

[共通セッション] 土木分野におけるAIの活用

土木分野における AIの活用 (4)

2023年9月14日(木) 15:10 ~ 16:30 CS-1 (広島工業大 五日市キャンパス三宅の森Nexus21 701 / 広島大 東広島キャンパス工学部講義棟 B218)

[CS14-23] マルチカメラでの同一人物照合に向けた基礎的検証 Verification aimed at identifying same person with multi-camera

*吉田 龍人¹、大久保 順一¹、高森 秀司¹、天方 匡純¹ (1. 八千代エンジニアリング株式会社)

*Ryuto Yoshida¹, Junichi Okubo¹, Shuji Takamori¹, Masazumi Amakata¹ (1. Yachiyo Engineering Co., Ltd.)

キーワード：深層学習、距離学習、人流解析、ArcFace

Deep learning, Metric Learning, People Flow Analysis, ArcFace

動画を対象とする人流解析は今後の活用が期待される技術であるが、複数カメラ間での同一人物の照合が課題となっている。同一人物は DNNの出力する特徴量の類似度によって照合することが一般的となっているが、カメラとの距離など対象人物の写り方の変化が算出される類似度に与える影響は明らかではない。

本研究では任意の距離に立つあらゆる方向を向いた同一人物を固定カメラで撮影し、画像認識での同一人物照合に影響する要素を評価するためのデータセットを作成する。さらに本データセットに対する DNNモデルの出力によって画像間の類似度を算出し、距離や向きの変化が類似度算出結果に与える影響を明らかにする。

マルチカメラでの同一人物照合に向けた基礎的検証

八千代エンジニアリング株式会社 正会員 ○吉田 龍人 高森 秀司 天方 匡純
非会員 大久保 順一

1. はじめに

動画を対象とする人流解析は今後の活用が期待される技術であるが、複数カメラ間での同一人物の照合が課題となっている。同一人物は DNN の出力する特徴量の類似度によって照合することが一般的となっている¹⁾が、カメラとの距離など対象人物の写り方の変化が算出される類似度に与える影響は明らかではない。

本研究では任意の距離に立つあらゆる方向を向いた同一人物を固定カメラで撮影し、画像認識での同一人物照合に影響する要素を評価するためのデータセットを作成する。さらに本データセットに対する DNN モデルの出力によって画像間の類似度を算出し、距離や向きの変化が類似度算出結果に与える影響を明らかにする。

2. 関連研究

複数カメラに写った同一人物の照合は Re-Identification¹⁾と呼ばれる画像認識タスクで様々な手法が検討されている。Re-Identification では距離学習で構築した DNN モデルを適用することが一般的となっており、各画像の推論結果である特徴量ベクトルの類似性によって同一人物を照合する。Re-Identification で使用されるモデルを構築する教師あり距離学習は、対照学習を行うアプローチと画像分類を行うアプローチに大別される。前者は入力データの特徴量ベクトルが同一人物であれば近く、非同人物であれば遠くなるように学習させる手法である。後者はデータセットの各人物を分類学習する初歩的手法であるが、角度距離を導入した Softmax Loss を損失関数に用いることで大きな精度向上を果たした。距離学習は学習時の入力データや損失関数の扱いに創意があり、推論に用いるモデル自体は ResNet など一般的なものである。本研究では ArcFace²⁾で示される Softmax Loss によって学習した ResNet50 を用いて、同一人物の照合を行い、その性能を評価する。ResNet50 には既往研究³⁾で提示した手法によって構築した学習済みモデルを用いる。

3. 実験方法

類似度に影響する要素を明らかにするため、同一人物を図-1 に示す様々なパターンで撮影した。カメラと人の水平距離は 2m 刻みで 2m~14m の計 7 パターンを撮影した。カメラに対する向きはカメラ正対を 0°として、45°刻みで計 8 パターンを撮影した。更に、見た目の違いによる類似度算出結果の違いを評価するため、表-1 に示す計 5 パターンを撮影した。立ち姿は全て直立である。

撮影した 7×8×5=280 枚の各画像を推論してそれぞれの特徴量ベクトルを取得し、各ベクトル間のコサイン類似度を計測した。入力画像は人の写る箇所のみを切り出し、元の画像の縦横比を維持するよう 128×384px にリサイズした。リサイズ時は幅が 128px、高さが 384px 以下という基準を満たしつつ、面積が最大となるよう画像の縦横を等倍し、不足する箇所はグレーで補った。なお図-1 では Gaussian Blur を実施した画像を示しているが、実験では Blur を実施していない元の画像を用いた。

表-1 見た目のパターン

条件	概要
①	ワイシャツ (ネクタイ無)
②	ワイシャツ (ネクタイ無, 袖まくり)
③	ワイシャツ (ネクタイ有) &手荷物 (右手持ち)
④	ジャケット (ネクタイ無)
⑤	作業着 (ネクタイ無)

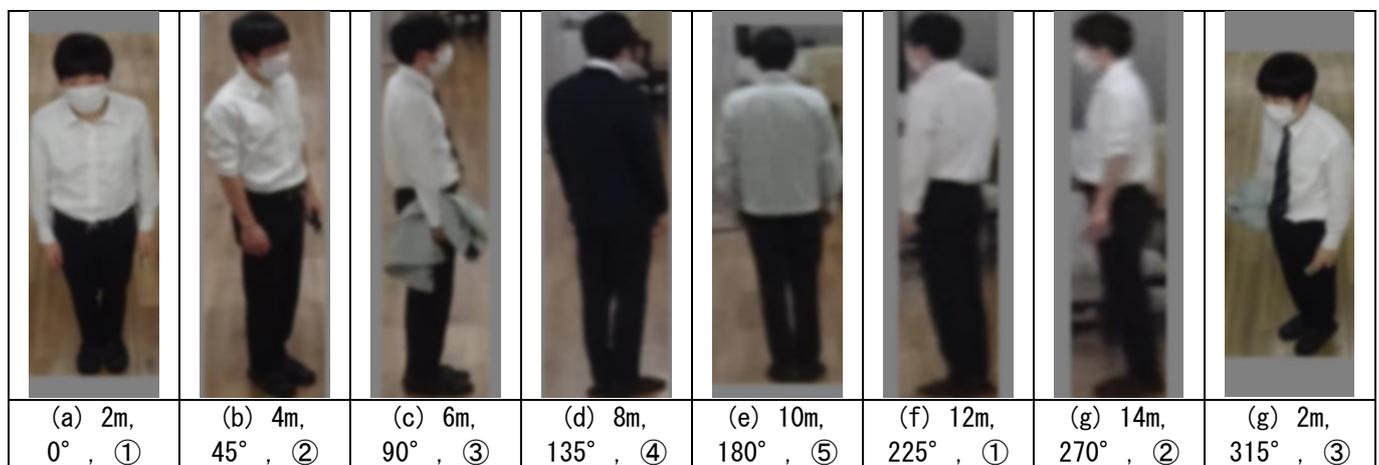


図-1 実験画像例 ※Gaussian Blur によるぼかしを実施

キーワード 深層学習, 距離学習, 人流解析, ArcFace

連絡先 〒111-8648 東京都台東区浅草橋 5-20-8 CS タワー 八千代エンジニアリング(株) TEL03-5822-6843

4. 実験結果

表-1 中の各条件に対して向きや距離ごとに類似度を算出した結果を図-2 に示す. ここでは(a) 0° と 0° , (b) 0° と 90° , (c) 0° と 180° の条件に絞って画像群の類似度を相関行列で示している. 図-2 (a)~(c)に示す①~⑤の番号は表-1 に示すパターンを意味する. 例えば図-2 (b)④の3行4列の要素は, 6m の距離で撮影した 0° を向いたジャケット着用画像と, 8m の距離で撮影した 90° を向いたジャケット着用画像間の類似度を表している.

図-2 (a)より向きが同じであれば距離が離れても比較的高い類似度となることを確認した. さらに(b) 0° と 90° と(c) 0° と 180° を比較したときに, (c)の方が高い類似度となることを確認した. また(b)の③ネクタイ&手荷物ありにて著しく類似度が低下することを確認した.

5. 考察

同じ向きの画像で距離に依らず高い類似度となったのは, 距離の変化により画像の精細さは変化するものの, リサイズによって画像の大域的な特徴量が保持されたためだと判断した. また(b)より(c)の方が高い類似度を示したのは, 全体的なシルエットが 180° の方が 0° の状態に類似するためだと考えた. (b)③にて精度が著しく低下した原因は, 90° を向いた際に右手に持った荷物が対象人物の服を覆い, 0° の画像とは大きく異なる特徴量の画像になったためだと推察した. これらの結果を踏まえて, 本実験で用いた DNN モデルは画像全体をぼんやりと捉えた形状情報をメインに抽出していると予測した.

6. おわりに

本研究では, 人の画像から算出される類似度に影響する要素を明らかにする目的で, 同一人物を様々な距離・向きから撮影した画像データセットを作成した. さらに既存の DNN モデルを用いて, 本データセット内画像の類似度を算出した. その結果, 既存の DNN モデルは撮影距離が変化した場合でも高い類似度を示したが, 向きがカメラ正面向きから横向きへと変化した場合に類似度が下がる傾向にあることが明らかとなった.

実験では同一の見た目をした画像同士の類似度のみに着目したが, 異なる見た目の画像など様々な観点で評価を行うことが今後の課題である. 人流解析では適用する現場に応じて, 表-1 に示す①と②のような若干の違いに対する許容の可否が異なると想定されるため, ニーズに即したモデルが選定できるよう容易にモデルの性能評価が可能となる環境の構築を目指す.

参考文献

- 1) M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, S.C.H.Hoi.: Deep Learning for Person Re-identification: A Survey and Outlook, arXiv: 2001.04193, 2021.
- 2) J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, S. Zafeiriou.: ArcFace: Additive Angular Margin Loss for Deep Face Recognition, arXiv: 1801.07698, 2018.
- 3) 吉田龍人, 菊池恵和, 堀井大輔, 大久保順一, 高森秀司: 深層距離学習のための MOT 解析を使った教師画像自動生成 "in press", 第 37 回人工知能学会全国大会, 2023.

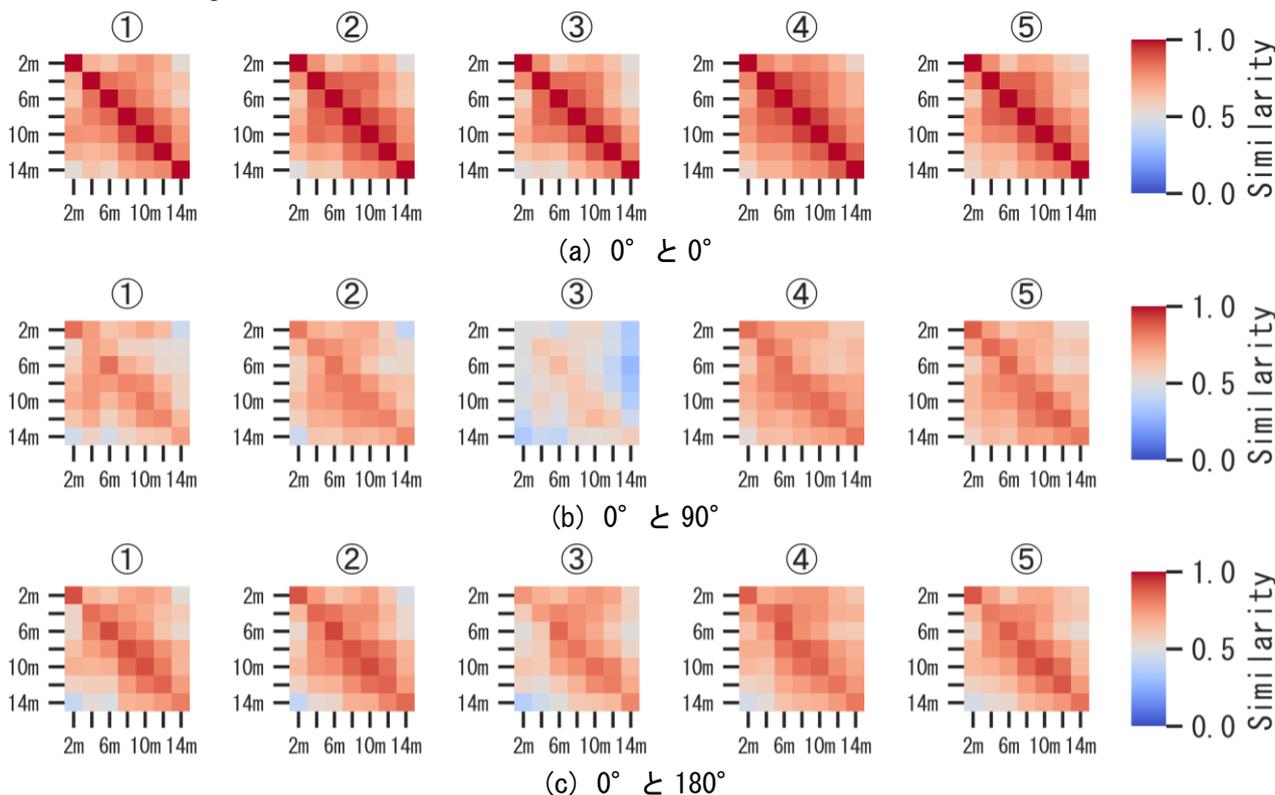


図-2 画像の類似度評価結果