

マンマシンインタフェースを対象としたCAMshiftによる実時間両手領域追跡

荒木 良介 池永研究室修士課程修了

研究背景

▶パソコンやスマートフォン、タブレット端末などを対象とした新しいタイプのマンマシンインタフェースが注目を浴びている

カメラを用いたマンマシンインタフェースに注目
身振り手振りで直感的な操作が可能

→ **手領域の検出・追跡が重要**

- 先行研究：
センサの装着・複数センサの併用 etc...
- 単眼カメラのみを用いた追跡手法は様々な場所・機器で利用可能



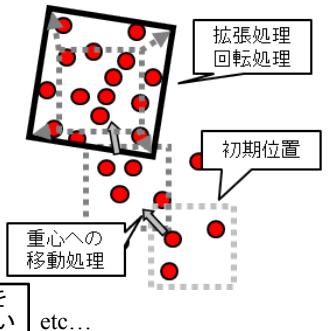
研究目的

□手領域の検出・追跡を単眼カメラのみで実現

- 実時間で処理
- 手の形に影響されない
- 両手を追跡可能

□ 追跡手法には**CAMshift法**を選択
→いくつかの問題を解決する必要

- 初期追跡対象の決定方法がない
- 背景と対象を区別できない



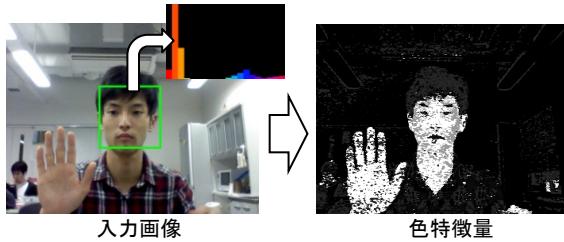
提案手法

顔領域からの色特徴量の作成

従来：色の特徴量は閾値や機械学習

提案：顔領域より基準となる肌色を取得

顔の認識にはHaar-Like特徴を使用
認識された領域から色相ヒストグラムを作成

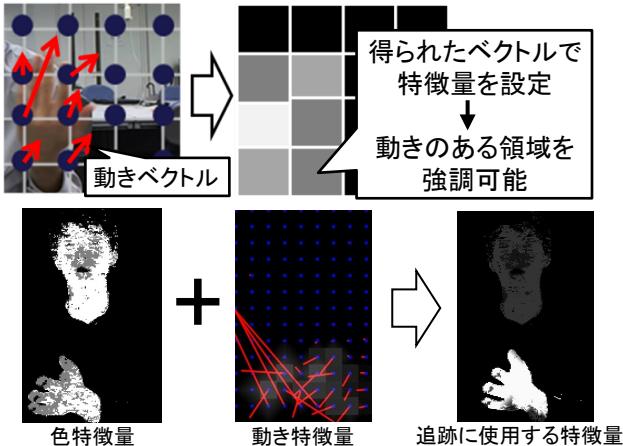


動きマスクの適用

従来：色の特徴量を使用
似た色が存在すると認識できない

提案：オプティカルフローによって得られたベクトルより特徴量を取得

→ **動きマスク**



動きベクトルを用いた追跡対象の自動決定

得られた領域が閾値以上の場合

追跡対象とすることで
対象決定の自動化



移動量に応じた追跡手法の切り替え

動きがない領域では動きマスクが機能しない

移動量が少なければ
Mean Shiftに切り替える



特徴量削減領域の設定

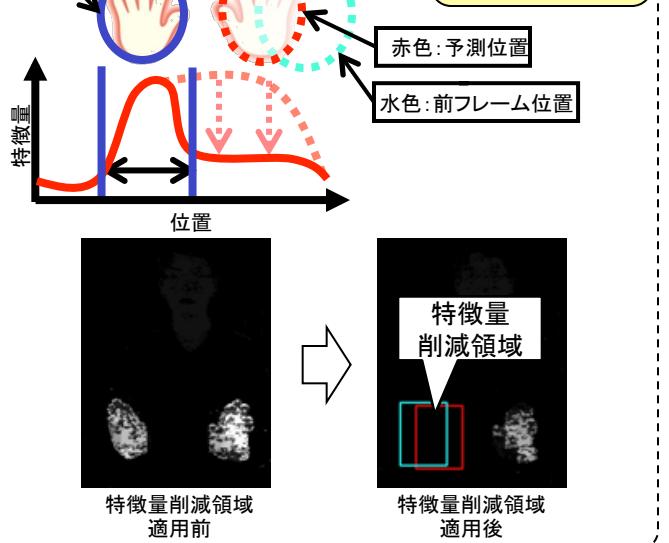
従来：類似対象の接近により特徴量の分布が広がると区別できない

提案：前フレームでの位置と動き予測による位置の特徴量を削減することでこの問題を回避

特徴量削減領域を設定
2次自己回帰モデルで動き予想

$$x(t) = A_1x(t-1) + A_2x(t-2) + Be(t)$$

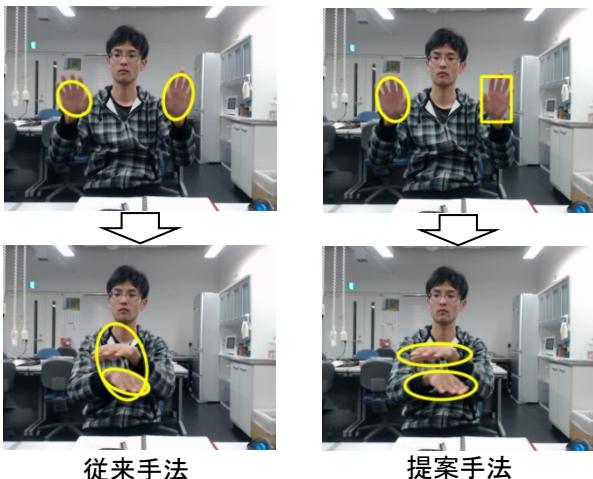
$x(t)$: t 秒での位置
 $e(t)$: t 秒でのノイズ
 A_1, A_2, B : 定数



実験結果

精度比較実験

- 従来手法と本提案手法で追跡精度を比較
- テスト動画サイズ: VGA(640×480)
- フレームレート: 30fps



実行時間の比較

- 従来手法と本提案手法で平均実行時間を比較
- 実験環境 → 「精度比較実験」と同様
- PCのスペック → CPU: Core 2 Duo (2.53GHz), メモリ: 4GB
- 色・動き特徴量取得時にOpenMPによる並列化

動画の内容	背景	フレーム数	従来手法 [ms]	提案手法 [ms]
両手を振る・近づける	物	647	8.37	15.6
両手を振る・近づける	物・人	1170	8.85	16.6
片手・両手を振る	物	482	8.71	15.8

- 従来手法と比較して2倍程度の処理量増加
→ しかし、提案手法では60 fps 程度の実行速度を実現
- 格子点の間隔を20×20 → 10×10にしても30 fps を超える処理が可能 → 追跡精度のさらなる向上が可能

ジェスチャ認識による評価

- 隠れマルコフモデル(HMM)によるジェスチャ認識を評価
- 手領域の追跡で認識が可能



Android端末での評価

- ARM系の低消費CPU環境で評価
- アプリケーションを作成した結果、平均処理速度が55[ms] → 改善する必要がある

