

複数カメラを用いた手書き文字認識システム

A Handwriting Recognition System Using Multi Cameras

保呂 毅 稲葉 雅幸*

Summary. We constructed a handwriting recognition system using multi cameras. This system reconstructs the human body from the multi camera images using volume intersection. It finds a human hand from the human body voxel data, and estimates a stroke of the handwriting, and recognizes the input character using DP matching in real time. This system doesn't depend on the user's position and the direction. The recognition rate was 96.6%, when it recognized 36 characters of numbers and alphabets.

1 はじめに

人と機械を繋ぐインタフェースとして、キーボードやマウスといった人にとって特殊な操作が必要なデバイスではなく、音声やジェスチャなどを用い、より人間の自然なコミュニケーションに近く、より直感的な操作を可能にするための研究が数多く報告されている。

ジェスチャ認識をインターフェースとして利用する研究例として、テレビのチャンネルを変えるといった操作を行うために、カメラからの画像を元に、ある特定の身振りや手振りを認識するものがある [1]。しかし、より複雑な処理を行うために、文字列もジェスチャによって入力できた方が便利である。カメラを用いて文字列を入力する研究として、装着型のカメラを用いて指の軌跡を検出して、空中の手書き文字認識を行う研究 [2] があるが、この方法だと、システム利用者にカメラを装着するという負担がかかる。

そこで我々は、利用者に装置を装着せずに、ある程度の広い空間内で、人がどの方向を向いていても、空中での手書き文字の認識が出来るように、まず、周囲に設置された複数のカメラの画像から視体積交差法を用いて人物の立体形状を復元し、その立体形状から指先の軌跡を検出し、空中に書かれた文字の認識を行うシステムを開発した。

2 システム構成

2.1 概要

システム構成図を図 1 に示す。認識する対象の空間の周囲に USB カメラを接続したカメラ PC を複数台設置する。各カメラ PC において、背景差分法で人のシルエットを切り出し、その結果を Gigabit Ethernet を通してメイン PC に送信する。そして、

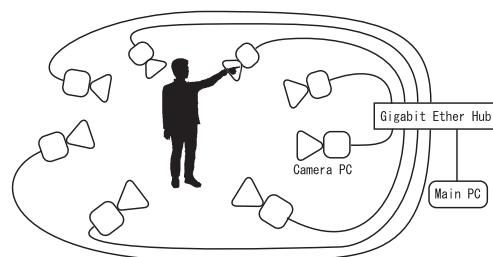


図 1. システム構成図

メイン PC において、人物の立体形状復元・指先の軌跡検出・文字認識の処理を行う。

2.2 人物の立体形状復元

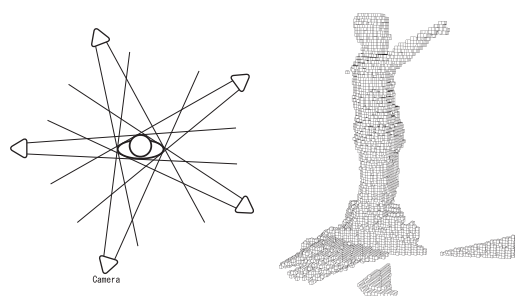


図 2. 立体形状復元

各カメラ PC で背景差分法を用いて切り出した人物のシルエット画像を元に、視体積交差法を用いて人物の立体形状を復元する。視体積交差法の処理は、認識する対象の 3 次元空間をボクセルで表現し、シルエット以外の部分をカットしていく、という方法で行う。具体的には、ボクセルの各点をカメラの投影平面に投影したときに、シルエットの内部がカメラ画像の外部に投影される場合は残し、それ以外の場合、つまり、カメラ画像内のシルエット以外の部分に投影される場合は除去する、という処理をすべ

Copyright is held by the author(s).

* Tsuyoshi Horo and Masayuki Inaba, 東京大学大学院情報理工学系研究科

てのカメラに対して行う．このようにして得られたボクセルの中に，人物以外のノイズとなる物体が現れることがあるので，ボクセル全体をラベリングし，一番大きな集合以外の部分を除去している．

2.3 指先の軌跡の検出

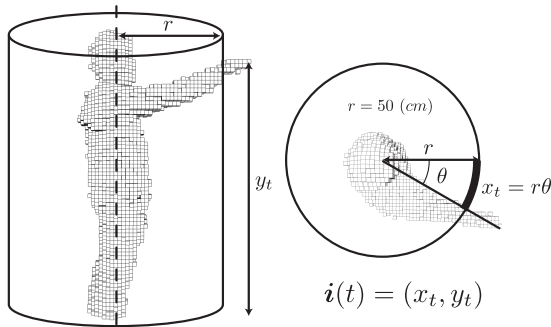


図 3. 指先の軌跡の検出

指先の軌跡を検出するために，まず，人物の立体形状全体の重心線を求め，重心線からの距離が一番遠いボクセルを指先とする．そして，その距離が50cm以上の時に空中に文字が書かれているとして，その期間の $i(t) = (x_t, y_t)$ を軌跡として検出する (図 3)．

2.4 文字認識

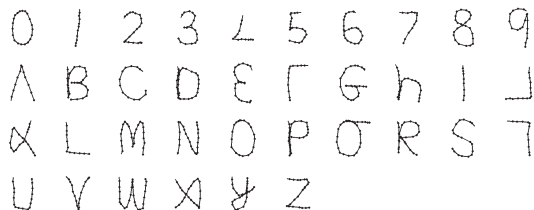


図 4. 事前に登録した文字の軌跡

本システムでは，事前に登録したそれぞれの文字の軌跡 (図 4) と，検出された指先の軌跡 $i(t)$ とを DP マッチングで照合し，一番似ている軌跡を求めることで文字認識を行う．なお，認識に用いる文字の軌跡は，Palm 社の Graffiti を参考にし，一筆書きで文字を認識できるように作成した．また，数字の 0 と英字の O，数字の 1 と英字の I は，互いを区別するのが困難なので，数字を認識するか英字を認識するかを事前に切り替えて行うことにした．

3 評価実験

今回作成したシステムの性能を測定するため，以下のような実験を行った．

3.1 実験方法

人物の周囲に 8 台のカメラを設置し，3 人の被験者に全英数字 36 文字を 3 回ずつ入力してもらい，システムの処理能力・認識率を測定した．認識の対象となる空間は， $200\text{cm} \times 200\text{cm} \times 180\text{cm}$ で，ボクセルの一辺は 2cm とした．カメラの解像度は QVGA (320×340) で，FPS は 30 で，カメラ PC の CPU は Intel Core Duo T2600 (2.16GHz)，メイン PC の CPU は AMD Athlon 64 FX-57 (2.81GHz) である．

3.2 実験結果

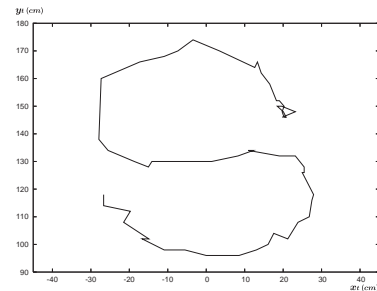


図 5. S を書いたときの指先の軌跡

視体積交差法を用いて人物の立体形状を復元し，指先の軌跡を検出する処理は約 22fps で実行でき，図 5 のような指先の軌跡を検出できた．DP マッチングを用いた文字認識の認識率は約 96.6% で，数字は 95.6% (86/90)，英字は 97.0% (227/234) であった．

4 まとめ

本稿では，周囲に設置された複数のカメラを用いて人体の立体形状を復元し，その形状を元に指先の軌跡を検出し，空中に書かれた文字を認識するシステムを提案した．また，このシステムの性能を検証し，すべての処理はリアルタイムで実行でき，文字認識は約 96.6% という高い認識率であることを確認した．今後は，このシステムを応用した様々なアプリケーションを作成していきたい．

謝辞

本システムの一部は情報処理推進機構 (IPA) の 2006 年度未踏ソフトウェア創造事業の支援を受けて開発された．

参考文献

- [1] 鈴木健一郎, 和田正樹, 梅田和昇. インテリジェントルームにおける家電機器操作の高度化. 日本機械学会ロボティクス・メカトロニクス講演会'06, 2006.
- [2] 園田 智也, 村岡 洋一. 空中での手書き文字入力システム. 電子情報通信学会論文誌, J86-D-II(7):1015 - 1025, 2003.