

拡張 K-means Tracker を用いた顔・髪の毛の追跡

鈴木 一正[†] 戚 意強[†] 呉 海元[†]

[†] 和歌山大学 〒640-8510 和歌山市栄谷 930 番地

E-mail: [†]{suzuki,seki,wuhy}@vrl.sys.wakayama-u.ac.jp

あらまし 本研究グループが提案している K-means Tracker [1] では、追跡対象と類似している色を持つ背景画素が追跡領域に混入した場合、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、追跡が不安定になるという問題がある。この問題を解決するために、追跡ターゲットとカメラ間の 3 次元空間の奥行き情報を追加した拡張 K-means Tracker [5] を提案している。本デモンストレーションでは、拡張 K-means Tracker の距離情報を kinect で獲得する手法を提案し、この手法を用いた顔や髪の毛の追跡システムを展示する。

キーワード kinect, 拡張 K-means Tracker, 顔・髪の毛追跡

1. K-means Tracker [1] の概要と問題点

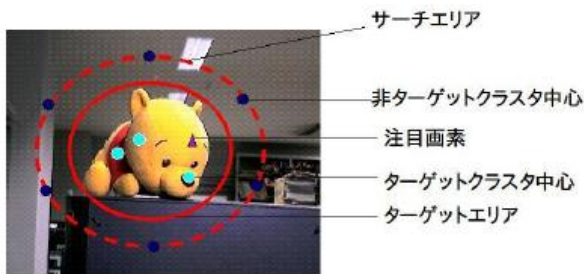


図 1 K-means Tracker の構成要素

K-means Tracker は、追跡対象と背景の両方に複数のクラスタ中心を割り当て、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分けることにより非剛体物体を追跡できる手法である。追跡対象を構成する画素の色の類似性と空間的近接性を同時に表現するために、各画素の特徴を 3 次元色空間 $c = [Y U V]^T$ と 2 次元座標空間 $p = [x y]^T$ からなる 5 次元ベクトル $f = [c p]^T$ として扱う。2 つの画素 a, b 間の距離 $d(f_a, f_b)$ は、色空間内のユークリッド距離と画像座標空間内のユークリッド距離を重み α により次式で定義している。

$$d(f_a, f_b) = \|c_a - c_b\|^2 + \alpha \|p_a - p_b\|^2 \quad (1)$$

ターゲットとしてクラスタリングされた画素集合に可変楕円を当てはめ、楕円のパラメータを毎フレーム更新することで、追跡対象の大きさや形状が変化しても安定な追跡が行える。

しかし、追跡対象と類似してい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を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、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問題がある (図 4, 5 参照)。

2. 拡張 K-means Tracker [5]

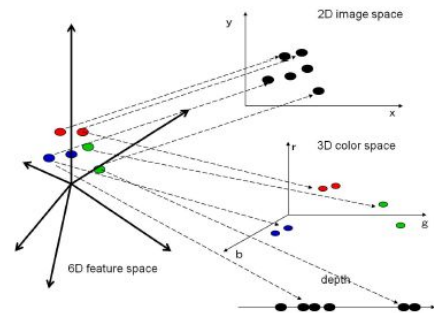


図 2 6 次元拡張特徴空間

K-means Tracker の類似背景問題を解決するために、K-means Tracker で定義した特徴空間を拡張する。3 次元ワールド座標系においてターゲットと背景が空間的に離れていると仮定すると、カメラからターゲット、背景までの 3 次元空間内の奥行き距離を利用することで、サーチエリア内に混入したターゲットと似たような色を持った背景画素を正しくクラスタリングすることができる。拡張 K-means Tracker では、従来の 5 次元特徴空間に奥行き特徴を新しく追加し、6 次元に拡張する (図 2 参照)。そこで、画素間の距離を次式のように変更する。

$$d(f_a, f_b) = \|c_a - c_b\|^2 + \alpha_1 \|p_a - p_b\|^2 + \alpha_2 \|d_a - d_b\|^2 \quad (2)$$

ここで、 α_1, α_2 はそれぞれ位置特徴と奥行き特徴の重み係数である。

3. 顔・髪の毛追跡システム

本デモンストレーションでは、拡張 K-means Tracker を用いた顔・髪の毛追跡システムを構築する。拡張 K-means Tracker で用いる距離情報は、これまでステレオカメラで取得していたが、ステレオ処理のコストが高く、サーチエリア内の全ての画素において距離を求めること

は難しかった．本システムでは，入力画像（640x480 画素）の距離情報が，kinect を用いることで高速（30fps）かつ高精度（11bit）に取得できる．図 3 は，kinect により得られた距離画像例である．明るいほど距離が短いことを意味する．

また本システムでは，顔や髪の毛といった追跡対象のクラスター情報を得るために，まず，距離情報を用いた高速な顔検出法 [2] を利用して顔検出を行う．この検出手法は，距離情報と実際の人物顔の大まかなサイズから画像上での顔サイズを推定することで，限られた位置とスケールのみでの探索を行い，ピラミッドスキャンする場合より高速に顔を検出することができる．図 3 の矩形は，探索する領域を表しており，各領域内を特定のスケールのみで探索する．探索を行う際の識別器には，OpenCV の顔検出ライブラリ関数 [3], [4] を用いている．

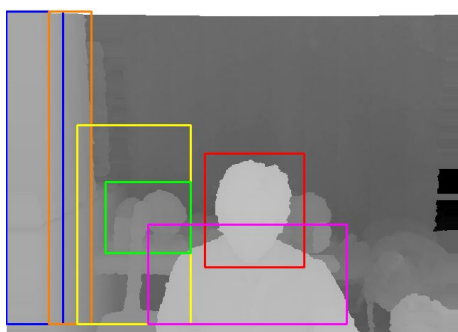


図 3 距離画像

顔検出に成功した場合，顔あるいは髪の毛の追跡に移行する．顔を追跡する場合は，顔検出によって得られた矩形の中心をターゲットクラスター，中心から顔のサイズだけ離れた円周上に非ターゲットクラスターを配置し，ターゲットクラスターから非ターゲットクラスターを半径とする円を初期フレームにおける可変楕円として追跡を開始する．髪の毛を追跡する場合は，顔を囲む矩形の上部にターゲットクラスター，そこから顔サイズ分だけ離れた円周上に非ターゲットクラスターを配置する．追跡時は，ターゲットとしてクラスタリングされた画素が少なくなった場合，対象の追跡に失敗したと判断し，再び検出に移行する．

図 4, 5 は従来手法（K-means Tracker）と提案手法（拡張 K-means Tracker）を比較した結果である．従来の K-means Tracker では背景がターゲット画素としてクラスタリングされてしまい，対象領域が広がって追跡が失敗するような場合においても，拡張 K-means Tracker では安定した追跡が行えていることがわかる．

4. おわりに

本論文では，kinect を用いることで，容易に高速・高精度な距離情報を得られることを利用し，拡張 K-means Tracker によって従来の類似色背景問題を解決できる手

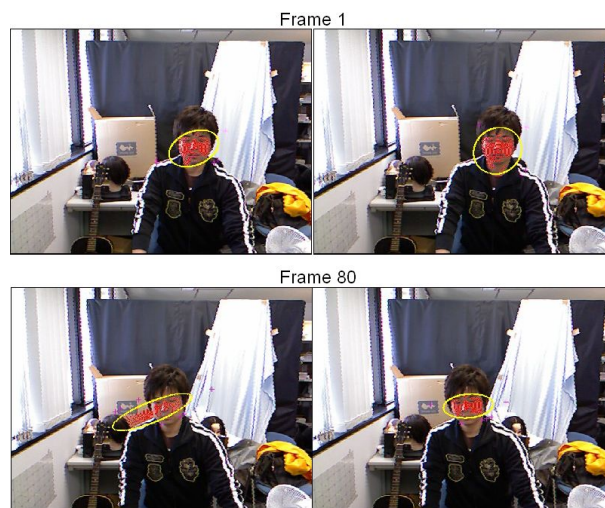


図 4 顔追跡結果例，(左) 従来手法，(右) 拡張 K-means Tracker



図 5 髪の毛追跡結果例，(左) 従来手法，(右) 拡張 K-means Tracker

法を提案した．また，顔検出と顔・髪の毛追跡システムを構築した．実機デモより，ターゲットとの類似色が背景に混入した場合でも，ターゲットと背景の奥行きの違いにより安定な追跡を行うことができることを示し，提案手法の有効性を確認した．

謝 辞

本研究の一部は，文部科学省科学研究費補助金基盤研究 (c)(一般)21500171 の補助を受けている．

文 献

- [1] Chunsheng Hua, Haiyuan Wu, Qian Chen, and Toshikazu Wada. K-means tracker: A General Algorithm for Tracking People. Journal of Multimedia, Vol. 1 Issue.4, pp. 46-53, July 2006.
- [2] Haiyuan Wu, Kazumasa Suzuki, Toshikazu Wada, Qian Chen. Accelerating Face Detection by Using Depth Information. In Proceedings of the 3rd Pacific Rim Symposium on Advances in Image and Video Technology, pp 657-667, Jan 2009.
- [3] Paul Viola and Michael Jones, Robust Real-Time Face Detection, IJCV, Vol.57, No.2, pp.137-154, 2004.
- [4] OpenCV, <http://www.intel.com/technology/computing/opencv>
- [5] Yiqiang Qi, HaiyuanWu, A Pixel-wise tracking algorithm using stereo camera, The 7th Int. Conf. on Ubiquitous Robots and Ambient Intelligence (URAI 2010), pp.198-201, 2010