Exemplar-DPMs による物体追跡に関する研究

TP15004 宇佐見 剛司　　指導教授：藤吉 弘亘

1.はじめに

Tracking-by-Detection アプローチによる物体追跡は，追跡対象カテゴリの検出器を用いて追跡する手法である．追跡対象と類似した物体の存在は追跡を困難にする要因の 1 つである．本研究では，類似した物体が存在する環境下においても，特定の対象を追跡可能な Tracking-by-Detection アプローチによる手法を提案する．提案手法は，Deformable Part Models(DPMs)[1] のパーツ検出器を Exemplar-SVM(E-SVM)[2] により学習されたパーツに特化したパーツ検出器に置換してスコア算出を行い，得られたスコアを用いて対象の追跡を行う．評価実験では，類似した対象物体に対して，誤追跡をしない高精度な追跡が可能であることが確認できた．

2.提案手法

本研究では，事前に学習した DPMs の検出器とオンラインで学習した E-SVM の検出器を用いた物体追跡手法を提案する．また，隠れの発生に対応するためスコア回帰によるオクルージョン対応を導入する．本章では，図 1 に示す提案手法の流れについての説明を行う．その後，E-SVM によるパーツ検出器の学習，スコア計算についてとスコア回帰によるオクルージョン対応の説明を行う．

2.1 提案手法の流れ

提案手法の流れを，以下に示す．

Step1. 前フレームにおける追跡対象のパーツを対象として E-SVM を用いてパーツフィルタを学習する．

Step2. 現在のフレームに対して HOG 特徴抽出を行い，特徴マップを得る．

Step3. 事前に学習してある DPMs のルートフィルタと Step1 にて学習した各パーツフィルタを Step2 で得た特徴マップにそれぞれ畳み込みスコアマップを算出する．算出したスコアマップの中から最も高いスコアを算出した座標を求める．

Step4. Step3 で得られた各パーツ位置におけるオクルージョン発生率を算出する．

Step5. Step3 で得た各スコアと Step4 で得たオクルージョン発生率を用いてスコアの回帰を行う．

Step6. Step5 で得た回帰後のスコアが閾値以上であれば，追跡成功として結果の出力を行い Step1 に戻る．閾値以下であれば，追跡対象が存在しないとして，追跡を打ち切る．

2.2 E-SVM によるパーツ検出器の学習とスコア計算

E-SVM によるパーツ検出器を学習する際，各パーツの位置情報が必要となる．そのため，フレーム $t-1$ における追跡対象のパーツの位置が不明の場合， DPMs を用いてパーツ検出を行い，フレーム $t-1$ におけるパーツ位置を求める． E-SVM の学習サンプルは，図 2 のようにパーツごとに作成する．

この際，ポジティブサンプルには，フレーム $t-1$ におけるパーツを使用し，ネガティブサンプルには，図 2 のフレーム $t-1$ の画像中の青い枠のように，ランダムに切り出したポジティブサンプルに被らない画像を使用する．ネガティブサンプルの切り出し時のサイズはポジティブサンプルのサイズと同じになるように切り出す．ネガティブサンプルを切り出す枚数は 100 枚とする．パーツ検出器はパーツごとに作成するため，1 つの追跡対象ごとに 6 つのパーツ検出器が学習される．Step3 のスコア算出では，あらかじめ学習されている DPMs のルート検出器と Step1 で学習した 6 つのパーツ検出器を使用し，現フレームにおける各スコアを式 (1)，式 (2) より求める．

$$score_root(p_0) = F_0' \cdot \phi(H, p_0) \tag{1}$$

$$score_part(p_i) = w_{E_i} \cdot \phi(H, p_i) - d_i \cdot \phi_d(dx_i, dy_i) \tag{2}$$

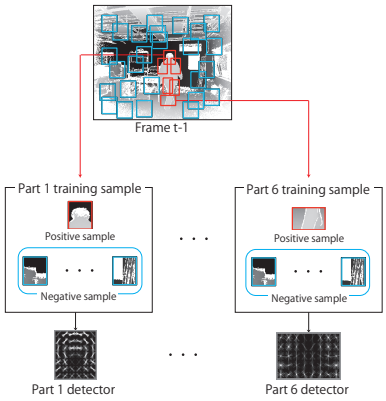

図 2：E-SVM によるパーツ検出器の学習

式 (1) の p_0 はルートフィルタの位置を示す．式 (2) の p_i はパーツフィルタの位置を表す．式 (1)，式 (2) を用いて DPMs の検出スコアを式 (3) により求める．

$$score(p_0) = score_root(p_0) + \sum_{i=1}^{6} \max_{p_i \in R(p_i')} (score_part(p_i)) \tag{3}$$

式 (3) の $R(p_i')$ は前フレームでのパーツ位置 p_i' の周辺領域の特徴マップを表す．これは前フレームのパーツ位置周辺でのみスコア算出を行うことにより計算コストの削減と精度の向上が期待できる．式 (3) において最も高いスコアを式 (4) にて求める．

$$score_dpm = \max_{p_0 \in R(p_0')} (score_root(p_0)) \tag{4}$$

式 (4) の $R(p_0')$ はフレーム $t-1$ におけるルートフィルタの位置 p_i' の周辺領域を表す．この時のルート位置 p_0 が追跡推定位置 p_0^* となる．また，式 (3) において追跡推定位置 p_0^* のときに選択されたパーツ位置が推定パーツ位置 p_i^* となる．

2.3 スコア回帰によるオクルージョン対応

スコア回帰によるオクルージョン対応は，DPMs において隠れの発生による性能の低下に対応するために提案された手法である．この手法は，DPMs の検出精度の低下の要因の 1 つであるオクルージョンの発生による検出対象の見えの大きな変動により低下してしまうスコアに対して，オクルージョンの影響を低減したスコアを回帰により求めることにより，オクルージョンが発生した物体であっても検出が可能となる手法である．

・オクルージョン発生率の計算

オクルージョンに対応した DPMs では，三次元空間上に配置した検出ウィンドウをラスタスキャンし，検出を行なっている．そのため，検出ウィンドウの三次元座標値と距離画像の画素値を比較することにより，注目画素においてオクルージョンが発生しているかの判定を行うことが可能である．この判定結果を用いることにより，パーツフィルタ位置 p_i^* におけるオクルージョンの発生率 $O(p_i^*)$ を式 (5) に示す式により求める．

$$O(p_i) = \frac{\sum_{u=x_i}^{x_i+xsize} \sum_{v=y_i}^{y_i+ysize} \alpha_O(u,v)}{xsize * ysize} \tag{5}$$

$\alpha_O(u,v)$ は，座標 (u,v) におけるオクルージョンの判定関数を表しオクルージョンが発生していた場合 1 を，発生していない場合 0 を返す．x_i,y_i はパーツの位置座標を表し，$xsize$,$ysize$ はフィルタのサイズを表す．

・DPMs のスコア回帰

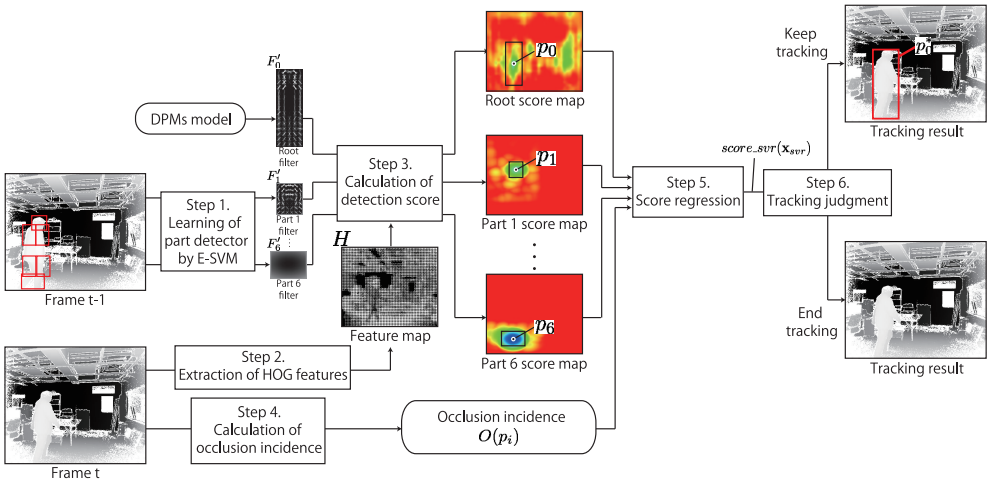

図 1：提案手法の流れ

説明変数 $\mathbf{x}_{svr}$ は，Step3 で得た DPMs の検出スコア，ルートスコア，パーツスコア，Step4 で得たオクルージョン発生率 $O(p_i^*)$ とパーツスコアの積とする．パーツ数 6 の場合の説明変数は式 (6) に示すように，14 次元となる．

$$\begin{aligned}\mathbf{x}_{svr} = [&score_dpm,\\ &score_root,\\ &score_part_1, ..., score_part_6\\ &O(p_1) * score_part_1, ...,\\ &O(p_6) * score_part_6]\end{aligned} \tag{6}$$

回帰の手法には Support Vector Regression(SVR) を用いる．SVR の識別関数を式 (7) に示す．

$$score_svr = \mathbf{w}_{svr}^T \cdot \mathbf{x}_{svr}; b_{svr} \tag{7}$$

$\mathbf{w}_{svr}^T$ は SVR の重みベクトル，b_{svr} は SVR のバイアスを表す．

3. 評価実験

提案手法の有効性を調査するために評価実験を行う．実験では提案手法を含めた 3 通りの手法の評価を行い結果の比較を行う．

3.1 実験概要

評価方法としては，フレーム t において矩形で囲ってある追跡対象をフレーム $t+1$ において追跡を行う．そして，フレーム $t+1$ において出力された矩形が正解矩形に 80%以上重なっていた場合，追跡成功と判断する．比較には追跡成功率を用いる．成功サンプル数と全体のサンプル数の除算で求められる．評価実験において，従来法として DPMs の検出器を用いた追跡，提案手法 1 として DPMs のルート検出器と E-SVM のパーツ検出器を用いた検出器である E-DPM を用いた追跡，そして，提案手法 2 として提案手法 1 にスコア回帰を導入した追跡の 3 つの手法で比較を行う．

3.2 データセット

データセットには，Kinect V2 により取得した距離画像を使用する．使用する距離画像には，隠れや類似物体への対応が可能かを調査するため，障害物による隠れや人物同士の重なりを一定割合で含んだ画像を使用する．DPM の学習用画像としてポジティブサンプル 903 枚，ネガティブサンプル 1,000 枚の計 1,903 枚を使用する．SVR の学習用画像としてポジティブサンプル 1,488 枚，ネガティブサンプル 800 枚の計 2,288 枚を使用する．評価用画像には，3,219 枚の画像を使用する．評価画像中にオクルージョンが発生していない追跡対象は 1,909 人，オクルージョンが発生している追跡対象は 1,207 人である．

3.3 実験結果

比較する 3 手法のオクルージョンの有無による追跡成功率を表 1 に示す．

表 1：提案手法における追跡成功率

比較手法	隠れなし	隠れあり
従来法 (DPMs)	78.3	43.0
提案手法 1 (E-DPMs)	94.3	64.9
提案手法 2 (E-DPMs+SVR)	96.4	83.2

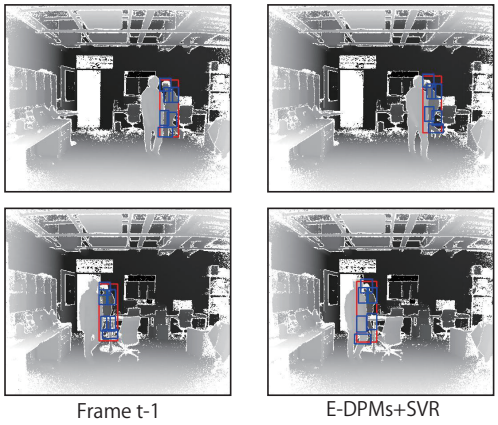

図 3：追跡成功例

表 1 より，オクルージョン発生時においてスコア回帰を行わない手法より約 13%，DPMs のみより 35%以上向上した．このことから，E-SVM により学習したパーツ識別器の導入は，類似した物体の交差が発生するシーンと隠れが発生するシーンに対して有効であると考えられる． 提案手法 2 における追跡の成功例を図 3 に示す．

4. おわりに

本稿では，既存の物体検出手法である DPMs と E-SVM の組み合わせによる高精度な物体追跡を提案した．E-SVM による対象パーツに特化した検出器の学習と DPMs による姿勢変動への対応により高精度の追跡性能を実現した．今後の課題として，リアルタイムで追跡を行うために，バイナリ近似の導入による高速化が挙げられる．

参考文献

[1] P.F.Felzenszwalb, *et al.*,“Object detection with discriminatively trained part based models”, PAMI, vol.32, no.9, pp.1627-1645, 2010.

[2] T. Malisiewicz, A. Gupta, and A. A. Efros,“Ensemble of Exemplar-svms for Object Detection and Beyond”, ICCV, pp.89-96, 2011.

研究業績

[1] T. Usami, H. Fukui, Y. Yamauchi, T. Yamashita, H. Fujiyoshi,“DPM Score Regressor for Detecting Occluded Humans from Depth Images”, FCV, 2016.

（他 学会発表 2 件　学会発表予定 1 件）