

## 色分布のクラスタリングによる所持物の検知 Detection of the Carrying Object by Clustering of Color Distribution

瀧田 悠一<sup>†</sup> 鈴木 輝彦<sup>‡</sup> 太原 育夫<sup>‡</sup>  
Yuichi Takita Teruhiko Suzuki Ikuo Tahara

### 1. はじめに

映像監視システムの目的は不審行動の発見である。しかし、限られた人数で複数の光景を監視し続ける事は困難である。そこで近年、不審行動の1つである所持物に関する行動(置き去り・持ち去り・受け渡し)を自動的に検知する手法の研究が行われている。

Haritaogluらは人体形状の左右対称性を利用して所持物を検知するシルエットに基づく手法を提案した[1]。しかしシルエットの特徴は服装や姿勢に影響されやすく、腰を屈めた場合などに誤認識が発生する。

これに対し、Chiらは同一人物のひたくりが起きる前後の色ヒストグラムを比較し、追加・除去された色を所持物とする手法を提案した[2]。この手法では、人体や服と似ている色の所持物の誤検知を避けるために髪的位置を仮定している。しかし、姿勢や服の色は多様に変化するため、この仮定では不十分である。

本稿では、同一人物の2枚の画像から得た色と座標のデータを併合してクラスタリングを行い、クラスタ中の1枚目と2枚目の画素数を比較することで、「所持物に関する行動」と「所持物の領域」を検知する手法を提案する(図1)。2枚の画像に現フレームと数フレーム前の画像を用いることで、姿勢を仮定することなく、片方の画像のみに出現する色分布を所持物領域として検知することができる。以下、提案手法について述べ、最後に実験を通して提案手法の有効性を示す。

### 2. 色分布を利用した所持物の検知

#### 2.1 人物領域の抽出と追跡

まず、人物領域のマスク画像を生成する。本研究では実環境における雑音と影を背景とみなすために、背景モデルとして適応背景混合モデル[3]を採用している。次に入力画像に対して、Mean-Shift法[4]による前フレーム中の人物領域の矩形との対応付けを行い、識別番号(ID)を引き継いだ人物領域の矩形を得る。ここで対応付けが行われなかった一定サイズ以上の連結領域は、新たな人物領域として新しいIDを割り当てている。

#### 2.2 特徴量データの抽出と併合

特徴量データには、現在の $t$ フレーム目と過去の $t-\varepsilon$ フレーム目における同一IDの人物領域から特徴量を抽出し、これらを併合したものを用いる。ただし、所持物の受け渡し行動などによって、異なる人物領域が1つの連結領域になった場合、 $t-\varepsilon$ フレームには連結直前時のフレーム、 $t$ フレームには分離時のフレームを設定する。

特徴量抽出の前処理として、人物領域をRGBから $L^*u^*v^*$ 表色系に変換した後に平滑化処理を行う。平滑化により、ストライプ柄の服など色変化の大きい領域の過分割を避けることができる。

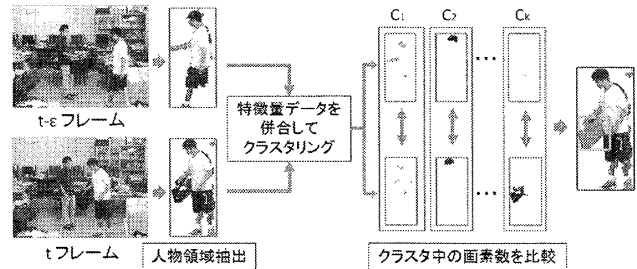


図1: 提案手法の流れ

特徴量としては以下を使用する。

#### (1) 色 $(L,u,v)$

知覚的に均等な色空間での色差を求めるために用いる。

#### (2) 座標 $(x,y)$

画素同士の連結性を反映させるために用いる。ただし、移動による座標のズレを無くすために、 $t-\varepsilon$ フレーム目の人物領域のマスク画像を $t$ フレーム目のマスク画像に対してテンプレートマッチングを行い、位置を補正した座標を特徴量として用いている。

#### 2.3 特徴量データのクラスタリング

本研究では、2枚の人物画像の対応する局所領域(髪・顔・腕など)と所持物領域の色分布を混合ガウス分布でモデル化する。混合ガウス分布の推定には、パラメータ(混合比 $\pi$ 、平均 $\mu$ 、共分散 $\Sigma$ )の初期値による過学習の問題を避けるために変分ベイズ法を用いる[5]。

変分ベイズ法では、パラメータを繰り返し推定する過程で、パラメータの事前分布と事後分布とのKLダイバージェンスが最小化されるため、正則化が自動的に行われる。そのため、混合要素数(クラスタ数) $K$ に大きな値を設定しても、どの画素も説明しない混合要素の要素数は0に近づき、過学習を避けることができる。事前分布として、混合比 $\pi$ にはディリクレ分布、平均 $\mu$ にはガウス分布、共分散 $\Sigma$ にはウィシャート分布を仮定した。

またパラメータの初期値には、色 $(L,u,v)$ のみを用いてK-means法を実行した結果を用いている。

#### 2.4 所持物領域の判定

クラスタリングによって得られた画素のクラスタそれぞれに対して、 $t$ フレーム目の画素数 $N_k^{(t)}$ と $t-\varepsilon$ フレーム目の画素数 $N_k^{(t-\varepsilon)}$ の比率を調べることで所持物の判定を行う。クラスタ $k$ の画素の集合を $C_k = C_k^{(t)} + C_k^{(t-\varepsilon)}$ 、画素数を $N_k = N_k^{(t)} + N_k^{(t-\varepsilon)}$ としたとき、フレーム間の比率を次式で求める。

$$\text{Ratio}N_k^{(t)(t-\varepsilon)} = \frac{N_k^{(t)}}{N_k^{(t-\varepsilon)} + 1}, \quad \text{Ratio}N_k^{(t-\varepsilon)(t)} = \frac{N_k^{(t-\varepsilon)}}{N_k^{(t)} + 1}$$

比率を用いた所持物の判定は次式で行う。ただしノイズ除去のために、所持物判定を行うクラスタの画素数は閾値 $N_{thld}$ 以上を満たすものとする。

<sup>†</sup>東京理科大学大学院 理工学研究科 情報科学専攻

<sup>‡</sup>東京理科大学 理工学部 情報科学科

$$C_k^{(t)} = \begin{cases} \text{所持物 (持ち去り, 受け取り)} & \text{if } \text{Ratio}N_k^{(t)(t-\varepsilon)} > R \\ \text{人体・服} & \text{otherwise} \end{cases}$$

$$C_k^{(t-\varepsilon)} = \begin{cases} \text{所持物 (置き去り, 渡す)} & \text{if } \text{Ratio}N_k^{(t-\varepsilon)(t)} > R \\ \text{人体・服} & \text{otherwise} \end{cases}$$

所持物が背景色と似ている場合、人物領域抽出の段階で所持物の画素の一部が背景に誤分類されてしまう。よって本手法では誤分類画素を補填するために、所持物と判定された画素集合の輪郭を求め、輪郭内部の画素を全て所持物とする後処理を行っている。

### 3. 実験

提案手法の有効性を検証するため実験を行った。カメラには Panasonic 製ネットワークカメラ BB-HCM581 を使用し、室内光景を毎秒 15 フレーム、解像度 320 × 240 で撮影したものを実験データとした。クラスタ数は  $K = 15$ 、フレーム間隔は  $\varepsilon = 5$ 、クラスタの最小画素数は  $N_{thld} = 80$ 、所持物と判定する比率は  $R = 10$  とした。

#### 3.1 実験結果と評価

まず、所持物領域の判定方法の有効性について述べる。図 2(a) の右人物を  $t - \varepsilon$  フレーム、図 2(b) の右人物を  $t$  フレームとしたとき、所持物の画素が殆どを占めるクラスタ  $C_9$  が生成される。このクラスタにおける比率は  $\text{Ratio}N_9^{(t)(t-\varepsilon)} = 75.6$  となり閾値  $R$  で明確に検知することができる (図 3)。

実験結果より、服と似ている色のカバンの「渡す」「受け取り」、小さい包丁の「取り出し」を検知できているため、提案手法は片方のフレームのみに存在する色分布を所持物として検知することができる (図 2 図 4)。また姿勢の異なる行動である、腰を屈めた「持ち去り」と歩行時の「置き去り」を検知できていることから、姿勢に関する仮定をすることなく所持物を検知することができる (図 5 図 6)。

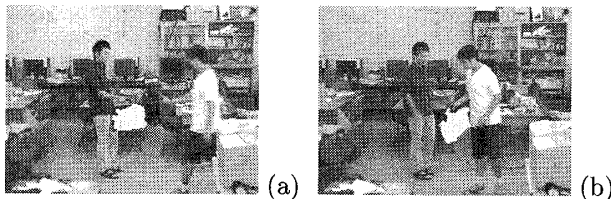


図 2: (a) 「渡す」 (b) 「受け取り」行動の検知結果

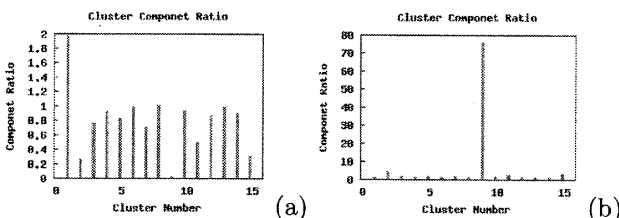


図 3: 「受け取り」行動のヒストグラム (a)  $\text{Ratio}N_k^{(t-\varepsilon)(t)}$  (b)  $\text{Ratio}N_k^{(t)(t-\varepsilon)}$

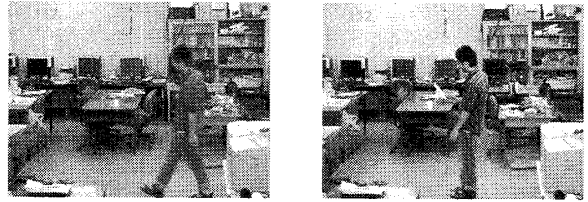


図 4: 「取り出し」行動の検知結果

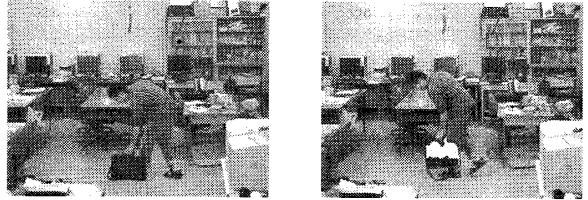


図 5: 「持ち去り」行動の検知結果

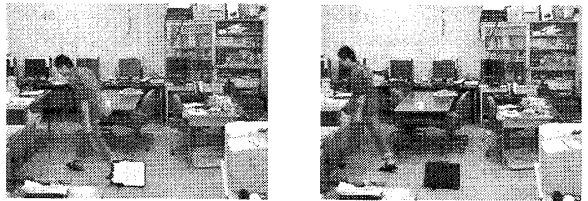


図 6: 「置き去り」行動の検知結果

所持物の誤検知は、フレーム間隔  $\varepsilon$  間に急激な姿勢や明るさの変化があった場合に発生している。これらの誤検知はクラスタ数を減らすことで減少するため、局所最適解に陥ることによる色分布の過分割が原因であると考えられる。そこで、固定個のクラスタで初期化するのではなく、1つのクラスタから開始し所持物クラスタが生成されるまで分割を繰り返すなど、局所最適解を克服するクラスタリング手法を取り入れる必要がある。

### 4. おわりに

本稿では、2枚の画像の併合した特徴量データをクラスタリングすることで、片方画像のみに存在する色分布を所持物として検知する手法を提案した。今後の課題として、急激な姿勢や照明変化に対応するために、クラスタ数の自動調節を行い色分布の過分割を避けることがあげられる。

### 参考文献

- [1] I. Haritaoglu, R. Cutler, D. Harwood, and L. Davis, "Backpack: detection of people carrying objects using silhouettes," *Computer Vision*, 1999. The Proceedings of the Seventh IEEE International Conference on, vol.1, pp.102-107, Sept. 1999.
- [2] C.H. Chuang, J.W. Hsieh, and K.C. Fan, "Suspicious object detection and robbery event analysis," *Computer Communications and Networks*, 2007. ICCCN 2007. Proceedings of 16th International Conference on, vol.1, pp.1189-1192, Aug. 2007.
- [3] P.W. Power, and J. A.Schoonees, "Understanding background mixture models for foreground segmentation," *Imaging and Vision Computing New Zealand 2002*, Nov. 2002.
- [4] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," *Computer Vision and Pattern Recognition*, 2000. Proceedings. IEEE Conference on, vol.2, pp.142-149, Sept. 2000.
- [5] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.