

# 複数カメラを用いたインタラクティブな個別情報支援システム

## Interactive Information Support System Using Multiple Cameras

岡谷 (清水) 郁子\*1    林翔子\*1    梅谷光\*1    冠城久芳\*1    西村 拓一\*2  
Ikuko Shimizu Okatani    Shoko Hayashi    Ko Umetani    Hisayoshi Kaburagi    Takuichi Nishimura

\*1 東京農工大学

Tokyo Univ. of Agri. and Tech.

\*2 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology

To realize a location-based information support system, we have been developing an information support system using Compact Battery-less Information Terminals (CoBITs). Our elementary system consists of multiple cameras for estimation of users' position with CoBITs, and multiple LED light sources on the controllable pan-tilt heads for sending the optical beam information to respective users. In this paper, we discuss about stable position estimation of multiple users with using multiple cameras, and users' gesture recognition to communicate with users and provide appropriate information for respective users.

### 1. はじめに

近年、「いつでも、どこでも、誰でも」情報にアクセスできるユビキタス情報環境が実現されつつあるのに伴い、ユーザの場所、状態、意図に応じた「いま、ここで、私が」ほしい情報が簡単な操作で提供することが重要になってきている。このような情報支援を実現するため、筆者らは、無電源小型通信端末 CoBIT[1]を用いたユーザの位置に基づく情報支援システムを提案している [2]。

本研究は、博物館、美術館などのイベント空間において、ユーザとインタラクションしながらユーザの欲しい情報を提供するための個別情報支援システムの開発を目的としている。基本システムでは、CoBIT を装着したユーザに対し、環境に設置した複数カメラ、方向制御 LED 光源から成る環境ユニットを用いて情報支援を行う。本稿では、このシステムにおける複数カメラを用いた複数のユーザの 3 次元位置の追跡、ユーザとのインタラクションのためのジェスチャ認識について検討する。

### 2. システムの概要

本システムは、図 1 のように、無電源小型通信端末 CoBIT と環境ユニットから構成する。CoBIT は、イヤホン、太陽電池、反射シートから構成された小型の端末である。一方、環境ユニットは、複数の赤外光投射・可視光カットカメラと、方向制御可能な LED 光源から構成する。LED 光源から送信された音声情報を変調した光波を CoBIT が受信し、ユーザはイヤホンから音声情報を聞くことができる。

イベント空間を動き回る CoBIT を装着したユーザに対し、イベント空間のひとつの展示品のおかれている場所について、ひとつずつ環境ユニットを設置し、情報支援を行うことを仮定している。ひとつの展示品の前にいる複数のユーザに対し、複数の赤外光投射・可視光カットカメラを用いて各ユーザの 3 次元位置を計測し、各ユーザに対して光源をひとつずつ割り当て、各ユーザの方向に向けて光源を制御し、個別の情報を送信する。また、ユーザのジェスチャを認識し、ユーザとのインタラクションを行い、ユーザの嗜好にあわせて送信する情報を変化させる。

連絡先: 岡谷 (清水) 郁子, 東京農工大学 大学院共生科学技術研究部, 東京都小金井市中町 2-24-16, e-mail: ikuko @ cc.tuat.ac.jp

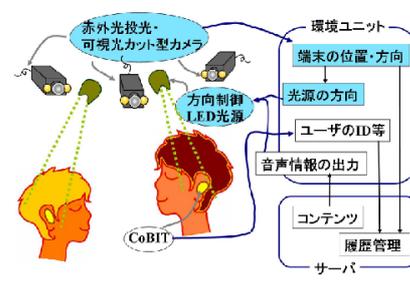


図 1: システムの概要。

### 3. 複数カメラと光源のキャリブレーション

赤外光投射装置を装着した可視光カットカメラで CoBIT を観測すると、カメラの位置から投射された赤外線を CoBIT 表面の反射シートで反射した像がうつるので、CoBIT のみが輝点として容易に検出でき、画像中での端末の位置の抽出は容易である。そこで、複数のカメラの画像中での位置から端末の空間での 3 次元位置を推定し、端末の方向に情報を送信するために重要となるのは、複数のカメラと LED 光源のキャリブレーション (校正)、つまり、カメラと LED 光源の環境中での位置、姿勢やカメラの焦点距離等のパラメータをあらかじめキャリブレーションしておくことである。

本システムでは、イベント空間での情報支援を行うことを考え、短時間で簡単かつ安定にシステムのセットアップを行うため、まず複数カメラのみを反射シートを用いてキャリブレーションし、キャリブレーションされたカメラシステムに対して光源のキャリブレーションを行う。

#### 3.1 複数カメラのキャリブレーション

複数カメラの組に対して、カメラの環境中での位置、姿勢やカメラの焦点距離等のパラメータを求めるキャリブレーションを行うために、キャリブレーションが容易なセルフキャリブレーションの手法 [3, 4] を適用する。セルフキャリブレーションでは、参照物体を利用せず、画像間の点の対応だけからカメラの位置、姿勢や焦点距離などのパラメータを求めることができる。ただし、絶対的な 3 次元座標を知るためには、別途長さの基準を与え、スケールの不定性を解消する必要がある。

本システムでは、長さが既知の棒の両端に反射シートを取り付けたものを利用してキャリブレーションを行う。これにより、この 2 点間の距離を用いてスケールの不定性を解消でき

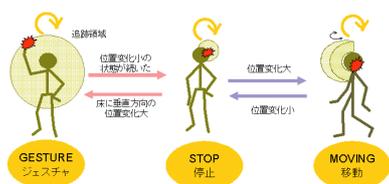


図 2: ユーザの状態に応じた時系列の対応付け.

る。ここで、1枚の画像中に2つの反射シートの像の対応付けが問題になるが、本システムでは、この棒を上下に反射シートがくるように立ててもち、空間を移動させながら各カメラで画像を得ることとし、画像中での上下関係から点を対応付ける。

### 3.2 光源のキャリブレーション

キャリブレーションされた複数カメラシステムに対し、光源をひとつずつキャリブレーションする。これは、光源からいろいろな方向に向けてビーム光を照射し、これを平面に投影した像を用いる。既にカメラはキャリブレーションされているので、平面に投影された像の3次元位置を推定することができる。そこで、ビーム光の平面上の像の3次元位置と、このときの光源の向きを用いて、光源の位置や方向のキャリブレーションが可能である。

## 4. 複数ユーザの追跡

空間を動きまわる複数のユーザの位置をそれぞれ正確に推定し、ユーザの行ったジェスチャを認識するために、複数のユーザの装着している CoBIT の像を複数のカメラで得た画像間に対応付け3次元位置を得、さらに、3次元位置を時系列で対応付ける必要がある。

本システムでは、赤外光投影・可視光カットカメラで CoBIT を観測しており、CoBIT の像のみが輝点としてうつるため、通常用いられるような、画像の輝度の変化や色などの画像特徴を手がかりに対応付けを行うことはできない。そこで、本研究では、画像間の対応付けにはカメラ間の幾何学的関係を、時系列での対応付けにはユーザの状態に応じた動きの連続性を、それぞれ用いる。

画像間の対応付けには、複数カメラから得られる画像の幾何学的関係から得られるエピポラ拘束 [3, 4] を用いる。これは、キャリブレーションされたカメラ間で利用できるもので、あるカメラにうつった点の他のカメラでの像は、直線上に拘束されるというものである。

時系列での対応付けには、動きの連続性の仮定を用いる。本システムではユーザがジェスチャを行うことがあるため、ユーザの動作がその状態によって非常に大きく異なる。そこで、ユーザの動作を停止/ジェスチャ/移動の三つの状態に分類し、各状態に応じた異なる追跡領域を設定する。図2に示すように、領域の重複や、オクルージョンが生じた場合を考慮し、各ユーザの向き、方向、移動スピードに基づき探索領域を適応的に変化させて対応付けを行い、ユーザの3次元位置を推定する。

## 5. インタラクションのためのジェスチャ認識

本システムでは、簡単なジェスチャでユーザとのインタラクションを行う。ジェスチャ認識には、従来手法 [5] と同様に連続 DP に基づくマッチング法を用いる。この手法は、リアルタイム性に優れ、少ない教示回数でも高い認識率が得られるという利点がある。本システムでは、多くのユーザが同じ場所にい

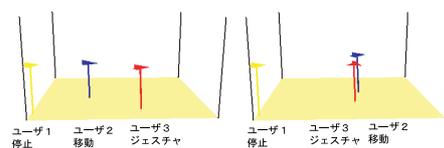


図 3: 複数ユーザの3次元位置推定結果.

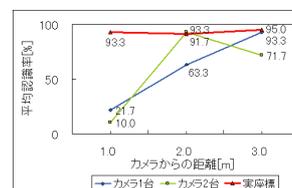


図 4: ジェスチャ認識結果.

ることが考えられ、教示位置と異なる場所でジェスチャを行う可能性が高い。そこで、特徴量として、画像上での2次元位置の変化ではなく、空間中での3次元位置の変化を用いる。

## 6. 実験

複数ユーザの3次元位置推定結果を図3に示す。3名のユーザは、それぞれジェスチャ、移動、停止しているが、各ユーザの位置を安定に推定していることがわかる。また、教示したジェスチャの認識実験を行った結果を図4に示す。比較のため、1つのカメラから得られた2次元の画像座標の変化を特徴量とした場合、2つのカメラから得られた2次元の画像座標の変化を特徴量とした場合の実験も行った。その結果、2次元の画像座標の変化を特徴量とした場合にはいずれも教示位置 (2m) 以外での認識率が低くなっているのに対し、3次元位置を特徴量として利用した場合には、教示位置以外でも高い認識率が得られた。

## 7. まとめ

本稿では、イベント空間におけるインタラクティブな個別情報支援システムの開発を目指し、複数カメラ、方向制御ビーム光源、CoBIT から成る基本的なシステムを構成した。簡単にセットアップを行うためのカメラと光源のキャリブレーション手法、複数のユーザの3次元位置の頑健な推定手法、ジェスチャ認識手法について検討した。

実用に向けて、ユーザの位置やジェスチャを有効に利用し、履歴管理・コンテンツサーバと連動した個別情報支援システムを構築することが今後の課題である。

## 参考文献

- [1] 西村他. 無電源小型通信端末を用いた位置に基づく状況支援システム. 情処研報, Vol. ICII-2, pp. 1-6, 2002.
- [2] I. S. Okatani et al. Location-based Information Support System Using Multiple Cameras and LED Light Sources with the Compact Battery-less Information Terminal (CoBIT), Proc. of Int. Workshop on CV in HCI, pp. 142-152, 2004.
- [3] D. Forsyth et al. *Computer Vision - A Modern Approach*. Prentice Hall, 2003.
- [4] 松山他 (編). コンピュータビジョン—技術評論と将来展望. 新技術コミュニケーションズ, 1998.
- [5] 西村他. 動作者適応のためのオンライン教示可能なジェスチャ動画のスポッティング認識システム. 信学論, Vol. J81-D-II, No.8, pp.1822-1830, 1998.