

Preferred Networks インターン選考 2019

コーディング課題レポート

Chainer (コンピュータビジョン) 分野

音賀優颯

法政大学院/理工学研究科/応用情報工学専攻 知的情報処理研究室 (彌富ゼミ)

動作環境・データセット

課題 1 は手元の Mac Book (OS X 10.14.4) で, 課題 2 は GPU サーバ (Ubuntu 16.04.4 LTS) で動作を確認した. データセットは 11k Hands¹ データセットの Hand images 先頭 200 データを用いた. 詳細な実行方法については README.md に記載した.

課題 1 手のポーズを推定するアノテーションツールを作成

opencv を用いてアノテーションツールを作成した. ツールはすべての指で共通する指の付け根の key point とそれぞれの指で 3 箇所ずつの key point の設定を行い画像ファイルと同一名の json ファイルを保存する. 以下のコマンドでツールを起動し Fig. 1 が表示される.

```
$ python annotate.py
```

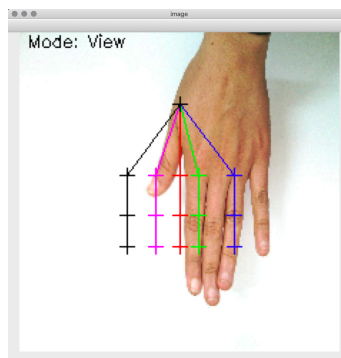


Fig 1: スタート画面

ツールは VIEW, SEQUENTIAL, EDIT, ROOT の 4 つのモードがあり VIEW モードのときに n を押下すると次の画像を表示, s を押下すると座標を保存する. その他のモードで画像をクリックすると key point を設定できるようになっており, SEQUENTIAL モードではすべての key point を連続して設定, EDIT モードでは 0-4 に応じて親指から小指の付け根以外を設定, ROOT モードでは指の付け根を設定することができる. それぞれのモードは a, 0-4, r を押下すると切り替わり, 該当する key point を設定し終わるか ESC を押下すると VIEW モードに戻る. 各モードの選択時は Fig. 2 のようになる.

¹<https://sites.google.com/view/11khands>

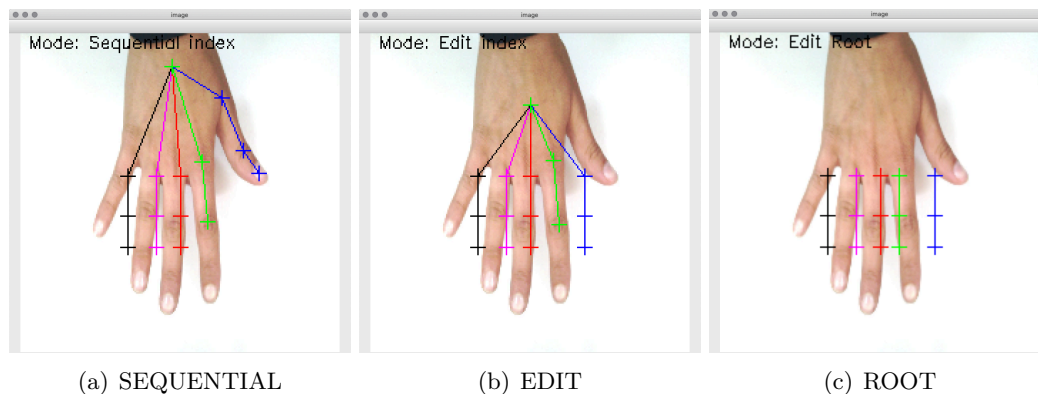


Fig 2: 各モード選択時

あると便利な機能

アノテーション作業を中断，再開できるように簡単なログ機能を追加した．座標を保存した画像は次回起動時には読み込まれなくなるようになっており，アノテーションを再開するにはログファイルから該当するファイル名を削除すればよい．

課題 2 CNN を用いて手のポーズを予測

実行方法は README.md の通りであり，

```
$ python predict.py -t [timestamp]
```

を実行後に

```
$ python annotate.py -p
```

を実行すると train データの予測結果がデフォルトの座標としてアノテーションツールが起動する．実行結果を Fig. 3 に示す．

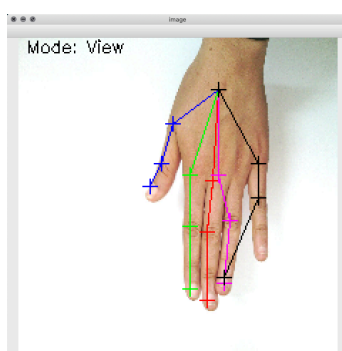


Fig 3: 予測結果の読み込み

課題で提示された手順に加えて適用して更に精度を上げるための手法

- key point だけでなくそれぞれの key point が結ぶエッジも合わせて予測するネットワークを追加する
予測されたエッジ端点や，エッジ同士の交点と対応する key point との平均を予測結果とすることでよりノイズに頑健でロバストな学習結果を得られると考えられる．