

1. はじめに

人の3次元位置や身長等を計測することは、映像監視システムの知能化において重要な課題である。画像から3次元情報を取得するにはキャリブレーションにより求めたカメラパラメータが必要である。しかし、カメラキャリブレーションを正確に行うには大変な労力を要する。そこで、本研究では人物領域のセグメンテーションに基づく自己カメラキャリブレーション手法を提案する。

2. 人物領域のセグメンテーション

提案手法では、単眼のカメラで撮影された映像から、画像中の物体の高さと位置、透視投影モデルにおける世界座標と画像座標の関係を用いて、カメラの自己キャリブレーションを実現する。画像中の物体の高さ情報を用いてカメラパラメータを高精度に推定するには、人物領域を正確にセグメンテーションする必要がある。本研究では、弱識別器の応答による類似シルエットの選択を用いた人物領域のセグメンテーション手法を行う[1]。この手法は、識別器の学習サンプルに対応する人体シルエット画像のスコアからハッシュテーブルを予め求めておき、入力画像をラスタスキャンして検出した人物領域をChamfer Matchingを用いてセグメンテーションを行う。図1に人物領域のセグメンテーション例を示す。

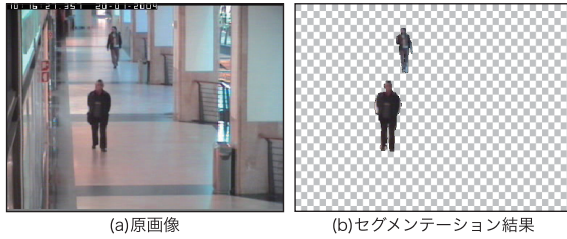


図1：人物領域のセグメンテーション例

3. カメラの自己キャリブレーション

人物領域のセグメンテーション結果を用いてカメラの自己キャリブレーションを行う。文献[2]を参考に、透視投影モデルにおけるカメラ位置と人の高さとの関係を定式化し、RANSACによりカメラパラメータの推定を行う。

3.1 カメラパラメータの導出

図2にカメラ位置と人の高さの関係を示す。

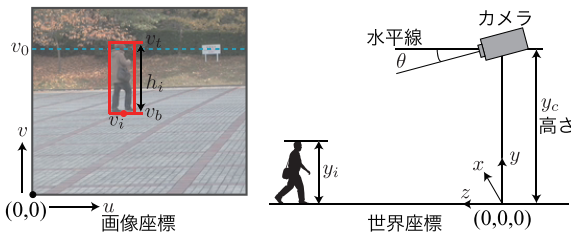


図2：カメラの位置と人の高さの関係

画像座標系を (u, v) 、世界座標系を (x, y, z) 、カメラのチルト角を θ 、焦点距離を f 、カメラ中心を (u_c, v_c) 、カメラの高さを y_c 、画像座標系の地面の消失ラインとして水平線を v_0 と定義する。このとき、カメラのチルト角（ラジアン）は次式で与えられる。

$$\theta = 2\arctan \frac{v_c - v_0}{2f} \quad (1)$$

カメラモデルとして透視投影モデルを使用すると、世界座標系から画像座標系への変換は以下の式で与えられる。

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f & 0 & u_c \\ 0 & f & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \\ y_c \end{bmatrix} \quad (2)$$

世界座標系は、カメラ位置を中心として、 $z_c = 0, x_c = 0$ 、接地平面を $y = 0$ と定義する。式(2)を変形し、画像中の n 人のサンプルとカメラパラメータの関係を行列で表現すると次式となる。

$$\begin{bmatrix} v_0 \\ y_c \end{bmatrix} = \begin{bmatrix} y_i & -h_i \\ \vdots & \vdots \\ y_n & -h_n \end{bmatrix}^{-1} \begin{bmatrix} y_i v_i \\ \vdots \\ y_n v_n \end{bmatrix} \quad (3)$$

式(3)より、画像座標での人の高さ h_i と位置 v_i 、世界座標系での人の高さ y_i の3つのパラメータが二組以上与えられると、カメラの外部パラメータであるカメラの高さ y_c と水平線の位置 v_0 を求めることができる。

3.2 RANSACによるカメラパラメータ推定

人物領域のセグメンテーションにより、画像座標での人の高さ h_i と位置 v_i を求めることができる。式(3)において、カメラパラメータの推定には世界座標系での人の高さ y_i を求める必要がある。しかし、画像から世界座標系の人の高さの真値を知ることは不可能である。そこで、提案手法では人の身長として、男性20-59歳の身長測定の結果を用いて、平均身長170.5[cm]、標準偏差5.3[3]の正規乱数により人の高さ y_i を与え、RANSAC(RANdam SAmple Consensus)を用いてカメラパラメータの推定を行う。RANSACを用いたカメラパラメータ推定の流れを以下に示す。

Step1 人検出により人物の候補領域を検出

Step2 類似シルエット画像の選択

Step3 人物領域のセグメンテーション

Step4 正規乱数を用いて全サンプルの y_i を仮定

Step5 ランダムに2点のサンプルを選択

Step6 最小二乗法により y_c と v_0 の推定

Step7 推定値から全サンプルとの誤差を算出

Step8 誤差の中間値を選択

Step9 Step4～9を繰り返し、誤差の中間値が最小となるパラメータを求め、最終的なカメラの高さ y_c と水平線位置 v_0 とする

4. シミュレーションによる評価実験

提案手法の精度をシミュレーション実験により評価する。実験には人体シルエットをOpenGLで描画したシミュレーションモデルを用いる。実験に用いたシミュレーションモデルの環境はカメラの高さ y_c は3m、生成する人の位置 (x, z) は乱数によって決定した。

4.1 チルト角による推定精度の評価

シミュレーションモデルを用いてカメラチルト角 θ を0～40度まで5度刻みで変化させた時のカメラパラメータ推定精度を比較する。各角度でカメラパラメータ推定に用いた人モデルのサンプル数は600とした。図3に生成したサンプルの例を示す。

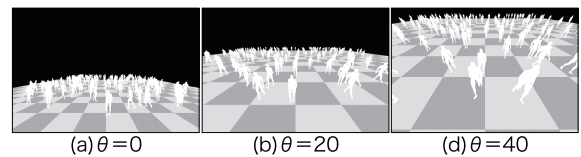


図3：生成したサンプルの例1

チルト角による推定精度の比較結果を図5(a)に示す。結果から、チルト角が40度のとき推定誤差は約11.0[degree]以内で推定できる。また、推定するシーンのチルト角が大きくなるにつれ推定誤差が増加することがわかる。これは、カメラパラメータの導出の過程でカメラのチルト角が小さいと仮定しているため、チルト角が大きくなるほど推定する際に誤差が生じると考えられる。

4.2 サンプルの分布による推定精度の評価

次に、提案手法で用いる人の身長分布とサンプルの分布の違いによるカメラパラメータの推定精度の比較を行う。カメラパラメータの推定に用いるサンプルに含まれる成人男性と子供のサンプルの割合を変化させた時の推定精度を比較する。成人男性のサンプルの割合を0%～100%まで10%毎に変化させ評価を行う。カメラのチルト角は0度、子供のサンプルは男子6歳の身長測定の結果である平均身長116.7[cm]、標準偏差4.8[3]の正規乱数によって高さを与えて生成した。図4にシミュレーションモデルにより生成したサンプルの例を示す。サンプルの分布による推定精度の比較結果を図5(b)に示す。

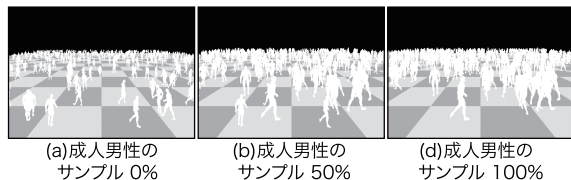


図4：生成したサンプルの例2

RANSACを用いて全サンプルからカメラパラメータ推定に最適なサンプルの組み合わせを選択して推定するため、成人男性の人モデルに当てはまらない身長のサンプル（子供のサンプル等）がノイズとして含まれているシーンに対しても、仮定した成人男性のモデルが60%以上含まれる場合、RANSACを用いることでカメラの高さを高精度に推定できる。

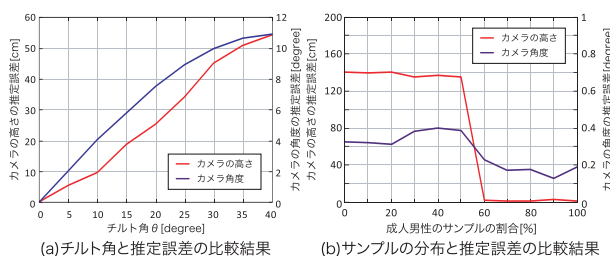


図5：シミュレーション実験の結果

5. 実画像による評価実験

提案手法の有効性を確認するために、実画像でのカメラパラメータと3次元位置の推定の評価実験を行う。

5.1 カメラパラメータ推定

撮影した映像からカメラの自己キャリブレーションを行い、カメラパラメータの推定精度の比較を行う。求めるパラメータはカメラの高さ y_c と水平線位置 v_0 で、水平線位置 v_0 からカメラの角度 θ を推定する。カメラパラメータの推定に用いたサンプル数を600である。図6にカメラパラメータの推定に用いたサンプルの人検出結果の分布を示す。表1にパラメータ推定結果を示す。

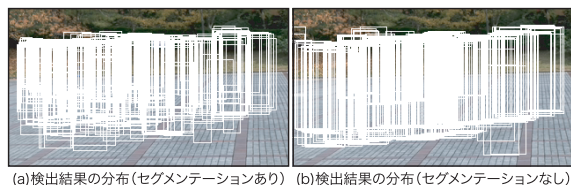


図6：人検出結果の分布

表1：カメラパラメータ推定結果

カメラパラメータ	真値	セグメンテーション	
		あり	なし
カメラの高さ y_c [cm]	184.6	186.7	205.2
カメラ角度 θ [degree]	10.0	13.4	16.9

表1より、セグメンテーションありの場合にカメラ高さが真値との誤差が約2.1[cm]、カメラの角度が真値との誤差は約3.4[degree]の精度で推定できた。

5.2 世界座標における位置推定

推定したカメラパラメータを用いて、検出した人の世界座標における3次元位置を推定をする。評価実験では、セグメンテーションあり、なしそれぞれのキャリブレーション結果を用いて、3次元位置の推定精度を評価する。推定精度の評価には15点のランドマーク点を用いて、世界座標の真値と3次元位置の精度結果の比較を行う。また、世界座標における位置の推定では地面の高さを $y=0$ として仮定している。図7に世界座標における位置の推定誤差の結果、図8に提案手法による3次元位置の推定例を示す。

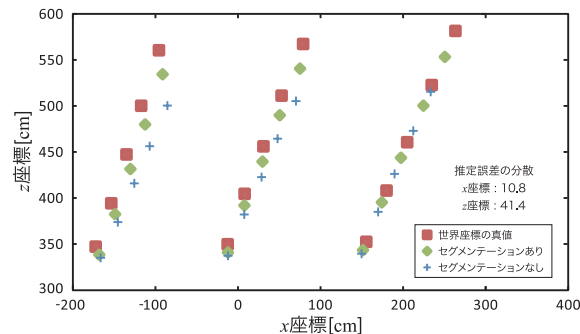


図7：世界座標における位置の推定誤差

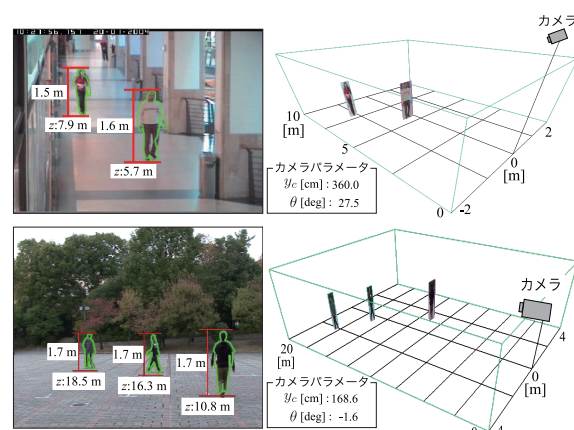


図8：3次元位置の推定結果例

図7より、セグメンテーション結果を用いてカメラパラメータを推定することで、世界座標における位置の推定誤差を小さくすることができた。図8より、カメラの自己キャリブレーション結果を用いることで、監視映像における検出した人の位置関係や身長を獲得することが可能となる。

6. おわりに

本研究では、人検出とセグメンテーション結果に基づく、カメラの自己キャリブレーション手法を提案した。セグメンテーションの結果から画像中の人の高さや、位置を自動的に抽出することで自己キャリブレーションを行い、カメラパラメータの推定を行った。これにより推定シーンにおける物体の3次元位置推定が可能となり、評価実験によりその有効性を確認した。

参考文献

- [1] 村井 陽介, 安藤寛哲, 藤吉弘亘, 伊藤誠也, 数井誠人, “弱識別器の応答に基づく類似シルエット画像選択による Chamfer Matching を用いた人領域のセグメンテーション”, 電子情報通信学会論文誌 Vol. J94-D, No. 4, 2011. (掲載予定)
- [2] D. Hoiem *et al.*, “Putting Objects in Perspective”, IJCV, vol. 80, No. 1, 2008.
- [3] 平成20年度体力・運動能力調査調査結果統計表：文部科学省, “http://www.mext.go.jp/b_menu/houdou/21/10/attach/1285568.htm”.

研究業績

- [1] 安藤寛哲, 藤吉弘亘, “人検出結果に基づく自己カメラキャリブレーションと3次元位置推定”, 電気学会論文誌 D Vol. 131, No. 4, 2011. (掲載予定)

(他 学会誌論文1件, 国際会議論文1件, 研究会技術報告2件, 学会口頭発表1件)