

固定カメラ映像を対象としたHOG人物検出器のシーン適応手法

服部 博憲[†] 満上 育久^{††} 椋木 雅之^{††} 美濃 導彦^{††}

[†] 京都大学大学院情報学研究科 〒 606-8501 京都市左京区吉田本町

^{††} 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

E-mail: [†]{hattori,mitsugami,mukunoki,minoh}@mm.media.kyoto-u.ac.jp

あらまし 本論文では、固定カメラにより撮影された映像を対象として、Histograms of Oriented Gradients(HOG)特徴量ベースの人物検出手法で発生する誤検出や検出漏れを、時間の経過とともに減少させる手法を提案する。このHOG人物検出手法では、学習した人物パターンの局所的な形状特徴のみに基づいて人物検出を行うため、偶然人物と類似する特徴量を示す背景領域のパターンを誤検出したり、逆にカメラやその設置条件によって人物の特徴を十分取得できず検出漏れを起こすことがある。そこで提案手法では、これらの検出結果群から正検出と誤検出を判別し、これらを用いてカメラに特化した追加知識を獲得して、以後の検出処理の性能を向上させる。ここでの検出の成功と失敗の判別は、それらの時系列的な出現パターンに注目することによって行う。提案手法によって誤検出や検出漏れが減少することを、実環境中の固定カメラ映像を用いた実験によって示す。

キーワード 人物検出, Histograms of Oriented Gradients, シーン適応

Scene Adaption of HOG-Based Human Detection for Fixed Camera

Hironori HATTORI[†], Ikuhisa MITSUGAMI^{††}, Masayuki MUKUNOKI^{††}, and Michihiko MINOH^{††}

[†] Graduate School of Informatics, Kyoto University

Yoshidahonmachi, Sakyo-ku, Kyoto, 606-8501 Japan

^{††} Academic Center for Computing and Media Studies, Kyoto University

Yoshidahonmachi, Sakyo-ku, Kyoto, 606-8501 Japan

E-mail: [†]{hattori,mitsugami,mukunoki,minoh}@mm.media.kyoto-u.ac.jp

Abstract In this thesis, we propose a method that reduces false detection of HOG-based human detection. Since the HOG-based method uses only local features of human pattern, it often causes the following errors; (i) it occasionally detects a false area since it might mistakenly detect a background area whose pattern is incidentally similar to human pattern, and (ii) it sometimes spoils a human region when it fails to acquire enough features in that region. The proposed method improves the performance by using additional information specific to each camera, which can be obtained by judging true or false detections from these results. This judgment is achieved by analyzing time series of detected results. Experimental results show the effectiveness of the proposed method; false detection were reduced in several scenes.

Key words People Detection, Histograms of Oriented Gradients, Scene Adaptation

1. ま え が き

駅やショッピングセンターなどの施設において人物の移動軌跡を獲得することができれば、様々な用途に利用することが可能である。例えば、移動軌跡を分析し施設内での人物の移動傾向を獲得することにより、施設内のレイアウト改善や施設自体の設計改善といったマーケティング面で活用することが可能である。また、移動軌跡から不審な徘徊を発見することで、セキュリティ面でも活用することが可能である。

移動軌跡は、環境中にカメラを設置し、その映像から時々刻々の人物位置を把握する方法が一般的である。ただし、時々刻々の映像から人物位置を把握する作業を人為的な方法で行った場合、作業に時間がかかることや、人件費が高くつくといった問題がある。そこで、移動軌跡を自動的に獲得できるようなシステムを構築することが望まれるが、その為にはカメラの映像中から人物を検出する画像処理技術が必要になる。よって、本研究ではこのような固定設置されたカメラの映像を対象とし、頑健な人物検出を実現する手法を提案する。

カメラの映像中から人物を検出する手法としては、近年様々なものが提案されているが、中でも人物の形状特徴量を用いたものが注目されている [1] ~ [4]。この手法は事前に人物に関する多数の正事例・負事例のサンプル画像から人物の局所形状特徴を学習しておくことで、任意のシーンにおいて比較的良好に人物を検出することが出来るものである。しかしながら、既存の検出手法は、学習した人物パターンの局所的な形状特徴のみに基づいて人物かどうかを識別しているため、固定背景であっても偶然人物と類似する局所特徴をもつ領域がある場合には人物と区別できずに誤検出を起こす。また、逆にカメラやその設置条件によって人物の特徴を十分取得できず、検出漏れを起こすことがある。

そこで本研究では、人物検出器が適用シーンにおける人物の見えに関する知識を獲得することで、そのシーンに適応し、既存手法により起こる誤検出や検出漏れを時間の経過とともに減少させる手法を提案する。ここでの適用シーンにおける人物の見えに関する知識は、検出器が出力する検出結果の正誤から獲得する。各フレームにおける検出結果の正誤は、そのフレームだけでは判断する術がないが、正検出・誤検出の時系列的な出現パターンに差があることを利用して推定することは可能である。そこで、提案手法では検出結果の時系列方向の変化に注目し、検出結果を時系列方向に解析することで、フレーム毎の検出結果に対して正検出・誤検出の判定及び検出漏れの推定を行う。そして、この検出結果の正誤から適用シーンにおける人物の見えに関する知識を獲得する。具体的には、正検出・誤検出、検出漏れの事例からそれらの形状特徴を獲得し追加学習を行うことで、検出器を更新する。また、正検出と判定された事例からはカメラキャリブレーションを行い、画像中での人物の見えを推定することで、以後の不要な探索を削減し、誤検出を軽減させる。

2. HOG 特徴量を用いた人物検出の原理

2.1 HOG 特徴量を用いた検出の概要

HOG 特徴量を用いた人物検出手法 [1] は、特徴量の学習過程と検出過程に分けることができる。検出過程は、図 1 に示されるように、探索候補の設定と検出対象らしさの判定の 2 段階に分けられる。1 段目の探索候補の設定では、画像に対して様々な向きや大きさで注目領域 (検出ウィンドウと呼ぶ) を設定し、画像を複数回ラスタスキャンすることで探索候補を獲得する。そして獲得した候補に対して、HOG 特徴量による評価を行うことで様々な大きさや向きに映る対象の検出を行う。

2.2 HOG 特徴量の算出

HOG 特徴量は検出ウィンドウによって切り出された領域に対して算出される。その算出過程は、図 2 に示されるように 3 つのステップからなる。まず、最初にピクセル単位で輝度勾配の強度・方向が算出される。次に、ここで算出された輝度勾配を近傍領域 (セル) でヒストグラム化することで、照明や影の影響を受けにくい特徴とする。最後に、近傍のセル同士 (ブロック) で正規化することで、ブロック内での検出対象の局所的な回転や平行移動が生じた場合でも、ほぼ変化しない特徴量が得

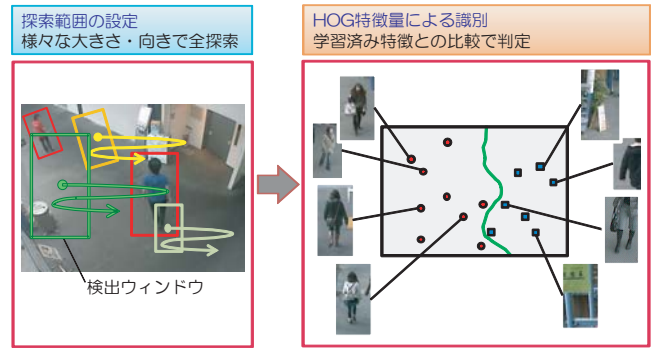


図 1 既存の HOG 特徴量ベースの検出手法の概要

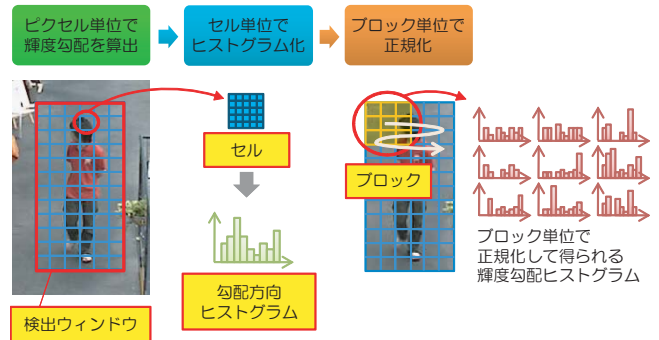


図 2 HOG 特徴量の算出

られる。以下では、これらの各ステップについて詳細を述べる。輝度勾配の強度・方向算出 画像中の位置 (u, v) におけるピクセルの輝度値を $L(u, v)$ とすると、 u, v 方向の輝度勾配 f_u, f_v は次式により算出できる。

$$\begin{cases} f_u(u, v) = L(u + 1, v) - L(u - 1, v) \\ f_v(u, v) = L(u, v + 1) - L(u, v - 1) \end{cases} \quad (1)$$

このとき、その勾配強度 m と勾配方向 θ は次式より算出できる。

$$m(u, v) = \sqrt{f_u(u, v)^2 + f_v(u, v)^2} \quad (2)$$

$$\theta(u, v) = \tan^{-1} \frac{f_v(u, v)}{f_u(u, v)} \quad (3)$$

ただし、ここで算出された勾配方向は $0^\circ \sim 360^\circ$ となるが、前景の明るさと隣接する背景領域の明るさの大小関係が逆転しても勾配方向が不変となるように、勾配方向を $0^\circ \sim 180^\circ$ に変換して用いる。

セルによるヒストグラム化 次に、上記で算出された勾配方向 θ を N_{bin} 方向に量子化し、それぞれに勾配強度 m の重みを与える。そして、これらを近傍領域 (セル, $W_{cell} \times H_{cell}$ ピクセル) において、ヒストグラム化することによって輝度の勾配方向ヒストグラム $h = (h_1, h_2, \dots, h_{N_{bin}})$ を作成する。ブロックによる正規化 最後に、各セルで作成した輝度の勾配方向ヒストグラム h を近傍のセル領域 (ブロック, $W_{block} \times H_{block}$ セル) で正規化する。ブロック内の i 行 j 列に位置するセルの勾配方向ヒストグラムを h_{ij} と表現すると、 h_{ij} はブロック内で次式により正規化される。

$$h'_{ij} = \frac{h_{ij}}{\sqrt{\sum_{i=1}^{H_{block}} \sum_{j=1}^{W_{block}} h_{ij}^2 + \epsilon}} \quad (4)$$

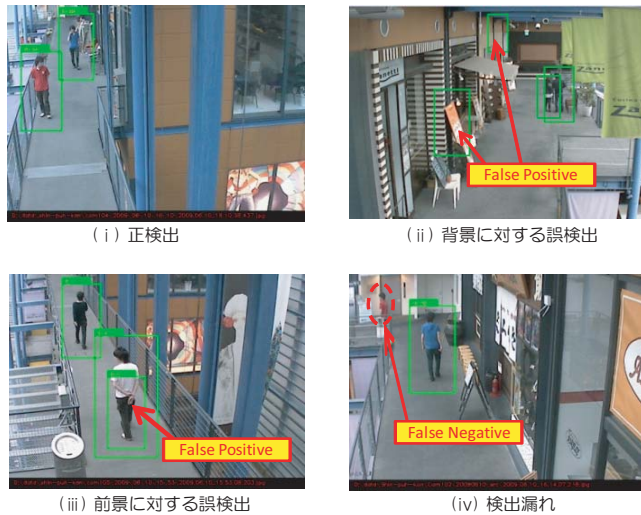


図 3 人物検出結果の例

ただし、 ε は分母が 0 の場合に計算不能になるのを防ぐ係数であり、ここでは $\varepsilon = 1$ とする。これにより、各ブロックからは $N_{bin} \times W_{block} \times H_{block}$ 次元の特徴量が得られることになる。そして、ブロックは 1 セルずつ移動しながら繰り返し正規化を行い、ブロックの移動回数を N_{block} とすると、最終的に抽出される特徴量は $N_{bin} \times W_{block} \times H_{block} \times N_{block}$ 次元となる。

2.3 既存の HOG 特徴量ベースの人物検出手法による検出結果

実環境で撮影された映像に対して、既存の HOG 特徴量ベースの人物検出手法 [1] を用いて人物検出を行った結果を図 3 に示す。このとき、様々な誤検出 (False Positive) や検出漏れ (False Negative) が生じる。これらの誤検出や検出漏れは、人物と類似する局所形状特徴をもつ固定背景に対すして発生する誤検出 (図 3(ii))、人物に対して発生する誤検出 (同図 (iii))、うまく特徴量が抽出できなかった瞬間に発生する検出漏れ (同図 (iv)) に分類できる。

図 3(ii) については既存の HOG 特徴量ベースの人物検出手法が有限個のサンプル画像から学習した局所形状特徴のみに基づいて人物かどうかを識別しているために、偶然人物と類似する局所特徴を持つ背景領域が存在したときに誤検出すると考えられる。また図 3(iii) や図 3(iv) については、カメラやその設置条件によって人物の特徴を十分取得できなかったために生じると考えられる。そして、図 3(ii) の誤検出は、常に画像中で決まった位置に出現するのに対して、図 3(iii) の誤検出や図 3(iv) の検出漏れは、ごく瞬間的にのみ発生する。このように、各誤検出や検出漏れは時系列方向にみるとそれぞれ正検出とは異なる特徴を持っていることが分かる。

そこで本研究では、これらの検出結果を時系列方向に解析することによって、それぞれの検出結果から正検出・誤検出の判定及び検出漏れの推定を行う事を考える。そして、その結果を検出器にフィードバックすることでシーンに適応し、誤検出や検出漏れを軽減させることのできる手法を提案する。

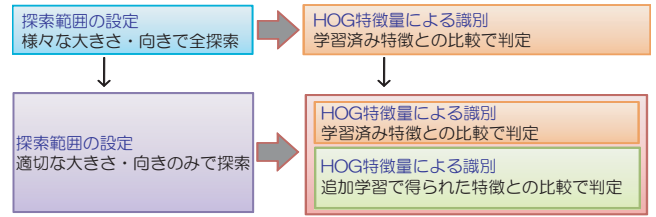


図 4 提案手法の概要

3. 検出結果の出現パターンに注目した検出器のシーン適応

2.3 節の考察をもとに、既存の HOG 特徴量ベースの検出器による検出結果を時系列方向に解析することで、その正誤を判定し、そのシーンにおける検出性能を向上させるための追加知識を獲得する手法を提案する。そして、このシーン依存の知識を利用することで、既存手法により起こる誤検出や検出漏れを時間の経過とともに減少させる。具体的には、正検出・誤検出と判定した事例からそのシーンに特化した人物形状特徴量を獲得し追加学習を行って生成した識別器と既存の検出器を併用して、識別性能を向上させる。

本研究で提案するシーン適応手法のアプローチは、図 1 に示される既存の HOG 特徴量ベースの人物検出器を、図 4 のように改善したことになる。すなわち、既存手法が探索範囲の設定を全探索としていたのに対して、提案手法では適切な探索のみを行うようにする。また、既存手法が事前に学習した局所特徴のみを用いて人物らしさの判定を行っていたのに対して、提案手法では追加学習によって生成された検出器による人物らしさの判定を併用する。

なお、この提案手法と同様に、時系列方向の情報を利用する手法としては、これまで時系列方向の動き特徴を用いるアプローチが提案されていた。例えば、Dalal ら [10] は HOG 特徴量にオプティカルフローを結合し、人物の見え特徴に加え、動き特徴を考慮することで人物検出の精度向上を行っている。また、山内ら [7] は HOG 特徴量に、ピクセルの輝度値に基づく状態判定 (背景・動状態・静状態) から得られる時空間特徴を結合することで、精度向上を行っている。しかしながら、両手法とも事前に学習した特徴量に基づいて検出を行っているという点は、既存の HOG 特徴量ベースの人物検出器と変わっていないため、学習シーンと異なるシーンに対して適用した場合、そのシーン毎に少なからず誤検出や検出漏れが発生すると考えられる。一方、本研究では、時系列方向の情報を利用してシーンに特化した形での学習を行うため、シーン毎の誤検出や検出漏れに対応することができると考えられる。

3.1 検出結果の位置変化に注目した解析に基づく HOG 特徴量の追加学習

本研究では、既存の HOG 特徴量ベースの人物検出結果に対して、時系列方向の位置変化に注目し解析することで、正検出・誤検出の判定及び検出漏れの推定を行う。そして、その判定結果に基づいて追加学習を行うことで、そのシーンにおける人物の見えを反映した検出器を獲得する。提案手法では、獲得され

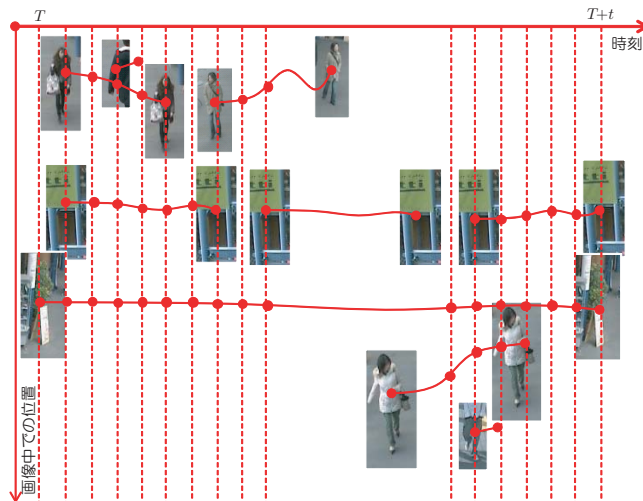


図 5 検出結果の時系列での解析

た検出器と既存の検出器を併用することで、誤った追加学習による精度の低下を招くことなく、誤検出・検出漏れを減少させるというアプローチをとる。

3.1.1 時系列方向の位置変化に注目した検出結果の解析

既存の HOG 特徴量ベースの人物検出器を用いた各フレームに対する検出結果を蓄積し、時系列方向に解析することで、正検出・誤検出の識別と検出漏れの推定を行う。

カメラの映像に対して、既存の HOG 特徴量ベースの人物検出器を適用し、時刻 T から $T+t$ の間に蓄積された検出結果をフレーム間で画像中の位置が近傍のもの同士対応付ける。ただし、映像のフレームレートは、フレーム毎の検出結果をフレーム間で対応付けることが可能なくらい十分小さいとする。このとき、図 5 の様な結果が得られ、これより、(i) 正検出は、画像中で滑らかに位置変化し、(ii) 固定背景に対する誤検出は画像中で常に同じ位置に生じ、(iii) 人物の一部だけに対する誤検出は正検出よりも短い時間で生じると言える。

よって、上記の特性を踏まえて本研究では以下の基準により検出結果の正誤判定を行う。

- 検出結果を対応付けたとき、それが k_1 フレーム以上に渡って対応付けられているならば、その検出結果は固定背景に対する誤検出であると判定する。ただし、 k_1 は人物が同じ場所に滞留している可能性のある時間よりも十分長い観測時間で得られるフレーム数に設定する。
- 検出結果を対応付けたとき、それが k_2 フレーム以上 k_1 フレーム未満に渡って対応付けられているならば、その検出結果は人物に対する正検出であると判断する。
- 検出結果を対応付けたとき、それが k_2 フレーム未満で対応付けられているならば、その検出結果は人物の一部だけに対する誤検出と判断する。

画像中で人物の位置変化は滑らかに起こるので、前後のフレームの検出結果から注目しているフレームの検出結果を推定する。ここで推定された検出結果と実際の検出結果が異なる場合に、検出漏れが発生していると推定する。

3.1.2 検出結果の解析による追加学習

正検出・誤検出の判定及び検出漏れの推定結果をもとに、正事例・負事例となる HOG 特徴量の獲得とその学習を行い、新たな検出器の生成を行う。ただし、ここでは上記で判定された結果を元に、2 つの識別器を構築する。

ここで獲得する識別器の片方は、固定背景に対する誤検出を学習させるものであり、この学習により生成される識別器の情報は常に保存されるとする。一方でもう片方は人物の一部だけに対する誤検出と検出漏れを学習させるものであり、この学習により生成される識別器の情報は古いものから削除されていくとする。この様に識別器を分ける理由は、固定背景に対する誤検出情報は時間が経っても変化しない情報であるが、人物の一部だけに対する誤検出や検出漏れは人物の服装の傾向などによって変わりうる情報であるからである。

3.2 検出結果の出現パターンに注目したカメラ位置推定

HOG 特徴量は注目領域に対して算出される特徴量であるため、回転やスケールに対して不変ではない。よって、既存の HOG 特徴量ベースの人物検出では任意の大きさや向きに映る人物を検出するために、検出ウィンドウの大きさや向きを変化させながら画像を複数回ラスタスキャンする必要があった。しかしながら検出ウィンドウに多大な自由度を許すと、偶然人物と同じような特徴量を示す領域が存在すると、その領域に対して誤検出を起こす。

従って、人物検出の精度を向上させるためには、検出ウィンドウの向きと大きさを適切に設定する必要があると考えられる。ただし、システムが事前に検出ウィンドウの適切な向きと大きさを知るためには、カメラと人物の位置関係を把握しておく必要があり、これにはカメラキャリブレーションを行う必要があるが、設置されている全てのカメラでこれを行うためには大変な労力が必要になる。そこで本研究では、画像中で検出された人物の位置と大きさ・向きを使うことで、人手でキャリブレーションを行わずに、システムが自動でキャリブレーションを行い、任意の位置での人物の向きや大きさを予測することで、検出ウィンドウの向きと大きさを適切に設定する手法を提案する。

3.2.1 人物サンプルの獲得による消失点推定

画像中で検出された人物の位置と大きさ・向きをもとに消失点・消失線推定を行う手法について述べる。ただし、本研究では単一の平面上に垂直に立った人物が往来している様子を上方から固定カメラで撮影しているとする。また、撮影された映像中には、人物の全身が映っていることを仮定する。

既存の HOG 特徴量ベースの人物検出器により検出された検出結果のうち正事例と判定されたものを、時系列方向に位置が近傍のもの同士対応付ける。そして、対応付けられた人物に対して、画像上の 2 地点で頭の位置 P_{head} 、足元の位置 P_{foot} を推定する。この推定は、図 6 のように検出された人物に対して、検出ウィンドウの上辺の中点に P_{head} 、下辺の中点に P_{foot} をとる。このとき、人物は床に対して垂直に立っており、かつ同一人物であるためその身長が同じであることを考慮すると、図 6 中に示された直線 $P_{head}P_{foot}$ と $P'_{head}P'_{foot}$ は、実世界における平行線である。同様に直線 $P_{head}P'_{head}$ と $P'_{foot}P'_{foot}$ も実

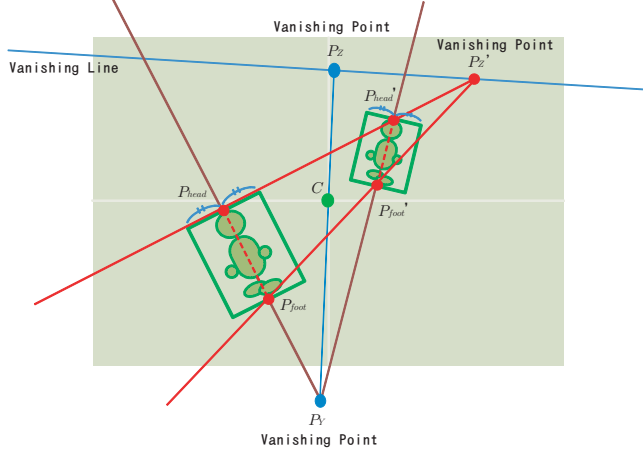


図 6 消失点と消失線

世界における平行線である．よって，これらの直線の交点から y 軸方向， z 軸方向の消失点 $P_Y(u_Y, v_Y)$ と $P_Z(u_Z, v_Z)$ が得られる．消失点と消失線の推定を安定に行うために，2 地点以上の多地点での頭の位置 P_{head}^i ，足元の位置 $P_{foot}^i (i \in U = \Omega + \bar{\Omega})$ から RANdom SAMple Consensus(RANSAC) によるロバスト推定を行う．

3.2.2 消失点を用いたカメラパラメータの推定

消失点の座標 $P_Y(u_Y, v_Y)$ ， $P_Z(u_Z, v_Z)$ から，カメラの焦点距離 f ，チルト角 α ，ロール角 γ を求める方法について述べる．図 7 のように座標系を設定しカメラの透視投影を考え，3 次元点 $M = (X, Y, Z)^T$ を射影して得られる画像点を $m = (u, v)^T$ で表現し，それぞれの同次座標を $\hat{M} = (X, Y, Z, 1)^T$ 及び $\hat{m} = (u, v, 1)^T$ とする．この時， \hat{m} と \hat{M} は次の透視投影の関係式を満たす．

$$s \hat{m} = K A \hat{M} \quad (5)$$

ここで， K, A はそれぞれカメラの内部パラメータ行列，外部パラメータ行列と呼ばれるカメラ固有の行列である．また， s は任意のスケールを表す．カメラの内部パラメータ行列 K は次式で定義される．

$$K = \begin{pmatrix} f_u & \nu & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (6)$$

ここで， f_u, f_v は焦点距離， ν は歪み， (u_0, v_0) は画像中心である．一方で，外部パラメータ行列 A は，次式のように 3×3 回転行列 R と 3×1 並進行列 T で表現できる．

$$A = (R|T) \quad (7)$$

ここで， X, Y, Z 軸周りの回転角を α, β, γ とすると R は次式で表現される．

$$R = \text{Rot}(z, \gamma) \text{Rot}(y, \beta) \text{Rot}(x, \alpha) \quad (8)$$

また，各軸方向の移動成分を T_X, T_Y, T_Z とすると T は次式で表現される．

$$T = (T_X \quad T_Y \quad T_Z)^T \quad (9)$$

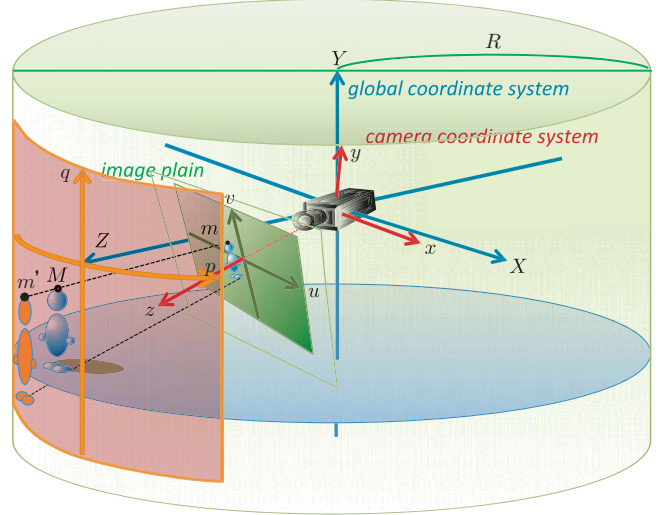


図 7 座標形

ここで，カメラのアスペクト比は $1 (f_u = f_v)$ ，以後 f と表記する)，画像に歪みはなく ($\nu = 0$)，画像中心 (u_0, v_0) は画像の中央の点 (画像が 640×480 pixel であるとする) $u_0 = 320, v_0 = 240$) として扱う．厳密には，これらのパラメータも推定する必要があるが，ここでは画像中のおおよその人物の見えを推定できれば十分である．また，世界座標の原点をカメラ位置を基準にとるとすると， $T_X = 0, T_Y = 0, T_Z = 0, \beta = 0$ となる．

ここで， $P_Y(u_Y, v_Y)$ と $P_Z(u_Z, v_Z)$ はそれぞれ Y 軸方向， Z 軸方向の無限遠点であるから， $Y \rightarrow \infty$ は (u_Y, v_Y) に対応し， $Z \rightarrow \infty$ は (u_Z, v_Z) に対応する．すなわち u_Y, v_Y と u_Z, v_Z は次式で表現される．

$$\begin{cases} u_Y = -f \frac{\sin \gamma \cos \alpha}{\sin \alpha} \\ v_Y = f \frac{\cos \gamma \cos \alpha}{\sin \alpha} \end{cases} \quad \begin{cases} u_Z = f \frac{\sin \gamma \sin \alpha}{\cos \alpha} \\ v_Z = -f \frac{\cos \gamma \sin \alpha}{\cos \alpha} \end{cases} \quad (10)$$

これらの式から f, α, γ が次式により推定される．

$$f = \sqrt{-u_Y u_Z - v_Y v_Z} \quad (11)$$

$$\alpha = \tan^{-1} \sqrt{-\frac{u_Z}{u_Y}}, \quad \gamma = \tan^{-1} \left(-\frac{u_Y}{v_Y} \right) \quad (12)$$

カメラキャリブレーションによって推定されたパラメータを用いることで，画像中に映る人物の向きと大きさを予測し，それに応じて適切な向きと大きさにのみ検出ウィンドウを配置する．そのために，本研究では図 7 のように画像を半径 R の円筒面に投影変換して扱う．これにより，検出ウィンドウを回転させることなく一定方向の検出ウィンドウのみで検出が可能になり，また人物の見えの縦横比が一定になるため検出に適していると考えられる．円筒面上に投影された画像上では，人物の大きさは画像の q 軸方向の位置に比例して変化するため，複数のサンプルを獲得しておくことで，画像上の任意の地点での人物の大きさを推定することが可能となる．

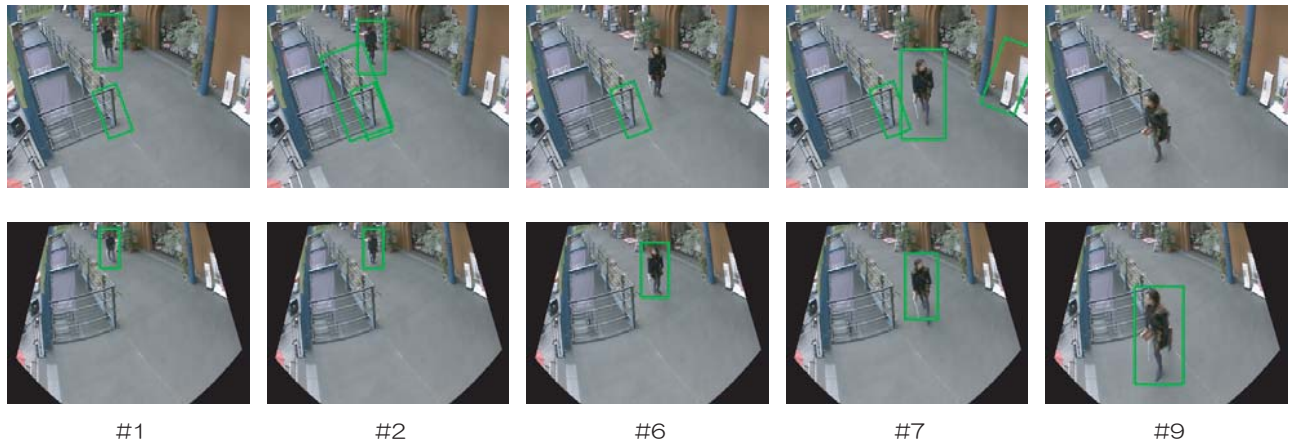


図 8 既存手法と提案手法の比較

4. 実験と評価

4.1 実験環境

日照の変動による照明変化が発生する屋外通路を固定カメラで撮影し、得られた映像に対して提案手法を適用した。映像はフレームレートが約 2.5fps、フレームサイズが VGA(320×240pixel) で撮影されたものである。HOG 特徴量については、 $N_{bin} = 9$, $W_{cell} = H_{cell} = 5$, $W_{block} = H_{block} = 2$ として扱った。また、検出ウィンドウのサイズを 30×60 ピクセルに正規化した場合、 $N_{block} = 40$ となり、最終的に抽出される HOG 特徴量は $9 \times 3 \times 3 \times 40 = 3,240$ 次元である。なお、HOG 特徴量の学習には AdaBoost を用いた。

4.2 既存手法と提案手法の比較

既存手法と提案手法による人物検出を行い、精度の比較を行う。ここでは、既存の HOG 特徴量ベースの人物検出器として Dalal ら [1] の手法を用いた。なお、探索候補は検出ウィンドウのサイズを 32×64 pixel ~ 96×192 pixel までの範囲で 1.05 倍ずつ大きくしながら、また同時に向きを -40 度 ~ $+40$ 度まで 4 度ずつ回転させながら全探索するように設定した。

30 分間で撮影された画像約 5,200 枚に対して、既存手法と提案手法を適用した結果を図 8 に示す。ただし、提案手法は別の 30 分間で撮影された画像を用いて人物の見えに関する知識を追加獲得している。このとき、提案手法を用いることで既存手法で生じる誤検出や検出漏れを軽減することが出来ていることが確認された。

5. 結 び

本研究では、固定設置されたカメラ映像を対象に人物検出を行う際に発生する誤検出や検出漏れを、人物検出器が追加学習することでシーンに適応し、誤検出や検出漏れを軽減させる手法を提案した。

今後の課題としては、提案手法を長時間稼働させた際の精度の安定性を調査する必要がある。また、提案手法は適用開始時の検出器の精度が適応性能に強く依存すると考えられるので、これらの関係についても考察することが求められる。

謝辞 本研究は、文部科学省の科学技術振興調整費(科学技術連携施策群の効果的・効率的な推進)による「センサ情報の社会利用のためのコンテンツ化」の一環として実施したものである。

文 献

- [1] N.Dalal and B.Triggs, "Histograms of Oriented Gradients for Human Detection", IEEE Computer Vision and Pattern Recognition, pp. 886-893, 2005.
- [2] F.Suard and A.Broggi, "Pedestrian Detection using Infrared images and Histograms of Oriented Gradients", IEEE Symposium on Intelligent Vehicule, pp. 206-212, Jun, 2006.
- [3] Q.ZhuS, Avidan, M.Yeh and K.Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients", IEEE Computer Vision and Pattern Recognition, Vol. 2, pp.1491-1498, Jun, 2006.
- [4] 小林 拓也, 日高 章理, 栗田 多喜夫, "Histograms of Oriented Gradients を用いた対象識別での特徴選択", 電子情報通信学会技術研究報告, PRMU[パターン認識・メディア理解], 106(606) pp.119-124 20070309.
- [5] 服部 寛, 関 晃仁, 西山 学, 渡辺 友樹, "車載ステレオカメラによる歩行者検出" 画像の認識・理解シンポジウム (MIRU2009), 2009.07.
- [6] 藤吉 弘亘, "Gradient ベースの特徴抽出 -SIFT と HOG-", 情報処理学会研究報告, CVIM[コンピュータビジョンとイメージメディア], 160, pp.211-224, 2007.
- [7] 山内 悠嗣, 藤吉 弘亘, Hwang Bon-Woo, 金出 武雄, "アピアランスと時空間特徴の共起に基づく人検出", 画像の認識・理解シンポジウム (MIRU2007), 2007.07.
- [8] 井手上 健一, 青木 工太, 長橋 宏, "夜間映像中の人物の検出と追跡", 画像の認識・理解シンポジウム (MIRU2009), 2009.07.
- [9] 池村 翔, 藤吉 弘亘, "距離情報に基づく局所特徴量によるリアルタイム人検出", 第 15 回 画像センシングシンポジウム, SSI109, IS4-05, 2009.
- [10] N.Dalal, B.Triggs and C.Schmid, "Human Detection Using Oriented Histograms of Flow and Appearance", IEEE European Conference on Computer Vision, vol.2, pp.428-441, May, 2006.
- [11] Imran Junejo, Hassan Foroosh, "Robust Auto-Calibration from Pedestrians", 2006 IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS'06)
- [12] F.Lv, T.Zhao, R.Nevatia, "Camera calibration from video of a walking human", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.28, no.9, pp.1513-1518, September 2006.
- [13] 安藤 寛哲, 藤吉 弘亘, "人検出結果に基づくカメラ自己キャリブレーションと 3 次元位置推定", 電気学会一般産業研究会, 2009.