

2者間における人の暴力検出の フレームワーク構築

黒羽子 純平[†] ティティズイン
[†] 宮崎大学工学部電気電子工学科

1. はじめに

近年、駅構内で1日に約2.36件の暴力事故が発生し、セキュリティカメラが多く設置されるようになった[1].本稿では、暴力検出の足掛けとして2者間における人の暴力検出のフレームワーク構築について検討する.

2. フレームワーク構築

フレームワーク構築とは、各シーケンスのフレーム間での行動要素の識別を目的とし、人物の入った動画像からフレーム毎に背景差分やノイズ除去等の前処理を行い、得られたシルエットから形状や状態から特徴量を計算することで、行動要素を分類できる. 以下に、人物の行動要素として次の5つを定義する(Punching, Pushing, Kicking, Handshaking, Hugging).

これらのシルエットを得られた特徴量から生成されるアルゴリズムによって分類する. また、これらの行動の予知検出を行い、その後、行為中における行動要素を正確に検出するために可能性として「Violence」と「Non-Violence」の2つに分類することで今後の行動検出の足掛けに繋がる. その結果、適切な分類結果を得ることができる.

3. 特徴検出

3.1 頭の重心の距離 まず、図1に示す頭の重心を用いて、各フレームにおける2者間での距離変化を測る. ここでは、ユークリッド距離を用いて式(1)から求められる.

3.2 フレーム変化による分散 図2に示す様に各行動要素においてフレーム間での変化をシルエットの分散によって算出していく. 暴力行為では、アクティブ側の変化が激しく2者間での分散変化は大きい、非暴力行為では両者の類似動作によって2者間での分散変化は小さい. ここで、式(2)より求められる.

$$V = \frac{(x_1 - m)^2 + (x_2 - m)^2 + \dots + (x_n - m)^2}{N} \quad (1)$$

ここで、 x_1, x_2, \dots, x_n : 平均値, N : 数, V : 分散 である.

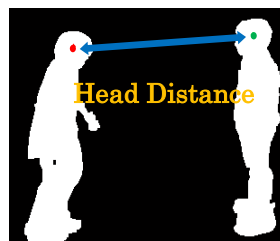


図1 頭の重心の距離

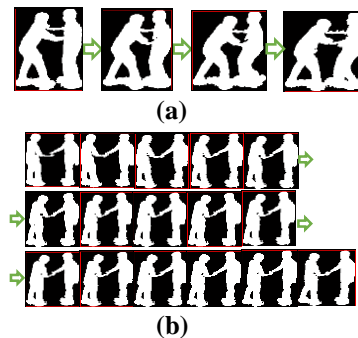


図2 行為中の frame 枚数:

(a)Pushing, (b)Handshaking

4. 5つの行動要素における分類結果

本稿では、各データセットにおいて定義した5つの行動要素での各フレームにおいて「暴力」と「非暴力」の可能性の分類について検討する. 2者間の行動解析を行い、各 frame に依存しない様におよそ10~30frameにおいて実験を行った. これらのシーンの行動を時系列で分類し、正しい結果が得られたかを確認する. なお、可能性が低いと分類された場合を誤りとしている. その結果を次の表1に示す.

表1 各行動における暴力・非暴力の発生確率の可能性
(a)Video1 (b)Video2

Action	Number frame	Error frame	Accuracy Rate[%]	Action	Number frame	Error frame	Accuracy Rate[%]
Punching	15	5	75	Punching	14	2	88
Pushing	16	3	84	Pushing	15	5	75
Kicking	18	3	86	Kicking	16	1	94
Handshaking	35	5	88	Handshaking	24	0	100
Hugging	36	2	95	Hugging	26	7	79
Total	120	18	85	Total	95	15	87

各行動要素においてエラー検出にばらつきが見受けられるが、これに関しては行為前の直立(Standing)のフレーム数が多少多く、誤って低確率であると認識されたと考えられる.

5. 今後の課題

今後の課題としては、より多くのデータを用いて人物の行動要素を正確に分類し、精度向上を目指す.

参考文献

[1]東京都交通局
http://www.kotsu.metro.tokyo.jp/newsevent/news/subway/2014/sub_p_201407074672_h.html<2015/1/29 Access>.