

# Joint HOG 特徴を用いた 2 段階 AdaBoost による車両検出

尾崎貴洋†、山内悠嗣†、藤吉弘亘†  
中部大学工学部情報工学科†

車載カメラの映像から後方車両を検出するための新しい手法として、複数の Histograms of Oriented Gradients(HOG)[6]特徴を組み合わせた Joint HOG 特徴と、2 段階に構築した AdaBoost[9]による学習法を提案する。Joint HOG 特徴は、複数の HOG 特徴を同時に観測するため、従来の単一の HOG 特徴のみでは捉えることのできない、車両特有の左右や上下の対称的な形状特徴や連続的なエッジを自動的に捉えることが可能となる。提案手法の有効性を確認するため、車載カメラの映像から構築したデータベースを用いて、従来法である単一の HOG 特徴のみを用いた手法と比較し、DET カーブにより評価した。評価実験の結果、誤識別率 2.0%において、提案手法は従来法より約 7.0%の識別率向上を実現した。

## 1. はじめに

現在、交通量の増加による交通事故の増加が深刻な問題となっている。この問題を対処するため、ITS(Intelligent Transport System : 高度道路交通システム)に関する研究が行われており、近年では、車載カメラの映像中から画像処理技術を用いて車両検出を行い、運転者の視覚支援に役立てる研究が多く取り組まれている。本研究では、車載カメラから後方車両を検出することを目的とする。車両は人工物特有の直線的なエッジを多く含み、その形状には左右や上下の対称性や連続性といった特徴がある。この特徴を捉えるための手法として、Sabzmeydani らにより Shapelet 特徴を自動生成する手法が提案されている[12]。この手法は、low-level な各ピクセルの勾配情報を組み合わせて新たな特徴量を選択することにより、識別に有効な特徴量を自動生成するものである。しかしこの手法は、検出ウィンドウをいくつかの局所領域に分割し、局所領域内の low-level 特徴を組み合わせて特徴量を生成するため、局所領域を越えた組み合わせの特徴量を生成することはできない。そこで、本手法では、位置の異なる領域における、車両の対称性や連続性を自動的に選択する Joint HOG 特徴と 2 段階 AdaBoost による後方車両検出法を提案し、その有効性を示す。

## 2. 従来の車両検出法

車両は、同一車両における形状の変化は少ないが、車種、車両の向き、積載物等により、エッジの方向や強度が大きく変化する。また、屋外を走行するた

め、天候や時間、季節の変化による影響を受けやすいという問題がある。このような問題に対し、形状モデルを用いた手法、特徴抽出による手法等の車両検出法が提案されている。

### 2.1 形状モデルによる車両検出

形状モデルを用いるアプローチとして、Sullivan らはワイヤーフレームモデルで車両の形状を記述して、マッチングを行う手法[1]を提案している。しかし、形状モデルを用いる手法は事例の収集にコストがかかるため、実用的ではない。

### 2.2 特徴抽出による車両検出

特徴抽出による車両検出には、時間的特徴と、空間的特徴を用いた手法が提案されている。時間的特徴を用いたアプローチとして、Iketani らはオプティカルフローを用いて複雑に運動する背景の中から移動対象を検出する手法[4]を提案している。しかし、オプティカルフローは移動中の車両の検出は可能であるが、静止状態の車両においては得られる特徴が減少し、検出精度が低下するという問題がある。

空間的特徴を用いるアプローチとして、三谷らはガボール変換を用いて車両の画像パターンの方向性を反映したモデルを構築する手法[3]を提案している。これは照明の変化を受けにくく優れた特徴量であるが、計算コストが非常に高いという問題がある。Hasegawa らは、背景差分により得られたシルエットのモーメントや、面積等を特徴量とすることで車両検出を行い、その車種も判別可能とする手法[2]を提案している。しかし、背景が常に変動する車載カメラの映像等には、背景差分を用いることはできない。

### 2.3 特徴抽出と識別器の選択による車両検出

複数の low-level な特徴と識別器の組み合わせにより車両検出を行うアプローチとして、Haselhoff らは、白領域の輝度値の和と黒領域の輝度値の和の差を特徴とした Haar-like 特徴を用い、AdaBoost により学習した車両の検出法[5]を提案している。また Han らは Histograms of Oriented Gradients (HOG)を用いて、SVM により学習した車両の検出法[8]を提案している。Haar-like 特徴や HOG 特徴は、車両の勾配を用いており、ノイズや照明変動に頑健である。特に HOG 特徴は局所領域における輝度の勾配をヒストグラム化するため、照明変動や影の影響を受けにくく局所的な幾何学変化に頑健である。HOG 特徴や Haar-like 特徴等の low-level な特徴は、単一の特徴のみでは識別に困難なパターンが存在する可能性がある。そのため、大量な low-level 特徴から有効な特徴をどのように選択するかは重要な課題である。

そこで本手法では、位置の異なる領域における、車両の対称性や連続性を自動的に選択する Joint HOG 特徴と 2 段階 AdaBoost による後方車両検出を行う。

### 3. Joint HOG 特徴を用いた 2 段階 AdaBoost

本手法は、2 段階に AdaBoost を構築することにより、位置の異なる 2 つの Low-level な特徴を組み合わせた Joint HOG 特徴の POOL を作成し、最終識別器により、POOL から自動的に車両検出に最適な Joint HOG 特徴を選択する。提案手法の概要を図 1 に示す。

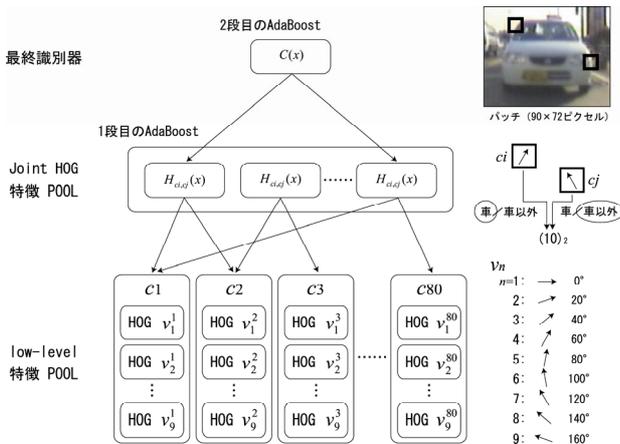


図 1 : 提案手法の概要

#### 3.1 low-level 特徴

本手法では、low-level 特徴として Dalal らが提案

した HOG 特徴を用いる。HOG は図 2 に示すように、セル (9×9 ピクセル) における勾配方向をヒストグラム化した特徴 ( $\mathbf{V} = [v_1, v_2, v_3, \dots, v_9]$ ) であり、物体の形状を表すことが可能である。隣接画素の勾配を局所領域毎にヒストグラム化して正規化するため、照明の変化を受けにくく、局所的な幾何学変化に頑健な特徴となる。以下に HOG 特徴の算出法を示す。

#### 3.1.1 輝度勾配の算出

画像  $L$  の輝度の勾配強度を  $m$ 、勾配方向を  $\theta$  とし、次式より算出する。

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad (1)$$

$$\theta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)} \quad (2)$$

$$\begin{cases} f_x(x, y) = L(x+1, y) - L(x-1, y) \\ f_y(x, y) = L(x, y+1) - L(x, y-1) \end{cases} \quad (3)$$

#### 3.1.2 セル領域におけるヒストグラムの作成

算出した輝度の勾配強度  $m$  と、勾配方向  $\theta$  よりセル (9×9 ピクセル) ごとに輝度の勾配方向ヒストグラムを作成する。勾配方向  $\theta$  は  $0^\circ \sim 360^\circ$  の値となるが、勾配の向きを考慮する必要がないため  $0^\circ \sim 180^\circ$  となる。勾配方向を  $20^\circ$  ずつに分割するため、9 方向の勾配方向ヒストグラムとなる。

#### 3.1.3 正規化

各セルより算出した勾配方向ヒストグラムを正規化する。本手法では検出ウィンドウ全体で正規化を行う。

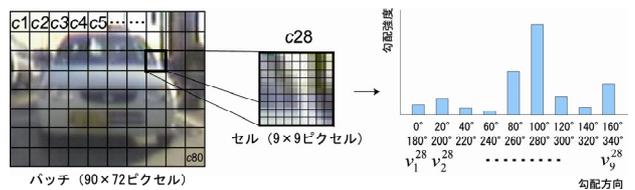


図 2 : HOG 特徴

### 3.2 Joint HOG 特徴

車両の対称的な形状特徴や連続的なエッジを捉えるため、位置の異なる 2 つのセル  $c_i$  と  $c_j$  内の HOG 特徴を組み合わせた Joint HOG 特徴を算出する。同時に 2 つの特徴量を観測することにより、単一の特徴量のみでは識別困難なパターンの識別精度を向上させる。

#### 3.2.1 Joint HOG 特徴の算出

Joint HOG 特徴を算出するために、各セル  $c_i$  と  $c_j$  から 1 つずつ HOG 特徴を選択し、次式から車か車以外を表す 2 値化符号  $s$  を算出する。

$$s(\mathbf{V}) = \begin{cases} 1 & p \cdot v_o > p \cdot \theta \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

ここで、 $\theta$  は閾値である。 $p$  は不等号の向きを決定する符号であり、+1 もしくは-1 をとる。 $o$  は勾配方向を表す。これにより、得られた2値化符号を2つ組み合わせることで共起[7]を表現した Joint HOG 特徴  $k$  を得る。例えばある学習サンプルにおいて、位置の異なる2つのセル  $c_i$  と  $c_j$  から HOG 特徴を2値化した符号  $v^{c_i} = 1, v^{c_j} = 0$  を観測したとき、Joint HOG 特徴  $k$  は  $k = (10)_2 = 2$  となる(図3)。Joint HOG 特徴  $k$  は2進表現された特徴の組み合わせのインデックス番号であり、本手法では2つの HOG 特徴の組み合わせとなるため、4通りの値をとる。

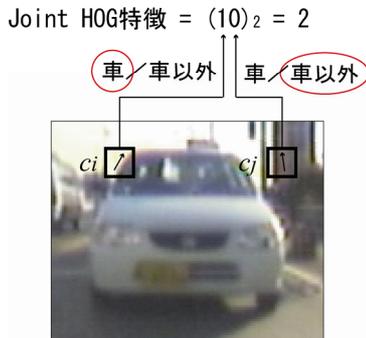


図3：Joint HOG 特徴

### 3.2.2 1段目の AdaBoost による特徴選択

位置の異なる2つのセル  $c_i$  と  $c_j$  において、算出した Joint HOG 特徴の中から、AdaBoost により識別に有効な特徴を選択する。入力画像  $x$  から Joint HOG を観測する関数を  $J_i(x)$  で表す。 $x$  からの特徴量  $J_i(x)=k$  を観測したとき、1段目の AdaBoost の弱識別器  $h_i(x)$  を次式により表す。

$$h_i(x) = \begin{cases} +1 & P_i(y=+1|k) > P_i(y=-1|k) \\ -1 & \text{otherwise} \end{cases} \quad (5)$$

ここで、 $y$  は正解ラベル  $y \in \{+1, -1\}$  を表し、 $P_i(y=+1|k)$  及び  $P_i(y=-1|k)$  は、それぞれ Joint HOG 特徴  $k$  を観測した際の車または車以外である条件付き確率であり、学習サンプル  $f$  の重み  $D_i(f)$  に基づいて次式により算出される。

$$P_i(y=+1|k) = \sum_{f: J_i(x_f)=k \wedge y_f=+1} D_i(f) \quad (6)$$

$$P_i(y=-1|k) = \sum_{f: J_i(x_f)=k \wedge y_f=-1} D_i(f) \quad (7)$$

図4に2つの HOG 特徴を選択した際の  $P_i(y=+1|k)$  及び  $P_i(y=-1|k)$  の分布を示す。

上記の過程をすべてのセルの組み合わせにおいて行う。本手法では、検出ウィンドウ(90×72 ピクセル)に対して 80 個のセルが存在するため、2つのセ

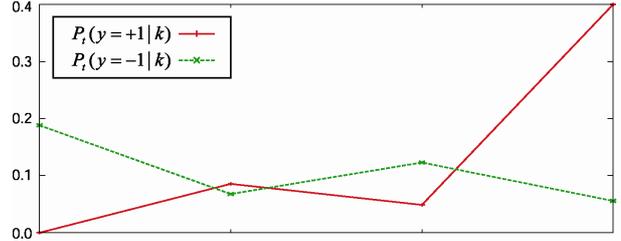


図4：各クラスの同時確率

ルの組み合わせは  ${}_{80}C_2=3160$  となり、次式に示すような 3160 個の強識別器  $H_{c_i, c_j}(x)$  を構築する。

$$H_{c_i, c_j}(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t^{c_i, c_j} \cdot h_t^{c_i, c_j}(x) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

ここで、 $T$  は学習回数、 $\alpha_t^{c_i, c_j}$  は1段目の弱識別器  $h_t^{c_i, c_j}(x)$  の重み、 $c_i, c_j$  はセルの組み合わせである。1段目の AdaBoost により、2つのセル間における Joint HOG 特徴が選択される。その組み合わせはセルの位置関係からエッジの対称性や連続性を自動的に捉えたものとなる。

### 3.3 最終識別器

2段目の AdaBoost では、Joint HOG 特徴で構築した強識別器  $H_{c_i, c_j}(x)$  の POOL の中から、最も識別に有効なものを選択して学習を行う。これは、車両の識別に有効なセルの位置関係を自動的に選択することになる。2段目の AdaBoost により得られる最終的な強識別器  $G(c)$  を次式に示す。

$$G(c) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t \cdot g_t(c) > \lambda \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

ここで  $g_t(c)$  は1段目の AdaBoost より得られる強識別器、 $c = \{c_i, c_j; 1 < i < 80, 1 < j < 80\}$ 、 $\lambda$  は検出器の閾値である。

## 4. 評価実験

本手法の有効性を示すため評価実験を行う。

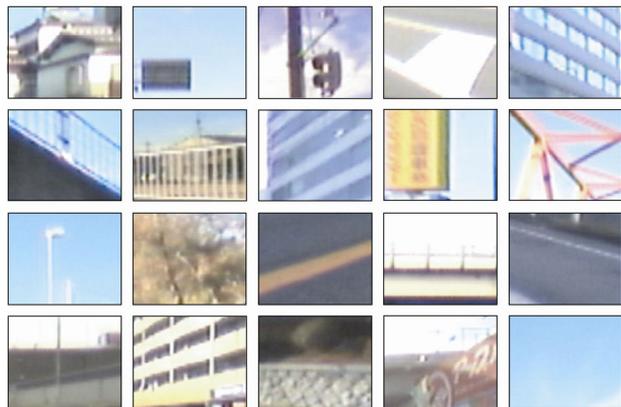
### 4.1 データベース

2段階 AdaBoost により学習するための学習用データベースと、評価実験を行うための評価用データベースを構築する。車載カメラの映像から後方車両領域を切り出したものを用いる。データベースは車両クラス(Positive クラス)と車両以外クラス(Negative クラス)をもつ。車両は捉える位置や角度により、形状が大きく変化する。そこで車両の位置や角度が変化しても、形状に変化の少ない車両の正面部分のみを用いて、車両クラスのデータベースを構築する。切り出しを行う基準として、車両のライトを中心にとり、上下左右に切り出す領域の15%をマージンと

した。また両隣1車線を走行中の車両のみを対象とする。一方、車両以外のクラスは、背景領域をランダムに切り出すことにより構築した。学習用データベースとして、車両クラス2464枚、車両以外クラス2415枚、評価用データベースとして、車両クラス1899枚、車両以外クラス2413枚を構築した。図5にデータベースの一部を示す。



Positiveクラス



Negativeクラス

図5：構築したデータベースの一部

#### 4.2 実験概要

評価用データベースを用いて、車両の識別実験を行う。提案手法 (Joint HOG + Two Stage AdaBoost) と HOG 特徴のみを用いた検出法 (HOG) [6]、Joint HOG 特徴のみを用いた検出法 (Joint HOG) との比較を行う。評価には Detection Error Tradeoff (DET) [10] を用いる。DET とは、横軸に誤検出率、縦軸に未検出率をとり、両対数グラフにより表したものである。原点に近いほど性能が良い。

#### 4.3 車両識別実験結果

図6に各手法のDETを示す。提案手法 (Joint HOG + Two Stage AdaBoost) は、従来法である HOG 特徴のみを用いた手法と比較して、識別性能が大幅に向上していることがわかる。誤検出率2.0%の場合、未

検出率が1.0%となり、約7.0%識別性能を向上させることができた。これは、HOG 特徴のみでは識別困難なパターンにおいても、位置の異なる2つのセル内の HOG 特徴を組み合わせることにより、新たに識別に有効な特徴量を自動的に捉えることができたためである。

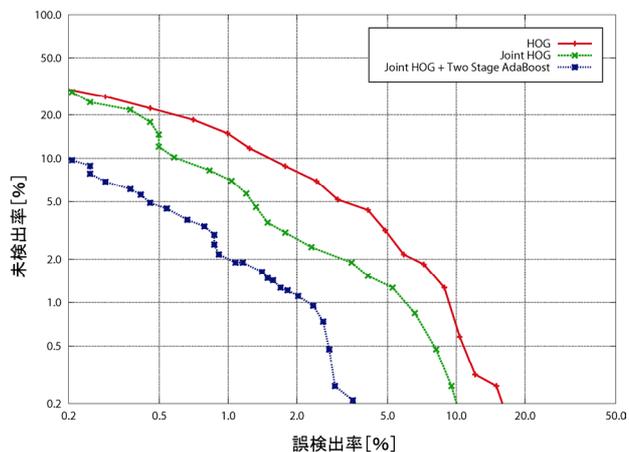


図6：DETによる比較結果

提案手法により、未検出したパターン (図7(a))は、高速に走行していることにより、カメラのピントを合わせることができずにぼけた画像や、太陽光の反射等により、テクスチャ情報(勾配)を得ることができない部分である。一方、誤検出をしたパターン(図7(b))は背景領域であるにもかかわらず、車両のエッジと似た直線的なエッジをもつものであり、主に建物や標識等の人工物である。未検出は、現状の車載カメラの性能では限界があるため、高性能なカメラで撮影する必要がある。誤検出は、提案手法に時系列データを組み合わせることにより、背景をリジェクトし、誤検出を減少させることが期待できる。



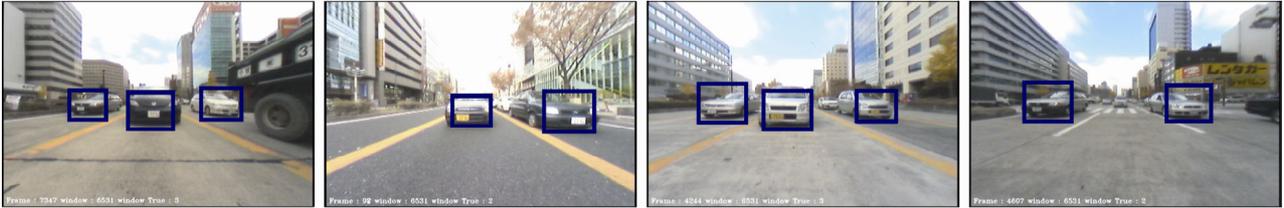
図7：未検出と誤検出例

#### 4.4 車両検出実験

構築した識別器を用いて、車載カメラの映像から車両の検出実験を行う。車両検出を行うために、検出ウィンドウのサイズを固定して、画像サイズを少しずつ拡大縮小しながら車両/車両以外の判別を行い、画像全体を複数回ラスタスキャンする。これにより車両のスケールが異なる場合でも検出可能となる。同一車両に対して、複数のウィンドウが車両として識別された場合、MeanShift クラスタリング[11]



従来法 (HOGのみ)



提案手法 (Joint HOG + Two Stage AdaBoost)

図 8 : 後方車両検出結果

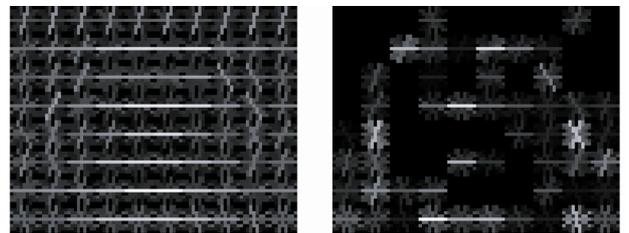
によるウィンドウの統合処理を行う。図 8 に車載カメラの映像から車両を検出した例を示す。提案手法は従来法と比較して、誤検出と未検出が少ないことがわかる。

## 5. 考察

車両の特徴である、形状の対称性や連続性が提案した Joint HOG 特徴に表現されているか考察する。

### 5.1 2 段階による特徴選択

図 9(a)に 1 段階目の AdaBoost において、Joint HOG 特徴により選択された HOG 特徴を可視化した結果を示す。また、図 9(b)に 2 段階目の AdaBoost において、最終識別器により選択された HOG 特徴を可視化した結果を示す。HOG 特徴の勾配方向を 9 方向で表現しており、白いほど各 AdaBoost により選択された頻度が高いことを示す。図 9(a)において、1 段階目の AdaBoost の Joint HOG 特徴により、車両内部の水平方向のエッジ、車両の輪郭部分に沿ったエッジが多く選択されている。図 9(b)では、図 9(a)で選択された HOG 特徴の中から、2 段階目の AdaBoost により得られる最終識別器により、車両の輪郭部分に沿った Joint HOG 特徴が選択されていることがわかる。これにより、車両の輪郭部分に沿った HOG 特徴が車両と背景を識別するのに有効な特徴であることがわかる。図 9(a)で多く選択されていた水平方向のエッジは一部しか選択されていない。これは、水平方向のエッジは車両に多く存在するが、背景領域の建物や看板、高架橋等にも多く存在するため、車両と背景を識別するのに有効ではないと判断され、選択されていないためである。以上から、2 段階に AdaBoost を構築することにより、車両識別に有効な特徴が自動的に選択されていることがわかる。



(a) 1 段階目の AdaBoost により 選択された HOG 特徴 (b) 2 段階目の AdaBoost により 選択された HOG 特徴

図 9 : 選択された HOG 特徴の可視化結果

### 5.2 車両検出に有効な特徴

次に、最終識別器の各弱識別器において、どのような位置関係のセルや Joint HOG 特徴が選択されているか考察する。図 10(a),(b),(c),(d)に最終識別器  $G(c)$  の第 1,第 2,第 3,第 15 ラウンドの弱識別器において選択された 2 つのセルの組み合わせを示す。また、図 10(e),(f),(g),(h)に最終識別器の各ラウンドにおけるそれぞれのセルの組み合わせにおいて、Joint HOG により選択された勾配方向と勾配強度を示す。第 1 ラウンドと第 2 ラウンドには図 10(e),(f)に示すように、垂直エッジと水平エッジの位置関係が選択されている。また第 3 ラウンドでは図 10(c)のように左右対称となる位置関係のセルが選択されている。さらに図 10(g)に注目すると、Joint HOG においても、それぞれのセルから左右対称となる勾配方向が選択されていることがわかる。第 15 ラウンドで選択された図 10(d)は連続性を捉えた位置関係のセルが選択され、図 10(h)より水平方向の特徴が選択されていることがわかる。提案した Joint HOG 特徴は、車両形状の対称性や連続性を捉える特徴量を予め用意しなくても、学習により自動的に対称性や連続性を表現したセルの位置関係を選択し、車両の識別に有効な特徴セットを得ることができる。

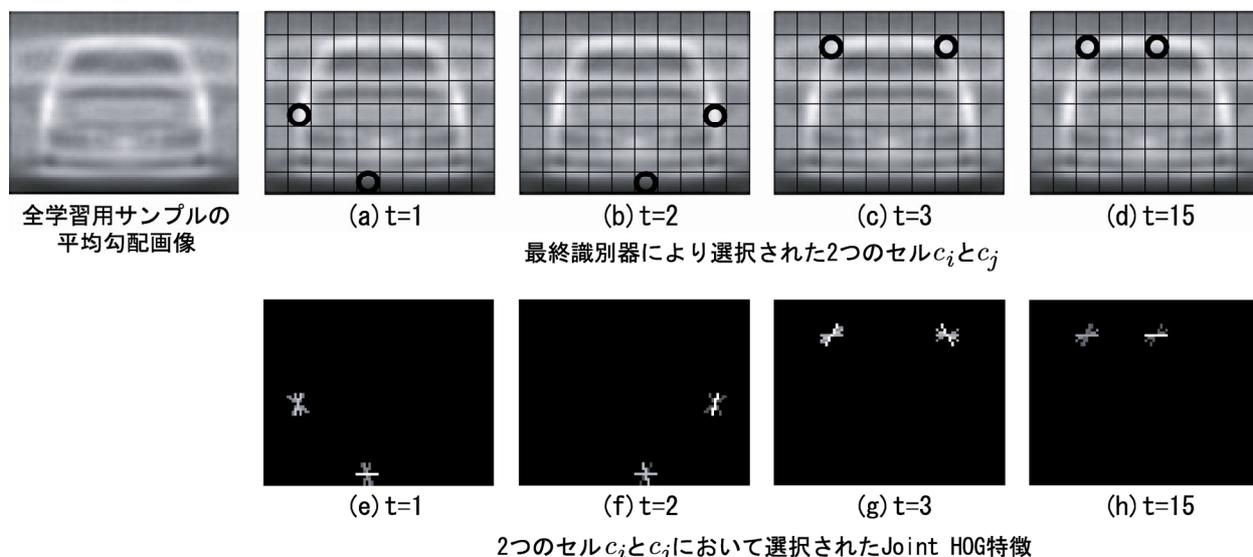


図 10：各弱識別器において選択された特徴

## 6. おわりに

本稿では、Joint HOG 特徴と 2 段階の AdaBoost による学習法を用いた車両検出する手法を提案し、その有効性を確認した。2 段階に AdaBoost を構築することで車両の対称的な勾配や連続的な勾配を自動的に選択し、識別に有効な特徴を得ることができた。今後は、時系列データの利用による高精度化を実現するとともに、今回提案した手法を利用し、車両領域のセグメンテーションを行う予定である。

## 文 献

- [1] G. Sullivan, A. Worrall, and J. Ferryman: "Visual Object Recognition Using Deformable Models of Vehicles", Context95, pp. 75-86, 1995.
- [2] O. Hasegawa and T. Kanade: "Type classification, color estimation, and specific target detection of moving targets on public streets", MVA, pp. 116-121. 2005.
- [3] 三谷重知, 海元, 塩山忠義: "ガボール変換による自動車検出", 信学論, pp. 2641-2651, 2000.
- [4] A. Iketani, A. Nagai, Y. Kuno, and Y. Shirai: "Real-Time Surveillance System Detecting Persons in Complex Scenes", IAPR, pp. 1112-1115, 1999.
- [5] A. Haselhoff, A. Kummert, and G. Schneider: "Radar-Vision Fusion with an Application to Car-Following using an Improved AdaBoost Detection Algorithm", ITSC, pp. 854-858, 2007.
- [6] N. Dalal and B. Triggs: "Histograms of Oriented Gradients for Human Detection", CVPR, pp. 886-893, 2005.
- [7] T. Mita, T. Kaneko and O.Hori: "Joint Haar-like Features for Face Detection", ICCV, pp. 1619-1626, 2005.
- [8] F. Han, Y. Shan, R. Cekaner: "A Two-Stage Approach to People

and Vehicle Detection with HOG-Based SVM", PerMIS, pp. 133-140, 2006.

- [9] Y. Freund and R. E. Schapire: "Experiments with a New Boosting Algorithm", ICML, pp. 148-156, 1996.
- [10] 鷲見和彦, 関真規人, 波部齊, "物体検出- 背景と検出対象のモデリング-", CVIM, pp. 79-98, 2005.
- [11] D. Comaniciu and P. Meer: "Mean shift analysis and applications", ICCV, pp. 1197-1203, 1999.
- [12] P. Sabzmeydani and G. Mori: "Detecting Pedestrians by Learning Shapelet Feature", CVPR, pp. 511-518, 2007.