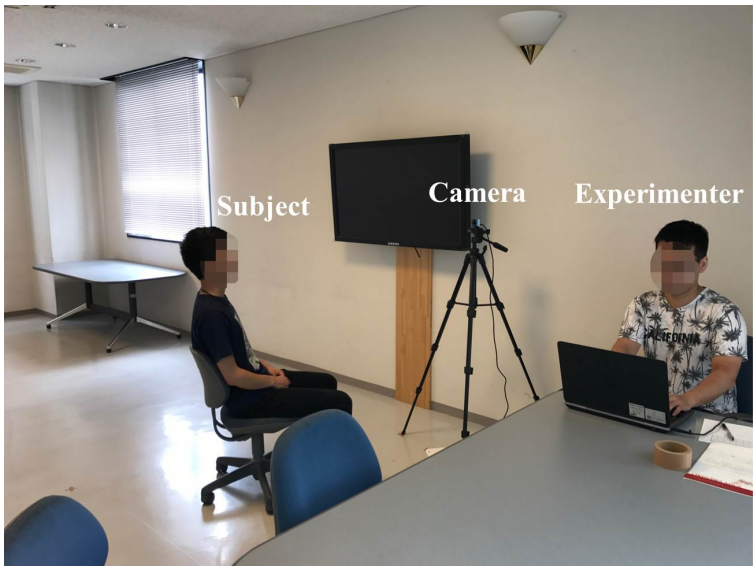


図 5.7 性能評価テストの実験環境

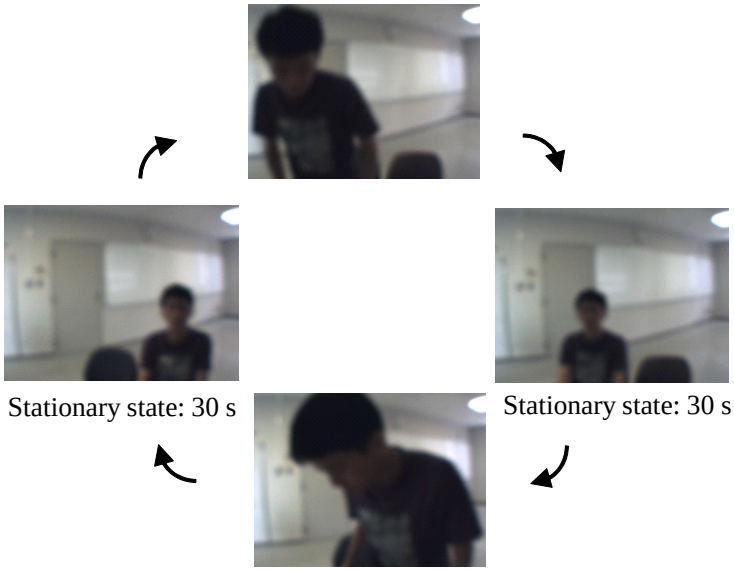


図 5.8 被験者の動作

提案手法のパーティクルの数は実時間処理を考慮して 200 個とし、探索矩形領域の最小値を 40 × 40 pixel 、最大値を 160 × 160 pixel とした。また、式 (5.16) で事前設定される脈拍数 HR_{re} は 70 bpm に設定し、DFT の時間窓の幅は 8 秒と設定した。画像撮影は被験者一人当たり 90 秒間行い、合計で 27000 枚の画像を使用した。

比較手法として Mean shift 法 [68, 69, 70] と Haar-like 特徴量を用いた顔検出手

法 [25] の二つを選択した。Mean shift 法は追跡の枠組みであり、画像処理では対象の色ヒストグラムを追跡する手法として使用されることが多く、顔の追跡にも用いられる (図 5.9)。この手法では画像の勾配やエッジ情報を必要としないため、不鮮明な画像でも動作すると予測される。しかしながら、Mean shift 法は指定された領域を追跡する手法であり、人を検出する能力が無いため、初期フレームにおいて人手で顔の位置を指定した。提案手法は人検出の機能と追跡の機能の両方を有しているため、提案手法との完全な比較とはならないが、追跡機能のみの比較対象として Mean shift 法を使用することとした。

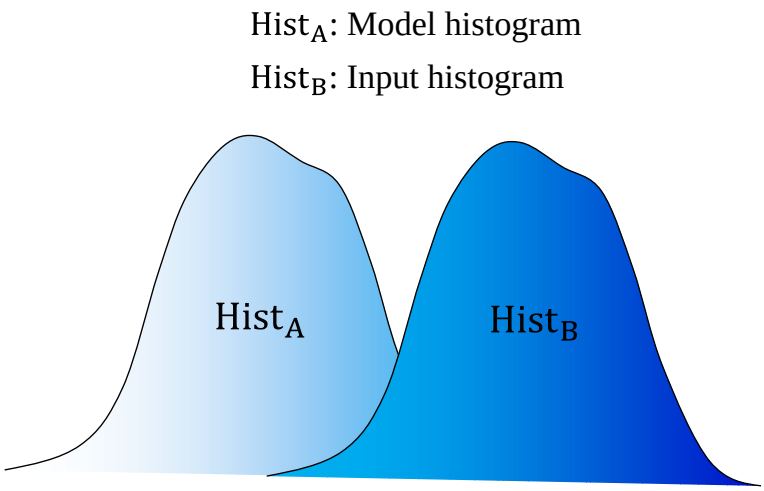


図 5.9 Mean shift 法概要

二つ目の比較手法である Haar-like 手法は、4.2 節で用いた手法と同様である。画像の明度差を利用して顔を検出する手法であり、顔検出の代表的な手法として知られている。全画面を対象領域として顔検出を行い、1 フレーム前と現フレームで検出された顔位置を距離の近さで対応付けることにより、顔の追跡を行った。Mean shift 法と Haar-like 手法は OpenCV (Ver. 2.4.3) で作成した。

検出の判定基準について説明する。まず撮影された画像に対して、人が目視にて真値となる人の領域を設定する。人追跡手法によって算出された人領域の矩形が真値の領域と少しでも重複していれば、人追跡は成功したと判定する。お互いの領域が重複していなければ、誤検出として判定する。今回の撮影では 1 フレームに一人しか存在していないため、1 フレームを 1 試行とする。また、位置誤差計測では、人手により真値として顔の中心点を入力する。そして、人追跡手法により検出された顔位置と真値との誤差を計測する。提案手法の検出領域と検出位

置については尤度における上位 20 個のパーティクルの平均値を代表値として用いた。

5.2.2 実験結果

提案手法と比較手法による人追跡の結果を表 5.1 に示す。成功率は全フレームに対して追跡が遂行されたフレームの割合を示している。成功率では提案手法が最も高い性能を示していることがわかる。Haar-like 手法は、ぼけた不鮮明な画像では顔の特徴を捉えることができず、多くのフレームで顔検出が失敗し成功率が低い値となった。一方で Mean shift 法は被験者の動きに追従することができ、高い成功率を示した。

図 5.10 に追跡位置の誤差を示す。顔の中心位置は画像中の画素位置として出力され、真値との誤差も一旦は画素数にて出力されるが、一般性を持たせるために、画角と解像度の関係を用いて角度に変換している。図 5.10 の縦軸が真値との誤差であり、角棒グラフの値は平均誤差を表している。また、誤差範囲として各値の標準偏差を示している。比較手法の Haar-like 手法は背景の壁を顔として検出することがあり、水平方向、垂直方向の両方で平均誤差が大きい値となった。提案手法と比較手法の Mean shift 法の水平方向の平均誤差は総じて同じであるが、標準偏差は提案手法の方が小さい値となっている。垂直方向の平均誤差では、比較手法の Mean shift 法と比べて提案手法の方が 2.6 度小さく、標準偏差も提案手法の方が小さい値となっている。比較手法の Mean shift 法は追跡対象の初期位置と領域を人手で入力した状態から追跡を開始しており、提案手法よりも有利な条件であったが、提案手法の方が高精度な性能を示した。

表 5.1 人追跡成功率

	Success rate
Proposed method [%]	90.1
Mean shift method [%]	84.8
Haar-like method [%]	10.0

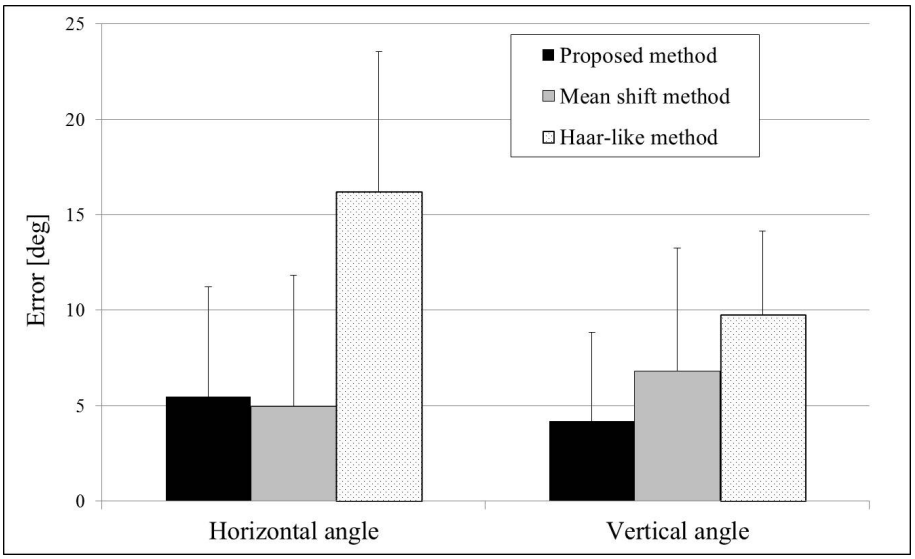


図 5.10 追跡位置精度計測結果

本項の実験結果で提案手法と比較手法の Mean shift 法で精度に差が現れたことについて考察を行う。一人の被験者データに対する水平方向と垂直方向の誤差の時間変化を図 5.11、図 5.12 にそれぞれ示す。図 5.11、図 5.12 における 200 秒、500 秒、800 秒付近で被験者は座る位置の移動を行っている。移動の瞬間には提案手法と比較手法の Mean shift 法はそれぞれ誤差が増大している。そして、その次の瞬間には誤差が減少している。これは立ち上がった瞬間に追跡が外れてしまったが、徐々に対象の位置に戻っていることを示している。しかし、比較手法の Mean shift 法は 800 秒前後で誤差が増大し、その後は一定値となっている。これは追跡が失敗して、その後も追跡が回復しなかったことを意味している。比較手法の Mean shift は人検出機能が無いため、対象から追跡領域が大きく外れてしまうと回復ができない。しかし、提案手法は動きと脈拍を検知することにより人検出を行っているため、追跡が外れても回復が可能である。図 5.11、図 5.12 の 800 秒以降では、被験者の移動直後に提案手法の誤差が増大し、その後も誤差が大きい状態が続くが、徐々に誤差を減らし追跡を回復させている。この回復機能の有無により、提案手法と比較手法の Mean shift 法で性能に差が現れたと考えられる。

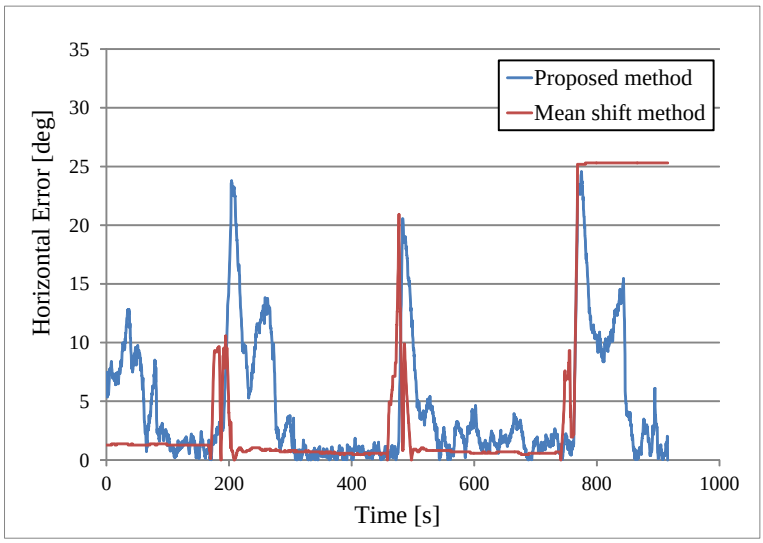


図 5.11 人追跡中の水平方向誤差の時間変化

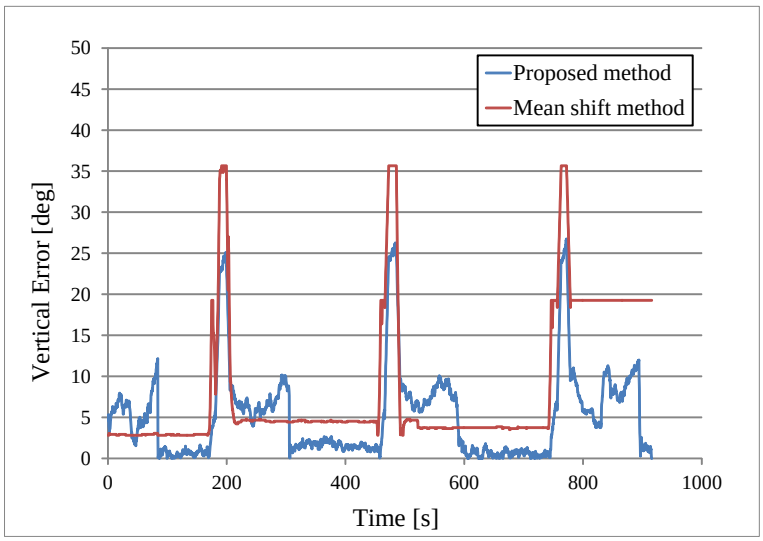


図 5.12 人追跡中の垂直方向誤差の時間変化

5.2.3 処理時間計測

提案手法は 5.1.2 項で積分画像を用いて高速化を行っている。その効果を確認するために、1 フレーム当たりの処理時間を計測した。計測には 5.2.1 項で撮影した一人の被験者のデータ 2700 枚の画像を使用した。計測された平均処理時間と標準偏差を表 5.2 に示す。提案手法で積分画像を使用しない場合は処理時間が

474.7 ms であり、積分画像を使用すると 32.1 ms となる。この結果から処理時間が大幅に短縮され、積分画像による高速化の効果が確認できる。積分画像を使用しない場合は標準偏差の値が大きくなっている。これは矩形の対象領域が大きくなったときに処理時間が増加していることが考えられる。積分画像では処理範囲の大きさによって処理時間が変わることがないため、標準偏差は小さい値となっている。また、32.1 ms はカメラ画像を 30 fps で取得する処理サイクルの 33.3 ms を下回っており、実時間処理が可能な処理時間である。3 章にて提案した脈拍計測手法は、3.2.4 項にて高速な処理時間を示していた。本章の人追跡手法は脈拍計測を繰り返し実行しているため、脈拍計測の高速性が人追跡の実時間処理に貢献していると考えられる。

表 5.2 人追跡 1 フレーム当たりの処理時間

	Average	Standard deviation
Proposed method [ms]	32.1	5.6
Proposed method without integral image [ms]	474.7	188.4

5.2.4 姿勢変化への対応

居間でテレビを見ている人を追跡の対象とすると、座っているだけではなくソファで横になってくつろぐ場面も想定される。そのような場面における提案手法の追跡性能の確認を行った。図 5.13 の (a)-(i) に提案手法の追跡結果画像を示す。図 5.13 において赤枠は動きを感知したパーティクル、緑枠は脈拍を検知したパーティクル、緑枠の中の数字は脈拍数を表している。また、青枠は各パーティクルの結果から推定された追跡結果領域を、青丸は追跡結果範囲の中心点を表している。図 5.13 (b) では、被験者が部屋に入ってきた瞬間であり、赤枠で示された動きに反応するパーティクルによって追跡されていることがわかる。図 5.13 (c) で被験者はソファに座るが、座った直後はパーティクルが分散しており、広範囲を探索していることがわかる。図 5.13 (d) では脈拍が検知され、パーティクルが顔の周辺に集中していることがわかる。また、図 5.13 (e)-(f) で被験者がソファで横になった場合には、動きを感知し、その後脈拍を検知することにより顔の位置を追跡していることがわかる。図 5.13 (g)-(h) では被験者が部屋から退出する動

きを検知して被験者の位置を追跡できていることがわかる。これらの結果から、提案手法では人が横たわっても追跡が可能であることが確認できる。

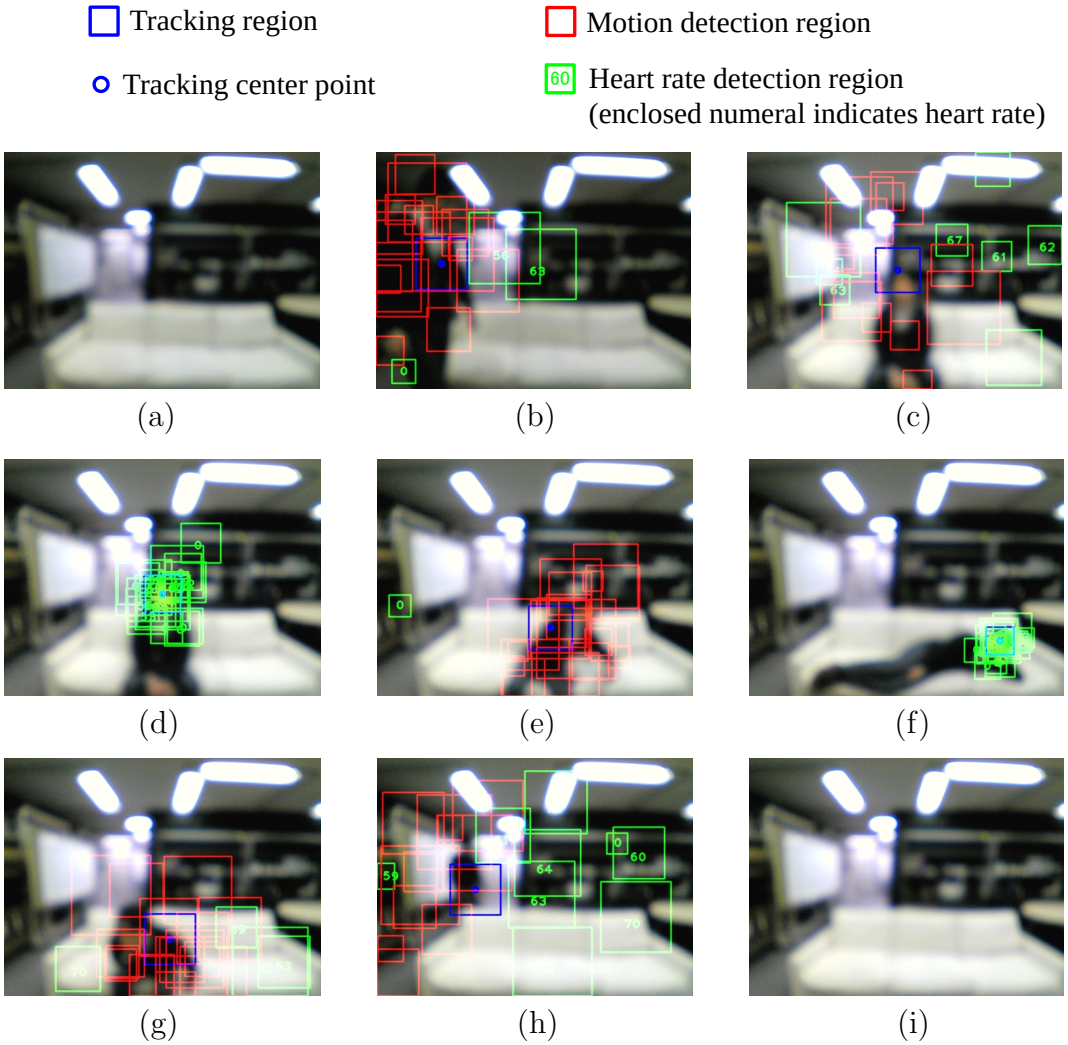


図 5.13 人がソファに横たわるシーンにおける追跡結果

5.3 匿名性検証

本研究では光学的に不鮮明にした画像を匿名化画像として扱ったが、本節にて匿名化画像の匿名性を検証する。5.3.1 項では、ソフトウェアの後処理にて不鮮明化した画像を用いて、個人特定が可能となる鮮明度合を調査する。5.3.2 項では、ソフトウェアにて不鮮明にした画像と光学的に不鮮明にした画像の比較を行い、光学的匿名化画像の匿名性を検証する。さらに、5.3.3 項では、撮影距離が変化し

た場合の匿名性について議論する。

5.3.1 個人特定実験

段階的に鮮明度合を変化させた画像を使用し、どの段階の鮮明度から個人が特定されるのかを検証した。個人特定用の画像として、世界の著名人の画像が掲載されている“Labeled Faces in the Wild”[71] のデータベースを用いた。このデータベースの画像は画像の大きさが 250 × 250 pixel で統一されており、画像の中央に著名人の顔が配置され、顔の大きさは 100 × 150 pixel から 150 × 200 pixel までの範囲で統一されている。被験者が既知である著名人を予測して、データベースから 20 人の著名人画像を抽出した。著名人にはスポーツ選手や政治家等が含まれる。5 人の被験者に対して実験を行い、被験者が著名人の事を既知ではない場合には、試行回数には含めず無効とした。著名人画像はガウシアンフィルタを用いて段階的に鮮明度合を変化させた。ガウシアンフィルタとは、式 (4.1)、式 (4.2) に示したように、ガウス関数を用いて画像をぼかす処理である。ガウシアンフィルタは OpenCV (Ver. 2.4.3) で作成した。OpenCV に実装されているガウシアンフィルタには、画像の縦方向と横方向のカーネルサイズと標準偏差がそれぞれパラメータとして設定されており、カーネルサイズは奇数値のみ入力できる。本検証において、標準偏差はカーネルサイズの半分の数値を入力し、段階的に鮮明度合を変化させた。図 5.14 に鮮明度を変化させた画像例を示す。ただし、画像例は著名人の画像ではない。

図 5.14 ガウシアンフィルタによる鮮明度変化

不鮮明な画像から順番に被験者に提示し、個人が特定できた時点でのガウシアンフィルタのカーネルサイズを記録した。被験者 5 名による結果を表 5.3 に示す。

表 5.3 個人特定実験結果

	Average	Maximum
Gaussian kernel size [pixel]	12.1 × 12.1	21 × 21

平均値のカーネルサイズは 12.1 × 12.1 pixel であった。平均値は図 5.14 の“13 × 13”の画像のぼけ具合に近い値を示しており、顔の器官である目、鼻、口の鮮明さが増すと個人が特定される可能性が高くなると考えられる。個人特定の能力に優れている被験者は、“21 × 21”の顔器官が鮮明でない画像から、推測により個人を特定することができた。この結果から、画像サイズが 250 × 250 pixel で中央に顔が配置されている画像において、ガウシアンフィルタのカーネルサイズが 21 × 21 pixel を超過する画像は匿名性が確保されていると言える。

5.3.2 光学的匿名化画像の匿名性

本研究における光学的匿名化画像は、レンズを 0.3 mm 前方に移動させて撮影した画像を用いた。その光学的匿名化画像と 5.3.1 項にて実施した処理にて鮮明度合を変化させた画像との対応関係を得ることにより、光学的匿名化画像の匿名性を検証する。図 4.17 のレンズ位置を制御できるカメラを用い、カメラから 1 m 離れた壁に写真を貼り付けて、撮影を行った。レンズは 0.3 mm 前方に移動させ、壁に貼る写真の大きさは人の顔の大きさと同じになるように調整を行った。画像は 640 × 480 pixel の大きさで撮影されるが、5.3.1 の実験と画像サイズを揃えるために、中心部分の 250 × 250 pixel を切り取って実験を行った。図 5.15 に焦点が合った画像を、図 5.16 にレンズを 0.3 mm 前方に移動させて撮影した画像をそれぞれ示す。

図 5.15 レンズ位置 0.0 mm 画像

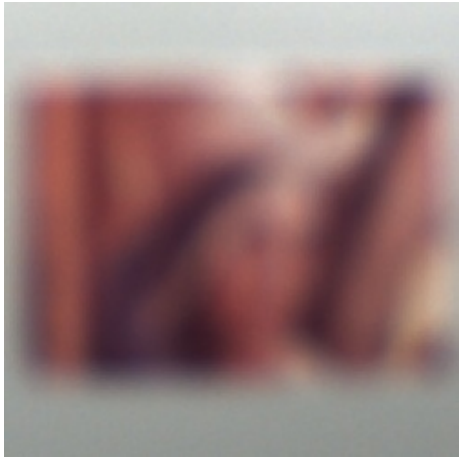


図 5.16 レンズ位置 0.3 mm 画像

図 5.15 のレンズ位置 0.0 mm の画像に対して、5.3.1 項と同様にガウシアンフィルタによる処理を行い、どのガウシアンフィルタのカーネルサイズが図 5.16 のレンズ位置 0.3 mm の画像に相当するかを調査する。図 5.15 のレンズ位置 0.0 mm の画像にガウシアンフィルタを用いて段階的に鮮明度合を変化させた画像を図 5.17 に示す。

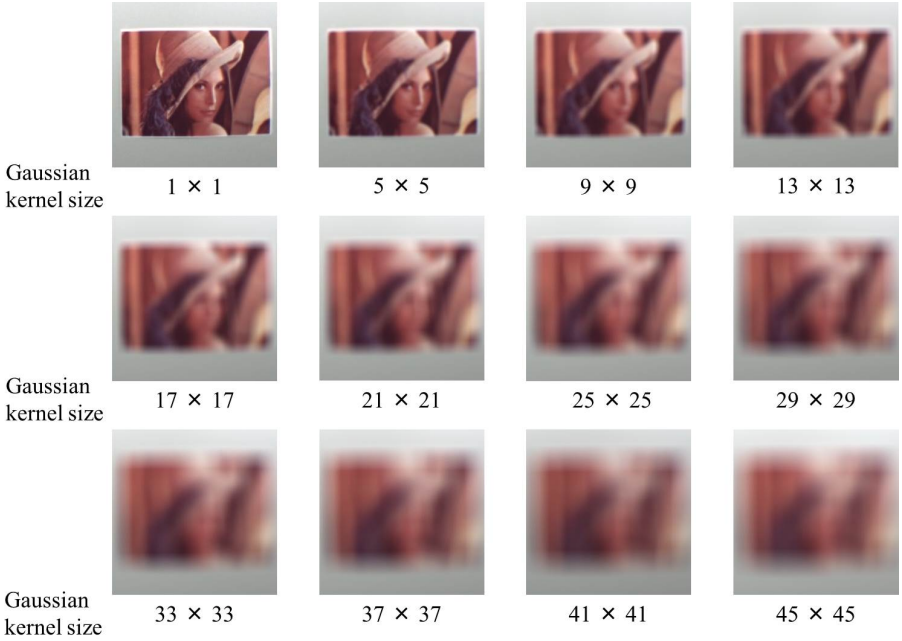


図 5.17 レンズ位置 0.0 mm 画像の鮮明度変化

図 5.16 のレンズ位置 0.3 mm の画像と図 5.17 の各画像の類似度を計測するために、SAD (Sum of absolute difference) を用いた。SAD は画素値の差分の絶対値の和を表しており、二つの画像の類似度が高ければ、SAD の値は低くなる。図 5.18 にレンズ位置 0.3 mm の画像とレンズ位置 0.0 mm の鮮明度を変化させた画像の類似度を示す。縦軸が SAD であり、横軸がガウシアンフィルタのカーネルサイズを表している。

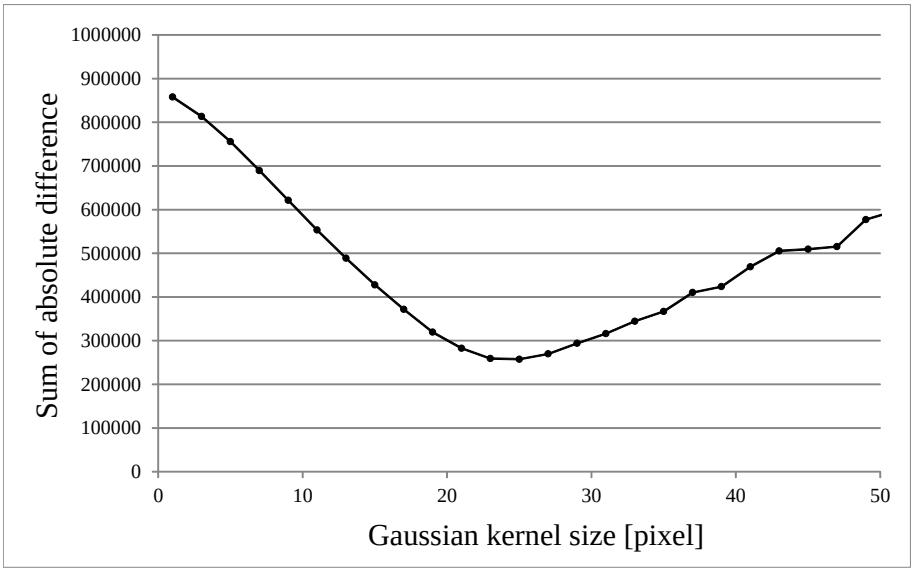


図 5.18 鮮明度変化による類似度

図 5.18 の結果から、カーネルサイズが 25 × 25 の時に SAD の値が最小であることが分かる。つまり、レンズ位置 0.0 mm の画像にカーネルサイズ 25 × 25 のガウシアンフィルタを適用した画像とレンズ位置 0.3 mm の画像は、カーネルサイズを変化させた画像群の中で類似度が最も高いという事を示している。5.3.1 項の結果から、カーネルサイズが 21 × 21 を超過する画像においては匿名性が確保されると言えるため、カーネルサイズ 25 × 25 の画像は匿名性が確保されていると言える。さらに本項の結果から、レンズ位置 0.3 mm の画像はカーネルサイズ 25 × 25 の画像に相当し、カーネルサイズ 21 × 21 の画像よりも不鮮明であると言えるので、レンズ位置 0.3 mm の画像は匿名性を有していると言える。

5.3.3 距離変化における匿名性

5.3.2 項において、カメラと被写体の距離は 1 m として匿名性の検証を行ったが、カメラと被写体の距離が極端に近い場合は焦点が合い、匿名性が確保できない可能性がある。本研究では、汎用レンズを前方に移動させることにより、不鮮明な画像を撮影したため、被写体の距離が変化すると焦点が合う可能性が生じるが、図 5.19 に示すような焦点の合わない撮像システムを使用することにより、この問題を防ぐことができる。対象とする距離を例えば 0.3 m から 5 m までと設定する場合は、その距離範囲内の不鮮明さが大きく変化しないように、レンズを設計することが可能である。専用レンズの使用により、距離変化による匿名性は確保できると考えられる。

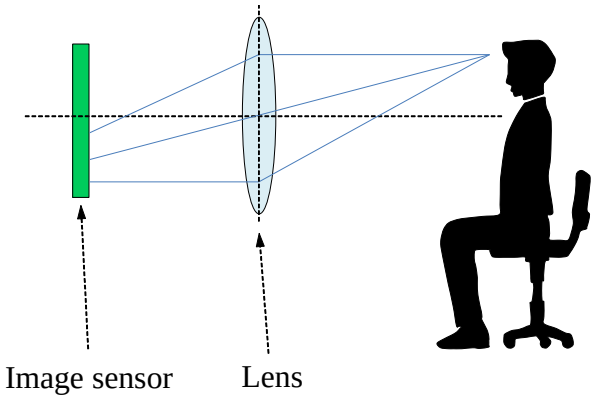


図 5.19 焦点の合わない撮像システム

5.4 まとめ

匿名化画像を用いた追跡手法について提案を行い、性能評価を行った。個人が特定できない不鮮明な画像に対して、動いている人の追跡を行うために、パーティクルフィルタの枠組みを導入して、動き検出と脈拍検出を利用した人追跡手法を提案した。性能評価では、レンズの焦点を移動させた不鮮明な画像に対して既存の Mean shift 法による追跡手法と Haar-like 手法の二つの手法を用いて比較を行った。その結果、Haar-like 手法は不鮮明な画像では、総じて顔を検出することができなかった。また、Mean shift 法には人を検出する仕組みが無いため人手にて初期位置を与えて、性能比較を行ったが、追跡成功率と位置誤差については提案手法の方が高い性能を示した。さらに、提案手法は積分画像を用いて高速化を行っ

ており、実時間処理が可能であることが確認できた。加えて、提案手法は人がソファに横たわるような姿勢変化する場面においても追跡が可能であることを確認した。性能評価の結果から、提案手法は匿名化画像において、有効な人追跡手法であることが確認できた。また、匿名化画像の匿名性の検証を行った。個人特定が可能となる鮮明度合を調査し、光学的匿名化画像がその鮮明度合よりも不鮮明であることを確認した。これにより、レンズを前方に 0.3 mm 移動させた匿名化画像には匿名性が確保されている事を確認した。

第6章

結論

本章では、6.1 節で本研究での提案と性能評価結果についてまとめを行い、将来の展望を述べる。6.2 節では、本研究の社会的貢献について述べる。

6.1 匿名化人センシングの実現

IoT 家電製品は、インターネットへの常時接続と人の位置に応じたサービスを提供することにより利便性を向上させている。情報量が多く安価であるため、センサはカメラが選択される場合が増加している。しかし、個人が特定できる画像を撮影するため、画像が意図せずに流出し、使用者のプライバシを侵害するという問題がある。この問題に対処すべく、本研究では使用者の匿名性を保ったまま人センシングを行う手法を提案した。カメラの前にフィルタを置いたり、レンズの位置を移動させて画像をぼかす事により匿名化画像を生成し、この匿名化画像から人センシングを行った。人検出の要素技術としてカメラ画像からの脈拍計測を用いた。

非接触脈拍計測システムは、顔領域の緑色の平均値を算出し、平滑化処理、差分処理を行うことにより体動雑音を除去し、脈拍成分だけを抽出した。運動状態での実験においては従来研究よりも誤差が小さく、高い性能を示した。

脈拍による顔検出システムでは、画像領域を分割し、各領域に対して非接触脈拍計測の処理を行い、人相当の脈拍数が計測された領域には人の顔が存在すると判定する手法を提案した。検出率の性能評価においては静止状態の人に対し、検出率 99.5%、誤検出率 2.7%と高い性能を示し、匿名化画像から顔検出が可能であることを確認した。また、検出された顔位置の周辺を再探索することにより顔位置精度を向上させた。評価実験では誤差が 1 度以下となり、匿名化画像から高精度に顔位置を検出できることを確認した。

移動する使用者を追跡できるように、パーティクルフィルタの枠組みを導入して、人追跡システムへと拡張を行った。尤度関数に動き成分と脈拍成分を用いることにより、静止状態と運動状態の両方で追跡が可能となった。性能評価においては既存の追跡手法よりも提案手法の方が高い追跡率を示した。人がソファに横たわる場面においても人追跡が可能であり、姿勢変化に対応できることを確認した。また、本研究で使用した匿名化画像が匿名性を有している事を確認した。

本研究は匿名化画像から人検出と追跡が可能であり、且つ高精度な位置検出が可能であることを示した。提案手法は使用者の匿名性を保ちながら、高い空間分解能と高い搭載性を実現する人センシング手法であることを確認した。これらの結果から、人検出装置の特性分布を表す図 6.1 の目標としていた右上の位置に提案手法を位置付ける。

本研究では、すべての環境における人センシングを実現しておらず、限定された環境における有効性の確認を行った。さらなる性能向上に向けて、実環境で起こりうる状況を想定して実験を重ねていく必要がある。

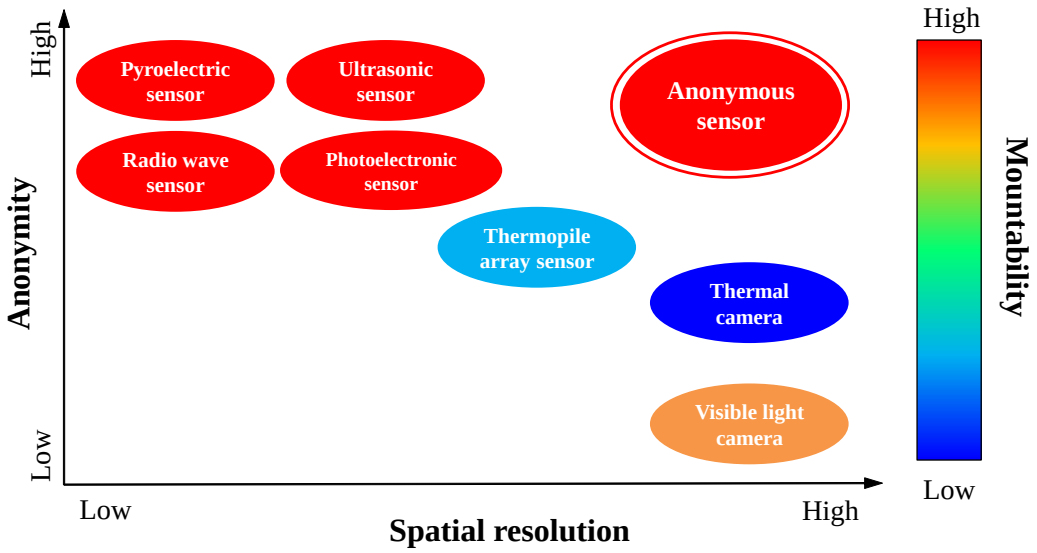


図 6.1 人検出装置特性分布図における提案手法

6.2 社会的展望

本研究の性能を向上させることにより、実用レベルの匿名化人センシングが実現できる。それにより、カメラ付き IoT 家電に人センシング技術を搭載すること

が可能となる。IoT 家電による画像流出を不安視して、購買を控えていた使用者に対して安心感を付与することができ、その結果、IoT 家電の普及に繋がる。IoT 家電は遠隔操作やソフトウェアの自動更新、クラウドサービス等の機能を有しており、非常に利便性の高い製品である。それらの普及を促進することにより、本研究が社会貢献に繋がると期待される。

また、家電製品だけでなく、常時ネットワークに接続された自動車や医療機器の普及が見込まれており、IoT 機器は今後増加すると予測される (図 6.2)。いずれの装置においても人センシング装置の需要は高く、プライバシの問題が発生する。さらに、IoT を発展させた IoE (Internet of everything) という“もの”だけではなく、人や空間、データ等のあらゆるすべてがインターネットに繋がる概念が提唱されており、近い将来、ネットワークによるプライバシの問題が深刻化すると予測される。この社会問題の解決に向けて、本研究が多くの人に活用され、人々の生活を豊かにする一助となることを期待する。

図 6.2 増加する IoT 機器

謝 辞

本研究は、大阪大学大学院基礎工学研究科 大城理 教授に、深い御理解と御指導をいただき進めてまいりました。業務との調整が難しく頻繁に通学できない中でも、社会人学生の立場を御配慮いただき、研究方針から論文執筆に至るまで多大なる御指導を賜りましたことに心より感謝の意を表すると共に、篤く御礼を申し上げます。

本研究を纏めるにあたって副査として多大なる御意見および御指導を賜りました大阪大学大学院基礎工学研究科 佐藤宏介 教授ならびに、大阪大学大学院基礎工学研究科 清野健 教授に深く感謝致します。

本研究を遂行するにあたり、多角的な視点から有益な御意見をいただきました、大阪大学大学院基礎工学研究科 黒田嘉宏 准教授に、深い感謝の意を表します。

日頃の研究活動から論文執筆に至るまで、様々な面において御助言をいただき、親身になって対応していただいた大阪大学大学院基礎工学研究科 吉元俊輔 助教に、深く感謝致します。

奈良先端科学技術大学院大学 名誉教授 木戸出正継 教授には、博士前期課程の時から御指導いただき、卒業後も進路に悩んだ時には相談に乗っていただきました。“企業人も博士号を取得すべきである”の教えが博士後期課程進学へと導いてくれました。篤く御礼申し上げます。

豊田工業大学 浮田宗伯 教授には、博士前期課程の時から御指導いただき、卒業後も博士後期課程進学について相談に乗っていただきました。御助言が背中を押してくれました。深く感謝致します。

株式会社サムスン日本研究所 ET-Center 高木剛 Center 長には、博士号取得に向けて激励の言葉をいただき、大変励みになりました。深い感謝の意を表します。

株式会社サムスン日本研究所 ET-Center ET-2 Lab 山下伸二 Lab 長には、私が在学中に ET-2 Lab に異動したにもかかわらず、引き続き在学することを奨励していただきました。篤く御礼申し上げます。

株式会社サムスン日本研究所 ET-Center ET-2 Lab ET-22 Part
村上エドワード Part 長には、業務が立て込んでいる状態でも優先的に学業に取り組めるようにと、頻繁に業務調整を行っていただきました。また、日頃から研究の進捗を気にかけていただき、論文掲載が決まった時には我が事のように喜んで下さりました。感謝の念で一杯です。

三星電子株式会社 崔相彦 Master には、企業での研究内容を論文発表する事に否定的な意見がある中、論文発表を行うことを奨励していただき、さらに後押ししていただきました。論文発表の成果が博士号取得に繋がりました。深い謝恩を申し上げます。

株式会社サムスン日本研究所 横井敦也 主幹、立花裕美 主任研究員、株式会社テクノプロ 長谷川智隆 氏、三星電子株式会社 Young-Kwang Seo 首席には、同じプロジェクトで共に闘っていただきました。順調にプロジェクトが遂行できたおかげで、学業に取り組めました。深く感謝致します。

同じ社会人学生として互いを鼓舞し、高め合った、オムロンヘルスケア株式会社 中川雄樹 氏、京都大学医学部付属病院 朝野美穂 氏に、深い感謝の意を表します。

大城研究室の皆様には、展示発表の準備を御協力いただき、さらに評価実験に快く御協力いただきました。深く感謝致します。

両親には、大学から大学院への進学を快諾いただき、博士前期課程修了まで育てていただきました。修士号あってこその博士後期課程進学でした。心より御礼申し上げます。

楽しい会話でいつも心に安らぎを与えてくれた、長女 莉奈に心より感謝致します、ありがとう。

屈託の無い笑顔で毎日、元気を与えてくれた、長男 快利に心より感謝致します、ありがとう。

最後になりましたが、家事と育児を一手に担い、仕事と学業に専念させてくれた、妻 奈央に心より感謝致します。

なお、本論文における被験者実験は、大阪大学大学院基礎工学研究科 「人を対象とした研究に関する倫理委員会」 の承認を得ており、被験者には実験内容について説明を行い、書面にて同意を得られた場合にのみ測定を行った (認証番号 27-5)。

参考文献

[1] 東京地方裁判所 昭和 36 年 (ワ) 第 1882 号 昭和 39 年 9 月 28 日 判決

[2] Kevin Ashton, “That ‘Internet of Things’ Thing”, RFiD Journal, 2009.

[3] Harald Sundmaeker, Patrick Guillemin, Peter Friess, and Sylvie Woelffle, “Vision and challenges for realising the Internet of Things”, CERP IoT, 2010.

[4] Chin-Chi Cheng, and Dasheng Lee, “Smart Sensors Enable Smart Air Conditioning Control”, Sensors 14, pp. 11179-11203, 2014.

[5] Asaad Elmoudi, Omar Asad, Melike Erol-Kantarci, and Hussein Talaat Mouftah, “Energy Consumption Control of an Air Conditioner Using Web Services”, Smart Grid and Renewable Energy, pp. 255-260, 2011.

[6] Hanne Grindvoll, Ovidiu Vermesan, Tracey Crosbie, Roy Bahr, Nashwan Dawood, and Gian Marco Revel, “A Wireless Sensor Network for Intelligent Building Energy Management Based on Multi Communication Standards - A Case Study”, Journal of Information Technology in Construction, pp. 43-62, 2012.

[7] Armando Roy Delgado, Rich Picking, and Vic Grout, “Remote-Controlled Home Automation Systems with Different Network Technologies”, IEEE Embedded Computing, 2008.

[8] Netalee Efrat, Piotr Didyk, Mike Foshey, Wojciech Matusik, and Anat Levin, “Cinema 3D: Large Scale Automultiscopic Display”, ACM Transactions on Graphics, Vol. 35, No. 4, 2016.

[9] 牧川 方昭, 吉田 正樹, 南部 雅幸, 塩澤 成弘, 岡田 志麻, “ヒト心身状態の計測技術”, コロナ社, pp. 118-131, 2010, 東京.

[10] Jaeseok Yun, and Sang-Shin Lee, “Human Movement Detection and Identification Using Pyroelectric Infrared Sensors”, Sensors 14, pp. 8057-8081, 2014.

[11] Ji Xiong, Fang-Min Li, and Jing-Yuan Zhang, “Human Tracking System Based on PIR Sensor Network and Video”, Advanced Technologies in Ad Hoc and Sensor Networks 295, pp. 13-25, 2014.

[12] Saravana Kumar K, Priscilla P, Germiya K Jose, and Balagopal G, “Human Detection Robot using PIR Sensors”, IJSETR 4, pp. 492-496, 2015.

[13] Sun Qiao, Yang Wei, and Zhang WenDong, “A new Pyroelectric Sensor System for Target Detection and Recognition”, Applied Mathematics and Information Sciences 9, pp. 2621-2626, 2015.

[14] Photoelectronic sensor principle, Available: https://www.keyence.co.jp/ss/sensor/aboutsensor/photoelectric/info/ [Accessed: Jan. 2, 2018].

[15] Min-Young Kim, and Tae-Sung Oh, “Thermoelectric Characteristics of the Thermopile Sensors with Variations of the Width and the Thickness of the Electrodeposited Bismuth-Telluride and Antimony-Telluride Thin Films”, Materials Transactions 51, pp. 1909-1913, 2010.

[16] Richard Wotiz, “Infrared Thermal Detectors”, Circuit Cellar 260, pp. 54-58, 2012.

[17] Huaiwen Wu, Arvin Emadi, Ger de Graaf, and Reinoud F. Wolffenbuttel, “Thin film encapsulated 1D thermoelectric detector in an IR microspectrometer”, Proceedings of SPIE, 2010.

[18] Ultrasonic sensor principle, Available: https://www.honda-el.co.jp/hb/4_1.html [Accessed: Jan. 2, 2018].

[19] Chenshu Wu, Zheng Yang, Zimu Zhou, Xuefeng Liu, Yunhao Liu, and Jiannong Cao, “Non-Invasive Detection of Moving and Stationary Human With WiFi”, IEEE Journal on Selected Areas in Communications Vol. 33, No. 11, pp. 2329-2342, 2015.

[20] Thermal camera principle, Available: https://www.apiste.co.jp/contents/technical_note/technical/infrared_light/microbolometer.html [Accessed: Jan. 2, 2018].

[21] Mauricio Correa, Gabriel Hermosilla, Rodrigo Verschae, and Javier Ruiz-del-Solar, “Human Detection and Identification by Robots Using Thermal and Visual Information in Domestic Environments”, Journal of Intelligent & Robotic Systems 66, pp. 223-243, 2012.

[22] Fu-Feng Lee, Feng Chen, and Jing Liu, “Infrared Thermal Imaging System on a Mobile Phone”, Sensors 15, pp. 10166-10179, 2015.

[23] Peter Christiansen, Kim Arild Steen, Rasmus Nyholm Jorgensen, and Henrik Karstoft, “Automated Detection and Recognition of Wildlife Using Thermal Cameras”, Sensors 14, pp. 13778-13793, 2014.

[24] Visible camera principle, Available: http://hatukame.web.fc2.com/satsuzososhi.html [Accessed: Jan. 2, 2018].

[25] Paul Viola, and Michael Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 511-518, 2001.

[26] Navneet Dalal, and Bill Triggs, “Histograms of Oriented Gradients for Human Detection”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 886-893, 2005.

[27] Guangzhen Li, Yoshimichi Ito, Xiaoyi Yu, Naoko Nitta and Noboru Babaguchi, “Recoverable Privacy Protection for Video Content Distribution”, EURASIP Journal on Information Security, 2010.

[28] Elaine Newton, Latanya Sweeney, and Bradley Malin, “Preserving Privacy by De-identifying Facial Images”, IEEE Transactions on Knowledge and Data Engineering, Volume 17, Issue 2, pp. 232-243, 2005.

[29] Ikuhisa Mitsugami, Masayuki Mukunoki, Yasutomo Kawanishi, Hironori Hattori, and Michihiko Minoh, “Privacy-Protected Camera for the Sensing

Web”, International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, 2010.

[30] Frederic Dufaux, and Touradj Ebrahimi, “Scrambling for video surveillance with privacy”, In Proc. of IEEE Computer Vision and Pattern Recognition Workshop on Privacy Research in Vision, 2006.

[31] Wei Zhang, Sen-ching S. Cheung, and Minghua Chen, “Hiding privacy information in video surveillance system”, In Proc. of International Conference on Image Processing, 2005.

[32] Absorption spectrum of main substances in living tissue, Available: http://www.hamamatsu.com/jp/en/technology/innovation/trs/index.html [Accessed: Nov. 18, 2017].

[33] Correspondence between wavelength and color, Available: http://www.ieij.or.jp/what/yougo.html [Accessed: Nov. 18, 2017].

[34] Pulse wave sensor principle, Available: http://www.rohm.com/web/global/pulse-wave-sensor [Accessed: Nov. 18, 2017].

[35] iHealth Wireless Pulse Oximeter, Available: https://ihealthlabs.com/fitness-devices/wireless-pulse-oximeter/ [Accessed: Nov. 18, 2017].

[36] Mio ALPHA Heart Rate Sport Watch, Available: https://www.mioglobal.com/ [Accessed: Nov. 18, 2017].

[37] Ming-Zher Poh, Daniel J. McDuff, and Rosalind W. Picard, “Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam”, IEEE Transactions on Biomedical Engineering, Vol. 58, No. 1, pp. 7-11, 2011.

[38] Cardio heart rate monitor software, Available: http://www.cardiio.com [Accessed: Nov. 18, 2017].

[39] Zephyr BioHarness 3 Wireless Professional Heart Rate Monitor, Available: https://www.zephyranywhere.com/ [Accessed: Nov. 30, 2017].

[40] iHealth Wireless Pulse Oximeter, Available: https://ihealthlabs.com/ [Accessed: Nov. 30, 2017].

[41] Mio ALPHA Heart Rate Sport Watch, Available: https://www.mioglobal.com/ [Accessed: Nov. 30, 2017].

[42] Daisuke Nagae, and Atsushi Mase, “Measurement of Heart Rate Variability and Stress Evaluation by Using Microwave Reflectometry”, 11th Cross Straits Symp. on Materials, Energy and Environmental Eng, pp. 282-283, 2009.

[43] Hirooki Aoki, Ryo Furukawa, Shinsaku Hiura, Ryusuke Sagawa, and Hiroshi Kawasaki, “Extraction and Visualization of Cardiac Beat by Grid-Based Active Stereo”, Lecture Notes in Computer Science, Vol. 7854, pp. 146-157, 2013.

[44] Marc Garbey, Nanfei Sun, Arcangelo Merla, and Ioannis Pavlidis, “Contact-Free Measurement of Cardiac Pulse Based on the Analysis of Thermal Imagery”, IEEE Transactions on Biomedical Engineering, Vol. 54, No. 8, pp. 1418-1426, 2007.

[45] Microsoft LifeCam Studio, Available:https://www.microsoft.com/accessories/ja-jp/products/webcams/lifecam-studio/q2f-00021 [Accessed: Dec. 16, 2017].

[46] TM-204 Light Meter, Available:http://www.tenmars.com/webls-en-us/TM-204.html [Accessed: Dec. 16, 2017].

[47] Web Camera Buffalo BSW20KM11BK, Available:http://buffalo.jp/product/multimedia/web-camera/bsw20km11bk/ [Accessed: Dec. 18, 2017].

[48] Kenko ZS-CYCLONE, Available: http://www.kenko-tokina.co.jp/discontinued/filters/ef_around/4961607349403.html [Accessed: Dec. 18, 2017].

[49] Kenko ZS-RADIAL, Available: http://www.kenko-tokina.co.jp/discontinued/filters/ef_around/4961607349397.html [Accessed: Dec. 18, 2017].

[50] Tom Fawcett, “An introduction to ROC analysis“, Pattern Recognition Letters 27, pp. 861-874, 2006

[51] John A. Swets, “Measuring the Accuracy of Diagnostic Systems”, Science, New Series 240, pp. 1285-1293, 1988

[52] Sho Ikemura, and Hironobu Fujiyoshi, “Real-Time Human Detection using Relational Depth Similarity Features”, 10th Asian Conference on Computer Vision, 2010

[53] Amnon Shashua, Yoram Gdalyahu, and Gaby Hayun, “Pedestrian Detection for Driving Assistance Systems: Single-frame Classification”, In IEEE IV, pp. 13-18, 2004.

[54] Michael Mayo, “Effective Classifiers for Detecting Objects”, The Fourth International Conference on Computational Intelligence, Robotics and Autonomous Systems, 2007.

[55] Hongming Zhang, Wen Gao, Xilin Chen, and Debin Zhao, “Object detection using spatial histogram features”, Image and Vision Computing 24, pp. 327-341, 2006.

[56] Focus control camera, DFK 72AUC02-F, Available: https://www.theimagingsource.com/products/autofocus-cameras/usb-2.0-color/dfk72auc02f/ [Accessed: Dec. 20, 2017].

[57] CBCL Face Database #1MIT Center For Biological and Computation Learning Available: http://cbcl.mit.edu/software-datasets/FaceData2.html [Accessed: Dec. 20, 2017].

[58] Genshiro Kitagawa, “Monte Carlo filter and smoother for non-Gaussian nonlinear state space models”, Journal of Computational and Graphical Statistics 5 (1), pp. 1-25, 1996.

[59] Neil Gordon, David Salmond, and Craig Ewing, “Bayesian state estimation for tracking and guidance using the bootstrap filter”, Journal of Guidance Control and Dynamics 18 (6), pp. 1434-1443, 1995.

[60] James Carpenter, Peter Clifford, and Paul Fearnhead, “An improved particle filter for non-linear problems”, IEE Proceedings Radar Sonar and Navigation 146 (1), pp. 2-7, 1999.

[61] Michael Isard, and Andrew Blake, “CONDENSATION—Conditional Density Propagation for Visual Tracking”, International Journal of Computer Vision 29 (1), pp. 5-28, 1998.

[62] 矢野 浩一, “粒子フィルタの基礎と応用：フィルタ・平滑化・パラメータ推定”, 日本統計学会誌 44 (1), pp. 189-216, 2014.

[63] 倉爪 亮, 石川 博, 加藤 丈和, 佐藤 淳, 三田 雄志, 八木 康史, 斎藤 英雄, “コンピュータビジョン最先端ガイド 1”, アドコム・メディア株式会社, pp. 75-95, 2008, 東京.

[64] Michael Isard, and Andrew Blake, “ICondensation: Unifying low-level and high-level tracking in a stochastic framework”, 5th European Conference Computer Vision, pp. 893-908, 1998.

[65] Norimichi Ukita, Toshihiro Kitajima, and Masatsugu Kidode, “Estimating the Positions and Postures of Non-rigid Objects lacking Sufficient Features based on the Stick and Ellipse Model”, Proc. of IEEE Workshop on Articlated and Nonrigid Motion, in conjunction with CVPR2004, 2004.

[66] Paul Viola, and Michael Jones, “Robust Real-Time Face Detection”, International Journal of Computer Vision 57 (2), pp. 137-154, 2004.

[67] 山下 隆義, 藤吉 弘亘, “特定物体認識に有効な特徴量”, 情報処理学会研究報告コンピュータビジョンとイメージメディア (CVIM), pp. 221-236, 2008.

[68] Keinosuke Fukunaga, “Introduction to Statistical Pattern Recognition”, Academic Press, Boston, 1990.

[69] Yizong Cheng, “Mean shift, mode seeking, and clustering”, IEEE Trans. Pattern Anal. Machine Intelligence 17, pp. 790-799, 1995.

[70] Gray Bradski, “Real time face and object tracking as a component of a perceptual user interface”, Applications of Computer Vision, 1998.

[71] Labeled Faces in the Wild database Available: http://vis-www.cs.umass.edu/lfw/ [Accessed: Jan. 10, 2018].

研究業績

関連論文

1. 北島 利浩, 浮田 宗伯, 木戸出 正継, ”スティックと楕円の統合モデルによる外見的特徴の少ない非剛体群の時系列位置・姿勢推定法”, 電子情報通信学会論文誌, Vol. J88-DII, No. 2, pp. 334-346, 2005

2. Toshihiro Kitajima, SangOn Choi, and Edwardo Arata Y.Murakami, ”Non-contact Heart Rate Measurement based on Camera Image”, Journal of Information Assurance and Security, Vol. 10, Issue 3, pp. 120-129, 2015

3. 北島 利浩, 村上 エドワード, 吉元 俊輔, 黒田 嘉宏, 大城 理, ”生体信号を利用したプライバシー保護画像からの顔検出手法”, 電気学会論文誌 C, Vol. 138, No. 3, 2018

国際発表

1. Norimichi Ukita, Toshihiro Kitajima, and Masatsugu Kidode, ”Estimating the Positions and Postures of Non-rigid Objects lacking Sufficient Features based on the Stick and Ellipse Model”, Proc. of IEEE Workshop on Articlated and Nonrigid Motion, in conjunction with CVPR2004, Jun., 2004

2. Toshihiro Kitajima, SangOn Choi, and Edwardo Arata Y.Murakami, ”Heart Rate Estimation based on Camera Image”, Proceedings of 14th International Conference on Intelligent Systems Design and Applications (ISDA 2014), pp. 50-55, Nov., 2014

3. Toshihiro Kitajima, SangOn Choi, Edwardo Arata Y.Murakami, Shunsuke Yoshimoto, and Osamu Oshiro, ”Contactless Measurement of Heart Rate

Using Single Channel Camera Images", 12th International Conference on Ubiquitous Healthcare (u-Healthcare 2015), Nov., 2015

4. Toshihiro Kitajima, Edwardo Arata Y.Murakami, Shunsuke Yoshimoto, Yoshihiro Kuroda, and Osamu Oshiro, "Human Detection Using Biological Signals in Camera Images with Privacy Aware", 16th International Conference on Intelligent Systems Design and Applications (ISDA 2016), Advances in Intelligent Systems and Computing, Vol. 557, pp. 175-186, Dec., 2016

国内発表

1. 北島 利浩, 浮田 宗伯, 木戸出 正継, "スティックと楕円の統合モデルによる外見的特徴の少ない非剛体群追跡法", 情報処理学会研究報告, コンピュータビジョンとイメージメディア研究会, No.2003-CVIM-143, Vol. 2003, No. 26, pp. 57-64, Mar., 2004

特許

1. 北島 利浩, 村上 エドワード, 崔 相彦, 心拍数測定装置及び心拍数測定方法, 特開 2015-100432

2. Toshihiro Kitajima, SangOn Choi, and Edwardo Arata Y.Murakami, Method And Apparatus For Measuring Heart Rate, US 2015/0148687 A1

3. Toshihiro Kitajima, SangOn Choi, and Edwardo Arata Y.Murakami, Method And Apparatus For Measuring Heart Rate, KRA20150059631

4. 北島 利浩, 脈拍検出装置及び脈拍検出方法, 特願 2016-212961