

ノープライバシーセンサアレイによるドア操作の観測

小野 堯[†]
山梨大学[†]

豊浦 正広[‡]
山梨大学[‡]

1 はじめに

人の行動を認識する技術は、スマートホームなどへの応用が期待されている。プライバシーの保護を目的として、フィルタ処理したカメラ映像や低画質画像によって行動が認識されることがある。Ugur らは、防犯カメラから得られる時系列低解像度画像から特徴量を抽出して行動認識を行う手法を提案した [1]。しかし、カメラ自体は使用していることから、撮影される側はプライバシーが十分に保護されていないと考えることがある。カメラの代わりに人感センサと呼ばれる照度センサを使用すると、ノープライバシーが実現されているとみなされ、社会的に受け入れられる場面は多い。一方で、照度センサを用いる場合には、人の行動認識のような高度な処理は考えられてこなかった。

プライバシー保護のためにはセンサ数は少ない方がよいが、逆に少なすぎると十分な認識ができない。本研究では、カメラで撮影した画像の解像度を極端に落とすことによって、照度センサアレイを模擬したデータセットを作成して、ネットワーク学習による人の行動認識の可能性を検証する。

2 ドアの開閉と人の有無の認識

2.1 データセット作成

図 1 にカメラ画像を加工して、照度センサアレイの模擬データセットを作成する方法を

示す。ドアを上方から観測したようなカメラ映像を撮影し、解像度を落とし、さらにグレースケールに変換する。各画素値として得られる値が 1 つの照度センサの出力とみなすことができ、これがアレイ上に並ぶことになる。

元の映像からは容易に任意のラベルを付与することができる。本研究では、「人が映っている」、「ドアを操作している」の 2 点について各フレームで判断し、ラベルを付与した。これによって、照度センサアレイのデータと、それに対応する 2 種類のラベルを持つデータセットが生成できたことになる。

上述の人の行動認識を少ない数のセンサの出力から推定することは難しいように思われたので、時系列信号の変化を表現するために直前のフレーム画像との差分を持つ画像を作成し、元の信号と合わせてセンサから得られるデータとした。

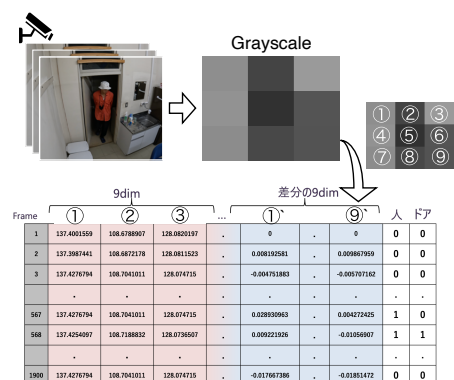


図 1 センサアレイデータの作成

2.2 認識ネットワーク

認識には Transformer を含むマルチタスク学習を行う認識ネットワークを使用した。図 2 に本研究で用いた認識ネットワークを示す。入力

Observation of door operation with no-privacy sensor array

[†] Yutaka Ono, University of Yamanashi

[‡] Masahiro Toyoura, University of Yamanashi

層は任意次元のセンサアレイデータを入力とする．解像度を $N \times N$ とするときには，直前のフレームとの差分を示す画像と合わせて $2N^2$ 次元の入力データが与えられることになる．出力層は，「人が映っているか」，「ドアの開閉動作」と異なるタスクを持つマルチタスク学習によって，同時に推定を行う．

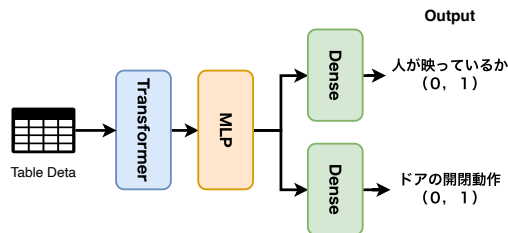


図2 センサアレイデータからの行動認識

3 実験

実験にあたって2つの部屋のドアを3方向からそれぞれ撮影し，計6本の映像を撮影した．被験者にはドアの前を通過する動作と，部屋への入退室の動作を行ってもらった．訓練には動画4本（総フレーム数:9201），テストには動画2本（総フレーム数:2623）を使用した．評価では，次元数ごとに認識ネットワークを訓練して，テストデータを入力するときの AUC を算出して比較を行った．

入力次元数を変えたときの AUC の結果を図3に示す．人の認識は 11×11 次元で 0.846，ドアの認識は 5×5 次元で 0.712 で最大であった．単一のセンサ (1×1 次元) の結果を比較すると，入力次元を増やすことで精度が向上する結果を得られた．一方，ドアの AUC は人の認識結果と比べて全体的に低い．0.5 に近い値であることは，認識器がランダムに推定を行っていることを示す．

ドア座標は各映像ごとに異なるため，次元ごとの値の持つ意味が変わるために，認識難易度が高く精度が低いことが考えられた．そこでさらに，ドアを基準とした座標に変換するホモグラフィを求めて，ドア座標を映像間で一致させ

る前処理を加えた．結果を図4に示す．座標変換前と比較し，ドア開閉動作に対する AUC は 5×5 次元を除く 3×3 次元から 16×16 次元の5項目で平均 17.8% 増加した．一方，人に対する認識は低下した．

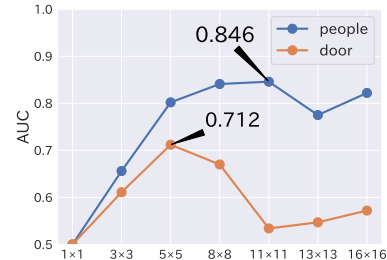


図3 各次元における AUC

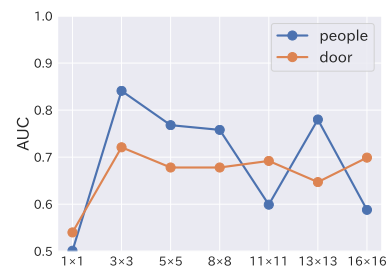


図4 ホモグラフィ変換後の AUC

4 おわりに

本研究では，カメラで撮影した画像の解像度を極端に落とすことによって，照度センサアレイを模擬したデータセットを作成して，ネットワーク学習による人の行動認識の可能性を検証した．単一のセンサを入力する場合と比較して，次元数を増やすことで人の認識精度が高くなることが確認できた．また，ホモグラフィによるドア基準座標系への変換が，ドア操作の認識に対し有効であることも確認できた．今後の課題は，人とドア操作の認識精度を同時に向上させることである．

参考文献

- [1] Demir, U., Rawat, Y. S. and Shah, M.: Tinyvirat: Low-resolution video action recognition, *Proc. ICPR*, pp. 7387–7394 (2021).