

1. はじめに

キャッシュレス決済や物品管理、広告、ロボットの認識機能等の分野で QR コードや AR マーカに代表される 2 次元コードが利用されている。2 次元コードには、数百から数千バイトの情報を埋め込むことができ、シンボルと呼ばれる特殊なパターンによって視点が変化しても高精度な検出が可能である。さらに、2 次元コードの大きさを事前に定義すればカメラの位置・姿勢を推定することができる。しかし、2 次元コードは平面に貼ることを前提条件としており、曲面に貼られた 2 次元コードは歪みによる見えの変化を引き起こすため、認識精度が低下する問題を抱えている。

そこで、本研究では機械学習により変形した AR マーカの認識及び姿勢推定を行う手法を提案する。変形した AR マーカの画像を Single Shot MultiBox Detector(SSD)[1] により学習することで、歪みを持つマーカを正確に検出し、ID、座標、大きさを推定し、Augmented Autencoder(AAE)[2] による学習で、歪んだ AR マーカの姿勢の推定を行う。

2. SSD と AAE を用いた姿勢推定

図 1 は、提案手法による変形 AR マーカの認識及び姿勢推定の流れを示す。本研究では AR マーカの検出と認識に物体検出手法である Single Shot MultiBox Detector (SSD) を学習することにより AR マーカの画面座標位置、カテゴリの種類を求める。AR マーカの姿勢を Augmented Autencoder (AAE) を学習することによって推定を行う事ができる。提案手法では SSD と AAE により変形 AR マーカの大きさ、姿勢を推定できるように拡張する。

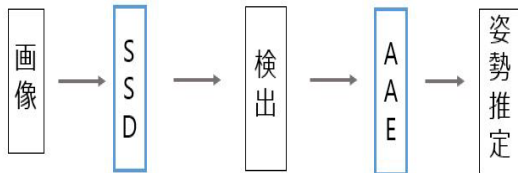


図 1：提案手法の流れ

3. SSD による検出

SSD のネットワークの説明する。始めは $300 \times 300 \times 3$ の入力画像から VGG-16 の畳み込み層の途中までを $38 \times 38 \times 512$ の特徴マップを抽出する。ここで抽出した特徴マップをネットワークの次の層の入力とし、VGG-16 の残りの畳み込み層と、 $3 \times 3 \times 1024$ 、 $1 \times 1 \times 1024$ の畳み込みで $19 \times 19 \times 1024$ の特徴マップを抽出する。同様に抽出した特徴マップを次の層に入力として特徴マップを抽出する。この方法を繰り返して合計 6 段階のサイズ違いの特徴マップを抽出する。そのため比較的小さい物体でも検出可能である。物体認識と位置決定の損失関数の和を小さくするように値を調整し学習する。次に SSD の損失関数の説明をする。

$$L(x, c, l, g) = \frac{1}{N} (L_{cls}(x, c) + \alpha L_{loc}(x, l, g)) \quad (1)$$

N は正解ラベルと位置がマッチしたバウンディングボックスの数、 L_{cls} は物体のクラス分けを行う condence loss を表し、 x は正解ラベルとマッチしたら 1 になる、 c はクロスエントロピー誤差である。 α は重み、 L_{loc} は物体の位置ずれを表す localization loss である。 l は予測したボックスの

中心の x, y , 幅, 高さを表し、 g は正解ボックスの中心の x, y , 幅, 高さを表す。これらの和が小さくなるように学習する。

4. AAE による姿勢推定

AAE は物体の三軸方向を推定するための Autencoder である。推定までの全体的な流れを図 (2) に示す。Autencoder により各姿勢情報を含むターゲット画像の潜在変数を学習時に蓄積しておき、検出時に獲得されたバウンディングボックスを切り取り、Autencoder に入力し得られた潜在変数と学習時に蓄積された潜在変数との、コサイン類似度を求めることによって最も近い値の潜在変数の姿勢情報を物体姿勢として推定する。

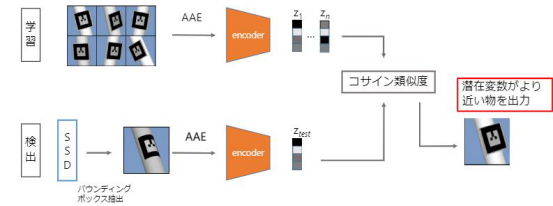


図 2：AAE による姿勢推定

ネットワークは、入力画像 128×128 の RGB 画像を 5×5 の畳み込み stride2 である、AAE の学習の流れを図 (3) に示す。学習したい元画像となる平面状の AR マーカの貼られた円柱の画像 (a) を正解画像として用意し、元画像と同じ姿勢の AR マーカを円柱に沿うように貼り付けた画像 (b) を入力として Autencoder にかけて出力 (c) は (a) と誤差を取り (a) を再現するように学習を行う。

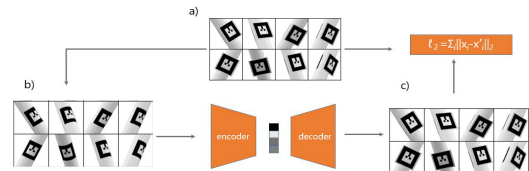


図 3：AAE の学習

5. 評価実験

今後行う予定

6. おわりに

本研究では、変形 AR マーカの認識及び姿勢推定を提案し、機械学習によって AR マーカの座標位置、姿勢を推定できる事を確認した。今後は、提案手法のよるリアルタイムでの三次元位置・姿勢の推定を研究を行う予定である。

参考文献

- [1] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg : "SSD: Single Shot MultiBox Detector", Proc. of ECCV, 2016.
- [2] Martin Sundermeyer *et al.* : "Implicit 3D Orientation Learning for 6D Object Detection from RGB Images", Proc. of ECCV, 2016.