

3W-8

## CMOS カメラを用いた小型状況認識システム

品川 泰英<sup>†</sup>  
慶應義塾大学 理工学部

長名 保範<sup>‡</sup>  
慶應義塾大学 理工学研究科

今井 倫太<sup>§</sup>  
慶應義塾大学 理工学部  
科学技術振興機構 さきがけプログラム

sinagawa@ayu.ics.keio.ac.jp, osana@am.ics.keio.ac.jp, michita@ayu.ics.keio.ac.jp

## 1 はじめに

本研究では、ロボット・センサネットワークにおける効率的な帯域利用を実現するために、人間の動作を認識する環境カメラノードシステムを提案する。

コミュニケーションロボットは人間の状況に応じて行動を決定するため、人間の動作を的確に把握する必要がある。そのためにはコミュニケーションロボット自身に取り付けられたセンサからの情報のみでは不十分であり、ロボットの周辺に配置した環境カメラからの情報を利用する必要がある。たとえば文献 [1] では、環境カメラからのデータ取得によるコミュニケーションロボット Robovie の状況把握の精度向上が実現されている。

従来は、環境カメラから得られた画像は、認識処理のためにネットワークを介してコミュニケーションロボットもしくは専用サーバ転送されてきた。しかし、カメラ数の増加に伴って転送画像データ量はネットワーク帯域を圧迫するほど増大する。例えば、1台のカメラからの VGA 画像 (3byte/ピクセル) を 30fps で画像処理する場合、ネットワーク帯域に占める画像データは 30Mbyte/秒となるから、環境に複数台のカメラを使用するシステムに必要なネットワーク帯域は、 $30 \times$  (カメラの台数) Mbyte/秒となってしまうため、カメラが 5 つ以上になれば 1Gbps のネットワークでも帯域を越えてしまう。

ところで、コミュニケーションロボットの行動決定に必要なのは画像データそのものではなく、画像処理の結果による人間の動作認識である。それゆえ、環境カメラ側で人間の動作を認識できれば、転送データ量を減らすことができ、ネットワーク帯域の圧迫という問題を解決できる。

そこで本研究では、環境カメラにおいてコミュニケーションロボットに必要な画像処理を実現する、環境カメラノードを提案する。

## 2 想定環境

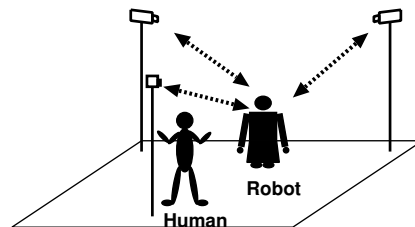


図 1 想定環境

本研究の想定環境を図 1 に示す。想定環境では、環境カメラの情報を利用してコミュニケーションロボットが人間を認識する精度を高める。

画像処理を用いて認識する人間の動作には、回転動作と注目動作がある。人間の回転を認識するには、フレーム間の肌色成分の増減値 (フレーム間差分) が必要である。また、人間がどの方向を向いているか向きを認識するには背景差分や肌色の大きさの値が必要である。

## 3 小型状況認識システム

回転動作と注目動作を環境カメラ側で認識するために、本研究では環境カメラノードである小型状況認識システムを提案する。

本システムは、ロボットからの入力信号を受け、コミュニケーションロボットの行動決定に必要な画像処理結果のみをネットワーク上に流す事が可能である。小型状況認識システムの外観を図 2 に示す。

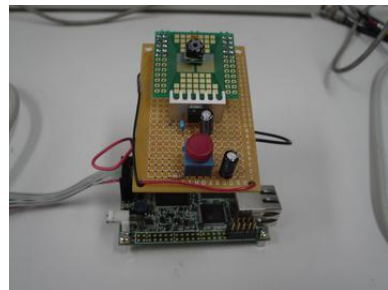


図 2 システム外観

## 3.1 システム構成

システム構成を図 3 に示す。本システムでは、画像処理回路を構成するのに FPGA を使用した。画像処理用の FPGA ボードには、Xilinx 社の spartan-3 を搭載するアットマークテクノ社製の SUZAKU を、CMOS カメラは Pixelplus 社の VGA 出力 CMOS イメージセンサ PO1030K を使用した。PO1030K は RGB、YUV、

Compact System for Situation Recognition with CMOS Camera

<sup>†</sup>Yasuhide SHINAGAWA

Faculty of Science and Technology, Keio University

<sup>‡</sup>Yasunori OSANA

Graduate School of Science and Technology, Keio University

<sup>§</sup>Michita IMAI

Faculty of Science and Technology, Keio University

PRESTO, JST

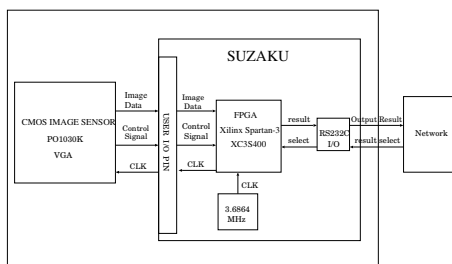


図 3 システム構成図

YCbCr 等のフォーマットで可能であるが、本研究では YCbCr を使用した。また、本システムの動作電圧は直流 3.3V である。

### 3.2 システム仕様

SUZAKU には色検出、フレーム間差分の 2 種類の画像処理を実装した。本システムの入出力は RS232C(115.2kbps) による。外部からの信号に応じて、色検出、フレーム間差分のいずれの結果を出力するかを選択する。

## 4 実装

環境カメラノードを開発するために、言語として verilog を使用し、論理合成ツールとして Xilinx ISE Foundation を使用した。

verilog によって、図 4 に示すモジュールを作成し、シミュレーションでの動作確認の後、spartan-3 上でコンフィグレーションを行った。

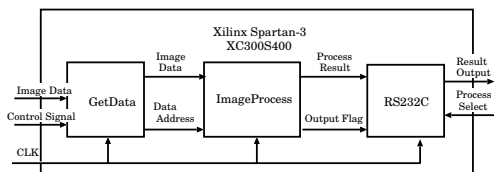


図 4 Spartan-3 内に実装したモジュール

図 4 には 3 つのモジュールがある。GetData モジュールは CMOS カメラから FPGA に画像データ (YCbCr 形式) を取得する。ImageProcess は GetData モジュールから画像データを受け取り、画像処理 (色検出、フレーム間差分) を行い、処理結果を RS232C に受け渡す。そして RS232C モジュールはシリアル通信モジュール。ネットワークを介し、コミュニケーションロボットからの信号を受け取る。その信号を判別して送信する処理結果の内容を決定し出力する。

## 5 認識処理の実現

本システムに実装した画像処理アルゴリズムを述べ、その実行例を示す。

### 5.1 色検出

640 ピクセル × 480 ピクセルのカメラ画像を 8 × 6 の計 48 ブロックに分割した。これを図 5 に示す。1 ブ

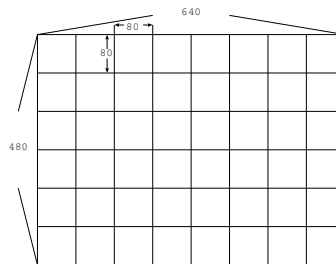


図 5 画像の分割

ロック中 6400(80 × 80) ピクセルのうち閾値以上のピクセルが検出すべき色の範囲であれば 1、閾値以下であれば 0 を出力する。1 画像分の出力結果は 8 × 6 のマトリクスで表される。(図 6) また、検出する色の範囲と閾値は容易に変更が可能である。これにより、肌の色の分布検出を実現できるから、人間の注目を認識できる。

```
00000111
00000110
00111000
00011000
00001110
00000011
```

図 6 画像処理結果の出力例

### 5.2 フレーム間差分

フレーム間差分は色検出により得られたマトリクスの各成分が一つ前のフレームの結果と比較し異なれば 1、同じならば 0 を出力する。1 画像分の出力結果は色検出と同様 8 × 6 のマトリクスで表される。

これにより、フレーム間差分処理が実現できるから、人間の回転動作を認識できる。

## 6 まとめと今後の課題

本論文では人間の動作認識を実行できる環境カメラノードを、CMOS カメラと FPGA を組み合わせることにより実現し、その動作確認をおこなった。そして提案システムは人間の回転動作および注目を認識できる。本システムにより、ネットワークに無用な画像データを流す事なしに、画像処理結果のみを送信することが可能となる。本システムをコミュニケーションロボットシステム中に適用するには、通信の無線化および、様々な人間認識処理で使われる多種多様な画像処理を追加する必要がある。

### 参考文献

- [1] 石井健太郎、今井倫太: “コンテキストを持つ環境センサ: ヒューマン・ロボットインタラクション支援システム COSPI の提案”, 第 18 回人工知能学会全国大会, CD-ROM 版 IB2-01, 2004.6