

複数方向から撮影された人物の手荷物所持判定に関する検討

浅井 康博[†] 西堀 研人[†] 高橋 友和^{††} 出口 大輔^{†††}

井手 一郎[†] 村瀬 洋[†]

[†] 名古屋大学 大学院情報科学研究科, 〒464-8601 愛知県名古屋市千種区不老町
^{††} 岐阜聖徳学園大学 経済情報学部, 〒500-8288 岐阜県岐阜市中鶯 1-38
^{†††} 名古屋大学 情報連携統括本部, 〒464-8601 愛知県名古屋市千種区不老町
E-mail: †{asa,y,nishiborik}@murase.m.is.nagoya-u.ac.jp, ††ttakahashi@gifu.shotoku.ac.jp,
†††ddeguchi@nagoya-u.jp, †{ide,murase}@is.nagoya-u.ac.jp

あらまし 近年, 公共空間への監視カメラの設置が進められている. 捜査などで監視カメラ映像を利用するにあたり, 膨大な映像中から対象の人物を探す作業が必要となる. しかし, 人手によるこれらの作業は膨大な時間と労力を要するため, 画像処理により人物検索を支援する技術が求められている. そこで我々は人物画像から手荷物所持の有無を判定することにより人物絞り込みの支援を行なうことを目的としている. 人物検索のためには検出漏れをいかに減らすかが重要となるが, 1枚の画像から手荷物所持の有無を判定する場合, 体により手荷物が隠れてしまい正しく判定できない場合がある. そこで, 本報告では複数方向から撮影された人物画像を利用した手荷物所持判定手法を提案する. これにより, 検出漏れの大幅な削減を実現した.

キーワード 監視カメラ, 手荷物, 所持判定, 複数方向

A study on detection of the type of a baggage utilizing multiple-viewpoints human images

Yasuhiro ASAI[†], Kento NISHIBORI[†], Tomokazu TAKAHASHI^{††}, Daisuke DEGUCHI^{†††},

Ichiro IDE[†], and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{††} Faculty of Economics and Information, Gifu Shotoku Gakuen University
Nakauzura 1-38, Gifu-shi, Gifu, 500-8288 Japan

^{†††} Information and Communications Headquarters, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

E-mail: †{asa,y,nishiborik}@murase.m.is.nagoya-u.ac.jp, ††ttakahashi@gifu.shotoku.ac.jp,
†††ddeguchi@nagoya-u.jp, †{ide,murase}@is.nagoya-u.ac.jp

Abstract Recently, the number of surveillance cameras in public space is increasing. We need to search a target person in an enormous amount of data. However, manual search consumes a lot of time and labor. Thus, image processing technology which supports manual search is required. Therefore, we aim at the detection of belongings from a human image so that human operators can filter specific persons. Though it is important for human searching system to reduce the failure of detection, baggage detection from a human image is sometimes difficult when a baggage is occluded by a human body. In this report, we propose a method for the detection of belongings utilizing multiple-viewpoints human images. We realized a large improvement of recall ratio.

Key words Surveillance camera, belongings, detection of belongings, multiple-viewpoints

1. まえがき

近年、防犯意識の高まりから、犯罪抑止効果として公共空間への監視カメラの設置が進められている。監視カメラ映像の解析は犯罪の防止や事件解決の糸口として期待されており、実際に犯人の特定や逮捕のきっかけとなった事件も多い。警視庁では繁華街等の防犯対策の一環として「街頭防犯カメラシステム」を導入している。平成 25 年中に警視庁本部において録画した 415 件の映像データのうち、320 件が検挙活動・事案の立件・解決等へ活用された [1]。監視カメラは実際に事件の発生した場所に設置されたものだけではなく、その近隣の監視カメラが逮捕につながる場合も増えている。このため、大量の監視カメラ映像の中から捜査対象の人物を探す作業が必要となる。監視カメラの数が増えるほど膨大な人手と時間を要するため、このような人物検索を画像処理により支援する技術が求められている。

性別や年齢などとともに、手荷物の情報は目撃情報などにもよく使われる属性であり、人物検索の重要な手がかりになると考えられる。たとえば、監視カメラ映像に映った大勢の人物のうち、リュックサックやキャリーバッグなど、特定の種類の手荷物を所持する人物のみを絞り込むことができれば、人物検索のコストを大幅に削減することができる。実際に、2013 年のボストン・マラソン爆発事件では、監視カメラ映像と手荷物の情報が事件の早期解決の手がかりになっている [2]。

手荷物の種類による人物の絞り込みが実現できた際の効果を検証するために、手荷物に関する簡単な統計情報調査を実施した。中部国際空港において、15 分間、延べ 198 人の手荷物について目視により調査した。この調査により、図 1 のような結果が得られた。このグラフからわかるように、手荷物の種類を限定するだけで人物の絞り込みの大幅なコスト削減が期待できる。たとえば、追跡対象の人物がリュックサックを持っていることがわかった場合、86%の候補者を削減できる。

手荷物を対象とした研究は数多く存在するが、多くは手荷物の検出に関連したものであり、種類まで判別する研究は少ない。Damen ら [3] は手荷物の検出として姿勢検出や歩容認識に用いられる Temporal template を用いて人物シルエットから突出した部分を手荷物として検出している。このため、服装による突出と区別ができないという問題点がある。また、映像中から手荷物が所有者から一定の距離以上離れたことを検出する置き去り検出に関する研究は盛んである。Smith ら [4] は置き去りにされた手荷物は動かない、人物より見た目が小さいといった仮定を利用している。そのため、手荷物の情報を用いた人物の絞り込みには適用できない。

そこで我々はこれまで、図 2 のように手荷物の種類や撮影方向の違いによる手荷物の位置や大きさの違いに着目することにより、1 枚の人物画像から種類の判別を含めて手荷物の所持判定を実現すべく取り組んできた [5]。人物検索支援では、いかに検出漏れを減らせるかが重要となる。しかし、体により手荷物が隠れてしまい、所持の有無を判定できない場合が存在した。そのため、複数方向からの人物画像を利用して判定を行った場合の判定精度を調査し、考察した結果を報告する。

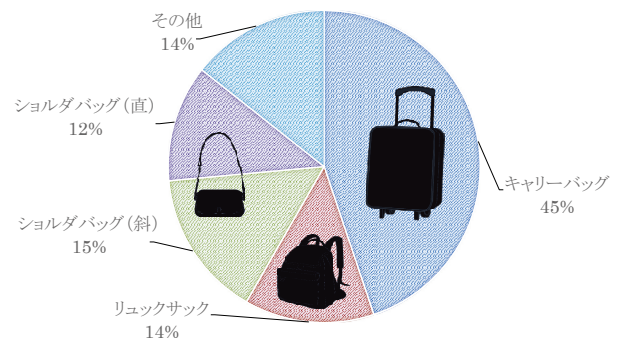


図 1 各手荷物の全体に占める割合

以降、図 2(d) のような斜め掛けのショルダバッグをショルダバッグ(斜)、図 2(e) のような片方の肩に真っ直ぐかけたものをショルダバッグ(直)として表記する。

2. 手荷物の所持判定手法

提案する手荷物の所持判定手法について述べる。図 3 に提案手法の処理の流れを示す。本手法は、(1) 判定対象領域の設定、(2) 識別器の学習、および (3) 手荷物の所持判定の 3 つの処理に分けられる。このうち、(1)、(2) はオフライン処理である。

各処理における具体的な処理は以下のとおりである。

2.1 判定対象領域の設定方法

手荷物を所持した人物画像を多数集め、手荷物の種類ごと、撮影方向ごとに所持判定の対象領域を設定する。そのために、各画像中の人物領域と手荷物領域の相対的な位置と大きさの情報を利用する。そこでまず手荷物を所持した人物画像を多数用意する。次に、これらの画像中の人物の大きさを正規化する。最後に、正規化した人物領域を基準として、全画像中の手荷物領域を包含するような領域を求め、これを判定対象領域とする。判定対象領域は手荷物の種類ごと、撮影方向ごとに独立して設定する。

2.2 識別器の構築

学習用画像を用いて手荷物の種類ごと、撮影方向ごとに識別器を学習する。すなわち、(手荷物の種類数 × 撮影方向数) 個の識別器を構築する。学習用画像として、正事例(判定対象となる手荷物を所持している人物画像)と負事例(判定対象となる手荷物を所持していない人物画像)を用意する。これらの画像に対して、判定対象となる手荷物の種類と撮影方向に応じた判定対象領域から画像特徴量を抽出し、それらを用いて識別器を学習する。

2.3 手荷物の所持判定

学習した識別器を用いて、異なる方向から撮影された 2 枚の入力画像中の人物が判定対象の手荷物を所持しているか否かを判定する。ここで、それぞれの入力画像の人物の撮影方向は既知であるとする。まず、学習と同様に判定対象領域から特徴量を抽出する。次に、その特徴量をその方向に対する識別器に入力し、手荷物所持の有無を判定結果として得る。これを 2 枚の入力画像それぞれに対して適用し、それらの結果を統合する。実際に犯罪捜査での利用を考えると、いかに再現率を高くする

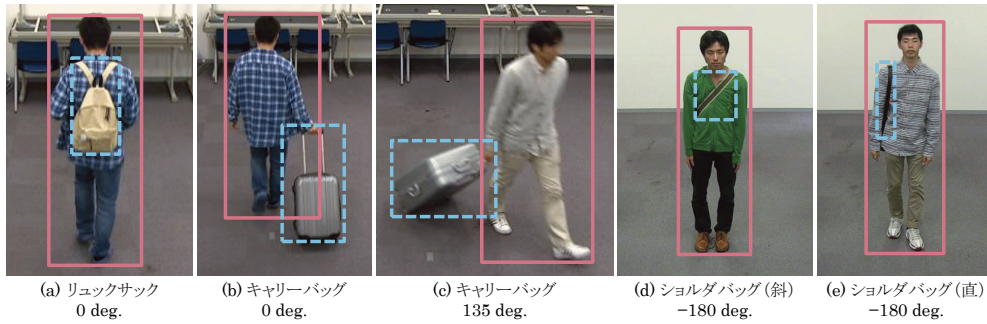


図 2 手荷物領域の種類と撮影方向による違い

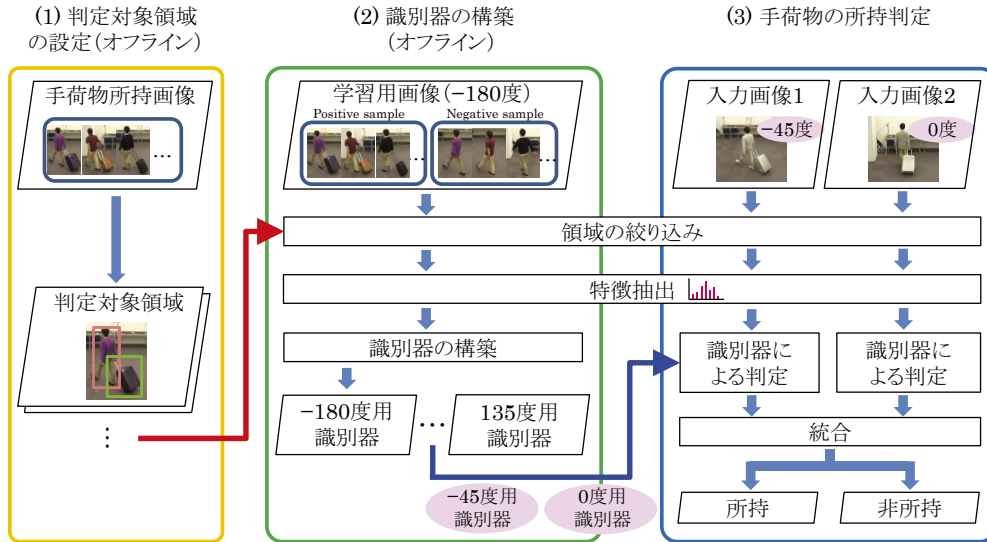


図 3 提案手法の処理の流れ (キャリーバッグを対象とした場合)

か、すなわち、いかに検出漏れを少なくするかが重要となるため、本手法では各方向の判定結果の OR をその人物の手荷物所持の有無の結果として採用する。

2.4 画像特徴量と識別器

本手法では手荷物としてキャリーバッグ、リュックサック、ショルダバッグ (斜), ショルダバッグ (直) の 4 種類を想定する。ここで、手荷物の種類により有効な画像特徴量が異なると考えられる。ただし、どの手荷物についても人物および手荷物の輪郭に最も特徴が表れやすいと考えられるので、エッジ特徴である HOG (Histograms of Oriented Gradients) [6] を利用する。キャリーバッグとリュックサックについては、持つ人や手荷物の大きさの違いにより、位置のばらつきが大きいと考えられる。そこで、位置ずれを吸収するために局所特徴として HOG を用いた BoF (Bag of Features) [7] を特徴量として採用する。BoF はベクトル量子化された局所特徴を用いることで画像を局所特徴の集合として考えるものである。ショルダバッグでは紐のラインが最もショルダバッグらしい特徴が出ると考えられる。また、人体領域を基準とすれば持つ人による位置ずれが発生しにくい。これらの理由により HOG 特徴をそのまま利用する。

識別器には、一般に 2 クラス識別問題に対して高い性能を持つとされる SVM (Support Vector Machine) 識別器を用いる。

3. 手荷物の所持判定実験

提案手法の有効性を確認するために、2 方向から撮影された人物画像を用いた、手荷物の所持判定実験を行なった。

3.1 実験条件

実験を行なうにあたり、データセットを作成した。駅等にある一般的な監視カメラは、3~4m の高さの場所に、30~40 度の俯角で設置されている。本実験で作成したデータセットにおいても、監視カメラ映像の利用を想定して、同様の条件になるようにカメラを設置し撮影した。手荷物としてキャリーバッグ、リュックサック、ショルダバッグ (斜), ショルダバッグ (直) の 4 種類を用意し、これらの手荷物の所持の有無、人物、撮影方向を変えながら、合計 2,112 枚の画像を撮影した。図 4 にデータセット中の画像の例を示す。上から順にキャリーバッグ、リュックサック、ショルダバッグ (斜), ショルダバッグ (直) である。これらの画像に対して、人手で手荷物所持の有無、人物領域と手荷物領域の位置、人物 ID、撮影方向の情報を付与した。なお、手荷物領域は図 2 の破線に示されるような領域とする。ショルダバッグについては紐のラインが本体部分の形によらず最もショルダバッグらしい特徴が出やすいと考えられるため、紐の部分のみを対象とした。

2 方向から撮影された人物画像から、手荷物の所持の有無を

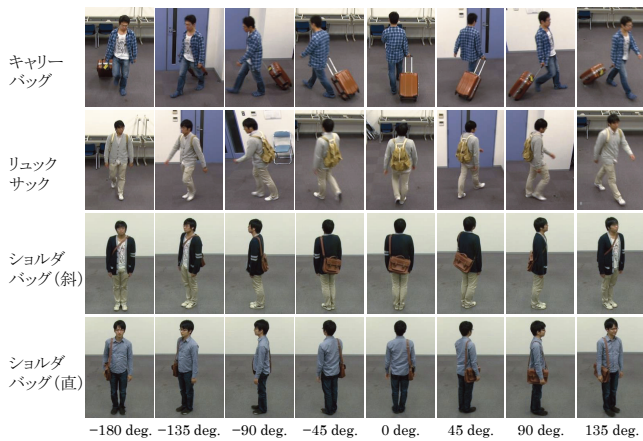


図4 データセット中の画像の例

表1 実験結果

手法	適合率	再現率	F 値
比較手法 (1 方向)	0.666	0.827	0.736
提案手法 (2 方向)	0.682	0.965	0.797

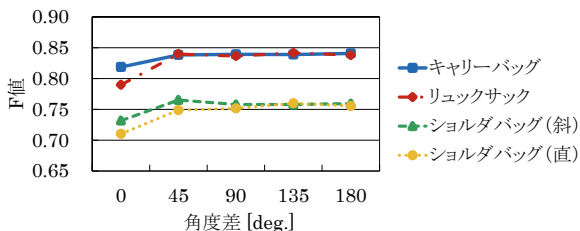


図5 角度差ごとの精度の違い

判定する実験を行なった。比較手法として、1方向の画像から手荷物所持の有無を判定する手法を用いた。評価には再現率と適合率、およびその総合的な評価であるF値を用いた。

3.2 実験結果および考察

表1に手法の手荷物の所持判定の正解率の比較を示す。比較手法は8方向の結果を、提案手法は2方向のすべての組み合わせの結果を平均したものである。提案手法による結果がすべての指標において最も高く、その有効性を確認した。特に再現率は、1方向のみを用いる比較手法の場合は0.827であるのに対し、2方向を用いる提案手法は0.965と大幅に向上した。これにより、人物検索システムの実現のためには不可欠な、検出漏れの削減を実現できたことを確認した。

図5は2枚の人物画像の角度差ごとの精度の違いを示す。角度差0度は1方向のみの画像から所持の有無を判定することと等価である。このグラフから、2方向からの人物画像を用いることにより、どの角度差でも1方向のみを用いる場合より高い精度であることがわかる。

図6に、リュックサック識別器において、同じ人物の2方向の結果を統合することにより判定に成功した例を示す。1方向のみで判定した場合、図6(b)ではリュックサックは服に埋もれてしまい、正しい判定は困難であったが、図6(a)のような比較的判定が簡単な別の方向からの結果を用いることで最終的な結果としては正しい判定を行なうことができた。ただし、この



(a) True Positive (1方向のみの結果) (b) False Negative (1方向のみの結果)

図6 リュックサック識別器での判定成功例

手法では検出漏れは大きく減らすことができるが、過検出は増加すると考えられる。そのため、それぞれの方向における判定精度を考慮したうえで2方向の結果を統合することにより、より正確な判定ができると考えられる。

4. まとめ

本報告では、監視カメラ映像中の人物検索を支援することを目的として、画像処理により人物画像から特定の手荷物の所持の有無を判定する手法を提案した。異なる方向から撮影された2枚の人物画像から、手荷物所持の有無を判定する実験を行ない、その結果から提案手法の有効性を確認した。具体的には、1枚では手荷物が体に隠れていることにより判定が難しい場合でも、異なる方向から撮影された複数の画像を用いることにより、高い精度で判定できた。

今後の課題として、人物の撮影方向を推定する手法の導入や、手荷物の色情報などのさらに詳細な情報による検索などがあげられる。

謝辞 本研究の一部は、文部科学省先導的創造科学技術開発費補助金、および科学技術研究費補助金による。

文献

- [1] “街頭防犯カメラシステム,” <http://www.keishicho.metro.tokyo.jp/seian/gaitoucamera/gaitoucamera.htm>, referred on 13 Dec. 2014.
- [2] M. Arsenault, B. Bender, M.J. Valencia, and M. Cramer, “Edging toward normal with healing still to do: FBI was warned 2 years ago of alleged bomber’s radical shift,” 2013. Boston Globe, referred on 21 Apr. 2013.
- [3] D. Damen and D. Hogg, “Detecting carried objects in short video sequences,” Proc. 10th European Conf. on Computer Vision (ECCV 2008), Part III, Lecture Notes in Computer Science, vol.5304, pp.154–167, Oct. 2008.
- [4] K. Smith, P. Quelhas, and D. Gatica-Perez, “Detecting abandoned luggage items in a public space,” Proc. 9th IEEE Int. Workshop on Performance Evaluation in Tracking and Surveillance (PETS ’06), pp.75–82, June 2006.
- [5] 浅井康博, 出口大輔, 高橋友和, 井手一郎, 村瀬洋, “手荷物の種類と向きを考慮した人物画像からの手荷物の所持判定に関する検討,” 信学技報, PRMU2013-87, Dec. 2013.
- [6] N. Dalal and W. Triggs, “Histograms of oriented gradients for human detection,” Proc. 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol.1, pp.886–893, June 2005.
- [7] G. Csurka, C. Bray, C. Dance, and L. Fan, “Visual categorization with bags of keypoints,” Proc. ECCV 2004 Workshop on Statistical Learning in Computer Vision, pp.59–74, May 2004.