

## 画像の学習に基づく姿勢推定手法の学習サンプル数と精度の関係性の検討

奥川 裕之<sup>†</sup> 原田 健吾<sup>††</sup> 玉木 徹<sup>†</sup> 天野 敏之<sup>†††</sup> 金田 和文<sup>†</sup>

<sup>†</sup>広島大学大学院工学研究科

<sup>††</sup>広島大学工学部

<sup>†††</sup>奈良先端科学技術大学院大学情報科学研究科

### 1. 背景

本研究では、ある物体の見え方が違う画像を学習することで、その物体の姿勢を推定する手法の推定精度を検討する。物体の見え方が違う画像とは、物体を図1のように様々な方向から撮影した画像である。その画像を学習サンプルとし、推定したい画像と学習サンプルを比較することで、物体の姿勢を推定する。

画像の学習による姿勢推定手法の精度は、学習サンプル数に依存している。よって、推定する際に重要となってくることは、どのぐらいの枚数を学習サンプルとするかである。学習サンプルが多いほど精度が良くなると考えられるが、様々な姿勢を推定するとなると、膨大な学習サンプルが必要である。そのため、なるべく学習サンプルは少なくしたい。しかし、どのくらいまで学習サンプルを減らして、精度良く姿勢推定できるかわかっていない。よって、学習サンプル数と推定精度の関係を明らかにし、学習サンプル数の基準を作ることを、この研究の目的とする。

### 2. 姿勢推定精度の検討

画像を学習して姿勢推定する手法として、EbC法[1]を用いる。学習サンプルには、COIL-20[2]を用いた。この学習サンプルは、物体の鉛直軸回転の1自由度である。また学習サンプルの間隔  $i$  は  $5[\text{deg}]$  刻みで、サンプル数は一物体に対して72枚である。そして物体数は20物体である。

#### 2-1 学習サンプル数変化

学習サンプル数が多いほど、精度が増すと予想できる。だが、実際どのように精度が変化するのは検討されていない。そこでサンプル数と推定誤差との関係を調査した。ここでは、学習サンプルの間隔を  $i = 5k$  ( $k=1,2,3,4,6,8,9,12,18,24$ ) $[\text{deg}]$  とする。つまり、サンプル数を72,36,24,18,12,9,8,6,4,3枚と減ることになる。推定誤差の評価にはRMSEを用いた。

結果を図2に示す。obj4では、予想通りサンプル数が多くなると推定精度が減っている。だがobj5のように予想

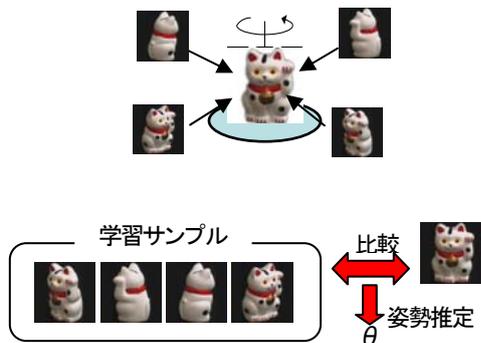


図1. 画像の学習に基づく姿勢推定

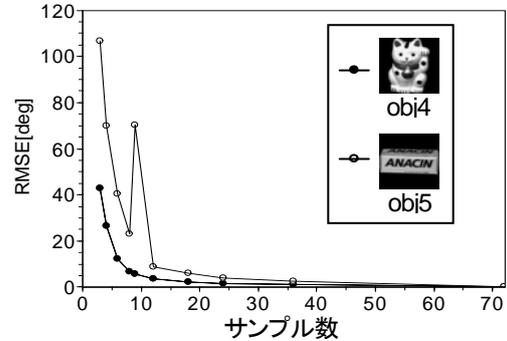


図2. 学習サンプル数に対する推定精度 (RMSE)

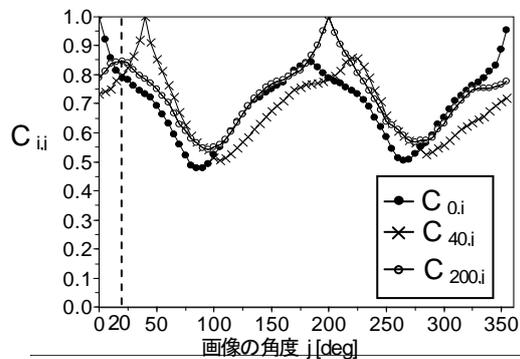


図3. obj5の画像の相関 ( $i=0, 40, 200$ )

と異なり、サンプル数が多い方が精度が悪くなる物体が存在した。

#### 2-2 画像間の相関の検討

学習サンプル数が多い方が精度が悪くなる原因は、学習した画像間の類似度が影響すると考えられる。そこで、obj5について、画像間の相関を計算した。画像  $\mathbf{x}_i$  と  $\mathbf{x}_j$  の相関  $C_{ij}$  は、1に近いほど相関が強くなり、0に近いほど相関が弱くなる。

検討の結果、精度が悪くなるサンプル数9(画像の間隔は40度)では、相関が高い画像同士(図3では200度の画像に対して20度)が学習画像に含まれていない。そのため、20度の画像に最も似た画像は200度の画像となり、推定誤差が増えている(図3の  $j=20$  度付近を参照)。それが、学習サンプル数が多い方が精度が悪くなる原因となったと考えられる。よって、相関が高い同士の画像は学習すべきであると考えられる。

#### 参考文献

- [1] 天野敏之, 玉木徹: 「EbC 画像対を用いた三次元物体の高速線形姿勢推定」, 電子情報通信学会論文誌 D, Vol. J90-D, No.8, pp.2060-2069 (2007).
- [2] S.A.Nene, S.K.Nayar, H.Murase: "Columbia Object Image Library (COIL-20)", Technical report CUCS-005-96, Columbia University (1996).