

SURF 特徴量を用いた指文字の型認識

1130342 清水 直人 【岡田研究室】

1 はじめに

指文字は手の形で文字を表す視覚言語の一種である。過去の研究でヒストグラムマッチング [1] などによる手法がある。本研究では新しい認識手法として画像の中身をより詳しく得ることができると言われている SURF 特徴量を利用した指文字の型認識を試みる。

2 SURF 特徴量

SURF 特徴量とは画像の回転やスケール変化に不変な特徴量である [2]。SURF 特徴量として算出されたキーポイントと呼ぶ特徴点は 128 次元ベクトル、X 座標、Y 座標、Hessian 値 (以下 H 値と表す) などからなる。座標は (X,Y) としたとき、左上を (0,0) としたときのキーポイントの位置を表す。H 値は輝度勾配を大まかに表す。

3 研究内容

3.1 キーポイントのマッチング

入力画像とテンプレート画像の 2 枚の画像に対して SURF 特徴量を算出する。入力画像の 1 つのキーポイントとテンプレート画像の全てのキーポイント間でキーポイントの 128 次元ベクトルのユークリッド距離を算出する。以下に 128 次元ベクトルのユークリッド距離の算出式 (1) を示す。 V_1, V_2 が入力画像、テンプレート画像の 128 次元ベクトルである。

$$M = \sqrt{\sum_{i=1}^{128} (V_{1i} - V_{2i})^2} \quad (1)$$

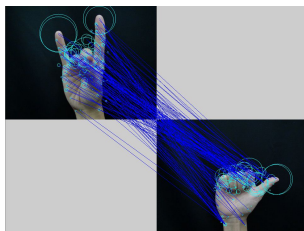


図 1 マッチング結果

3.2 認識手法

認識にはキーポイントの X 座標、Y 座標、H 値の 3 次元ユークリッド距離を利用する。入力画像とテンプレート画像のキーポイントはマッチングしている。テンプレート画像はグー型、キツネ型、パー型をそれぞれ 2 枚ずつ、計 6 枚用意する。入力画像 1 枚に対しテンプレート画像 6 枚との間に 3 次元ユークリッド距離をそれぞれ求める。型ごとに 3 次元ユークリッド距離の平均値を取る。平均値が最小の型を入力画像の型として認識する。以下に 3 次元ユークリッド距離の算出式を示す。 X', Y', H' はそれぞれの最大値で正規化してある。 N は入力画像のキーポ

イントの個数である。

$$D = \sqrt{\sum_{i=1}^N \{(X'_{1i} - X'_{2i})^2 + (Y'_{1i} - Y'_{2i})^2 + (H'_{1i} - H'_{2i})^2\}} \quad (2)$$

4 実験内容

今回の実験では下の表のテンプレートと入力画像を用意した。入力画像の型に対して正しく認識された回数をカウントし、認識率を算出する。座標 (X,Y) のみを用いた 2 次元ユークリッド距離での認識も行う。

表 1 使用画像

	グー型	キツネ型	パー型	人数	合計枚数
テンプレート	あ、さ	い、き	け、て	1	6
入力画像	あ、え、さ	い、き、ち、つ、 や、め、ろ	け、て、ほ	5	65

5 実験結果と考察

下の表は各式を利用したときの型別の認識率である。2 次元ユークリッド距離を使用するとキツネ型の認識率が極端に悪い。H 値を加えた 3 次元ユークリッド距離を使用するとキツネ型の認識率が上がり、全体の認識率も上がった。またグー型の認識率は下がった。グー型の誤認識の多くは「え」の指文字であり、「え」のパー型に近い形状が原因と思われる。キツネ型の誤認識は「め」と「ろ」が多い。これは他の指文字が正面を向いているのに対して横を向いていることが原因だと思われる。パー型に関しては両方において 100% の認識率を示した。

表 2 認識率 (%)

	グー型	キツネ型	パー型	全体
2 次元ユークリッド距離	73.3%	51.4%	100%	67.7%
3 次元ユークリッド距離	53.3%	82.9%	100%	80.0%

6 まとめ

本研究では、入力画像 1 枚とテンプレート画像 6 枚の間で 3 次元ユークリッド距離を求めることで型の認識を行った。一部の指文字を除いて高い認識率が見られた。しかし、「え」、「め」、「ろ」の指文字での認識率は悪かった。今後の課題は横を向いている指文字への対応と指文字認識へ導入である。

参考文献

- [1] 富田努武, "ヒストグラムマッチングを用いた指文字認識", 平成 21 年度 学士學位論文, 2010.
- [2] 中部大学 藤吉研究室, "SURF: Speeded Up Robust Features", 2009.