

## 拡張 K-means Tracker による物体の追跡

戚 意 強<sup>†1</sup> 吳 海 元<sup>†1</sup>

本論文では、拡張 K-means Tracker という非剛体対象を追跡できるアルゴリズムを提案する。本研究グループが提案している K-means tracker では二つの問題点が残っている。1) 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された場合、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、追跡が不安定になる可能性が高い。2) 入力画像内の追跡対象が極端に大きい場合、ターゲット中心から離れたターゲット画素の画像空間内の距離が大きくなり、間違っって背景画素としてクラスタリングされる可能性が高くなる。拡張 K-means Tracker では、二つの拡張によって従来の問題を解決する。1) 3次元空間内の奥行き情報を追加し、画素の特徴空間を5次元から6次元に拡張する。1) マハラノビス汎距離による画像内の画素位置の情報を表現する。比較実験により提案手法の有効性を確認できた。

### Object tracking using Extended K-means Tracker

YI QIANG QI<sup>†1</sup> and HAIYUAN WU<sup>†1</sup>

In this paper, we propose a non-rigid object-tracking algorithm named extended K-means tracker. K-means tracker has two problem remained. 1) K-means tracker tracks unstably in a scene that a search area mixed by background pixels which have a similar to target pixels. 2) When target has a big size, those target pixels which far from the target center will have a big distance to the target center in image space, and they will be miss-classified into non-target cluster at a high possibility. Our algorithm solves those problems of K-means tracker by two extensions. 1) Add a depth feature to extend the 5D feature space to a 6D feature space. 2) Calculate Mahalanobis distance to describe the position distance. The effectiveness of the proposal algorithm was confirmed from comparison experiments.

<sup>†1</sup> 和歌山大学システム工学部  
 Faculty of Systems Engineering, Wakayama University

## 1. K-means Tracker の概要<sup>1)</sup> と問題点

### 1.1 K-means Tracker の概要

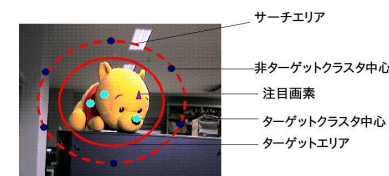


図 1 K-means Tracker の構成要素

K-means Tracker では、入力画像内の各画素の位置を表現する 2次元座標空間  $\mathbf{p} = [x \ y]^T$  と色を表現する 3次元色空間  $\mathbf{c} = [Y \ U \ V]^T$  の合わせた 5次元特徴ベクトル  $\mathbf{f} = [\mathbf{c} \ \mathbf{p}]^T$  で画素の情報を表現している。

二つの画素  $a, b$  間の距離  $d(\mathbf{f}_a, \mathbf{f}_b)$  はユークリッド距離で定義している。

$$d(\mathbf{f}_a, \mathbf{f}_b) = \|\mathbf{c}_a - \mathbf{c}_b\|^2 + \alpha \|\mathbf{p}_a - \mathbf{p}_b\|^2 \quad (1)$$

クラスタが未知である注目画素  $\mathbf{f}_u$  がターゲットクラスタ中心  $\mathbf{f}_T$  への最短距離を  $D_T$ 、非ターゲットクラスタ中心  $\mathbf{f}_{NT}$  への最短距離を  $D_{NT}$  を以下のように定義される。

$$D_T = \min_{i=1 \sim n} \{d(\mathbf{f}_{Ti}, \mathbf{f}_u)\} \quad (2)$$

$$D_{NT} = \min_{j=1 \sim m} \{d(\mathbf{f}_{NTj}, \mathbf{f}_u)\} \quad (3)$$

ここで、 $m, n$  をそれぞれターゲットクラスタ中心と非ターゲットクラスタ中心の個数とする。 $D_T$  と  $D_{NT}$  を比較することによって、注目画素  $\mathbf{f}_u$  をターゲットクラスタか非ターゲットクラスタにクラスタリングされる。

ターゲットとクラスタリングされた画素の集合を正規分布と仮定し、ターゲット画素点の分布のマハラノビス距離より可変楕円のパラメータ（長軸，短軸，傾き，重心位置）を求め、現フレームにおけるターゲットエリア楕円を推定する。この楕円を  $r$  倍拡大し、次フレームにおけるサーチエリア楕円、非ターゲットクラスタ中心を同時に更新する。

## 1.2 K-means Tracker の問題点

### 1.2.1 類似色背景問題

K-means tracker は色をメイン特徴とするので、ターゲットと類似色を持った背景画素は 5 次元特徴空間において、ターゲット色に相似なので、 $\|c_T - c_u\|^2 \approx 0$ 。また、ターゲットに近づくと、 $\|p_T - p_u\|^2 \approx 0$ 。その結果、 $d(f_T, f_u) \approx 0$  になる。つまり、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された場合、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なる。その影響で現フレームのターゲット領域と、次フレームのサーチエリアを正確に決められず、追跡が不安定になり、場合によっては追跡失敗になる。

### 1.2.2 ターゲットサイズ問題

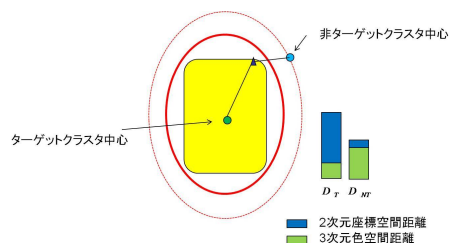


図 2 ターゲットサイズ問題

図 2 で示すように、入力画像内の追跡対象が極端に大きい場合、ターゲット中心から離れたターゲット画素の画像空間内の距離が大きくなる。この際、2 次元画素の位置特徴ベクトルのユークリッド距離が 5 次元特徴空間内における距離の割合が大きくなり、3 次元色空間の特徴が軽視される。ターゲット画素であるにも関わらず、間違っって背景画素としてクラスタリングされる可能性が高くなる。

## 2. 拡張 K-means Tracker

### 2.1 K-means Tracker における特徴空間の拡張

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。本論文は、3 次元ワールド座標系において、ターゲットと背景の間距離的に離れていると仮定する。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積極的に利用することで、サーチエリアに混入され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できると考える。拡張 K-means Tracker では、従来の 5 次元特徴空間を図 3 のように、奥行き特徴を新しく追加し、6 次元に拡張する。そこで、画素間の距離は以下のように変更する。

$$d(f_a, f_b) = \|c_a - c_b\|^2 + \alpha_1 \|p_a - p_b\|^2 + \alpha_2 \|d_a - d_b\|^2 \quad (4)$$

ここで、 $\alpha_1, \alpha_2$  はそれぞれ位置特徴と奥行き特徴の距離計算する際の重み係数である。

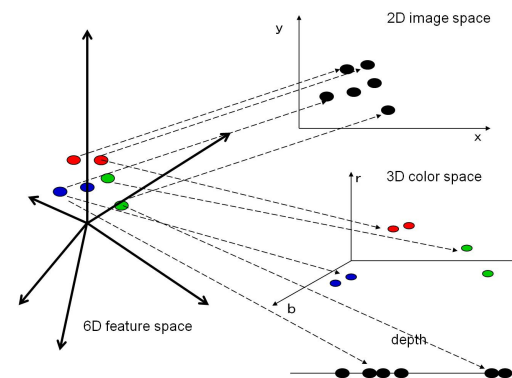


図 3 6次元拡張特徴空間

### 2.2 ROI の動的配置法

ステレオカメラは、カラー画像と奥行き情報を同時に提供でき、奥行き特徴を得るには特殊なデバイスを追加しなくてもよいというメリットがあるので、本研究では Point Gray Research 社の BumbleBee というステレオカメラを用いて追跡アルゴリズムを実装する。

BumbleBee ステレオカメラは付属の Triclops SDK<sup>2)</sup> を用いて、高速に画素単位で奥行き情報を計測することができる。しかし、我々の予備実験より、1 フレームに対して、すべての入力画素を計測するには約 20ms かかることがわかった。その後、追跡処理を行うと、ビデオレートでの追跡は難しい。ここで、我々は精度を保つ上、計算量を削減する手法として、ROI の動的配置法を提案する。

ROI とは Region of Interest の略で、画像上に指定された矩形領域のことである。我々の予備実験により、ビデオレートで奥行きを計測できる画素数が 1 万点以下であり、ROI の最多個数が 100 個であることがわかった。提案システムはビデオレートで動作するために、ROI は 100 個とする。また、できるだけ多い画素の奥行き情報を推定するために、

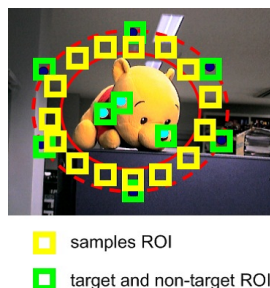


図 4 ターゲット・非ターゲットクラスタ中心とサンプル画素の ROI 配置

各 ROI の面積は  $10 \times 10$  画素とする。本論文では、正しくクラスタリングを行うために、限られた数の ROI を効率的に配置する方法を提案する。

クラスタリングする際、ターゲットと非ターゲットクラスタ中心の持っている特徴が最も重要な情報なので、ターゲットと非ターゲット中心に優先的に ROI を配置する。K-means tracker はターゲット色を最大 10 色と仮定し、均等に非ターゲットクラスタ中心を 8 個配置するため、ターゲットと非ターゲットクラスタ中心に最大 18 個の ROI を配置することになる。

ターゲットエリアはターゲット画素により更新するため、ターゲットエリア内の画素はターゲット画素である確率は高いと考えられる。ここで、ターゲットエリア内の画素の奥行きはターゲットセンタの値と同じであると仮定する。サーチエリアとターゲットエリア間における画素のクラスタリングが誤り発生しやすく、これらの画素が正しくクラスタリングできるか否かは追跡の安定さに影響する。そのため、残った数の ROI を全てサーチエリアとターゲットエリア間の領域に配置する。

K-means tracker はターゲットエリア楕円の中心 (ターゲット中心でもある) から、10 度ずつ、3 画素間隔に画素をサンプリングするため、より多くのサンプル画素をカバーするように、ROI をそれに合わせ、10 度ずつ、5 画素間隔に配置する。また、カバーされないサンプル画素はターゲットと同じ奥行き特徴を持たず。

ROI が図 4 で示すように配置され、毎フレームターゲット・非ターゲットクラスタ中心、ターゲットエリア楕円中心の更新と共に更新していくことによって、動的に配置される。

### 2.3 マハラノビス汎距離による位置特徴距離の計算

前章 1.2.2 で述べたように、ターゲットのサイズが大きくなると、ターゲット中心と離れ

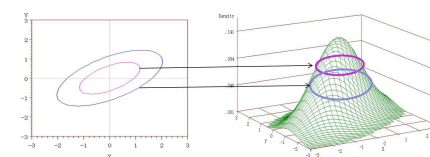
たターゲット画素が非ターゲット画素としてクラスタリングされる問題がある。

従来の K-means tracker は位置特徴距離の計算をユークリッド距離を用いて計算したため、画像において、離れた分だけ位置特徴距離が大きくなる。その結果、位置的に離れたターゲットと同じ色の画素 (ターゲット画素) であっても、位置特徴が無制限に大きくなり、ターゲットとの特徴空間における距離全体が大きくなり、非ターゲット画素としてクラスタリングされてしまう。

ここで、我々は位置特徴距離の計算に、マハラノビス汎距離を導入する。マハラノビス汎距離は以下のように定義される。

$$D_M(\mathbf{x}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})} \quad (5)$$

ここで、 $x$  はデータ集合、 $\mu$  がデータの平均、 $\Sigma$  はデータの共分散行列である。



確率等高線となり、中心への距離＝確率

図 5 マハラノビス汎距離

マハラノビス汎距離は図 5 で示すように、ユークリッド距離とは違い、ここでのターゲット座標の分散共分散を考慮し、同距離線は確率等高線となり、確率意味での距離となる。つまり、画像上の (ユークリッド) 座標距離によらない特性がある。

ここで、サンプル画素の位置特徴ベクトルがターゲット中心の位置特徴ベクトルへのマハラノビス汎距離を以下のように定義する。

$$D_{MT}(\mathbf{P}_u) = \sqrt{(\mathbf{P}_u - \boldsymbol{\mu}_e)^T \boldsymbol{\Sigma}^{-1} (\mathbf{P}_u - \boldsymbol{\mu}_e)} \quad (6)$$

ここで、 $x$  はクラスタリングされたターゲット画素の座標集合、 $\mu$  がターゲット画素座標の平均、 $\Sigma$  はターゲット画素座標の共分散行列である。

また、ターゲットと背景は同じ画像にあり、同じ確率分布に従うので、非ターゲットクラスタ中心のマハラノビス汎距離  $D_{MNT}$  を以下のように定義する。

$$D_{NTi}(\mathbf{P}_u) = \sqrt{(\mathbf{P}_u - \mathbf{P}_{NTi})^T \Sigma^{-1} (\mathbf{P}_u - \mathbf{P}_{NTi})} \quad (7)$$

ここで、 $\mathbf{P}_{NTi}$  は  $i$  番目の非ターゲットクラスタ中心の座標特徴ベクトルである。非ターゲット中心までの最短マハラノビス汎距離は以下になる。

$$D_{MNT} = \min\{D_{NTi}\} \quad (8)$$

最終的に、拡張 K-means Tracker の画素間の距離は以下のように定義する。

$$d(\mathbf{f}_a, \mathbf{f}_b) = \|\mathbf{c}_a - \mathbf{c}_b\|^2 + \alpha_1 D_M(\mathbf{P}_u) + \alpha_2 \|d_a - d_b\|^2 \quad (9)$$

6次元特徴空間において、クラスタリングする際、従来 K-means tracker と同じく、式(2)と式(3)により、サンプル画素の  $D_T$  と  $D_{NT}$  を比較することによってクラスタを決める。

ターゲットエリア楕円フィッティングする際、ターゲット画素集合の分散共分散行列が計算されている。この行列を用いてマハラノビス汎距離を求める計算量と、従来手法のユークリッド距離の計算量との差が無視できる。

### 3. 実験

#### 3.1 類似色背景の比較実験

本論文 2.1 で提案した K-means tracker における特徴空間の拡張について、有効性を確認するため、従来の K-means tracker との比較実験を行った。図 6 で示すように、フレーム 180 で、ターゲットと似たような色(黄色)を持った背景はサーチエリアに混入した際、従来の K-means tracker はそれらの画素をターゲットクラスタにクラスタリングした。それに対して、提案手法は奥行き特徴を用いるため、ターゲットと類似色を持つ背景画素を正しく非ターゲットにクラスタリングできた。また、ROI を動的に配置することによって、ビデオレート (30fps) での追跡ができることを確認できた。

#### 3.2 大きいターゲットの場合の比較実験

図 7 で示したように、大きいターゲットを追跡する場合、提案手法がターゲットを含む楕円を正しく推定できた。一方、従来手法では、ターゲットエリアを示す楕円がターゲットの内部に入り込み、ターゲットの一部しか囲めなかった。本論文 2.3 で提案したマハラノビス汎距離による位置特徴の距離計算の有効性を確認できた。また、追跡はビデオレートで行えることも確認できた。

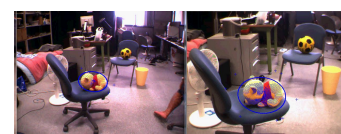
### 4. Conclusion

本論文では、従来 K-means tracker の類似色背景問題と ターゲットサイズ問題を解決す

るために、入力画素の情報を表現する特徴空間を拡張し、拡張された 3次元空間内の奥行き情報をビデオレートで計算するため、ROI の動的配置方法を提案した。また、マハラノビス汎距離による画素位置の特徴距離を計算する方法を提案した。比較実験により、提案手法の有効性を確認できた。

### 参考文献

- 1) C. Hua, H. Wu, Q. Chen, T. Wada: "A General Framework for Tracking People", Journal of Multimedia, ISSN: 1796-2048, Vol.1, Issue 4, pp. 46-53, July 2006
- 2) Point Grey Research, "TORICLOPS Software Development Kit", Version 3.1



frame 30



frame 180



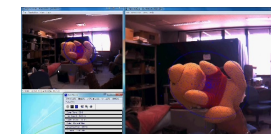
frame 210



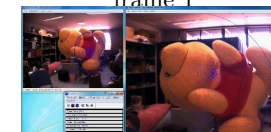
frame 240

左:提案手法 右:従来の K-means tracker

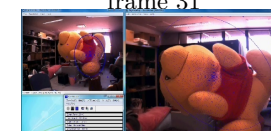
図 6 類似色背景の比較実験



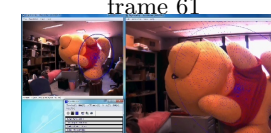
frame 1



frame 31



frame 61



frame 91

左:従来の K-means tracker 右:提案手法

図 7 ターゲットが大きくなった場合の比較実験