

Shinpuhkan データセットの拡張と関連ツールの公開に向けて

川西 康友[†] 飯山 将晃^{††} 椋木 雅之^{†††} 美濃 導彦^{††} 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

^{††} 京都大学 学術情報メディアセンター 〒606-8501 京都市左京区吉田本町

^{†††} 宮崎大学 工学教育研究部 情報システム工学科 〒889-2192 宮崎市学園木花台西1丁目1番地

E-mail: †{kawanishi,murase}@is.nagoya-u.ac.jp, ††{iiyama,minoh}@mm.media.kyoto-u.ac.jp,

†††mukunoki@cs.miyazaki-u.ac.jp

あらまし 人物照合やカメラ間追跡の研究を目的とした大規模データセットの需要は高いが、カメラ台数、画像枚数、人数の全てについて規模の大きいデータセットは存在しない。本発表では、これまでに公開した Shinpuhkan 2014 から人数と画像枚数を大幅に増加させた、現在準備中の新しい Shinpuhkan データセットを紹介し、データセットについてアノテーションの質と量の観点から議論する。また、多カメラ間人物追跡アノテーションツール及び、既存のデータから仮想的なデータセットを生成するツールについても紹介し、必要な機能について議論する。

キーワード カメラ間人物追跡, データセット, アノテーションツール, 仮想データセット生成

Toward Extension of Shinpuhkan Dataset and Release of Related Tools

Yasutomo KAWANISHI[†], Masaaki IIYAMA^{††}, Masayuki MUKUNOKI^{†††}, Michihiko MINOH^{††},
and Hiroshi MURASE[†]

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

^{††} Academic Center for Computing and Media Studies, Kyoto University Yoshidahonmachi, Sakyo-ku,
Kyoto, 606-8501, Japan

^{†††} Department of Computer Science and System Engineering, Miyazaki University 1-1, Gakuen
Kibanadai-nishi, Miyazaki, 889-2192, Japan

E-mail: †{kawanishi,murase}@is.nagoya-u.ac.jp, ††{iiyama,minoh}@mm.media.kyoto-u.ac.jp,

†††mukunoki@cs.miyazaki-u.ac.jp

Abstract Some public datasets for person re-identification or tracking across cameras are published but there is no dataset which has a large number of cameras, images, and pedestrians. In this paper, we introduce our current activity for a new dataset, which has more pedestrians and images than our previous Shinpuhkan 2014 dataset, and discuss requirements for the dataset in terms of the amount and quality of the dataset. We also introduce our annotation tool for pedestrian tracking across cameras and virtual dataset generation tool and discuss their desirable functions.

Key words Person Re-identification, Tracking across Cameras, Annotation Tool, Virtual Dataset Generation

1. はじめに

我々の身の回りには多くの防犯カメラが設置されている。カメラ台数は膨大になっており、全てを人手でチェックし監視することは不可能である。したがって、監視の自動化に対する需要が高まっている。

防犯において、人の移動を把握することは最も重要なことの

ひとつである。カメラに視野の重複があれば、物理的な位置関係の制約から、複数カメラにわたる人の自動的な追跡は容易であるが、視野重複がない場合、複数カメラ視野間での画像の照合により、同一人物か否かの同定をする必要がある。この技術は“人物再同定 (Person Re-identification)”と呼ばれており、監視の自動化のための重要な技術のひとつである。

視野重複のないカメラ視野間における人物追跡・照合につい

ては、近年様々な手法が提案されている。特に、各カメラ視野内での人物追跡によって得られる複数枚の画像系列を利用した人物照合手法が注目されている。また、2台のカメラ視野間だけでなく、多数のカメラによって構成されるカメラネットワーク上での人物追跡も研究が盛んである。これは、複数カメラ視野にわたる人物追跡と呼ばれる。

画像処理・コンピュータビジョンの分野においては、大規模な公開データセットを用いた従来手法との比較が盛んに行われている。例えば一般物体認識の研究では ImageNet [1] などの大規模なデータセットが公開され、多くの研究がなされている。

歩行者検出などを目的とした幾つかのデータセット [2], [3] は存在するものの、プライバシー問題があるため、多数のカメラで多くの人を合法的に撮影したデータセットは少ない。また、複数のカメラ映像に映る多数の人物について全てアノテーションを付与することは非常に手間がかかるため、複数カメラを用いたアノテーション付きの既存のデータセットは、カメラ台数や人数が少ないものが多い。

複数のカメラ視野にわたる人物追跡・照合を評価する場合、様々な人が様々なルートで歩き回っているのを撮影した複数のデータセットがあるべきである。しかし、1回撮影しただけのデータセットでは、その1パターンしかデータがない。同じ人物を、異なるルートで歩かせた複数の歩行パターンがあれば良いが、そうしたデータセットは存在しない。

人物照合や複数カメラにわたる人物追跡のためのよいデータセットは、カメラ台数が多く、写っている人数が多く、1カメラ1人当たりの画像枚数が多いことが重要である。また、複数の歩行パターンのデータがあることが望ましい。本稿では、我々が以前に公開した Shinpuhkan 2014 [4] を拡張した新たなデータセット構築に向けての取り組みの紹介と、そのデータセット構築のためのアノテーションツールについて述べる。また、人物の歩行をシミュレートすることにより、1つのデータセットから、歩行パターンの異なる複数のデータセットを仮想的に生成するツールについても紹介する。

2. 関連研究

これまでに、複数のカメラ視野にわたる人物追跡・照合の研究を目的とした人物画像データセットがいくつか公開されている。これら既存のデータセットについてまとめる。

最もよく使われている人物画像データセットのひとつとして、VIPeR (Viewpoint Invariant Pedestrian Recognition) と呼ばれるデータセットがある [5]。このデータセットは 632 人の人を対象として、各人について異なるカメラ視野で撮影された 2 枚の画像があり、画像の大きさは 48×128 に正規化されている。このデータセットには 1 人 1 カメラ視野につき 1 枚ずつしか画像がないため、画像集合や系列を用いる手法には使えない。

ETHZ データセット [6] は歩行者検出に用いられるデータセットであるが、Schwartz ら [7] が人物照合に使えるようにアノテーションを付与したことで、人物照合の評価に用いられるようになった。このデータセットは移動カメラによる 3 つの映像系列からなり、それぞれ 83 人を含む 1,000 フレーム、35 人

を含む 451 フレーム、28 人を含む 354 フレームの画像からなる。このデータセットでは、各人物は 1 カメラでのみ撮影されているため、その人物の映り始めと映り終わりの方からそれぞれ画像を選択し、人物照合の評価に用いられている。

2008 i-LIDS Multiple-Camera Tracking Scenario (MCTS) [8] というデータセットに対し、Zheng ら [9] が人物照合用にアノテーションを付与して公開したものが、i-LIDS と呼ばれるデータセットである。このデータセットは 119 人の人物画像 476 枚からなる。各人物について、視野重複のない 4 台のカメラで撮影された画像があり、それぞれの画像は 64×128 画素に正規化されている。このデータセットも、各視野について限られた枚数の画像しかないのが欠点であるといえる。

以下の 2 つのデータセットは、同じ論文 [10] で公開されたものであり、前述の i-LIDS MCTS データセットに対して異なるアノテーションを付与したものである。iLIDS-MA データセットは 2 台のカメラで撮影された 40 人の人物画像からなる。各人物について 46 フレームずつが手動でアノテーションされており、全部で 3,680 枚の画像からなる。手動でアノテーションする代わりに、iLIDS-AA データセットでは、HOG 特徴量に基づく人物検出とトラッキングによる自動的なアノテーションがなされている。このデータセットは 100 人の人物を含み、合計 10,754 枚の画像からなる。自動アノテーションであるため、アノテーションは不完全であり、ノイズを含むため、iLIDS-MA よりも実践的なデータセットとなっている。

3DPeS データセット [11] も人物照合に用いられるデータセットである。このデータセットは 6 台のカメラ映像からなり、200 人の人物について合計 200,000 枚の画像を含む。個々の人物の画像はかなり多いが、各人物は平均で 2 台のカメラにしか写っていない。そのため、2 カメラ間人物照合には利用できるが、多カメラ間人物追跡に利用することは難しい。

Person Re-Id データセット [12] は、2 カメラ間の人物照合を目的として作成されたデータセットである。2 台のカメラ映像から切り出された人物画像列からなり、一方のカメラからは 475 人が、もう一方のカメラからは 856 人が切り出されており、245 人が両方のカメラに写っている。大きく遮蔽された人物を除いて、一方のカメラからは 385 人、もう一方のカメラからは 749 人が評価に利用され、両方に写った 200 人のみを精度良く対応付けすることができるかという評価をする。

CAVIAR4REID データセット [13] は、ショッピングセンターで撮影された人物検出・追跡の評価に用いられるデータセットである CAVIAR データセットに対し、人物照合用のアノテーションを施したものである。このデータセットは、視点の異なる 2 台のカメラで撮影された 26 本の映像からなり、歩行者や買い物中の人物が写っている。そのなかから、両方のカメラに写っている 50 人と、片方のカメラにしか写っていない 22 人の合計 72 人を選び、その人物を切り出した画像の中から、解像度、照明条件、オクルージョン、姿勢などが様々に異なる画像を選択してある。このデータセットは実環境で収集したデータセットであり、かつ、VIPeR などのデータセットと異なり各人物について複数枚の画像を含む、非常にチャレンジングなデー

タセットである。

SAIVT-SoftBio データセット [14] は 8 台のカメラ映像からなるデータセットである。各映像は 25fps で撮影され、サイズは 704 × 576 画素である。このデータセットには 150 人の人物が含まれ、各人物は殆ど全てのカメラに撮影されている。各人物の画像はおよそ 400 枚である。

Dana36 データセット [15] は 36 台のカメラで撮影された 23,000 枚以上の画像からなるデータセットである。36 台のカメラは、27 台のカメラが屋外、9 台のカメラが屋内に設置されている。撮影対象は 15 人の人物と 9 台の自動車であり、様々な解像度で撮影されている。また、15 人の人物は決められたルートに沿って歩いている。

実環境でデータセットを収集することは、プライバシーの問題があり非常に困難である。また、膨大な映像に対してアノテーションを手動で行うのも非常に大変な作業である。Starzyk ら [16] は仮想環境中の 3D モデルのシミュレーションにより、人が歩行している環境を撮影したデータセットを生成する仕組みを提案しており、Python による実装が公開されている。これを用いれば、いくらでも好きなだけ画像を生成することが可能であるが、現状 5 人分の人物モデルしか公開されていない。もし人数を増やしたい場合は、テクスチャを自分で追加する必要がある。

実環境で膨大な人数のデータセットを構築した事例として、CUHK01 [17], CUHK02 [18], CUHK03 [19] という、Li らによって公開された人物照合用データセットがあり、徐々に拡充されてきている。CUHK01 は Campus データセットと呼ばれていたもので、VIPeR データセットと同じく、971 人に対して 2 台のカメラで 1 枚ずつ撮影された画像からなるデータセットである。CUHK02 は 1,816 人を、5 つのカメラペア（計 10 台のカメラ）で撮影したデータセットである。1 つのカメラペアを構成する 2 台のカメラは同一地点を観測しており、人物 1 人カメラペア 1 つにつき、2 枚の画像がある。CUHK03 は 1,360 人を、6 台のカメラで撮影したデータセットであり、人物 1 人あたり、各カメラにつき平均 4.8 枚の画像がある。これらのデータセットは人数が多いことから、人物のバリエーションが豊富であり、Deep Learning など最近の機械学習を用いた手法でよく利用されている。

Shinpuhkan 2014 データセット [4] は、我々が以前に公開したデータセットである。このデータセットは、16 台のカメラで観測された 24 人の画像約 22,000 枚からなる。個々の人物が、全てのカメラの視野内を、異なる方向に何度も通過したのを撮影したデータセットであることが、このデータセットの特色である。このデータセットについて詳細は次節で述べる。

3. データセットの拡張

人物検索や複数カメラにわたる人物追跡、大規模な人物照合を行う場合、複数のカメラで撮影された多数のトラックレットから構成されるデータセットがあると良い。ここで、トラックレットとは、ある人物がカメラ視野内に現れてからカメラ視野外へ消えるまでに観測される画像列のことである。我々がこれ



図 1 データセット中の人物。

表 1 カメラ毎の 1 人あたりのトラックレット数。

cameraID	Cam01	Cam02	Cam03	Cam04
#Tracklets/person	4	8	4	6
cameraID	Cam05	Cam06	Cam07	Cam08
#Tracklets/person	4	4	6	4
cameraID	Cam09	Cam10	Cam11	Cam12
#Tracklets/person	8	8	8	4
cameraID	Cam13	Cam14	Cam15	Cam16
#Tracklets/person	4	6	4	4

までに発表した Shinpuhkan 2014 データセットは、トラックレットの数が多いことが特色であるが、さらにこのデータセットを拡張し、新たなデータセットを構築しているところである。以下では、Shinpuhkan 2014 データセットの概要を述べ、拡張すべき点についてまとめる。

3.1 Shinpuhkan 2014 データセットの概要

Shinpuhkan 2014 データセットは、16 台のカメラが設置されたショッピングモール内を、募集した 20 歳前後のおよそ 100 人の被験者が、決められたルートに従って歩行しているのを撮影し、そのうち 24 人についてアノテーションを付与したものである。Shinpuhkan 2014 データセットに含まれる全人物の画像を図 1 に示す。

収録は、ショッピングモールがオープンする直前の一般客がいない状況で、予め決められたルート（図 2）に従って歩く、というシナリオで 2 回行った。1 人の人物がこのルートを 1 周すると、16 台のカメラの視野を合計 43 回横切ることになる。全ての人物が同一ルートを 2 周しているため、データセット内には 1 人あたり 86 個のトラックレットが入っている。各カメラでの 1 人あたりのトラックレット数は表 1 の通りである。

トラックレット内の各画像は人物を正しく囲うようにバウンディングボックスを設定し、すべて手動で切り出し、48 × 128 画素に正規化してある。オクルージョンを含む画像は全て除去

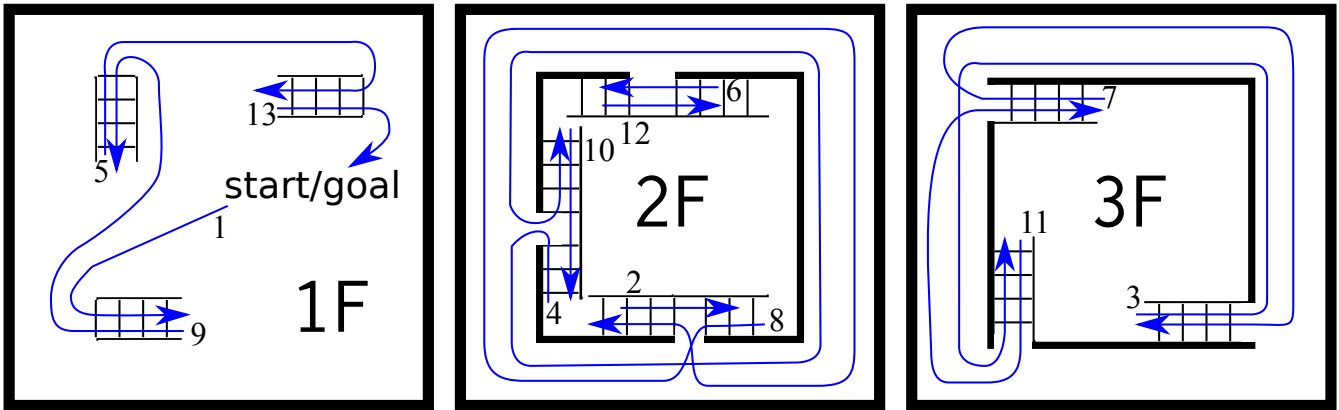


図2 データセットの歩行シナリオ. 全人物は start から出発し、矢印の番号順に施設内で goal まで移動した。

し、各トラックレット中から8枚程度の画像を選び出しデータセットに収録してある。

3.2 Shinpuhkan 2014 データセットの不足点

Shinpuhkan 2014 データセットには、以下の不十分な点がある。

- 人数が少ない (24 人)
- トラックレットごとの画像枚数が少ない (8 枚程度)

人物照合の評価をする場合、人物間の十分なバリエーションがあることが望ましいが、そのためには、データセットに含まれる人数が多いことが必要である。Shinpuhkan 2014 データセットでは 24 人の人物しか含まれていないため、評価に用いることが難しい、という報告 [20] がある。

また、Multi-shot re-identification と呼ばれる、複数枚の画像集合同士の照合に基づく人物照合をする場合、ある人物のバリエーションを表現するために1つの画像集合内、ここでは1つのトラックレットに含まれる画像の枚数が多い方が望ましい。Shinpuhkan 2014 データセットでは、1トラックレットにつき8枚程度が公開されているが、高い次元数の特徴空間中で、人物のバリエーションを表現するためには十分とは言い難い。

この2点について、今後のデータセット拡張により、対処していきたいと考えている。

3.3 拡張の方針

データセットに含まれる人数が少ない点については、現在、未アノテーションの人物についてアノテーション作業を進めている。今後、100人程度のデータセットになる予定である。

各トラックレットに含まれる画像については、人物を囲む矩形をできるだけ正確に全てのフレームに対して付与することが理想的だと考えている。正確に矩形のアノテーションを付与しておけば、後の加工により、自動的な検出による検出枠のズレを再現することもでき、難しいデータセットにすることも可能である。しかし、全てのフレームに正確なアノテーションを付与することは、作業コストが膨大であり困難である。

そこで、数フレームおきにアノテーションを付与し、補間によって残りのフレームに自動的にアノテーションをすることで、少ない工数で多数のフレームへのアノテーションを実現する。

現在は数フレームおきに付与した手動アノテーションのキュービック補間によりフレーム間の自動アノテーションを実現している。しかし、補間による自動アノテーションは必ずしも正確であるとは限らない。アノテーションの精度と手動アノテーションの作業量はトレードオフであるため、精度を高めるためには、手動でアノテーションするフレーム数を増やす、もしくは補間アルゴリズムの改善が必要である。

アノテーションの質でもう1つ問題になるのが、オクルージョンである。手動でアノテーションを付与する場合、その人物がオクルージョンされているか、他の人をオクルージョンしているかの判断ができるため、オクルージョンの有無をアノテーションに追加することは容易である。しかし、補間による自動的なアノテーションをした場合、オクルージョンの有無を判定することは難しい。そのため、オクルージョンの有無の情報をアノテーションに追加するためには、全てのフレームにわたり目視で確認する必要がある。

これらの問題があるので、データセットを公開する際には、アノテーションの質を保つため、各フレームに対して、手動で付与したアノテーションなのか、補間により自動で付与したアノテーションなのかを表すフラグを付与する予定である。

4. データセット関連ツール

4.1 アノテーションツール

複数カメラにわたる人物追跡や照合用のデータセットを作成する場合、多数のカメラの映像に対し、各カメラ視野内での多人数の追跡及び、カメラ間を移動する多人数の対応付けを手動で行い、アノテーションを付与する必要があるが、これは大変な作業である。

そこで我々は、こうしたアノテーション作業を効率化するために、アノテーションツールを作成した。このアノテーションツールがもつ機能を以下にまとめる。

- 各カメラ視野内での人物追跡
 - 人物を囲む矩形の付与
 - 各人物に対する ID の付与
 - トラックレットに対する ID の付与

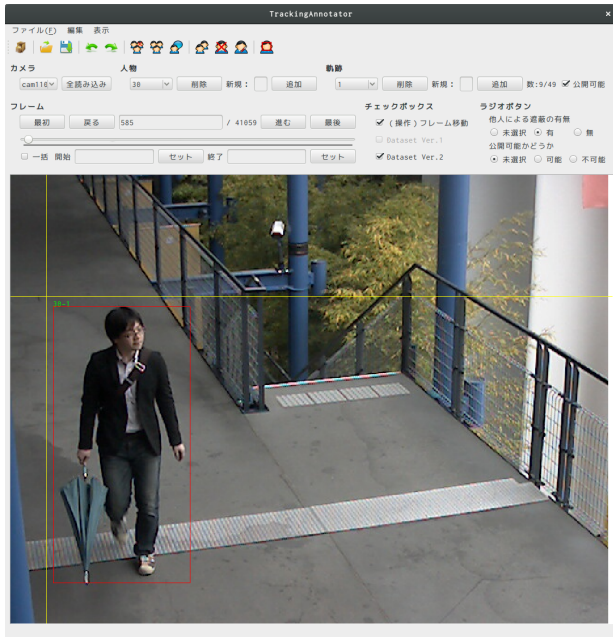


図 3 あるカメラ映像に対する人物追跡のアノテーション作業の様子。



図 4 あるトラックレットを表示した時の表示。

- オクルージョンの有無を付与
- フレーム間の補間
- 複数カメラで観測された同一人物に対する同一 ID の付与
- 各人物をデータセットに収録するか否かのフラグの付与
- 可視化
- 人物一覧の表示
- トラックレット中の画像一覧の表示

アノテーション作業中の画面を図 3 に示す。また、トラックレット中の画像一覧を表示した画面を図 4 に示す。この一覧を見ることで、人物を囲う矩形が大きくずれているものを見つけ、修正することが出来る。

このアノテーションツールは、Qt を使って実装されており、Windows, Mac OS X, GNU/Linux のどの環境でも、Qt 等のライブラリがインストールされていれば動作する。今後、ソースコード及びドキュメントを整理したうえでオープンソースで公開する予定である。

4.2 仮想データセット生成ツール

複数カメラ視野にわたる人物追跡・照合でポイントとなるのは、画像の類似度及び、カメラ視野間の移動にかかる時間と、カメラ視野の隣接関係である。そのため、複数カメラ視野にわたる人物追跡・照合の手法を評価する場合、様々な配置でカメラが設置された環境を、様々な人が様々なルートで歩き回っているのを撮影した複数のデータセットを用いるべきである。し

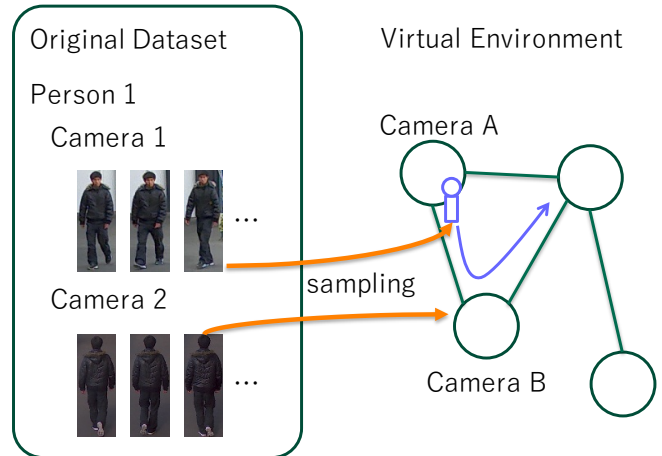


図 5 仮想環境中でのシミュレーションによる仮想データセット生成。

かし、一般に公開されているデータセットは、一度に撮影した 1 パターンしかデータがない。同じ人物を、異なるルートで歩かせ、かつそれを様々な配置パターンでカメラで撮影したデータがあれば良いが、そうしたデータセットは存在しない。

そこで、我々は 1 つの実在するデータセットから、仮想的なデータセットを生成するツールを作成した。この生成ツールでは、仮想環境中でランダムに仮想カメラを配置し、その環境内で歩行者をランダムに歩かせる。そして、各歩行者と既存のデータセット中の人物を、仮想環境中のカメラとデータセット中のカメラを対応付けておき、ある歩行者がある仮想カメラ視野内に入ったときに、対応する人物の画像をデータセットからサンプリングする (図 5)。

この処理により、様々な歩行パターンでのデータセットを仮想的に生成することができる。

このツールは PRMU アルゴリズムコンテスト用に既に公開してあるが、ドキュメントを整えた上で一般に向けて公開する予定である。

5. おわりに

本稿では、Shinpuhkan 2014 データセットを拡張した Shinpuhkan 2017 データセット構築に向けた取り組みについて紹介した。また、多カメラ間人物追跡のために必要なアノテーションについて議論し、我々が開発したデータセット構築のためのアノテーションツールを紹介した。さらに、カメラ配置と人の移動をシミュレーションにより仮想的に生成することで新たなデータセットを生成するデータセット生成ツールの紹介をした。このデータセット生成ツールは、PRMU アルゴリズムコンテストでも利用したが、その詳細については PRMU アルゴリズムコンテストの報告にて詳述する。

今後、データセットのアノテーション作業を進め、公開する予定である。また、その公開データセットを用いたコンテストを開催したい。

謝辞 本研究は文科省、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「環境適応型で実用的な人物照合システム」の一環として実施したものである。

- [1] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg and L. Fei-Fei: “ImageNet Large Scale Visual Recognition Challenge”, *International Journal of Computer Vision (IJCV)*, **115**, 3, pp. 211–252 (2015).
- [2] M. Enzweiler and D. Gavrilu: “Monocular pedestrian detection: Survey and experiments”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**, 12, pp. 2179–2195 (2009).
- [3] P. Dollár, C. Wojek, B. Schiele and P. Perona: “Pedestrian detection: an evaluation of the state of the art.”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**, 4, pp. 743–61 (2012).
- [4] Y. Kawanishi, Y. Wu, M. Mukunoki and M. Minoh: “Shinpuhkan2014: A Multi-Camera Pedestrian Dataset for Tracking People across Multiple Cameras”, *Proc. of the 20th Korea-Japan Joint Workshop on Frontiers of Computer Vision, FCV2014* (2014).
- [5] D. Gray, S. Brennan and H. Tao: “Evaluating Appearance Models for Recognition, Reacquisition, and Tracking”, *Proc. of 10th International Workshop on Performance Evaluation for Tracking and Surveillance (PETS)*, Vol. 3, pp. 41–47 (2007).
- [6] A. Ess and B. Leibe: “Depth and Appearance for Mobile Scene Analysis”, *Proc. of International Conference on Computer Vision* (2007).
- [7] W. R. Schwartz and L. S. Davis: “Learning Discriminative Appearance-Based Models Using Partial Least Squares”, *Proc. of 2009 XXII Brazilian Symposium on Computer Graphics and Image Processing, Ieee*, pp. 322–329 (2009).
- [8] H. Office: “iLIDS datasets”. <https://www.gov.uk/imagery-library-for-intelligent-detection-systems>.
- [9] W.-S. Zheng, S. Gong and T. Xiang: “Associating Groups of People”, *Proc. of the 20th British Machine Vision Conference (BMVC)* (2009).
- [10] S. Bak, E. Corvee, F. Bremond and M. Thonnat: “Boosted human re-identification using Riemannian manifolds”, *Image and Vision Computing* (2011).
- [11] D. Baltieri, R. Vezzani and R. Cucchiara: “3dpes: 3d people dataset for surveillance and forensics”, *Proc. of the 1st International ACM Workshop on Multimedia access to 3D Human Objects, Scottsdale, Arizona, USA*, pp. 59–64 (2011).
- [12] M. Hirzer, C. Beleznaï, P. M. Roth and H. Bischof: “Person Re-identification by Descriptive and Discriminative Classification”, *Proc. of Scandinavian Conference on Image Analysis*, pp. 91–102 (2011).
- [13] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani and V. Murino: “Custom Pictorial Structures for Re-identification”, *Proc. of the 22nd British Machine Vision Conference (BMVC)*, pp. 68.1—68.11 (2011).
- [14] A. Bialkowski, S. Denman, S. Sridharan, C. Fookes and P. Lucey: “A Database for Person Re-Identification in Multi-Camera Surveillance Networks”, *Proc. of International Conference on Digital Image Computing : Techniques and Applications (DICTA)* (2012).
- [15] J. Per, V. S. Kenk, R. Mandeljc, M. Kristan and S. Kovacic: “Dana36: A Multi-camera Image Dataset for Object Identification in Surveillance Scenarios”, *Proc. of 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance*, pp. 64–69 (2012).
- [16] W. Starzyk, A. Domurad and F. Z. Qureshi: “A virtual vision simulator for camera networks research”, *Proc. of the 2012 9th Conference on Computer and Robot Vision (CRV 2012)*, pp. 306–313 (2012).
- [17] W. Li, R. Zhao and X. Wang: “Human Reidentification with Transferred Metric Learning”, *Proc. of Asian Conference on Computer Vision* (2013).
- [18] W. Li and X. Wang: “Locally Aligned Feature Transforms across Views”, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3594–3601 (2013).
- [19] W. Li, R. Zhao, T. Xiao and X. Wang: “DeepReID : Deep Filter Pairing Neural Network for Person Re-Identification”, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (2014).
- [20] T. Xiao, H. Li, W. Ouyang and X. Wang: “Learning Deep Feature Representations with Domain Guided Dropout for Person Re-identification”, *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1249–1258 (2016).