2R-07

Deformable Part Model を用いた変形物体領域検出に関する検討

冨田大輝† 森本正志‡

愛知工業大学†‡

1 はじめに

ウェアラブルデバイスや携帯端末の普及により，そのカメラで撮影された画像をクエリとする画像検索サービスが出現している．機械学習技術の発展に伴いカメラ画像からのクラス分類や画像検索は高性能化し，また画像内における対象物体の変形は，その変形を含む大量の学習画像によりある程度対応可能となってきた．しかし，のぼり旗や衣服などのように大きく変形する物体に対しては，その性能を確保するために検討すべき課題があると考えられる．

我々は変形物体例として国旗を用い，変形を許容するための検索手法に関する研究を行ってきた[1]．本稿では Deformable Part Model[2](以下，DPM)を用いて変形を伴う国旗をモデル化し，動画から画像検索に適したフレーム画像検出および物体領域検出を行う手法の検討を行う．

2 提案手法

2.1 従来研究における国旗領域検出

従来研究[1]における国旗領域検出処理の流れは以下のとおりである．まずカメラで撮影した国旗画像に対しエッジ検出を行い，閉輪郭を取り出す．面積が最も大きい閉輪郭を国旗領域とし，その外接長方形を求めて閉輪郭と接する 4 頂点座標を求め，傾きを正規化する．なお，このようにして得られた正規化国旗領域に対し色特徴と局所特徴を算出し，その照合処理により検索結果を算出している．

2.2 Deformable Part Model による国旗領域検出

前節の国旗領域検出手法では，大きな変形が発生すると国旗領域検出やその正規化が困難になり，結果として検索結果も悪化することが判明した．そこで，閉輪郭や面積などの物理的特徴だけではなく，国旗としての特徴を用いた領域検出手法が必要であると考えた．本研究では国旗領域検出に機械学習手法である DPM を使用する．DPM とは，物体を全体像とパーツの集合体で評価するモデルで，全体像やパーツの相対位置の妥当性でスコアを決定する手法である．本稿ではまず文献[2]における DPM をそのまま用い，変形物体に対するクラス分類および領域検出の有効性を検証することとする．

2.3 動画からの検索対象画像検出

これまでは単一の画像から国旗検索を行ってきたが，そもそも対象国旗が大きく歪んだ際に撮影された画像では，国旗領域検出もおぼつかないと考えられる．そこで本研究では，動画から画像検索に適したフレーム画像を検出し，その画像からさらに国旗領域を検出する方法を提案する．提案手法のイメージを図 1 に示す．ここでも DPM を用い，変形の少ない国旗画像および国旗領域で学習を行う．動画像の各フレーム画像に対して DPM でクラス分類を行い，算出スコアから検索対象となるフレーム画像検出および国旗領域検出を行う．

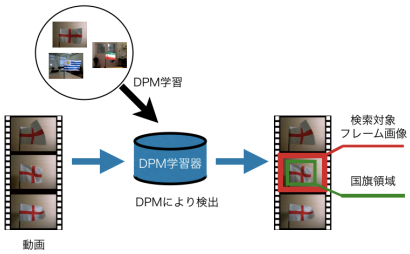

図 1:提案手法のイメージ

3 評価実験

3.1 学習データ

DPM の学習には正解画像であるポジティブサンプルと不正解画像であるネガティブサンプルに加え，対象物体の位置とクラスを示したアノテーションデータが必要である．本研究では，学習データ用として 15 種類の国旗と背景の異なる屋内撮影場所 10 カ所を用意した．モーションブラーの影響を抑えるために，動画の撮影にはハイスピードカメラを使用した．無作為な変形を行うため，動画撮影中は横からサーキュレータによって送風を行った．また，蛍光灯のフリッカー現象の影響を抑えるために LED ライトで国旗を照射した．1 つの動画から 10 枚ずつ画像を選出し，計 1500 枚の画像をポジティブサンプルとした．アノテーションデータも同数の 1500 個用意

A Study on region detection for non-rigid objects using Deformable Part Model

† Daiki Tomita, Aichi Institute of Technology

‡ Masashi Morimoto, Aichi Institute of Technology

Copyright ©2018 Information Processing Society of Japan.
All Rights Reserved.

した．ネガティブサンプルには画像認識コンペ VOC2011 の画像 11540 枚を使用した．

3.2 国旗領域検出性能評価

まず国旗領域検出性能を評価するために，交差検証を行った．前述のポジティブサンプル画像 1500 枚を国旗 3 カ国ずつでグループ分けする．1 グループ分の画像 300 枚をテスト用データとし，残り 1200 枚の画像で学習を行った結果生成された識別器を用いてテスト用データの評価を行う．テストグループを変えて 5 回検証を行い，正答率を求める．DPM は対象クラスとして判定された被写体箇所を内包するバウンディングボックス（以下，Bbox）の座標とそのスコアが出力されるので，正のスコアを記録した Bbox のみを対象に検出領域の正誤を判断し，正答率を国旗ごとに算出する．

国旗領域の正検出例を図 2 に，誤検出例を図 3 に示す．図 3 の誤検出は，国旗の中の線を国旗の国旗の輪郭線と間違えて判断したものと思われるが，この類の誤検出は多く見られた．また，性能評価結果を表 1 に示す．全体での平均正答率は 66.33%となった．グループ 4 の正答率が低く，これはグループ 4 に含まれる国旗（ブラジル）と似た特徴を持つ国旗が学習データ内に少なかった事が原因と考えられる．また，背景によっても正答率の差が見られ，図 2 のように背景に他の物体が写っていない画像では正答率が下がる傾向が見られた．これは背景に多くの物体が写っている学習画像を多く用いたためと考えられる．

図 2:正検出例

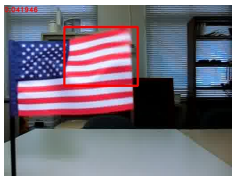

図 3:誤検出例

表 1:交差検証結果

テスト画像	正答率	
グループ 1	66.33%	66.33%
グループ 2	64.00%	
グループ 3	82.00%	
グループ 4	58.67%	
グループ 5	60.67%	

3.3 動画からのフレーム画像検出性能評価

次に，動画から変形の少ない国旗画像と適切な国旗領域を検出可能かどうか確認する性能評価を行った．学習用データは 3.1 で示したポジティブサンプル画像 1500 枚を使用する．テスト用画像は，ポジティブサンプルに使用した国旗 15 カ国と異なる 5 カ国を追加した計 20 カ国の国旗を，3.1 と異なる屋内 2 カ所および屋外 2 カ所の計 4 カ所で撮影した．ハイスピードカメラを用いているため，隣接フレーム画像がほぼ同じ内容となってしまう．そこで，予め撮影動画から 30 フレームごとに 1 枚の画像を取り出し，1 つの動画につき計 100 枚の画像を用いた（合計 8000 枚）．3.2 と同じく正のスコアを記録した画像および Bbox のみを対象に正誤を判断し，正答率を背景ごとに算出する．ここでは変形が少なく，正しく国旗領域が取れているものを正解とした．

評価結果を表 2 に示す．全体での正答率の平均は 65.39%となった．正答率が一番低かった背景 1 では検出画像は 2000 枚中 890 枚と多かったものの，そのうち正解と判断した画像の割合が少なかった．国旗別にみると，ポジティブサンプルに登録されている 15 か国の国旗は多くのフレーム画像が検出されたが，登録されていない 5 か国の国旗は画像が検出されないことが多かった．

表 2:評価結果

テストデータ	正答率	
背景 1	53.71%	65.39%
背景 2	68.30%	
背景 3	79.57%	
背景 4	63.46%	

4 まとめ

本稿では DPM を用いた国旗画像検出および国旗領域検出に関する検討を行った．評価実験により，交差検証では 66.33%の正答率を記録した．また，動画からのフレーム画像検出では 65.39%の正答率を記録した．いずれも学習データに入っていない国旗に対しては性能が劣化すると考えられるので，学習データに全国旗を登録した上で様々な背景などの条件に対して検出が可能であるかを調査する必要がある．また，検出として今回は正のスコアを記録した Bbox を対象としたが，その閾値に関する検討も必要である．

参考文献

[1] 冨田，森本，“非一様な形状変形を許容する国旗画像検索手法に関する検討，”第 14 回情報学ワークショップ 2016(WiNF2016),B-26,2016.

[2] P. Felzenszwalb et al.,“Object Detection with Discriminatingly Trained Part-Based Models”, IEEE Trans. on PAMI, Vol.32, No.9, pp.1627-1645, 2010.

Copyright ©2018 Information Processing Society of Japan.
All Rights Reserved.