

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年 4月 27日現在

機関番号：34419

研究種目：若手研究（B）

研究期間：2011～2012

課題番号：23700210

研究課題名（和文） 多カメラ映像監視のための場所固有コンテキストのモデル化

研究課題名（英文） Context Modeling for Multi-Camera Video Surveillance

研究代表者

波部 斉 (HABE HITOSHI)

近畿大学・理工学部・講師

研究者番号：80346072

研究成果の概要（和文）：

防犯カメラなどの複数のカメラで広域監視を行う状況を想定し、映像中の人物を精度よく検出する手法の検討を行った。既存の手法では、広く一般に共通した情報を用いて検出を行うが、それだけでは不十分な場面も多い。そこで「場所依存コンテキスト」すなわち、場所によって変化する事前知識を適切にモデル化することで、人物などの対象物体の有無を正しく判断する。この手法の応用例として、人物検出や物体追跡システムを開発した。

研究成果の概要（英文）：

We propose an object detection framework especially for wide-area visual surveillance system. Existing methods usually make use of features which target objects usually have in common. However, those common features are sometimes insufficient for robust detection. Therefore, in this project, we model contextual information which is particularly associated with a certain location. This enables us to detect targets robustly and efficiently.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
交付決定額	3,300,000	990,000	4,290,000

研究分野：コンピュータビジョン

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：画像情報処理，人物検出，アンサンブル学習，物体検出，物体追跡

1. 研究開始当初の背景

コンピュータビジョンの研究において、映像中の人物検出やその行動の理解は基本技術として古くから研究が行われている。

背景差分などによって映像中の変動を検知して追跡を行う事例（文献[a]）に続き、人物の見え方をモデル化し、ある領域の中に人物が存在するか否かを高精度で識別する手法（文献[b]）が提案され、映像監視など実用システムとしても用いられはじめている。

さらに、近年では検出対象の人物だけではなく、対象とするシーンの構造、付近の物体や人物との相互関係などの事前知識を利用し

て精度を向上させる手法が提案されている。このような事前知識は一般にコンテキストと呼ばれている。例えば、権威ある国際会議 ICCV で最優秀論文となった[c]は物体相互の位置関係をモデル化して過検出や未検出を低減する手法である。たとえば、混雑している状況でも、人物を過不足無く検出したり、他の物体（自転車など）に重なっていても人物を正しく検出できることを示している。

研究代表者も、同じような観点から、シーン中に人物が存在し易い領域を自動的に求めて人物検出の精度を向上させる手法を提案していた。このように、コンテキストを利用する研究が活発に行われ人物検出の精度

を大幅に向上できることが示されている。

このように基礎技術が確立されつつある中で、次のステップとして、実際に応用される場面を想定して必要な技術を確立していく必要がある。そこで、街中に多数設置されている防犯カメラから得られた膨大な映像から人物の動きを獲得・解析する応用を想定し、そこでの人物検出・行動理解の効率化・高精度化を最終目標とする本研究に着手した。

■参考文献

- [a] N.Oliver et.al. A Bayesian Computer Vision System for Modeling Human Interactions. PAMI, 2000.
- [b] N.Dalal et. al. Histograms of Oriented Gradients for Human Detection. CVPR, 2005
- [c] C.Desai et. al. Discriminative Models for Multi-class Object Layout, ICCV, 2009
- [d] P.M.Roth et. al. Classifier Grids for Robust Adaptive Object Detection, CVPR, 2009

2. 研究の目的

一般に、映像中に存在する人物は様々な見え方（姿勢）をし、シーン中の人の流れは場所や時間によって変化する。よって、コンテキスト情報は時と場合によって動的に変化すると考えられる。その変化の要因としては、場所（対象人物の位置やカメラの位置）の違い、対象とする人物の個人差、日照などの環境変動などが考えられるが、その中で大きく影響するのは対象人物が存在する場所の違いである。例えば、カメラに対する位置が違えば画像中の人間の形状は大きく変化し、人間同士の重なり具合（相互位置関係）も変化する。

先行研究では、これらの変化を扱うために、様々なカメラ設置場所や環境変動などを網羅したデータを収集し、すべてに共通なコンテキストを学習している。しかし、カメラの設置箇所のバリエーションが増えるにつれ変化も多様になっていくため共通したコンテキストのみでは効率的に表現できない。そこで、それぞれの場所に固有のコンテキスト（以降、場所固有コンテキスト）を効率よく表現できるモデルを導入し、共通のコンテキストと併用することで、効率的かつ高精度な人物検出・行動理解が可能になると期待される。

本研究では、このような観点から場所に固有の事前知識（コンテキスト）と共通の事前知識（コンテキスト）のモデル化・獲得・利用のフレームワークの構築を目的とした。

3. 研究の方法

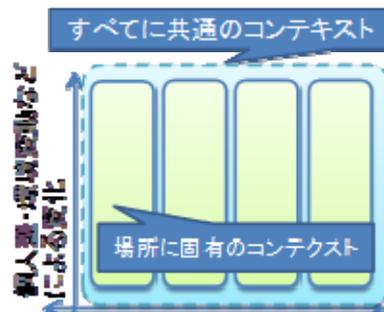


図 1 場所固有コンテキスト

本研究では図 1 に実線で示すような「場所固有コンテキストを考慮した人物モデルの獲得」によって頑健な人物検出を実現し、さらに「場所固有コンテキストを考慮した人物相互関係モデルの獲得」によってその動作のより深い理解を実現する。前者は画像上での人物の大きさや見え方を記述したモデルであり、後者は通行している人物同士の関係を記述したものである。

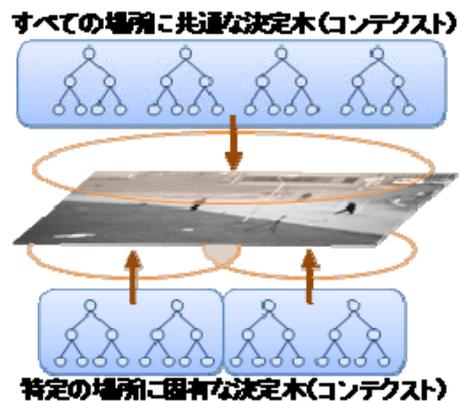


図 2 Random Forests によるモデル化

場所固有コンテキストのモデル化にはアンサンブル学習の概念を利用し、その中でも特に Random Forests を用いたモデル化を検討する。Random Forests は多数の決定木を組み合わせることで高精度の学習を実現する。並列化による高速演算が可能であることや他クラス識別や回帰への適用が容易であるため、コンピュータビジョンの分野で注目を集めている。Random Forests を用いた人物モデル表現手法としては、図 2 のように全ての場所に共通なコンテキストを表現する決定木（共通決定木）と、特定の場所に固有な決定木（場所固有決定木）を用意し、ある場所について人物の有無を識別する際には、共通決定木と、その場所に対応する場所固有決定木の両方を用い、その結果を統合して最終的な結果を得る。

また、人物間の関係性のモデル化については、研究代表者がすでに実施していた、移動

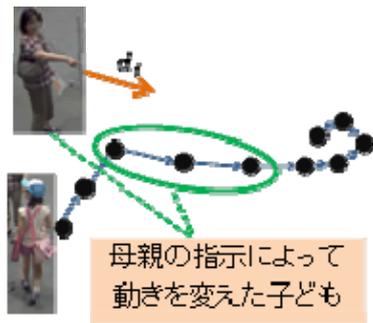


図 3 指差しによる人物関係モデル

軌跡を用いた人物関係モデリングの手法を
発展させ、移動軌跡だけではなく、人物の仕
草やジェスチャーも加味した方法を検討す
る。

これは図3のようにある一方の人物が他方に
指示をするときに典型的に現れ、図では母親
の指差しにしたがって、子どもがその歩行方
向を変える様子を示している。このような、
人物間の指示する-される関係をモデル化で
きれば、その社会的な関係の推定の第一歩と
なると考えられる。

これらの事項を検討するにあたり、研究初
年度には既存のデータセットを用いた基礎
検討を重点的に行い、それぞれのアイデアの
有効性を確認し、万が一不十分な場合は再検
討を行うこととした。

2年目には、1年目に検討したアルゴリ
ズムの検証・改良を行うため、映像データを
収集するシステムを構築した。研究代表者は、
既に屋外で映像を収集するシステムを保持
しているが、設置条件のバリエーションが十
分ではないため、それを補うための追加デー
タ取得を計画した。

4. 研究成果

- (1) まず、「場所依存コンテキストを用いた人
物検出」については、当初計画どおり、
Random Forests を用いたアンサンブル学
習を行い、決定木それぞれが「コンテク
スト」を表して人物検出を行う手法を考
案してその評価を行った。これにより、
当初計画していたそれぞれの木に違う情
報を表現させるアイデアが有効であるこ
とを示した。

そのための具体的なアプリケーション
として、オンラインでの物体追跡を行う
手法の検討・評価を行った。オンライン
物体追跡では、物体の見えの変化をどの
ように扱うかが問題になるが、ここでも
様々な見えに対応する決定木を用意し、
それらの組み合わせを動的に変更するこ
とで、頑健な追跡を実現している。

図4がその結果の一例であり、赤枠で囲
んでいるのがその時点での追跡位置であ

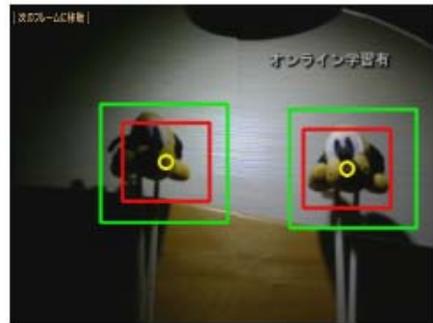


図 4 オンライン追跡結果

る。照明条件変化による見えの変化に対
しても頑健に追跡できていることがわか
る。

これらの検討の結果、当初検討してい
た決定木それぞれで場所共通、場所固有
のコンテキストを表現する方式だけでは
なく、一つの決定木で場所固有のコンテ
キストと場所共通のコンテキストを表現
する方式(図5:縦結合方式)も取りう
ることがわかってきた。この方式につい
ては当初想定していなかったため、論文
発表の形式とはなっていないが、今後の
研究につながるものとして特筆すべき事
項である。

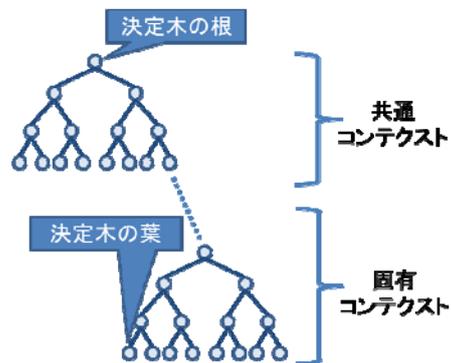


図 5 縦結合方式による表現

- (2) 本研究の評価を行うために非常に重要な
位置を占める大規模データベースの構築
については、計画当初は研究代表者が所
属していた奈良先端科学技術大学院大学
構内での取得を予定していた。しかし、
実質途中で代表者が研究機関を異動した
ため、当初予定とは異なる形態でデー
タを収集した。

研究代表者が本課題と別に参加してい
る JST CREST プロジェクトと協力し、
CREST プロジェクトのアウトリーチ活動
として行った科学館での展示風景をカメ
ラで撮影し、そこでの来場者映像を収集
した。5日間の展示で述べ1000人以上
の来場者があったため、評価映像とし
ては十分なボリュームとなっている。

そこで収集した映像に対して、人物位置を人の手で入力して、本研究で用いるデータとして利用可能とした。その例を図6に示す。



図 6 人手で与えた人物位置

- (3) 「人物相互関係モデリング」については、当初想定どおり、人物関係を推定するモデルを設計しその学習方法を検討した。その手法では、指示動作などの人物同士のインタラクションに着目して、それが多く発生しているペアは何らかの関係があるとするものである。
- (2) で示したデータについて実験を行ったところ、意図したとおり、親子などの人間関係を推定できることが示された。今後の課題としては、指示動作などははっきりとした動きが見えない場合の対処があげられる。
- (4) これらの研究から派生した成果として、Random Forests のチュートリアル講演を行った。Random Forests は近年コンピュータビジョン分野で広く用いられるようになってきているため、初学者の関心も高くなっている。そこで、初学者の理解を助けるためのチュートリアルを情報処理学会 CVIM 研究会 (2012 年 5 月) で行った。これにより、コンピュータビジョン分野全体の技術レベルが上がることを期待され、科研費としても大きな成果であると考えている。

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)

[学会発表] (計 3 件)

- ① 田口 弘貴, 波部 斉: 弱識別器の適応的学習と選択によるオンライン物体追跡, 情報処理学会 CVIM 研究会, 東京, 2013 年 5 月 30 日
- ② 波部 斉: ランダムフォレスト, 情報処理学会 CVIM 研究会, 愛知, 2012 年 5 月 23 日
- ③ 前 佑樹, 波部 斉, 柴田 智広: 機械学習を用いた屋外カメラによる駐車場監

視システムの開発, 電子情報通信学会 ITS 研究会, 北海道, 2012 年 2 月 20 日

[その他]

ホームページ等

<http://www.habe-lab.org/habe/>

6. 研究組織

(1) 研究代表者

波部 斉 (HABE HITOSHI)

近畿大学・理工学部・講師

研究者番号: 80346072