村 井 泰 裕<sup>†1</sup> 藤 吉 弘 亘<sup>†1</sup> 金 出 武 雄<sup>†2</sup>

本論文では,Space-Time Patch 特徴を用いて物体の移動方向識別とセグメンテー ションを同時に行う手法を提案する.提案手法は,複数の学習サンプルを用いて作成 した Space-Time Patch 特徴のコードブックを参照したベクトル量子化により,移動 方向識別と対象物体の重心位置推定を行う.そして,推定された物体の重心位置から, 学習サンプルのマスク画像を用いた確率計算に基づいて物体領域のセグメンテーショ ンを行う.これにより,移動方向が異なる物体が交差する際にオクルージョンが発生 していても,移動方向を識別しているためセグメンテーションが可能となる.評価実 験の結果,アピアランス情報を用いた従来法より,オクルージョン発生時における対 象物体の検出率を向上させることができた.また,Lucas-Kanade 法による移動物体 検出法との比較実験より,提案手法の有効性を示す.

> Classifying Direction of Movement and Segmenting Regions of Objects by Using Space-Time Patches

Yasuhiro Murai,<sup>†1</sup> Hironobu Fujiyoshi<sup>†1</sup> and Takeo Kanade<sup>†2</sup>

We developed a method for classifying the direction of movement and for segmenting objects simultaneously using features of space-time patches. Our approach uses vector quantization to classify the direction of movement of an object and to estimate its centroid by referring to a codebook of the space-time patch feature, which is generated from multiple learning samples. We segment the objects' regions based on the probability calculated from the mask images of the learning samples by using the estimated centroid of the object. Even though occlusions occur when multiple objects overlap in different directions of movement, our method detects objects individually because their direction is more accurate with our method than with the conventional method, which is only based on appearance features. And, we show the effectiveness of our approach by comparing with the movement object detection using Lucas-Kanade method.

## 1. はじめに

近年,監視映像からの不審者の発見や,ITS における歩行者の安全のため,カメラでと らえた画像から自動的に対象物体の識別を行う技術の実現が期待されている.Fujiyoshi ら は,屋外に設置したカメラ映像からフレーム間差分により検出された移動物体の形状から, ニューラルネットワークを用いて人や車に識別し,その結果を地図上へマッピングすること により,歩行者であるユーザの移動支援を行う RCT システム(Robotic Communication Terminals)<sup>1)</sup>を提案している.また,藤本らは,店舗内の全方位視覚センサから得た映像 から顧客の位置や移動方向を推定し,その情報を用いて空間マップを作成することにより, 顧客全体の店内における分布,各顧客の移動軌跡および停留状況等をユーザに提示するシス テム<sup>2)</sup>を提案している.このような画像処理技術を用いたシステムでは,画像上の移動物 体の動きの情報を得ることや,対象物体の領域を検出することは非常に重要な問題である.

画像上の移動物体の検出・動きの解析を行う手法として,物体の移動量をベクトル表現し たオプティカルフロー<sup>3),4)</sup>があげられる.オプティカルフローでは,ノイズが発生しやすい ことや,非剛体の物体に対する動きの解析,テクスチャの異なる物体どうしの動きの評価が 困難である等の問題がある.これらの問題に対して Shechtman らは,テンプレート映像と 入力映像の時空間画像から取得された2つの Space-Time Patch (ST-patch)の相関関係 を複数用いることにより,テンプレート映像の物体と同じ動きをしている物体を入力映像か ら識別する手法<sup>5)</sup>を提案している.この手法で用いられているST-patchの相関関係(動き の類似度)は,テクスチャが異なる対象物体の動きや,非剛体の物体のように複雑に動く物 体に対しても動きの評価を行うことができる.また,このような局所的な時系列情報の学習 を行うことにより物体の動きの識別を行う手法<sup>6)</sup>も提案されている.

一方,対象物体の検出法には,物体の局所的な特徴を用いた part-based approach があ

<sup>†1</sup> 中部大学工学部情報工学科

Department of Computer Science, Chubu University

<sup>†2</sup> カーネギーメロン大学ロボティクス工学研究所

The Robotics Institute, Carnegie Mellon University

る.Agarwalらは,類似したアピアランス情報を持つ学習サンプル画像パッチをクラス分 けをしたコードブックを作成し,入力画像の画像パッチとコードブックをマッチングさせる ことにより物体検出を行う手法<sup>7)</sup>を提案している.Leibeらは,各画像パッチに物体の重心 位置,物体のマスク画像を格納したコードブックの作成を行い,コードブック内の学習サン プル画像パッチに格納してある物体の重心位置,物体のマスク画像から,物体の識別とセグ メンテーションを行う手法<sup>8),9)</sup>を提案している.Opeltらは,物体の輪郭の学習サンプルか ら物体の重心位置の投票により重心位置推定を行い,重心位置に投票された物体の輪郭の 位置関係からAdaBoostを用いて物体検出を行う手法<sup>10)</sup>を提案している.また,局所領域 における輝度の勾配方向をヒストグラム化した特徴ベクトルである HOG (Histograms of Oriented Gradients)を用いた人検出法<sup>11)</sup>,Haar-like特徴を用いた顔検出法<sup>12)</sup>が提案さ れている.このように,part-based approachを用いた手法は,物体の部分的な特徴から学 習サンプルを用いて対象物体の検出を行うことで,オクルージョンへの対応が可能となるメ リットがあり,数多くの手法が提案されている.

part-based approach を用いた手法では,主に1枚の画像のアピアランス情報に基づいて 対象物体の検出を行っているため,対象物体どうしの重なりが生じると検出が難しいという 問題がある.そこで本論文では,局所的なアピアランス(見え)とモーション(動き)の情 報を持つ ST-patch 特徴を用いて,物体の移動方向識別とセグメンテーションを同時に行う 手法を提案する.提案手法では,まず複数のシーンにおける学習サンプルからあらかじめ作 成した ST-patch 特徴のコードブックを参照し,入力画像から得られる ST-patch 特徴のベ クトル量子化を行う.次に,学習サンプル ST-patch 特徴に格納してある物体の重心位置を 投票し,物体の重心位置推定を行う.そして,推定された物体の重心位置から,学習サンプ ル ST-patch 特徴に格納してある物体のマスク画像を用いることで物体のセグメンテーショ ンを実現する.これにより,オクルージョン発生時における対象物体の検出率の向上が期待 できる.

## 2. ST-patch

提案手法では, part-based approach で用いられている画像パッチの代わりに, Shechtman らにより提案された ST-patch 特徴を使用する.ST-patch 特徴は,時空間画像における画素 の動きが異なる場合(たとえば,左移動と右移動)では,それぞれ異なる特徴が得られる.ゆ えに,複数のシーンにおける学習サンプルより ST-patch 特徴を取得し,類似した ST-patch 特徴をクラスタリングすることにより,アピアランスとモーションの情報ごとのクラスタで



Fig. 1 Overview of the ST-patch.

分けられたコードブックを作成することができる.本章では,ST-patch 特徴の概要とその 性質,学習サンプルST-patch 特徴の取得方法,コードブックの作成方法について述べる.

## 2.1 ST-patch の概要

ST-patch 特徴は,画像を時間方向に重ねた3次元データである時空間画像の局所領域から得られる特徴である.図1にST-patchの概念を示す.x,yは画像の座標軸,tは時間軸,3つのラインは個々の画素の動き, $[u v w]^{T}$ はST-patch中の動き方向ベクトル, $\nabla P_i$ は個々の画素の勾配方向ベクトルを表している.

2.2 ST-patch 特徴

ST-patch 特徴を得るため, x 軸, y 軸, t 軸の勾配を求める.ここで,画像中の動きが一定の場合,画素 i における  $P_{x_i}$ ,  $P_{y_i}$ ,  $P_{t_i}$  の 3 つの勾配を持つ時空間勾配  $\nabla P_i$  は,画素の動き方向ベクトル  $[u v w]^T$  に対して垂直となるため,式 (1)の関係が成り立つ.

$$\nabla P_i \begin{bmatrix} u \\ v \\ w \end{bmatrix} = 0 \tag{1}$$

画素数が n の場合,式 (1) は式 (2) となる.

$$\begin{bmatrix} P_{x_1} & P_{y_1} & P_{t_1} \\ P_{x_2} & P_{y_2} & P_{t_2} \\ \vdots & \vdots & \vdots \\ P_{x_n} & P_{y_n} & P_{t_n} \end{bmatrix}_{n \times 3} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n \times 1}$$
(2)

ST-patch 中の n 画素の  $\nabla P_i$  からなる  $n \times 3$  の行列を G とし, 行列 G<sup>T</sup> を掛けると式 (3) となる.

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\begin{bmatrix}u\\v\\w\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}_{3\times 1} \tag{3}$$

ここで,行列  $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ を  $\mathbf{M}$ とすると式 (4) のように表すことができる.

$$\mathbf{M} = \begin{bmatrix} \sum P_x^2 & \sum P_x P_y & \sum P_x P_t \\ \sum P_y P_x & \sum P_y^2 & \sum P_y P_t \\ \sum P_t P_x & \sum P_t P_y & \sum P_t^2 \end{bmatrix}$$
(4)

式 (4) より求められる行列 M が 1 つの ST-patch から得られる特徴となる.行列 M に は,左上の 2 × 2 の行列にアピアランスの情報,3 行目と3 列目にはモーションの情報を持 つ.この行列 M を 9 次元ベクトルとして表現した行列 e を ST-patch 特徴とする.

$$\mathbf{e} = \left(\sum P_x^2, \sum P_x P_y, \cdots, \sum P_t^2\right) \tag{5}$$

## 2.3 ST-patch 特徴の性質

図2に背景(Background),左移動(Leftward),右移動(Rightward)する物体の局所 領域の勾配値 $P_{x_i}$ , $P_{y_i}$ , $P_{t_i}$ をそれぞれプロットした図を示す.それぞれの画像中の矩形領 域は勾配値を求めたST-patchの領域(50×50[pixel]×3[frame]),図中の線と数値は主 成分分析により求められた第1主成分の軸とその寄与率を表している.また,表1に水平 方向勾配値と時間勾配値,表2に垂直方向勾配値と時間勾配値における第1主成分(First Principal Component: 1st PC)と第2主成分(Second Principal Component: 2nd PC) の寄与率を示す.第1主成分の軸はプロットされた勾配値の分布の傾向を示し,寄与率が高 いほどその軸方向への分布が強いことを意味する.背景の図を見ると,横方向に分布の傾向 を表す軸が出ており,寄与率も高いことが分かる.これは,動きのない背景では時間勾配値 が低いためである.右移動と左移動の垂直方向勾配値と時間勾配値の図を見ると,縦方向へ 分布の傾向を表す軸が出ているが,寄与率は低い値となっている.これより,垂直方向勾配 値と時間勾配値の分布は広く分散しているといえる.次に,右移動と左移動の水平方向勾 配値と時間勾配値の図を見ると,第1主成分の軸がそれぞれ対角方向になっていることが 分かる.また,その寄与率はともに高い値となっており,第1主成分方向へ強く分布して



#### 表 1 水平方向勾配値と時間勾配値における寄与率

Table 1 Contribution ratio of horizontal gradients and temporal-spatial gradients.

	Background	Leftward	Rightward
1st PC	0.9358	0.7839	0.7754
2nd PC	0.0642	0.2161	0.2246

#### 表 2 垂直方向勾配値と時間勾配値における寄与率

Table 2 Contribution ratio of vertical gradients and temporal-spatial gradients.

	Background	Leftward	Rightward
1st PC	0.9964	0.5724	0.5692
2nd PC	0.0036	0.4376	0.4308

いることが分かる.このように,移動方向の異なる物体の勾配値からは,それぞれ違った特徴を得ることができる.このような性質を9次元ベクトルとして表した特徴量がST-patch 特徴である.

2.4 学習サンプル ST-patch 特徴の取得

ST-patch 特徴は x 軸, y 軸, t 軸の勾配値より求められる.そのため, ST-patch 特徴を 算出するポイントは, これらの勾配情報がより現れていることが望ましいと考えられる.そ こで提案手法では,学習フェーズにおいて, ST-patch 特徴を算出するポイントとしてハリ スのコーナ検出<sup>13)</sup>を用いる.ハリスのコーナ検出によって得られた特徴点の周辺領域から ST-patch 特徴を算出し,学習サンプルとして取得する.図3に学習サンプル ST-patch 特 徴の取得例を示す.

2.5 コードブック作成

ST-patch 特徴のコードブックを作成するために,本研究では LBG アルゴリズム<sup>14)</sup> を用 いる.これにより,学習サンプルの特徴ベクトルを N 個の代表ベクトルの組にクラスタリ ングする.取得された学習サンプル ST-patch 特徴を用いてコードブックを作成する流れを 以下に述べる.

Step1 複数のシーンより学習サンプル ST-patch 特徴を取得.

Step2 学習サンプル ST-patch 特徴に移動物体ラベル,物体の重心位置,物体のマスク画像の情報を格納.

Step3 LBG アルゴリズムにより N 個のクラスタに分けたコードブックを作成.



Step4 コードブックのクラスタに属する移動物体ラベルから移動物体確率  $p(o_d \mid I)$  を 計算.

学習サンプル ST-patch 特徴を取得した後,移動物体ラベル  $o_d = \{\text{Left, Right, Bg}\}$ を各 ST-patch 特徴に割り当てる.ここで,LBG アルゴリズムにより ST-patch 特徴のコード ブックを作成した場合,クラスタ内のラベルは必ずしもすべて同じラベルが含まれている とは限らない.しかし,クラスタ内には同じ移動物体ラベルの割合が高くなるという傾向 がある(図4).そこで,クラスタ内の移動方向ラベルの割合から,移動物体確率  $p(o_d | I)$ を計算する.ここで,I はある1つのクラスタを表す.この移動物体確率により,各クラス タの信頼性を表現する.また,物体の重心位置推定と物体のセグメンテーションのため,各 学習サンプル ST-patch 特徴に物体の重心位置,あらかじめ作成しておいた物体のマスク画 像の情報を格納しておく.



図 4 移動物体確率例 (N = 8) Fig. 4 Probability for direction of movement (N = 8).

25 Space-Time Patch を用いた物体の移動方向識別とセグメンテーション



3. 物体の移動方向識別とセグメンテーション

学習サンプルより作成した ST-patch 特徴のコードブックを参照することにより,入力画 像から取得された ST-patch 特徴のベクトル量子化を行う.そして,Leibe らにより提案され た手法を応用し,識別された移動方向結果と学習サンプル ST-patch 特徴から物体の重心位 置の投票を行い,そこから重心位置の推定をする.このように推定された物体の重心位置周 辺に投票された学習サンプル ST-patch 特徴に対応したマスク画像を用いて,物体のセグメ ンテーションを行う.図5に物体の移動方向識別とセグメンテーションの流れを示す.本章 では,物体の移動方向識別とセグメンテーションを行うための各処理の詳細について述べる.

3.1 ST-patch 特徴のベクトル量子化

入力画像より取得された ST-patch 特徴を,作成したコードブックを用いてベクトル量子 化する.入力 ST-patch 特徴のベクトル量子化による移動方向識別の流れを以下に述べる. Step1 入力画像のラスタスキャンにより取得されたパッチを学習サンプルと同じサイズに ダウンサンプリングし,ST-patch 特徴を取得.

Step2 入力 ST-patch 特徴を e, 作成したコードブックの各クラスタの代表ベクトルを c とし, 2 つの特徴ベクトル間のユークリッド距離を計算し, ユークリッド距離が最小の クラスタ *I* を式 (6) により算出.

$$I = \operatorname{argmin}_{\mathbf{c}} \| \mathbf{e} - \mathbf{c} \|^2 \tag{6}$$

Step3 入力画像から取得するパッチのサイズを変更.



Fig. 6 Example of classifying direction of movement by vector quantization.

# Step4 画像全体の処理が終了するまで Step1 ~ Step3 を繰り返す.

このようにパッチのサイズを変更し、繰り返して処理を行うことにより、物体のスケール 変化に対応させることができる、図 6 に入力 ST-patch 特徴のベクトル量子化により求め られクラスタ *I* の移動物体確率  $p(o_d | I)$  を基に、移動方向識別を行った例を示す、このよ うに、作成した ST-patch 特徴のコードブックは移動方向ごとにクラスタリングされている ことが分かる、



#### 3.2 物体の重心位置推定

入力 ST-patch 特徴をコードブックを参照してベクトル量子化を行う.これにより最も類 似した代表ベクトルを持つクラスタを求め,そのクラスタに属する学習サンプルに格納され ている重心位置を投票していく.そして,投票点の最も密な位置を Mean-Shift クラスタリ ング<sup>15)</sup> により求めることで,物体の重心位置推定を行う.

**3.2.1** 重心位置の投票

物体の重心位置を推定するため,物体の重心位置の投票を行う.画像座標とスケール l よ り取得された入力 ST-patch 特徴 e のベクトル量子化を行うことでクラスタ I が得られる. ここで, p(I | e, l) の確率は,入力 ST-patch 特徴 e とコードブッククラスタ I の相対的な マッチングコストを用いる.そして,求められたクラスタ I からは,移動物体  $o_d$  の重心位 置が x である確率  $p(o_d, x | I, l)$  を得ることができる.よって,移動物体  $o_d$  の重心位置 xの確率は式 (7) で求めることができる.

$$p(o_d, x \mid \mathbf{e}, l) = p(o_d, x \mid \mathbf{e}, I, l) p(I \mid \mathbf{e}, l)$$
(7)

ここで,式(7)は式(8),式(9)と表すことができる.

$$p(o_d, x \mid \mathbf{e}, l) = p(o_d, x \mid I, l) p(I \mid \mathbf{e})$$
(8)

$$= p(x \mid o_d, I, l) p(o_d \mid I, l) p(I \mid \mathbf{e})$$
(9)

 $p(x | o_d, I, l)$ は,入力 ST-patch 特徴をベクトル量子化して求められたクラスタ *I* と移動 物体  $o_d$  から得られる重心位置 x を表す.そして, $p(o_d | I, l)$ は,コードブックによりベク トル量子化して求められる移動物体  $o_d$  を表し,p(I | e)は,入力 ST-patch 特徴 e とコー ドブッククラスタ *I* の相対的なマッチングコストを表している.複数の ST-patch (k 個) から得られるウィンドウ W(x) 内の投票点の確率のスコアは式 (10) により得られる.

$$score\left(o_{d}, x\right) = \sum_{k} \sum_{x_{j} \in W(x)} p\left(o_{d}, x_{j} \mid \mathbf{e}_{k}, l_{k}\right)$$

$$(10)$$

## **3.2.2** 重心位置の推定

図7 に重心位置推定の流れを示す.入力画像より取得されたパッチをダウンサンプリン グし,そのパッチから取得された ST-patch 特徴(図7(a))のベクトル量子化(図7(b)) を行い,重心位置の投票を行う.次に,識別された移動方向ごとに投票点が最も密な位置を 3 次元空間における Mean-Shift クラスタリング(x,y,scale)により探索する(図7(c)). そして,探索された位置の周辺に投票された点の重みの合計(式(10))が閾値以上の場合, 物体の重心位置推定結果とする(図7(d)).

3.2.3 Mean-Shift クラスタリング

注目点を x , 各投票点の位置を x<sub>i</sub> (i = 1, ..., n), バンド幅のパラメータを h, カーネル 関数を k とすると , Mean-Shift ベクトル  $\mathbf{m}_h(\mathbf{x})$  は式 (11) により求められる .

$$\mathbf{m}_{h}\left(\mathbf{x}\right) = \frac{\sum_{i=1}^{n} \mathbf{x}_{i} k\left(\left\|\frac{\mathbf{x}-\mathbf{x}_{i}}{h}\right\|^{2}\right)}{\sum_{i=1}^{n} k\left(\left\|\frac{\mathbf{x}-\mathbf{x}_{i}}{h}\right\|^{2}\right)} - \mathbf{x}$$
(11)

カーネル関数 k は式 (12) のように定義する.

$$k\left(\mathbf{x}\right) = \exp\left(-\frac{1}{2} \| \mathbf{x} \|^{2}\right) \tag{12}$$

このようなカーネル関数 k を用いることにより, 注目点と投票された点が近いほど大き

情報処理学会論文誌 コンピュータビジョンとイメージメディア Vol. 1 No. 2 21-31 (July 2008)

© 2008 Information Processing Society of Japan





$$\mathbf{x}^{t+1} = \mathbf{x}^t + \mathbf{m}_h \left( \mathbf{x}^t \right) \tag{13}$$

求められた Mean-Shift ベクトル  $\mathbf{m}_h(\mathbf{x}^t)$  が閾値以下に収束するまで繰り返すことにより,物体の重心位置推定を行う.

3.3 物体領域のセグメンテーション

図 8 に物体領域のセグメンテーションの流れを示す.推定された物体の重心位置周辺に 投票した学習サンプル ST-patch 特徴と,そこに格納されている物体のマスク画像を用いて 物体領域を推定し,物体のセグメンテーションを行う.

3.3.1 ST-patch 特徴の逆投影

推定した物体の重心位置周辺へ投票した ST-patch 特徴の逆投影を行う.これにより,重心位置外の投票点(アウトライア)の除去を行い,信頼性の高い情報を選択する.次に, 学習サンプル ST-patch 特徴の各投票点が,ある移動物体  $o_d$  の重心位置が x である確率  $p(o_d, x \mid \mathbf{e}, l)$ の事後確率を式 (14) により求める.

$$p(\mathbf{e}, l \mid o_d, x) = \frac{p(o_d, x \mid \mathbf{e}, l) p(\mathbf{e}, l)}{p(o_d, x)} = \frac{p(o_d, x \mid I, l) p(I \mid \mathbf{e}) p(\mathbf{e}, l)}{p(o_d, x)}$$
(14)

ここで, $p(o_d, x | \mathbf{e}, l)$ は式 (7)より得られる.

#### **3.3.2** 物体領域の推定

物体領域の推定を行うために、ピクセル p が移動物体である確率  $p(\mathbf{p} = obj. | o_d, x)$  を 求める.これは式 (15) で定義するように、各クラスタに属する学習サンプル ST-patch 特 徴に格納されているマスク画像から得られる各ピクセルの情報と、式 (14) で定義した事後 確率を掛けることにより求める.

$$p\left(\mathbf{p} = obj. \mid o_d, x\right) = \sum_{num} p\left(\mathbf{p} = obj. \mid o_d, x, \mathbf{e}, l\right) p\left(\mathbf{e}, l \mid o_d, x\right)$$
(15)

ここで, *num* は逆投影された ST-patch の数を表している.よって, 各ピクセルが移動物 体である確率は式 (16) により表すことができる.

$$p\left(\mathbf{p} = obj. \mid o_{d}, x\right) = \sum_{num} p\left(\mathbf{p} = obj. \mid o_{d}, x, \mathbf{e}, I, l\right) p\left(\mathbf{e}, I, l \mid o_{d}, x\right)$$
$$= \sum_{num} p\left(\mathbf{p} = obj. \mid o_{d}, x, I, l\right) * \frac{p\left(o_{d}, x \mid I, l\right) p\left(I \mid \mathbf{e}\right) p\left(\mathbf{e}, l\right)}{p\left(o_{d}, x\right)}$$
(16)

式 (16) では,各ピクセルにおいて移動物体である確率の加重平均が求められる.このように求められた各ピクセルの物体である確率から物体のセグメンテーションを行う.この確率が閾値以下のピクセルを背景,閾値以上のピクセルを移動物体とすることにより,物体領域の矩形によるセグメンテーションを行う.

情報処理学会論文誌 コンピュータビジョンとイメージメディア Vol. 1 No. 2 21-31 (July 2008)

## 4. 評価実験

提案手法を用いて物体の移動方向識別とセグメンテーションの実験を行う.評価実験では提 案手法の有効性を示すため,アピアランス情報を用いて物体の識別とセグメンテーションを行 う手法<sup>8)</sup> との比較,オプティカルフローを求める代表的な手法の1つである Lucas-Kanade 法<sup>4)</sup> による移動物体検出法との比較を行う.

## 4.1 実験概要

表3に歩行者,車両それぞれの学習サンプルから取得したST-patch特徴の数を示す.このような数の学習サンプルを用いてそれぞれのコードブックを作成し,物体の移動方向識別とセグメンテーションを行う.実験では,ST-patchのサイズを15×15[pixel]×3[frame]とし,LBGアルゴリズムにより作成するST-patch特徴のコードブックサイズを512とする.

従来法と提案手法の比較実験に用いるシーケンスには,固定カメラにより学習サンプルと 異なる場所で撮影された,歩行者が左右に移動している映像,車両が左右に移動する2車線 の道路の映像を用いる.

#### 4.2 実験結果

図9にアピアランス情報を用いた物体検出法の結果,Lucas-Kanade 法による移動物体 検出結果,提案手法による物体の移動方向識別とセグメンテーション,右移動の物体領域推 定,左移動の物体領域推定の結果を示す.以下に提案手法の検出結果,検出率の定量的評 価,Lucas-Kanade 法との比較,提案手法の問題点について述べる.

**4.2.1** 提案手法の検出結果

図9より,提案手法を用いることで,物体の移動方向の識別と物体のセグメンテーション を同時に行うことができていることが分かる.

図 9(a) に示すように,提案手法は異なる方向へ移動している物体が重なる場合において も,個々に物体領域のセグメンテーションを行うことが可能である.これは,移動方向識別

表 3 学習サンプル ST-patch 特徴の数 Table 3 Number of learning sample ST-patch features.

	Sequence	
	Pedestrian	Vehicle
Background	36,982	20,047
Leftward	10,220	9,968
Rightward	10,198	9,885

結果を基に物体の重心位置の推定を別々に行い,物体のセグメンテーションをしているため である.また,提案手法は影の影響を受けにくいことが分かる.

図9(b)では,足の部分にオクルージョンが発生した歩行者や,柱に体の一部が隠れている歩行者においても,物体領域を考慮して物体のセグメンテーションが可能であることが分かる.これは,画像上に写っている歩行者の部分的な情報を基に物体の重心位置を求め,そこから物体のセグメンテーションをしているためである.

図 9 (c) は,物体のスケールが変化した場合の結果である.提案手法では物体のスケール 変化に対応した物体の移動方向の識別と物体のセグメンテーションを行うことができる.

図 9 (d) では、学習サンプルとして車両を用いた場合の結果である.提案手法では、学習 サンプルをそれぞれ用意し、そのコードブックを作成することで、様々な物体の移動方向の 識別と物体のセグメンテーションを行うことが可能である.

4.2.2 検出率の定量的評価

提案手法とアピアランス情報を用いた物体検出法との検出率を比較することにより,定量 的評価を行う.表4に提案手法とアピアランス情報を用いた物体検出法の検出率を示す.比 較実験には,歩行者を対象とした図9(a)のシーケンス,車両を対象とした図9(d)のシー ケンスを用いた.識別対象フレームは画像中に物体が存在しているフレームのみとし,物体 の検出と領域のセグメンテーションが成功したフレームを正解とする.

表4より,提案手法はアピアランス情報のみを用いる従来法と比べ,歩行者,車両ともに 検出率が向上していることが分かる.従来法では,物体が重なった状態の検出を行うことは 困難であるが,提案手法では移動方向が異なる物体が重なった状態においても検出すること ができるため,検出率を約16.7%向上させることができた.

4.2.3 Lucas-Kanade 法との比較

図 9 に示すように, Lucas-Kanade 法による移動物体検出法では,移動物体領域の検出 とその移動方向の識別を行うことが可能である.しかし,物体のオクルージョン領域や,物

表 4 検出率 [%] Table 4 Detection rate [%].

	$\frac{\text{Conventional}}{\text{method}^{8)}}$	Proposed method
Pedestrian	64.3%	74.7%
Vehicle	70.7%	93.3%
Average	67.3%	84.0%

情報処理学会論文誌 コンピュータビジョンとイメージメディア Vol. 1 No. 2 21-31 (July 2008)



Fig. 9 Example of classifying direction of movement and segmenting the objects' regions.

情報処理学会論文誌 コンピュータビジョンとイメージメディア Vol. 1 No. 2 21-31 (July 2008)

© 2008 Information Processing Society of Japan



体どうしが重なった場合はフローの検出が不可能である.提案手法では,学習サンプルとした ST-patch 特徴を用いて,物体の重心位置を求め,マスク画像から物体領域を推定するため,Lucas-Kanade 法におけるこれらの問題点に対応することが可能となる.

4.2.4 提案手法の問題点

提案手法の問題点として,図10(a)のように同一方向へ移動している物体が重なった場合,投票された点から物体の重心位置の推定を行うことが困難であるため,物体のセグメンテーションに失敗することがあげられる.これは,個々の物体の重心位置投票の精度が低く,投票点が分散してしまうためである.

また,図10(b)のように,学習サンプルに存在しない移動物体(バス等)の場合,物体の 識別を行うことができないため,誤識別を起こすことがある.これらの問題は,ST-patch特 徴にアピアランス情報が少ないことが原因であると考えられる.そのため,このST-patch 特徴にアピアランスを表す情報を加えることによって,これらの問題に対処することが可能 となると考えられる.

5. おわりに

本論文では,ST-patch 特徴を用いて物体の移動方向識別とセグメンテーションを同時に 行う手法を提案した.提案手法では,物体にオクルージョンがある場合においても,物体領 域を考慮したセグメンテーションが可能である.このように,移動方向識別結果を基に物体 のセグメンテーションを行うことで,移動方向が異なる物体が重なった状態においても物体 のセグメンテーションが可能となり,アピアランス情報を用いた従来法より,オクルージョ ン発生時における対象物体の検出率を向上させることができた.

今後は,提案手法に用いた ST-patch 特徴にアピアランス情報を付加することで,より正確な物体の重心位置の投票を行い,物体のセグメンテーション精度を向上させる予定である.

# 参考文献

- Fujiyoshi, H., Komura, T., Yairi, I.E. and Kayama, K.: Road Observation and Information Providing System for Supporting Mobility of Pedestrian, *IEEE International Conference on Computer Vision Systems*, pp.37–44 (2006).
- 2) 藤本喜彦,小原ゆう,柴田史久,馬場口登,八木康史:店舗内に設置した全方位視覚 センサによる顧客の行動解析,情報処理学会研究報告,No.142, pp.17-22 (2004).
- Horn, B.K.P. and Schunck, B.G.: Determining optical flow, Artificial Intelligence, Vol.17, pp.185–203 (1981).
- 4) Lucas, B.D. and Kanade, T.: An Iterative Image Registration Technique with an Application to Stereo Vision, *Proc. 7th International Joint Conference on Artificial Intelligence*, pp.674–679 (1981).
- Shechtman, E. and Irani, M.: Space-Time Behavior Based Correlation, Computer Vision and Pattern Recognition, Vol.1, pp.405–412 (2005).
- 6) Niebles, J.C., Wang, H. and Fei-Fei, L.: Unsupervised learning of human action categories using spatial-temporal words, *British Machine Vision Conference*, Vol.3, pp.1249–1258 (2006).
- Agarwal, S. and Roth, D.: Learning a Sparse Representation for Object Detection, European Conference on Computer Vision, pp.113–130 (2002).
- 8) Leibe, B., Leonardis, A. and Schiele, B.: Interleaved Object Categorization and Segmentation, *British Machine Vision Conference*, Norwich, pp.759–768 (2003).
- 9) Leibe, B., Leonardis, A. and Schiele, B.: Combined Object Categorization and Segmentation with an Implicit Shape Model, *European Conference on Computer* Vision, Prague, pp.496–510 (2004).
- Opelt, A., Pinz, A. and Zisserman, A.: Incremental learning of object detectors using a visual shape alphabet, *Computer Vision and Pattern Recognition*, Vol.1, pp.3–10 (2006).
- Dalal, N. and Triggs, B.: Histograms of Oriented Gradients for Human Detection, IEEE Computer Vision and Pattern Recognition, pp.886–893 (2005).
- 12) Viola, P. and Jones, M.: Rapid Object Detection using a Boosted Cascade of Simple Features, *Computer Vision and Pattern Recognition*, Vol.1, pp.511–519 (2001).
- 13) Harris, C. and Stephens, M.: A Combined Corner and Edge Detector, Proc. 4th

Alvey Vision Conference, pp.147–151 (1988).

- 14) Linde, Y., Buzo, A. and Gray, R.M.: An Algorithm for Vector Quantizer Design, *IEEE Trans. Communications*, Vol.28, No.1, pp.84–95 (1980).
- Comaniciu, D. and Meer, P.: Mean Shift Analysis and Applications, *International Conference on Computer Vision*, Vol.2, pp.1197–1203 (1999).

(平成 19 年 9 月 22 日受付)(平成 20 年 3 月 10 日採録)

## (担当編集委員 斎藤 英雄)



# 村井 泰裕(学生会員)

2007年中部大学工学部情報工学科卒業.同年より同大学大学院工学研 究科情報工学専攻に在籍.画像情報処理・コンピュータビジョンに関する 研究に従事.



会員.

# 藤吉 弘亘(正会員)

1997年中部大学大学院博士後期課程修了.1997~2000年カーネギーメ ロン大学ロボット工学研究所 Postdoctoral Fellow.2000年中部大学講師 を経て2004年より同大学助教授.2006年カーネギーメロン大学ロボット 工学研究所客員研究員.工学博士.計算機視覚,動画像処理,パターン認 識・理解の研究に従事.2005年度ロボカップ研究賞.電気学会,IEEE各



## 金出 武雄

1974 年京都大学大学院工学研究科電子工学専攻博士課程修了(工学博 士).1976 年同大学助教授.1980 年カーネギーメロン大学計算機科学科 ロボット研究所高等研究員.1985 年同大学教授.1992~2001 年カーネ ギーメロン大学ロボット研究所所長.1993 年カーネギーメロン大学 U.A. and Helen Whitaker 記念全学教授.2001 年産業技術総合研究所デジタル

ヒューマン研究センター長(非常勤).計算機視覚,マルチメディア,センサ,医用ロボット, 自律ヘリコプタ等の研究に従事. National Academy of Engineering 外国特別館員, IEEE
フェロー, ACM フェロー, American Association of Artificial Intelligence フェロー, C&C
賞, Joseph Engelberger 賞, FIT2004 船井業績賞, JARA 賞,人工知能学会業績賞, Marr
賞等各受賞.