

複数台カメラを用いたグラウンド利用者カウントシステムの開発

Development of the multi camera human detection system for counting users in playground

廣田 隼[†] 服部 公央亮[†] 田口 亮[†] 保黒 政大[‡] 梅崎 太造[†]

Shun Hirota[†], Koosuke Hattori[†], Ryo Taguchi[†], Masahiro Hoguro[‡] and Taizo Umezaki[†]

名古屋工業大学(Nagoya Institute of Technology)[†] 中部大学(Chubu University)[‡]

1. はじめに

近年、防犯カメラは学校や街頭など至る所に設置されている。しかし、多くのカメラシステムは録画機能のみを持ち、人数カウントなどの機能は有していない。そのため、マーケティングなどにおいては、撮像した動画像をもとに手動で人数のカウントを行う必要がある。また、防犯目的としている場合、常に人が撮影画像に対して注視しなければならない。

本研究室では人数計測を自動で行うグラウンド利用者カウントシステムの開発を進めている。本システムはグラウンドに設置したカメラを使用して人を検出し、利用者のカウントを行うことを目的としている。1台のカメラでは死角になる領域が多いため、複数台のカメラを用いて検出を行う。本稿では人検出のための動体検出法に関して述べる。

2. 計測環境

本システムでは名古屋工業大学所有のグラウンドを計測環境とし、カメラを18台設置して撮影する。カメラのフレームレートは3 [fps]である。最終的な目標は人の追跡結果をカメラ間で統合し、利用者をカウントすることである。図1にカメラの設置環境の概観を示す。図1に示すように275 [m]×125 [m]の面積を持つグラウンドに6本のポールを設置し、それぞれのポールにカメラを3台ずつ設置してグラウンド全体を部分ごとに撮影する。

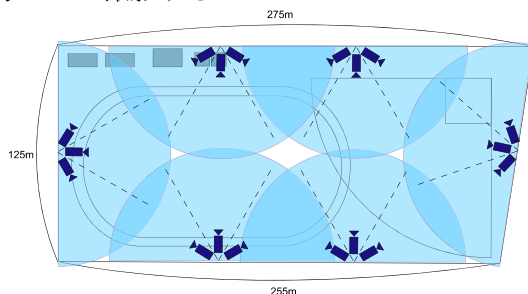


図1 グラウンドの概略図

3. 動体検出

画像中の人検出の代表的な手法として、多数生成した検出窓中のHOG特徴量を導出し、検出する手法が挙げられる。しかし、この手法は処理量が多く、18台のカメラで検出をリアルタイムに行うことは困難である。そこで、本システムではより処理量の少ないブロックベースのCodebookによる動体検出[1]を使用する。

Codebook アルゴリズムはCodewordと呼ばれる背景情報を学習期間に記録し、新たな入力画像とそれを比較することで、動体検出を行う手法である。処理量を減らすため画像全体を20×20 [pixel]のブロック単位で検出する。背景情報にはブロック内の輝度値の平均値を用いる。ただし、ブロック内に明るい物体、暗い物体が混在していると、ブロックを一つの平均値だけで表現できない。そのため、[1]ではブロック内で四つの平均値 $\mu_{ht}, \mu_{hb}, \mu_{lt}, \mu_{lb}$ を導出し、背景情報としている。ブロックの平均値を μ としたとき、それより大きい値の平均値を μ_h とする。 μ_h より大きい値の平均値が μ_{ht} 、 μ より大きいかつ μ_h より小さい平均値が μ_{hb} である。 μ_{lt}, μ_{lb} も μ より小さい値の平均値を μ_l とし、同様に導出される。

それら平均を並べたベクトルを r とする。入力画像中のブロックを r_{src} 、背景情報を r_{ref} とし、その差分を以下のように計算する。

$$d = r_{src} - r_{ref} \quad (1)$$

その後、match関数により、背景か動体かが判定される。

$$match(r_{src}, r_{ref}) = \begin{cases} \text{背景, if } \frac{d^T d}{\dim(d)} < \lambda^2 \\ \text{動体, O.W.} \end{cases} \quad (2)$$

なお、 λ^2 はしきい値、 $\dim(d)$ は次元数を表す。

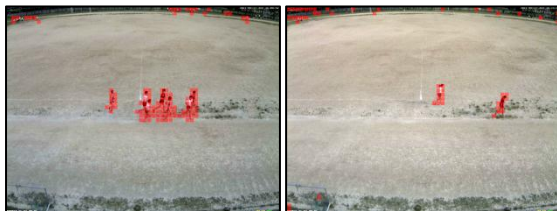
しかし、この手法は明るさの変動や、影の影響を受けやすいため、屋外で使用する本システムには不向きである。そこで、それらの影響を軽減するため、式(3)、(4)で輝度値の正規化

を行う.

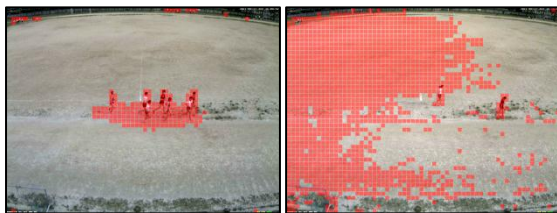
$$r' = \frac{r}{r+g+b+c} \quad (3), \quad g' = \frac{g}{r+g+b+c} \quad (4)$$

ただし, r, g, b は各色の輝度値, c は定数.

正規化した場合, 正規化しない場合の動体検出例を図 2 に示す. 影の影響, 明るさの変動による影響を軽減することを確認した. また, 図 2 の左の示すように, 複数人がランニングしている動画に対しての, 動体検出精度を図 3 に示す. なお, 検出精度は tp を True Positive 数, fp を False Positive 数としたとき, $tp/(tp+fp)$ で計算される. 非正規化の場合の平均精度 67.2 [%] が正規化することにより 92.9 [%] となり, 検出精度が向上することを確認した.



(a) 提案手法 (正規化)



(b) 従来手法 (非正規化)

図 2 動体検出結果 (赤: 動体)

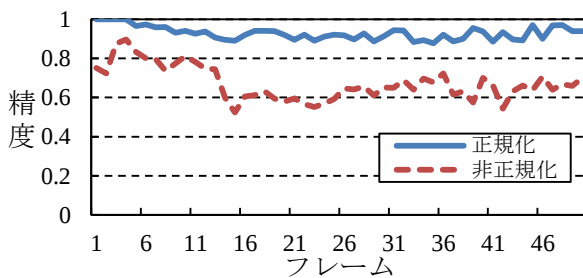


図 3 検出精度

4. カメラ画像と衛星写真の対応付け

本システムは複数台のカメラを使用しているため, 同一人物に対する検出結果をカメラ間で統合する必要がある. そのためには, 各カメラ画像間でマッチングを行い, 重複箇所を特定することが考えられる. しかし, 18 台のカメラ間でマッチングを行うと組み合わせ方が非常に多くなり扱いにくい. そこで, 本システムでは各カメラ画像とグラウンドの衛星写真 (Google Map) との対応付けを行う.

カメラ画像と衛星写真の対応付けはベンチなど固定物体を元にして手動で求める. 対応関係を線で図示したものが図 4 である. しかし, 画像によっては固定物体があまり映り込んでいない場合がある. そのような場合には GPS を併用することで対応関係を求める.

対応関係を求めた後, 射影変換行列を導出[2]する. 図 5 は図 4 のカメラ画像を射影変換した結果であり, カメラ画像中の検出された人がグラウンドのどの位置にいたかが特定可能となる. 複数カメラで同一の人が検出された場合, ほぼ同じ位置に人がいると判定されるため, 統合することが可能である.

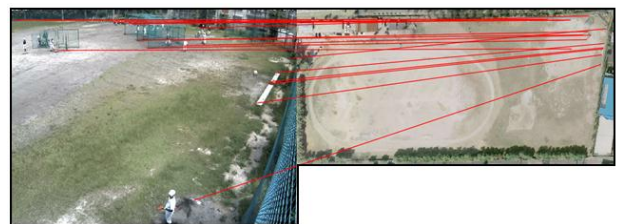


図 4 カメラ画像と衛星写真の対応付け



図 5 射影変換結果 (点線内)

5. まとめ

本稿では人検出を目的とした動体検出法の検討を行った. 従来法による動体検出の平均精度 67.2 [%] が輝度値の正規化を行うことで 92.9 [%] に向上することを確認した. しかし, 動体検出のみでは木の揺れなどに対し, 誤検出が発生する. 今後, 人かどうかの判定処理を組み込むことで人検出精度の向上を目指す.

参考文献

- [1] J.-ming Guo, S. Member, Y.-fu Liu, S. Member, and C.-hsien Hsia, "Hierarchical Method for Foreground Detection Using Codebook Model," IEEE Transactions on Circuits and Systems, vol. 21, no. 6, pp. 804-815, 2011.
- [2] H. Niitsuma, P. Rangarajan and K. Kanatani, "High accuracy homography computation without iterations", 第 16 回画像センシングシンポジウム (SSII 2010) 講演論文集, 2010.