

身体の動作軌跡のクラスタ化に基づく特徴量を用いた バレーボールの動作検出

久保田 栄次郎 池永研究室 修士課程終了

■ スポーツ競技

選手の動き情報をデータベース化



人間が目視でデータ取得



戦略立案

要望

カメラ動画からの自動取得

■ 従来研究

動画からの動作検出
標準的フレームワーク

特徴量
抽出

BoF
Bag of
Features

SVM
Support
Vector
Machine

特徴量: 特徴点を追跡した軌跡

■ Dense Trajectories

高密度な特徴量



- ① 密なオプティカルフロー計算
特徴点(例: 目)
- ② 軌跡追跡 軌跡長=15
オプティカルフロー変位
- ③ 軌跡近傍特徴量取得
HOG (Histograms of Oriented)
人物などのおぼろげな輪郭の特徴量
HOF (Histograms of Optical Flow)
物体の動きを表した特徴量
MBH (Motion Boundary Histogram)
フローの変化部分を表した特徴量
オプティカルフローにsobelフィルター(縦・横)を適用したもの

■ 特徴量 (1軌跡あたり426次元)

軌跡(相対値)	HOG	HOF	MBHx	MBHy
30次元	96次元	108次元	96次元	96次元

高密度な特徴量 → 認識に重要な特徴量を捉える可能性が高まる → 検出性能向上

■ 問題点

高密度な特徴量 → 動画の場合さらに特徴量増大 → 学習器適用困難

ランダムサンプリングで特徴量削減

しかし、動作認識に重要な特徴量が間引かれる可能性

動作認識に重要でない特徴量(ノイズ)を多くサンプリングする可能性

■ 提案手法

選手の身体動作軌跡に基づき特徴量集約
全特徴量

クラスタ化の
ための
絶対値軌跡

絶対値軌跡

腕や足などの身体の特徴点が描く軌跡は、
部位ごとに似ていることに着目

軌跡
クラスタ化

密度
クラスタ化

ノイズ除去

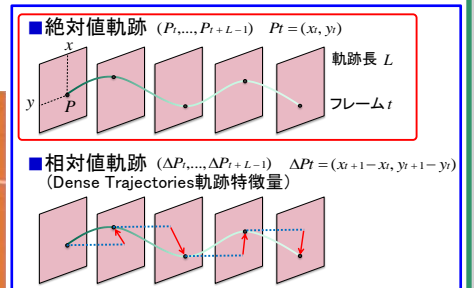
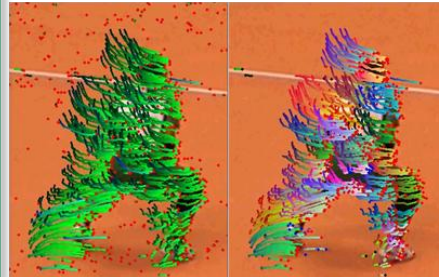
特徴量集約
(平均・分散)

提案手法の処理

提案手法の処理

1. 軌跡クラスタ化

絶対値軌跡を基にk-meansでクラスタ化



k-means アルゴリズム

$$\sum_{i=1}^k \sum_{Trajectories \in c_i} (Trajectories_j - centroid_i)^2$$

k : クラスタ数

$centroid$: クラスタ重心

$Trajectories$: 軌跡の特徴量

式の値が最小化するように処理

2. 密度クラスタ化

軌跡クラスタを

軌跡の最新座標(絶対値)を基にさらにクラスタ化



密度クラスタ1

密度クラスタ2

クラスタ不成立
(ノイズ除去)

最新
クラスタ化 DBSCAN アルゴリズム

$$x \in N_{Eps}(nx)$$

$$|N_{Eps}(nx)| \geq MinPts$$

Eps : ($Eps=10$) クラスタ成立可能距離

$MinPts$: ($MinPts=3$) クラスタ成立可能な

最低の座標(点)数

Nx : 座標(点)

$N_{Eps}(nx)$: 距離 Eps 内の座標(点)の集合

条件を満たす点
(例: 距離10以内に点3個)
を次々に結合しクラスタ化

3. 特徴量集約

密度クラスタ単位で特徴量を集約

平均 密度クラスタ内の全特徴量の平均を特徴量として集約

分散 平均では失われる密度クラスタ内の全特徴量のばらつきを
特徴量として集約

クラスタ内の 特徴量	軌跡(相対値)	HOG	HOF	MBHx	MBHy	426次元 × n (例: n=4)
	軌跡(相対値)	HOG	HOF	MBHx	MBHy	
	軌跡(相対値)	HOG	HOF	MBHx	MBHy	
	軌跡(相対値)	HOG	HOF	MBHx	MBHy	
	軌跡(相対値)	HOG	HOF	MBHx	MBHy	

各次元(426次元)
ごとに平均・分散を
計算

$$\text{平均 } \bar{x}_i = \frac{1}{N} \sum_{j=1}^N x_{ij}$$

$$\text{分散 } \sigma^2 = \frac{1}{N} \sum_{j=1}^N (x_{ij} - \bar{x}_i)^2$$

平均	軌跡(相対値)	HOG	HOF	MBHx	MBHy	426次元
分散	軌跡(相対値)	HOG	HOF	MBHx	MBHy	× 2

実験結果

■ 評価環境

使用動画

平成26年度全国高等学校総合体育大会

男子バレーボール競技大会

コート右斜め後ろから動作の画像を切り出し

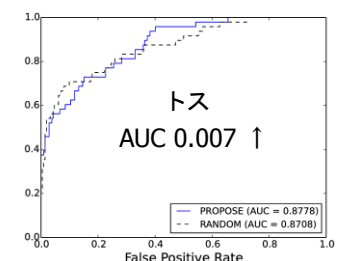
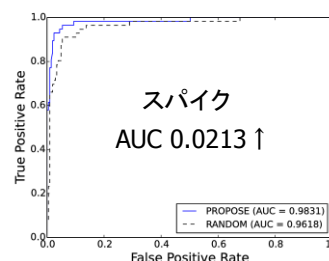
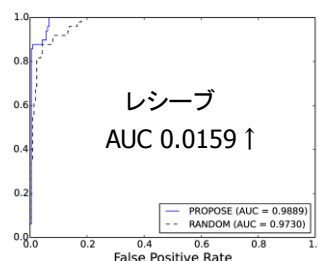
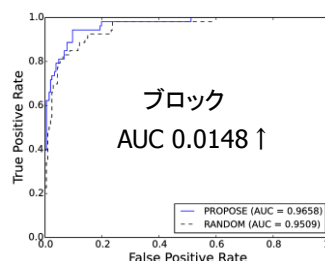
切り出し動画 640×480pixel 60fps

学習動作 「ブロック」「レシーブ」「スパイク」

「トス」「その他」(各10動画)

検出動作 「ブロック」「レシーブ」「スパイク」「トス」

(各約50動画)



ROC (Receiver operating characteristic) グラフ

縦軸: 真陽性率 (True Positive Rate) = $\frac{TP}{TP + FN}$

横軸: 偽陽性率 (False Positive Rate) = $\frac{FP}{FP + TN}$

グラフ下側面積 AUC (Area under ROC curve) の大きさを
性能を評価。

■ 結論

「ブロック」「レシーブ」「スパイク」提案手法優勢
(AUC平均0.017367向上)

「トス」提案手法とランダムサンプリング間で明確な差異なし
(AUCは提案手法がわずかに優勢)



早稲田大学大学院情報生産システム研究科