

複数カメラを用いたユーザーインターフェースの開発

1. 背景

コンピュータの小型・高性能化が進み情報通信技術が発達普及した現在、Weiser によって提唱されたユビキタスコンピューティングという概念、つまり、生活環境のあらゆる場面にネットワークで結合されたコンピュータが存在し、人がその存在を意識することなく恩恵をうけることができる世界の実現が期待されている。しかし、現状では人と情報世界を繋ぐユーザーインターフェースとして、キーボードやマウスといった、特殊なデバイスに拘束され、非直感的で特殊な操作が必要なものが広く普及している。カメラを用いて人のジェスチャを認識することによって、人に何も装置を把持させる必要が無いユーザーインターフェースシステムはすでに提案されてきているが、それらは使用する範囲・方向が限られているものがほとんどである。

2. 目的

本プロジェクトでは、上記背景をふまえて、人に位置的な拘束を与えることなく、コンピュータやカメラの存在を意識せずに使えるユーザーインターフェースシステムの開発を目的としている。

3. 開発の内容

本プロジェクトでは、環境に設置した複数のカメラからの画像を用いて視体積交差法によって人の立体形状を復元し、その立体形状を元に人の指先の位置を認識し、それを入力インターフェースとして用いるシステムの開発を行った。

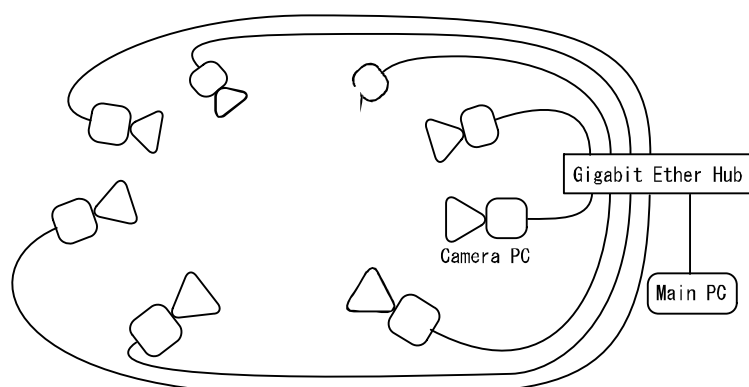


図 1 システム構成

まず、複数のカメラ画像から背景差分法でシルエットを切りだし、ボクセルで表現された認識対象空間をシルエットの情報を基に視体積交差法によって立体復元する。そして、復元された立体形状から人の領域を切りだし、人の立体形状から頭の位置や指先の位置などを認識し、ジェスチャを認識する。



図 2 入力画像

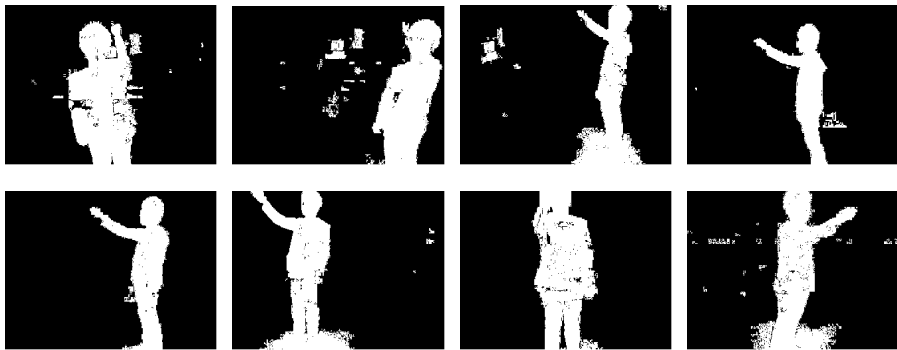


図 3 シルエット画像

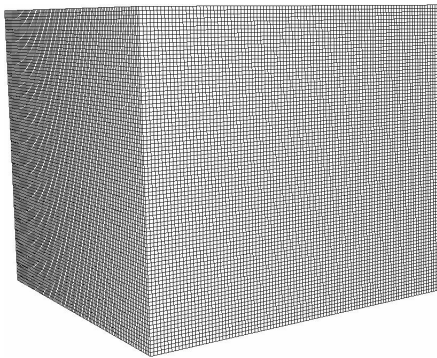


図 4 認識対象空間のボクセル

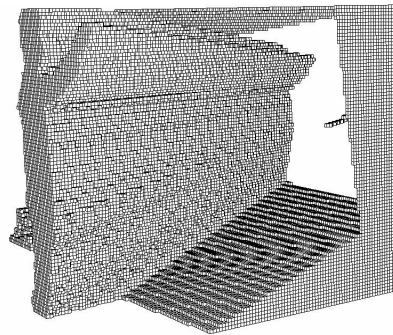


図 5 視体積交差法の途中状態



図 6 立体復元結果

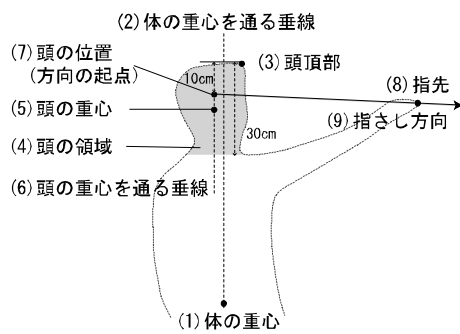


図 7 指さし方向の認識

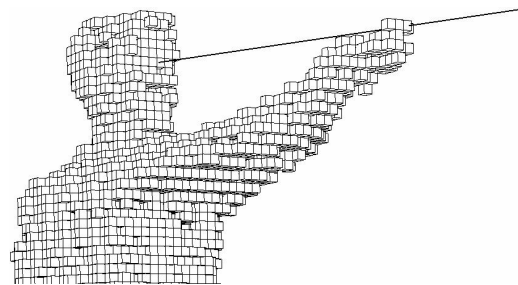


図 8 指さし方向の認識結果

複数カメラ画像から得た立体形状を基にジェスチャ認識を行うので、人は装置を装着及び把持する必要がなく、コンピュータの存在を意識することなく、認識空間内であればどこにいても、どの方向を向いていても利用可能である。認識対象空間を $2\text{m} \times 2\text{m} \times 1.8\text{m}$ とし、ボクセルの一边を 2cm として実験を行った結果、このシステムはリアルタイム処理可能であることを確認した。また、応用システムとして、ポスター解説システム・腕振り地図操作システム・空中手書き文字認識システム・腕の振りを利用したゲーム・移動ロボット操作システムを作成した。なお、本システムは、ネットワーク経由で外部のソフトウェアに認識結果を提供する機能を実装しており、本ユーザーインターフェースを使った応用アプリケーションを容易に作成可能である。



図 9 ポスター解説システム



図 10 腕振り地図操作システム



図 11 ゲーム風景

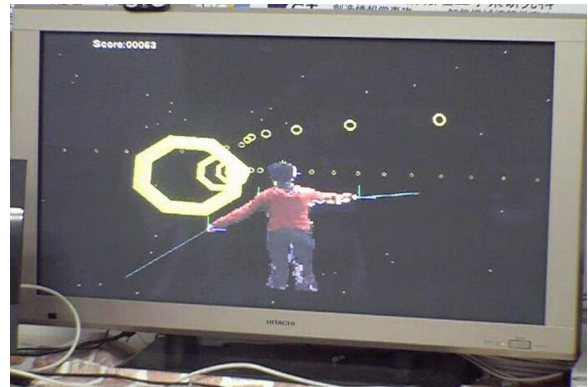


図 12 ゲーム画面

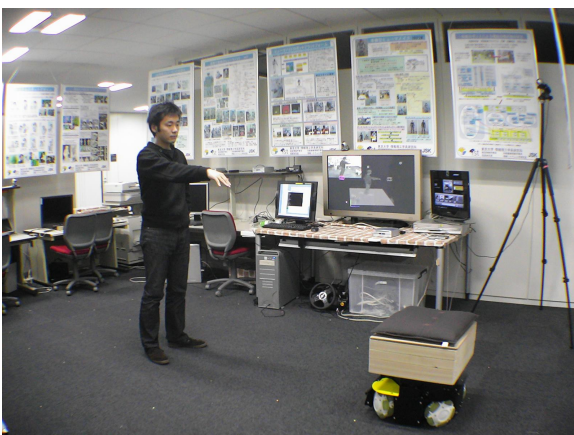


図 13 遠隔操作型インターフェース

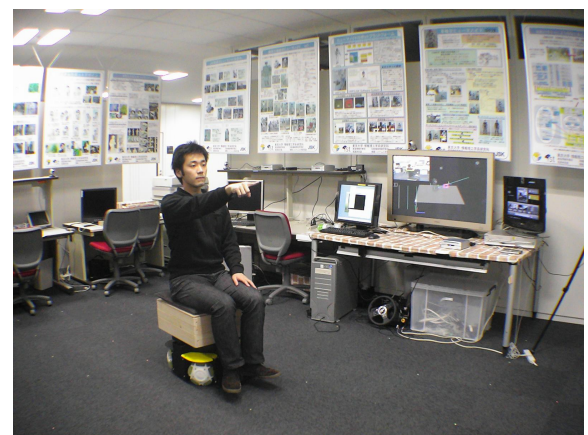


図 14 搭乗型インターフェース

4. 従来の技術(または機能)との相違

装置を装着及び把持する必要がなく、コンピュータの存在を意識することなく、認識空間内であればどこにいても、どの方向を向いていても利用可能であるという特徴を持ったジェスチャによるユーザーインターフェースシステムは今まではあまり存在していなく、本プロジェクトで開発したシステムは非常に画期的なシステムであると思われる。

5. 期待される効果

非常に画期的なユーザーインターフェースシステムであるので、今まで想像もつかなかったようなまったく新しい応用アプリケーションの開発に繋がっていくかもしれない。特に、ゲームなどのアミューズメント用途に親和性が高そうであるので、今までにないような新しいアミューズメントに繋がっていくことが期待される。

6. 普及(または活用)の見通し

本プロジェクトで開発を行ったソフトウェアを開発者ホームページにて誰でもダウンロードできる状態で公開している。詳細は下記開発者URLを参照のこと。

7. 開発者名(所属)

保呂 毅(東京大学大学院 情報理工学系研究科)

(参考)開発者URL

<http://www.jsk.t.u-tokyo.ac.jp/~horo/>