

氏名（本籍）	鈴木 一正（大阪府）
学位の種類	博士（工学）
学位授与番号	甲第77号
学位授与日付	平成27年9月30日
専攻	システム工学専攻
学位論文題目	奥行き情報を利用した顔・髪の検出と追跡に関する研究
学位論文審査委員	（主査）教授 呉 海元 （副査）教授 和田 俊和 教授 入野 俊夫

## 論文内容の要旨

### 1 研究概要

本研究では、奥行き情報を利用した顔や髪領域の検出・追跡システムを提案する。画像中に含まれる人の顔を検出・追跡する技術は、コンピュータビジョンの中でも重要な研究分野であり、監視カメラや、マシンインタフェース、ロボットとの対話など様々な応用がある。髪領域に関する研究報告は少ないが、髪領域は個人を特徴付ける重要な要素であり、個人識別の特徴として用いたり、髪型を利用したアプリケーションなどの応用が考えられる。また、実際に様々な応用システムを開発するために、高速かつ安定なシステムが要求されている。

本研究の目標は、高速かつ安定な顔や髪の検出・追跡システムを構築することであり、奥行き情報を用いた顔検出の高速化と追跡処理の安定化について議論する。提案システムでは、顔検出と追跡手法を組み合わせ、顔検出で得られる情報に基づいて環境に応じた追跡モデルを自動的に構築し、追跡を開始することができる。本論文で提案する手法は以下の3点である。

提案手法1 奥行き情報を用いた顔検出の高速化

提案手法2 顔検出と色弁別度追跡法を組み合わせた顔追跡システム

提案手法3 顔検出と6D K-means Tracker を組み合わせた髪追跡システム

### 2 奥行き情報を用いた顔検出の高速化

一般的に顔検出を行う場合、画像上から位置や大きさが未知の顔を検出するため、位置や大きさを変えて抜き出した数十万のサブウィンドウに対し、識別器を用いた判定を行う必要がある（ピラミッドスキャン）。また、ステレオカメラで顔検出を行った場合、視点や視野の増加により検出精度の向上が期待できるが、単純にピラミッドスキャンを行うと処理コストが倍増してしまう。

本研究では、ステレオ処理によって得られる奥行き

情報を活用し、識別回数を削減することで顔検出の高速化を行う。また、スパースに配置されたサンプル点のみで距離計算を行うことで、ステレオ処理のコストを抑制し、ビデオレートでの顔検出を実現する。

人によって顔の大きさにはそれほど大きな違いがないため、ワールド座標空間内の顔の大きさをあらかじめ決めておくことで、カメラから顔までの距離がわかれば画像上での顔の大きさを推定することができる。画像内の各領域を距離に応じたサイズのみで顔を探査することによって識別回数を大幅に削減することができる。また、2枚の画像を探査することで検出率を向上させながら、識別回数の削減にともなって誤検出を減少することができるため、検出精度を向上させることができる。

このようにして、提案手法では、ステレオカメラを用いることで、検出精度を向上させながら単眼カメラよりも高速なビデオレートでの顔検出が可能となる。また、写真など実際の顔と大きさが異なるような顔を検出せず、実物の顔だけを検出するといったことが可能となる。

### 3 顔検出と色弁別度追跡法を組み合わせた顔追跡システム

本研究では、提案手法1による高速顔検出法と色弁別度追跡法を統合することで、自動的に追跡が開始できる高速かつ安定な顔の検出・追跡システムを構築する。色弁別度追跡法は、入力画像の各画素をターゲットらしさを表す弁別度に変換した弁別性マップを構築し、弁別度の高い領域に楕円フィッティングすることで追跡を行う手法である。画素毎に弁別度を算出するため、ワイヤオブジェクトの追跡やオクルージョンに対する頑健性を持っている。また、テーブル化することで高速化した弁別度算出によって弁別性マップを構築しているため、非常に高速な追跡処理が行える。また、弁別度は再近傍識別によるターゲット検出としても利用でき、検出処理と追跡処理を統合することでターゲットの検出・追跡システムを構築して

いる。

しかし従来システムでは、あらかじめ肌色情報(ターゲットプロトタイプ)を教示し、ターゲット色に基づく検出を行っているため、教示した色と似た色が画像中に入ってくると誤ってターゲットとして検出し、追跡されてしまう問題がある。提案システムでは、色情報ではなく顔識別器を用い、入力画像中から顔だけを正確に検出することでこの問題を解決する。また従来手法では、照明環境の変化や人種、個人差による肌色の違いなどによって、あらかじめ教示したターゲット色と異なる追跡対象者を検出できなかったり、検出に成功したとしても追跡が安定に行えない問題がある。提案手法では、検出結果の情報を基に色の学習を自動で行う。顔検出で得られた顔領域内外の色情報を用いてターゲット(肌色)、非ターゲット(背景の色)のヒストグラムを構築し、背景には少なく追跡しやすいターゲット色を選択しながら弁別度 LUT の構築を行う。また、追跡中にも顔検出を行い、顔が検出された場合に弁別度 LUT の更新を随時行うことで、周辺環境の変化に対する頑健性を向上させる。

また、従来手法では、背景に対象と非常に似ている色が現れ対象と隣接してしまうと、追跡対象周辺の画素も高い弁別度を持ち、追跡範囲である楕円が広がりすぎてしまうという問題がある。提案手法では、従来研究によりステレオ視によって追跡対象までの距離を測ることができる利点を活かし、カメラから対象までの距離と画像内の楕円サイズを対応付けることで、より安定な探索範囲の決定ができる。提案手法では、人の顔検出・追跡問題に特化しているため、顔の大きさを予め仮定しておく、ステレオ視によって得られた距離情報から画像上での顔の大きさを推定することができる。この推定された顔サイズにより、追跡楕円がこれよりも大きくならないように制限することで、楕円が広がりすぎることなく安定な追跡が行える。

#### 4 顔検出と 6D K-means Tracker を組み合わせた髪追跡システム

髪領域は、テクスチャや色情報が少なく形状も変化するため、モデル化や安定な特徴を抽出することが難しく、これまでビデオレートでの追跡を行っている例は見当たらない。

本研究グループが提案している K-means Tracker は、画像上の追跡対象と背景の両方に対して複数のクラスタ中心を割り当て、5 次元特徴空間(3 次元の色、2 次元の画素位置)における K-means クラスタリン

グによって、サーチエリア内の各画素をターゲットと非ターゲットにラベリングすることにより追跡を行う手法である。特徴空間内でクラスタ中心と入力画素の距離を計算し、その距離に基づいてターゲットか非ターゲットのラベルが入力画素に付けられる。追跡対象が動けば、クラスタリングの結果、画像上でのターゲットクラスタ中心の重心位置が更新され、ターゲットの追跡が行われる。また、K-means Tracker は画素単位でのクラスタリングを行い、可変楕円をターゲット画素にフィッティングさせることで、追跡対象の大きさや形状の変化にも追従することができるため、様々な形状の髪領域を追跡するのに有効である。

しかし、5 次元の K-means Tracker には次のような問題点がある。追跡対象と類似する色を持つ背景画素がサーチエリアに混入した場合、5 次元特徴空間ではその背景画素が誤ってターゲット画素としてクラスタリングされる可能性が高く、その影響によりターゲット領域やサーチエリアが不安定となり、追跡が失敗してしまう場合がある。また、空間的距離を画像上の画素単位で計算するため、ターゲットが小さく写っている時は 2 次元座標における距離は小さくなり、ターゲットが大きく写っている時は 2 次元座標における距離は大きくなる。このように、画像内のターゲットの大きさによって、5 次元特徴空間内における 2 次元位置特徴ベクトルの割合が変わることで、大きく写っているターゲットの画素が、すべてターゲット画素としてクラスタリングされない可能性が高くなる。さらに、クラスタ中心の初期位置を手動で指定しなければならないという問題がある。

本研究では、2 次元の画素位置を奥行き情報により 3 次元の空間位置情報とした 6D K-means Tracker を用い、3 次元化された位置情報によって従来の K-means Tracker の問題を解決しながら追跡の安定化を行い、色彩情報の乏しい髪領域の追跡を実現する。また、髪領域を追跡することで、頭部の姿勢や顔の見え方にかかわらず頑健な頭部追跡が可能となる。

以上。



## 論文審査の結果の要旨

本論文は、奥行情報と色を用いた顔と髪を検出と追跡の高速化と安定化に関する研究について述べられたものである。第一の手法は、ステレオカメラで得られる2枚の画像それぞれで多重解像度画像上の顔検出を実行すると計算時間がかかりビデオレートでの処理が行えなくなる問題を、奥行きと画像上の顔の大きさの関係を利用し、先にステレオカメラで奥行きを大まかに計算して、画像上で探索する顔領域の大きさを限定することで速度と検出率の向上を実現している。第二の手法は、追跡途中、一定期間ごとに顔検出を行い、検出された顔の色のアップデートを行う手法であり、照明等の変動に強い追跡手法を提案している。第三の手法は、センサとして kinect を用い、奥行きを含めた 6D k-means Tracker によって髪の追跡を行う手法である。これは、画素の位置を実シーン中の座標で扱うため、元の k-means Tracker が持っていた画像上での対象の大きさの影響を受けやすいという問題を解決している。これにより顔・髪の両方を同時に追跡でき、センサに対して後ろを向いた際にも追跡が継続できるようになり、大まかな顔向きの推定も実現できる。第一の手法は、被引用回数 15 回であり、研究内容に関する注目度も高い。予備審査段階で指摘された論文の加筆修正なども適切になされており、最終的に若干の字句を修正すれば博士論文として評価できる内容になっている。

## 最終試験の結果の要旨

公聴会（平成27年8月26日）は全審査委員と20名の出席により開催され、1時間の研究発表と30分の質疑応答を行った。発表に対する質問に対しては、一部を除いてほぼ正確な回答が得られ、的確さと明確さの両面で良好な結果であった。

上記の結果を総合的に判断し、最終試験に合格したものと判定する。