7章 人映像解析による未来犯罪捜査支援

八木康史^{†1}, 榎原 靖^{†1}, 村松大吾^{†1}, 岡田隆三^{†2}, 黄瀬浩一^{†3}, 和田俊和^{†4}

キーワード：広域監視, バイオメトリクス, 行動理解, 犯罪捜査

1. まえがき

平成25年6月「世界最先端IT国家創造宣言」が閣議決定された。本宣言では、2020年までに、世界最高水準のIT利活用社会の実現を目標として、①IT・データの利活用による新産業創出、②世界一安全で災害に強い社会への貢献等が掲げられている。政府は、2016年度から、携帯電話やカーナビの位置情報、人工衛星画像、「〇〇地区で火災発生」といったTwitterにおけるつぶやき等の利活用を想定し、企業協力を呼びかけている。

われわれの暮らす生活空間には、固定防犯カメラ、各種ライブカメラ、車載カメラ、ウェブカメラ等、膨大で多様なカメラが溢れ、これらのカメラ映像の利活用は、次世代映像セキュリティを考える上で極めて重要である。

一方で、膨大で多様なカメラの存在は、新たな問題を生み出す。特に、数百万台のカメラから毎日生み出される、エキサバイトクラスの大規模映像データを人手でモニタリングすることは非現実的である。また、容疑者が映像に映っている場合でも、映像品質が低いことで活用されないケースは多い。カメラの多様性、台数の増加が進む中で、それらの映像からいかに有用な情報を取出すかは重要な課題といえよう。

映像のセキュリティ利用を考えた場合に、これまでの、映像の鮮鋭化、超解像処理など、画像処理の問題として扱われてきた。しかし、犯罪捜査においては、容疑者の探索、特定、足取り追跡が主であり、大規模イベント等の警備目的では、常習犯・テロ容疑者の早期発見、追跡、要人の周囲状況把握、不審行動の発見、万引き、窃盗、傷害等の現

認支援、急病人、迷子等の発見、震災等の災害時は、避難誘導、災害状況の把握など、コンピュータビジョンの問題として捉えた場合、発見、追跡、認識が関係してくる。

本稿では、2010年～2014年度で進めてきた文部科学省の科学技術振興調整費（途中より、社会システム改革と研究開発の一体的推進に改名）によるプロジェクト研究での成果を中心に、人映像解析による科学捜査について展望する。具体的には、ビッグデータ時代の到来により、重要度が増した、低品質、大量映像データに対する顔認証技術ならびに人物認証技術について紹介する。

2. 低品質、大量映像データに対する顔認証技術

犯罪やテロの捜査において個人特定を行う際、人物の顔は大きな手掛かりであり、多くの事件の解決に貢献している。捜査機関には、多数の人物の顔写真を収録したデータベースも存在することから、対象人物の顔画像をデータベース中から自動で検索する機能が実現すると、捜査の効率化が期待される。そのようなシステムを実際の捜査で活用するためには、検索の高速性、および、顔画像に生じるさまざまなタイプの劣化への頑健性が重要となる。

本プロジェクトでは、「劣化を伴う1枚の人物画像を検索キーとし、データベース内の同一人物の顔画像を高速に検索」する高速顔画像検索システムを開発し、顔正面に対して左右 $\pm 90^\circ$ 、上下 $\pm 45^\circ$ の範囲に接する楕円内の角度までの顔向き変化に対応し、照明変動・マスクなどによる顔の隠蔽・低解像度化（最小で 25×25 画素）といったさまざまな劣化タイプの顔画像に対して、上位10位に正解が含まれる精度98%で、1,000万枚/秒以上の画像照合を目指した。

顔向き変化への対応については、まず、顔検出や顔特徴点検出手法として、学習データから自動的に検出に有効な特徴を抽出する統計的アプローチを実現した。提案手法により、顔向きのみならず、多様な変動や劣化を、学習データを追加することで包括的に扱うことができる。また、横顔の学習データの拡充により、左右 $\pm 90^\circ$ 、上下 $\pm 45^\circ$ の範囲に接する楕円内の角度を含むテストセットにおいて平均顔検出率94.3%を達成し、顔特徴点検出については、人手の入力点（真値）に対して誤差5%以内を達成した。顔向

^{†1} 大阪大学 産業科学研究所

^{†2} 株式会社東芝 研究開発センター

^{†3} 大阪府立大学 大学院工学研究科

^{†4} 和歌山大学 システム工学部

"The Future of Video Surveillance for Criminal Investigation" by Yasushi Yagi, Yasushi Makihara, Daigo Muramatsu (The Institute of Scientific and Industrial Research, Osaka University, Osaka), Ryuzo Okada (R&D Center, Toshiba Corporation, Kanagawa), Koichi Kise (Graduate School of Engineering, Osaka Prefecture University, Osaka) and Toshikazu Wada (Faculty of System Engineering, Wakayama University, Wakayama)

図1 劣化顔画像に対する性能評価

図2 低照度顔画像に対する性能評価

きの正規化処理については3次元顔モデルを用いて複数の顔向き画像を生成する手法を開発した。

一方、顔検出・正規化後の認証処理において、精度と高速化を両立するには、検索方法の工夫が極めて重要である。そこで、超高速にラフに絞り込む概略検索と、ある程度絞り込まれた候補に対して、劣化映像に対してもロバストな詳細検索の二段階探索を行う。

具体的に概略検索においては、局所特徴量と近似最近傍探索の組合せにより、認識時間の大幅な削減を目指す。近似最近傍探索による個々の特徴量の照合精度は高くないが、照合結果の統合に投票を用いることで、精度良く絞り込みができる。ただし、低解像度画像では、局所特徴量検出で得られる特徴量の数が減少するため、認識率が低下する。そこで、画像上から特徴点のDense samplingによる特徴抽出を行うことで認識性能の低下を防いでいる。この手法では、犯罪捜査に用いられる顔データベースの規模を1,000万人と想定し、その中から候補を1,000人に絞り込む検索タスク(概略検索)に対して、1000位累積認識率98.5%を達成した¹⁾。

一方、詳細検索では、「画像を直交展開した際の項の欠落または混入」によって劣化をモデル化する考えに基づき、欠落の起きていない部分の展開係数が一致すれば、劣化画像は比較した画像に近いとの考えに基づき、まず、任意の直交展開に対して2枚の画像の展開係数の一致数ができるだけ多くなるように劣化画像の輝度値を調整する。劣化の影響を受けにくい画像間相違度²⁾を提案し、劣化画像に対しても、大幅に精度を向上できることを示した(図1)。

また、防犯カメラは、夜間も稼働し、また、比較的照明環境の悪い場所に設置されている場合も多いことから、照度に対する性能評価も重要である。図2の実験結果から256階調にて、平均輝度20程度までであれば、性能が維持できていることがわかる。

最終的に、概略探索、詳細探索を合わせて、10位正解率98%、毎秒1,000万画像探索を実現した。

3. 大量防犯カメラ映像からの人物照合

犯罪捜査においては、犯人と被疑者の同一性の判定や、カメラ渡りの人物追跡による犯人の逃走経路の割り出しなど、個人認証が重要な役割を果たすこととなる。現在の捜査において、多くの場合、捜査員の目視確認による人海戦術に頼って同一人物を探し出していることから、捜査の効率化のためには自動での個人認証技術が切望されている。

3.1 歩容・頭部・身長によるマルチモーダル個人認証

防犯カメラ映像からの個人認証の手がかりとしては、カメラから離れた場所でも利用可能な歩容による個人認証が有望視されている。実際に、英国においては、歩容鑑定が裁判所で強盗犯の証拠として認められたケースもある³⁾。また、日本においても、われわれの研究グループによって開発された歩容鑑定システム⁴⁾が、平成26年度の警察白書において防犯カメラ映像に基づく新たな個人識別法として紹介されている。また、歩容に加えて、これまでも鑑定に利用されている身長情報や、低解像度ながらも顔や髪型を含めた頭部領域のテクスチャ情報は個人認証の有力な手がかりである。そこで、われわれが提案したマルチモーダル個人認証手法⁵⁾では、歩容・頭部・身長という遠隔からでも観測できる、異なる性質のモダリティをスコアレベル統合することで、認証精度の向上を目指す(図3)。

各モダリティに対する照合器として、歩容に対しては歩容エネルギー画像⁶⁾(平均シルエット⁷⁾)、頭部に対しては解像度が低いことから相関法によるテンプレートマッチング、身長についてはキャリブレーション済みカメラと地面制約に基づく足下点と頭頂点間の距離を用いて差の絶対値を相違度とする手法を各々実装している。

表1は、約1,000人の歩行映像データベースに対する評価実験結果である。人物高さ180画素の実験では、等価誤り率(EER)0.37%という、通常の顔認証と変わらない精度が得られた。また、人物高さ60画素以下では、頭部の解像度が低すぎ有効に利用できないことから、頭部領域を用いなかったが、いずれも統合効果から性能向上が見られている。

図3 スコアレベル統合によるマルチモーダル個人認証

表1 マルチモーダル個人認証の精度

画像中での人物高さ [画素]	EER [%]	10位正解率 [%]	使用特徴
30	1.14	98.74	歩容・身長
60	0.91	98.95	歩容・身長
約180	0.37	99.70	歩容・身長・頭部

なお、ハイビジョンクラスの防犯カメラを想定した場合、標準レンズを想定すると、約50mの遠景に相当する。

さらに、歩容・頭部・身長による個人認証精度は、いずれも時空間解像度による影響を受け、かつ、その影響度合いがモダリティによって大きく異なることを考慮し、時空間解像度のある種の品質尺度と考えて、各モダリティに対してスコアレベル統合時の適切な重みを与える手法を提案した。特に、すべての時空間解像度に対する重みを事前計算することが現実的ではないことから、限られた時空間解像度に対する重みを用いて、任意の時空間解像度に対する重みをガウス過程回帰によって求める手法を開発した⁸⁾。

3.2 人物検出

膨大な映像データすべてをモニタリングすることは不可能であるが、人のモニタリングを目的とする場合は、人が映っているシーンのみモニタリングすれば良い。人の動きは、人々のライフスタイルに合わせて、ある程度規則性があり、防犯カメラに映し出される人も時間帯によりその密度が異なる。したがって、防犯カメラ映像より、人が映し出されるカットのみを抽出することで、モニタリングの効率向上につながる。例えば、人が1分間に平均して1秒程度しか映っていないとすれば、1日のモニタリング時間を1/60の24分へと短縮することができる。そのための基礎技術として、人物検出は必要不可欠である。

人物検出として、①複数の半径の輝度勾配抽出フィルタを適用する CoHOG 特徴量のマルチスケール化、②輝度勾

図4. 人物検出性能

配方向計算の適応的閾値の導入、③Joint Haar-like特徴を用いた識別器のアンサンブル化、の三つの技術を新たに開発し、これらの技術の導入により、タイムラー・クライスラー (DC) データにおいてEERが3.0%と改善されている(図4)。なお、人物検出においては、人型看板等との見間違いを軽減する策として、時間方向の特徴(例えば、CHLAC特徴やSTHOG特徴)の利用も有効である⁹⁾¹⁰⁾。

3.3 カメラ間での渡りを含む人物行動追跡

防犯カメラ上で特定人物の足取りを追跡するためには、カメラ間での渡りを含む人物行動を追跡する必要がある。カメラ渡りに関する研究は、Person Re-identificationとして、国際的にも盛んに研究されている。それらの研究の多くは、対象者の色分布特徴を活用する手法である。カメラ渡りにおいては、同一日のほぼ同じ時間帯に別カメラで撮影された映像間の照合を行うことから、衣装は同一であると仮定している。

本プロジェクトにおいても、異なるカメラによって撮影された人物の対応付けとして、服の色に注目し、画像中の人物領域の色分布に基づく特徴量を利用し、さらに、本人の見え方の変動を抑え、個人間の差異を強調する計量空間を用いた手法を開発した。

特徴量は、人物検出および背景差分によって得られた画像中の領域の色分布を、色の照明等の影響で起こる変動に強いと考えられるHSV空間とYCbCr空間の2種類の色空間のヒストグラムとして表現し、それらを連結したベクトルを特徴ベクトルとして抽出する。その際、ヒストグラム生成では、人物追跡によって得られる画像中の移動軌跡から身体の向きを推定し、対応付けする人物同士の向きから推定される共通部分の色分布により高い重みを付ける手法を開発した。計量空間として、学習データをランダムに分けた二つのグループから学習した識別器を複数生成し、そ

表2 構築されたデータベース一覧

種 類	データの種類	人数・公開の有無	計測情報
顔画像	主に正面に近い顔画像	約1,000万枚	ウェブサイトから収集
直線歩行	2次元映像（斜め前から側方、側方から斜め後ろの計2台）	4,016人（公開、利用申請件数: 42件）	日本科学未来館（2009年度、2011年度等）
直線歩行（往復、荷物あり情報含む）	2次元映像（14方向）	約7万人（計測中、予定）	日本科学未来館（2015年度～）
トレッドミル歩行	2次元映像（25台カメラ）、仮想3次元データ	約600人（一部公開、利用申請件数: 93件）	八木研究室内（2013年度（仮想レンジデータ作成））
意図を含む歩容	2次元映像、2次元レンジデータ、3次元レンジデータ	約900人（公開）	日本科学未来館（2011年度）
万引き犯の行動	2次元映像（店内、頭部、胸部）	36人	大阪大学生協内（2014年度）
集団歩行	2次元映像	約140人	大阪大学構内（2014年度）

の出力を特徴量とする REMetric¹¹⁾ を使用した。

実験では、屋内および屋外で歩行中の50名の人物を正面および横から撮影したデータを使用した。25名を計量空間の学習、残りの25名を評価に使用し、環境や身体の向きが異なる4台のカメラ画像による100人（＝25人×4台）に対して3秒以内のカメラ渡りの照合を行った。その結果、1位正解数が99人で対応誤り率1%であった。

歩行映像を対象とした場合、人物の形状（体型）と動き情報も人の個性を表現する重要な要素である。ここでは、人検出に提案された、形状と動きの両方を表現する STHOG 特徴を用いて、カメラ間の人物照合を行う手法を紹介する。STHOG 特徴自体は、観測方向による影響を受けることから、歩行方向とカメラ姿勢を考慮した幾何補正による人物ウィンドウ領域の補正と STHOG 特徴の勾配補正、さらに、位相に非依存な方向変換モデルによる STHOG 特徴の変換を組合せて、人物照合を行う。また、色特徴と STHOG 特徴を、観測角度差に応じて適応的に統合する手法についても提案した¹²⁾。初期性能評価として、実証実験場所（某小学校のデータセット）に対して、3秒以内のカメラ渡りではほぼ0%の対応誤り率を実現した。

4. データベース

これまでに述べてきたような人映像解析による犯罪捜査支援技術の評価においては、共通のデータベースを利用することが望ましいため、そのようなデータベースについても構築した（表2）。顔画像については、ウェブサイトを中心にして、1,000万枚を収集し、上述の高速顔画像検索の評価実験にも用いている。また、歩行映像については、地面上の直線歩行のデータを日本科学未来館での展示イベント等と並行して収集し、その被験者数の規模は世界の既存データベースの10倍以上と、世界最大のものとなっている¹³⁾。さらに、現在、同じく日本科学未来館にて常設展示と並行したデータ収集も実施しており、期間終了までに約7万人の歩行データが収集される予定である。一方、条件変化の歩行映像としては、トレッドミル（歩行機）上での歩行として、観測角変化・速度変化・服装変化を伴うようなデータベースを構築しており¹⁴⁾、視体積交差法による3次元デー

図5 人物対応付けの評価に使用した登録と照合データ

タについても作成している。他にも、対象から逃げる、対象を追いかけるといったような意図を含む歩行映像のデータ、万引き犯の行動を模した歩行映像のデータ、また、集団での歩行映像データなど、さまざまなバリエーションを持つデータベースを構築することで、多面的な映像セキュリティの性能評価が可能となる。

5. むすび

本稿では、2010年～2014年度に進めてきた文部科学省の科学技術振興調整費（途中より、社会システム改革と研究開発の一体的推進に改名）プロジェクトでの成果を中心に、人映像解析による科学捜査について紹介した¹⁵⁾¹⁶⁾。低品質、大量映像データに対する顔ならびに人物認証技術は、ビッグデータ時代に必要不可欠な技術であり、例えば、オリンピック等の世界的イベントにおいては、全世界から犯罪者、テロリストも集まり犯罪リスクが高くなり、防犯カメラ設

置台数もますます増加することを考えると、効率的絞り込みが極めて有効であり、必要とされる技術である。

一方で、一般市民の中には防犯カメラ映像設置や防犯カメラ映像を対象とした人映像解析に対しネガティブな反応を示す人も少なからず存在する。このネガティブな反応の多くは、個人を特定されたり、年齢性別等の属性情報を取得されたり、特定されなくても一人ひとり同定され追跡されたりすることへの嫌悪感や、監視社会への恐怖感に基づくものと考えられる。科学技術の倫理的・法的・社会的問題(ELSI)に対する課題解決においては、市民協働による研究開発プロセスが必要不可欠である。例えば、ELSI専門家によるワークショップ等を通じたアセスメントでは、一般市民が想像し得ないことも含めたあらゆる可能性を検討、さらに、一般市民参加型ワークショップでは、実際に市民が何に不安を感じるか議論を行い、その結果を踏まえての研究開発の改善を行い、そして、不安解消策の提示、新技術紹介による透明性の確保が、両者の距離を埋める手段となると考えられる。そのような新しいスタイルでの研究開発が行える市民生活の場を行政とも一体となり整備していくことも、今後の映像セキュリティ分野の発展において重要である。

(2015年11月10日受付)

〔文 献〕

- 1) 内海ゆづ子, 坂野悠司, 前川敬介, 岩村雅一, 黄瀬浩一: "局所特徴量と投票処理を用いた大規模データベースに対する高速顔認識", 信学誌, J97-D, 8, pp.1263-1272 (Aug. 2014)
- 2) 岡 藍子, 和田俊和: "欠損による劣化に対して頑健な画像間相違度に関する研究", 信学誌, J94-D, 8, pp.1269-1278 (2011)
- 3) I. Bouchrika, M. Goffredo, J. Carter and M. Nixon: "On Using Gait in Forensic Biometrics", Journal of Forensic Sciences, 56, 4, pp.882-889 (2011)
- 4) H. Iwama, D. Muramatsu, Y. Makiyara and Y. Yagi: "Gait Verification System for Criminal Investigation", IPSJ Trans. On Computer Vision and Applications, 5, pp.163-175 (2013)
- 5) 村松大吾, 岩間晴之, 木村卓弘, 横原靖, 八木康史: "一歩行映像から取得される複数特徴を用いた個人認証", 信学誌, J97-A, 12, pp.735-748 (2014)
- 6) J. Han and B. Bhanu: "Individual Recognition Using Gait Energy Image", IEEE Transactions on Pattern Analysis and Machine Intelligence, 28, 2, pp.316-322 (2006)
- 7) Z. Liu and S. Sarkar: "Simple Representation Yet for Gait Recognition: Averaged Silhouette", Proceedings of the 17th International Conference on Pattern Recognition, 4, pp.211-214 (2004)
- 8) T. Kimura, Y. Makiyara, D. Muramatsu and Y. Yagi: "Quality-dependent Score-level Fusion of Face, Gait and the Height Biometrics", IPSJ Trans. On Computer Vision and Applications, 6, pp.53-57 (2014)
- 9) T. Kobayashi, N. Otsu: "Action and Simultaneous Multiple-Person Identification Using Cubic Higher-order Local Auto-Correlation", Proc. International Conference on Pattern Recognition (ICPR), pp.741-744 (2004)
- 10) C. Hua, Y. Makiyara and Y. Yagi: "Pedestrian Detection by using a Spatio-temporal Histogram of Oriented Gradients", IEICE Trans. On Information and Systems, E96-D, 6, pp.1376-1386 (2013)
- 11) T. Kozakaya, S. Ito and S. Kubota: "Random Ensemble Metrics for Object Recognition", Proc. of 13th International Conference on Computer Vision, pp.1959-1966 (2011)
- 12) R. Kawai, Y. Makiyara, C. Hua, H. Iwama and Y. Yagi: "Person Re-identification using View-dependent Score-level Fusion of Gait and

Color Features", Proc. of the 21st Int. Conf. On Pattern Recognition (ICPR 2012), pp.2694-2697, Tsukuba, Japan (Nov. 2012)

- 13) H. Iwama, M. Okumura, Y. Makiyara and Y. Yagi: "The OU-ISIR Gait Database Comprising the Large Population Dataset and Performance Evaluation of Gait Recognition", IEEE Trans. On Information Forensics and Security, 7, 5, pp.1511-1521 (2012)
- 14) Y. Makiyara, H. Mannami, A. Tsuji, M.A. Hossain, K. Sugiura, A. Mori and Y. Yagi: "The OU-ISIR Gait Database Comprising the Treadmill Dataset", IPSJ Trans. On Computer Vision and Applications, 4, pp.53-62 (2012)
- 15) 安全・安心な社会のための犯罪・テロ対策技術等を実用化するプログラム: "人物映像解析による犯罪捜査支援システム", 平成24年度成果報告書 (2013)
- 16) 安全・安心な社会のための犯罪・テロ対策技術等を実用化するプログラム: "人物映像解析による犯罪捜査支援システム", 平成26年度成果報告書 (2015) (公開予定)

やぎ やすし 1985年, 三菱電機(株)入社。同社産業システム研究所にて, ロボットビジョンの研究に従事。1990年, 大阪大学基礎工学部情報工学科助手。同学部システム工学科講師, 同大学大学院助教授を経て, 2003年, 同大学産業科学研究科教授。2012年, 同大学研究所長。2015年より, 大阪大学理事・副学長。1995年~1996年, 英オックスフォード大学客員研究員。2002年, 仏ピカルディ大学招聘助教授。全方位視覚センシング, 画像理解, 知能ロボットに関する研究に従事。工学博士。

ひらばら やすし 2005年, 大阪大学産業科学研究科特任助手。2006年, 同大学研究所助手。2007年, 同大学研究所助教。2014年, 同大学研究所准教授となり, 現在に至る。歩容認証を含むパターン認識に関する研究に従事。博士(工学)。

むらまつ だigo 1999年, 三菱商事(株)入社。2004年, 早稲田大学理工学部助手。2007年, 成蹊大学理工学部助手。2008年, 同学部助教。2011年, 大阪大学産業科学研究科特任講師となり, 2015年, NICT主任研究員。現在, 大阪大学産業科学研究科准教授。オンライン署名認証・歩容認証を含むバイオメトリクスに関する研究に従事。博士(工学)。

おかだ りょういち 1999年, 大阪大学大学院工学研究科博士課程修了。同年, (株)東芝研究開発センター関西研究所に入社。2006年~2008年, カリフォルニア大学ロサンゼルス校客員研究員。車載画像認識, 映像監視を中心にコンピュータビジョン技術, システムの研究開発に従事。博士(工学)。

きせい こういち 1991年, 大阪府立大学工学部電気工学科助手。同大学講師, 助教授, 准教授を経て, 現在, 同大学大学院工学研究科教授。2000年~2001年, ドイツ人工知能研究センター客員教授。2013年より, 日本学術振興会学術システム研究センター研究員。現在, IAPR TC11 (Reading Systems) Chair, International Journal of Document Analysis and Recognition (IJ DAR) Editor-in-Chief。文書画像解析, 大規模パターン認識などの研究に従事。博士(工学)。

わだ しゅんろ 1990年, 岡山大学工学部助手。1995年, 同大学講師。1997年, 京都大学大学院情報学研究科助教授。1999年~2000年, カナダ・ブリティッシュコロンビア大学客員助教授。2002年, 和歌山大学システム工学部教授となり, 現在に至る。パターン認識, コンピュータビジョンの研究に従事。工学博士。