

# ジェスチャ識別のためのガボールフィルタを用いた特徴点自動抽出

## Automatic feature point extraction using Gabor filter for gesture recognition

高橋 佑輔<sup>\*1</sup>  
Yusuke Takahashi

井上 聡<sup>\*1\*2</sup>  
Satoru Inoue

<sup>\*1</sup> 埼玉工業大学大学院 工学研究科  
Graduate School of Engineering, Saitama Institute of Technology

<sup>\*2</sup> 埼玉工業大学  
Saitama Institute of Technology

The method using the Elastic Bunch Graph Matching and SHIFT features are well known gesture recognition. But several weakness has been reported in these two methods. Elastic Bunch Graph Matching is not good at real-time recognition and method using SHIFT feature is not guaranteed to pick up the fingertips as feature points. In this paper, we proposed the automated system extracting the feature points on fingertips for real-time gesture recognition.

### 1. はじめに

画像処理において、対象とする精度の高い物体の抽出や追跡、リアルタイム性を高める事は重要な課題であり、虹彩や耳形などの生体認証やジェスチャ識別など多くの分野で期待されている。手話をコンピュータに識別させるジェスチャ識別の研究もそのひとつであり、健常者がリアルタイムに手話を学習することが容易となる。手話を識別するには入力されたデータから特徴を取得し、得られたデータに基づき識別する必要がある。ジェスチャには上半身全体の動作を用いて表現をする手法や、指で文字を表す指文字がある。

るため、フレーム間差分を用いる。そして、得られた HVS 表現のデータに対して閾値処理を適用する事により肌色のみに着目、Canny エッジ抽出をする。エッジ抽出をすることで外光に対する影響を排除した、指を含めた手の形状情報のみを取得できる。抽出したデータに対して 8 方向のガボールフィルタを畳み込む事で、同ステップで手の向きを識別、それぞれの波形の方向に対して顕著性の高いデータを得ることが可能になる。それらのデータに対して学習で得られた重み係数による重み付けを行い、2 次元ガウス関数による探査で指先から特徴点抽出を試みる。

### 2. 手指領域抽出と抽出されたデータの細分化

#### 2.1 フレーム間差分・肌色抽出

手話動画の各々のフレーム間データを比較すると、背景や体の他のパーツも含め最も手指に差分が現れる。本研究では、着目しているフレームと、それ以前の 5 フレームの平均画素値による差分で動きが多く見られる領域のみを抽出する。しかし、フレーム間差分によるデータには手指だけでなく、次に差分の大きくみられる腕部の服飾などのデータも含まれており、それらのデータは有用ではないため、肌色抽出をすることで手指のみの領域が抽出可能となる。

#### 2.2 手の向き検出・データの細分化

手の向きを知ることは、指の方向に関する情報が絞り込まれ精度向上やシステムの高速化につながる。本研究では、22.5 度、45 度、67.5 度、90 度、112.5 度、157.5 度、180 度の 8 方向のガボールフィルタを畳み込み積分する事で手の向きを検出し、同時に畳み込み積分で得られたデータを 8 方向の顕著性の高いデータ群として扱う。

ガボールフィルタは以下の式で表される。

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (1)$$

$$x' = x \cos \theta + y \sin \theta \quad (2)$$

$$y' = -x \sin \theta + y \cos \theta \quad (3)$$

ここで、 $\lambda$  は波長の余弦成分、 $\theta$  は角度、 $\psi$  は位相オフセット、 $\sigma$  は分散、 $\gamma$  は縦横比を表す。

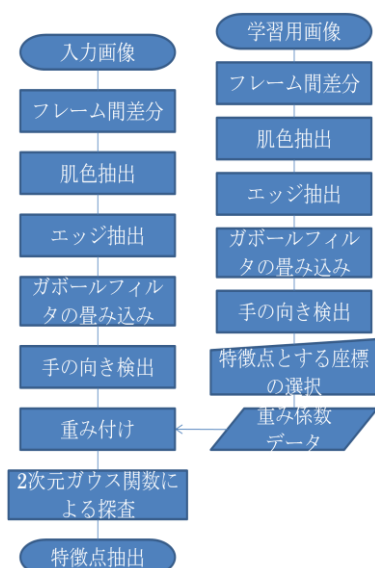


図1 本研究で提案するシステムのフローチャート

本研究では、識別システムを構築するには欠かす事の出来ない、特徴点自動抽出システムを提案する。しかし、SHIFT 特徴量などとは違い、指先からのみ特徴点自動抽出する事を目的とする。システムの汎用性を考慮し、ウェブカメラにて得られた動画データから特徴データを取得する。すべての画素値を特徴データとすると計算量が増えるために、必要のないと思われるデータを省く処理が必要である。リアルタイム処理を指向す

連絡先: 〒369-0293 埼玉県深谷市普濟寺 1690 埼玉工業大学大学院 m0004ucs@sit.ac.jp

畳み込む対象がエッジ抽出された画像であり、各々の角度で輪郭形パターンを把握する事を目的としている為、 $\sigma=1.5$ ,  $\phi=3.0$ ,  $\lambda=1.0$ ,  $\gamma=1.0$  とすることでフィルタのサイズを小さく設定する事が可能であり、計算量を減らすことができる。

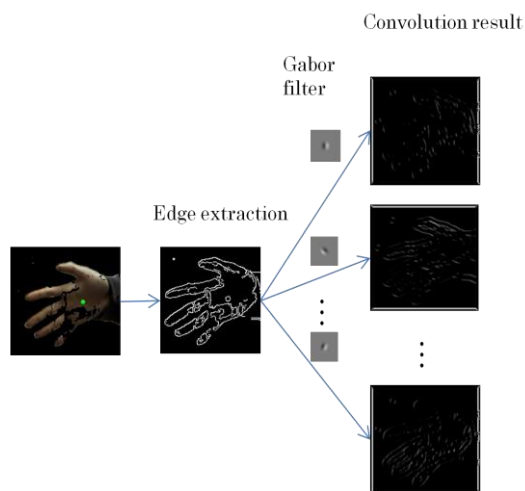


図2 ガボールフィルタによる畳み込み

### 3. 重み付けと2次元ガウス関数によるラスタスキャン

各々の畳みこまれたデータに対して重み係数を掛け、各々のデータを1つに複合する。複合したデータに対し、2次元ガウス関数を特徴点の探索フィルタとしてラスタスキャンを行う事により、特徴点抽出を試みる。

Convolution result

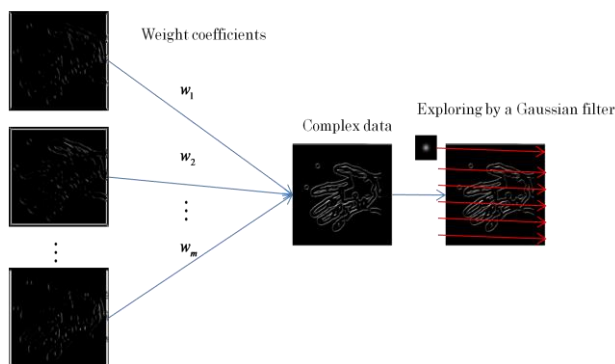


図3 畳み込んだデータ群に対する各々の重み付け、2次元ガウス関数による探索

#### 3.1 重み係数

重み係数は2次元ガウス関数 $G(X)$ で探索した際、指先と思われる元画像座標 $X=(x,y)$ に対し閾値 $\theta$ を超えるよう学習される。よって重み係数 $w_m$ は各々の細分化されたデータ $I(X)_m$ に対して教師データ $h^*(x)$ が満足できる値である。本研究では単層ネットワークで誤差逆伝搬法を用いた学習によって重み係数を作成した。フィルタの横幅と縦幅をそれぞれ $R_H pixels$ と $R_V pixels$ とする。

$$C_m(X) = \sum_{q=-R_V/2}^{R_V/2} \sum_{p=-R_H/2}^{R_H/2} I_m(x+p, y+q)G(p, q) \quad (4)$$

$$h^*(X) = \text{sign}\left(D \sum_{m=0}^M w_m C_m(X) - \theta\right) \quad (5)$$

$D$ は正規化係数であり、入力画像のサイズで識別結果が著しく変化しないよう設定された値である。

#### 3.2 2次元ガウス関数

2次元ガウス関数はガウス関数を利用して作られた釣鐘型の関数であり、平滑化フィルタとして用いられる事が多い。対象画素に近い画素に重みを、対象画素から離れた画素には小さい重みを付ける事ができる。対象画素データを情報の稠密なデータとして扱うのではなく、対象画素周辺の広いデータを扱う事が可能な為、探索フィルタとして採用した。

### 4. 実験結果

本研究ではジェスチャ識別のための指先を特徴点と置くための特徴点自動抽出システムについて研究を行った。

実験には、CPU: Intel(R) Core i5 動作周波数 3.33GHz メモリ 4096MB を使用した。重み係数は学習用画像 10 枚から学習データ 100 個を取得し作成した。図4が本研究の結果である。今回の実験では高いリアルタイム性を確認する事は出来なかったが、指先から特徴点を得られる結果を得られた。結果より次の2点の問題点が考えられる。フレーム差分を使用した際に現れる雑音データに対して脆弱である点や、本研究で使用した探索フィルタは不変である為、手の大きさに対するロバスト性が低く、探索フィルタのスケールと一致しない可能性のある点である。



図4 本研究の実験例

### 5. おわりに

本研究の結果より、手の向き検出アルゴリズムや雑音データの削除、手のサイズに合わせた可変な探索フィルタの設計を今後の課題とし、重みづけデータなどは AdaBoost や SVM などを導入することでより精度の高い特徴点自動抽出ができるのではないかと考えられる。また動画処理のリアルタイム性に関しては高速フーリエ変換を用いた畳み込みにて対処できると考えている。

### 参考文献

- [Fang 07] Yikai Fang, Kongqiao Wang, Jian Cheng and Hanqing Lu: "A REAL-TIME HAND GESTURE RECOGNITION METHOD", IEEE Comput, 2007.
- [Triesch 02] jochen Triesch, Christoph von der Malsburg: "Classification of hand postures against complex backgrounds using elastic graph matching", Image and Vision Computing 20(2002)937-943, 11 July 2002.
- [今井 08] 今井 彰博, 島田 伸敬, 白井 良明: "複雑背景下におけるモデルの照合誤りを考慮した手指形状推定", 電子情報通信学会誌 Vol. J91-D No.3 pp784-792, 2008