

画像処理とゲームを組み合わせた エンターテインメントシステムの提案

A Proposal for Entertainment System Combining Image Processing with Game

林健太郎, 加藤邦人, 山本和彦, 大田紘高
Kentaro HAYASHI, Kunihito KATO, Kazuhiko YAMAMOTO and Hirotaka OHTA

岐阜大学 工学部

(〒501-1193 岐阜県岐阜市柳戸 1-1, {hayashi, ohta}@yam.info.gifu-u.ac.jp, {kkato, yamamoto}@info.gifu-u.ac.jp)

Abstract: We are developing a person detection and person tracking system by the entertainment robot, especially friendly robots using the image processing. We considered that the first step of the communication between the person and the robot is the eye contact. Thus, the person's face should be detected at first. In order to detect a person's face by the template matching, there is a problem how and when a robot learns a person's face. In this paper, we describe how to learn a person during the easy game. By using the dictionary made by this game, we realized the system that has communication ability between the human and the robot by detecting the person and turning robot's eyes to a person.

Key Words: Entertainment Robot, Person Learning, Four Directional Feature Fields

1. はじめに

近年、エンターテインメントを目的としたロボットの開発が盛んに行われている[1]。しかし、これらのロボットの多くは、あらかじめ決められた動作を行うだけのものが多く、その動作を人間が一方的に見て楽しむものがほとんどである。そこで、本研究では、人間とロボットとがより深くコミュニケーションをとることを目指す。ロボットが人間とコミュニケーションをとるために、ロボットの両眼にステレオカメラを搭載し、カメラから得られる情報により人間とのコミュニケーションを行う。

ところで、人間同士のコミュニケーションを考えた場合、最初に行われるのはお互いが視線を合わせることである。そこで本システムではロボットの両眼にアクティブカメラを搭載し、人物を検出、追跡するシステムを構築し、視線をあわせることでより親近感のあるエンターテインメントロボットを構成した。ところで、このシステムでは、テンプレートマッチングにより人物を検出するため、検出する人物をいつ学習するかという問題があった[2]。そこで本稿ではロボットと簡単なゲームを行うことでロボットがアイコンタクトする人物の辞書を作成する方法について提案する。ゲームのルールとして、ゲームの得点が高いほどより良い辞書が作成されるようなゲームを構成した。ゲームで作成した辞書を用いて人物を追跡し、人物に対して顔を向け、視線を合わせることにより、エンターテインメント性を高めたシステムを実現した。

2. システムの概要

2.1 システム構成

本システムで使用するロボットの概観図を図1に示す。本研究では新たにロボットの頭部を作成し、両眼にステレオカメラを搭載した。図2に頭部の概観図を、図3にその内部構造を示す。また、両眼のカメラ間の距離は7cmでカメラの光軸は平行に設置した。

ロボットの両眼に搭載したステレオカメラの制御にはサーボモータを使用し、コンピュータからRS232-Cによって制御している。それぞれのカメラは同期をとって制御しているため、カメラ間の距離は常に一定であり光軸は平行である。眼は上下左右に動くことが可能であり、眼の可動範囲は上下に $+10 \sim -10^\circ$ 、左右に $+45 \sim -45^\circ$ である。



図1 ロボットの概観図



図2 頭部の概観図

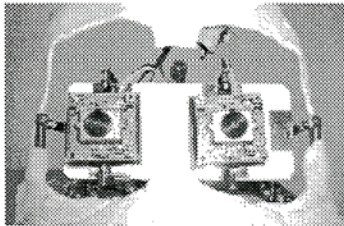


図3 頭部の内部構造

3. 人物の学習

一般的に人物追跡システムでは、あらかじめ追跡する人物の辞書を学習しておく必要がある。しかし、エンターテインメントロボットでは、いつ追跡する人物を学習するかが問題となる。また、テンプレートマッチングにより人物を追跡する場合、図4に示すように追跡する人物が移動することで照明環境などにより人物の見え方に変化が生じる。そこで、マッチングの精度をあげるには、人物が様々な場所にいる画像で人物を学習することによって、環境の変化に強い辞書を作る必要がある。



(a) 右側から照明が当たった例



(b) 左側から照明が当たった例

図4 照明による人物の見え方の変化

そこでこれらの問題を解決するために、ロボットと簡単なゲームを行うことで、様々な状況を意図的に作り出し、パリエーションの多い人物の画像を得る方法を提案する。

3.1 ゲームの内容

エンターテインメントロボットを用いた人物の学習において、ユーザ側からはロボットが学習していることを意識させないように、かつユーザが楽しむことができるゲームを構成する必要がある。ロボット側からはゲームの得点が高い場合、人物の追跡に有効な辞書が作成できるようなゲームを構成することが望ましい。そこで、次のようなゲームにより、人物の学習を行う。

まず、図5に示すように、ロボットの前方の4ヶ所にエリアを用意する。そして、ロボットの近くにあるモニタにゲームの画面を映し、ユーザは4ヶ所のうちのあるエリアに移動することによりゲームのキャラクターを動かす。人物の移動とキャラクターの移動をリンクさせ、人物に移動して欲しい位置をゲーム側で要求する。要求した位置に人物が移動したら得点を加算するというゲームにより、4ヶ所の位置で均等に画像を得る。これにより、ゲームの得点が高い人物は、様々なパリエーションの画像をかたよりにく取得することができる。

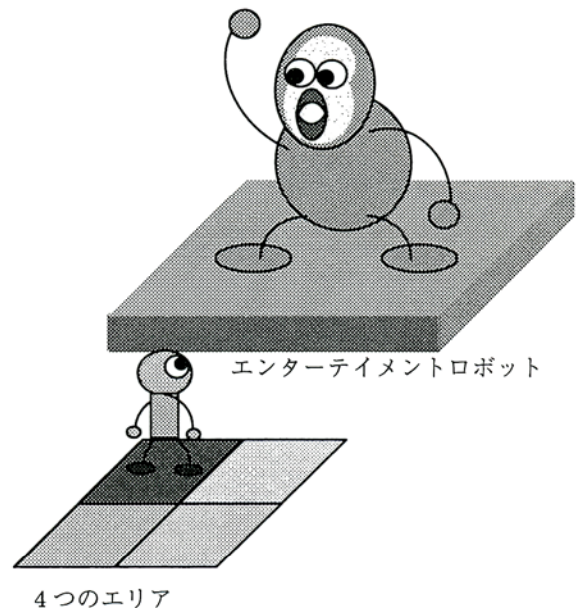


図5 システムの概要図

3.2 辞書作成方法

ゲームを行うことで得た複数枚の人物の画像から、テンプレートを作成する。図6にテンプレートの作成のための顔領域抽出方法を示す。

まず、背景差分により抽出した差分領域を人物領域とする。抽出した差分領域から、y軸についての射影ヒストグラムを作成する。一般的に、人間は、頭部の幅や肩の幅よりも首の部分が細くなっているため、射影ヒストグラムの谷となる部分の最小値をもとめることで胴体部分と顔領域を分割することとした。分割した領域のうち、上部にあ

る領域を顔領域とした。このようにして抽出した顔領域から、四方向面特徴[3]を抽出する。図7に四方向面特徴の抽出例を示す。

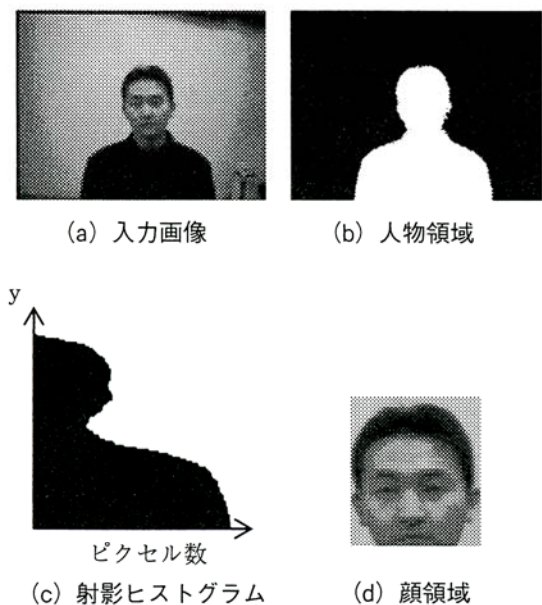


図6 顔領域の抽出

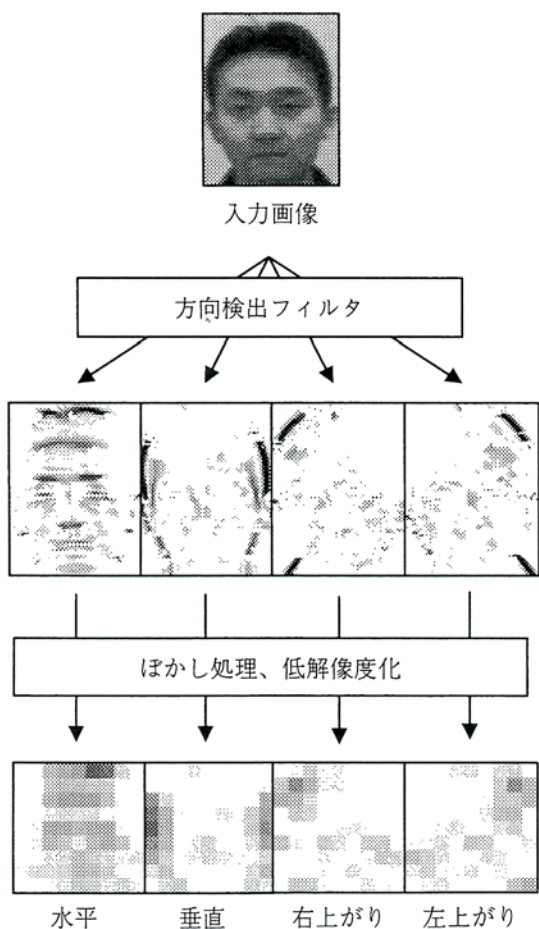


図7 四方向面特徴抽出例

この特徴は、文字認識や人物認識の分野で良く用いられている手法のひとつで [4]、低解像度化しても性能が落ちにくく、ノイズなどの影響を受けにくいという特長をもっている。四方向面特徴の抽出には、濃淡の入力画像から、エッジ検出フィルタを用いて水平方向、垂直方向、右上がり方向、右下がり方向の各方向のエッジ強度を求め、その強度に応じた4つの方向面を作成する。その後、ぼかし処理、低解像度化を行う。ゲームを行うことで得た複数枚の画像から抽出した特徴を足しこむことで人物のテンプレートを作成する。このテンプレートを用いて人物の追跡を行う。

4. 実験

4.1 人物の学習

3章で説明した人物の学習方法により、ロボットの要求したとおりに人物が動作した場合、すなわち、様々な場所で人物を学習した場合に、人物の検出に有効なテンプレートが作成されているかどうかの比較実験を行った。比較するテンプレートは、図5に示す4ヶ所のエリアのうち、1ヶ所のエリアのみでの画像を用いて学習したもの、2ヶ所で学習したもの、4ヶ所で学習したものの比較を行った。学習に用いた画像枚数はどのテンプレートも100枚で、1ヶ所のエリアで作成したものはそのエリアでの画像100枚、2ヶ所のエリアで作成したものは1ヶ所のエリアで50枚ずつの計100枚、4ヶ所のエリアで作成したものは1ヶ所のエリアで25枚ずつの計100枚となっている。

4.2 人物検出実験

人物検出の実験データとしては、人物が2人でカメラからの距離が90cm~170cmの範囲を人物同士が重ならないように自由に前後に動いたものを1組として使用した。図8にロボットの左右の目から見た入力画像例を示す。被験者は7人で、実験画像枚数は1組につき400枚、被験者1人につき6通りの組合せができ、更に人物が左、右にいるときの2通りの場合が存在するので追跡する人物1人につき $400 \times 6 \times 2$ の合計4800枚を使用した。被験者は7人である。表1に異なった条件で検出した結果を示す。ここで、1ヶ所のエリアで学習したテンプレートは4種類、2ヶ所のエリアで学習したテンプレートは6種類できるので、表1ではこれらの平均検出率を示す。



図8 入力画像例

表1 平均検出率

学習した エリア数	1ヶ所	2ヶ所	4ヶ所
平均追跡率	97.9%	98.3%	98.8%

結果より、学習するエリアの数を増やすに従い追跡率が向上した。これにより人物が様々な位置にいる画像を用いて人物を学習することで、照明環境の変化や人物が移動することによって起こる形状の変化などを吸収したテンプレートを作成できることが分かった。

ロボットの要求したとおり行動した人物はゲームの得点が高くなり、その人物に対する辞書を選ぶことによって追跡に有効な辞書が構成されることを確認した。

5. まとめ

本研究では、追跡に有効な人物学習を、ロボットと簡単なゲームを行うことによって実現する手法を提案した。そして、実験結果より、人物が様々な位置にいる画像で学習することで、照明環境の変化や形状の変化を吸収した辞書が作成されることにより、より高精度な人物追跡が可能となった。これにより、ロボットの要求する動作を行った人

物は、ゲームの得点が高く、かつ人物追跡に有効なテンプレートが作成される。

今後の課題として、よりエンターテイメント性を高めたシステムの構築や、より追跡に有効なテンプレート作成方法の検討などがあげられる。

参考文献

- [1]SONY “Entertainment Robot AIBO”
<http://www.aibo.com/>
- [2]平光, 大田, 山本, 加藤, 林: “エンターテイメントロボットによる人物検出システム”, 信学技報, PRMU-2001-25, pp.41-46, 2001
- [3]K.Yamamoto: “Present State of Recognition Method on Consideration of Neighbor Points and Its Ability in Common Database”, IEICE Trans. Inf. & Syst. Vol. E79-D, No. 5, pp. 417-422, 1996
- [4]S.Kuriyama, K.Yamamoto, H.Hongo and K.Kato: “Evaluation of A Face Recognition Method Using Four-Direction Feature Fields”, Proc. of FCV2001, pp. 128-133, 2001