環境変化に頑健な逐次更新型マルチカラーHS ヒストグラム手法を用いた 対人追従視覚機能の開発*

木村 祐太^{*1}, 根元 善太郎^{*2}, 竹村 裕^{*3}, 溝口 博^{*3}

Development of Target Person Detecting and Tracking Method for Mobile Robots Based on Gradual Updating Multi Color HS-Histogram

Yuta KIMURA, Zentaro NEMOTO, Hiroshi TAKEMURA^{*3} and Hiroshi MIZOGUCHI

^{*3} Tokyo University of Science. Dept. of Mechanical Engineering
2641 Yamazaki, Noda, Chiba, 278-8510 Japan

This paper presents that the novel target person detecting and tracking method for mobile robots based on updating multi color HS-histogram. To detect and track a specific person, the image processing with color camera is common method for vision based navigation system of mobile robot. The image processing for the target person detecting and tracking by robot vision has many problems such as the image from the vision device located on robot is often influenced by the change of target size, target moving direction and lighting condition. Moreover, it is difficult to detect a target person in the crowd. To solve these problems, we propose "GRUMIC-HS-hist method (Gradual Updating Multi Color HS-histogram method)". In the proposed method, the colors of the tracking object are extracted by HSV color space. To decrease the influence of the illumination changes, a HS-histogram is proposed. The extracting colors on HS-histogram are clustered by Nearest Neighbor Method. The colors on HS-histogram are weighted by the anisotropic asymmetric Gaussian distribution to update the HS-histogram without spreading. This proposed "GRUMIC-HS-hist method" achieves the robust target person detecting and tracking by a mobile robot. In addition, the quantitative evaluation method is also proposed to evaluate the proposed target person detecting and tracking method. The quantitative evaluation method is composed with "Extraction Rate", "Object Detecting Rate" and "Tracking Error". The quantitatively experimental results show that the proposed "GRUMIC-HS-hist method" is the robust method for many problems.

Key Words : Robot Vision, Detecting, Tracking, Image Processing, Quantitative Evaluation

1. はじめに

近年, 少子高齢化の進行や, 労働人口の減少が社会的課題となっている. また, 「安心」, 「安全」, 「快適」といった日常生活の質的向上に対する関心が高まっている. この様な社会的背景から工場などで普及している産業用ロボットだけでなく, 一般の生活環境に溶け込み, 人と共存・協調し, さらに補助が出来るというようなサービスロボットの研究が様々なところで行われている⁽¹⁾⁻⁽³⁾. 日立の EMIEW2⁽¹⁾は, 機敏な動作や障害物の回避など, 一般生活環境で必要な運動能力を持ち, 離れた場所 (1m 程度) からでもマイクなしで人との対話をして行動するロボットであり, 富士通フロンテックの enon⁽²⁾は案内・誘導, 搬送, 巡回・見回りなどのサービスを提供するロボットである. また竹下ら⁽³⁾は天井カメラ画像を用いて家庭用掃除ロボットの研究を行っている. 人との共存・協調可能なロボットの研究の一環として対人追従ロボットすなわち特定人物を発見し・追従するロボットの研究がある⁽⁴⁾⁻⁽⁷⁾. 人を追跡する方法としては, 赤外線センサーを用いた手法⁽⁸⁾や画像位置取得可能な光 ID タグとフロ

* 原稿受付 2011 年 9 月 22 日

^{*1} 学生員, 東京理科大学 (〒278-8510 千葉県野田市山崎 2641)

^{*2} 東京理科大学

^{*3} 正員, 東京理科大学理工学部

E-mail: takemura@rs.noda.tus.ac.jp

アセンサを用いた手法⁽⁹⁾，レーザーレンジファインダによる距離情報を用いる手法など様々な手法⁽¹⁰⁾がある．しかし，人間とロボットの共存・協調実現を目指す上で，ロボットから「人間が見えている事」・「人間に合わせて動く事」が必要不可欠であり，これらの手法では特定人物一人のみを追跡するという事は難しい．

そこで，対人追従ロボットの追跡技術の中で，カラーステレオカメラを用いた画像処理により追跡するものがある．特定人物に付いていくにはカラーステレオカメラを用いた画像処理により特定人物を認識しなければならない．本研究では人間とロボットの共存・協調実現へ向け，ロボットに人が見えていることの象徴として「対人追従視覚機能」の実現を目指す．(図 1)．対人追従視覚機能はロボット単体でも荷物の自動運搬・ペットロボット等の実用性が考えられるが，対人追従視覚機能を核として音声発声・音声認識によるコミュニケーション機能，アーム制御による把持機能等と組み合わせていくことが出来れば，ロボットは人への有用なサービスを実現することが可能となると考えられる．

対人追従視覚機能は特定の場所・限定的な環境のみではなく，屋内・屋外を含めた日常空間の中で実現されなければ有用性がない．すなわち，屋内・屋外を含めた日常空間において人間を常に追いかける為には，環境変化に対して頑健な対象人物追跡視覚機能が必要不可欠である．カラーステレオカメラを用いた画像処理の場合，環境変化，特に照度変化に大きく影響され，その変化に対応するのが困難になってしまう．また，追跡対象となる人物の大きさや向きの変化や複数人の人物がカメラ画面上にいる場合での人物の特定が困難である．これらの課題を解決するために，本研究では「逐次更新型マルチカラー HS ヒストグラム手法」(GRUMIC-HS-hist method, Gradual Updating Multi Color HS-histogram method)を提案する．

これまで，環境変化の一つである照度変化に頑健な発見・追跡手法として様々な手法が提案されている^{(11)~(13)}．しかし，これらの手法はそれぞれ照度変化に頑健であると評価されているが，定量的な評価を充分には行っておらず，それぞれの手法の有用性を追跡結果画像の呈示等により定性的に主張しているに過ぎない．また，統一された評価基準などがなく再現性に欠けているため，既存の手法との比較を定量的に行っているものもない．映像監視システム⁽¹⁴⁾や交通映像異常事象検出システム⁽¹⁵⁾などにも照度変化に頑健な追跡技術が用いられているが，これらも追跡が成功している画像など定性的な評価は示されているが，定量的な評価が示されていない．そこで，本提案手法の有用性を定量的に示すために「画素抽出率」「対象発見率」「追跡誤差」という定量的評価基準を考案し，定量的評価を行う．また，「追跡誤差」を用いて提案手法と既存手法との比較を行い，定量的に提案手法の優位性を示す．

Fig. 1 Overview of Person Following Robot

本論文の構成は以下のとおりである．

第 2 章では，対人追従視覚機能として用いる画像処理における課題とその課題を解決するための手法を提案する．第 3 章では，課題を解決するために提案した「逐次更新型マルチカラー HS ヒストグラム手法」について述べる．第 4 章では，提案手法の有用性を定量的に評価し，第 5 章では，本論文のまとめについて述べる．

2. 対人追従視覚機能

対人追従実現の為に、まず対象人物を発見しなければならない。画像処理を用いた人発見に関する研究は、形状特徴を用いて顔を検出する手法、人間の全身の形状特徴を用いる手法、背景差分による対象人物の抽出手法等を初めとして、多種多様な方向性から行われている^{(16)~(18)}。しかしながら、移動ロボットにおける対人追従のための画像処理には以下のような課題が存在する。図2はそれぞれの課題を説明するための基準の状態画像とする。

- (1). 対象人物と移動ロボット間の位置関係が変化する(図3)。移動ロボットの走行もしくは対象人物の歩行によってカメラ画像上の対象人物の大きさが変化してしまう。
- (2). 対象人物と移動ロボットの向きが変化する(図4)。顔特徴等の特定の方向からのみ観測できる特徴量を用いたのでは、向きの変化に対応することができない。
- (3). 複数人のヒトが存在している可能性がある(図5)。複数の”ヒト”を見分けられなければ、どちらの人物を追跡すれば良いのか判断できない。
- (4). 実環境においては、照度変化が発生する。照度変化によって、色の情報が変化してしまう(図6)。蛍光灯の有無による明るさの変化・屋外における日当や日陰などの変化によって、対象人物を見失ってしまうのは実用性がない。
- (5). 実時間以内での処理を実現しなければならない。対人発見処理に数十秒・数分の計算コストを掛かってしまうのでは、対象人物を追いかけることはできない。3次元形状マッチング等の現状の計算機においては膨大な時間を要するアルゴリズムは適さない。

Fig.2 Base Image

Fig.3 Varying Target's Position

Fig.4 Varying Target's Direction

Fig.5 Multiple Person in View

Fig.6 Varying Illumination

これらの課題へ対処する為、本研究では伊藤ら⁽¹⁹⁾、西村ら⁽²⁰⁾の照度変化に対して頑健な対象人物発見手法をベースに拡張した対人追従視覚機能を提案する。伊藤らはカラステレオカメラからの距離情報と単一の逐次更新型 HS 色ヒストグラムを用いた色分布マッチングを併用する手法を用いた。図3、図4のように変化する対象を追跡する場合に対象の色分布を特徴量として用いることが有効とされている⁽²¹⁾。本提案手法では、HSV 色空間において、基準ヒストグラムを複数個持ち、複数色の識別が可能な、「逐次更新型マルチカラーHS ヒストグラム手法」を用いることで、より汎用的な環境での対象人物発見を実現する。色ヒストグラムを用い、照明変化に頑健な手法として、MeanShift 法⁽²²⁾⁽²³⁾がある。この手法は、密度分布の勾配を計算し、重心を移動して極大となる位置を求める方法である。本提案手法は、ヒストグラムにより位置を求めるのではなく、H-S 平面上の色ヒストグラムを色抽出のために利用し、抽出した色領域を用いて対象の検出を行う。色ヒストグラムを更新する場合、対象以外の色によって色ヒストグラムが発散するという問題がある。そこで、色ヒストグラムにカーネル関数として非等方非対称ガウス分布による重みづけを用いて、逐次更新する色ヒストグラムの発散を防ぐ。次章において、本提案手法である「逐次更新型マルチカラーHS ヒストグラム手法」について述べる。本提案手法は、基準ヒス

トグラムを複数個持ち、複数色の識別が可能な「マルチカラーHS ヒストグラム」と非等方非対称ガウス分布を用いた「重み付けによるヒストグラムの逐次更新」によって構成される。

3. 逐次更新型マルチカラーHS ヒストグラム手法

3・1 マルチカラーHS ヒストグラム

まず、対象人物の複数方向からの色情報を登録する。これは、通常私達が身につける服の着色は、前後・左右に異なる模様をしている事が多く、単一方向からの服装の登録では、対象人物の向きが変化してしまった際に対応することが出来なくなってしまうからである。図7は正面からの対象人物の服着色によるテンプレートヒストグラムの生成を表す。このテンプレートヒストグラムでは縦軸に HSV 色空間における H 成分、横軸に HSV 色空間における S 成分、高さ方向に対応する H 値と S 値を持つピクセルの総数を最大数で割った数値を表している。図8は同様に後背面からの対象人物の服着色によるテンプレートヒストグラムの生成を表す。この対象人物の服着色は正面からは、青・紫と二種類の色をしている為二つのヒストグラムの山があるが、後背面からでは紫の単色である為一つのヒストグラムの山がある。正面画像と後背面画像とで対象人物の服着色に差異がある事が示されている。

Fig.7 Template Histogram (two color)

Fig.8 Template Histogram (single color)

2つのテンプレートヒストグラムを統合すると、図9の基準となるマルチカラーHS ヒストグラムが得られる。この基準となるヒストグラムは複数のテンプレートヒストグラムを加算し、最大値で割ることで正規化されている。マルチカラーHS ヒストグラムもテンプレートヒストグラムと同様に縦軸に HSV 色空間における H 成分、横軸に HSV 色空間における S 成分、高さ方向に対応する H 値と S 値を持つピクセルの総数を最大数で割った数値を表している。左のグラフでは、ピクセル数を高さ方向の高さで現しており、右のグラフでは、ピクセル数を輝度で表している。

Fig.9 Basic Histogram

得られた基準となるマルチカラーHS ヒストグラムに対し、NN(Nearest Neighbor)法を用いたクラスタリングを行う。NN 法は最も近い距離にあるクラスタにパターンを分類していく手法で、距離が閾値以上であると新規のクラスタとしてパターンを分類していく。クラスタリングを行った結果が図 10 である。前述のように対象人物は単色の服を着ているとは限らない為、対人発見処理のためには様々な色分布・模様の服装飾へ対応する必要がある。クラスタリングにより、色毎に分類する事ができる。今回の対象人物の服着色は、青色・紫色からなるので、2 個のクラスタに分類されている。NN 法によるクラスタリングを行った結果において、中央部の赤いクラスタが青色へ対応し、左上部の青いクラスタが紫色へと対応している。このように、クラスタリングにより複数色が分類されている。

Fig.10 Clustered Histogram by NN Method

3・2 重み付けによるヒストグラムの逐次更新

クラスタリングにより分類された代表色毎に非等方非対称ガウス分布に基づく重み付けを行う。非等方非対称ガウス分布は、最大値 1 をピークとする単峰性の 2 次元分布であり、正規分布に基づいた分布であるが方向に依存して分散が変化する特徴を持つ。非等方非対称ガウス分布により重み付けされたヒストグラムを図 11 に示す。非等方非対称ガウス分布で重み付けをする主な利点は、更新によるヒストグラムの発散を防ぐためと、H 値より S 値の変化が大きいという特徴を利用するためである。ヒストグラムの更新を行うと対象以外の色を含むようになり発散し、ノイズに弱くなる性質がある。そこで、更新に重みを付けることにより発散を防ぐ。さらに、HSV 色空間を用いた場合、理想的には照明変化には影響しない H, S 値が、実際にはハードウェアの制約のため変化する。その上、H 値より S 値の変化が大きいという特徴を利用するために、H 値より S 値の変化が大きくなるように非等方非対称ガウス分布により重み付けを実施する。これにより、HS ヒストグラムが対象外の色やノイズを含むことなく、安定して更新することが出来、照明変化に対応することが出来るようになる。非等方非対称ガウス分布[7]を式(1)に示す。各クラスタのクラスタ中心を基準として放射線状に中心から分布距離 $f(\theta)$ を求める。さらに、クラスタ中心からの各要素の距離 r を求め、重み W を求める。

$$W(h,s) \equiv W(r,\theta) = e^{-\frac{r^2}{2f(\theta)}} \quad (1)$$

Fig.11 Weighted Histogram by Anisotropic Asymmetric Gaussian Distribution

次に、ヒストグラムの更新について述べる．照度変化に頑健な色抽出を行うためには，照度変化に応じてヒストグラムを更新しなければならない．そこで対象人物領域の色情報を元にヒストグラムの更新を行う．前項にて得られた基準となる非等方非対称ガウス分布を W ，前フレームにおいて色抽出に用いた HS ヒストグラムを $HS(t-1)$ ，対象人物発見領域の HS ヒストグラムを $HS_{new}(t)$ ，時系列を t とすると，色抽出に用いる更新された $HS(t)$ ヒストグラムは式(2)のように表される．

$$HS(t) = \frac{HS(t-1) + W \times HS_{new}(t)}{\max(HS(t-1) + W \times HS_{new}(t))} \quad (2)$$

以下に非等方非対称ガウス分布を用いた重み付けによるヒストグラムの逐次更新のプロセスを示す．①～④の各プロセスにおける結果画像を図 12 に示す．

- ①. カラーステレオカメラから得られた視差画像もしくは 3 次元座標の情報を用いてカメラから対象までの距離を求め，閾値以上離れた領域を背景と見なし分離する．マッチングを画像全体に対して行うと計算コストが大きくなってしまいが，領域を限定することにより計算コストを削減する事が可能となる．
- ②. 各ピクセルの H と S の値が逐次更新型ヒストグラム上で閾値以上のもののみを抽出する．逐次更新型マルチカラー HS ヒストグラムは照度変化に合わせて，変動を吸収するように変化する為，このヒストグラムを用いて対応する最適色のみを取り出すことができる．
- ③. 抽出されたピクセルからプロファイル処理を行いて，領域を求める．領域内の距離情報の平均値から，対象の位置を特定する．プロファイル処理とは画像上の x 軸方向及び y 軸方向に抽出したピクセル数を加算し，ピクセル数が最大の位置から閾値以上（画像上の x 軸方向及び y 軸方向に，それぞれ 30[pixel]・50[pixel]以上）の位置までの領域を探索することである．
- ④. 求められた対象人物領域から，マルチカラー HS ヒストグラムの更新を行う．ヒストグラムを更新することで，照度の変化を反映したマルチカラー HS ヒストグラムへと更新する．

Fig.12 Target Person Area Determination

本提案手法では、追跡対象の色情報テンプレートを取得し、このテンプレートを基に基準のテンプレートヒストグラムを作成し、発見・追跡を行う。よって、追跡対象の色情報のみを含んだテンプレートが大きいほど正確なテンプレートヒストグラムを作成することが出来る。図 13 はカラーステレオカメラから追跡対象までの距離が 3.66[m]のときのカメラ画像である。このとき追跡対象人物の服の色情報のみを含んだテンプレートは最大で約 30[pixel]×約 50[pixel]で取得することが出来る。追跡対象人物の服装によってテンプレートの取得サイズが変わってしまうが、図 14 では肌面積が多い服装の場合であるため、おおよその服装でも 30[pixel]×50[pixel]以上の色情報のテンプレートを取得することが出来ると思われる。また、カラーステレオカメラから追跡対象までの距離が 3.66[m]以上離れてしまった場合は 30[pixel]×50[pixel]という大きさのテンプレートを満たすことが出来ない可能性があるが、本研究において、カラーステレオカメラから追跡対象までの距離の上限は 3.66[m]とするため 30[pixel]×50[pixel]以上の色情報のテンプレートを取得することが出来る。以上のことから、ヒストグラムに基づいて抽出した色情報に対してプロファイル処理を行う際、ピクセル数が画像上の x 軸方向及び y 軸方向にそれぞれ 30[pixel]・50[pixel]抽出されている場合をプロファイル処理の閾値とした。

カラーステレオカメラから追跡対象までの距離の上限は 3.66[m]については文化人類学者エドワード・ホールによる対人距離⁽²⁴⁾を参考にする。対人距離とは、人物間の関係性によってコミュニケーションを図る際の距離が異なることから、密接距離、個体距離、社会距離、公衆距離の 4 つに分類される。「密接距離」(0.15～0.46[m])では、愛撫や格闘など接触的なコミュニケーションを意味する。「個体距離」(0.46～1.22[m])では、1 対 1 の会話など個人的なコミュニケーションを意味する。「社会距離」(1.22～3.66[m])では、業務上の対話など形式的なコミュニケーションを意味する。「公衆距離」(3.66～7.62[m])では、演説など 1 対多数のコミュニケーションを意味する。対人追従を行う場合、ロボットと追跡対象人物との関係性は親しい間柄で 1 対 1 のコミュニケーションに相当する。また、特定人物を認識する必要があるため、互いを個々に認識する距離の限界をロボット(カラーステレオカメラ)と追跡対象人物との距離の上限としなければならない。これらのことを考慮すると、対人距離の分類において「個体距離」、「社会距離」を用いるのが適切である。ゆえに、ロボット(カラーステレオカメラ)と追跡対象人物との距離の下限、上限をそれぞれ、0.46[m]、3.66[m]とする。

Fig.13 Definition of Profile Processing Threshold

3・3 画像処理結果

図 13 に対人追従の結果を示す。この図は図上部に基準となる HS ヒストグラムと各画像処理結果によって構成されている。(a)は提案手法によって対象人物を発見した画像処理結果を表す。入力画像に前フレームの対象人物領域を緑矩形として表している。(b)更新 HS ヒストグラムに対応する色を抽出した結果である。(c)は対象人物領域から作成した HS ヒストグラムである。(d)は基準となる重み付け HS ヒストグラムと新たに作成した HS ヒストグラムから更新された HS ヒストグラムである。各画像処理結果の下には経過時間[sec]が記されている。各画像処理結果について述べる。緑矩形は取り出した色からプロファイル処理によって、対象人物の矩形領域を求めた結果を表す。各画像処理結果の(a)画像処理結果を時系列順に追っていくと、対象人物の移動による大きさの変化・対象人物の向きの変化・照度変化に対してロバストに対象人物をトラッキングできていることがわかる。また、各画像処理結果の(d)新たに更新された HS ヒストグラムを時系列順に追っていくと、照度にあわせて HS ヒストグラムが変移している。これは、各画像処理結果の(b)更新 HS ヒストグラムに対応する色を抽出した結果 HS ヒストグラムと合わせて見ていくと、常に色抽出にて最適な色が抽出されるように、HS ヒストグラムを変化させていった結果であることが示されている。

Fig.14 Results of Target Person Detection Image Processing

4. 定量的評価

4・1 照度変化に関する定量的評価

照度変化状況における定量的評価手法を作成するために、まず照度変化状況を作成する。図 15 は照度変化状況を作成するための実験環境を示す。発見・追跡を行う対象とカメラの位置は動かさず、照明機材によって対象付近の照度を変え、27 種類の照度環境を撮影する（図 16, 17）。図 16, 17 は同じ実験環境下で、対象を単色のものと複数色のものとで撮影を行った各照度における画像である。撮影された 27 枚の照度の異なる画像を画像処理プログラミングに順次読み込ませることにより、連続的に照度が変化する照度変化状況を再現した。

Fig.15 Experiment environment

Fig.16 Illumination Images (Single color target)

Fig.17 Illumination Images (Multi color target)

4・1・1 画素抽出率

本提案手法である対人追従視覚機能では、色情報に基づいて対象の発見・追跡を行っている。そのため色情報を抽出出来ているかということが対象の発見・追跡に大きく関わってくる。そこで色情報を抽出してくる際の「画素抽出率」を考える。あらかじめ撮影しておいた 27 枚の照度画像の中で一般環境における平均的照度である 600[lux]程度の照度画像を基準とする。基準となる照度画像において、対象の色テンプレートを作成し、それに基づいてヒストグラムを作成する。順次読み込ませている照度画像の中で、作成したテンプレートと同一座標内において作成したヒストグラムに一致する画素数つまり色を抽出出来ている画素数を取得する。画素数を抽出する場合、1 枚の照度画像に対して 1 回のみ抽出処理を行うと、ヒストグラムが安定せず画素数を抽出しきれない。ゆえに何回かの抽出処理によってヒストグラムを更新し、追跡対象を安定して抽出してから次の照度画像を読み込むようにする。図 18 は 202.3[lux]の照度画像から 128.8[lux]の照度画像に変わった際、および 128.8[lux]の照度画像中での、20 回の抽出処理の結果を示す。またその際に抽出出来た画素数について表 1 に示す。さらに、表 1 に示した各抽出処理における抽出画素数をグラフに表す (図 19)。

Fig.18 Result of Extracting Target Color (Illumination = 128.8[lux])

Table 1 Number of Pixels while Extracting Target Color (Illumination = 128.8[lux])

Roop	1	2	3	4	5
Extracting Object[pixel] (max number = 5162[pixel])	3112	5198	5233	5224	5216
Roop	6	7	8	9	10
Extracting Object[pixel] (max number = 5162[pixel])	5203	5194	5192	5181	5168
Roop	11	12	13	14	15
Extracting Object[pixel] (max number = 5162[pixel])	5162	5156	5141	5136	5130
Roop	16	17	18	19	20
Extracting Object[pixel] (max number = 5162[pixel])	5124	5124	5124	5124	5124

Fig.19 Result of Number of Pixels while Extracting Target Color (Illumination = 128.8[lux])

図 18, 19 のように, 202.3[lux]から 128.8[lux]に変わった直後 (Roop 1) の追跡対象の抽出画素数は, その後のループ処理のときに抽出出来ている画素数の 60%程度しか抽出できていない. また表 1 からは, 16 回目から, 抽出出来ている画素数が同じ値を取り, 安定していることがわかる. この画素数をその照度画像の抽出画素数とする. 抽出画素数を基準とすると 2~15 ループの間での抽出画素数との誤差は最大で 109[pix]になる. これを抽出画素数 5124[pix]で割り, 誤差率を求めると, 約 2.127[%]の誤差率となっている. 誤差としては最初の 1 ループ目に比べるとはるかに小さいので, 2 回以上ループ処理を行えばより良い抽出が出来るが, 定量的評価を目的としているので, 安定している常態となったときの抽出画素数を用いる. 安定している状態とは, 同じ抽出画素数が 3 回以上連続で取得出来ている状態とする. ここでは 1 枚の照度画像 (128.8[lux]) についてのみ考えているが, 全 27 枚の照度画像がある. しかし, これらの画像においても 3 回以上連続で同じ抽出画素数を取得出来ている状態を安定とすれば 20 回目には全て安定した状態になっている. よって, 抽出処理は 20 回行い, 20 回目の抽出画素数をその照度画像における抽出画素数とする. その抽出画素数を色テンプレートの画素数 (全 2496[pixel]) で割ったものを画像抽出率とした.

評価対象として, 色抽出の際, 表色系に RGB, HSV, YCrCb を用いた場合の 3 種類, さらに, 各表色系それぞれの場合においてヒストグラムの非更新・更新についても考えるため, 計 6 種類を考える. YCrCb は輝度信号 Y と 2 つの色差信号 Cr・Cb からなるため, HSV のときと同じように照度変化に影響される輝度 Y 以外の Cr・Cb で CrCb ヒストグラムを作り実験を行う. RGB については HSV・YCrCb のようなヒストグラムを作成していないため, 色テンプレートの RGB 値をそれぞれ平均し, それを基準として, テンプレートと同一座標内での RGB 値の差が ± 20 未満のものを抽出画素数とした. RGB における更新は RGB 値の平均を更新することとする. 図 20, 21, 22, 23 に評価の結果を示す.

Fig.20 Extraction Rate under illumination changes (Single color target, without updating histogram)

Fig.21 Extraction Rate under illumination changes (Single color target, with updating histogram)

Fig.22 Extraction Rate under illumination changes (Multi color target, without updating histogram)

Fig.23 Extraction Rate under illumination changes (Multi color target, with updating histogram)

図 20, 21 では単色の対象についてヒストグラムの非更新・更新の場合における結果を, 図 22, 23 では複数色の対象についてヒストグラムの非更新・更新の場合における結果を示している. 3 つの表色系の中で HSV 表色系を用いている場合が一番画素抽出率の高い値を維持していることを示している. しかし, 図 20, 22 に示すようにヒストグラムを更新していない場合, 照度の低い範囲に関しては画素抽出率が急激に落ちるなど不安定な状態になっている. 図 20 と 21, 図 22 と図 23 のようにヒストグラムの非更新・更新で比べてみると, HSV 表色系に関しては照度が低いときの画素抽出率も高い値に達し, 照度に変化しても安定した画素抽出率を得ることが出来ている. これらの結果より, HSV 表色系を用い, ヒストグラムを更新する本提案手法が照度変化に頑健であることを定量的に示すことが出来た. また, 複数色を対象としている図 22, 23 において RGB 表色系による色抽出が全く出来なかった. これは色情報を抽出する際に, 基準のテンプレートの RGB 値を平均して抽出しているため, 複数色になった場合に色情報が対象と全く異なるものになってしまうことが原因である. このことから, 複数色に対応するためにクラスタリングを行った提案手法が様々な条件に対応出来る手法であると言える.

4・1・2 対象発見率

本提案手法では, 色情報に基づき抽出されたピクセルからプロファイル処理を行い, 対象領域を求める. この対象領域の発見の繰り返しによって追跡を行っている. 先の画素抽出率はテンプレートと同一座標内での評価だったが, 画像全体から対象を発見出来ているかどうかを評価していく. そこで「対象発見率」を考える. 実験状況については画素抽出率のときと同様である. 27 種類の照度画像について 20 回抽出処理を行うため, 計 540 フレーム分の処理を行っている. この 540 フレーム分の抽出処理の結果, プロファイル処理の閾値を満たし, 対象を発見出来ているとされたフレーム数の割合を対象発見率とした. 評価対象としては, 先の画素抽出率での評価において RGB 表色系では複数色に全く対応出来ていなかったため, RGB 表色系を用いる 2 種類を除く 4 種類について考える. 表 2, 3 に評価の結果を示す.

Table 2 Object Detecting Rate (single color target)

Color space (without updating)	HSV	YCrCb	Color space (with updating)	HSV [Proposed method]	YCrCb
The number of detecting (max number = 540[frame])	420	260	The number of detecting (max number = 540)	540	440
Object Detecting Rate[%]	77.8	48.2	Object Detecting Rate[%]	100	81.5

Table 3 Object Detecting Rate (multi color target)

Color space (without updating)	HSV	YCrCb	Color space (with updating)	HSV [Proposed method]	YCrCb
The number of detecting (max number = 540[frame])	320	0	The number of detecting (max number = 540)	520	0
Object Detecting Rate[%]	59.3	0	Object Detecting Rate[%]	96.3	0

表 2 は単色の対象についてヒストグラムの非更新・更新の場合における結果, 表 3 は複数色の対象についてヒストグラムの非更新・更新の場合における結果を示している. 画像中において, プロファイル処理により加算されたピクセル数が最大の位置から閾値以上 (画像上の x 軸方向及び y 軸方向に, それぞれ 30[pixel]・50[pixel]以上) の位置までの領域を対象と認識する.

ここで画素抽出率における結果についても考えてみる. 画像全体より範囲の小さいテンプレートと同一座標内 (全 2496[pixel]) という範囲において, 30×50[pixel], つまり 1500[pixel]以上抽出出来ている場合においては必ず対象を発見出来ると言える. 1500[pixel]は 2496[pixel]のおよそ 60%である. 図 21 (単色の対象に対して, ヒストグラムを更新する場合) において HSV 表色系を用いた場合, すべて 60%以上の画素抽出率を誇っている. よって, 表 2 においても HSV 表色系を用い, ヒストグラムを更新する場合の対象発見率は 100%となっている. 対象発見率は画像全体で考えるので, 表 2 の YCrCb の結果は, 図 20, 21 において YCrCb 表色系の 60%以上の画素抽出

率を誇っている割合よりも高い値になっている。これらのことから、色情報を抽出してくることが対象の発見・追跡に大きな影響を及ぼしていることがわかった。また、表 2, 3 からヒストグラムを更新すると対象発見率が上がっていることが見て取れる。さらに、表 3 において、YCrCb 表色系を用いた場合、対象発見率が 0% となっている。このことから、複数色のヒストグラムを作成する場合、YCrCb 表色系では対応出来ていないということが出来る。

以上のように「画素抽出率」「対象発見率」の照度変化に関する定量的評価結果から、提案手法は照度変化に頑健な手法であるということが示された。

4・2 追跡に関する定量的評価

前項までに色抽出を評価する画素抽出率と、対象の発見を評価する対象発見率について提案した。本研究においては、対象を発見出来るということは対象を追跡出来ているということなので対象発見率と追跡出来ているかどうかを表す追跡率はほぼ同義と考えられる。そこで、追跡の精確さを評価する。追跡に関する定量的評価を行うために、まず対象が移動している動画を撮影する。評価を行う動画は、単色の対象を人が持って移動することによって対象が移動している様子を撮影したものと、人の顔を対象として撮影したものの 2 種類の動画を用いる。動画は両方とも 250 フレーム分撮影する。その各 250 フレームにおいて対象の中心座標を求めておく。

4・2・1 追跡誤差

追跡の正確さを評価するために「追跡誤差」について考える。あらかじめ撮影しておいた 2 種類の動画において追跡を行わせ、画像処理により算出される追跡対象領域の中心座標と各フレームにおいて求めておいた中心座標とを比較したときの誤差を追跡誤差とする。ここで追跡を行う手法は提案手法と、同様に色情報を基に追跡を行う MeanShift 法とを用いて、比較を行う。図 24 に 2 種類の動画に対して追跡を行った際の各手法における追跡の画像処理結果を示す。また図 25, 26 はその際の追跡誤差の結果を示す。

Fig.24 Tracking result by each method

Fig.25 Tracking Error (single color target)

Fig.26 Tracking Error (human face)

図 25 は単色の対象について追跡を行わせた際の追跡誤差の結果について示している. 対象の大きさが画像上で約 106×135 [pix]であり対角線の長さが約 172 [pix]なので, その半分である 86 [pix]以上中心から離れてしまうと追跡出来なくなってしまう. MeanShift 法による追跡では誤差が $60 \sim 70$ [pix]のときがあり, 極めて対象を見失ってしまう可能性が高いと言える. それに提案手法の追跡誤差は 30 [pix]未満で MeanShift 法の約半分程度の誤差であり, 誤差の値も極端にばらつきがあるわけではなく, 安定した追跡手法であると言える. 図 26 は顔を対象として追跡を行わせた際の追跡誤差の結果について示している. このとき両手法とも対象の中心付近を追跡出来ていて対象を追跡出来なくなる可能性が低いと考えられる.

ここで, 追跡誤差の平均値を考えてみると, 図 25 の場合は, 提案手法・MeanShift 法それぞれの平均追跡誤差は 6.7 [pixel], 39.1 [pixel], 図 26 の場合はそれぞれ 8.85 [pixel], 14 [pixel]となっている. このことから, 提案手法が MeanShift 法に比べて最大で 5 分の 1 に平均追跡誤差を抑えていることがわかる.

以上のように「追跡誤差」の追跡に関する定量的評価結果から提案手法と MeanShift 法とを比較し, 提案手法の方が MeanShift 法より正確に対象を追跡出来る手法であるという優位性を示した.

5. おわりに

本研究では, ロボットのための環境変化に頑健な対人追従視覚機能の開発を行った. 画像処理を用いた視覚機能の実現のためには 5 つの課題があった. まず, 「対象の大きさが変化する問題」と「対象の向きが変化する問題」. これらの問題は, 色情報を基に対象を発見・追跡することによって解決出来た. 次に, 「複数人のヒトが存在する問題」. この問題は, 視差画像もしくは 3 次元座標からの情報を用いてカメラから対象までの距離を求め, 閾値以上離れた領域を背景と見なし分離することにより, 複数人のカメラからの距離がバラバラの場合, 個人を特定することが出来る. また, 複数人のカメラからの距離が一緒でも, 「逐次更新型マルチカラー HS ヒストグラム手法」における, クラスタリングによる複数色に対応した対象の色情報の取得によって個人を特定することが出来る. さらに, 距離情報を用いて背景領域を差分することによって「計算コストの問題」も解決し, 実時間での追跡が可能となっている. 最後に, 「照度変化の問題」に関しては, 「逐次更新型マルチカラー HS ヒストグラム手法」において, HS ヒストグラムを用い, 非等方非対称ガウス分布に基づく重み付けをし, ヒストグラムを更新することによって照度変化に頑健な対人追従視覚機能とした.

さらに, 提案手法の有用性を示すために定量的評価基準を考案した. 「画素抽出率」や「対象発見率」を用いて, 提案手法が照度変化に頑健であると定量的に評価した. 「対象発見率」については, 照度変化する状況下でも提案手法が単色の追跡対象を 100%, 複数色の追跡対象を 96.3%発見・追跡出来ていることを示しており, 複数色の追跡対象を全く追跡することが出来なかった RGB・YCrCb 表色系を用いた手法に比べ, 提案手法の有用性が高いことがわかる. また, 「追跡誤差」を用いて提案手法と既存手法とを定量的に評価・比較し, 提案手法が既存手法に比べて最大で 5 分の 1 に平均追跡誤差を抑えていることがわかった. このことから提案手法が既存手法より正確に対象を追跡出来る手法であるという優位性を示した.

ゆえに, 本研究において提案した対人追従視覚機能は環境変化に頑健な有用な手法であると言える.

本研究において, 画像処理を用いた視覚機能の実現に向けた課題として 5 つの課題を挙げ, それらの課題を解決する手法を提案した. しかし, 視覚機能に影響を及ぼす特別な状況が他にも存在する. 例えば, オクルージョンの問題が挙げられる. また, 追跡対象の服の色が単色の黒色や白色であったりする場合はヒストグラムを形成する過程やプロファイル処理を行う過程で困難が生じてしまう可能性が高い. そこで, このような特別な状況に対しても頑健性を示すことが出来るように, 今後, 本提案手法の機能拡張を目指していく.

謝辞

本研究はNEDO「次世代ロボット知能化技術開発プロジェクト」との成果の一部である.

文 献

- (1) 戸上 真人, 住吉 貴志, 川口 洋平, 小窪 浩明, 天野 明雄, 大淵 康成, “人間共生型ロボット EMIEW2 における音源方向推定機能”, 日本ロボット学会誌, Vol. 28, No. 1 (2010), pp. 35-36.

- (2) 神田 真司, 村瀬 有一, 沢崎 直之, 麻田 務, “サービスロボット「enon」の開発”, 日本ロボット学会学会誌, Vol. 24, No. 03 (2006), pp.12-15.
- (3) 竹下勉, 大矢晃久, 油田信一, “家庭用掃除ロボットの研究—天井カメラ画像を用いた動作指示”, 日本機械学会ロボティクス・メカトロニクス講演会 2004, (2004), pp. 2P1-L2-7.
- (4) 佐竹 純二, 三浦 純, “移動ロボット制御のための人物シルエットの重なりを考慮した複数人物追跡”, 情報処理学会研究報告, 2010-CVIM-172(3) (2010), pp.1-8.
- (5) D. Schulz, W. Burgard, D. Fox and A.B. Cremers, “People Tracking with a Mobile Robot Using Sample-based Joint Probabilistic Data Association Filters”, *International Journal of Robotics Research*, Vol.22, No.2 (2003), pp.99-116.
- (6) H. Kwon, Y. Yoon, J. B. Park and A. C. Kak, “Person Tracking with a Mobile Robot using Two Uncalibrated Independently Moving Cameras”, *IEEE International Conference on Robotics and Automation* (2005), pp. 2888-2894.
- (7) H. Sidenbladh, D. Kragic, H. I. Christensen, “A Person Following Behavior for a Mobile Robot”, *IEEE International Conference on Robotics and Automation*, No.1 (1999), pp. 670-675.
- (8) 本田 誠一, “赤外線センサーネットワークによる人物追跡”, 人工知能学会 全国大会 (第 20 回), (2006).
- (9) 大田 周, 坂本 竜基, 小暮 潔, 藤波 務, “画素位置取得可能な光 ID タグとフロアセンサを統合した複数人物追跡”, *IPSJ SIG Notes. Computer Vision and Image Media*, (2006), pp.309-314.
- (10) 奥迫 伸一, 坂根 茂幸, “レーザーレンジファインダを用いた移動ロボットによる人の追跡”, 日本ロボット学会誌, Vol. 24, No.5 (2006), pp.605-613.
- (11) H. Uemura, T. J. Kooi, H. Kim, S. Ishikawa, “Adaptive Color Extraction Employing an Update Type Two-Dimensional Color Histogram Model”, *IPSJ SIG Notes*, No. 102 (2006), pp.19-24.
- (12) T. Fujita, Y. Mukaigawa, T. Shakunaga, “3D Shape Reconstruction and Analysis of Dance Motion using Multiple Camera System”, *IPSJ SIG Notes. Computer Vision and Image Media*, No.26 (2002), pp.95-102.
- (13) Z. Nemoto, H. Takemura, H. Mizoguchi, “Development of Small-sized Omni-directional Laser Range Scanner and Its Application to 3D Background Difference”, *Proceedings of 2006 IECON* (2007), pp.2284-2289.
- (14) 伊藤 渡, 山田 浩正, 上田 博唯, 千葉 隆広, “映像監視システムにおける差分用背景画像の照度変化追従性向上の検討”, *IPSJ SIG Notes. Computer Vision and Image Media*, No. 79 (1998), pp.1-7.
- (15) 上條 俊介, 原田 将弘, 坂内 正夫, “統計モデルと意味階層の結合による交通映像異常事象検出システム”, 電子情報通信学会誌 A, Vol.88, No. 2 (2005), pp.152-163.
- (16) A. Ohya, T. Munekata, “Intelligent Escort Robot Moving together with Human -Interaction in Accompanying Behavior-,” *Proceedings 2002 FIRA Robot World Congress* (2002), pp.31-35.
- (17) Y. Negishi, J. Miura, and Y. Shirai, “Mobile Robot Navigation in Unknown Environments Using Omnidirectional Stereo and Laser Range Finder,” *Proceedings of 2004 IEEE/RSJ IROS* (2004), pp. 2737- 2742.
- (18) H. Uemura, Tan Joo Kooi, S. Ishikawa, “A Color Tracker Employing a Two-dimensional Color Histogram Under Changeable Illumination”, *Proceedings of 2006 IECON* (2006), pp.3273-3278.
- (19) K. Itoh, T. Kikuchi, H. Takemura, H. Mizoguchi, “Development of a Person Following Mobile Robot in Complicated Background by Using Distance and Color Information,” *Proceedings of 2006 IECON* (2006), pp.3839-3844 .
- (20) S. Nishimura, H. Takemura, H. Mizoguchi, “Development of Attachable Modules for Robotizing Daily Items -Person Following Shopping Cart Robot-,” *Proceedings of 2007 ROBIO* (2007), pp.1506-1511.
- (21) K. Nummiaro , E. Koller-meier , L. V. Gool, “Color Features for Tracking Non-Rigid Objects”, *Chinese Journal of Automation* , Vol. 29 , (2003) , pp. 345-355.
- (22) D. Comaniciu and P. Meer, “Mean Shift Analysis and Applications”, *IEEE International Conference on Computer Vision* (1999), pp.1197-1203.
- (23) D. Comaniciu, V. Ramesh, and P. Meer, “Real-Time Tracking of Non-Rigid Objects using Mean Shift”, *IEEE Computer Vision and Pattern Recognition* (2000), pp.142-149.
- (24) 内田 頼望也, 荒井 亮磨, 木村 祐太, 奥村 亮, 竹村 裕, 溝口 博, “次世代ロボット知能化プロジェクト・リファレンスハードウェアによる対人追従”, 第 11 回 SICE システムインテグレーション部門講演会講演論文集 (2010), pp.1B2-5(1)-(4).