

## 差分ステレオを用いた EOH 特徴量による車両検出

中央大学/JST CREST ○有江誠, 鈴木一尉, 生形徹, モロアレッサンドロ, 寺林賢司, 梅田和昇

### Car Detection Using Subtraction Stereo with EOH Features

Chuo Univ. / JST CREST Makoto ARIE, Kazui SUZUKI, Toru UBUKATA,  
Alessandro MORO, Kenji TERABAYASHI, Kazunori UMEDA

It is important to detect humans from an image automatically for the measurement of human flow with a surveillance camera. When measuring a human flow, cars appear in many cases. Therefore, it is necessary to separate human and car. We have proposed a stereo vision method called "Subtraction Stereo" which focuses distant measurement on foreground regions. A fast and stable car detection is achieved using the subtraction stereo with Edge of Orientation Histograms (EOH) features. The effectiveness of our method is verified by experiments.

#### 1. 序論

監視カメラからの人流計測のために、カメラ画像から自動的にリアルタイムで人物を検出することが期待されている。人物を対象とするシーンでは車両も映り込むことが多く、人物と車両を切り分けて検出する必要がある。物体検出によく用いられる Edge of Orientation Histograms (EOH)特徴量[1]は、計算量が多くリアルタイムでの処理が困難である。これまでに我々は、前景領域のみに距離計測を限定する差分ステレオを提案している[2]。そこで本研究では、差分ステレオより得られる前景領域と、その領域の三次元情報を用いて、EOH 特徴量による高速で誤検出の少ない車両検出手法を行うことを目的とする。

#### 2. 差分ステレオの概要

差分ステレオの基本アルゴリズムと計測例を Fig.1 に示す。通常のステレオでは、左右のカメラの画像をマッチングすることで視差画像を得る。これに対し、差分ステレオではまず左右カメラそれぞれで差分処理によって前景領域を抽出し、その後に抽出された前景領域をマッチングすることで距離情報を得る。この手法は、差分で得られた前景領域しか計測できないという欠点がある反面、対応点探索を左右カメラの前景領域に限定することができ、計算量の削減とともに、マッチングをロバスト化することができる。

#### 3. 車両検出

特徴量をベースとした物体検出に関する先行研究[3]では、特徴量を算出するウィンドウを様々なサイズで走査し、画像全体から物体検出を行っている。しかし、この手法では、大量の計算時間が必要となりリアルタイムでの車両検出は難しく、誤検出も多くなる。そこで、差分ステレオを用いて前景領域を抽出し、その領域での距離情報に応じて走査領域と走査するウィンドウの大きさを動的に変化させ、ウィンドウ毎に EOH 特徴量を算出することで計算時間と誤検出の問題を解

決する。また、統計的手法である Real AdaBoost 識別器を用いて、ウィンドウ内の前景領域の車両判定を行う。

#### 3. 1 ウィンドウの走査手法

本研究では、差分ステレオより得られる前景領域のみに対し、ウィンドウを走査する。また、カメラから前景領域までの距離と物体の進行方向よりウィンドウのサイズを決定する。ウィンドウの高さと幅は、擬似透視投影を仮定した以下の式より求める。

$$height = \frac{k_h}{W_Y} (\cos \theta - y \sin \theta) \quad (1)$$

$$width = \frac{k_w}{C_Z} \quad (2)$$

ここで、 $k_h, k_w$ は比例定数、 $W_Y$ は世界座標系での距離、 $C_Z$ はカメラ座標系での距離、 $\theta$ はカメラの仰角、 $y$ は焦点距離を 1 に正規化した画像座標である。前景領域に対しウィンドウサイズを動的に変化させることでウィンドウの走査回数を減少することができる。以上より、誤検出の少ないリアルタイムでの車両検出が可能になる。

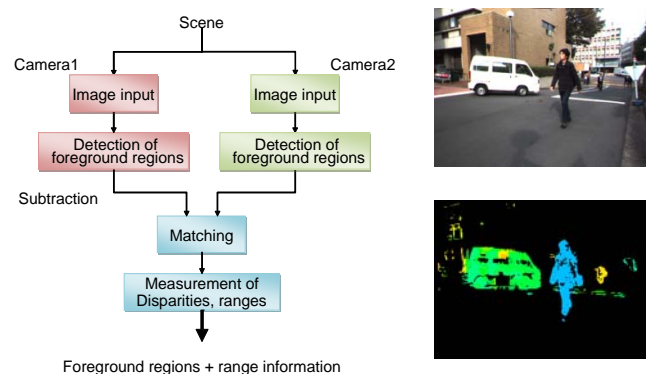


Fig.1 Schema of subtraction stereo

### 3. 2 EOH 特徴量の算出

EOH 特徴量とは、1 つの局所領域内におけるエッジ強度の勾配に着目した特徴量[1]である。勾配方向ごとの累積エッジ強度の比により与えられ、人や車などの形状変化を含む物体に有効である。また、積分画像を用いることで、高速に演算することが可能である。本研究では、320×240 の画像内の前景領域において EOH 特徴量を求めている。

### 3. 3 車両方向別識別器の構築

車両の形状は、向きによる見え方の違いにより多様である。多様な形状は車両の特徴を捉えにくくし、検出精度向上の妨げになる。従って、車両の進行方向に対応した識別器を用意する必要がある。本研究では、ステレオカメラから得られる三次元情報を用いて、車両の三次元における進行方向を 8 方向に分割している。具体的には、進行方向の前後、左右、斜めである。また、前後、左右、斜めそれぞれでの車両の見え方は変化が少ないため、作成する識別器は 4 つとした。

EOH 特徴量を用いた際の Boosting による学習の流れは次の通りである。まず検出対象と非検出対象の学習サンプルを多数用意し、各サンプルに重みを設定する。初期時はすべて同じ重みとする。そして、サンプルの中から人物検出に最も有効な EOH 特徴量を弱識別器として選択する。車両として正しく識別できるサンプルの重みは小さくし、誤って識別したサンプルの重みを大きくする。これらの処理を繰り返し行い、強識別器を学習する。Boosting は、次次に示すように、検出対象であるかを出力する性能の低い  $T$  個の弱識別器  $h_t(x)$  を線形結合することで、検出性能の良い強識別器  $H(x)$  を学習する。

$$H(x) = \text{sign} \left( \sum_{t=1}^T h_t(x_i) \right) \quad (3)$$

このように、Boosting を用いることで多数の特徴量候補の中から車両検出に有効な特徴量を学習し、識別器を構築することができる。本研究では、識別器の出力値が 2 値でなく実数である Real AdaBoost を用いる。識別器の出力に実数を用いることで、各サンプルに対し効率的な重みの更新が可能になる。

### 4. 車両検出精度評価実験

3.3 節で示した 4 つの車両方向別識別器を用いて車両検出の有効性を検証するため、条件の異なる 2 つのシーンにおいて、それぞれ 200 フレーム分で精度評価を行った。差分ステレオより前景領域を限定してウィンドウを走査させた場合を Controlled scan (proposed) とし、差分ステレオを用いずに画像全体に対してウィンドウを走査させた場合を Sliding scan とする。実験に使用したステレオカメラは Bumblebee2 (Point Grey Research, カラー,  $f=3.8\text{mm}$ ,  $b=120\text{mm}$ , VGA, 48fps), PC は DELL XPS420, Core2Duo 3.00GHz を使用した。Real AdaBoost 識別器用の学習用画像は車両である positive クラスが 2,066 枚、車両以外の negative クラスを 2,654 枚である。Controlled scan(proposed)と Sliding scan での検出率(TP), 誤検出率(FP), 適合率(TP/(TP+FP)), 処理速度(PT)を比較したものを Table 1,2 に示す。また、車両検出結果を Fig.2 に示す。

Table 1,2 より差分ステレオと EOH 特徴量を組み合わせることで、FP と PT の性能が向上しているのが分かる。これは、ウィンドウの走査領域を前景領域に限定し、サイズを動的に変化させたためだと考えられる。また、画像 1 枚当たりでのウィンドウの数と EOH 特徴量の次元数が大幅に減少したため、リアルタイムでの車両検出が可能だと考えられる。

### 5. 結論と今後の展望

差分ステレオと EOH 特徴量による車両検出手法を提案した。また、提案手法の有効性を条件の異なる複数のシーンで行う事で検証した。

今後の展望としては、カメラの設置角度を考慮した識別器を用意し、識別器構築用のサンプル画像を増やすことで車両検出の精度向上させることが挙げられる。また、オクルージョンが生じた際の対処も今後の課題である。

Table 1 Experimental results of scene 1 (gate)

	TP (%)	FP (%)	Precision (%)	PT (ms)
Controlled scan (proposed)	67.5	4.0	94.4	45.7
Sliding scan	65.0	41.0	61.3	272.7

Table 2 Experimental results of scene 2 (local street)

	TP (%)	FP (%)	Precision (%)	PT (ms)
Controlled scan (proposed)	81.5	3.5	95.9	43.1
Sliding scan	80.5	27.0	74.9	260.1



(a) Scene 1 (gate)



(b) Scene 2 (local street)

Fig.3 Car detection results

### 参考文献

- [1] Y.T. Chen, C.S. Chen, "A Cascade Feed-Forward Classifiers for Fast Pedestrian Detection", ACCV 2007, Part I, LNCS 4843, pp.905-914, 2007.
- [2] K. Umeda, et al, "Subtraction Stereo-A Stereo Camera System That Focuses on Moving Regions-", Proc. of SPIE-IS&T Electronic Imaging, Vol. 7239 Three-Dimensional Imaging Metrology, 723908, 2009.
- [3] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection", International Conference on Computer Vision and Pattern Recognition, Vol.2, pp 886-893, June 2005.