

## Eigen Image Decomposition に基づく複合的画像処理システム

小川達也<sup>†</sup> 田村安孝<sup>†</sup> 柳田裕隆<sup>†</sup>

山形大学大学院理工科学研究科情報科学専攻

## 1 はじめに

画像センサと処理システムの低コスト化が進んだことにより、画像処理は様々な分野で応用されるようになった。画像処理においては、画像センサの出力のみが処理されるのではない。対象の構造についてのモデルや知識が、明示的あるいは暗黙的に組み込まれている。また、近年 Kinect など、奥行きや音響信号など多様なデータを画像に関連して取得するセンサが利用可能になった。これらのセンサが出力する分解能・精度が異なるデータや、対象に関する先験的な知識を統合して扱う方法の確立が重要となってきた。

本研究は、濃淡画像のデータに加え、手描きのスケッチで得られた輪郭線など、多様なデータを統合する手法の確立を目標とする。実験を行う対象として、人物の頭部のような1つのカテゴリに所属するものに限定し、複数のセンサのデータを単一のベクトルに結合し、Eigen Image Decomposition<sup>[1]</sup>を用いた手法により、劣化あるいは欠落した濃淡情報、輪郭の形状の復元などを行う手法を開発し、性能を評価する。

## 2. Eigen Image Decomposition

Eigenimage (固有画像) の作成と、それを用いるデータの復元は以下のように行われる。

## 2.1 トレーニング行列の作成

濃淡画像を画素数  $N$  の次元を持つベクトルとして表す。画像から輪郭を  $M$  個の点を連結した閉曲線として抽出し座標値を要素とする  $M$  次元ベクトル  $\mathbf{x}$ ,  $\mathbf{y}$  として表す。 $K$  組のトレーニングデータは、 $L$  行  $K$  列 ( $L=N+2M$ ) のトレーニング行列  $T$  として、(1) 式のように定義される。

$$T = \begin{bmatrix} \mathbf{g}_1 - \bar{\mathbf{g}} & \mathbf{g}_2 - \bar{\mathbf{g}} & \cdots & \mathbf{g}_K - \bar{\mathbf{g}} \\ \mathbf{x}_1 - \bar{\mathbf{x}} & \mathbf{x}_2 - \bar{\mathbf{x}} & \cdots & \mathbf{x}_K - \bar{\mathbf{x}} \\ \mathbf{y}_1 - \bar{\mathbf{y}} & \mathbf{y}_2 - \bar{\mathbf{y}} & \cdots & \mathbf{y}_K - \bar{\mathbf{y}} \end{bmatrix} \\ = \begin{bmatrix} \mathbf{f}_1 - \bar{\mathbf{f}} & \mathbf{f}_2 - \bar{\mathbf{f}} & \cdots & \mathbf{f}_K - \bar{\mathbf{f}} \end{bmatrix} \quad (1)$$

ここで、 $\bar{\mathbf{g}}, \bar{\mathbf{x}}, \bar{\mathbf{y}}$  は画像のグレー値と輪郭座標のトレーニングデータに関する平均値である。

## 2.2 固有画像の計算

トレーニング行列  $T$  から、以下の計算により共分散行列  $C$  を求める。

$$C = \frac{1}{K} TT^* = \frac{1}{K} \sum_{k=1}^K (\mathbf{f}_k - \bar{\mathbf{f}})(\mathbf{f}_k - \bar{\mathbf{f}})^* \quad (2)$$

ここで、“\*” は共役転置を表す。行列  $C$  の固有ベクトルが固有画像である。 $C$  のサイズは  $L \times L$  となり計算量が膨大になるため、 $T$  を特異値分解した式、

$$T = U\Lambda^{1/2}V^* \quad (3)$$

を用いる。式(3)の  $U, V$  を用いて、

$$C = U\Lambda^{1/2}V^*V\Lambda^{1/2}U^* = U\Lambda U^* \quad (4)$$

となる。 $U$  は  $L$  行  $L$  列の直交行列 ( $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_L]$ )、 $\Lambda$  は  $r$  個 ( $r \leq \min(L, K)$ ) の非零の対角成分を持つ行列である。したがって、 $r$  個の固有ベクトル  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$  が求められる。

## 2.3 固有画像によるデータの復元

雑音で劣化、あるいは輪郭線データが欠落した画像データ  $\mathbf{f}$  から、以下の式を用いて復元画像  $\hat{\mathbf{f}}$  を求める<sup>[2]</sup>。

$$\hat{\mathbf{f}} = \sum_{n=1}^r w_n \mathbf{u}_n + \bar{\mathbf{f}} \\ w_n = \mathbf{u}_n^* (\mathbf{f} - \bar{\mathbf{f}}) \quad (5)$$

## 3. 実験環境

## 3.1 データ取得システム

Kinect を用いてグレイスケール情報を取得するシステムを構築した。ソフトウェアの開発環境は以下のとおりである。

- Visual C++ (VisualStudio2010)
- Kinect for Windows SDK
- Open CV ライブラリ

Multisensor image processing system based on Eigen Image Decomposition

Ogawa Tatsuya<sup>†</sup>, Yasutaka Tamura<sup>†</sup> and Hirota Yanagida<sup>†</sup>

<sup>†</sup>Department of Informatics, Graduate School of Science and Engineering, Yamagata University

Kinect for Windows SDK の仕様のため取得画像の寸法は 480×640 に固定され、撮像距離も 80cm よりは短くできない。また、輪郭の画像はトレーニング画像から手描きで作成する。

### 3.2 画像復元のソフトウェア

収集した画像データの輪郭画像の作成及び輪郭の  $x-y$  座標のデータ抽出には OpenCV を用いる [3]。トレーニング画像作成などの処理、固有画像の計算、画像の復元などの処理には MATLAB (Image Processing Tool Box) を用いた。

## 4 実験結果

### 4.1 固有値画像の計算

人物の正面からのグレイスケール画像を取得し、顔の部分トリミングして 147×179 画素のトレーニングデータを作成する。輪郭データは、顔全体、両目、鼻、口の 5 つの輪郭で構成している。今回の実験ではトレーニング画像の個数は 10 とした。トレーニングデータの画像の例を図 1 に示す。



図 1 トレーニング画像と手描きの輪郭

手描きデータから openCV を用いて取得した輪郭座標の個数は画像ごとに変化するのでスプライン補間によってデータ数を揃える。

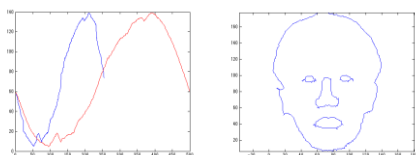


図 2 スプライン補間の結果

図 2 で目や口はデータ数を 100 に外側の輪郭はデータ数を 500 にする。2 章の Eigen Image Decomposition を用いて固有画像を求める。図 3 はグレイ値の平均と第 1 固有画像になる。



図 3 平均画像と第 1 固有画像

### 4.2 輪郭情報の復元

輪郭情報を欠落したデータから輪郭の復元を試みた。データの輪郭部を平均データに置き換えたものを  $f$  としてデータを復元した。

図 4 に示すように、トレーニングに使用した画像からは、元のデータとほぼ一致する輪郭が

復元される。一方、トレーニングに使用していない画像からの復元では精度が低下する (図 5)。

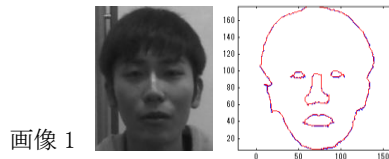


図 4 トレーニング画像から輪郭を復元

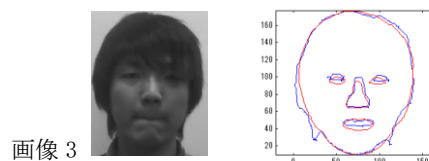
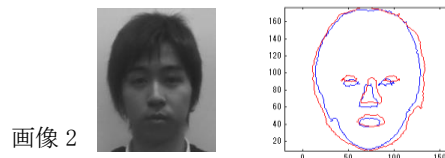


図 5 トレーニングに使用していない画像からの輪郭復元

3 つの画像について測定誤差は表 1 のようになった。

表 1 復元した輪郭の測定誤差

	画像 1	画像 2	画像 3
ずれの平均	-0.161	0.869	-0.109
分散	0.118	21.358	29.571
標準偏差	0.344	4.622	5.438

## 5 おわりに

濃淡画像に輪郭情報を追加して拡張した Eigenimage を用いて、輪郭復元の可能性を確認した。輪郭情報が欠落した濃淡画像を入力することで、5 つのパーツを持つ顔輪郭の復元が可能であった。トレーニング集合に含まれない画像からの復元で精度が大きく低下した。これは、トレーニングに使用する画像数が少なく、撮影条件のばらつきや前処理も適切でなかったためと考えられる。

トレーニングデータ取得と前処理のシステムの整備が課題となる。また、顔輪郭を構成する部分輪郭の詳細化や、顔の立体的構造の知識の組み込みなども検討したい。

## 6 参考文献

- [1] Xiaou Tang and Xiaogang Wang: Face Sketch Synthesis and Recognition, (2003)
- [2] Subhasis Chaudhuri and Joshi Manjunath: Motion-free super-resolution, pp148-172, Springer (2005)
- [3] 中川信行, OpenCV プログラミングブック, pp103-104 (2012)