

任意視点からの指文字の認識

唐 志峰[†] 諸角 建^{††}

† 拓殖大学大学院 工学研究科 †† 拓殖大学工学部

1. はじめに

人間の視覚による認識には視点の独立性があることはよく知られている。しかし、視覚に基づく指文字認識アルゴリズムの大部分はカメラの視点を正面視に限定している。視点が変化すると、画像の特徴が変化する。視点の影響を受けずに正確に特徴を抽出することは困難な作業である。

2. 研究方法

あらゆる視点からの認識を達成するために、本研究では3D スキャナを使用して、41 文字から成る指文字の3次元のテンプレートを構築し、様々な方向から見た画像を機械学習させ、指文字認識における3次元識別のデータベースを作成する方法を考えた。さらに、3次元識別のデータベースの画像について機械学習をして認識を行った。認識の結果を分析して、その精度を上げる方法を提案した。

3. データベースの確立

3.1 モデルのスキャン

本研究で使用しているスキャナはEinScan-SEである。図1に示すように、このスキャナのスキャンサイズは限られているため、人の手をスキャンすることはできない。そこで、木製の手モデルを使用した。3D スキャナで数字1, 2, 3, 4, 5のジェスチャの手モデルをスキャンした。

3.2 モデルの処理

数字1~5のジェスチャの3Dモデルを読み取る^[1]。モデルを、垂直軸を回転軸として50秒で360°回転させる。回転させながら、50ミリ秒ごとに自動スクリーンショットプログラムを使用して、ジェスチャ画像を1000枚取得する。さらに、垂直仰角が30°、45°、60°の3Dモデルを取得する。

4. 学習画像

TensorFlow^[2]を参考にネットワークで学習を行う。それから学習結果をテストした。図2はテストしたジェスチャ画像の認識確率のグラフである。学習の回数が3000回より少ない場合には、認識確率が低い。3000回を越えても認識確率は基本的に変化しない。認識確率が90%以上は成功であるとして、テストを行った。学習回数90000回のモデルを用いて、100枚のジェスチャ画像について認識を行った結果の成功率を図3に示す。

5. モデルの利用

学習結果に基づいてスマートフォン用の2つのアプリケーションを作った。1つは、自動的な認識アプリケーションである。もう1つは、手動で写真を撮影してから認識をする形式のものである。

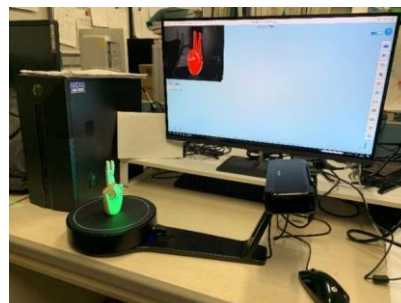


図1 スキャン風景

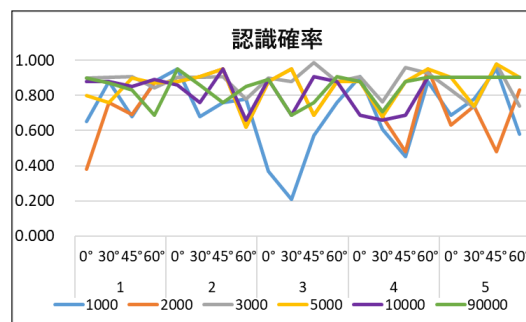


図2 認識確率

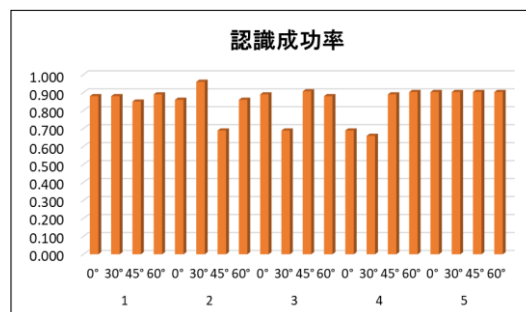


図3 認識成功率

6. まとめ

視点が異なるジェスチャ画像を取得した。自動スクリーンショットプログラムで各視点からジェスチャを自動的に取得した。TensorFlowを参考にして学習を行った。学習回数がおおよそ3000回以上では、認識の精度が変わらないことが判明した。最後に、検証済みの学習結果に基づいて2つのアプリケーションを作って認識テストを行った。手動で認識するアプリケーションの方が良好な認識率を得ることができた。

参考文献

- [1] Java3D应用实例-加载Obj 三维模型 (Java3D应用例-入力Obj 三次元モデル), <https://www.cnblogs.com/dennisit/archive/2013/05/07/3065126.htm>
- [2] TensorFlow is an end-to-end open source platform for machine learning, <https://www.tensorflow.org/overview?hl=en>