

平成 30 年 6 月 23 日現在

機関番号：13901

研究種目：挑戦的萌芽研究

研究期間：2016～2017

課題番号：16K12460

研究課題名(和文) 監視カメラ映像を用いた人物詳細認識

研究課題名(英文) Fine-Grained Classification of Person Attributes with Surveillance Cameras

研究代表者

加藤 ジェーン (Kato, Jien)

名古屋大学・情報科学研究科・准教授

研究者番号：70251882

交付決定額(研究期間全体)：(直接経費) 2,700,000円

研究成果の概要(和文)：本研究の目的は、膨大な監視映像に含まれている人物の詳細認識を実現することである。そのため、車載カメラ映像中の歩行者を認識対象とし、4種類の属性(計14クラス)の詳細認識を目指した。認識精度を高めるために、以下の手法を考案し開発した。(1)超解像処理、(2)特徴抽出時のパッチ分割、(3)部分と全体に着目したCNNの融合、(4)姿勢ごとの識別器の構築、及び(5)グラフィカルモデル。これらすべての手法を組み合わせることによって詳細認識を実現した。大規模歩行者データセットであるCRPデータセットを用いて評価した結果、提案手法はstate-of-the-artの手法による精度を上回ったことが確認された。

研究成果の概要(英文)：The objective of this research is to recognize fine-grained categories of person attributes from enormous surveillance videos. To achieve this goal, we focus on 4 kinds of pedestrian attributes including 14 categories. We developed five methods to enhance the accuracy of fine-grained classification, including (1) super-resolution based image processing, which helps to recover the image details; (2) patch dividing based feature extraction, which extracts features in a way that preserves the spatial layout of inputs; (3) fusing multiple CNN models, which acquires more detailed features; (4) pose-wise classifier sharing, which learns robust classifiers and makes robust predictions; and (5) graphical model based inference, which utilizes the interdependence between different subcategories to update raw estimations to better ones. We conducted experiments on a pedestrian data set, and confirmed superior performance of our approach based on these methods over the state-of-the-art.

研究分野：画像・映像の内容理解

キーワード：歩行者属性の詳細認識 歩行者認識 部分と全体に着目したCNN融合 超解像処理 グラフィックモデル
姿勢推定

1. 研究開始当初の背景

監視カメラの急速な普及に伴い、監視カメラは防犯、防災、計測・記録等の分野で益々重要な役割を果たしている。しかし、現状では蓄積された膨大な映像から必要な情報を抽出するためには多くの人手に頼らざるを得ない。その莫大な人的コストは監視映像の有効利用を妨げていることから、監視映像の自動解析へのニーズが高まっている。

監視映像の自動解析に関する研究では、現在、人物の同定(全身のアピランス、歩容等による)が注目されている。しかし、この種の技術はインスタンス・ベースであり、例えば、アピランスによる人物同定では、対象人物の服装が変わるだけで正しい認識結果を得られなくなる。もちろん、「中年中肉男性」のような属性による高度な検索は困難である。

2. 研究の目的

本研究では、映像自動解析の次世代基盤技術として、映像に含まれている人物の詳細認識技術を開発する。人物詳細認識は、人の詳細な属性(性別、年齢、服装、体重等)に対して詳細な属性値の分類(年齢の場合:少年、青年、中年、高年等)を行うことを意味する。これまで、人の少数の属性を対象に識別する研究例があった。扱う属性値は少ないため、一般物体認識手法(「椅子」と「車」の分類など基本カテゴリの識別手法)が問題解決に有効であった。しかし、一般物体認識手法は、基本カテゴリの大ざっぱな統計情報に基づいているため、属性数または属性値の種類が大幅に増える場合、カテゴリ間の視覚的差異が小さくなり、認識能力が大幅に低下してしまい、実用性が完全に失ってしまう。本研究では、人物詳細認識手法(属性数:4-8,各属性の取り得る値:2-5)を開発し、実世界データ上でその有効性を検証することを目的とする。

3. 研究の方法

本研究は、車載カメラ映像中の歩行者を研究対象とし、歩行者の4種類の属性(性別2クラス:男性・女性;年齢5クラス:子ども・若年・成年・中年・老年;体格3クラス:痩せ型・健康型・肥満型;服装4クラス:薄着・運動着・カジュアル・ドレス)の詳細認識を目指す。

歩行者の詳細認識精度を向上させるために、我々は、5つのアプローチを開発する。(1)超解像処理:画像細部の特徴を保持するため、(2)特

徴抽出時におけるパッチ分割:画像の空間的なレイアウトを保持するため、(3)部分と全体に着目した CNN の融合:全体のアピランスと詳細な特徴表現を学習するため、(4)姿勢ごとの識別器の構築:同種の姿勢別に識別器を学習し、頑健な推定を行うため、(5)グラフィカルモデル:複数の属性間の依存関係を利用して推定結果を補整するため。

詳細認識では歩行者の画像から、属性ごとにクラスラベルを推定する。超解像処理(SR)とパッチ分割(PD)、部分と全体に着目した CNN の融合(SF)は、より高度に特徴を表現するためのアプローチである。SRでは画像細部の形質を保持したままスケールアップができ、PDでは歩行者の空間的なレイアウトを反映した特徴の獲得を可能にする。また、SFでは細部と全体の両方のアピランスを考慮した特徴の獲得を可能にする。一方、姿勢ごとの識別器の構築(PS)とグラフィカルモデル(GM)は、認識性能を向上させるためのアプローチである。PSでは姿勢別に識別器を学習することで、各識別器の頑健性を上げる。更にGMでは認識精度を向上させるために、複数の属性間の依存関係を利用する。これら五つのアプローチは高い適応性を持っており、各アプローチを個別に取り入れることも、複数のアプローチを組み合わせることも可能である。

(1)超解像処理

一般に、歩行者の詳細認識手法では、歩行者の検出器の出力結果を認識の入力として扱う。通常歩行者の識別器では、出力される歩行者のインスタンスはサイズやアスペクト比が異なるため、特徴抽出の際には画像の正規化が必須となる。他の研究においては、最近傍やバイリニア、バイキュービックなどの補間手法が最も汎用的に使用されている。しかしこれらの手法では、低解像度画像の拡大において不鮮明の画像が出力されてしまう。また特徴抽出の際は一様のサイズに正規化されるため、各画像におけるアスペクト比の変化率が異なる。これらのことは詳細認識の性能に悪影響を与えていると解釈できる。そこで本研究では、超解像処理とパッチ分割により、これらの問題の解決を図る。本研究では超解像処理に、画像のスパース性に基づいた手法を適用する。パッチ分割では、3種類の正方形のパッチにより、歩行者画像から画像上部、中部、下部から領域抽出する。

(2)部分と全体に着目した CNN の融合

一般に CNN の学習において、認識対象の部分画像をネットワークの入力とした場合、該当部分に対しては詳細な特徴が表現できるが、画像全体の情報が欠落する。一方、認識対象の全体画像を入力とした場合は、画像全体のアピアランスを考慮することができるが、各部分に対しては、部分画像を用いて学習した場合と比較して詳細な特徴が表現できない。この問題を解決するため、本研究では、複数の部分画像と全体画像、それぞれのパッチ別に CNN モデルを学習する。そして、各 CNN モデルを統合したネットワークに対して再学習をすることで、相互の欠点を排し、より高精度な認識を実現する。

(3)姿勢毎の識別器の構築

本節では、認識に人体の姿勢情報を用いる。姿勢情報は、14 箇所のキーポイントによって表現される。本研究では、これらのキーポイントのうち、下半身位置を示す 6 箇所を使用して姿勢ごとの識別器を構築する。

歩行者の姿勢には多様な種類が存在する。本研究では、人物の姿勢ごとに複数の識別器を学習し、それぞれの識別結果を統合してクラスを予測する。次のような流れで識別が行われる：

人物の姿勢を推定する、人物の向きを右向きに統一する、脚部のキーポイントを用いて学習データをクラスタ分類する、クラスタごとに識別器と識別器の重み導出パラメータを学習する、

テストデータに対して各識別器の推定確率と重み導出パラメータを使用して認識クラスを決定する。

(4)グラフィカルモデル

歩行者の詳細認識では性別、年齢、体格、服装の属性のクラス分類に焦点を当てている。これらの属性間には相関が存在する(例：子供は年配と比較して痩せ型である傾向が強い等)。グラフィカルモデルによる手法では、画像による識別結果に加えて、このような傾向による補正を加えることで精度をさらに向上させる。

本アプローチでは、4種類の属性の中で性別は他の属性と比べて相関が弱いと判断したため、年齢、体格、服装の属性を用いて補正を行う。

4. 研究成果

歩行者は歩行者特有の振る舞いが存在することから、それらの特徴を詳細認識に利用する5つの効果的なアプローチを考案した。実用性の観点からして、これらのアプローチはほかの手法への追加適用が容易である。つまりこれは、既

存の詳細認識手法との統合を行うことで、更なる精度の向上が期待できることを意味する。

評価実験では、近年公開された大規模歩行者データセット(CRP データセット)を用いて state-of-the-art の精度との比較を行った。その結果、5つのアプローチ全てがベースラインの精度を上回った。更に、各アプローチを統合する提案手法では、state-of-the-art の値を超えることが確認された。また、深層学習を取り入れた提案手法による評価実験でも、更なる精度向上を示す結果が得られた。

今後の課題としては、大きく三つの事柄が挙げられる。まず、第一に、他の既存手法に対して提案手法を組み合わせ、性能の変化を確認することである。実用性の観点からして、本研究で提案した5つのアプローチは他の手法への追加適用が容易である。つまりこれは、既存の詳細認識手法との統合を行うことで、更なる精度の向上が期待できることを意味する。精度の変化を確認するため、既存の歩行者認識手法を提案手法のパイプラインへ組み込み、同様の実験を行う。

次に、より多くの属性、クラスを含むデータセットを使用して実験を行い、提案手法の拡張性を評価することである。提案手法では、性別、年齢、体格、服装に焦点を当て実験を行ったが、これは CRP データセットで定義された属性がこれら4種類であるためであり、提案手法はより多くの属性、クラスに対しても有用であると推測される。CRP データセット以上に詳細な属性、クラスを定義したデータセットが公開され次第実験を行い、提案手法の有用性を再確認する。

最後に、手法の演算効率を向上させることで、実世界のアプリケーションへの適用を目指すことである。本研究では、基礎研究として精度向上に重きを置いており、演算効率については考慮していない。最終的な目的である、実世界アプリケーションへの適用のためには、演算効率を向上させ実用的な認識速度に収める必要がある。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計6件)

加藤ジェーン、王彧、小久保嘉人、「深層学習アプローチに基づいた歩行者の詳細認識」、画像ラボ、日本工業出版、Vol.29, No.2, pp.15-24, 2018.2. (招待論文)

小久保嘉人、王彧、加藤ジェーン、「単眼カ

- メラ映像からの路面位置推定手法の開発」, 車載センシング技術の開発と、自動運転システムへの応用, 第 7 節, pp.228-234, (株)技術情報協会, 2017. (招待論文)
- Yuan Cao, Heng Qi, Jien Kato and Keqiu Li, "Hash Ranking with Weighted Asymmetric Distance for Image Search", IEEE Transactions on Computational Imaging, Vol.3, No.4, pp.1008-1019, 2017. (査読有)
- 加藤ジェーン, 張冠文, 王彧, 「フィルター・パラメータの重要度に着目した小規模タスクのための CNN 初期化手法」, 人工知能学会論文誌, Vol.32, No.3, pp.1-11, 2017. 3. (査読有)
- 小久保嘉人, 王彧, 張冠文, 加藤ジェーン, 間瀬健二, 「歩行者の詳細認識精度を向上させるための追加型手法」, 電子情報通信学会論文誌, Vol.J100-D, No.2, pp. 265-276, Feb. 2017. (査読有)
- Yu Wang and Jien Kato, "Efficient Local Feature Encoding for Human Action Recognition with Approximate Sparse Coding", IEICE Transaction on Information and Systems, Vol.E99-D, No.4, pp.1212-1220, 2016. (査読有)
- [学会発表] (計 20 件)
- Yu Wang and Jien Kato, "Collision Risk Rating of Traffic Scene from Dashboard Cameras", Proc. of International Conference on Digital Image Computing: Techniques and Application (DICDA 2017), Sydney, Australia, pp.519-524, Nov.29-Dec.1, 2017. (査読有)
- Dichao Liu, Yu Wang and Jien Kato, "Evaluation of Triple-Stream Convolutional Networks for Action Recognition", Proc. of International Conference on Digital Image Computing: Techniques and Application (DICDA 2017), Sydney, Australia, pp.513-518, Nov.29-Dec.1, 2017. (査読有)
- Longjiao Zhao, Yu Wang and Jien Kato, "Local Features Augmenting for Better Image Retrieval", Proc. of International Conference on Digital Image Computing: Techniques and Application (DICDA 2017), Sydney, Australia, pp.683-687, Nov.29-Dec.1 2017. (査読有)
- Cao Cong, Yu Wang, Jien Kato, Guanwen Zhang and Kenji Mase, "Solving Occlusion Problem in Pedestrian Detection by Construction Discriminative Part Layers", IEEE Winter Conference on Applications of Computer Vision (WACV 2017), Santa Rosa, U.S.A., pp.91-99, March 27-29, 2017. (査読有)
- Yoshihito Kokubo, Yu Wang, Jien Kato, Guanwen Zhang and Kenji Mase, "Add-on Strategies for Fine-grained Pedestrian Classification", Proc. of International Conference on Digital Image Computing: Techniques and Application (DICDA 2016), Gold Coast, Australia, pp.532-537, Nov.30-Dec.2, 2016. (査読有)
- 小平美沙季, 王彧, 加藤ジェーン, 「マルチレベルのパーツマイニングを用いた歩行者検出」, 電子情報通信学会技術報告(PRMU), 117(514), pp.97-102, 青山学院大学, 2018.3.18. (査読無)
- Longjiao Zhao, Yu Wang and Jien Kato, "Image Retrieval with Augmented Fine-tuned CNN Features", 電子情報通信学会技術報告(PRMU), 117(514), pp.115-119, 青山学院大学, 2018.3.18. (査読無)
- Ruisheng Qin, Yu Wang and Jien Kato, "BOJ253: A Large-scale Dataset for Butterfly Recognition", 電子情報通信学会技術報告(PRMU), 117(362), pp.19-24, 慶応大学, 2017.12.16. (査読無)
- 三井弘希, ワンユ, 加藤ジェーン, 「車載カメラ及び Lidar を用いた多操作状況の推定」, 電子情報通信学会技術報告(PRMU), 117(362), pp.59-63, 慶応大学, 2017.12.16. (査読無)
- Dichao Liu, Yu Wang and Jien Kato, "Hierarchical Multi-stream STNs for Fine-grained Action Recognition", 電子情報通信学会技術報告(PRMU), 117(362), pp.13-17, 慶応大学, 2017.12.16. (査読無)
- Dichao Liu, Yu Wang and Jien Kato, "Spatial Attention Selection for Fine-grained Action Recognition", 電気関係学会東海支部連合大会講演論文集, 名古屋大学, G4-3, 2017.9.8. (査読無)
- IEEE Nagoya Section Student Paper Award*
三井弘希, 加藤ジェーン, 王彧, "車載カメラ映像を用いた煩雑状況の推定", 電気関係

学会東海支部連合大会講演論文集, 名古屋大学, G4-3, 2017.9.8. (査読無)

小久保嘉人, 王彧, 加藤ジェーン, 間瀬健二, 「部分と全体を注目した CNN の融合による歩行者の属性の詳細認識」, 電子情報通信学会技術報告 (PRMU), 116(461), pp.87-90, 北海道大学, 2017.2.18. (査読無)

Cong Cao, Yu Wang, Jien Kato and Kenji Mase, "Part-aware CNN for Pedestrian Detection", 電子情報通信学会技術報告 (PRMU), 116(461), pp.87-90, Hokkaido University, 2017.2.18. (査読無)

Dichao Liu, Yu Wang, Jien Kato and Kenji Mase, "Evaluation of Triple-stream Convolutional Network for Action Recognition", 電子情報通信学会技術報告 (PRMU), 116(461), pp.91-94, Hokkaido University, 2017.2.18. (査読無)

Longjiao Zhao, Yu Wang, Jien Kato and Kenji Mase, "More Convolution Filters for Better Image Retrieval", 電子情報通信学会技術報告 (PRMU), 116(461), Hokkaido University, pp.211-212, 2017.2.18. (査読無)

Yingying Di, Yu Wang, Jien Kato and Kenji Mase, "A Dataset of Dashboard Camera Videos for Driving Danger Detection", 情報処理学会第 79 回全国大会, pp.2-447-2-448, Nagoya University, 2017.3. 16-18. (査読無)

(査読無)

三井弘希, 王彧, 加藤ジェーン, 間瀬健二, 「車載センサと歩行者の属性による歩行者の行動予測」, 情報処理学会第 79 回全国大会, pp.3-251-3-252, 名古屋大学, 2017.3.16-18. (査読無)

北垣修, 王彧, 加藤ジェーン, 間瀬健二, 「車載カメラ映像における危険の検知」, 情報処理学会第 79 回全国大会, pp.2-455-2-456, 名古屋大学, 2017.3.16-18. (査読無)

Cong Cao, Guanwen Zhang, Yu Wang, Jien Kato, Kenji Mase, "Pedestrian Detection using Mid-level Body Parts", 電子情報通信学会技術研究報告 (PRMU), 116(39), pp.91-94, Nagoya University, 2016.5.19. (査読無)

(査読無)

Monthly Best Presentation Award

[図書] (計 0 件)

[産業財産権]

出願状況 (計 0 件)

[その他]
ホームページ等
無し

6. 研究組織

- (1) 研究代表者
加藤 ジェーン (KATO Jien)
名古屋大学・大学院情報学研究科・
准教授
研究者番号: 70251882
- (2) 研究分担者
王 彧
名古屋大学・大学院国際開発研究科・
助教
研究者番号: 60724169
- (3) 連携研究者
無し
- (4) 研究協力者
無し