

マルチバンドカメラを用いた顔検出システム

宮原 知也¹ 引間 湧司¹ 姜 有宣^{1,a)}

概要: 近年、生体情報がパスワードとして使われる時代に向けて、様々な生体情報を用いた認識システムの研究が実現されている。特に顔はすべての人間が持ち合わせており、人間にとって非常に身近で取り扱いやすい生体情報である。そこで、本研究ではカラー画像と近赤外画像を同時に取得できるマルチバンドカメラを用いて顔を検出する新しいシステムを提案する。カラー情報だけでなく、近赤外の分光反射特性を利用することで識別性能の向上を図る。本稿ではマルチバンドカメラの構成を紹介するとともに、それを用いた顔検出手法について説明する。評価実験により、マルチバンドカメラを利用した顔検出システムの有効性を示す。

1. はじめに

近年、自動車の安全性や快適性の向上を目的として、様々な運転支援システムの実用化が進められている。多くのシステムにおいて走行レーンや障害物の存在などを認識するために、カメラやレーダなどの車載センサが利用されている。特に、カメラは周辺状況の映像を直接ドライバーに提供したり、画像処理によって特定の認識対象を検出したりすることができるため、様々なシステムで有用されている。しかしながら、システムごとに測定範囲や検出対象が異なるため、カメラの仕様（撮像波長帯、画角、解像度など）を統一することは難しい。今後、ますます多くのシステムが車両に搭載されることが予想され、搭載性やコストの面からもカメラの統合化に対する要求が高まると考えられる。そのため、波長帯の異なるカメラの統合化を目的として、マルチバンドカメラの開発が要求されている [1][2][3]。

マルチバンドカメラは従来のカラーカメラと近赤外カメラを1台に統合したものであり、昼夜を問わず多目的なアプリケーションに利用可能である。そこで、本研究では、その多目的なアプリケーションの一つとしてマルチバンドカメラを用いた顔検出システムを提案する。カラー画像と近赤外画像を同時に取得できるマルチバンドカメラは近赤外光の分光反射特性を利用できるため、マルチバンド画像を利用すると樹木や人肌の識別に有効であると考えられる。図1にマルチバンドカメラを用いた顔画像を示している。マルチバンドカメラは使用する目的によって図1(a)のようにカラー画像、図1(b)のように近赤外線画像、図

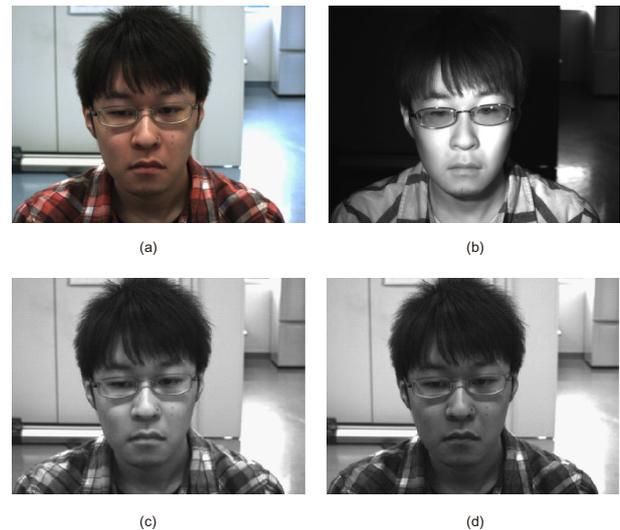


図1 マルチバンドカメラを用いた多様な顔画像 (a) カラー画像 (b) 近赤外線画像 (c) 赤い (R) バンドの画像 (d) 緑 (G) バンドの画像

1(c), (d) のようにグレースケール画像など多様な入力画像の抽出が出来る。

生体情報がパスワードとして使われる時代を向け、音声や指紋、掌紋、網膜など、様々な生体情報を用いられた認識研究が注目されている。人間の生体情報の中に顔は、すべての人が持ち合わせており、非常に身近で取り扱いやすい情報である。特に指紋認証、静脈認証などと違って非接触で認識が行われているため、ユーザーの負担が少なくなり、本人の気づかないうちに認識が可能である。本研究では、マルチバンドカメラを利用し顔を検出する時、カラー情報だけでなく、近赤外の分光反射特性を利用することで識別性能を向上するのが本研究の目的である。

¹ 東京工芸大学
1583 Iiyama, Atsugi, Kanagawa 243-0297, Japan
^{a)} yskang@cs.t-kougei.ac.jp

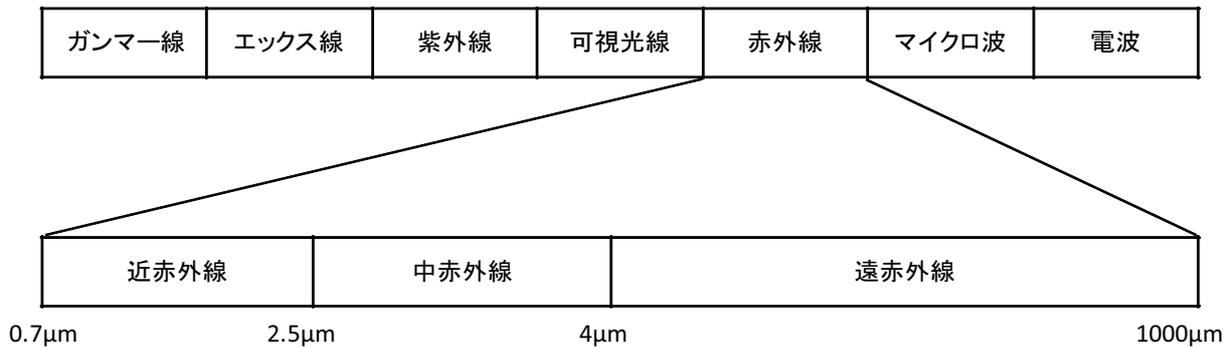


図 2 波長帯による赤外線の種類

本研究でカラーと近赤外両方を使う理由は、カラー画像は近赤外画像と違い人が見てわかりやすい、近赤外線画像はカラー画像に比べ情報が少なく照明変化に強いためである。入力画像から顔の特徴を抽出するため用いる特徴抽出手法は、Haar-Like 特徴 [5] である、それを識別する関数は AdaBoost[9] を使う。本稿にはマルチバンドカメラの構成と Haar-Like 特徴を説明するとともに、それをを用いた顔認識手法について説明する。評価実験では、照明変化により顔検出の結果を求め、マルチバンド画像の有効性を示す。

2. 赤外線の種類

赤外線は赤色光よりも波長が長く、ミリ波長の電波よりも波長の短い電磁波全般を指し、波長ではおよそ $0.7\mu\text{m} \sim 2.5\text{mm}(=2500\mu\text{m})$ に分布する。すなわち、可視光線と電波の間に属する電磁波と言える。赤外線は波長によって、近赤外線、中赤外線、遠赤外線に分けられる。本研究で利用する近赤外線は、およそ $0.7\mu\text{m} \sim 1.1\mu\text{m}$ の可視光(赤)にほど近い電磁波である。この近赤外線は可視光線に近い性質を持ち、直接肉眼では見えないが、デジタルカメラやビデオカメラのモニターなどで見ることができる。

赤外線について詳しく説明すると、1800年にドイツの天文学者ハーシェル (Sir William Herschel) が、プリズムを通り抜けた太陽光のスペクトルがそれぞれどんな熱作用を持っているかを調べるため、寒温計を紫から赤の方へ動かしていき、赤を通り過ぎて少しも光がないところに最も強い熱作用があることを発見し、これを赤外線 (Infrared rays) と名づけた。熱作用が強いため熱線ともいわれる。1853年になってフランスの物理学者アンペール (Andre-Marie Ampere) はこれが可視光線と同じ光波であることを示した。その後、1900年にいたってはドイツの物理学者プランク (Max Karl Ernst Ludwig Planck) が黒体放射のスペクトル分布に関する量子論を確立し、赤外放射に関する理論の基礎が固まった。

人の目で捉えられる光、すなわち可視光線の波長は、短波長側で $360\text{nm} \sim 400\text{nm}$ 、長波長側は $760\text{nm} \sim 830\text{nm}$ と

いわれている。可視光線より波長が短くても長くても人の目には見るができなくなり、可視光線より波長の短いものを紫外線、長いものを赤外線と呼ぶ。可視光線に対し、赤外線と紫外線を指して、不可視光線(ふかしこうせん)と呼ぶこともある。近赤外線とは赤外線の種類の一つである。赤外線の種類は波長帯によって図2のように三つに分けられている。

2.1 近赤外線

近赤外線は波長がおよそ $0.7 \mu\text{m} \sim 2.5 \mu\text{m}$ の電磁波で、赤色の可視光線に近い波長を持つ。性質も可視光線に近い性質を持つため人の目では見るのできない光とも言われている。近赤外線に感光するカメラを用いることで人には見えないものを見ることができるため、夜間に近赤外線光源で照射し被写体に気付かれない撮影が可能のため防犯カメラとして使われている。他にも赤外線通信、家電用のリモコンなどに応用されている。

2.2 中赤外線

中赤外線は、波長がおよそ $2.5 \mu\text{m} \sim 4 \mu\text{m}$ の電磁波で、近赤外線の一部として分類されることもある。赤外線光の分野では、単に赤外と言うとこの領域を指すことが多い。有機化学物質などの分光分析に利用される。

2.3 遠赤外線

遠赤外線は、波長がおよそ $4 \mu\text{m} \sim 1000 \mu\text{m}$ の電磁波で、電波に近い性質も持つ。赤外線は物体からは必ず放射されていて、この現象を黒体放射と呼ぶ。高い温度の物体ほど赤外線を強く放射し、放射のピークの波長は温度に反比例する。主な用途としては暖房(電気ストーブ・パネルヒーター・電気コタツ・電気カーペット)、加熱乾燥(電気オーブン・電子レンジ・電気炊飯器)などがある。遠赤外線が物を温めるのは、遠赤外線の周波数が分子の振動と一致すると遠赤外線のエネルギーが吸収され、分子の振動を増加させることから温度が上がる。

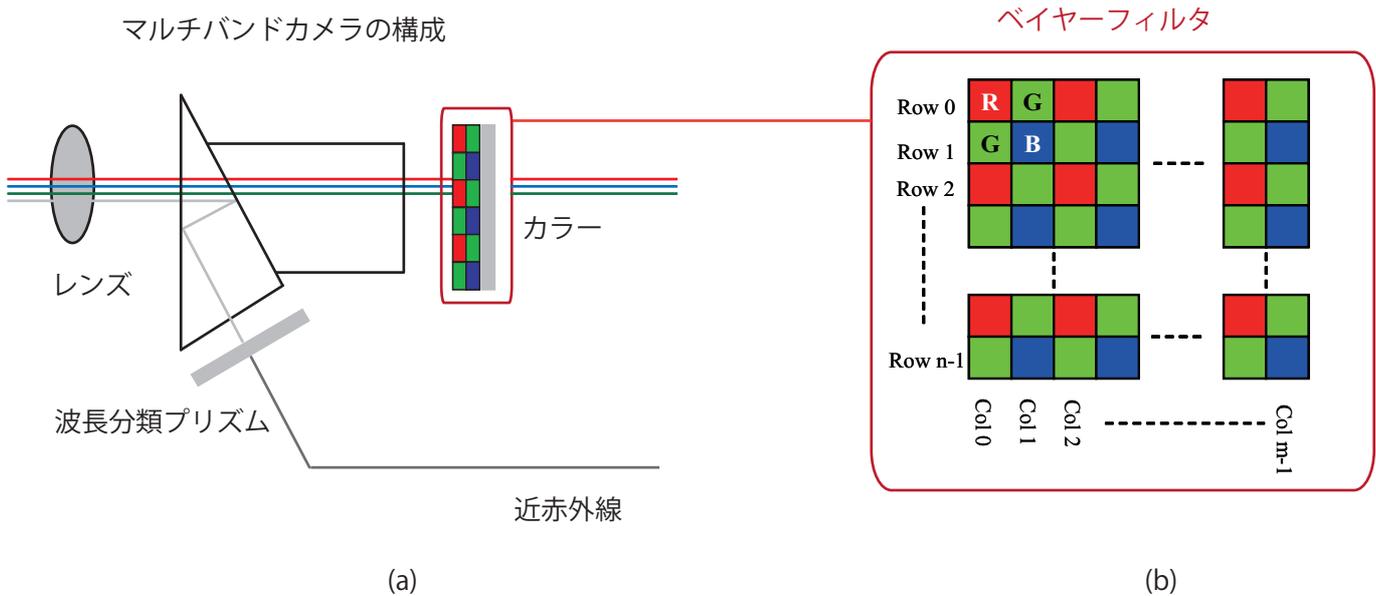


図 3 マルチバンドカメラの構成 (a) マルチバンドカメラ内部の構成 (b) ベイヤーフィルタ

3. マルチバンドカメラ

3.1 マルチバンドカメラの構成

本研究で使うマルチバンドカメラは、4つの波長帯の光を同時に受光するために、カメラの中に特殊なプリズムが備えている。図3にそのマルチバンドカメラの構成を示す。マルチバンドカメラは可視領域から赤外領域までの複数波長の光を同時に分離撮像できるカメラである。図3(a)はマルチバンドカメラの内部構成を、図3(b)は単板式のカラーカメラでよく利用されるベイヤー (Bayer) 配列 [4] と呼ばれるフィルタを示したものである。

図3(a)にマルチバンドカメラの内部の構成を示す。レンズから入ってきた光は赤外光を反射し赤外光と可視光を分離する。透過した可視光はベイヤー配列によりカラー画像として取得する。そして、波長分類プリズムから可視光がカットされ近赤外線の画像が生成出来る。これにより、4波長帯の光を時間差なく取得することができ、カラー画像や近赤画像など、目的に応じて様々な画像を提供することが可能となる。

図3(b)のベイヤー配列に記号, R, G, Bは、それぞれ光の三原色、赤色、緑色、青色の波長帯の光を透過するフィルタを示し、画素1個に1つのフィルタが配置されている。総画素数に対して各色の解像度は低下するものの、簡易な構成でありカメラの小型化に適している。またデモザイキング [4] と呼ばれる適切な補間演算を施せば、総画素数とほぼ同サイズのカラー画像を生成することもできる。デモザイキング処理はベイヤー画像をRGB画像に変換する際に使われる処理のことで、単板の状態から三板の状態

を作り出すことである。ここで緑の画素が他の画素の2倍になっているのは、人間の目の分光感度が緑付近をピークとしており、緑の解像度が見かけ上の解像度を向上させるためである。

3.2 マルチバンド画像の抽出

マルチバンドカメラからカラーと近赤外画像を抽出する方法は図4に示す。本研究で使うマルチバンドカメラと赤外線投光機は図4(a)に示している。4バンドの画像を同時に受けるためカラー画像の情報と近赤外画像の情報は図4(b)のように一つのビデオシーケンスから入力させる。その時、カラー画像は図4(c)のようにビデオシーケンスの緑チャンネルからベイヤー配列の形で入力されている。緑チャンネルから取得したベイヤー画像はデモザイキング処理によりカラー画像の生成が出来る。図4(d)は拡大したベイヤー画像の例とカラー画像の生成のためデモザイキング処理が行われることを示している。一方、入力されたビデオシーケンスから赤チャンネルを抽出すると図4(e)のように近赤外画像が生成される。

4. 顔検出

4.1 特徴抽出

顔検出のため用いられた画像の特徴は2001年にviolaとJonesの2人が考案したHaar-Like特徴 [5] である。本来、顔検出はデジタルカメラなどではリアルタイムに顔の位置を検出してくる必要がある処理である。従って、できるだけ計算量の少ない特徴量を用いながら、かつパターン認識精度の高い識別器により各探索窓ごとの処理を行う必要があると考えられる。この目的を達成するために

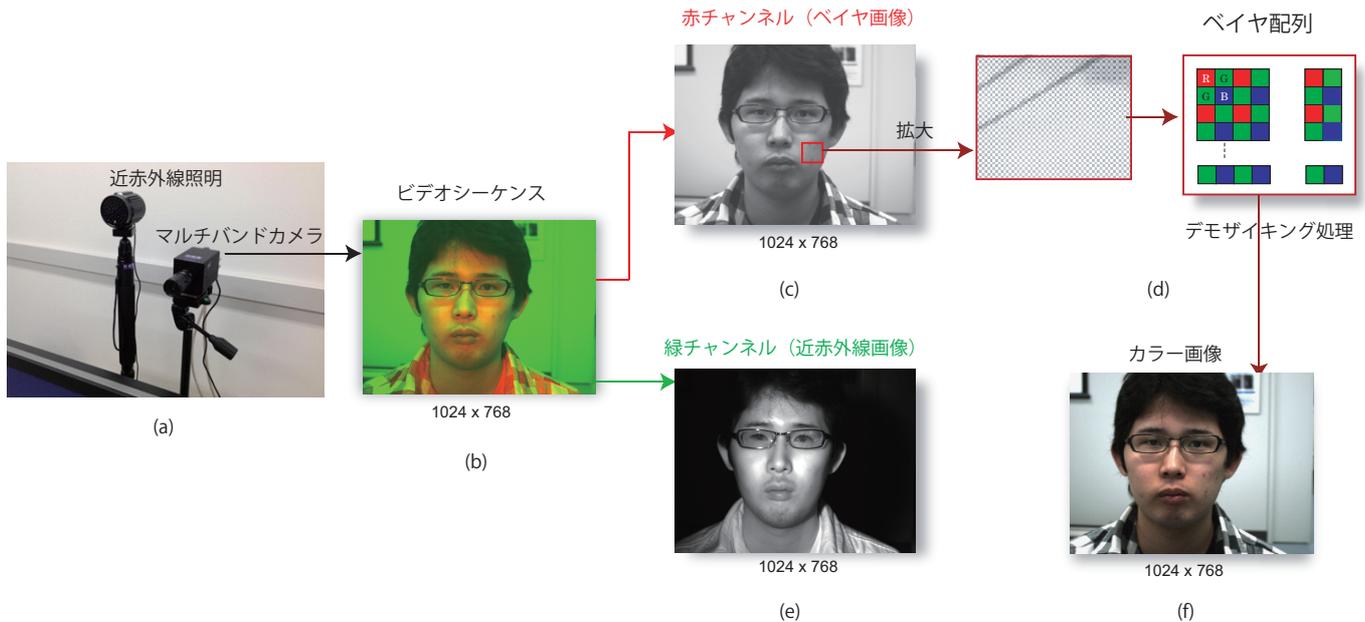


図 4 マルチバンドカメラからマルチバンド画像の抽出 (a) マルチバンドカメラと赤外線投射機 (b) 入力ビデオシーケンス (c) ベイヤ画像 (d) 拡大したベイヤ画像とベイヤ配列によるデモザイキング処理 (e) 生成された近赤外画像 (f) 生成されたカラー画像

Haar-Like 特徴量がシンプルな特徴量なので Boosting の学習により弱識別器を作成し、顔を検出するのが一般的な顔検出アルゴリズムである。

Haar-like 特徴量は、探索窓の中で計算対象である矩形中の黒色の領域のピクセル値の和の値から白色の領域のピクセル値の和の値を引いただけの非常にシンプルな値である。この矩形の配置場所は探索窓のどこにどう配置しても良いことから、探索窓中に数万通りの配置場所の可能性があることになる。オブジェクト検出特徴の例を図 5 の上に示す。矩形特徴のパターンを図 5 の下に示す。入力画像上の任意の場所に、図 5 のような検索窓をおいた場合、検出窓の中の画像の特徴量は、以下の式で計算される。

特徴量 = 矩形特徴の黒領域の画素値の合計 - 矩形特徴の白領域の画素値の合計

Haar-like 特徴量は矩形領域の平均明度の差分値として求められるスカラ量であり、その値は明度勾配の強度を表しているから絶対的な明度値に依存せず、テクスチャに相当する特徴量を抽出することができる利点がある。でも、顔によく似た非顔パターンに対しては誤検出も行っている。一般的に Haar-like 特徴を使った顔検出は、特定の姿勢に限定され、例えば側面顔を検出することは難しくなることがある [6]。

従って、検索窓が 20×20 ピクセルだとすると、Haar-Like 特徴の計算対象の矩形の位置は数万通りの候補があるわけだが、AdaBoosting により検索窓内の各弱識別の重みづけを学習で決定しておくことが可能となる。これにより、数

万通りある Haar-Like 特徴量の弱識別器の内から、重要度が低いものは強識別全体の性能に影響が出ないので使用せず、重要度が上位の数十個の弱識別器のみを選択して強識別器を作ることも可能である [8]。

4.2 識別関数

AdaBoost とは Boosting 学習の一種であり、2 クラス分類の判別器を構築する学習アルゴリズムである。一つ一つの識別関数 (一般に弱識別関数または弱判別器と呼ばれる) が 50% 以上であればよいという決して能力が高い訳ではない識別関数を線形結合し、まるで 1 つの識別関数 (強識別関数または強識別器) とみなすものである。AdaBoost はその Boosting 学習に「重み」をいう概念を加えたものである。学習を進めていく上で識別対して最も適した特徴を識別関数として採用し、最終的にはそれらを線形結合する訳だが、学習毎に前回採用された最も能力の高い特徴で誤っ

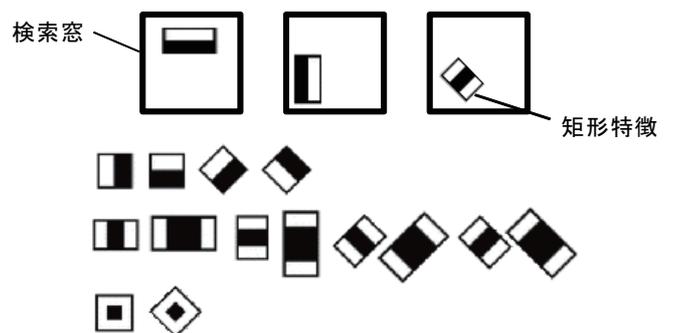


図 5 オブジェクト検出特徴の例と矩形特徴のパターン

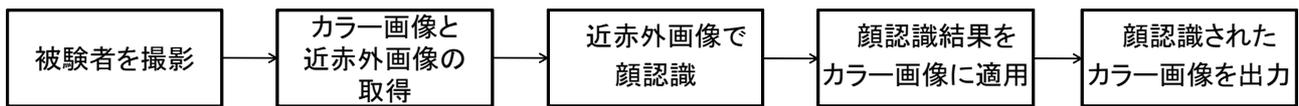


図 6 全体の流れ

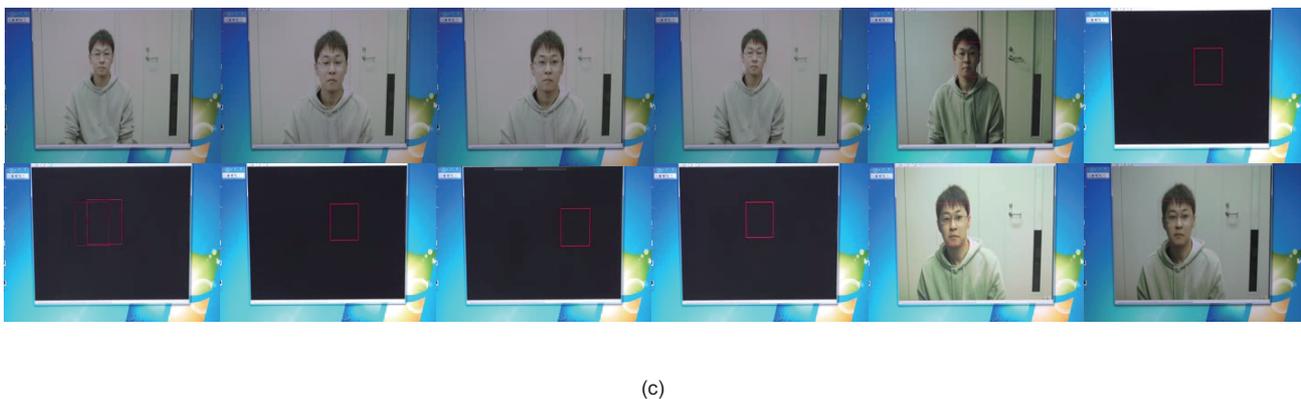


図 7 マルチバンドカメラを用いた顔検出システムの結果 (a) カラー画像でも顔が確認出来る明るさで撮影した場合 (b) カラー画像では顔が確認できない明るさで撮影した場合 (c) 動画で出力された結果

て識別してしまった画像に対して重みを大きくし、次のステップでは要注意すべきサンプルとして注意を促すもので、重みを逐次更新することによって効率良く学習をすすめていくものである。学習の回数は任意になるが、一般的に学習回数と識別関数の個数は同値であり、当然多数の識別関数で構築された強識別関数 (判別器) の方が優秀な結果を出力する傾向がある [10]。本研究ではオープンソースのグラフィックライブラリである OpenCV[7] を用いることで AdaBoost 学習器を実現した。

5. 実験結果

マルチバンドカメラを用いた顔検出システムの実験の流れを図 6 に示す。まず、被験者をマルチバンドカメラで撮影する。そして、マルチバンドカメラからカラー画像と近赤外画像を取得する。取得された近赤外画像から Haar-Like

特徴を抽出し、AdaBoost 学習器で検出を行う。すなわち、近赤外画像で顔を検出事を意味する。後は検出の結果をカラー画像に適用し、出力する流れで実験を行う。

本研究のため備えているマルチバンドカメラシステムは、Jai Inc. の AD-080CL の 2 CCD カメラである。マルチバンドカメラは赤外線投光機が必要である。今回の研究では、OpenCV と Matrox Imaging Library を使用している。OpenCV とは Intel Open Source Computer Vision Library の略で米 Intel 社で開発された画像処理・画像認識用の C 言語ライブラリを言う。それは、オープンソースであり、商用・非商用を問わず無料で使用できる。静止画にも動画にも対応しているため、特にロボット工学やバイオメトリクスの研究によく使われている。推奨プラットフォームは Linux (Intel C Compiler, GNU C) と Microsoft Visual C/C++ である。あと、Matrox Imaging



図 8 様々な照明環境に対応できる顔検出システムの結果

Library (マトロックスイメージングライブラリ) とは、カナダ Matrox 社の Imaging 部門が開発した総合画像処理ライブラリで、産業用として必要な画像処理機能をほぼ全て備えており、医療・薬品・半導体・自動車・印刷など多岐にわたる分野で使用されている。略称は「MIL (ミル)」と言う。

実験はカメラから 1.5m~2.0m 離れた距離で顔画像を撮影し顔検出を行った。照明の変化はカラー画像でも顔が確認出来る明るさで撮影した場合とカラー画像では顔が確認できない明るさで撮影した場合で実験を行った。顔検出の結果は図 7 に示す。カラー画像でも顔が確認できる明るさで撮影した場合の結果を図 7(a) に示す。右の近赤外画像で顔認識をし、結果をカラー画像に適用している。カラー画像で顔が確認できない明るさで撮影した場合の結果を図 7(b) に示す。すべての顔検出の結果は提案したシステムの中に動画像として出力されるので、その動画像の結果は図 7(c) に示している。その結果を見るとカラー画像では顔が確認できないくらい暗いところでも、近赤外画像から顔検出ができるため、顔検出の結果は正確に出ている。提案したシステムはカラー画像にすべての結果を出力し、人の目には見えなくなる照明環境でも顔検出が出来る利点がある。

そして、様々な照明の変化を持ち顔検出を行いその結果を図 8 に示している。図 8 はカラー画像の結果のみ表示している。その結果により提案したマルチバンドカメラを用いた顔検出システムは人の目に顔が見えても見えなくても照明環境に影響を受けないロバストな顔検出システムになる事を確認した。そして、提案したカメラシステムは照明変化が激しい場所のアプリケーションや昼夜を問わず使用する多目的なアプリケーションなどへの幅広い応用が期待されている。

6. まとめ

本研究ではマルチバンドカメラを用いた顔検出システムの構築を行った。この顔検出システムのメリットは、近赤外画像から認識した結果をカラー画像に表示するため、人が見て違和感がないことである。近赤外画像はカラー画像では撮影できない暗闇でも撮影でき、顔を検出することが

出来ることである。マルチバンドカメラを用いた顔検出システムは、カラー画像と近赤外画像を同時に取得することで照明変化に強い性能を持っていることがわかった。今後の課題としてはこの顔認識システムは正面の顔しか認識することができないため横顔などでも認識出来るようにすることが挙げられる。

謝辞 本研究は JSPS 科研費 12861561 の助成を受けたものです。

参考文献

- [1] Kidono K. and Ninomiya Y.: itVisibility Estimation under Night-time Conditions using a Multiband Camera, Proc. of IEEE Intelligent Vehicles Symposium (2007).
- [2] Du H., Tong X., Cao X., and Lin S.: *A prism-based system for multispectral video acquisition*, Proc. IEEE International Conference Computer Vision (2010).
- [3] 姜有宣, 城殿清澄, 木村好克, 内藤貴志, 二宮芳樹: マルチバンドカメラに基づいたイメージセグメンテーション, 第 19 回画像センシングシンポジウム (2008).
- [4] Gunturk B. K., Altunbasak Y., and Mersereau R.: *Color plane interpolation using alternating projections*, IEEE Trans. Image Process, vol. 11, no. 9, pp. 997-1013 (2002).
- [5] Viola P. and Jones M.: *Rapid object detection using a boosted cascade of simple features*, Computer Vision and Pattern Recognition, vol. 1, pp. 511-518 (2001).
- [6] 塚田 真敬: Haar-like 特徴を使ったリアルタイム顔検出
- [7] 入手先 (<http://opencv.jp/>)
- [8] 入手先 (<http://news.mynavi.jp/>)
- [9] 三田 雄志, AdaBoost の基本原理と顔検出への応用: CVIM 研究会 チュートリアルシリーズ
- [10] 武田修平, AdaBoost による顔画像の属性判別, 修士論文 (法政大学)
- [11] Kang Y., Kidono K., Kimura Y., and Ninomiya Y.: *Texture-Based Objects Recognition for Vehicle Environment Perception Using a Multiband Camera*, International Symposium on Visual Computing, LNCS No.4842, Springer-Verlag(II), pp. 582-591 (2009).
- [12] Kang Y., Yamaguchi K., Naito T., and Ninomiya Y.: *Multiband Image Segmentation and Object Recognition for Understanding Road Scenes*, IEEE Trans. on Intelligent Transportation Systems, vol.12, no.4, pp.1423-1433 (2011).
- [13] Kang Y., Kidono K., Naito T., and Ninomiya Y.: *Multiband Image Segmentation and Object Recognition Using Texture Filter Banks*, Proc. of International Conference on Pattern Recognition (2008).