

カメラ映像を用いた人物間の関係性推定手法の検討 Study of People Relationship Estimation Method Using Camera Images

丸尾 元一[†]
Genichi Maruo

尾崎 敦夫[†]
Atsuo Ozaki

1. はじめに

デジタルサイネージにおけるパーソナライズ広告は進化しているが、その多くは個人単位の最適化に留まる。一方で、小売業界などでは AI・DX 技術を活用した顧客体験向上の取り組みが進められており [1]、複数人視聴時の「関係性」を考慮した広告配信は、ターゲティング精度や広告効果、体験価値の向上に繋がる重要な要素であるが、現状では不十分である。そのため、「関係性推定技術」の導入が次なる一手として求められている。本稿では、この課題に対し、単眼カメラ映像から人物の非言語的行動（対人距離、体の向き）を分析し、リアルタイムに人物間の関係性（知人同士か他人同士か）を推測するアルゴリズムを提案する。

2. 関連研究

人物間の関係性推定において、角谷ら [2]は複数のステレオカメラと顔認証技術を用い、視線や表情から対人関係構造を抽出する手法を提案している。同手法は高精度な推定を実現するが、システム構成が複雑であるという課題を有する。また、[2]では取得データに基づくオフライン処理によって関係性が検証されている。これに対し本研究では、単眼カメラから得られる対人距離と身体の向き情報のみを分析対象とし、軽量な処理によるリアルタイムな関係性推定の実現を目的とする。このリアルタイム性は、デジタルサイネージなど即時性が要求される応用を可能にする点において、既存研究のアプローチとは異なる。

3. 提案手法

本研究では、単眼カメラ映像から人物の向きと地面上の座標をリアルタイムに推定し、これらの時系列データを統合して人物間の関係性（知人／他人）を推測する。図 1 に提案手法におけるデータ累積から関係性推論までの処理フローを示す。まず、単眼カメラで取得された映像から人物の位置、向きを推定し、フレームごとのデータをデータベースに蓄積する。次に、蓄積されたデータを用いて、人物間の関係スコアを計算し、最終的に知人関係と他人関係を推論する。

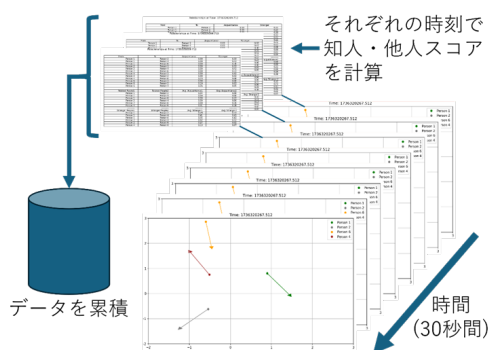


図 1 データ累積と関係性推論の流れ

3.1 人物の向きと位置推定

人物の向きは、YOLO11 のポーズ推定機能 [3]により検出された両肩のキーポイント座標から算出する。具体的には、左右の肩を結ぶベクトルを求め、これを正規化した後、 $-\pi/2$ ラジアン回転させることで向きベクトルを得る。人物の地面上の位置は、ホモグラフィ変換を用いて推定する。事前にカメラ視野内の地面に設置した基準点（例：1m 四方の正方形の頂点）の画像座標と実空間座標の対応からホモグラフィ行列を算出する。検出された人物のバウンディングボックス下端中央を足元位置とし、この画像座標にホモグラフィ行列を適用して地面上の実座標に変換する。

3.2 関係性スコアの算出

人物間の関係性を定量的に評価するため、各人物ペアに対して「知人らしさ」と「他人らしさ」を示すスコアを算出する。このスコアは、人間が日常的に他者との関係性を判断する際に無意識に考慮されると思われる物理的な距離と、互いの注意の向きに着目して設計した。具体的には、人物 A と人物 B の関係性を評価するため、以下の要素を段階的に計算し、最終的なスコアを導出する。

距離 d : 人物 A の座標 (x, y) と人物 B の座標 (x', y') 間のユークリッド距離。

$$d = \sqrt{(x - x')^2 + (y - y')^2} \quad (1)$$

相対位置ベクトル \vec{r}_{norm} : 人物 A から B へ向かう正規化された方向ベクトル。

$$\vec{r}_{norm} = \left(\frac{x' - x}{\|\vec{r}\|}, \frac{y' - y}{\|\vec{r}\|} \right) \quad (2)$$

ただし、

$$\vec{r} = (x' - x, y' - y) \quad (3)$$

向きの類似度 similarity: 人物 A の向きベクトル \vec{d} と \vec{r}_{norm} の内積。0 未満は 0 とする。

$$\text{similarity} = \max(\vec{d} \cdot \vec{r}_{norm}, 0) \quad (4)$$

値が 1 に近いほど A が B の方向を向いていることを示す。

距離に基づく重み f_a : 距離が近いほど影響が大きくなる減衰関数。

$$f_a = \frac{1}{d + 1} \quad (5)$$

知人スコア S_a と他人スコア S_s :

$$S_a = \text{similarity} \cdot f_a \quad (6)$$

S_a は向き合い、かつ近い場合に高くなり、 S_s は向き合っていない、または遠い場合に高くなる。

[†] 大阪工業大学 Osaka Institute of Technology

3.3 関係性の最終判断

各フレームで全人物ペアに対し、一方から他方への S_a と S_s を計算する。これらのスコアを過去 30 秒間にわたり累積し、平均値を算出する。人物 A と B の関係性判断には、以下の相互平均スコアを用いる。

- 相互平均知人スコア: $(A \rightarrow B \text{ の平均 } S_a) + (B \rightarrow A \text{ の平均 } S_a)$
- 相互平均他人スコア: $(A \rightarrow B \text{ の平均 } S_s) + (B \rightarrow A \text{ の平均 } S_s)$
- 相互平均知人スコアが相互平均他人スコアを上回る場合「知人」、逆の場合は「他人」と判断する。

4. 評価方法

提案手法の有効性を検証するため、以下の実験を行った。

- **使用ツール:** CPU: Intel Core i9 13900HX, GPU: NVIDIA RTX 4070 Laptop GPU, OS: Windows 11 Home. Python 3.12, PyTorch 2.5.1+cu124 [4], YOLO11x-pose
- **実験設定:** 協力者に「知人同士」と「他人同士」の 2 種類の関係性を模倣して行動してもらう (例: 知人同士として自然に会話、他人同士として無関心に行動)。
- **評価基準:** システムが推測した関係性と協力者の演技設定 (実際の関係性) との一致度を正解率として評価する。

5. 結果

合計 23 の異なる場面で撮影されたデータを用いて評価を行った。これらの場面で発生した人物ペアの総数は 132 通りであった。

提案システムが 132 通りの対人関係のうち、正しく関係性を推測できたのは 107 通りであった。これにより、システム全体の正解率は約 81.1% (107/132) となった。

図 2 に各シーンにおける対人関係推測の正解率を示す。

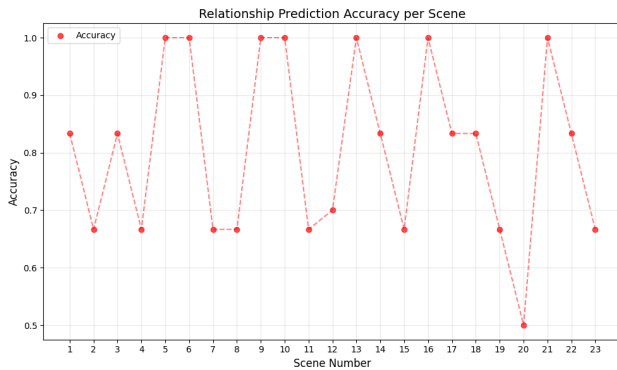


図 2 各シーンにおける対人関係推測の正解率

図 2 から、シーンによって推測精度にばらつきが見られることがわかる。例えば、シーン 5 や 9 などでは正解率が 1.0 と高い精度を示している一方、シーン 2 や 20 などでは正解率が 0.5 程度まで低下する場合もあり、特定の状況や行動パターンが推測の難易度に影響を与えることを示唆している。

6. 考察

実験結果から、提案手法は一定の精度で関係性を推測できる可能性が示された。しかし、約 18.9% の誤判定も確認された。主な誤判定の要因として以下の 2 点が挙げられる。

- **一時的な対面の誤判定:** 短時間向かい合って座る、あるいは通路ですれ違い様に視線が合うなど、一時的に近距離で対面する状況において、実際には関係性が浅い、ま

たは全くない場合でも「知人同士」と誤判定する傾向があった。これは、提案手法が体の向きや距離といった瞬間的な物理的特徴を主要な手がかりとするため、短期的なインタラクションや偶然的な配置を、持続的な関係性として捉えてしまうことに起因すると考えられる。

- **背後の人物との誤判定:** 図 3 は、知人の背後に存在する他人が知人と誤判定された場面の一例である。このように、向き合って話している二人の人物の背後に位置する第三者が、これら二人と関係があると誤判定されるケースがあった。これは、向き合う二人の相対的な位置関係や向きが強く影響し、偶然その範囲に入った第三者を含めてしまうためと推察される。

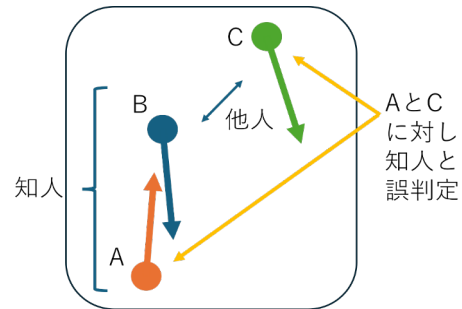


図 3 知人の背後にいる他人が知人と誤判定された場面

提案手法の強みは、単眼カメラという比較的簡易な設備で、リアルタイムに人物間の関係性を視覚情報のみから推定できる点である。一方で、上記の誤判定要因に見られるように、向きや距離といった要素への依存度が高く、複雑な社会的文脈や短時間の意図しないインタラクションの識別には限界がある。

7. まとめと今後の課題

本研究では、単眼カメラ映像から人物の向きと位置を推定し、これらに基づくスコアの時系列分析により、人物間の関係性 (知人/他人) をリアルタイムに推測するアルゴリズムを提案した。評価実験の結果、提案手法は模擬的な状況において約 81.1% の正解率を示し、その有効性の一端を確認した。今後の課題としては、まず、人間による関係性判断との比較検証を行い、提案手法の推測精度をより客観的に評価する必要がある。次に、深層学習モデルを導入し、より多様な特徴量から関係性を学習することで、精度向上を目指す。さらに、より多くの人数が参加する複雑な実環境での検証を行い、手法のロバスト性を高め、実用化に向けた検討を進める予定である。

参考文献

- [1] “三菱商事・KDDI・ローソン、「未来のコンビニ」への変革に向けた取り組みを開始,” Lawson, Inc, 18 9 2024. [オンライン]. Available: https://www.lawson.co.jp/company/news/detail/1491927_2504.html. [アクセス日: 10 1 2025].
- [2] 隆, 角谷, 真, 山本, 博, 芳賀, 重, 金田, “複数ステレオカメラと顔認証技術を用いた集団の対人関係抽出手法の提案,” 情報処理学会研究報告, Vol.2009-UBI-22, No.17, pp. 1-8, 2009.
- [3] “Ultralytics YOLO11,” Ultralytics Inc, 2025. [オンライン]. Available: <https://docs.ultralytics.com/models/yolo11/>. [アクセス日: 20 1 2025].
- [4] “PyTorch documentation,” [オンライン]. Available: <https://pytorch.org/docs/stable/index.html>. [アクセス日: 20 1 2025].