

大西 研究室	氏 名	川 野 隆 亮
論 文 題 目	視聴覚情報を用いた複数人物の追跡	

### 背景と目的

人間は動いているものを目で追い、耳で音の発生している方向を定位することができる。例えば、複数の人物から話者を特定し、その人物を見続けることができる。このような機能の実現は、ビジョンシステムにおいて有益な応用である。そこで、我々はカメラとマイクの2種類のセンサを利用して、複数の人物を追跡するシステムを構築する。

人物の情報をよりたくさん得ようと考えた場合、顔を撮影することが望ましい。そこで、人物領域の特徴として頭部を検出する。また、複数の人物を洩れなく追跡するために、カメラに「追跡」と「監視」の2種類の役割を与える。これらの役割をカメラの連携を用いて動的に変化させることで、注目している人物の詳細な情報を獲得でき、他の人物も見逃すことなく追跡を行なうシステムを構築する。

### システム概要

システムは、視聴覚センサ部においてカメラから頭部領域、マイクから音信号を検出する。検出結果をセンサ統合部において人物情報の更新や音源位置の特定を行なう。これらの結果から、どのカメラがどの人物を追跡するかを決定する。あるシーンにおいて何か話している人物を撮影できれば、そのシーンを理解するのに役立つ。そこで、音源位置情報から音を発生している人物を特定し、その人物を優先的に追跡する。追跡する人物が決定したら、その人物を追跡するカメラを決定する。それは、人物を高解像度で撮影するためできるだけ近いカメラで、かつ、正面から人物を撮影できるカメラを割り当てる。さらに、オクルージョンの発生により割り当てられたカメラでは人物を捉えられない場合があるので、新たに他のカメラに追跡を委託する。以上の処理を繰り返すことにより人物の追跡を行なう。

### 実験結果

3人の人物を追跡する実験を行なった。追跡結果を図1、追跡の様子を図2に示す。各人物の位置情報を獲得し、人物の顔を捉えて効果的に追跡することが確認できた。また、音情報を得た場合、音を発した人物を優先的に追跡し、追跡中の人物がオクルージョン領域に入った場合に他のカメラが追跡することを確認した。

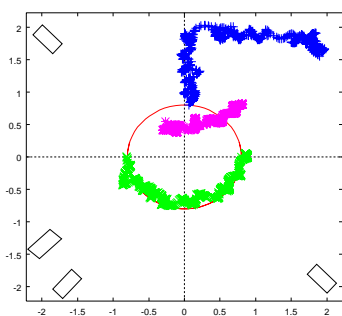


図 1 追跡結果



図 2 追跡の様子

### 発表実績

- 川野隆亮, 岡島正道, 竹内義則, 大西昇, “複数のパン・チルト・ズームカメラによる複数人物の監視と追跡”, SICE 第 2 回 SI 部門学術講演会, 2001.12
- 川野隆亮, 竹内義則, 大西昇, “視聴覚情報を用いた複数人物の追跡”, 2003 年 3 月 PRMU 研究会において発表予定.
- 他、国内 2 回発表.