

色情報と差分ステレオを組み合わせた人物トラッキング

中央大学/JST CREST ○星川佑磨, 中西達也, 橋本優希, 寺林賢司, 梅田和昇

Human Tracking Using Color Information and Subtraction Stereo

Chuo Univ. / JST CREST Yuma HOSHIKAWA, Tatsuya NAKANISHI, Yuki HASHIMOTO,

Kenji TERABAYASHI, Kazunori UMEDA

We propose a method for human tracking using a stereo camera system called “Subtraction Stereo” and a color information. The tracking system using the subtraction stereo, which focuses its stereo matching algorithm to regions obtained by background subtraction, is realized using Kalman filter. To make the tracking system more robust, the new method also uses color information as an another distinctive information of person. The effectiveness of the new proposed method is verified in the scene which is difficult to realize without color information.

1. 序論

近年, ステレオカメラを用いた人物の検出, トラッキングに関する研究が多く行われている[1]. これらの多くが, カラー画像上における人物の二次元情報を, 視差空間における三次元情報と統合することで人物を検出, トラッキングする手法である. 中でも, ステレオマッチングをロバスト化, 高速化する差分ステレオという手法が提案されており, 本手法とカルマンフィルタ[2]を用いた人物トラッキング手法が提案されている[3]. 本研究では, この差分ステレオを用いた人物トラッキング手法に, 新たに人物の色情報の特徴量として加えることで, トラッキング性能を向上させる手法を提案する.

2. 差分ステレオ

差分ステレオでは, まず左右の画像それぞれにおいて, 背景差分により前景領域を抽出し, 対応点探索の領域を前景に限定しマッチングを行うことで視差画像を得る. 本手法で得られる視差画像から, 検出物体の三次元座標と幅, 高さを取ることが出来る. 得られる視差画像の例を Fig.1 に示す.

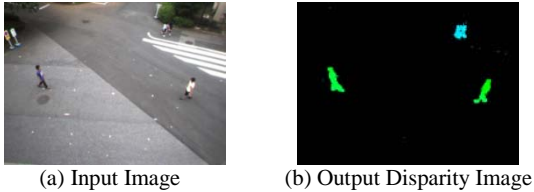


Fig.1 Disparity Image Obtained by Subtraction Stereo

3. 人物トラッキング手法

3.1 カルマンフィルタを用いた人物トラッキング

人物の動きが概ね等速直線運動であるモデルを立て, カルマンフィルタに与える運動モデルとする. 状態変数 \mathbf{x} は, 世界座標系での人物の位置 (x, y, z) および速度 $(\dot{x}, \dot{y}, \dot{z})$ を用い, 次式で表わされる.

$$\mathbf{x} = [x \ \dot{x} \ y \ \dot{y} \ z \ \dot{z}]^T \quad (1)$$

カルマンフィルタの状態方程式は次式で与えられる.

$$\mathbf{x}_{i+1} = \Phi \mathbf{x}_i + \Gamma_i \boldsymbol{\omega}_i \quad (2)$$

Φ は状態遷移行列であり

$$\Phi = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

で与えられる. $\boldsymbol{\omega}_i$ はプロセスノイズである. 観測変数には,

$$\mathbf{z} = [u \ v \ d]^T \quad (4)$$

を用いる. u, v は画像上の座標, d は視差である. 観測変数と状態変数の関係は次式で表わされる.

$$\mathbf{z}_i = \mathbf{f}(\mathbf{x}_i) + \mathbf{v}_i, \quad \mathbf{f}(\mathbf{x}_i) = \begin{bmatrix} x_i \cdot f & y_i \cdot f & b \cdot f \\ z_i & z_i & z_i \end{bmatrix}^T \quad (5)$$

f はカメラの焦点距離, b は基線長, \mathbf{v}_i は観測ノイズである. これらの式からカルマンフィルタを更新し予測値を得る.

カルマンフィルタの予測値と, 人物の観測値を対応づけることでトラッキングを行う. カルマンフィルタの予測値の集合を K , その要素を k_i とし, 観測値の集合を M , その要素を m_j とする. i 番目の予測値から j 番目の観測値までの三次元空間でのユークリッド距離 $D_E(k_i, m_j)$ を計算し, ユークリッド距離が閾値 D_{Th} 以内にあり, かつ最小のものを対応づける.

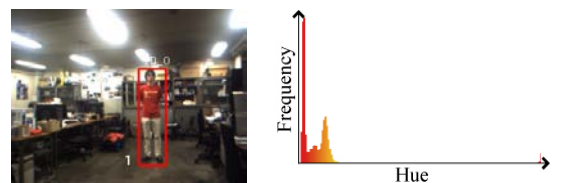
人物のすれ違いなど, 人物領域の重なりが起こるシーンでは, 複数の予測値に対応づける観測値が一致してしまうためトラッキングが出来ない. このような場合は, 観測値を用いずカルマンフィルタの更新を行い, 人物をトラッキングする.

3.2 色情報を組み合わせたトラッキング手法

3.1 節で示した手法は, 人物の領域同士に重なりが生じている間に, 人物の動きが予測に沿わない場合や, 人物間の距離が近い場合には安定したトラッキングが難しい. そこで本研究では, 人物の色情報を用いてこの問題を解決する. 色情報には, 色相値から得られる色ヒストグラム間の Bhattacharyya 距離を用いる. この Bhattacharyya 距離を三次元空間のユークリッド距離に重み付けし加え, 新しい距離を定義し, この距離をカルマンフィルタの予測値と観測値の対応付けに用いる.

色情報を用いるには, それぞれの人物の特徴となる色情報を取得する必要がある. そこで, ある人物が計測されてから 10 フレーム人物領域の色ヒストグラムを計算し, その平均をその人物のモデル色ヒストグラムとする. モデルヒストグラムは, その人物をトラッキングするカルマンフィルタに対応する. そして, 各モデル色ヒストグラムと, 毎フレーム計算される人物領域の色ヒストグラム間の Bhattacharyya 距離を計算する. 人物のモデル色ヒストグラムの例を Fig.2 に示す.

Bhattacharyya 距離は, 二つの正規化したヒストグラム間の類似度を表す Bhattacharyya 係数を距離値として表したものである. Bhattacharyya 係数 ρ は次式で表される.



(a) Scene (b) Model Color Histogram
Fig.2 Example of Model Color Histogram

$$\rho = \sum_{u=1}^{m+1} \sqrt{p_u q_u} \quad (6)$$

ここで、 p, q は比較対象となる正規化された色相分布 ($\sum_{u=1}^m p_u = 1, \sum_{u=1}^m q_u = 1$), u は色相番号, m は色相の成分数を表す。 ρ を用いて、Bhattacharyya 距離 D_B は次式で表される。

$$D_B = \sqrt{1 - \rho} \quad (7)$$

式(7)の Bhattacharyya 距離を、三次元空間でのユークリッド距離に重み付けし加えた距離 $D_{EB}(k_i, m_j)$ を定義し、予測値と観測値の対応付けに用いる。 $D_{EB}(k_i, m_j)$ は次式で定義する。

$$D_{EB}(k_i, m_j) = \alpha_i \cdot \frac{D_E(k_i, m_j)}{D_{Th}} + (1 - \alpha_i) \cdot D_B(k_i, m_j) + \delta(\alpha_i) \cdot U(D_E(k_i, m_j) - D_{Th2}) \quad (8)$$

ここで $D_B(k_i, m_j)$ は、 i 番目の予測値に対応するモデル色ヒストグラムと、 j 番目の観測値の色ヒストグラム間の Bhattacharyya 距離とする。また、観測値の集合 M の中で、 i 番目の予測値の閾値 D_{Th} 以内にある観測値の集合を $M_{Th,i}$ とする。式(8)の α_i は、集合 $M_{Th,i}$ の要素数 n に応じて次式で定義する。

$$\alpha_i = \begin{cases} 0 & (n=0) \\ 1 & (n=1) \\ \min_{s,t \in M_{Th,i}} \left(\frac{|D_E(k_i, m_s) - D_E(k_i, m_t)|}{D_{Th}} \right) & (n \geq 2) \end{cases} \quad (9)$$

α_i は、ユークリッド距離のみでの対応付けが曖昧になると、色情報に重みをおいて対応付けを行うよう定義した。式(8)において D_{Th2} は、 $\alpha_i = 0$ で観測値の対応付けを色情報のみで行う際に、予測値から極端に離れた観測値を用いないようにするためのユークリッド距離の閾値である。式(8)で定義する距離が閾値 D_{Th3} 以内にあり、かつ予測値までの距離が最小のものを観測値としカルマンフィルタを更新する。

4. 実験

3.2節で示した色情報を用いたトラッキング手法の有効性を示すため、被験者実験を行った。実験は、Fig.3に示すような、人物の領域同士に重なりが生じている間に、各人物が歩く方向を変え、予測に沿わない動きをするシーンにおいて行った。

ステレオカメラは Bumblebee2(Point Grey Research, カラー, $f=3.8\text{mm}$, $b=120\text{mm}$, VGA, 48fps)を使用した。カメラの位置姿勢は地面から高さ 8.3[m], 仰角 50°とした。プロセスノイズを速度に対して 0.2[m/frame], 観測ノイズを画面上の座標 u, v に対して 0.3[pixel], 視差に対して 0.04[pixel]に設定した。

実験結果を Fig.4 に示す。人物の枠の色はそれぞれの ID を表す。実験結果より、色情報を用いることで人物のトラッキングが正しく行われたことが分かる。また、それぞれの人物の x, y 座標(世界座標)の計測結果を、実験シーンを俯瞰した形で Fig.5 に示す。計測結果より、人物の歩いた軌跡が正しく得られていることが分かる。

5. 結論

カルマンフィルタを用いたトラッキング手法に色情報を加え、トラッキング性能を向上させた。今後の展望は、更に複雑な場面において正確なトラッキングを実現することである。

参考文献

- [1] 鈴木達也, 岩崎慎介, 佐藤洋一, 杉本晃宏: “環境モデルの導入による人物追跡の安定化”, 画像の認識・理解シンポジウム MIRU2004 講演論文集, Vol.1, pp.243-248, 2004.
- [2] Sebastian Thrun, Wolfram Burgard, Dieter Fox, 上田隆一 訳: “確率ロボティクス”, 毎日コミュニケーションズ, 2007.
- [3] 中西達也, 寺林賢司, 橋本優希, 星川佑磨, 梅田和昇: “差分ステレオを用いた複数人物計測”, 日本機械学会ロボティクス・メカトロニクス講演会2009講演論文集, 1A1-D16, 2009.

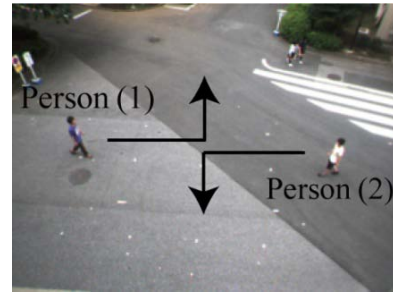


Fig.3 Experimental Scene

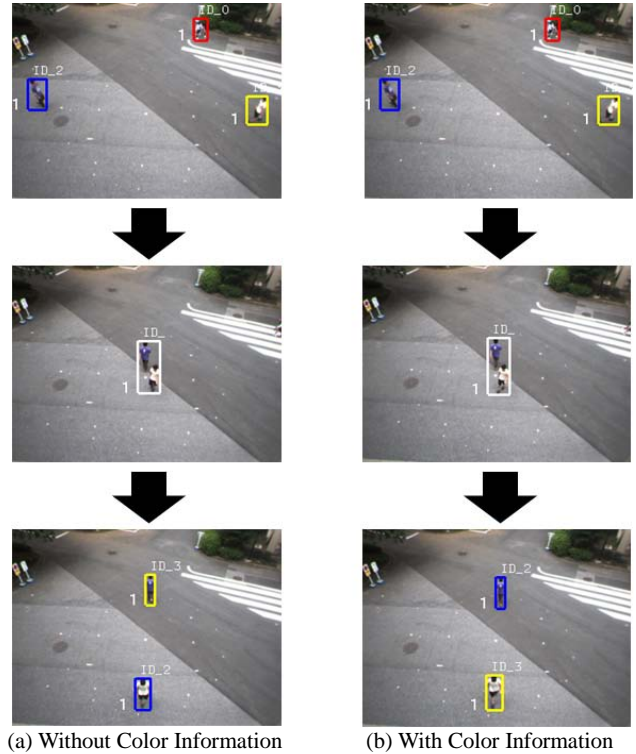


Fig.4 Experimental Result

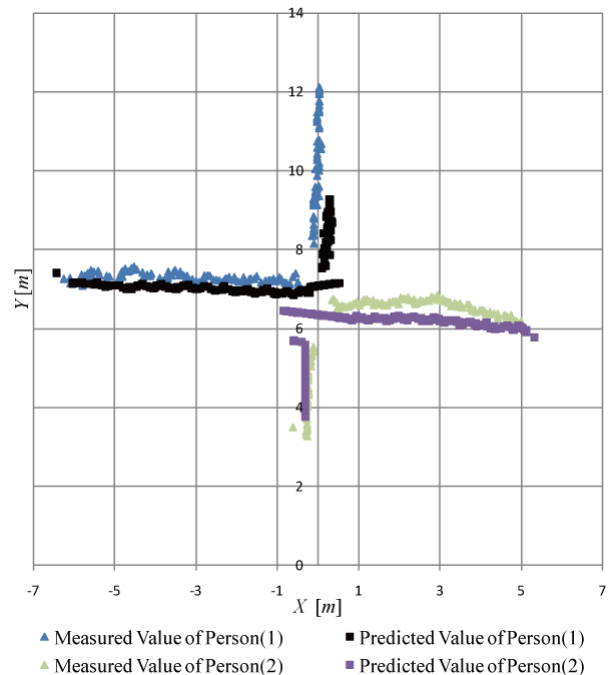


Fig.5 Measured Values of Each Persons with Color Information