1N-06

フォトグラメトリを用いた物体検出のための学習データ生成手法

小嶺愛紀菜* 赤嶺有平† 根路銘もえ子‡

琉球大学理工学研究科情報工学専攻* 琉球大学工学部† 沖縄国際大学経済学部‡

要旨

特定物体に対して重畳表示を用いたキャプション提示を行うためには物体検出を行う必要がある．深層学習による検出手法が主流になりつつあるが，特定物体検出の場合，既存の学習データを利用することができないため，大量の学習データを利用者自身が用意する必要がある．本稿では，学習データの準備に要するコストの低減を目的として，フォトグラメトリを用いて多数視点の画像を生成し，最小限の撮影画像から大量の学習データを生成する手法を提案し，その評価について述べる．

1 はじめに

近年，博物館等の展示施設においてモバイルデバイスを用いたディジタルガイダンスシステムが注目されている．特に，拡張現実技術を用いたディジタルガイダンスシステムでは，文字のみの説明だけではなく，動きのある映像や画像など視覚的な情報を来館者に提供することができる．

現在，本研究室において博物館と共同で拡張現実技術を用いた展示物のガイダンスシステムの開発を行っている．深層学習を用いることで比較的ロバストな重畳表示が実現可能である．しかし，博物館には多数の展示物がありそれらのインスタンス名を予測するために，それぞれの検出対象（重畳表示の対象）を様々な角度から撮影した大量の学習データを準備する必要がある．本稿では，学習データの準備に要するコストの低減を目的として，フォトグラメトリを用いて多数視点の画像を生成し，最小限の撮影画像から大量の学習データを生成する手法を提案する．さらに，多視点画像生成によって作成したデータセットを用いて学習を行い，その評価について述べる．

2 提案手法

まず，カメラの高さは変えずにそのまま横へ移動しながら検出対象を撮影した画像を複数枚用意する．次に，SfM(Structure from motion)技術を用いて，撮影した画像のカメラ位置姿勢を特定し，三次元構造を取得する．撮影した画像のカメラ位置を元に，少しずつ仮想カメラの位置を上下にずらして，違う位置・角度から撮影したような画像を生成する．このようにすることで，実際に撮影するのは同じ高さの画像のみでよく，上下の視点移動に関してはフォトグラメトリを活用して画像生成する．最小限の画像から大量の学習画像を自動生成でき，画像分類・検出モデルの学習コストを大幅に削減できる．

2 実験

3.1 多視点画像生成

実際に，検出対象を撮影した画像の一部を図 1 に示す．



図 1: 展示品を実際に撮影した画像

撮影した画像を元に，図 2 のような多視点画像を生成した．図 2 で生成した画像の元視点の画像も図 3 に示す．

A Training Data Generation Method Using Photogrammetry for Object Detection

*Akina Komine, Ryukyu University, Graduate School of Engineeringa and Science, Information Engineering

†Yuhei Akamine, University of the Ryukyus, Faculity of Engineering

‡Moeko Nerome, Okinawa International University, Department of Economics

Copyright ©2021 Information Processing Society of Japan.
All Rights Reserved.

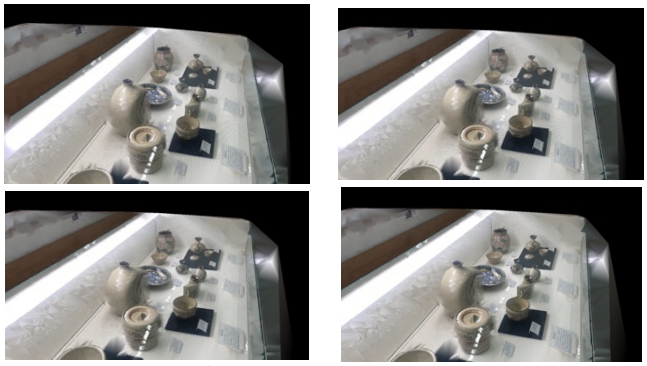


図 2: 多視点画像生成による画像



図 3: 元画像

4.2 分類

提案手法を用いた場合と用いない場合で 2 クラス分類を行った．提案手法を用いる場合では，多視点画像生成によって作成された画像を学習データとして使用した．多視点画像生成によって作成した学習データの一部を図 4 に示す．1 クラスあたり 4 枚の実写画像があり，多視点画像生成によってそれを 40 枚に増やす．一方，提案手法を用いない場合は，学習データに同じ実写の画像 4 枚を使用し，それらを回転・反転・拡大縮小・チャネルシフトなどを行い，提案手法を用いた場合と同じである 40 枚になるようにデータ拡張を行った．両者，同じ実写画像をテストデータとして分類結果の比較を行う．



図 4: 提案手法によって作成した学習データ

3 結果

提案手法を用いた場合と用いない場合それぞれのテストデータに対する分類結果を表 1，表 2 に示す．

		モデルの予測結果	
		壺	酒器
ラベル	壺	16	1
	酒器	5	10

表 1: 提案手法を用いた場合の分類結果

		モデルの予測結果	
		壺	酒器
ラベル	壺	17	0
	酒器	12	3

表 2: 提案手法を用いない場合の分類結果

4 考察

多視点画像生成によって作成した画像でモデルを学習させても実際のカメラで撮影した写真を十分に分類できることが分かった．比較的遠距離にある物体は多視点画像生成が上手くいかない場合があり，深度推定が上手く行われていない可能性がある．そのため，写真の撮り方には工夫が必要だと考えられる．また，展示品が置かれている棚はカメラで撮影すると光を反射したような写りになることが多いため反射光を除去することで SfM の結果がより良くなるのではないかと考えられる．

5 まとめと今後の課題

フォトグラメトリを用いたデータ生成手法を提案し，実験を行った．多視点画像生成をすることによって実データにない角度から見た物体の分類や検出の精度向上に期待ができる結果となった．今回は 2 クラス分類を行ったが，今後，写真の撮り方の工夫や反射光の除去など改善を行ったのちに，物体検出の学習データとして用い，一般的な Data Argumentation との比較検証を行う必要があると考える．

謝辞

本研究は JSPS 科研費 JP19K01142，JP19K01145 の助成によるものである．

Copyright ©2021 Information Processing Society of Japan.
All Rights Reserved.