

# 差分ステレオを用いた人物セグメンテーションの検討

寺林 賢司\* 生形 徹 梅田 和昇 (中央大学/JST CREST)

Object Segmentation with Subtraction Stereo for Human Detection

Kenji Terabayashi\*, Toru Ubukata, Kazunori Umeda (Chuo University / CREST, JST)

## Abstract

Object segmentation with a stereo camera is improved by changing coordinate system and kernel type for mean shift clustering. Experimental results show two main findings: (i) polar coordinate system is better than orthogonal coordinate system for making histograms of measured three-dimensional points from a stereo camera; (ii) Gaussian kernel can improve true positive rate of object segmentation without increasing the false positive rate.

キーワード: ステレオカメラ, セグメンテーション  
(Stereo camera, segmentation)

## 1. はじめに

画像中から人物を検出することは、人流計測や監視カメラなど様々なアプリケーションにおいて重要な技術である。画像中において人物全体が観測できる場合に対して、遮蔽によって一部のみ観測可能な状況での検出は難しく、ステレオカメラを使ったセグメンテーションが検討されてきた<sup>(1)(2)</sup>。本稿では、差分ステレオ<sup>(3)</sup>を使った人物セグメンテーションにおける座標系とクラスタリングに利用するカーネルの検討について報告する。

## 2. 差分ステレオを用いた人物セグメンテーション

先行研究<sup>(2)</sup>の人物セグメンテーション方法について、概要を述べる。差分ステレオ<sup>(3)</sup>によって計測した前景領域の三次元計測点を地面と水平な平面へ投影し (図 1 参照), その平面上における計測点の分布を Mean Shift クラスタリングで分類することで人物セグメンテーションを行う。このクラスタリングは、計測点群そのものではなく、平面をセル分割して作成した計測点のヒストグラムに対して行う。これによって計算時間の削減と、誤検出の低減を図っている。この際、平面上での座標系は、直交座標系であり、Mean Shift クラスタリングで利用するカーネルは矩形で一樣重みを用いている。

## 3. 座標系およびカーネルの検討

ステレオカメラで人物を計測した際に、その計測点群は、計測誤差の影響を受け、カメラと人物をむすぶ方向へ主に

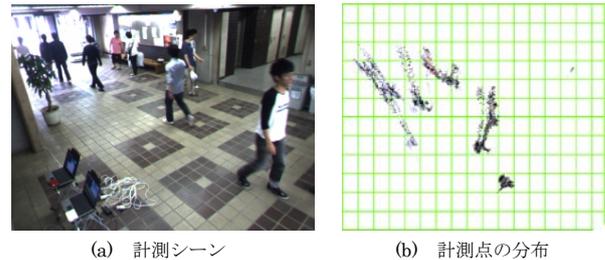


図 1 差分ステレオを用いた人物セグメンテーション  
Fig. 1. Object segmentation with subtraction stereo.

分布し、その広がり距離の二乗に比例する (図 1 (b)参照)。このことを考慮し、極座標系によるセグメンテーションを提案する。極座標系をとることで、人物の計測点分布の主軸が座標軸方向と沿い、複数人物が重なるようなオクルージョン環境下においても Mean Shift クラスタリングによる人物セグメンテーションの向上が期待できる。また、Mean Shift クラスタリングのカーネルをガウシアンカーネルとすることで、より混み合った状況でのセグメンテーションの改善が期待できる。

本稿では、座標系とカーネルの違いによる検討に加え、ヒストグラムを作成するときのセルサイズについても検討する。前述のように、ステレオカメラからの計測点は、距離の二乗に比例して誤差の影響を受け分布する。したがって、計測距離に応じて動的にセルサイズを変化させることを行う。

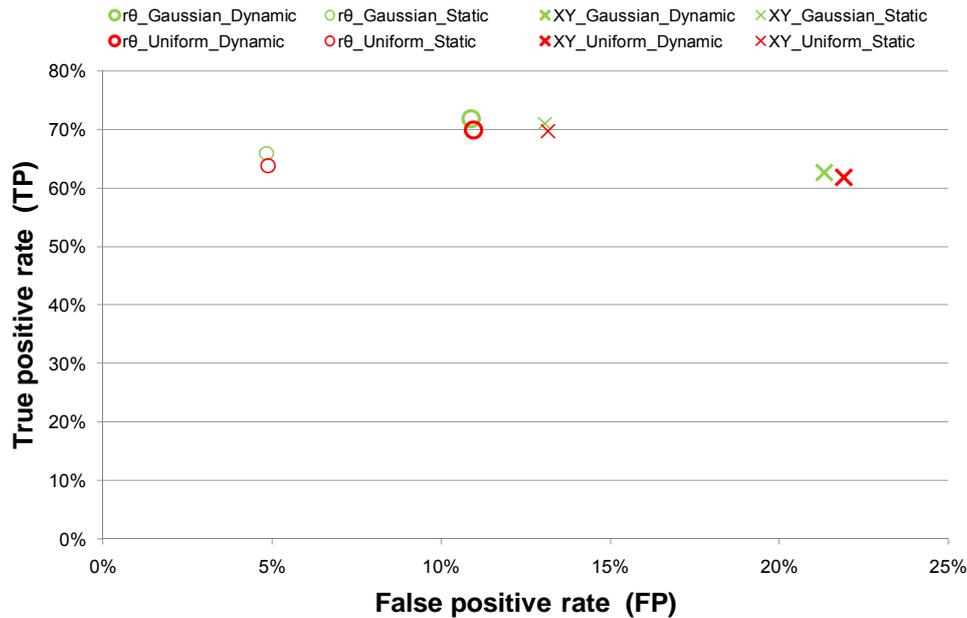


図 2 人物セグメンテーションの比較結果

Fig. 2. Comparative results of human segmentation.

#### 4. 実験

座標系，カーネル，セルサイズの違いが人物セグメンテーションに与える影響を検証する。座標系は，直交座標系 (XY) と極座標系 (rθ) の比較を，Mean Shift におけるカーネルは，矩形一様 (Uniform) とガウシアンカーネル (Gaussian) の比較を，セルサイズについては，固定値 (Static) と距離に応じた変化 (Dynamic) を比較する。計測シーンは図 1 (a) を含む 3000 フレームとし，このうち等間隔で抽出した 1000 フレームを用いて比較評価を行った。このシーンでは，複数の人が行き交い，人物同士のオクルージョンが多数発生している。

検出率 (TP) と誤検出率 (FP) をまとめたものを図 2 に示す。本グラフでは，左上に行くほど，検出率が高く，また誤検出率が低く，より良いセグメンテーションができていくことを表す。

座標系の比較では，その他の条件によらず，極座標系 (rθ, ○印) が直交座標系 (XY, ×印) より優れた結果を示した。特に，誤検出の低下に極座標系の表現が寄与しており，計測点の分布方向を考慮することの重要性を反映した結果となった。

Mean Shift クラスタリングのカーネルについての検討では，ガウシアンカーネル (Gaussian, 緑) が矩形一様カーネル (Uniform, 赤) より良い結果を示した。多くのケースにおいて，検出率の向上にガウシアンカーネルが寄与しており，遮蔽領域が多いオクルージョンの時でもセグメンテーションできていることが特徴であった。

セルサイズの効果については，座標系の取り方に依存する結果となった。直交座標系 (XY, ×印) では，距離に応

じたセルサイズの変化 (Dynamic, 太線) は，固定値 (Static, 細線) に対して悪い結果を示しており，極座標系 (rθ, ○印) では，検出率と誤検出率のトレードオフとなった。この点は，距離にも依存する結果を示しているため，今後より詳しい検討が必要である。

#### 5. おわりに

本稿では，差分ステレオを用いた人物セグメンテーションについて，座標系，カーネル，セルサイズの検討を行った。実験結果から，極座標表現が誤検出の低下に寄与し，ガウシアンカーネルを用いることで，より複雑なオクルージョンにおける検出率の向上が確認された。セルサイズの検討については，距離に依存した結果が得られたことから，今後更なる詳しい検討を行う必要がある。

#### 文 献

- (1) S. Bahadori, L. Iocchi, G. Leone, D. Nardi, and L. Scozzafava: "Real-time people localization and tracking through fixed stereo vision", *Applied Intelligence*, Vol. 26, No. 2, pp.83-97 (2007)
- (2) T. Ubukata, K. Terabayashi, A. Moro, K. Umeda: "Multi-object segmentation in a projection plane using subtraction stereo", *Proceedings of the 20th International Conference on Pattern Recognition (ICPR2010)*, pp.3296-3299 (2010)
- (3) 梅田和昇, 寺林賢司, 橋本優希, 中西達也, 入江耕太: 「差分ステレオ—運動領域に注目したステレオ視—の提案」, *精密工学会誌*, Vol. 76, No. 1, pp. 123-128 (2010)