

## レンジイメージを用いた空間内の時系列人物認識とモデル化

土屋 友寛<sup>†</sup> 大石 浩徳<sup>†</sup> 西村 直久<sup>†</sup> 有澤 博<sup>‡</sup>

<sup>†</sup> 横浜国立大学大学院環境情報学府 〒240-8501 横浜市保土ヶ谷区常盤台 79-7

<sup>‡</sup> 横浜国立大学環境情報研究院 〒240-8501 横浜市保土ヶ谷区常盤台 79-7

E-mail: <sup>†</sup> {tsuchiya, thomyorke, nishimura}@arislabs.ynu.ac.jp, <sup>‡</sup> arisawa@ynu.ac.jp

あらまし 空間内の複数の人を認識し自動で追尾撮影することや、人物の挙手・起立・着席などのイベントを検出し、検索可能な状態で映像と共に蓄積しておくことは、監視カメラ、遠隔講義などの分野で非常に有用である。人物認識の手法として一般的な画像認識のアプローチでは、対象となる環境に制約があり、安定して認識できない問題がある。そこで本研究では3次元スキャン型レーザレンジファインダからのレンジイメージを用いて人物を認識し、人物のイベント情報を映像と共に蓄積する手法を提案し、実際にシステムを構築して優位性を確認した。

キーワード レーザレンジファインダ、レンジイメージ、人物認識、イベント検出

### Time series person recognition and modeling in the real world by using range images

Tomohiro TSUCHIYA<sup>†</sup> Hironori OHISHI<sup>†</sup> Naohisa NISHIMURA<sup>†</sup> and Hiroshi ARISAWA<sup>‡</sup>

<sup>†</sup> Graduate School of Environment and Information Science, Yokohama National University

<sup>‡</sup> Faculty of Environment and Information Sciences, Yokohama National University

79-7 Tokiwadai, Hodogaya-ku, Yokohama, 240-8501 Japan

E-mail: <sup>†</sup> {tsuchiya, thomyorke, nishimura}@arislabs.ynu.ac.jp, <sup>‡</sup> arisawa@ynu.ac.jp

**Abstract** We can get the results of recognizing several people in a space and purchasing the motion of them automatically, detecting one's raising hand, standing up, sitting down and other events. It is very useful to store the index-state results with the images, in the field of monitoring camera and remote lecture and etc. Normally, because of the restriction of the circumstance, it is difficult to do the recognition stably by using the images of the cameras. In this research, we propose a new approach. We did the recognition by using a laser measurement which can scan the three-dimension. We also constructed the system and proved the superiority of it.

**Keyword** Laser Range Finder, Range Image, Person Recognition, Event Detection

#### 1. はじめに

現実世界から多種多様な情報を取り込み、実世界の現象をありのままに蓄積・再現し、また評価・シミュレーションを行うことができるシステムを構築することは非常に広い応用範囲を持つと考えられる。

そこで我々は、複数のカメラから得られる多視点同期映像や、多様なセンサからの時系列データを取り込み、データベース化、モデルを用いた解析・解釈、評価・シミュレーションを行うリアルワールドデータベース(Real World Database)[1]を提唱し研究開発を行っている。

この中でもビデオカメラから得られる映像は適切な解析を行うことで非常に豊富な情報が引き出せるため、検索可能な状態で蓄積することができれば、非常に広い応用分野が考えられる。例えば、監視カメラ・遠隔講義などの分野において、人物を認識し空間内の位置を知る事ができれば、自動追尾による撮影が可能

となる。またさらに、人物の挙手・起立・着席などのイベントを検出し、映像と共に蓄積しておく事ができれば、そのようなイベントを条件とした映像の検索が可能となる。

カメラによって撮影された画像中の人物を認識する手法としては、背景画像との差分を取る方法や、肌の色を利用する方法、顔のパターンや人体の3次元モデルとのマッチングを行う方法などが考えられる。しかし先に挙げたような応用例で実際に使用する事を考えると、それらのカメラ画像を解析する手法を単に用いただけでは、設置場所が制限される・観測視野(カメラの画角)が狭い・照明条件の変化に弱いなどの問題が残ってしまう。

そこで本研究では監視カメラや遠隔講義などの分野を想定し、以下の項目が実現できるシステムの構築を目標とした。

- ・ 複数人物の認識、3次元位置の時系列取得
- ・ 人物の自動追尾撮影
- ・ イベント（人物の挙手・起立・着席動作）検出
- ・ イベント等に基づく映像検索

この目的を達成するために、まず我々は、レーザ光のパルスの伝播時間により距離を測定することのできるレーザレンジファインダに着目した。この装置は、ビデオカメラと比較して、ダイレクトな距離計測、高精度計測、広角計測、長距離計測が可能である。この装置と、我々が当研究室で以前から開発している電動チルト架台を組み合わせることにより、空間の3次元スキャンが可能となる。本研究では、この装置によって短時間で空間をスキャンし、距離画像（レンジイメージ）を取得・解析することで人物を認識し、算出した3次元位置に基づいて人物を自動追尾・撮影するシステムを構築した。また、本システムでは人物に関するイベント（挙手・起立・着席）を検出し、人物の空間内位置およびカメラで撮影された映像と共に蓄積する。データの蓄積に際しては、後の解析や検索が効率的に行えるように、前もって定義したスキーマに基づいて蓄積した。このように前述の目的を達成する手法を提案し実際にシステムを構築している。

以降2章では既存の空間内の人物認識手法とその問題点について述べ、3章では提案手法、4章では実際に構築したシステムの詳細、5章では実験と考察、6章で本稿のまとめを述べる。

## 2. 既存の人物認識取得手法

既存研究における空間内の人物認識する取得する手法は、大きく分けてカメラから得られる映像を用いて人物を認識する手法や、カメラ以外の機器を用いた手法に分けられる。以下、それぞれの手法について問題点を整理する。

### 2.1. カメラから得られる映像を用いた手法

カメラを用いて人物を認識する様々な手法が提案されている。しかし、どの手法もカメラの特徴である観測視野が狭く、照明条件の変化に弱いという問題を含んでいる。また、これらを解決するために複数のカメラの設置が必要となるが、複数カメラのキャリブレーションは容易ではなく、前述の目的を達成することは難しい。以下、個別の手法について問題点をまとめる。

#### ① 背景差分による手法

計測空間の背景画像を先に取得しておいて、計測時に背景との差分から空間内の人物を抜き出し、抜き出した画像をもとに人物を認識する手法[2]である。この手法では、明るさが変化する環境や、背景が変化するような環境では人物をうまく認識することができない問題がある。

#### ② 色情報を用いる手法

人の肌色情報などの色情報を用いて人物の認識を行う手法[3]がある。人の肌色領域のRGBやHSVパターンによって、画像上の顔領域や手の領域を認識する。この手法も①の手法と同じような環境制約がある。また、人物認識のために服装の色情報などを用いる場合には、認識する人物の服装の色情報を先に取得しておかなければならないので、様々な場面で不特定多数の人物が存在する場面には適していない。

#### ③ モデル情報を用いる手法

計算機内に認識対象の人物のモデルを事前に持っておき、そのモデルと画像を比較することで人物を認識する手法[4]である。この手法はモデルベースマッチと呼ばれ、計算機内のモデルの方向や姿勢を変えて、2次元に透視投影し2次元の画像にして、カメラと取得した画像と比較する。この手法では3次元的な人物の位置・姿勢がわかる。しかし、この手法も②の手法と同じようにモデルを先に取得しておかなければならず、不特定多数の人物が存在する場合には適していない。

#### ④ その他カメラによる人物認識

ステレオカメラを用いて3次元的に解析し人物を認識する手法[5]や、非常に多くのカメラを用いて人物の位置、姿勢を計測する視体積交差法[6]と呼ばれる手法がある。これらの手法に関しても、他の手法と同じような環境制約がある。

## 2.2. カメラ以外の機器を用いた手法

### ① 人物にセンサをつける手法

人物にセンサを付けることで、人物の位置情報を取得し、人物を認識しようとする手法もいくつか提案されている。GPSやジャイロセンサーを人物に付け位置や姿勢を求める手法やRFIDなどのタグを人物に付けておき、そのデータを受信することで位置情報を取得・人物を認識する手法もある。これらの手法も、あらかじめ人物にセンサを付けておかなければいけないので、不特定多数の人物が存在する場合に適していない。また、取得できる人物の位置精度があまりよくないため、人物の自動追尾撮影には向かない。

### ② レーザレンジファインダを用いる手法

既存のレーザレンジファインダを用いて人物を認識する手法は、平面スキャン型のレーザレンジファインダで人物の足を計測し認識する手法[7]や、3次元スキャン型のレーザレンジファインダで人物の形状を抜き出し、人物を認識する手法がある。しかし、前者の手法では人物の2次元的位置しか分からない。また、後者では、既存の3次元スキャン型のレーザレンジファインダは形状計測を目的としている装置のため、一度のスキャンに非常に時間がかかってしまう。

このように、既存手法には我々の設定した目標を満たすものは存在しない。そこで以下では、目標を達成するために当研究室で開発した動的に3次元スキャン可能なレーザレンジファインダを用いたシステムを提案する。

### 3. 提案手法

#### 3.1. 使用する機器

カメラの代わりに人物を認識するために使用するセンサはレーザレンジファインダを用いる。このセンサは、レーザ光のパルス伝播時間により距離を測定するものであり、ビデオカメラと比較して、ダイレクトな距離計測、高精度計測、広角計測、長距離計測が可能という特徴を持つ、またレーザ光の拡散が少ないため解像度が高く、高角度分解能を持つ。このレーザレンジファインダと、当研究室で開発した電動チルト架台を組み合わせることで3次元スキャンを可能にした。電動チルト架台はレーザレンジファインダをチルト方向に自在に制御することができ、レーザレンジファインダの回転ミラー・ホイールの中心とチルトの回転軸が一致した構造で、レーザレンジファインダから得られる距離角度情報 ( $r, \theta$ ) とチルト架台から得られる角度情報 ( $\phi$ ) を統合することができる。この装置を用いることで、短時間で空間を3次元スキャンすることができ、チルト架台を常に動かし続ければ時系列の情報を得られる。

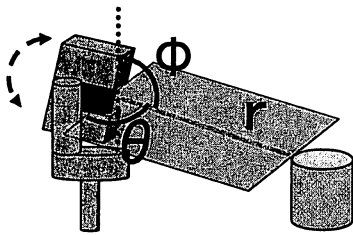


図1：レーザレンジファインダと電動チルト架台

本研究では、この装置から得られた距離情報をレンジイメージとして捉え、画像処理の手法を利用することで、空間内の人物を認識する。この際、レーザレンジファインダを用いてダイレクトに距離計測を行っているため、複数人物の3次元位置、大きさを時系列で取得できる。この位置と大きさの時系列データを利用することで人物の自動追尾撮影、イベント（人物の挙手・起立・着席動作）検出が可能となる。これらの解析結果により得られた情報とビデオカメラの映像を蓄積し、イベント等による映像検索を可能にした。

以下、人物検出の詳細な方法と、イベントによる映像検索のためのデータの蓄積方法を示す。

#### 3.2. 人物検出方法

レンジイメージを用いた人物認識の大まかな流れは事前準備：レーザレンジファインダのキャリブレーション、背景レンジイメージの取得、1. 電動チルト架台をチルトさせて空間のレンジイメージを取得する。2. 背景差分を行い、背景から人物に相当するデータを抜き出す。3. 差分データに移っている人物を分類し、各人物の位置、大きさを計算する。以上を毎チルトごとに行い、毎チルト完了時の人物の位置と大きさの変化を連続して試みていくことで、人物の大きさの変化から人物の挙手・起立・着席動作をイベントとして検出する。

次に処理の各フェーズで行う処理の詳細を述べる。

##### ① レーザレンジファインダのキャリブレーション

複数台のレーザレンジファインダのデータを統合する際や自動追尾撮影のためのカメラ制御には、各機器の座標系を統一座標系に変換しなければならない。そのため各機器の統一座標系における位置・姿勢をあらかじめ求めておく必要がある。

今回レーザレンジファインダの統一座標系での位置・姿勢は計測空間内にキャリブレーション物体を置いて、統一座標が既知の3点との対応付けにより求めた。

##### ② 背景差分

レンジイメージから人物を抽出するために、カメラ画像と同じように、事前に背景データを取得しておき背景差分を行った。人物は背景よりもレンジファインダに対して近くに移るので、背景よりも近くなったピクセルのみを残すことにより実現できる。



a: レンジイメージ

b: 背景

c: 差分

図2：背景差分

##### ③ ノイズ除去

背景差分を行った後の差分データにはノイズが含まれることが多い。これらは特に物体のエッジ部分に現れることが多く、差分データには直線として現れる。これはレーザレンジファインダの特性から現れるものと考えられる。本手法では差分データにHaff変換をかけることでノイズを除去した。

##### ④ レンジイメージの領域分割

得られた差分データに何人の人物が存在するのかわ、解析するためにレンジイメージの領域分割を行う必要がある。今回は、レンジイメージの輝度値から、シードフィルアルゴリズムを用いた。領域を分割する

ために人物を構成する距離データは連続的に変化をするため、閾値に幅を持たせる必要がある。実験では閾値 100mm とした。



図 3 : レンジイメージの領域分割

⑤ 人物の位置・大きさの計算

次に人物の大きさについてその人物に含まれる点データのうち各成分で最大値、最小値の差で求めた。1台のレーザレンジファインダから得られるのは人物の片面だけであるので正確な大きさが必要な場合は人物の全周囲分のデータがあることが望ましい。そのためにはレーザレンジファインダの配置を工夫する必要がある。

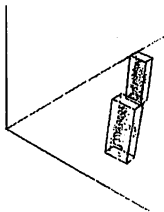


図 4 : Bounding Box

⑥ フレーム間対応

対象空間内に複数の人物が存在する場合、あるフレームで検出された各々の人物が他のフレームではどの人物と同一のものであるかを判別する必要がある。そのために各フレームにおける人物の位置を比較し、最も近接しているものを同一の人物であると判定した。

⑦ イベントの検出 (大きさの変化の検出)

手を挙げた等のイベント検出法については、人物の Bounding Box のサイズの変化を用いて検出した。垂直成分 (高さ) の大きさをフレーム間で比較を行い、変化量の大きさを計算し、その変化量によってイベントであると判定した。

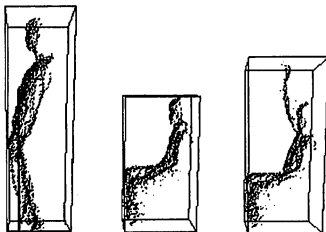


図 5 : イベント検出

⑧ 複数台のレーザレンジファインダの使用

広範囲な空間を対象としたい場合には複数台のレーザレンジファインダを用いる。それぞれ1枚ずつレンジイメージを領域分割した後、人物を外含する Bounding Box を考える。Bounding Box に重なりが生じた場合、同一人物であるというような手法で複数台のレンジファインダに映し出された人物のデータの統合を行った。

3.3. データ蓄積方法

以上のような、相互に関連する様々なデータを蓄積する形式として、当研究室が提案しているデータ構造 BOE(Binary Object Expression)を用いた。BOE ではモデリングの対象となる事柄を、属性名を持つオブジェクトの入れ子として表現する。すなわち

```
オブジェクト := 属性名:値
                | 属性名:{値,値,...}
                | 属性名:タプル
                | 属性名:{タプル,タプル,...}
```

タプル := [オブジェクト, オブジェクト, ...]

なるBNFによって定義される構造である。

上記の定義に基づいて、具体的なデータ構造を記述するためのスキーマが既に提案されており、実際に BOEデータをプログラム上で読み込み・操作・保存するためのライブラリも実装されている。表 1 に、本システムが取得・蓄積するデータのスキーマを示す。

表 1 : BOEスキーマ

| 階層構造 | 属性名           | データ型    |
|------|---------------|---------|
| □    | 撮影            |         |
| └─   | LRF ID        | int     |
| └─   | 時刻データ         |         |
| └─   | 静的データ         |         |
| └─   | 回転行列: R       | matrix  |
| └─   | 平行移動行列: T     | matrix  |
| └─   | 背景レンジデータ      |         |
| └─   | 距離: r         | double  |
| └─   | パン角: $\theta$ | double  |
| └─   | チルト角: $\phi$  | double  |
| └─   | 時系列データ        |         |
| └─   | 時間: t         | double  |
| └─   | レンジデータ        |         |
| └─   | 距離: r         | double  |
| └─   | パン角: $\theta$ | double  |
| └─   | チルト角: $\phi$  | double  |
| └─   | 解析データ         |         |
| └─   | オブジェクト        |         |
| └─   | オブジェクトID      | int     |
| └─   | 時点            |         |
| └─   | 時間: t         | double  |
| └─   | 位置            | 3Dpoint |
| └─   | 大きさ           | 3Dsize  |
| └─   | 形状レンジデータ      |         |
| └─   | 距離: r         | double  |
| └─   | パン角: $\theta$ | double  |
| └─   | チルト角: $\phi$  | double  |
| └─   | カメラID         | int     |
| └─   | 静的データ         |         |
| └─   | 回転行列: R       | matrix  |
| └─   | 平行移動行列: T     | matrix  |
| └─   | 内部パラメータ       |         |
| └─   | 焦点距離: f       | double  |
| └─   | レンズ歪み係数: k    | double  |
| └─   | 画像中心          | 2Dpoint |
| └─   | 時系列データ        |         |
| └─   | 時間: t         | double  |
| └─   | 撮影オブジェクトID    | int     |
| └─   | 画像            | image   |
| └─   | バルス値          |         |
| └─   | パン値           | int     |
| └─   | チルト値          | int     |
| └─   | ズーム値          | int     |

## 4. 実験

ここでは実際に構築したシステムで行った複数人物を認識する実験の内容と結果について述べる。

### 4.1. 実験環境

6m四方の室内において数名の人を配置し、全体が十分に見渡せる位置にレーザレンジファインダ、カメラを設置し実験した。事前にレーザレンジファインダおよびカメラはキャリブレーションを行った。

以下にシステム構成図、実際に使用した機器それぞれの仕様を示す。

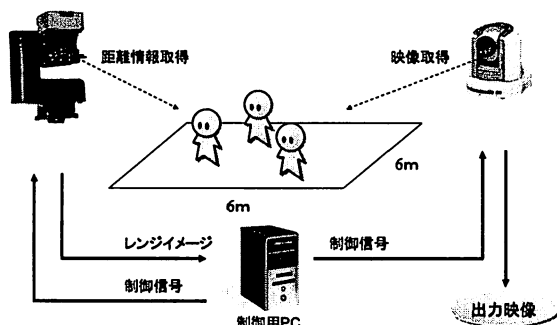


図 6 : システム構成図

表 2 : レーザレンジファインダ仕様

|              |              |
|--------------|--------------|
| 名称           | LMS200-30106 |
| スキャン角度       | 180度         |
| 角度分解能        | 0.5度         |
| 応答時間         | 53ms         |
| 測定分解能        | 10mm         |
| システム誤差       | 代表値±15mm     |
| 統計的誤差        | 5mm          |
| データ・インターフェース | RS422        |
| 伝送速度         | 500kbaud     |
| レーザ保護クラス     | 1(本質的に安全)    |
| 重量           | 4.5kg        |

表 3 : チルト架台仕様

|        |                    |
|--------|--------------------|
| 名称     | 電動チルト架台            |
| 最大搭載質量 | 5kg                |
| 動作範囲   | +90度~-45度          |
| 最大速度   | 30度/秒              |
| 加減速制御  | S字制御               |
| 制御分解能  | 153600pulse/360度   |
| 絶対位置精度 | ±10pulse           |
| 送信間隔   | 52ms               |
| 通信     | USB ver2.0         |
| 寸法     | W292.5×D140×H340mm |
| 質量     | 6kg                |

### 4.2. 実験内容

この実験では構築したシステムが前述の

- ・ 複数人物の認識、3次元位置の時系列取得
- ・ 人物の自動追尾撮影
- ・ イベント(起立・着席動作)の検出

が可能かどうかを確認するために、以下の3点について実験を行った。

- 複数人物の認識  
対象空間内にいる人を正確に区別しカウントする。
- イベントの検出  
人物のイベント(起立・着席動作)を検出しイベントが起こった人物にカメラをズームし撮影する。
- フレーム間の対応  
複数の動いている人物を認識し各フレーム間で同一人物であると認識できるか。

### 4.3. 実験結果

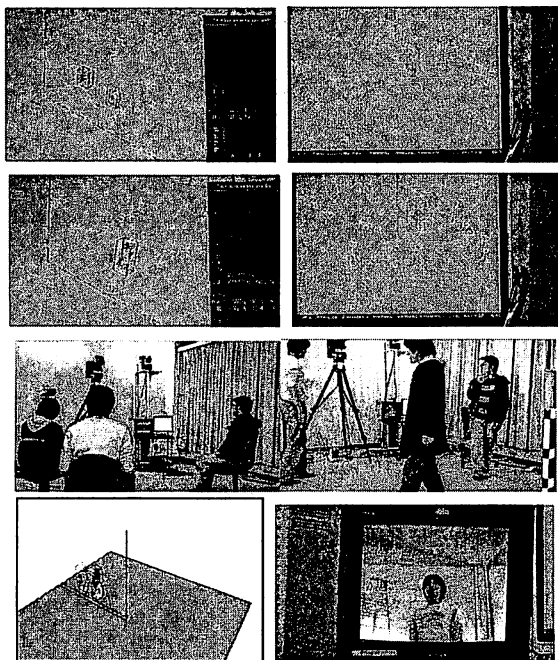


図 7 : 実験結果

- 人物の認識  
空間内にいる人物を正確に区別し、カウントすることが確認できた。
- イベントの検出  
人物のイベント(起立・着席動作)を検出し動作した人にカメラをズームし撮影することができた。
- フレーム間の対応  
複数の動いている人物を認識し各フレーム間で同一人物であると認識できた。

## 5. まとめ

我々は監視カメラ・遠隔講義などの分野において重要な、複数人物の認識、3次元位置の時系列取得・人物の自動追尾撮影・イベント（人物の挙手・起立・着席動作）検出・イベント等に基づく映像検索が可能となるシステムを目指した。本稿では、本研究の有用性について述べ、既存手法と、その問題点をまとめた。また、目標としたシステムの構築を達成すべく、3次元スキャン型レーザレンジファインダから得られるレンジイメージを用いた時系列人物認識手法を提案した。さらに提案した手法を用いたシステムを実際に構築、複数人物の認識・自動追尾撮影・イベント検出を行う実験を行った。

今後、蓄積したデータを検索・表示を行う部分の実装や、より多数の人物がいる環境下（実際の遠隔講義の場面）での実験、複数のレンジイメージの統合処理によるオクルージョンの改善、レーザレンジファインダの高速な水平スキャンデータをを用いた高速な自動追尾撮影などを行うことで、より実用的で有用なシステムを構築できると考えている。

## 文 献

- [1] 有澤 博, “リアルワールドデータベースとその実現技術,” bit, 共立出版, 1996.9/1996.10/1996.11, Vol.28.29.30, No.9.10.1.
- [2] 物部 祐亮, 和田俊和, 松山隆司, “視点固定型パン・チルト・ズームカメラを用いた実時間対象検出・追跡,” 情報処理学会研究会資料, CVIM109-2, 1998.1
- [3] 松村 朱里, 岩井 儀雄, 谷内田 正彦, “肌色情報を用いた複数人物追跡,” 情報処理学会研究報告. CVIM, IPSJ SIG Notes. CVIM Vol.2002, No. 34 (20020509) pp.133-138
- [4] 坂木和則, 有澤博, “モデル知識を用いた遺伝的アルゴリズムによる人間の姿勢推定,” 情報科学技術フォーラム2002(FIT2002), 第3分冊, pp.45-46, Sep. 2002
- [5] 鈴木 康平 “リアルタイム画像解析による被写体自動認識システムの実現” 横浜国立大学院 環境情報学府 2005 修士論文
- [6] K.M. Cheung, S. Baker, and T. Kanade “Shape-From-Silhouette Across Time Part I: Theory and Algorithms” International Journal of Computer Vision, Vol. 62, No. 3, May, 2005, pp.221 - 247.
- [7] 中村克行, ほか, “マルチレーザスキャナを用いた歩行者の抽出,” 第10回画像センシングシンポジウム講演論文集, 横浜, 2004-06, 画像センシング技術研究会, pp.409 - 414.